UNIVERSITY OF CALIFORNIA
RIVERSIDE

Collaborative Estimation of Regional Atmospheric Delay Errors for Rapid Precise Point Positioning

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Jean-Bernard Uwineza

June 2024

Dissertation Committee:

Dr. Jay A. Farrell, Chairperson
Dr. Matthew J. Barth
Dr. Wei Ren

Copyright by
Jean-Bernard Uwineza
2024

The Dissertation of Jean-Bernard Uwineza is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I am deeply grateful to my advisor, Professor Jay Farrell, for his unwavering encouragement, guidance, and support throughout this journey. His expertise and constant motivation have been instrumental in my success. I would like to thank Professor Matthew Barth and Professor Wei Ren for serving on the committee of this dissertation defense. Their work on cooperative intelligent transportation systems and distributed consensus algorithms, respectively, has been a major inspiration to the direction of this research.

This work has greatly benefited from fruitful discussions with current and former colleagues in the Controls and Robotics research group. I am particularly grateful to Wang Hu, Zeyi Jiang, Farzana Rahman, and Mohammad Billah for their assistance in running data collection experiments for various projects and for their insightful discussions on the theoretical foundations of our shared research interests. Many thanks also to Elahe Aghapour, Kathryn Hammar, Mike Stas, and others for years of engaging meetings and discussions in the lab.

I am also grateful to other researchers in the field who have generously shared their open-source software projects and data archives, making my research easier. I especially thank Marcus Wareyka-Glaner, whose software package raPPPid has been the foundation for the GNSS processing software developed for this dissertation.

I am deeply indebted to the many friends and colleagues who have supported me personally and professionally throughout my PhD journey. I am especially grateful to Leeann, my girlfriend and best friend, whose support has been instrumental in my success.

Finally, I am immensely grateful to my family for their sacrifices, which have allowed

me to pursue higher education far from home. Though many years and a great distance separate us, your love and support is always with me. I am also grateful to the Shaffer family for providing a home away from home.

To my parents, for their resolute support.

ABSTRACT OF THE DISSERTATION

Collaborative Estimation of Regional Atmospheric Delay Errors
for Rapid Precise Point Positioning

by

Jean-Bernard Uwineza

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, June 2024
Dr. Jay A. Farrell, Chairperson

Recent advances have established Precise Point Positioning (PPP) as a preferred technique for providing accurate and precise position solutions for users anywhere on the globe. PPP relies on the availability of accurate corrections of satellite orbits, clocks, and hardware biases, in addition to other corrections which are estimated using a global network of reference stations. These global corrections allow PPP to achieve centimeter-level positioning accuracy using pseudorange and phase observations. However, PPP suffers from long convergence times, mainly due to residual errors from atmospheric delay corrections. Therefore, the current challenge for PPP lies not in improving its accuracy, but rather in improving the convergence time. This dissertation presents two novel approaches to improve PPP convergence time by focusing on the estimation of regional atmospheric delays.

First, a new approach is developed to enable collaboration between satellites in estimating regional atmospheric delays using measurements from a single receiver. This approach takes advantage of the increasing availability of multi-GNSS signals by using uncombined

measurements from all usable satellites. By estimating the ionospheric delay as a regional Vertical Total Electron Content (VTEC) map parameterized using polynomial B-splines, the approach allows satellites to collaborate in estimating each other's slant delays. Results show that this collaboration improves the stability and precision of ionospheric delay estimates, leading to faster convergence of PPP solutions. The method is not limited to multi-frequency users and also benefits single-frequency users.

Furthermore, the dissertation proposes a collaborative PPP approach that utilizes estimated regional atmospheric corrections from a network of local agents. This approach combines corrections from global models with estimates from local collaborating agents to derive more accurate regional corrections. The resulting collaborative estimation problem is formulated as an average consensus problem, and a distributed information-weighted consensus algorithm is developed to solve it. The algorithm allows agents to share *only* their atmospheric delay estimates and uncertainties with their neighbors. This distributed approach eliminates the need for a central processing center and enables agents to instantaneously achieve network accuracy upon joining the network, thereby significantly improving PPP convergence time. The collaborative PPP approach is shown to achieve instantaneous convergence for some users under favorable conditions, demonstrating its potential for real-time precise positioning applications.

Contents

List of Figures

List of Tables

List of Algorithms

Acronyms

AC Analysis Center 2, 17, 102, 112

APC Antenna Phase Center . . . . . 52 f.

AR Ambiguity Resolution 3, 47, 62, 155

ARP Antenna Reference Point . . . . . 20 f.

BDS BeiDou 1, 10, 17, 19

CAV Connected Automated Vehicles 1

CDF Cumulative Distribution Function 150

CDMA Code Division Multiple Access 19, 43

CORS Continuously Operating Reference Stations . . . . . 2, 4, 102 f., 108, 112, 115, 175

DCB Differential Code Bias . . . . . 17 f., 37, 43, 50, 54, 56 f., 59 ff., 64, 83, 86, 99

DGNSS Differential GNSS . . . . . 42 f., 45

DGPS Differential Global Positioning System 2, 42

ECEF Earth-Centered Earth-Fixed . . . . . 50–53, 59, 111 f.

ECMWF European Centre for Medium-Range Weather Forecasts 29, 32

ENU East-North-Up/Height . . . . . 111 ff.

FDMA Frequency Division Multiple Access 19

FKP Flächen–Korrekturparameter 48

GAL Galileo 1, 10, 17, 81, 102, 104, 143

GDV Group Delay Variations . . . . . 15, 19 f.

GIM Global Ionospheric Model 17 f., 37 ff., 70, 79, 82, 84, 101 f., 107 ff., 114, 119, 142, 152

GLONASS Global'naya Navigatsionnaya Sputnikovaya Sistema 1, 10, 19

GM Gauss-Markov . . . . . 69, 89 f., 94–99

GNSS Global Navigation Satellite Systems . 1, 7, 10 f., 13, 15, 18 ff., 22, 24, 26 ff., 33 ff., 41 f., 45, 54 f., 59, 62, 69, 75, 79, 81, 83 f., 86 f., 102, 114, 119, 125, 139, 154, 156 f.

GPS Global Positioning System 1, 10, 16, 42, 81, 83, 102, 104

GPT Global Pressure and Temperature 29, 68

IF Ionosphere-free 57–60, 64

IFLC Ionosphere-free Linear Combination . . . . . 57 ff., 61, 64

IGS International GNSS Service . . . . . 17 f., 21, 37, 51 f., 101–104, 112

INS Inertial Navigation System 156

IONEX Ionosphere Exchange 17

IRNSS Indian Regional Navigation Satellite System 10

ITS Intelligent Transportation Systems 1, 119, 132, 152, 155

KF Kalman Filter . . . . . 89, 91 f., 94, 113, 115, 138 ff., 142, 144

LAM Local Atmospheric Modeling 107, 109, 155

LAM-PPP PPP with Local Atmospheric Modeling . 7, 62, 101 f., 104, 107 ff., 111–117, 119, 139 f., 143, 145 f., 149, 154, 157

LOS line-of-sight . . . . . 20 f., 24, 80, 84, 90

MAC Master-Auxiliary Concept 48

MAD median absolute deviation 111

MAP maximum-a-posteriori . . . . . 129, 132, 135 f.

NLOS non line-of-sight 24

NRTK Network RTK . . . . . 2, 4, 47 ff., 51, 62, 119, 125, 127

OSB Observable-specific Code Biases . . . . . 54 f.

OSR Observation Space Representation 2, 49

PAR Partial Ambiguity Resolution 155

PCO Phase Center Offset . . . . . 20 f.

PCV Phase Center Variation . . . . . 20 f.

PPP Precise Point Positioning vii f., 2–7, 15, 49 ff., 57, 59 f., 62 ff., 66, 68, 70, 82, 89, 101, 117, 119, 121, 125, 144, 146, 154–157

PSD Power Spectral Density . . . . . 95 f., 98

PVA Position-Velocity-Acceleration 156

QZSS Quasi-Zenith Satellite System 10

RINEX Receiver Independent Exchange Format 53, 102

ROTI Rate of TEC Index 70

RTK Real-time Kinematic . . . . . 2, 45 ff., 49, 51, 62 f., 125

RTS Real Time Service 52

SA Selective Availability 42

SH Spherical Harmonics 69

SINEX Solution (Software/Technique) INdependent EXchange 17, 54

SP Service Provider 47–50, 62, 125

SP3 Standard Product 3 Orbit 52

SPS Standard Positioning Service . . . . . 41 f., 49, 51

SSR State Space Representation . . . . . 2, 51 f., 55, 62, 125

STEC Slant TEC . . . . . 34 f.

TEC Total Electron Content . . . . . 33 f.

TECu Total Electron Content units 33, 35, 70

VMF3 Vienna Mapping Functions 3 . . . . . 32 f.

VRS Virtual Reference Station . . . . . 48 f.

VTEC Vertical TEC 34 f., 37, 39, 67, 69–73, 75, 79 f., 82, 88, 94, 101 f., 104, 106 f., 114, 116 f., 119, 121 f., 127, 136, 142 f., 151 f., 157

ZHD Zenith Hydrostatic Delay 28, 31, 39

ZTD Zenith Total Delay . . . . . 28, 30, 59, 68, 99, 102 ff., 139 f.

ZWD Zenith Wet Delay . . . . . 28, 30 f., 39, 68 f., 80, 99, 104, 119 f., 139, 151

Chapter 1

Introduction

The availability of highly precise positioning has become a necessity to many aspects of modern life. A wide array of existing and upcoming applications are enabled by the existence of accurate and reliable positioning and location information. One such application is automotive navigation, which is enabled by Intelligent Transportation Systems (ITS). An emergent direction in ITS is towards Connected Automated Vehicles (CAV), which are the future of mobility for people and goods. Some of the foremost navigation performance requirements for CAV are: decimeter-level accuracy, high availability, and integrity [1, 2]. No individual technology can currently offer that level of performance, and a suite of sensors needs to be used, with Global Navigation Satellite Systems (GNSS) as the main providers of absolute positioning [2]. This work presents improvements to a GNSS positioning technique used in ITS.

GNSS positioning services have been around in various forms since the 1980s [3]. Initially there was only one satellite constellation, the American Global Positioning System (GPS), followed by the Russian Global'naya Navigatsionnaya Sputnikovaya Sistema (GLONASS),

and more recently, the European Galileo (GAL) and the Chinese BeiDou (BDS). The status of all GNSS constellations will be further discussed in Section 2.1.

In the intervening decades, the positioning accuracy offered by these systems has improved from tens of meters to a few meters for standard positioning using only pseudorange measurements from a single receiver. In the late 1990s, positioning accuracy was further improved by using differencing techniques to eliminate various common measurement errors due to the satellite signals' propagation. The first of these techniques was Differential Global Positioning System (DGPS), which achieves decimeter-level accuracy by using the pseudorange measurements from a nearby GNSS reference (base) station with precisely known coordinates. The base stations are typically Continuously Operating Reference Stations (CORS), and the measurements are communicated in the Observation Space Representation (OSR) format. Further still, centimeter-level accuracy was soon achieved by the Real-time Kinematic (RTK) technique, which relies on adding precise carrier-phase measurements to DGPS. Since both DGPS and RTK rely on eliminating common-mode atmospheric delays errors, their performance heavily depends on the separation (baseline) between the base receiver and the user. For longer baselines, the atmospheric delays at both the user and the base station are longer the same. To address this problem, Network RTK (NRTK) was then developed as a technique of using a network of base receivers to interpolate and send the necessary corrections to the user. The NRTK measurements are processed at an Analysis Center (AC), and the resulting corrections are disseminated to users as State Space Representation (SSR) corrections. These SSR corrections include satellite orbits, clocks, and biases, as well as ionospheric and tropospheric delays. The current performance of various positioning services will be discussed in Chapter 4.

Although DGPS and NRTK deliver exceptional precision and accuracy, they require extensive infrastructure in the form of dense networks of local base stations and ACs. This severely narrows the application of these techniques, since their service area is very limited. Additionally, ACs need to compute and communicate corrections for each user in the network. This severely limits the scalability of the network size, since the increase of users increases computation and communication load on the ACs. To address this problem, PPP was developed as a similar technique to NRTK, but with the ability to provide precise positioning to users anywhere on the globe. PPP, similar to NRTK, uses a network of reference stations spread around the globe to estimate global SSR corrections. Due to the sparsity of the base stations, and unlike NRTK, PPP does not provide precise regional atmospheric corrections as part of the SSR corrections. As a consequence, users have to approach PPP by either estimating ionospheric delay unknowns, or eliminating them via linear combinations of measurements. These approaches add more unknown states to estimation, or reduce the number of available measurements, respectively. Inevitably, PPP take several tens of minutes to converge to the precision and accuracy of NRTK, primarily due to the lack of precise local/regional atmospheric corrections. Hence, there is a need for novel techniques that reduce the convergence time, while preserving other advantages of PPP.

1.1 Motivation

Long convergence time has been the main drawback of PPP techniques since their inception. For the case of GPS dual-frequency, convergence to centimeter-level accuracy

typically takes about 30 minutes to more than an hour. To speed up convergence, research has focused on making progress on Ambiguity Resolution (AR) techniques to produce the so-called *fixed* PPP solutions, as opposed to the typical *float* PPP solutions. (Fixed solutions use AR to take advantage of the more precise phase measurements.) These AR techniques have proven to be beneficial for high-quality receivers which are able to generate measurements free of cycle-slips in open-sky environments. However, due to residual local atmospheric delay errors, the float and fixed solutions are typically similar during the first few minutes and after float solutions has converged. In fact, it has been noted that during the first few minutes the float solution might be slightly better than the fixed solution[1], i.e., a specific time period might be required for the float solution to converge to ensure correct integer ambiguity resolution [4]. This is because very accurate estimates of receiver clocks and biases are necessary to resolve float ambiguities into integer ambiguities. In that sense, PPP convergence time not only means the duration it takes the float solution to converge, but also the duration it takes to ensure successful integer ambiguity resolution. For that reason, speeding up the converge of the float solution directly affects the success of ambiguity resolution, and leads to faster fixed solutions.

Generally, local atmospheric information has also been shown to speed up convergence [5]. Indeed, recent research has demonstrated that near-instantaneous convergence for PPP hinges on the availability of highly accurate local atmospheric delay corrections, supplementing the standard satellite and site-related corrections [6]. There are strict requirements on accuracy for these local corrections, and techniques for estimating such cor-

[1]↑It was noted in [4] that the float solution is better than the fixed solution for the first 15 minutes. The duration of this period is not universal, and is likely to depend on a number of factors, such as the quality of the receiver or the environment in which it operates.

rections have thus far required dense CORS infrastructure and processing similar to that of NRTK. This use of extensive local infrastructure has been the main limitation of estimating regional atmospheric delays. Consequently, there is a need for novel techniques that improve PPP convergence time efficiently with less dense networks. These local corrections must be very accurate, and current estimation techniques necessitate dense CORS infrastructure and processing akin to NRTK. The reliance on extensive local infrastructure has been one of the primary obstacles to estimating regional atmospheric delays effectively. Therefore, there is a pressing need for innovative techniques that can enhance PPP convergence time efficiently without relying on dense networks.

There has been recent promising attempts at estimating local atmospheric delays. For example, the work in [7] leverages a network of nearby reference stations and used PPP to estimate slant ionospheric delays, then interpolate between the estimates to obtain the actual slant delay at the user. A more recent proposal [8] involves partitioning a dense network of receivers into smaller sub-networks to ensure the similarity of tropospheric and slant ionospheric delays within each sub-network. The sub-networks are then used to estimate satellite clocks and biases, as well as atmospheric delays. The slant ionospheric delays are then interpolated for each user in the network, which introduces non-negligible interpolation errors. This work aims to develop approaches that enable users in a network to share ionospheric delay estimates without the need for interpolation.

1.2 Goals & Contribution

The main goal of this work is to reduce PPP float convergence time by using collaborative approaches of estimating the regional atmospheric delays. This is accomplished by: (1) enabling collaboration between satellites in estimating regional atmospheric delays, and (2) combining corrections from global models and estimates from a network of local receivers to derive more accurate regional atmospheric corrections. The research presented herein focuses on developing an approach of estimating regional atmospheric delays, which will be shown to result in better estimates for other unknowns such position, clocks, and ambiguities.

Our approach ensures that the local collaborating agents are not only high-quality reference stations, but could also include automotive-grade receivers, smartphones, and other GNSS-equipped devices.

This dissertation's main contributions are:

- Develop a novel technique that enables collaboration between satellites when estimating local atmospheric delays in single-receiver PPP. The technique leverages the developing multi-GNSS capabilities (multiple frequencies, new precise signals, additional constellations, etc) by using uncombined measurement from all usable satellites. This contribution addresses the limitations of traditional PPP methods that rely on global atmospheric models, which may not accurately capture local variations.

- Develop a method for distributed collaboration between local agents to generate precise regional atmospheric delay corrections. This method eliminates the need for agents to share raw measurements or their full estimated states. Instead, agents only

share their atmospheric delay estimates, the only common subset of their full states.

This approach differs from previous methods because it eliminates the need for a regional analysis center and allows each agent to maintain its own precise estimate of regional atmospheric delay corrections, supplementing global corrections. The key advantage is the network's scalability: adding more users enhances correction accuracy without increasing individual computational or communication loads.

- Present comprehensive experimental results using real-world data to demonstrate the effectiveness of the proposed approaches in improving PPP convergence time and accuracy. These results provide valuable insights into the practical applicability and benefits of the developed methods.

1.3 Dissertation Overview

The remainder of this dissertation is structured as follows. Chapter 2 provides background information on GNSS fundamentals, including signal structure, measurement models, and error sources. Chapter 3 provides details of residual atmospheric delay errors, discussing their impact on GNSS signals and methods for correction. Chapter 4 explores various GNSS positioning services, comparing their performance and highlighting the need for improved PPP convergence time. Chapter 5 introduces a novel PPP approach with local atmospheric modeling PPP with Local Atmospheric Modeling (LAM-PPP) to enhance convergence time. Chapter 6 presents a collaborative PPP approach that leverages a network of agents to further reduce convergence time. Finally, Chapter 7 summarizes the findings, discusses potential issues, and suggests future research directions.

1.4 Notation

This dissertation adopts the notation previously defined by the author in [9], with some modifications. Different equality symbols are used to distinguish between definitions, computations, and theoretical models used for analysis. The symbol '$=$' indicates that an equation involves a theoretical model used for analysis and physical interpretation of quantities. The symbol '$\doteq$' indicates that the equation involves a computation, i.e. the quantity on the left-hand side could be directly computed from the quantities on the right-hand side. The symbol '$\triangleq$' indicates a *definition* of a specific model or quantity.

When necessary, the actual and estimated values of a quantity are distinguished by x and $\hat{x}$, respectively. Vectors and matrices are written in boldface, with vectors in lowercase and matrices in uppercase. For instance, $\mathbf{v} \in \mathbb{R}^n$ is an n-element vector and $\mathbf{V} \in \mathbb{R}^{n \times n}$ is an $n \times n$ matrix. Positive definiteness (i.e., $\mathbf{v}^\top \mathbf{V}\, \mathbf{v} > 0,\ \forall \mathbf{v} \neq \mathbf{0}$) and semi-definiteness (i.e., $\mathbf{v}^\top \mathbf{V}\, \mathbf{v} \geq 0,\ \forall \mathbf{v} \neq \mathbf{0}$) of a matrix $\mathbf{V}$ are indicated by $\mathbf{V} \succ 0$ and $\mathbf{V} \succeq 0$, respectively. The matrix $\mathbf{I}_{n \times n} = \mathbf{I}_n \in \mathbb{R}^{n \times n}$ is the $n \times n$ identity matrix. The notation $\mathrm{tr}(\cdot)$ is the trace operator and $\|\cdot\|$ is the L^2 norm. For a random vector $\boldsymbol{\zeta}$, its mean squared value will be denoted as $\|\boldsymbol{\zeta}\|_M^2 = E\langle \boldsymbol{\zeta}^\top \boldsymbol{\zeta} \rangle$.

The notation $\overrightarrow{\mathbf{A}}$ indicates the *vectorization* operation on matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, obtained by stacking the columns of $\mathbf{A}$, and defined as:

$$\overrightarrow{\boldsymbol{A}} = \mathrm{vec}(\boldsymbol{A}) = [a_{1,1}, \ldots, a_{1,m}, a_{1,2}, \ldots, a_{m,2}, \ldots, a_{1,n}, \ldots, a_{m,n}]^\top \in \mathbb{R}^{mn \times 1}. \tag{1.1}$$

The notation $\mathbf{A} \otimes \mathbf{B}$ indicates the *Kronecker product* of matrices $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{p \times q}$,

defined as:

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{1,1}\boldsymbol{B} & \cdots & a_{1,n}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{m,1}\boldsymbol{B} & \cdots & a_{m,n}\boldsymbol{B} \end{bmatrix} \in \mathbb{R}^{mp \times nq}. \tag{1.2}$$

A vectorized matrix product can be written in terms of the Kronecker product [10]:

$$\overrightarrow{\boldsymbol{ABC}} = \left(\boldsymbol{C}^\top \otimes \boldsymbol{A}\right) \overrightarrow{\boldsymbol{B}}. \tag{1.3}$$

Other notations will be defined on a case-by-case basis.

Chapter 2

GNSS Fundamentals

This chapter presents an overview of GNSS signals: their status, measurement models, and error sources. It starts with details of current GNSS constellations, then provides general measurement models of their signals, and ends with a thorough exploration of various error sources considered in the measurement models.

2.1 Multi-GNSS Status

The use of signals from multiple GNSS constellations (multi-GNSS) is very beneficial to all kinds of positioning services, not least due to the increased number of measurements, but also due to the increased diversity of the measurements. There are currently four operational global navigation satellite constellations: the American GPS, the European GAL, the Chinese BDS, and the Russian GLONASS. In addition, there are other operational regional constellations offering supplementary signals to those of global constellations, albeit with a reduced number of satellites. The Japanese Quasi-Zenith Satellite System (QZSS) and

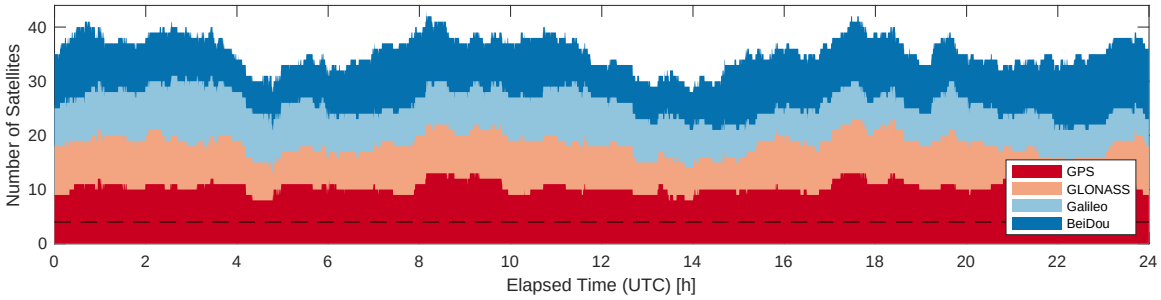


Figure 2.1: The number of visible satellites for the GNSS reference station `BILL` (in Southern California) on July 15, 2023.

Indian Indian Regional Navigation Satellite System (IRNSS) are the two most prominent of these regional constellations. All GNSS systems are equipped to transmit signals in multiple carrier frequency bands in the lower and upper L-band.

As of this writing, there are currently more than 130 operational GNSS satellites in different types of orbits around the globe. Most satellites transmit support at least two frequency bands, with the future modernization goal of supporting three or more bands on each satellite. The range of carrier frequencies ranges from 1176.45 MHz for the GPS L5/GAL E5a/BDS B2a band, to 1575.42 MHz for the GPS L1/GAL E1/BDS B1 band. Designing a receiver that covers all frequency bands poses both technical and financial challenges, and hence most low-cost receivers support one or two bands, whereas survey-grade or geodetic receivers may support three or more bands. In North America, on average around 30 satellites from global constellations are simultaneously visible at any time. Fig. 2.1 shows the typical number of visible satellites at any given time of day for `BILL`, a GNSS reference station in Southern California. In other regions of the world, the number of visible satellites can be much higher. For instance, in the Southeast Asia regions there is an average of around 50 satellites simultaneously visible at any time, due to satellites in geosynchronous orbits and regional constellations. A receiver that supports at least two

frequency bands and can generate both code and phase measurements will have 100 to 200 measurements at any epoch.

2.2 Measurement Models

Consider the following models for the code (pseudorange) and carrier-phase (phaserange[1]) measurements for the signal of frequency f_j, transmitted from satellite s, received by a user with receiver r:

$$\begin{aligned} \rho^s_{r,j} &= R^s_r + c(dt_r - dt^s + \Delta d^s_{r,j}) \\ &\quad + I^s_{r,j} + T^s_r + E^s + c\left(\delta t_{r,j} - \delta t^s_j\right) \\ &\quad + \xi^s_{r,j} - \varrho + m^s_{r,\rho} + \eta_{r,\rho} \end{aligned} \tag{2.1}$$

$$\begin{aligned} \phi^s_{r,j} &= R^s_r + c(dt_r - dt^s + \Delta \delta^s_{r,j}) \\ &\quad - I^s_{r,j} + T^s_r + E^s \\ &\quad + \left(\lambda_j N^s_{r,j} + \psi^s_r + h_{r,j} - h^s_j\right) \\ &\quad + \zeta^s_{r,j} - \varrho + m^s_{r,\phi} + \eta_{r,\phi} \end{aligned} \tag{2.2}$$

[1] In the literature, the terms carrier-phase, phase, and phaserange interchangeably refer to a range measurement derived from carrier-phase. In some cases, the distinction is made that the first two terms refer to a number of carrier-phase cycles, whereas the last term refers to the corresponding range obtained by multiplying the carrier-phase cycles to the speed of light. In the following text, both carrier-phase and phase will refer to a range quantity, unless otherwise stated. Similarly, the terms code and pseudorange will be used interchangeably to a range measurement obtained from the transmitted code.

Similarly, the code and phase measurements for a signal of frequency f_k are modeled as:

$$\begin{aligned}\rho^s_{r,k} &= R^s_r + c\,(dt_r - dt^s + \Delta d^s_{r,k}) \\ &\quad +\gamma_k\, I^s_{r,j} + T_r + E^s + c(\delta t_{r,k} - \delta t^s_k) \\ &\quad +\xi^s_{r,k} - \varrho + m^s_{r,\rho} + \eta_{r,\rho}\end{aligned} \tag{2.3}$$

$$\begin{aligned}\phi^s_{r,k} &= R^s_r + c\,(dt_r - dt^s + \Delta \delta^s_{r,k}) \\ &\quad -\gamma_k\, I^s_{r,j} + T_r + E^s \\ &\quad +\left(\lambda_k N^s_{r,k} + \psi^s_r + h_{r,k} - h^s_k\right) \\ &\quad +\zeta^s_{r,k} - \varrho + m^s_{r,\phi} + \eta_{r,\phi},\end{aligned} \tag{2.4}$$

where $\lambda_j = \frac{c}{f_j}$ is the f_j signal's wavelength, c is the speed of light, and $\gamma_k = \left(\frac{f_j}{f_k}\right)^2$ is an ionospheric coefficient. Details on the remaining symbols are given in Table 2.1. Section 2.3 provides more details on all the errors in eq. (2.1)–(2.4).

2.3 Measurement Error Sources

The sources of GNSS measurement errors can be grouped into three categories based on the three GNSS segments: space, control, and user segments, respectively. GNSS measurement errors are compensated for using a combination of the following four options:

- *Elimination*: Some errors can be eliminated by combining measurements from multiple signals or receivers. Errors (such as those due to satellite clocks and orbits) are common to all signals in the same receiver, and they can be eliminated by combining measurements from different signals. Atmospheric delay and site-displacement errors

Symbol	Unit	Magnitude	Uncertainty	Description and Notes
R_r^s	m	–	–	Geometric range between the antenna phase centers of satellite s and receiver r
dt_r, dt^s	s		20 ps – 3 ns	Receiver and satellite clock offsets
$\delta t_{r,j}, \delta t_j^s$	s		–	Receiver and satellite code hardware biases at frequency j
$\Delta d_{r,j}^s, \Delta \delta_{r,j}^s$	s	up to 5m		Code and phase Inter-channel code bias (ICB), respectively, for FDMA signals (GLONASS only)
$I_{r,j}^s$	m	up to 30 m	up to 1 m	Ionospheric delay at at frequency j. Higher order terms of the ionospheric delay can contribute up to 2 cm.
T_r^s	m	2.6 m	up to 0.40 m	Tropospheric delay. The dry tropospheric component contributes the most (2.3 m), but its uncertainty is only 5 mm. The wet component contributes about 0.3 m, but its uncertainty can be up to 0.3 m (100%).
E^s	m	1–2 m	2.5–5 cm	Ephemeris (orbit) errors
$\delta b_{r,jk}$	s	up to 100 ns	–	Receiver Differential Code Bias (DCB) of frequency k relative to frequency j.
δb_{jk}^s	s	up to 20 ns	0.1 - 1 ns	Satellite Differential Code Bias (DCB) of frequency k relative to frequency j
$h_{r,j}, h_j^s$	m			Frequency-dependent receiver and satellite phase hardware biases for frequency j
ψ_r^s	cy	0–1 cy	–	Phase wind-up
$\zeta_{r,j}^s$	m	up to 30 cm	10 cm	Phase center offset (PCO) and phase center variations (PCV)
$\xi_{r,j}^s$	m	1mm–1.5m	–	Group delay variations (GDV) for frequency j.
$N_{r,j}^s$	m	–	–	Phase ambiguity for frequency f_j
ϱ	m	up to 0.4 m	5 cm	Site displacement
$m_{r,\rho}^s$	m	up to 100 m	–	Code multipath
$m_{r,\phi}^s$	m	up to 10 cm	–	Phase multipath
$\eta_{r,\rho}$	m	up to 1 m	–	Code random noise
$\eta_{r,\phi}$	m	up to 2 mm	–	Phase random noise

Table 2.1: Definition of measurement model symbols.

are, to some extent, common to all users in the same geographic area, and they can be eliminated using a combination of measurements from two or more receivers separated by a small distance.

- *Estimation*: Errors which are common to all users (e.g. satellite clocks and orbits) can be precisely estimated by networks of receivers, whereas errors which are unique to each user (e.g. receiver clock) can be estimated individually alongside the receiver coordinates.

- *Modeling*: The GNSS navigation message contains parameters to model the satellite clock and orbit. It also contains parameters to model the ionospheric delay error. Other errors such as the tropospheric delay and site displacements can be modeled with sufficient accuracy by empirical models in standard positioning applications.

- *Omission*: Applications that do not require high precision can afford to omit error sources with relatively smaller magnitudes, such as Group Delay Variations (GDV) or site displacements.

Different GNSS positioning approaches are classified mainly based on which of the above options they use to mitigate and account for these error sources. In positioning services like PPP, errors from space and control segments can be corrected by estimates and empirical models from external sources, whereas some errors from the user segment can only be estimated by the user, if possible, or omitted in other cases. The following subsections provide more details on the source of different kinds of errors and how they affect GNSS measurements. Chapters 3 and 4 will discuss how these errors are handled in various positioning services, especially in PPP–the positioning service of interest in this dissertation.

2.3.1 Clock Errors

The receiver clock error, dt_r, is usually composed of the reference constellation's receiver clock bias, plus the receiver clock offset of any additional constellation. When multiple GNSS constellations are processed, a receiver clock offset is added to the observation model and estimated for each processed GNSS other than the reference GNSS. Therefore, the reference clock error and a certain number of receiver clock offsets are estimated as:

$$dt_r = dt_r^{GNSS_{ref}} + \Delta t_r^g, \tag{2.5}$$

where $dt_r^{GNSS_{ref}}$ is the reference system's clock bias and Δt_r^g is the clock offset for any additional GNSS system g (for example GAL.) An equivalent option would be to estimate the receiver clock error for each processed GNSS[2]. Typically, the reference GNSS in eq. (2.5) is GPS because of its maturity and ubiquity.

2.3.2 Hardware Delays

Hardware biases are delays to the signal propagation induced by the receiver and satellite hardware, and interpreted by the receiver as variation in the measured range. These biases are due to the physical characteristics of the used hardware: satellite and receiver circuitry, cables, and antennae. and/or signal tracking implementation. Consequently, they are relatively stable over time. For CDMA signals, receiver hardware biases are frequency-dependent and are estimated per GNSS constellation.

[2]↑For simplicity, eqs. (2.1)–(2.4) assume that clock errors are estimated separately for each processed GNSS system, i.e. there's no reference clock. In practice, GPS is selected as the reference clock and clock errors of any additional system are estimated as offsets to the reference clock. These two approaches are equivalent.

Code hardware biases ($\delta t_{r,j}, \delta t^s_j$) and phase hardware biases ($h_{r,j}$, h^s_j) are difficult to estimate in their undifferenced form, as they are highly correlated with other terms, such as clock errors for code biases or integer ambiguities for phase biases. Thus, only differences between biases can be estimated directly from code and phase observations [11]. For some positioning techniques such as PPP, it is often sufficient to know only the differences between certain biases. The calculated positions are not influenced because common offsets to the absolute biases might be absorbed by other estimated terms (e.g., the receiver clock error).

2.3.2.1 Differential Code Biases

The Differential Code Biases (DCB) or inter-frequency biases (IFB) for both receiver and satellite are the per-frequency difference between code hardware biases. The satellite and receivers DCBs between frequencies f_j and f_k are defined as:

$$\delta b_{r,jk} = \delta t_{r,j} - \delta t_{r,k}, \tag{2.6}$$

$$\delta b^s_{jk} = \delta t^s_j - \delta t^s_k. \tag{2.7}$$

Satellite DCBs are initially determined by the satellite manufacturer before launch and can be revised by the constellation's control segment or continuously estimated as a by-product of ionospheric monitoring [12]. For example, DCBs are estimated by International GNSS Service (IGS) ACs and disseminated as standalone DCB products in the format known as Solution (Software/Technique) INdependent EXchange (SINEX), or they're provided along with Global Ionospheric Model (GIM) products in the format known as Ionosphere Exchange (IONEX). The magnitude of satellite DCBs is usually less than 20 ns, and stable for several hours (typically assumed stable for one day) [13].

The magnitude of receiver DCBs varies significantly across different receiver models. Within the same receiver family (e.g. Leica survey-grade receivers), the DCBs typically vary in a range of 10–15 ns for different frequencies within the same GNSS system. For instance, the Differential Code Bias (DCB)s between GPS L1 C/A and L2C for specific models of Leica receivers used in IGS ACs was found to range from 15 ns to 25 ns [13]. Across different types of receivers, the DCBs typically vary in a range of 20–40 ns. For instance, the DCBs between GAL E1 and E5b on *different* Leica receiver models has been found to cover a range of -55 ns to -45 ns, whereas they cover a range of 5 ns to 15 ns for Septentrio receivers. However, it is possible for the DCBs for different receivers to differ as much as 100 ns, especially for combinations of new signals on BDS and GAL. (See Table 3 in [13] for a comprehensive treatment of receiver DCB ranges for various receivers used by IGS.)

The receiver DCBs, $\delta b_{r,jk}$, can be estimated from code measurements in eqs. (2.1) and (2.3), using ionospheric delay error estimates and satellite DCBs:

$$\delta b_{r,jk} = \left(\frac{f_k^2 - f_j^2}{f_j \, f_k} \right) I_r^s - (\rho_{r,j}^s - \rho_{r,k}^s) - \delta b_{jk}^s, \tag{2.8}$$

where I_r^s can be obtained as part of state estimation or from external sources such as IGS GIM (as will be discussed in Section 3.2), and the satellite DCBs δb_{jk}^s can be obtained from external sources.

2.3.2.2 Phase Biases

Phase biases are often omitted from GNSS measurement models as they cannot be distinguished from phase ambiguities. This is fine for applications in which there is no intention to resolve float ambiguities into integers, but phase biases have to be modeled and

corrected for successful ambiguity resolution. Both satellite and receiver phase biases are frequency-dependent, and are typically not constant as they can vary with temperature. In this work, satellite phase biases h_j^s are corrected using external products, and receiver phase biases $h_{r,j}$ are lumped together with ambiguities.

2.3.3 Inter-channel Biases

Inter-channel bias errors (ICBs), $\Delta d_{r,j}^s$, are due to the presence of different channels for different frequencies in constellations using Frequency Division Multiple Access (FDMA). These are not present when processing measurements from constellations using Code Division Multiple Access (CDMA) signals, i.e $\Delta d_{r,j}^s = \Delta \delta_{r,j}^s = 0$. Currently, GLONASS is the only GNSS constellation using FDMA. This dissertation assumes that only CDMA signals are used, unless otherwise stated.

2.3.4 Group Delay Variations

Group Delay Variations (GDV), $\xi_{r,j}^s$, are code measurement errors which depend on antenna type and signal frequency. They vary with the nadir angle of the transmitted signal and the elevation angle of the received signal [14]. GNSS constellations exhibit different GDV characteristics based on: transmitted frequency, but also on the generation (age) of satellite hardware, and orbit type. For example, the second generation of BDS satellites (BDS-2) have been shown to introduce much larger GDV error compared to GPS satellites [15]. Moreover, GDV errors are more pronounced for satellites in medium earth orbit (MEO) than for those in inclined geosynchronous orbit (IGSO) or geosynchronous earth orbit (GEO) satellites [16]. GDV errors for BDS B1 frequency can reach up to 1.5 m

for BDS-2 MEO satellites, compared to 0.8 m for BDS-2 IGSO satellites (see Fig. 4 in [14]). The new generation of BDS-3 satellites do not exhibit the same large GDV errors, and are comparable to other constellations (less than 20 cm) [14, 17].

GDV depends on the quality of the receiver antenna, and it can be characterized using elaborate antenna calibrations (see Fig 1. in [14] and [18]). It was shown that for a combination of low-cost antennas and low-elevation satellites, the GDV error could be as much as 6 times that of high-quality antennas [19].

For PPP processing using high-quality antennas, this error is sometimes omitted because its magnitude is typically smaller than, or comparable to, code noise magnitude. Hence, GDV errors can be absorbed into code measurement noise. However, as already mentioned, care should be taken when using measurements from low-quality antennas and low-elevation satellites, since GDV errors are more pronounced in this situation.

2.3.5 Phase Errors

2.3.5.1 Phase Center Offset/Variations

GNSS signals are emitted and received at the phase center of the antenna. For both satellites and receiver, the location of the phase center is not fixed (it can be either inside or outside the antenna body), and it depends on the signal's line-of-sight (LOS) vector. The phase center is anisotropic, i.e. it is a function of the signal's frequency, azimuth, and elevation [20]. In precise applications, the phase center needs to be accurately modeled since phase measurements at the receiver are made at the fixed location called the Antenna Reference Point (ARP). The anisotropy of the phase center is modeled by a constant offset of the phase center from the ARP is called the Phase Center Offset (PCO). Additionally, for

each satellite there is a corresponding variation called the Phase Center Variation (PCV) which depends on the signal's frequency, in addition to the elevation and azimuth. The combined range errors $\zeta^s_{r,j}$ induced by both satellite and receiver PCO/PCV can amount to a few centimeters up to 30 centimeters,

Parameters for satellite PCO/PCV corrections for all constellations are estimated and readily provided by various sources such as the IGS[3]. For receivers, the ARP is usually a point well-defined by the manufacturer, and given the azimuth, elevation, and frequency of a signal, the vector for receiver PCO/PCV can be approximated to determined the corresponding error.

2.3.5.2 Phase wind-up

The phase wind-up errors, ψ^s_r, depend on relative orientation of the receiver and satellite antennas, and the direction of the LOS vector. GNSS satellites usually emit electromagnetic waves with right-hand circular polarization (RHCP), containing two orthogonal sine wave signals with a 90° relative phase shift.[4] When a receiver antenna rotates, the electric field's relative orientation shifts, which shifts in the measured carrier phase, i.e. causes a range error. For instance, a full 360° rotation of the receiver antenna about its boresight results in a 1 cycle shift in the measured carrier phase for all signals (for GPS L1 frequency, this corresponds to 19 cm range error.) Similarly, if the satellite's antenna rotates, the receiver measures the relative orientation shift as phase range error due to a carrier phase shift.

[3]↑The antenna information for both satellites and receivers used in the IGS network is regularly updated and the most recent is provided at: https://files.igs.org/pub/station/general/igs20.atx

[4]↑In practice, a portion of the emitted signals has left-hand circular polarization (LHCP) due to reflection, but this portion is only significant to users outside a 16° cone of the transmitting (satellite) antenna's boresight. Earth-bound and low-Earth orbit users are *within* this cone, and need not worry about the effects of LCHP [20].

For a stationary receiver, the phase wind-up is mainly due to *continuous* satellite maneuvers performed to keep the satellite solar panels pointed towards the sun for maximum power. During these maneuvers, GPS satellites have a maximum yaw rate of $0.2^{\circ}/s$. However, satellites can also perform abrupt rotation maneuvers. If the satellite and receiver orientations are known, the phase wind-up correction can be modeled and the resulting errors corrected [20]. Range errors due to the phase wind-up from transmitting antennas is different for all satellites but similar for signals of different frequencies from the same satellite. As a result, they can be removed by inter-frequency differencing.

For moving receivers, phase wind-up errors can be due to a combination of satellite maneuvers and receiver dynamics. In these cases, the phase wind-up errors are absorbed into the phase ambiguity, as they're difficult to model if the satellite and receiver orientations are not precisely known. Range errors due to the phase wind-up from receiver antennas is the same for all signals, and they can be removed by between-satellite differencing.

2.4 Site Displacement Errors

The Earth experiences various loading forces that result in periodic deformations of the Earth surface. The loading come from various sources, the primary of which are: solid Earth tides and ocean tides [21]. Other negligible loading effects are due to polar tides resulting from centrifugal polar motion, atmospheric pressure loading which is (stronger at the equator, negligible at the poles), and surface hydrology (resulting from accumulation of ground water or snow build up, for example). The effects of these deformations on the computed coordinates of GNSS receiver on the Earth's surface are subtracted from the

measured range as a displacement error ϱ.

The same lunar and solar gravitation forces that generate ocean tides are strong enough to deform the "solid" Earth, resulting in the so-called solid Earth tides. These tidal deformations consist of a permanent latitude-dependent component, and a periodic part with largely diurnal and semi-diurnal periods The vertical and horizontal displacements resulting from these deformations can be represented by spherical harmonics [22, 23]. The solid Earth tidal corrections can reach magnitudes of up to 30 cm and 5m in the height and horizontal directions, respectively. Receivers separated by baselines of up to 1000 km have nearly identical deformations due to solid Earth tides.

Ocean loading is due to the ocean tides periodically moving the weight of the ocean water. This periodic weight redistribution results in the deformation of the Earth's crust, which is not completely rigid. These deformations are more localized to coastal areas and negligible for areas far from coastlines (more than 1000 km inland). Similar to those from solid Earth tides, ocean loading effects have diurnal and semi-diurnal periods; however, they have no significant permanent part. Areas most affected by ocean loading are near shallow seas with large tidal amplitudes, such as the North Sea, seas around Indonesia, and the Hudson bay in Northern America. Ocean loading effects are at least an order of magnitude smaller than solid Earth tides, even for receivers on the coastlines. When the atmospheric delay or clock errors are estimated, the ocean loading effects have to be taken into account. Otherwise, these effects will be absorbed into the estimates of atmospheric delays and receiver clocks [24]. Corrections for ocean loading can be accurately derived from global and regional ocean tide models, whose parameters can be obtained from services such

Displacement source	Magnitude [cm]		Period
	Vertical	Horizontal	
Solid Earth tides	30.0	5.0	Diurnal, semi-diurnal[a]
Ocean loading	5.0	2.0	Diurnal, semi-diurnal
Polar motion/tides	2.5	0.7	∼430 days
Atmospheric pressure loading	2.0	0.3	Latitude-dependent[b]

[a] Earth tides have a latitude-dependent, non-periodic component.
[b] Low-latitude regions show strong annual and semi-annual periods, whereas mid-latitude regions show periods of 5–10 days [25]. Pressure loading in polar regions is nearly non-existent.

Table 2.2: Periodic site displacement sources, the maximum magnitude of their amplitudes, and their corresponding periods.

as the one provided electronically by researchers at Chalmers University of Technology[5].

Table 2.2 shows the most significant sources of periodic site displacements, their maximum magnitudes in vertical and horizontal directions, and their corresponding periods. The values in Table 2.2 are gathered from Chapter 7 of [23] and from [25].

2.5 Multipath and Receiver Measurement Noise

Multipath errors, $m^s_{r,\rho_j}, m^s_{r,\phi_j}$, are encountered when the received signal is a composite of LOS and non line-of-sight (NLOS) components. The NLOS components might be caused by reflection and/or diffraction of the original signal onto nearby surfaces, which is then combined with the direct signal, thus introducing range errors. Code multipath errors are significantly larger than phase multipath errors. The former can be up to tens of meters, whereas the latter can have a maximum of $\lambda/4$ for signals with wavelength λ.

The receiver measurement noise refers to random errors induced by the interaction of the GNSS radio waves with various electrical components in the signal processing hardware

[5] Free Ocean Tide Loading Provider, provided by M.S. Bos and H.-G. Scherneck. Available at: http://holt.oso.chalmers.se/loading/index.html. A convenient summary of the eighteen ocean tide models offered by the service is provided in Table 7.5 of [23].

(antenna, cables, and receiver itself). These random errors are background noise with a certain power level. A ubiquitous measure of the received GNSS carrier signal's strength is the carrier-to-noise-power-density ratio C/N_0 (in dB-Hz), which is the ratio of power levels of the carrier signal (C) to the background noise (N_0) per unit bandwidth. The measurement noise of both the pseudorange and phase is assumed to be zero-mean Gaussian noise, whose spectral density is inversely related to the C/N_0; the higher the C/N_0, the lower the measurement noise. Code noise has decimeter-level standard deviation in modern receivers tracking modernized signals, whereas the phase noise is typically millimeter-level.

Chapter 3

Residual Atmospheric Delay Errors

This chapter explores characteristics of the range errors which arise when GNSS signals pass through the Earth's atmosphere. First, the chapter shows that signals' delay through the neutral atmosphere can be quantified by local meteorological measurements such as pressure, temperature, and humidity. Second, it shows that the delay through the ionized atmosphere has spatio-temporal variations with diurnal regularity. For both cases, the chapter discusses various methods of compensating for the delays. Finally, the chapter discusses how after compensation, there are local residual atmospheric delays which have to be estimated for high precision GNSS applications presented in later chapters.

3.1 Tropospheric Delay

The tropospheric delay error generally refers to the range error resulting from the delay of GNSS signals as they travel through the neutral atmosphere from the satellite to the receiver. The neutral (non-ionized) atmosphere is made up of both the troposphere and the

stratosphere. It is a non-dispersive medium, i.e. signals of different frequencies are delayed by the same amount.

The troposphere is the layer of the atmosphere closest to ground-based receivers. It can reach a height of up to of 20 km, depending on latitude (it is thickest at the equator and thinnest at the poles), season (highest in winter, lowest in summer), and time-of-day [26,27]. It contains around 99% of the atmosphere's water vapor, moist gases, and aerosols (such as dust particles, salt grains, and ice crystals); and for that reason, it is where most of the weather happens [28]. About 75% of the *tropospheric delay* comes from the troposphere. The CO_2 gas is the only constituent of the troposphere which has diurnal variations due to human activity. However, gases due to human activity, such as CO_2, have negligible effects on the tropospheric delay. Although CO_2 is mostly concentrated a few kilometers from ground-level in the tropospheric layer, it only accounts for 0.033% [29] of the delay error.

The stratosphere sits on top of the troposphere and extends to a distance of up to 50 kilometers. It is composed of a mixture of dry gases, of which more than 99.9% is nitrogen (N_2), oxygen (O_2) and Argon (Ar) [30] by volume[1]. Most of the stratosphere is both spatially and temporally homogeneous and constant. Hence, the stratosphere can be accurately modeled with the use of both empirical models and current meteorological observations.

From the above discussion, the tropospheric delay term T_r in eqns. (2.1)–(2.4) is divided into two distinct parts:

$$T_r^s = T_{r,d}^s + T_{r,w}^s, \tag{3.1}$$

[1]↑The stratosphere also contains the well-known ozone layer (O_3), which, incidentally, has no effect on GNSS signal propagation.

where $T_{r,w}$ is the wet component and $T_{r,d}$ is the dry component. Both the troposphere and stratosphere contribute to each of the above components.

As in most of the GNSS literature, the *tropospheric delay* error T_r^s used in this dissertation refers to the slant version of the Zenith Total Delay (ZTD), defined as:

$$T_r^s = MF(t, \alpha, \beta)\, ZTD, \tag{3.2}$$

where $MF(t, \alpha, \beta)$ is a mapping function that depends on time (t), satellite azimuth (α), and elevation (β). Subsection 3.1.3 will explore different kinds of mapping functions used. The ZTD is in turn also separated into two distinct components, namely the Zenith Hydrostatic Delay (ZHD) and Zenith Wet Delay (ZWD): $ZTD = ZHD + ZWD$. The tropospheric delay in eq. (3.1) can then be expressed as

$$T_r^s = MF_h(\beta)\, ZHD + MF_w(\beta)\, ZWD, \tag{3.3}$$

where $MF_h(\beta)$ and $MF_w(\beta)$ are the dry and wet mapping functions, respectively. The next subsections discuss how each of these components is estimated using various empirical and numerical weather models.

3.1.1 Dry Component (ZHD)

The ZHD is also referred to as the dry component of the tropospheric delay, since it is caused by dry gases in the troposphere and stratosphere. Because the ZHD depends on the quantity of these gases (their weight in all atmospheric layers), it can be precisely estimated using on-site surface pressure measurements. The first empirical model of the ZHD using pressure measurements is the *Saastamoinen* model, which was proposed in [31] and was

later improved by [32]:

$$ZHD = \frac{0.0022768\,p}{1 - 0.00266\,\cos 2\varphi - 0.28 * 10^{-6}\,h_e}, \tag{3.4}$$

where p is the air pressure, φ and h_e are the receiver's latitude and ellipsoidal height, respectively. There are three sources that could provide pressure values: (1) local/on-site pressure measurements, (2) numerical meteorological (weather) pressure models, and (3) empirical pressure models.

On-site pressure measurements would be ideal, but are impractical for most applications. Empirical pressure models capture historical trends, since they are based on historical empirical observations. The Hopfield model in eq. (3.5) is the most widely used empirical pressure model [33, 34]. It is expressed as:

$$p = 1013.25 \left(\frac{T_k - \alpha_T\,h_0}{T_k}\right)^{\frac{g}{\alpha_T\,R_d}}, \tag{3.5}$$

where $T_k = 291.16$K is the atmospheric temperature at sea level, $\alpha_T = 4.5$ K/km is the normal temperature lapse rate with elevation, h_0 is the receiver's orthonometric height in meters, $g = 9.7867$ m^2/s is the Earth's gravity at the surface, and $R_d = 287.0464$ J/K/kg is the gas constant for dry air. There are more sophisticated empirical models such as UNB3m, which is based on meteorological parameters (pressure, temperature, humidity, temperature lapse rate, and water vapor pressure height factor) at five latitude bands which are symmetric with respect to the equator [35]. It should be noted that, regardless of their sophistication, empirical models cannot fully account for local spatiotemporal variations.

Pressure models based on numerical meteorological models are more accurate than empirical models, and hence preferable for high-precision applications. The first of such models is the Global Pressure and Temperature (GPT), which is based on spherical harmonics of

up to degree and order nine (resolution of $2.5^\circ \times 5^\circ$) [36]. GPT was subsequently improved by GPT2w [37] and GPT3 [38], which do away with spherical harmonics in favor of a grid model of resolution up to $1^\circ \times 1^\circ$. These models require information about time and location, and their pressure data is based on numerical weather models from the European Centre for Medium-Range Weather Forecasts (ECMWF).

3.1.2 Wet Component (ZWD)

The ZWD is also referred to as the wet component of the tropospheric delay. The ZWD is harder to estimate using only surface-based measurements, since it depends on dynamic weather-related phenomena happening kilometers off the surface. Several empirical models have been proposed to obtain an approximate estimate of the ZTD. These models are mostly based on water vapor pressure and temperature. For example, the aforementioned *Saastamoinen* model [31] uses the ideal gas law to derive a simplified estimate:

$$ZWD = 0.0022768 \ (1255 + 0.05\, T_s)\, \frac{p_{ws}}{T_s}, \tag{3.6}$$

where T_s is the on-site surface temperature and p_{ws} is the on-site measurement of partial water vapor pressure at the Earth's surface. The partial water vapor pressure is related to the humidity via the following relationship:

$$p_{ws} = h_{rel}\,(p_{sat}), \tag{3.7}$$

where h_{rel} is the relative humidity and p_{sat} is the partial pressure of saturated air. *Relative humidity* measures the current amount of water vapor in the air as a percentage of the saturation water vapor, i.e. the current humidity as a percentage of the maximum humidity at the same temperature.

Conditions	h_{rel}	Measurements		Zenith Delay	
		p [mbar]	T_s [K]	ZHD [m]	ZWD [m]
Cold + Dry	10%	950 (940)	230 (220)	2.16 (2.14)	0.00 (0.00)
Hot + Humid	90%	1050 (1040)	310 (300)	2.39 (2.37)	0.75 (0.40)

Table 3.1: Magnitudes of Zenith Total Delay (ZTD) components under extreme environmental conditions.

An improved model for ZWD was proposed in [28] as:

$$ZWD = 10^6 \left(k_2' + \frac{k_3}{T_s} \frac{R_d \left(p_{ws} \right)}{g_m \left(\lambda + 1 \right)} \right), \tag{3.8}$$

where k_2' and k_3 represent empirically determined constants, $g_m = 9.80665$ m^2/s is the mean gravity, and λ is the water vapor decrease factor. The model in eq. (3.8) is used by both GPT2w and GPT3.

Table 3.1 shows examples of the dry and wet components of the zenith tropospheric delay, under extreme conditions[2]. The values in parentheses are for the scenario in which both the temperature and air pressure have been simultaneously decreased by 10 K and 10 mbar, respectively. Note that under very dry conditions, the ZWD is very small due to low quantities of water vapor in the atmosphere. This implies that ZWD depends heavily on the humidity. Under humid conditions, the decrease in air pressure and temperature leads to a drastic decrease in ZWD for the same relative humidity.

3.1.3 Mapping Functions and Gradients

The mapping functions for the dry and wet components in eq. (3.3) depend on the refractivity of the neutral atmosphere, and can be thought to represent the thickness thereof [39]. Similar to ZHD and ZWD, the mapping functions can be estimated using both empirical

[2]↑The hot and humid conditions in the table can be experienced regularly during summers in tropical to mid-latitude regions, whereas the cold and dry conditions can be experienced in polar and sub-polar regions.

models and numerical weather models. There has been simplistic empirical models based only on satellite elevation, such as those proposed in [26, 40]. More sophisticated models take into account other aspects of the atmosphere. For example, the most widely used is formulated in [39] as:

$$MF(\beta) = \frac{1 + \left(\frac{a}{1+\frac{b}{1+c}} \right)}{\sin\beta + \left(\frac{a}{\sin\beta + \frac{b}{\sin\beta+c}} \right)}, \tag{3.9}$$

where a, b, c are parameters which depend on the integrals of refractivity through the neutral atmosphere. These parameters can be determined empirically (e.g. [41]) or using numerical weather models (e.g. [27,42–44]). In more precise models such as Vienna Mapping Functions 3 (VMF3) [38], these parameters are determined using weather models from ECMWF for both the dry and wet components, i.e. the parameter a becomes a_h and a_w for MF_h and MF_w, respectively.

The models discussed so far are global, empirical or otherwise. Recently, there has been interest in developing regional models for high-precision applications. Examples include AFRC-Trop [45] over the African region, and an improved GPT2w over the Chinese region [46].

The discussion of mapping functions has so far assumed that the tropospheric delay has azimuthal symmetry around the receiver, i.e. the delay is the same for all azimuth angles α. Due to local and regional climatic and weather conditions, atmospheric delays at a constant elevation angle will have slight variations with the horizontal (azimuth) direction [47]. To account for these variations in high precision positioning applications, azimuth-dependent directional gradients are necessary. Horizontal (North and East) gradients for both dry and wet components can be determined empirically [47,48], or through ray-tracing

using meteorological models as the mapping functions (e.g. GRAD gradients are provided alongside VMF3 mapping functions [38]).

For applications requiring high precision, the data necessary to compute the total tropospheric delay T_r in eq. (3.3) is often disseminated as a single product. An example is the VMF3/GPT3 product [38, 49], which provides modeled: surface pressure, temperature, water vapor pressure, and other parameters needed for computing the ZHD and ZWD. It also includes parameters to determine the corresponding dry and wet mapping functions, along with their horizontal gradients. This product is used later in Chapters 5 and 6.

3.2 Ionospheric Delay

The ionosphere is the upper layer of the Earth's atmosphere, which lies above the troposphere and stratosphere, and ranges from 80 km to 1500 km. Unlike the neutral atmosphere, the ionosphere is a dispersive and anisotropic medium, i.e. it affects GNSS signals differently depending on their frequency. This means that GNSS electromagnetic waves experience various levels of absorption, refraction, scattering, and polarization shift as they propagate through the ionosphere. The effects of absorption and polarization are negligible for GNSS signals due to their long wavelengths and circular polarization, respectively [20].

The refraction causes signal delays, carrier advances, and ray bending, and it typically accounts for most of the ionospheric effect on GNSS signals [20]. The amount of signal delay through the ionosphere depends on Total Electron Content (TEC), which is defined as the total number of free electrons in a column of 1 m^2 cross-section. It is expressed as:

$$TEC = \int_r^s N_e(l)dl, \tag{3.10}$$

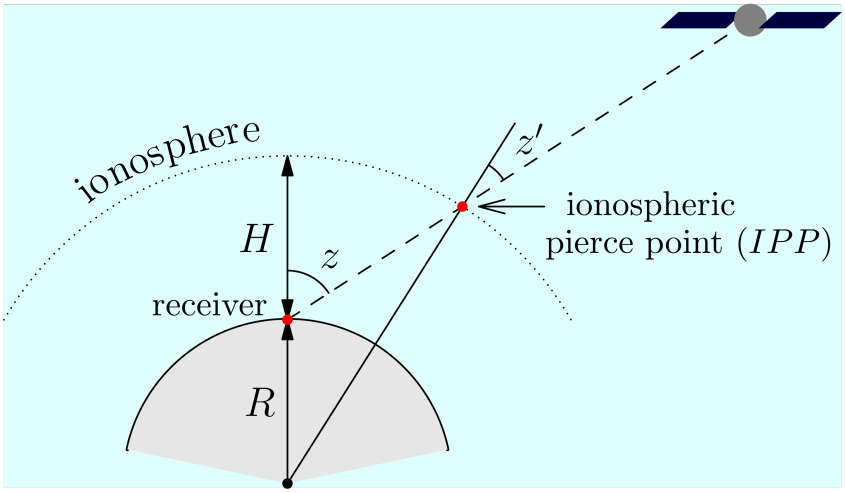


Figure 3.1: Illustration of the ionospheric thin-shell model.[3]

where $N_e(l)$ is the total number of free electrons along a signal path l from satellite s to receiver r. The TEC is measured in Total Electron Content units (TECu), where 1 TECu $= 10^{16}$ electrons/m^2. The number of free electrons in the ionosphere is affected by ionizing solar radiation due to solar winds and other space weather phenomena. As a consequence, the dynamic behavior of the ionospheric delay is closely related to space and solar weather phenomena.

For GNSS applications, the ionosphere is typically modeled as being compacted into a thin shell at a specific height above the Earth's surface, as shown in Fig. 3.1. The TEC at the point where the signal pierces the thin shell is referred to as Vertical TEC (VTEC). The TEC in a column along a signal path from the satellite to the receiver is referred to as Slant TEC (STEC), and it used to determine the ionospheric delay. The VTEC and STEC are related through the following elevation-dependent mapping function:

$$STEC = M(\beta)\big(VTEC\big), \tag{3.11}$$

[3]Illustration by Wikimedia user YuryKirienko, distributed under a CC BY-SA 4.0 DEED license: https://en.m.wikipedia.org/wiki/File:IPP.svg.

where $M(\beta)$ is an elevation-dependent mapping function defined as:

$$M(\beta) = \frac{1}{\sqrt{\left(1 - \left(\frac{R_E \cos\beta}{R_E + H}\right)^2\right)}} \tag{3.12}$$

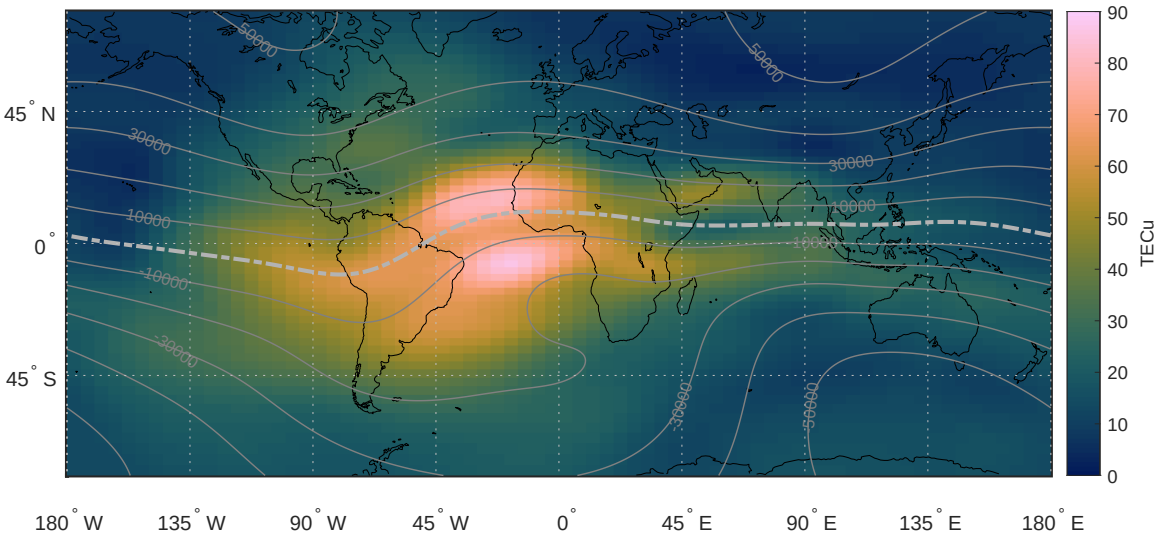
$$= \frac{1}{\cos z'}, \tag{3.13}$$

where β is the satellite elevation, R_E is the Earth radius, H is the thin-shell height, and z' is the thin-shell zenith angle at the pierce point. The total ionospheric delay in eq. (2.1)–(2.4) can be computed terms of STEC as:

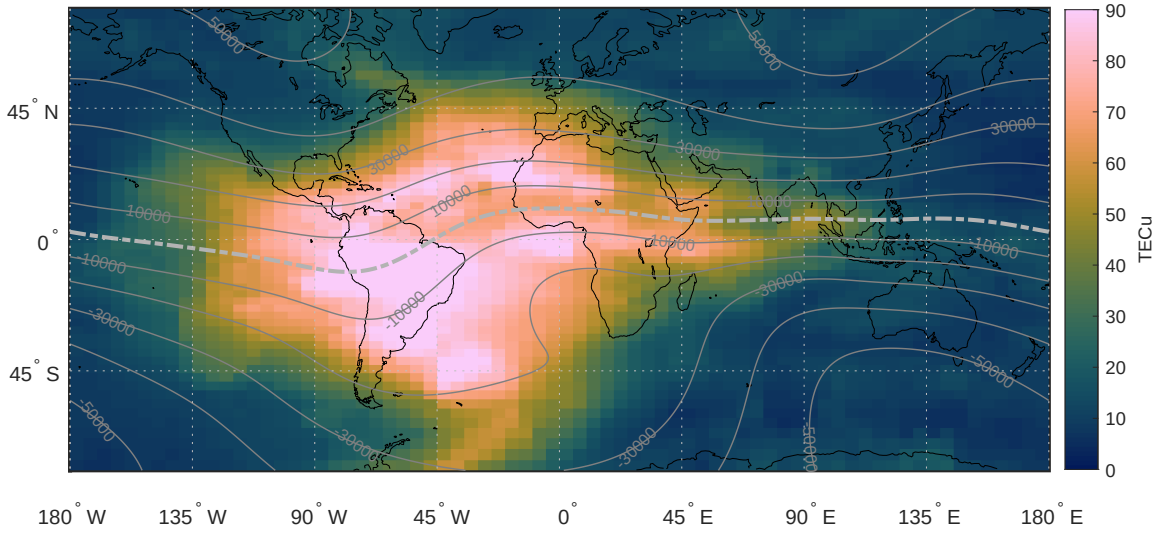
$$I^s_{r,j} = \frac{40.3}{f_j^2}(STEC), \tag{3.14}$$

where f_j is the signal's frequency. For GPS L1 frequency, 1 TECu of VTEC accounts for an approximately 16.2 cm delay when the satellite is directly above a receiver located at the equator, i.e. $z' = 0°$ or $\beta = 90°$.

Besides signal delay, ionospheric activity can also have other adverse effects on GNSS signals. Solar geomagnetic activity causes plasma irregularities, which lead to a scattering effect that results in intermittent variations in GNSS signal amplitude and phase, referred to as *ionospheric scintillation*. The scintillation causes increased measurement errors, carrier-phase cycle slips, and loss-of-lock for the carrier tracking loop [50]. Intense scintillations are more prevalent in high and low-latitude regions and can last for hours, but they can also occur unpredictably in short bursts at any location on the globe [51]. Consequently, for applications requiring high precision, it is important to eliminate or estimate the ionospheric delay error using *local* GNSS observations.



(a) VTEC map (with WMM2015 contours) for quiet ionosphere conditions on January 1, 2023, at 18:00 UTC.



(b) VTEC map (with WMM2020 contours) for the St. Patrick's Day geomagnetic storm on March 17, 2015, at 18:00 UTC.

Figure 3.2: Examples of final IGS VTEC maps during quiet typical ionospheric conditions, and during a major geomagnetic storm. The gray undulating lines are contours of the Earth's magnetic field down component, as modeled by the World Magnetic Model (WMM).

3.2.1 Estimation from Code and Phase Measurements

The ionospheric delay $I^s_{r,1}$ can be estimated from either phase or code measurements. It is preferred to estimate it using code measurements, due to the fact that one does not have to estimate the phase ambiguities, which can be discontinuous. Subtracting eq. (2.3) from eq. (2.1), we get the following observable:

$$\begin{aligned}\rho^s_{r,1} - \rho^s_{r,2} &= I^s_{r,1} - \gamma I^s_{r,1} - c\left((\delta t_{r,1} - \delta t_{r,2}) - (\delta t^s_1 - \delta t^s_2)\right) + \Delta\epsilon \\ &= I^s_{r,1} - \gamma I^s_{r,1} - \delta b_{r,12} + \delta b^s_{12} + \Delta\epsilon \\ &= \left(\frac{f_2^2 - f_1^2}{f_1\, f_2}\right) I^s_{r,1} - \delta b_{r,12} + \delta b^s_{12} + \Delta\epsilon, \end{aligned} \tag{3.15}$$

where $\Delta\epsilon$ encapsulates the combined time-varying effects of group delay variations, multipath, and random noise; $\delta b_{r,12}$ and δb^s_{12} are receiver and satellite DCBs, respectively. (In the above and subsequent analyses, we assume CDMA signals, which implies the absence of ICBs: $\Delta d^s_{r,j}, \Delta\delta^s_{r,j}$.) Rearranging eq. (3.15), the ionospheric delay is then modeled as:

$$I^s_{r,1} = \frac{f_1\, f_2}{f_2^2 - f_1^2}\left(\rho^s_{r,1} - \rho^s_{r,2} + \delta b_{r,12} - \delta b^s_{12} - \Delta\epsilon\right), \tag{3.16}$$

The model in eq. (3.16) contains the so-called ionosphere-free DCB from both the satellite and the receiver. When both the receiver and satellite DCBs are known, then eq. (3.16) is a noisy estimate of the ionospheric delay, whose precision is dictated by that of the code measurements.

3.2.2 Global Models

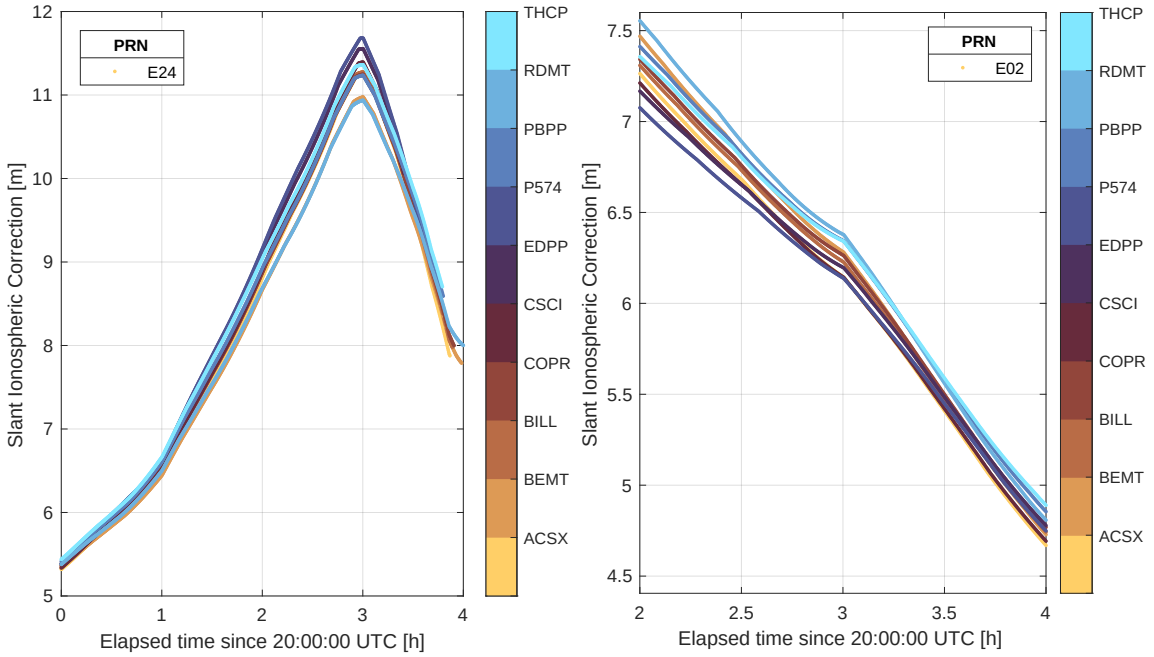
The estimate of the ionospheric delay can be obtained from external sources. One such source is the Global Ionospheric Model (GIM) VTEC maps from the International GNSS

Service (IGS) [52]. Fig. 3.2 shows examples of the GIM VTEC map at 18:00:00 UTC for: (a) quiet ionospheric conditions, and (b) highly active geomagnetic storm conditions. The gray lines are the main down component (Z) of the Earth's magnetic field from the World Magnetic Model (WMM) [53] in contour intervals of 10,000 nT, with the dashed middle line being the magnetic equator of 0 nT. These lines illustrate the VTEC's dependence on the Earth's magnetic field. Fig. 3.2a shows the VTEC map from IGS on January 1, 2023, which had relatively quiet ionospheric conditions. Fig. 3.2b shows the VTEC map from IGS on March 17, 2015 during a geomagnetic storm, the so-called St. Patrick's Day storm whose impact on the ionosphere was detailed in [54]. Note that, compared to Fig. 3.2a, VTEC values in Fig. 3.2b are: (1) much larger in tropical latitudes (up to more than three times), (2) irregular (blobs in mid-latitudes), and (3) rapidly changing locally (which is not captured by the two-dimensional map). As a consequence, for high-precision applications during conditions depicted in Fig. 3.2b, the rapidly changing ionosphere must be estimated locally, so as to maintain the same positioning performance as in quiet conditions.

The typical spatial resolution of GIM maps is 2.5° latitude by 5° longitude, and the typical temporal resolution is 2 hours. For some applications, this resolution is inadequate to handle situations in which the ionospheric delay is rapidly changing in space and time, as depicted by in Fig. 3.2b.

3.3 Local Residual Delay Errors

Local residual delays are defined as the remaining atmospheric delay errors after applying corrections from atmospheric global models. These residual delay errors are due to the



(a) GAL E24 E1 correction

(b) GAL E24 E1 correction

Figure 3.3: GIM slant ionospheric corrections for the first frequency of two GAL satellites. Each curve represents the correction for one receiver, for a total of ten receivers in the same local area.

fact that global tropospheric and ionospheric corrections cannot capture local atmospheric conditions. To achieve maximum precision, the residual delay errors need to be estimated as unknowns.

As discussed in Section 3.1, the ZHD component of the tropospheric delay error is corrected by global models with very good accuracy, whereas the ZWD cannot be accurately corrected since it depends on local atmospheric conditions. As result, the residual ZWD contributes the majority (and in most cases all) of the residual tropospheric delay error. The magnitude of the local residual delay will depend on the accuracy of the global corrections source, as well as the local environmental conditions. As shown in Table 3.1, the ZWD (and consequently its residual) has the largest magnitude under hot and humid conditions. On the other hand, residual ZWD has a negligible magnitude under cold and dry conditions.

Similar to residual ZWD, the residual ionospheric delay errors are due to the inability of global models to capture local spatiotemporal variations of VTEC. As discussed in Section 3.2.2, interpolation is necessary when applying ionospheric delay corrections because GIM VTEC maps have limited temporal and spatial resolutions. Fig. 3.3 shows the slant ionospheric corrections for two GAL satellites obtained from a GIM model. The modeled corrections are for the E1 frequency of GAL. The temporal interpolation artifacts are visible at around 23 h (UTC) in both subplots. These artifacts are marked by abrupt changes in the modeled delay, which cannot correspond to physical behavior of the atmosphere. Since the particular GIM model has a temporal resolution of two hours, the abrupt changes are visible around the hour marks in both Fig. 3.3a and Fig. 3.3b.

Chapter 4

GNSS Positioning Services

GNSS positioning services refer to categories of approaches used to compute the user position coordinates using GNSS signals. Positioning services can be divided into three main groups, depending on the techniques used to correct measurement errors. Namely: *standard positioning*, which only uses information transmitted from the satellites and that which is already available to the receiver; *relative positioning*, which takes advantage of information from nearby reference stations (or other receivers); and *precise point positioning*, which uses external information disseminated from a global network of reference stations. This dissertation is primarily concerned with precise point positioning (PPP); and this chapter discuses other positioning services to provide context on the PPP discussion.

4.1 Standard Positioning

The Standard Positioning Service (SPS) is the simplest, most ubiquitous, and inexpensive GNSS positioning service. It consists of approaches which use only pseudorange

measurements from a single receiver to determine the receiver's position. The pseudorange measurements can be from a single or multiple GNSS constellations, using single or multiple frequencies. To correct some of the errors discussed in Table 2.1, SPS uses information from broadcast messages, or other empirical models. These models are based on ephemeris information that is broadcast by the GNSS satellites. SPS also uses coarse broadcast information about global atmospheric conditions, which has been obtained from external sources such as global meteorological and empirical models discussed in Chapter 3.

The specified accuracy for SPS is meter-level [55], mainly due to uncompensated errors in eqs. (2.1)–(2.3). For example, the 95% horizontal position accuracy of 3.0 meters has be observed for various GPS SPS receivers located in North America [56]. Users of multiple GNSS systems and multiple frequencies can expect horizontal accuracy of 1-3 meters [57].

4.2 Relative Positioning

Relative positioning services compute the user's coordinates relative to the known coordinates of another GNSS receiver. This requires the exchange of measurements between the user and other receiver(s). This section provides an overview of the three types of relative positioning approaches.

4.2.1 Differential GNSS

Differential positioning was initially developed to sidestep the effects of Selective Availability (SA), i.e. the purposeful degradation of GPS accuracy in the early days of satellite navigation [58]. After SA was turned off in May 2000, DGPS prevailed because it retained its benefits of reducing clock, hardware, and atmospheric biases. Due to these benefits,

the DGPS technique remains important to this day. With the emergence and evolution of multiple operational GNSS systems, DGPS has evolved into Differential GNSS (DGNSS).

The simplest form of DGNSS involves two receivers, the base and the rover, separated by a certain baseline (distance) which are able to exchange their code observations. The base receiver's position is precisely known. The rover position is then estimated as a *differential* offset from the known base position. For receivers $r = \{A, B\}$, the between-receiver code observations for satellite p and frequency j are modeled using eqs. (2.1) and (2.3):

$$\begin{aligned} \rho^p_{AB,j} &= \rho^p_{A,j} - \rho^p_{B,j} \\ &= R^p_{AB} + c\,(dt_{AB} + \delta t_{AB,j,\rho}) - \gamma_j\, I^p_{AB,j} + T^p_{AB} \\ &\quad + m^p_{AB,j,\rho} + \eta^p_{AB,\rho}\,, \end{aligned} \tag{4.1}$$

where dt_{AB} is the differential receiver clock error, $\delta t_{AB,j,\rho}$ is *a* differential code bias between two receivers[1], γ_j is the ionospheric coefficient for frequency j, $I^p_{AB,j}$ and T^p_{AB} are the *differential* ionospheric and tropospheric biases, $m^p_{AB,j,\rho}$ is the differential code multipath, and $\eta^p_{AB,\rho}$ is the differential code noise. It should be noted that the inter-channel code biases $\Delta d^s_{r,j}$ have been omitted from the model in eq. (4.1). As discussed in Section 2.3.3, these biases are absent for CDMA signals. These biases should be appropriately modeled when processing GLONASS signals.

The atmospheric delays $I^p_{A,j}$ and T^p_A are spatially correlated, and their corresponding differential biases will always be much smaller than the corresponding individual delays. For short baselines, the differential ionospheric bias $I^p_{AB,j}$ will be smaller, increasing with the baseline length. The magnitude of the differential ionospheric bias varies depending on

[1]↑This differential bias $\delta t_{AB,j,\rho}$ is not to be confused with *the* differential code biases (DCBs), as discussed in 2.3.2.1. The latter is a difference between code biases of two different frequencies on the same receiver, while the former is a difference between code biases of the same frequency on two different receivers.

the location on Earth: it is smaller in mid-latitudes and larger in tropical regions [59]. It also depends on the stage of the solar cycle, since ionospheric delay is highly correlated to solar activity, as discussed in Section 3.2. For instance, the worst-case $|I^p_{AB,j}|$ for a mid-latitude location with 100 km of baseline will be around 0.1 m during the minimum stage of the solar cycle, and rise up to 1.1 m during the maximum stage of the solar cycle. For a tropical location, the numbers will be 0.4 m and 6 m, respectively [59, 60].

Similarly, the size of the differential tropospheric bias T^p_{AB} depends on baseline length. Unlike its ionospheric counterpart, T^p_{AB} will also different on other factors such as: the atmospheric conditions and height difference between receivers. As discussed in eq. (3.1) of Section 3.1, the tropospheric delay T^p_A consists of a dry component $T^p_{A,d}$ which can be corrected accurately, and a wet component $T^p_{A,w}$ which depends on local conditions. Consequently, after the dry component has been corrected, T^p_{AB} mostly consists of the differential wet component. For instance, for a 100 km baseline after dry component correction, $|T^p_{AB}|$ varies between 0.0 m for identical cold and dry conditions, and can rise up to 0.19 m for identical hot and humid conditions. If the atmospheric conditions are not identical, $|T^p_{AB}|$ can rise dramatically to over 1 m [59].

The receiver noise, multipath and other unmodeled residual biases cannot be estimated, and should be minimized by other means, such as choosing a reference station with small multipath, or using a receiver with better noise characteristics. The receiver differential clock and hardware biases are not spatially correlated and hence cannot be reduced or cancelled. These biases have to either be estimated as unknown parameters, or be eliminated via an additional between-satellite differencing to form double-differenced observations.

After the between-receiver difference of eq. (4.1), the between-satellite difference pro-

duces the code *double-difference* for receivers $r = \{A, B\}$ and satellites $s = \{p, q\}$ on frequency j, which is modeled as:

$$\begin{aligned} \rho^{pq}_{AB,j} &= \rho^{p}_{AB,j} - \rho^{q}_{AB,j} \\ &= R^{pq}_{AB} - \gamma_j \, I^{pq}_{AB,j} + T^{pq}_{AB} \\ &\quad + m^{pq}_{AB,j,\rho} + \eta^{pq}_{AB,\rho}. \end{aligned} \tag{4.2}$$

The remaining double-difference biases $I^{pq}_{AB,j}$ and T^{pq}_{AB} can be ignored for short baseline applications with less stringent precision requirements.

After performing the double-differencing in eq. (4.2), the position of receiver B is then estimated as a "differential" offset vector from the known position of base receiver A. For long baselines (more than 100 km, for example) and other situations in which the differential atmospheric biases might not be eliminated, the state vector could be augmented to include either one (or both) of $I^{pq}_{AB,j}$ and T^{pq}_{AB} as unknowns. Further details of the estimation process for GNSS can be found in various references in the literature, such as Chapter 8 of [61] and other references therein.

4.2.2 Real-Time Kinematic Positioning

Real-time Kinematic (RTK) positioning is another name for phase-based DGNSS. It combines simultaneous code and carrier phase observations from a user receiver and a nearby base station to compute very precise position solutions with a near-instantaneous convergence of a few seconds [62].

For receivers $r = \{A, B\}$, the between-receiver phase observations for satellite p and

frequency j are modeled using eqs. (2.3) and (2.4):

$$\begin{aligned}\phi^p_{AB,j} &= \phi^p_{A,j} - \phi^p_{B,j} \\ &= R^p_{AB} + c\,(dt_{AB} - \delta t_{AB,j,\phi}) - \gamma_j\, I^p_{AB,j} + T^p_{AB} \\ &\quad + \lambda_j\, \widetilde{N}^p_{AB,j} + m^p_{AB,j,\phi} + \eta^p_{AB,\phi},\end{aligned} \tag{4.3}$$

where

$$\widetilde{N}^p_{AB,j} = N^p_{AB,j} + \frac{1}{\lambda_j}\left(\psi^p_{AB} - h_{AB,j}\right). \tag{4.4}$$

In contrast to its code counterpart in eq. (4.1), the phase difference is still ambiguous due to the ambiguity term $\tilde{N}^p_{AB,j}$.

The between-receiver, between-satellite phase double-difference for receivers $r = \{A, B\}$ and satellites $s = \{p, q\}$ on frequency j can then be modeled as:

$$\begin{aligned}\phi^{pq}_{AB,j} &= \phi^p_{AB,j} - \phi^q_{AB,j} \\ &= R^{pq}_{AB} - \gamma_j\, I^{pq}_{AB,j} + T^{pq}_{AB} \\ &\quad + \left(\lambda_j\, N^{pq}_{AB,j} + \psi^{pq}_{AB}\right) + m^{pq}_{AB,j,\Psi} + \eta^{pq}_{AB,\phi}.\end{aligned} \tag{4.5}$$

The double-difference ambiguities can be resolved using various methods such as the popular LAMBDA method [63, 64]. After the ambiguities $N^{pq}_{AB,j}$ have been resolved to their integer value, the double-differenced phase observation $\phi^{pq}_{AB,j}$ provides very precise range measurements. RTK Positioning solutions have been reported to reach mm-level accuracy for short and medium baselines, whereas cm-level accuracy can be reached for long baselines.

However, RTK has some limitations in real-time applications because:

1. It needs high bandwidth for transmitting and receiving large amounts of differential corrections data.

2. The service area of one reference (base) station is limited to a local area, which implies the need for deploying extensive infrastructure to service larger areas.

3. It requires a bidirectional communication link between the user receiver and the reference receiver (or the RTK service provider).

4.2.3 Network RTK

The single-base RTK technique discussed thus far has some drawbacks and limitations. The foremost limitation is the maximum base–rover baseline for which the distance-dependent biases can be reliably eliminated, which is typically 10–20 km. This limitation is addressed by using a network of multiple reference stations (bases) around the rover, thus extending RTK into Network RTK (NRTK). The use of NRTK allows bigger separation between bases, while still enabling successful double-difference ambiguity resolution. Typical commercial NRTK networks deliver the same performance as single-base, short-baseline RTK, with average base separation of 50 to 80 km [65], and even more than 100 km [66].

There are several approaches to NRTK processing, but all of them involve a fundamental four-step process:

1. Ambiguity resolution for reference stations. This AR differs from the one performed at the rover, since the precise position of each reference station is known. Additionally, a priori station analysis might provide estimates of multipath and receiver antenna characteristics, which are crucial to fast AR. Reference observations are used only after successful AR to ensure the quality of computed biases in the next step.

2. Modeling and computation of distance-dependent biases. After successful AR, phase observations from each station are used by the central Service Provider (SP) to estimate atmospheric delays, which are distance dependent. This is done primarily using interpolation between stations. Ionospheric and orbit biases are estimated per satellite, whereas tropospheric biases are estimated per station.

3. Generation of Virtual Reference Station (VRS) observations. After the successful estimation of biases, virtual observations are derived from real observations of a reference station and the computed biases. The reference station is often the one closest to the rover, but can also be chosen using other heuristics (e.g. quality of observations). These virtual observations form the so-called VRS, the computation of which represents the main difference of NRTK approaches. The most prominent techniques used to compute VRS are:

    - Direct VRS [67]: The rover sends its approximate position to the SP, which computes and sends back interpolated VRS observations.

    - Master-Auxiliary Concept (MAC) [68]: The rover obtains from the SP the observations of the *main* reference station, along with respective differences between auxiliary reference stations and main reference station. The rover interpolates the information to compute VRS observations.

    - Flächen–Korrekturparameter (FKP) [69]: Literally, area correction parameters. The rover obtains from the SP the observations of the *main* reference station, along with *area* coefficients for the distance-dependent biases. The VRS observations are generated by combining the main reference observations, and corrections

are obtained by applying the coefficients to the coordinate differences between the rover and the main station.

Recently, it has been proposed that VRS observations can also be generated without requiring any local infrastructure around the user. This involves leveraging global PPP corrections (discussed in Section 4.3.1) to synthesize OSR observations [70], thereby implementing a *global* NRTK.

4. Baseline processing and rover position computation. After VRS observations have been obtained, they are used to compute the rover position using the same methods as single-base RTK (discussed in Section 4.2.2).

Generating VRS observations incurs interpolation error, which leads to additional positioning error for NRTK, compared to similar processing for RTK using a very short-baseline. However, for well-maintained networks, the performance of ambiguity-resolved NRTK is comparable to that of short- to medium-baseline RTK, thus delivering the same performance with less density of reference stations.

4.3 PPP

On the one hand, PPP extends SPS whereby the broadcast information (e.g. satellite clocks and orbits) is replaced with respective precise estimates provided by global SPs. On the other hand, PPP is similar to RTK in that very precise carrier-phase observations are used together with the code observations, usually on two or more frequencies, which allows to either eliminate or estimate the ionospheric delays. The use of carrier phase observations also introduces unknown ambiguities, which need to be resolved to integers to take full

advantage of phase observations. The phase observation on each frequency of each satellite has an unknown integer ambiguity which, along with other residual errors, lead to PPP requiring lots of observations and a long time to converge to a precise solution.

Unlike relative and standard positioning techniques, PPP is also able to provide, in addition to precise coordinates, the precise estimates of receiver clocks, tropospheric delays, and ionospheric delays. For this reason, PPP is widely used for applications such as geodynamics and meteorology in the disciplines of remote sensing and Earth monitoring [71, 72].

Generally, PPP estimates four kinds of unknowns, namely: (1) receiver position, (2) receiver clock errors and biases, (3) phase ambiguities, and (4) atmospheric delay errors. The complete PPP state vector of unknowns can be written as:

$$\mathbf{x} = [\mathbf{p}_r, \mathbf{c}, \mathbf{n}, \mathbf{a}]^\top , \tag{4.6}$$

where $\mathbf{p}_r$ is the Earth-Centered Earth-Fixed (ECEF) receiver position coordinates, $\mathbf{c}$ is a vector of receiver clock and bias parameters (clock errors, offsets, and DCBs), $\mathbf{n}$ is a vector of phase ambiguities, and $\mathbf{a}$ is a vector of atmospheric delays (troposphere and ionosphere). Different PPP approaches differ on how the non-random measurement errors are handled, and consequently which unknowns are estimated. The next subsections describe the two main PPP approaches and detail their respective state vectors.

The next subsection describes various PPP corrections required to get from the measurement models described in eqs. (2.1)–(2.4) to those in eqs. (4.16)–(4.19).

4.3.1 PPP Corrections Sources and Formats

To perform PPP, the non-random errors in eqs. (2.1)–(2.4) have to be accounted for. As discussed in Chapters 2 and 3, some of the error sources in PPP can be modeled, while others have to be estimated, either by the user as part of PPP processing, or by a network of reference stations before being disseminated to the users from a SP.

In post-processing, the corrections are provided as products in various formats defined by IGS[2]. In real-time, these products are provided in the standard SSR format [73], in which case they are usually referred to as SSR corrections[3]. The SSR corrections include parameters to compute:

- Satellite orbit error, $\delta\mathbf{p}^s$
- Satellite clock error, dt^s
- Satellite signal biases, $\delta b^s_{jk}, b^s_j, b^s_k$
- Satellite ionospheric delay, I^s_r
- Satellite tropospheric delay, T^s_r

The following subsections contain details on the sources and formats of common-mode corrections in real-time and post-processing.

[2]↑IGS Formats and Standards: https://www.igs.org/formats-and-standards.

[3]↑There exists other related formats for real-time PPP corrections. For example, the commercial company Geo++ has recently proposed SSRZ, a new open SSR format that offers added flexibility. See https://www.geopp.de/ssrz/ for more information.

4.3.1.1 Orbit Error Correction

For SPS and RTK/NRTK, the ECEF estimates of satellite position $\hat{\mathbf{p}}^s(t)$ and velocity vectors $\dot{\hat{\mathbf{p}}}^s(t)$ at time t are computed from the broadcast ephemeris. The computed satellite orbits are typically only precise to a few meters. The resulting residual ephemeris error is eliminated via between-receiver differencing for RTK (see Section 4.2.2), or ignored for SPS. For PPP, precise satellite orbits are provided by external sources, since between-receiver differencing is not possible. The typical accuracy of the precise orbits is a few centimeters.

Post-processed precise satellite orbits in ECEF frame are provided in the IGS Standard Product 3 Orbit (SP3) format (*.`sp3`). The SP3 format contains the three-dimensional coordinates of each satellite with a temporal resolution of 5–15 minutes. Since satellite orbits are sufficiently smooth, the satellite position at any time t can be computed with sufficient accuracy via Lagrange interpolation. This dissertation uses post-processed satellite orbits.

In real-time, the satellite orbit corrections are disseminated in the SSR format by various sources such as the IGS Real Time Service (RTS). The orbit correction parameters $\delta\mathbf{O}_0$ and rate $\delta\dot{\mathbf{O}}_0$ are provided in Antenna Phase Center (APC) APC frame along with a reference time t_0 every 60 seconds. Given these items, the correction computation involves four steps [73, 74]:

1. The satellite orbit correction parameters $\delta\mathbf{O}_0$ and rate $\delta\dot{\mathbf{O}}_0$ are provided every 60 seconds along with a reference time t_0. The orbit correction at any time t is computed as

$$\delta\mathbf{O}(t) = \delta\mathbf{O}_0 + \delta\dot{\mathbf{O}}_0\,(t - t_0). \tag{4.7}$$

In the APC frame, the vector $\delta\mathbf{O} = \begin{bmatrix}\delta O_r & \delta O_a & \delta O_c\end{bmatrix}^\top$ has radial, along-track, and

cross-track components, respectively.

2. The radial $\mathbf{e}_r$, along $\mathbf{e}_a$, and cross-track $\mathbf{e}_c$ unit vectors in the ECEF frame are respectively computed as:

$$\begin{aligned} \mathbf{e}_a &= \frac{\dot{\hat{\mathbf{p}}}^s}{\|\dot{\hat{\mathbf{p}}}^s\|}, \\ \mathbf{e}_c &= \frac{\hat{\mathbf{p}}^s \times \dot{\hat{\mathbf{p}}}^s}{|\hat{\mathbf{p}}^s \times \dot{\hat{\mathbf{p}}}^s|}, \\ \mathbf{e}_r &= \mathbf{e}_a \times \mathbf{e}_c. \end{aligned} \tag{4.8}$$

3. The orbit correction $\delta\mathbf{O}(t)$ is then transformed from APC frame to ECEF frame as

$$\delta\mathbf{p}^s(t) = \left[\mathbf{e}_r \ \mathbf{e}_a \ \mathbf{e}_c\right] \delta\mathbf{O}(t). \tag{4.9}$$

4. The precise orbit $\hat{\mathbf{p}}^{s_p}$ is finally computed as

$$\hat{\mathbf{p}}^{s_p}(t) = \hat{\mathbf{p}}^s(t) - \delta\mathbf{p}^s(t). \tag{4.10}$$

After the computation of the satellite's precise orbit $\hat{\mathbf{p}}^{s_p}(t)$, it can be used to correct pseudorange and phase measurements in two ways that are nearly equivalent. For users with access to the receiver navigation code, the precise satellite position can be used for computing the estimated range $R_r^s = R(\hat{\mathbf{p}}_r, \hat{\mathbf{p}}^{s_p}) = \|\hat{\mathbf{p}}_r - \hat{\mathbf{p}}^{s_p}\|$ and satellite-to-receiver unit vector. In this case the ephemeris error $E_r^s = 0$. For users supplying corrections to a receiver, the receiver will still use the broadcast orbit $\hat{\mathbf{p}}^s$, the ephemeris error portion of the corrections E_r^s is

$$E_r^s = \hat{E}_r^{s_p} = \frac{(\hat{\mathbf{p}}_r - \hat{\mathbf{p}}^{s_p})}{\|\hat{\mathbf{p}}_r - \hat{\mathbf{p}}^{s_p}\|} (\delta\hat{\mathbf{p}}^s) \tag{4.11}$$

where $\hat{\mathbf{p}}_r$ is the estimated user position.

4.3.1.2 Clock Error Correction

The post-processed satellite clocks are provided in the Clock Receiver Independent Exchange Format (RINEX) format (*.`clk`), with temporal resolution of 5 seconds, 30 seconds, and 5 minutes. Due to the typical stochastic nature of clock variations, linear interpolation between consecutive epochs is sufficient to obtain higher rate clock corrections.

For real-time applications, IGS and Geo++ SSR data products provide satellite Low Rate SSR clock corrections using three polynomial parameters (a_{c_0}, a_{c_1}, a_{c_2}) and the message reference time t_0. The clock correction $c\,dt^s(t)$, in meters, is computed using the following equation [73, 74]:

$$c\,dt^s(t) = a_{c_0} + a_{c_1}(t - t_0) + a_{c_2}(t - t_0)^2. \tag{4.12}$$

Both IGS and Geo++ High Rate Clock SSR and SSRZ messages support higher update rates of the clock state. These messages are updated roughly every 5 seconds, compared to 30 seconds for the Low Rate Corrections [74].) When these higher resolution SSR messages are used, the High Rate Clock correction $\delta C_{High\,Rate}$ is added to the right-hand side of (4.12) [73, 74]. The high rate clock correction becomes:

$$c\,dt^s(t) = a_{c_0} + a_{c_1}(t - t_0) + a_{c_2}(t - t_0)^2 + \delta C_{High\,Rate}. \tag{4.13}$$

4.3.1.3 Signal Bias Correction

In post-processing, the signal biases are given in the SINEX BIAS format (*.`bia`). This format allows dissemination of biases specific to: (1) GNSS system, (2) satellite, and (3) receiver. Additionally, it allows biases to be *relative*/differential (e.g. DCB), or *absolute* (e.g. Observable-specific Code Biases (OSB)).

Differential biases are the difference between two specific signals from a particular GNSS system. For example, the DCB bias between GPS L1C/A (C1W) signals and L2C (C2W) signals is:

$$DCB_{G,C1W-C2W} = B_{G,C1W} - B_{G,C2W}, \tag{4.14}$$

where $B_{G,*}$ are the specific GPS code biases.

Absolute biases are observation-specific biases that can be applied directly to uncombined observations, such as the ones in eqs. (2.1)–(2.4). However, these biases are referenced to a common signal, and must be interpreted as *pseudo-absolute*, since truly absolute GNSS biases are inaccessible. Hence, any OSB bias B_O should be expected to be shifted by an arbitrary offset ΔB, with respect to the true (unknown) bias B, i.e.:

$$B = B_O + \Delta B. \tag{4.15}$$

Consequently, to apply OSB biases, users need to use consistent GNSS clock and bias products. This typically means that one has to use clocks and biases from the same source.

The IGS SSR message provides parameters that are directly equal to the satellite code and phase bias corrections b_j^s in meters. These corrections can be directly applied to their respective uncombined measurement models in eqs. (2.1)–(2.4), when the corresponding SSR clock messages are applied. Additionally, a yaw angle is provided to compute the phase wind-up correction of the satellite (ψ_r^s) using appropriate satellite yaw models.

4.3.1.4 PPP-Corrected Measurement Models

After all the available satellite corrections have been applied, the measurement models in eqs. (2.1)–(2.4) can be simplified. When only two frequencies are processed, the

aforementioned equations for satellite ℓ and GNSS system k are:

$$\rho^s_{r,1} = R^s_r + c\,(dt_r + \delta t_{r,1}) + T^s_r + I^s_{r,1} + \epsilon_\rho \tag{4.16}$$

$$\phi^s_{r,1} = R^s_r + c\,(dt_r) + T^s_r - I^s_{r,1} + (\lambda_1\, N^s_1 + \psi^s_r + h_{r,1}) + \epsilon_\phi \tag{4.17}$$

$$\rho^s_{r,2} = R^s_r + c\,(dt_r + \delta t_{r,2}) + T^s_r + \gamma_{1,2}\, I^s_{r,1} + \epsilon_\rho \tag{4.18}$$

$$\phi^s_{r,2} = R^s_r + c\,(dt_r) + T^s_r - \gamma_{1,2}\, I^s_{r,1} + (\lambda_2\, N^s_2 + \psi^s_r + h_{r,2}) + \epsilon_\phi \tag{4.19}$$

where ϵ_ρ and ϵ_ϕ represent the combined effects of multipath random measurement noise. The coefficient $\gamma_{1,2}$ is the squared ratio of frequencies f_1 and f_2 for GNSS system k, i.e.

$$\gamma_{1,2} = \left(\frac{f_1}{f_2}\right)^2.$$

The receiver clock and code biases in eqs. (5.22) and (5.24) are correlated and cannot be separated. Hence, the estimable receiver clock is typically the combination of the reference frequency's clock and its corresponding code bias. The code biases of other frequencies are represented through their DCBs. For instance, if the clock of f_1 is the reference clock, then the pseudorange and phase measurements in eqs. (4.16)–(4.19) are written as:

$$\rho^s_{r,1} = R^s_r + \widetilde{dt}_r + T^s_r + I^s_{r,1} + \epsilon_\rho, \tag{4.20}$$

$$\phi^s_{r,1} = R^s_r + \widetilde{dt}_r + T^s_r - I^s_{r,1} + (\lambda_1\, N^s_1 + \psi^s_r + h_{r,1} - c\,\delta t_{r,1}) + \epsilon_\phi, \tag{4.21}$$

$$\rho^s_{r,2} = R^s_r + \widetilde{dt}_r + db_{r,12} + T^s_r + \gamma_{1,2}\, I^s_{r,1} + \epsilon_\rho, \tag{4.22}$$

$$\phi^s_{r,2} = R^s_r + \widetilde{dt}_r + db_{r,12} + T^s_r - \gamma_{1,2}\, I^s_{r,1} + (\lambda_2\, N^s_2 + \psi^s_r + h_{r,2} - c\,\delta t_{r,2}) + \epsilon_\phi, \tag{4.23}$$

where

$$\widetilde{dt}_r = c\,(dt_r + \delta t_{r,1}) \tag{4.24}$$

is the estimable reference clock, and $db_{r,12} = c\,(\delta t_{r,2} - \delta t_{r,1})$ is the receiver DCB between f_1 and f_2. Note that phase measurements on the second frequency in eq. (4.23) can be

equivalently modeled using only the absolute receiver code biases from the first frequency, instead of receiver DCBs and absolute biases from the second frequency, i.e.:

$$\phi^s_{r,2} = R^s_r + \widetilde{dt}_r + T^s_r - \gamma_{1,2}\, I^s_{r,1} + (\lambda_2\, N^s_2 + \psi^s_r + h_{r,2} - c\,\delta t_{r,1}) + \epsilon_\phi. \tag{4.25}$$

Eq. (4.23) shows that when receiver DCBs are modeled for the second phase, the corresponding float ambiguities absorb both the code and phase biases from the second frequency. On the other hand, eq. (4.25) shows that without modeling the receiver DCBs on the second phase, the corresponding float ambiguities absorbs the absolute receiver code biases from the first frequency.

Each satellite may transmit signals on more than one frequency. Since the atmospheric delays on each signals are different, it is important to distinguish between the many signals from one satellite contributing measurements. A signal ℓ is henceforth defined as a measurement from satellite s with frequency f_j.

4.3.2 Ionosphere-free PPP

The dual-frequency Ionosphere-free Linear Combination (IFLC) approach uses a linear combination of code and phase observations to form new observations without ionospheric delays. For dual-frequency observations, one approach to forming the code and phase

Ionosphere-free (IF) observation using eqs. (4.20)–(4.23) is [75]:

$$\rho_{IF} = \frac{f_1^2}{f_1^2 - f_2^2}\rho_{r,1}^s - \frac{f_2^2}{f_1^2 - f_2^2}\rho_{r,2}^s \tag{4.26}$$

$$= R_r^s + c\,\widetilde{dt}_{r,IF} + T_r^s + \epsilon_{\rho,IF} \tag{4.27}$$

$$\phi_{IF} = \frac{f_1^2}{f_1^2 - f_2^2}\phi_{r,1}^s - \frac{f_2^2}{f_1^2 - f_2^2}\phi_{r,2}^s \tag{4.28}$$

$$= R_r^s + c\,\widetilde{dt}_{r,IF} + T_r^s + \lambda_{IF}\,\widetilde{N}_{r,IF}^s + \epsilon_{\phi,IF}, \tag{4.29}$$

where $d\tilde{t}_{r,IF}$ is the estimable ionosphere-free receiver clock, defined as:

$$\widetilde{dt}_{r,IF} = \widetilde{dt}_r - \frac{f_2^2}{f_1^2 - f_2^2}\,\delta b_{r,12}. \tag{4.30}$$

The IF integer ambiguity is combined together with the IF phase biases from both the receiver and the satellite, resulting in a *float* (non-integer) ambiguity $\widetilde{N}_{r,IF}^s$:

$$\widetilde{N}_{r,IF}^s \triangleq \lambda_{IF} N_{r,IF}^s + \frac{1}{\lambda_{IF}}\left(h_{r,IF} - c\delta t_{r,IF}\right). \tag{4.31}$$

The measurement error terms $\epsilon_{\rho,IF}$ and $\epsilon_{\phi,IF}$ are the *amplified* multipath and measurement noise terms. The measurement noise for GPS L1–L2 pseudorange and phase IF observations in eq. (4.27) and (4.29) is amplified to about three times the original noise of respective uncombined observations. For GPS L1–L5 or GAL E1–E5a, the noise is amplified by a factor of about 2.6. The bigger the separation of frequency bands, the better. This is because the IFLC noise is amplified dramatically when the observations are from frequency bands close together. For instance, the noise is amplified by a factor of about 16.6 for GPS L2–L5, and about 27.5 for E5a–E5b. The increased measurement noise affects the convergence performance. To overcome the amplification of noise, other ways of forming IF combinations have been proposed. For example, for single-frequency receivers, [75]

proposed averaging the pseudorange and phase measurements in eqs. (4.16) and (4.17) to form a new IF observation with half the noise of the original pseudorange observation.

Critically, the dominant first-order ionospheric effects are eliminated[4] from the IF observations. However, the main downside of the IFLC approach is the reduced measurement precision due to the amplification of measurement noise. Moreover, the addition of a third or a fourth frequency in multi-GNSS also requires the careful and complicated handling of satellite DCB corrections.

Assuming dual-frequency observations from k constellations, and ℓ satellites per constellation, then estimated state vector from eq. (4.6) can be rewritten as:

$$\mathbf{x} = [\mathbf{p}_r, \mathbf{c}, \mathbf{n}, \mathbf{a}]^\top \in \mathbb{R}^{(4+k+k\,\ell)\times 1}, \tag{4.32}$$

where $\mathbf{p}_r$ is the ECEF position vector for the receiver:

$$\mathbf{p}_r = [x, y, z]^\top \in \mathbb{R}^{3\times 1}, \tag{4.33}$$

the vector $\mathbf{c}$ is a vector containing the receiver reference IF clock error for the k processed systems:

$$\mathbf{c} = \left[\widetilde{dt}^{\,(1)}_{r,IF}, \widetilde{dt}^{\,(2)}_{r,IF}, \cdots, \widetilde{dt}^{\,(k)}_{r,IF}\right]^\top \in \mathbb{R}^{k\times 1}, \tag{4.34}$$

the vector $\mathbf{n}$ contains the IF ambiguities of satellites from all k systems:

$$\mathbf{n} = \left[\widetilde{N}^1_{r,IF}, \widetilde{N}^2_{r,IF}, \cdots, \widetilde{N}^s_{r,IF},\right]^\top \in \mathbb{R}^{(k\,\ell)\times 1}, \tag{4.35}$$

and $\mathbf{a}$ is the ZTD defined in eq. (3.2).

[4]↑There remains second-order ionospheric effects, but these typically account for less than 1% of the total ionospheric delay [76].

The ambiguity term in eq. (4.31), which is no longer an integer, cannot be separated from the combined effects of receiver code and phase biases. As shown in Table 4.2, the single-epoch measurement redundancy of IFLC is negative, which means that the corresponding linear system is underdetermined for a single epoch. Hence, it takes more than one epoch to solve for all unknowns. Initially, the performance of dual-frequency the IFLC approach depends on the IF code observations, as there is still large uncertainty in the float ambiguity terms in eq. (4.29). Additionally, the redundancy of IFLC does not depend on the number of GNSS systems or satellites used, since the addition of more satellites or systems adds an equal number of ambiguity unknowns. This contributes to the observed slower convergence of IFLC compared to other PPP approaches.

For a static dual-frequency GPS observations, the convergence time of IF PPP is reported to be 1 to 1.5 hours in [77] to reach 10 cm accuracy, while 2 to 3 hours time is reported in [78] to reach 5 cm or better.

4.3.3 Uncombined, Undifferenced PPP

The uncombined, undifferenced approach directly uses the code and phase observations in eqs. (4.20)–(4.23). For frequency f_i of each constellation, the measurement models become:

$$\rho^s_{r,i} = R^s_r + \widetilde{dt}_r - \delta b_{r,1i} + T^s_r + \gamma_{1,i}\, I^s_{r,1} + \epsilon_\rho \tag{4.36}$$

$$\phi^s_{r,i} = R^s_r + \widetilde{dt}_r - \delta b_{r,1i} + T^s_r - \gamma_{1,i}\, I^s_{r,1} + \lambda_i\, \widetilde{N}^s_{r,i} + \epsilon_\phi, \tag{4.37}$$

where the float ambiguity is defined as:

$$\widetilde{N}^s_{r,i} \triangleq N^s_{r,i} + \frac{1}{\lambda_i}\left(\psi^s_r + h_{r,i} - c\delta t_{r,i}\right). \tag{4.38}$$

The measurement noise for uncombined observations is a combination of the original multipath and measurement noise terms which are, in contrast to IF observations, *not* amplified. The absolute receiver code biases $\delta t_{r,i}$ in eqs. (2.1) and (2.3) are not directly estimable since they cannot be distinguished from clock errors. Hence, the receiver DCBs $\delta b_{r,1i}$ defined in eq. (2.6) are used instead. These receiver DCBs are usually estimated as unknowns as part of PPP processing, although they can be corrected with known values for some reference stations, since they are intrinsic properties of the used receiver hardware (receiver-cable-antenna).

Similar to receiver DCBs, the slant ionospheric delays $I^s_{r,1}$ are either estimated or corrected by external sources. In the case when they are estimated, the corrections from external sources can act as pseudo-observations, that are used as tight constraints to the estimated ionospheric delays, which has shown to be very beneficial [79].

The total number of unknowns in the state vector from eq. (4.6) is $n = 4+2k+2k\ell+k\ell$. Assuming once again dual-frequency observations from k constellations and ℓ satellites per constellation, the elements of the state vector are detailed below. The clock parameters include both the clock errors and receiver DCBs for each frequency of each system used:

$$\mathbf{c} = \left[\widetilde{dt}_r^{\,sys_1}, \delta b_{r,1i}^{sys_1}, \Delta t_r^{sys_2}, \delta b_{1i}^{sys_2}, \cdots, \Delta t_r^{sys_k}, \delta b_{1i}^{sys_k}\right]^\top \in \mathbb{R}^{(2k)\times 1}. \tag{4.39}$$

Unlike, IFLC, multiple frequency-dependent ambiguities are estimated per satellite, if two or more frequencies are used. The ambiguity vector is:

$$\mathbf{n} = \left[\widetilde{N}_{r,i}^1, \widetilde{N}_{r,i}^2, \cdots, \widetilde{N}_{r,i}^\ell\right]^\top \in \mathbb{R}^{2k\ell} \tag{4.40}$$

The estimated atmospheric delay states include the residual zenith wet delay and the slant

ionospheric delays for ℓ used satellites:

$$\mathbf{a} = \left[\text{ZWD}, I_{r,1}^1, I_{r,1}^2, \cdots, I_{r,1}^\ell\right]^\top \in \mathbb{R}^{(k\ell+1)\times 1}. \tag{4.41}$$

The approach which estimates the unknowns in eq. (4.41) is henceforth referred to as PPP with Local Atmospheric Modeling (LAM-PPP) because it estimates the local atmospheric delays (both tropospheric and ionospheric delays). To emphasize the fact that the ionosphere is modeled through slant delays, the approach is specifically referred to as "LAM-PPP Slant".

The measurement redundancy of the LAM-PPP Slant approach depends on the number of satellites used, as shown in Table 4.2. For any number of GNSS systems, the corresponding linear system is overdetermined when at least four satellites are used. The addition of more satellites or more GNSS systems increases the redundancy of the system, making this approach more suitable for multi-GNSS PPP.

4.3.4 PPP-RTK

Networks of receivers can be used to enable Ambiguity Resolution (AR) in PPP. The PPP approaches that depend on networks similar to NRTK are referred to as PPP-RTK or Network PPP. For these approaches, receivers in a network send their SSR estimates to a central SP that computes corresponding corrections and interpolations, then sends improved SSR corrections back to users. Similar to NRTK, the PPP-RTK approach augments PPP with external ionospheric and tropospheric corrections derived from a ground-based network of nearby receivers [69]. Similar to PPP, the PPP-RTK corrections are shared as SSR corrections compatible with existing implementations of PPP. With precise atmospheric

Service	Mode	Accuracy [cm]	Convergence[a] [min]	Service area	Implementation
SPS	–	< 300	–	Global	–
RTK	Float	< 5	Instantaneous	Local/Regional[b]	CORS networks, peer-to-peer, RTCM, NTRIP
	AR	< 2	Instantaneous		
PPP	Float	< 10	30.0	Global	Global SPs, SSR
	AR	< 5	5		
PPP-RTK	Float	< 10	30.0	Global/Regional	Regional SPs, SSR
	AR	< 5	5		

[a] Time required to reach the horizontal position error less than 10 cm.
[b] Assuming a baseline of around 30 km; NRTK can have an extended service area, albeit at the cost of slight degradation in performance.

Table 4.1: Typical performance of different positioning services. Results assume dual-frequency measurements from two GNSS systems.

corrections, PPP-RTK enables reliable AR, which makes its positioning performance similar to that of RTK. (The discussion of AR is out of scope of this work.)

4.4 Summary

Table 4.1 gives a simplified performance summary of different positioning services. The convergence time (applicable only to PPP) assumes measurement sampling interval of 15 sec; convergence time increases with larger sampling intervals. The last two columns contain a general characterization of areas that can be serviced, and which communication implementations and data formats are used.

4.4.1 Discussion

Although RTK offers very fast and precise position solutions through double-differencing, it does not offer accurate *absolute* solutions of atmospheric delays. For instance, the RTK

tropospheric delay solution might be biased by an unknown constant bias due to the clock error. On the other hand, in addition to precision position solutions, PPP offers precise clock and *absolute* atmospheric delay solutions (in the case of uncombined PPP). This makes PPP suitable for collaborative applications in which agents in a defined geographic area have to share their estimates of common unknown atmospheric parameters.

The IFLC approach offers a convenient way to eliminate the ionospheric delay and receiver DCB unknowns, but it comes with a few disadvantages compared to the uncombined model. Foremost of these disadvantages is the aforementioned amplification of the measurement noise, which increases the convergence time. Although IF combinations exists for triple-frequency and multi-frequency processing, observations from only two frequencies can be combined at a time [80]. For instance, if GPS L1/L2/L5 frequencies are available, the following IF combinations can be made: L1–L2, L1–L5, and L2–L5. This not only complicates the modeling and handling of various corrections, it also complicates the stochastic modeling and ambiguity resolution. This is in direct comparison to the uncombined model for which the presence of additional frequencies does not increase the complexity of the modeling, but increases the number of available measurements. Another important disadvantage is that users cannot collaborate on estimating the ionospheric delays, since these delays are not provided by IF solutions. Therefore IFLC is not suitable for collaborative estimation of atmospheric delays. Instead, the use of uncombined signals can be greatly beneficial to PPP since the constraining of ionospheric delays using external sources speeds up convergence [79, 81].

Model	m	n	Redundancy[a]	Atm. states[b]
IFLC	$2k\ell$	$4+k+k\ell$	$k(\ell-1)-4$	1
LAM-PPP Slant	$4k\ell+k\ell$	$4+2k+2k\ell+k\ell$	$2k(\ell-1)-4$	ℓ
LAM-PPP Spline[c]	$4k\ell+K$	$4+2k+2k\ell+K$	$2k(\ell-1)-4$	K

[a] Single-epoch redundancy.
[b] Number of estimated atmospheric states. For IFLC only the tropospheric state is estimated, whereas both the tropospheric and ionospheric states are estimated for uncombined approaches.
[c] Proposed approach in Chapter 5.

Table 4.2: Redundancy comparison for different dual-frequency PPP approaches.

4.4.2 Desired Improvements

External global PPP corrections for both tropospheric and ionospheric delay errors are adequate for applications requiring centimeter-level accuracy after a long convergence time. However, they are not adequate for precise positioning applications, such as automotive transportation, which requires real-time centimeter-level accuracy. Hence, there is a need for novel approaches to reducing convergence time via the use of local atmospheric corrections. Chapter 5 presents a novel approach of generating regional PPP atmospheric corrections from individual receivers. Chapter 6 presents a new collaborative PPP approach that achieves rapid convergence time using atmospheric corrections from a network of receivers.

Chapter 5

PPP with Local Atmospheric Modeling

5.1 Introduction

Chapter 4 has shown that for multi-GNSS, multi-frequency observations, using the uncombined PPP measurements allows the use of twice as many measurements, each with lower noise levels. The trade-off is that the ionospheric error must be removed removed by other means. As shown in Section 4.3.3, it is not possible to eliminate atmospheric delay errors in uncombined PPP; they have to be handled using the remaining three options discussed in Section 2.3: (1) estimation, (2) modeling, and (3) omission. For high-precision PPP applications, omission is not viable, mainly because these errors have large magnitudes. Modeling is also not viable because they have unpredictable spatio-temporal variations. Hence, the only viable option is estimation.

Previous work has shown that having precise local ionospheric and tropospheric delays significantly improves PPP convergence time [5,7,8,82]. Global ionospheric models often do not have the required precision; therefore the per-satellite slant ionospheric delays are often estimated along with other unknown parameters. In such cases, the external global models are used to constrain ionospheric slant delays. The main limitation of this approach is that measurements from each satellite can only be used to constrain the slant delay constraints for the corresponding satellite, i.e. there is no possibility for satellites to collaborate in estimating each other's slant delays. This chapter explores a novel approach for satellite collaboration in estimating ionospheric delays.

Recall that the slant ionospheric delay is related to VTEC via the satellite's elevation, as shown in eqs. (3.11) and (3.14). Each satellite's pierce point has a VTEC value that corresponds to an elevation-dependent slant delay. For a user receiver, the multiple slant delays correspond to multiple pierce points distributed geographically around the user. If the area around a user is divided into a grid with predefined knots, then multiple pierce points can contribute to estimating the VTEC value at a particular knot. Fig. 5.1 shows an illustration of satellites collaborating to estimate the VTEC values across the grid. The blue dots (•) and the red dot (•) are satellite pierce points. The gray dots (•) are grid knot points, and the green dot (•) is the user's receiver location. Each pierce point can contribute to estimating the VTEC values for the nearest four grid knots. Furthermore, an unlimited number of pierce points can contribute to estimating the VTEC value of one grid knot. This approach takes advantage of the larger number of satellites available in multi-GNSS without increasing the number of estimated ionospheric delay parameters. Fig. 5.1 shows two examples of three pierce points (•) collaborating to estimate VTEC values for one

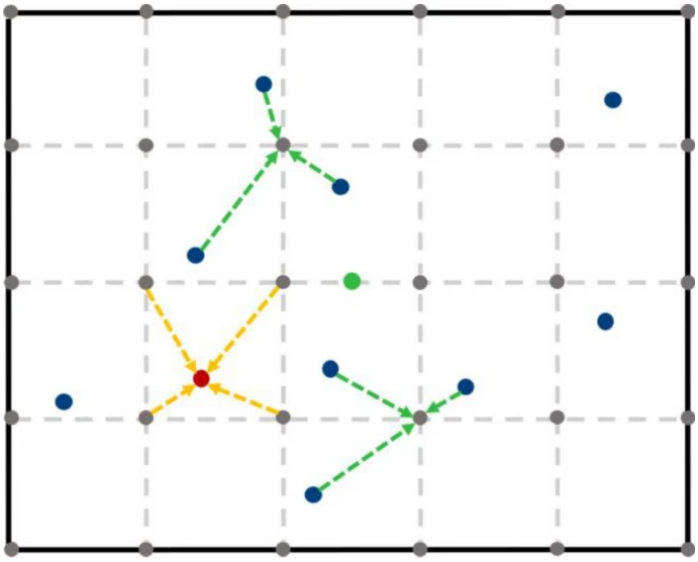


Figure 5.1: An illustration of satellite collaboration in estimating the regional VTEC map. The gray dots are the VTEC grid knots, the blue dots are the satellite pierce points, the red dot is the user receiver.

grid knot (•). The between-satellite collaboration is shown by dashed green arrows (- - -›): multiple pierce points are collaborating to estimate the VTEC at one grid knot. The dashed yellow arrows (- - -›) also shows that multiple grid knots (•) contribute to the VTEC value at each pierce point, illustrated here by red (•). This chapter will explore, among other things, an approach of estimating the ionospheric delay as a VTEC grid covering a specific region, instead of estimating the individual satellites' slant delays. The primary goal of this chapter is to achieve a specified accuracy performance at least as fast as existing methods, using an approach that can be extended to collaboration between nearby receivers (i.e. agents).

5.2 Modeling Local Tropospheric Delay

As discussed at length in Section 3.1, the tropospheric delay is modeled as two components with distinct behavior. The dry component (ZTD) is homogeneous and constant, and

hence it can be computed precisely using standard atmosphere models. On the other hand, the wet component (ZWD) is dynamic and depends on weather phenomena. If precise in-situ meteorological measurements are available, the ZWD can be computed very precisely by empirical models such as those represented by eq. (3.6) or eq. (3.8). However, since it is not always possible to obtain accurate meteorological measurements, PPP uses global models of the standard atmosphere such as GPT3 or GPT2w, described in Sections 3.1.1 and 3.1.2. Because the wet tropospheric delay depends on dynamic local weather phenomena, global models cannot account for the entire ZWD. As a consequence, after applying corrections from global models, there remains a *local* residual ZWD delay (defined in eq. (3.3)) that will be estimated as an unknown parameter.

The ZWD has a temporal variation in the order of a few centimeters per hour. Hence, the local residual ZWD delay is modeled as a Gauss-Markov (GM) random walk. See Section 5.4.5 and Table 5.3 for more details on GM processes and the stochastic modeling of all other estimated states.

5.3 Modeling Regional VTEC using B-Splines

The Spherical Harmonics (SH) approach has been the main technique for modeling Global VTEC, due to its suitability to modeling well-behaved phenomena on spherical surfaces. SH models use special functions to fit sparse measurements on the surface of a sphere. Such functions, with appropriate interpolation, can then be used to infer the behavior of the phenomenon in regions of the sphere that have no measurements [83]. However, despite their ability to handle sparse data, SH models are only optimal for representing measure-

ments distributed globally [84]. As result they are not optimal for estimating the local VTEC using sparse GNSS measurements distributed in a smaller regional area.

Multiple studies have shown that techniques used to model the global VTEC are not fit for the purpose of modeling regional VTEC, due to a lack of the *localizing* property [84,85]. The localizing property, also referred to as compact support, means that the value of VTEC at a point depends only on a few control points localized nearby. One of the alternative functions that has been proposed is polynomial B-splines, which are further investigated in this subsection.

Before proceeding, it is important to state two important assumptions about the ionosphere and the available global ionospheric corrections. The regional area is defined as follows:

Definition 5.3.1. *The regional area is defined as the rectangular area which covers all possible pierce points around the user's receiver for a specific elevation cut-off.*

Assumption 5.3.2. *The temporal and spatial resolution of the GIM is insufficient to capture regional ionospheric variations necessary for rapid instantaneous convergence in PPP applications.*

In mid-latitudes (30°), the typical GIM resolution of $2.5° \times 5°$ corresponds to an area of around 280 km $\times$ 480 km. The temporal sampling resolution of 2 hours also exceeds the desired temporal resolution of around 5 minutes.

Assumption 5.3.3. *The ionosphere changes quicker than GIM temporal resolution.*

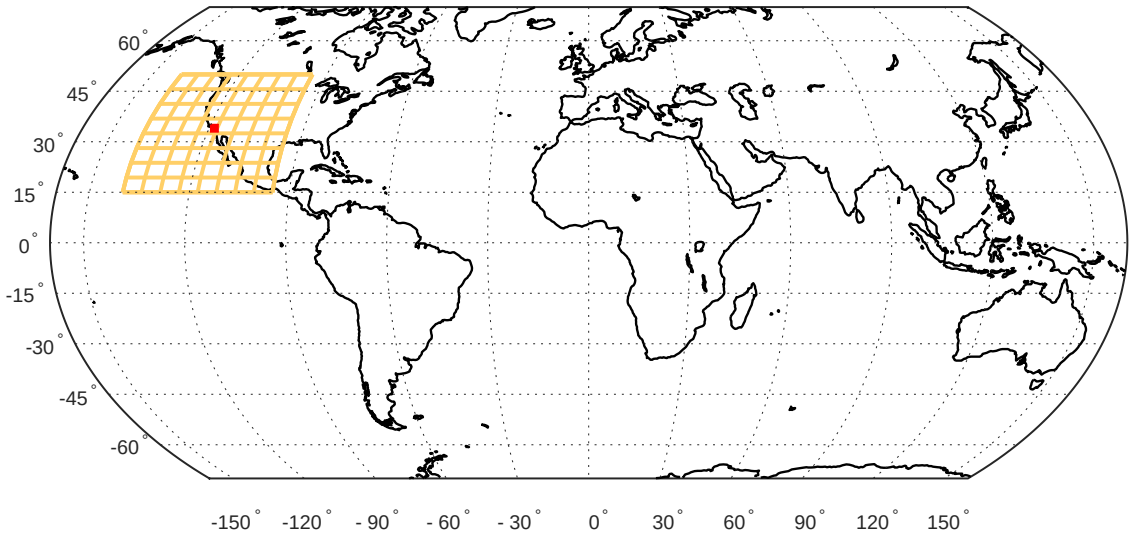


Figure 5.2: Global map showing the regional area of interest, assuming an elevation cutoff of 15° and the ionospheric height of 450 km.

Temporal ionospheric irregularities are characterized by the Rate of TEC Index (ROTI) index, which is defined as the standard deviation of the rate of TEC (ROT) [86]. During normal ionospheric activity in mid-latitude regions, Global ROTI values typically have a mean of 0.0–0.75 TECu/min and a maximum of up to 2 TECu/min for a one-minute period. During periods of heightened ionospheric activity (e.g. scintillation periods), ROTI values can be much higher. For instance, during the St. Patrick's Day storm (whose sample of VTEC values is shown in Fig. 3.2b), ROTI values were as high as 5 TECu/min for extended periods of time in high-latitude regions [87].

5.3.1 Modeling Regional VTEC Using B-splines

The three-dimensional VTEC map is a function of time (t), and the geographic latitude (φ_{pp}) and longitude (λ_{pp}) of the ionospheric pierce point. It can be represented by an infinite

series:

$$VTEC(t, \varphi_{pp}, \lambda_{pp}) = \sum_{k=0}^{\infty} d_k(t)\, \Psi_k(\varphi_{pp}, \lambda_{pp}), \tag{5.1}$$

where the quantity $\{d_k(t) \,|\, k = 1, \ldots, \infty\}$ is a set of time-dependent scalar coefficients, and $\{\Psi_k(\varphi_{pp}, \lambda_{pp}) \,|\, k = 1, \ldots, \infty\}$ is set of space-dependent two-dimensional 2-D basis functions. Let the latitude of pierce points be defined as $\{\varphi_{pp} \,|\, \varphi_{min} \leq \varphi_{pp} \leq \varphi_{max}\}$. Similarly, let the longitude of pierce points be defined as $\{\lambda_{pp} \,|\, \lambda_{min} \leq \lambda_{pp} \leq \lambda_{max}\}$. Define the range of latitudes and longitudes as $W_\varphi = [\varphi_{min}, \varphi_{max}]$ and $W_\lambda = [\lambda_{min}, \lambda_{max}]$, respectively. Then, eq. (5.1) can be used to conceptually form a representation of a set of time-dependent two-dimensional VTEC maps covering the geographical area defined by $G = \{W_\varphi \times W_\lambda\}$. Fig. 5.2 shows an example of such an area $G = \{[-15°, 50°] \times [-95°, -140°]\}$, centered around a user in a mid-latitude region. For each time t instant, eq. (5.1) can be approximated by:

$$VTEC(\varphi_{pp}, \lambda_{pp}) = \sum_{k=0}^{K} d_k\, \Psi_k(\varphi_{pp}, \lambda_{pp}) + \varepsilon, \tag{5.2}$$

where ε is the error resulting from truncating the infinite series in eq. (5.1) to K series. This truncation error is defined as:

$$\varepsilon = \sum_{k=K+1}^{\infty} d_k\, \Psi_k(\varphi_{pp}, \lambda_{pp}). \tag{5.3}$$

The value of K, the number of basis functions, is determined by the desired *resolution* of the represented VTEC map. Compared to lower resolution maps, higher resolution maps require larger values of K, resulting in smaller values of truncation error ϵ.

Let K_φ and K_λ be the number of basis functions sufficient to represent the latitude and longitude dimensions, respectively. For a given region G, there are K_φ latitude and K_λ longitude basis functions, each with local support over a resolution of $w_\varphi = W_\varphi / K_\varphi$ and

$w_\lambda = W_\lambda / K_\varphi$, respectively. Then, the total number of basis functions is $K = K_\varphi \, K_\lambda$. As K_φ and K_λ increase, the resolution increases and w_φ and w_λ approach 0, and K increases towards infinity. Choosing values for either w_φ or K_φ (or their longitude counterparts) involves a trade-off between computation and resolution. (The next subsection discusses an approach to determine the appropriate values for K_φ and K_λ.)

For a sufficient number of series K, the truncation error ε is assumed to be negligible [88], and it will be omitted in subsequent analysis. Eq. (5.2) with ϵ set to 0 forms a two-dimensional VTEC map for values of latitude and longitude in region G defined before eq. (5.2). The two-dimensional, space-dependent basis functions can be decomposed into one-dimensional functions. Then, eq. (5.2) becomes:

$$VTEC(\varphi_{pp}, \lambda_{pp}) = \left(\sum_{k_1=0}^{K_\varphi - 1} d_{k_1}^{J_\varphi} \Psi_{k_1}^{J_\varphi}(\varphi_{pp}) \right) \otimes \left(\sum_{k_2=0}^{K_\lambda - 1} d_{k_2}^{J_\lambda} \Psi_{k_2}^{J_\lambda}(\lambda_{pp}) \right) \tag{5.4}$$

$$= \sum_{k_1=0}^{K_\varphi - 1} \sum_{k_2=0}^{K_\lambda - 1} d_{k_1}^{J_\varphi} \, d_{k_2}^{J_\lambda} \left(\Psi_{k_1}^{J_\varphi}(\varphi_{pp}) \otimes \Psi_{k_2}^{J_\lambda}(\lambda_{pp}) \right) \tag{5.5}$$

$$= \sum_{k_1=0}^{K_\varphi - 1} \sum_{k_2=0}^{K_\lambda - 1} d_{(k_1,k_2)}^{J_\varphi, J_\lambda} \; \Psi_{k_1,k_2}^{J_\varphi, J_\lambda}(\varphi_{pp}, \lambda_{pp}), \tag{5.6}$$

where $d_{k_1,k_2}^{J_\varphi, J_\lambda} = d_{k_1}^{J_\varphi} \, d_{k_2}^{J_\lambda}$ are scaling coefficients, and $\Psi_{k_1,k_2}^{J_\varphi, J_\lambda}(\varphi_{pp}, \lambda_{pp})$ are the 2-D basis functions defined as tensor products of the individual 1-D scaling functions, i.e.:

$$\Psi_{k_1,k_2}^{J_\varphi, J_\lambda}(\varphi_{pp}, \lambda_{pp}) = \Psi_{k_1}^{J_\varphi}(\varphi_{pp}) \otimes \Psi_{k_2}^{J_\lambda}(\lambda_{pp}) \tag{5.7}$$

The representation in eq. (5.6) can be written as a matrix product. Define 1-D basis vectors

$$\mathbf{\Psi}^{J_\varphi}(\varphi_{pp}) = \left[\Psi_0^{J_\varphi}(\varphi_{pp}), \Psi_1^{J_\varphi}(\varphi_{pp}), \cdots, \Psi_{K_\varphi - 1}^{J_\varphi}(\varphi_{pp}) \right]^\top \in \mathbb{R}^{K_\varphi \times 1} \tag{5.8}$$

$$\mathbf{\Psi}^{J_\lambda}(\lambda_{pp}) = \left[\Psi_0^{J_\lambda}(\lambda_{pp}), \Psi_1^{J_\lambda}(\lambda_{pp}), \cdots, \Psi_{K_\lambda - 1}^{J_\lambda}(\lambda_{pp}) \right]^\top \in \mathbb{R}^{K_\lambda \times 1} \tag{5.9}$$

Similarly, define a coefficient matrix $\boldsymbol{D}_{J_\varphi,J_\lambda} \in \mathbb{R}^{K_\varphi \times K_\lambda}$

$$\boldsymbol{D}_{J_\varphi,J_\lambda} = \begin{bmatrix} d_{(0,0)} & d_{(0,1)} & \cdots & d_{(0,K_\lambda-1)} \\ d_{(1,0)} & d_{(1,1)} & \cdots & d_{(1,K_\lambda-1)} \\ \vdots & \vdots & \ddots & \vdots \\ d_{(K_\varphi-1,0)} & d_{(K_\varphi-1,1)} & \cdots & d_{(K_\varphi-1,K_\lambda-1)} \end{bmatrix}, \tag{5.10}$$

where the superscripts J_φ, J_λ has been omitted for clarity. Then eq. (5.6) can be written as:

$$VTEC(\varphi_{pp}, \lambda_{pp}) = \left(\boldsymbol{\Psi}^{J_\varphi}(\varphi_{pp})\right)^\top \boldsymbol{D}_{J_\varphi,J_\lambda} \left(\boldsymbol{\Psi}^{J_\lambda}(\lambda_{pp})\right) \tag{5.11}$$

The quantity in eq. (5.11) is a one-element 1-D vector, i.e. a scalar. It can be thought of as a dot product of two 1-D vectors. Using the definition of the vectorized matrix product in eq. (1.3), the product in eq. (5.11) can be written as a product of two vectors:

$$\begin{aligned} VTEC(\varphi_{pp}, \lambda_{pp}) &= \left(\left(\boldsymbol{\Psi}^{J_\lambda}(\lambda_{pp})\right)^\top \otimes \left(\boldsymbol{\Psi}^{J_\varphi}(\varphi_{pp})\right)^\top\right) \vec{\boldsymbol{d}}_{J_\varphi,J_\lambda} && (5.12) \\ &= \left(\boldsymbol{\Psi}^{J_\lambda}(\lambda_{pp}) \otimes \boldsymbol{\Psi}^{J_\varphi}(\varphi_{pp})\right)^\top \vec{\boldsymbol{d}}_{J_\varphi,J_\lambda}, && (5.13) \\ &= \left(\boldsymbol{\Psi}^{J_\lambda,J_\varphi}\right)^\top \vec{\boldsymbol{d}}_{J_\varphi,J_\lambda}, && (5.14) \end{aligned}$$

where $\vec{\boldsymbol{d}}_{J_\varphi,J_\lambda} \in \mathbb{R}^{(K_\varphi K_\lambda)\times 1}$ is the vectorized form of matrix $\boldsymbol{D}_{J_\varphi,J_\lambda}$, defined by stacking the columns of $\boldsymbol{D}_{J_\varphi,J_\lambda}$ on top of one another:

$$\begin{aligned} \vec{\boldsymbol{d}}_{J_\varphi,J_\lambda} &= \mathrm{vec}\left(\boldsymbol{D}_{J_\varphi,J_\lambda}\right) \\ &= \left[d_{(0,0)}, \ldots, d_{(K_\varphi-1,0)}, d_{(0,1)}, \ldots, d_{(K_\varphi-1,1)}, \ldots, d_{(0,K_\lambda-1)}, \ldots, d_{(K_\varphi-1,K_\lambda-1)}\right]^\top, \end{aligned}$$

and $\boldsymbol{\Psi}^{J_\lambda,J_\varphi} \in \mathbb{R}^{(K_\varphi K_\lambda)\times 1}$ is a vector defined as:

$$\boldsymbol{\Psi}^{J_\lambda,J_\varphi} = \left[\Psi_0^{J_\varphi}(\varphi_{pp})\left(\boldsymbol{\Psi}^{J_\lambda}(\lambda_{pp})\right)^\top, \Psi_1^{J_\varphi}(\varphi_{pp})\left(\boldsymbol{\Psi}^{J_\lambda}(\lambda_{pp})\right)^\top, \ldots, \Psi_{K_\varphi-1}^{J_\varphi}(\varphi_{pp})\left(\boldsymbol{\Psi}^{J_\lambda}(\lambda_{pp})\right)^\top\right]^\top.$$

(For clarity, the following sections drop the "J_φ, J_λ" subscript and "J_λ, J_φ" superscript in the above two equations, respectively.)

There exists multiple candidates that could be used as basis functions in eqs. (5.8) and (5.9). Some of the most commonly used in the GNSS literature are: spherical cap harmonics, trigonometric B-splines, and polynomial B-splines. The former is suitable for modeling global phenomena on a spherical surface. The latter two are suitable for modeling regional phenomena, due to their local support property, which means that each basis function has non-zero values only in a finite sub-interval of the region being modeled. Functions with this property are also said to have *compact support*. This property allows for handling input data gaps resulting from uneven input data distribution [84, 89, 90].

Polynomial B-splines have been chosen for this work because they have local support, and they are relatively simpler to define and analyze. Other local approximators (e.g., radial basis functions, wavelets) could also be used. It should noted that both latitude and longitude basis functions need not be of the same kind. For modeling global VTEC, it has previously been shown that both polynomial and trigonometric basis functions can be used simultaneously for latitude and longitude, respectively [89].

5.3.2 Construction of Polynomial B-splines

The polynomial B-splines are constructed using the Cox-de Boor recursion formula [88]. The degree n and resolution level J_φ of the B-splines are chosen beforehand, according to criteria such as: expected VTEC curvature and geographic area to be covered.

The basis functions of n-th degree for latitude values $\varphi_{pp} = \{\varphi_{pp} \,|\, \varphi_{min} \leq \varphi_{pp} \leq \varphi_{max}\}$ can be constructed using the following steps:

1. Define the degree n and *resolution level* J_φ of the B-splines.

2. Using the extent of the chosen area, define the minimum and maximum latitude values: φ_{min} and φ_{max}, respectively.

3. Define a non-decreasing sequence of $K_\varphi = 2^{J_\varphi} + 2$ knot points $\varphi_{k_1}^{J_\varphi}$:

$$\varphi_{min} = \varphi_0^{J_\varphi} = \varphi_1^{J_\varphi} = \varphi_2^{J_\varphi} < \varphi_3^{J_\varphi} < \cdots < \varphi_{K_\varphi}^{J_\varphi} = \varphi_{K_\varphi+1}^{J_\varphi} = \varphi_{K_\varphi+2}^{J_\varphi} = \varphi_{max}. \tag{5.15}$$

   The knots for $k_1 = 2, \cdots, K_\varphi - 1$ are called interior knots, whereas the remaining knots $k_1 = \{0, 1, K_\varphi, K_\varphi + 1\}$ are called endpoint or pole knots. There are multiple endpoint knots to ensure that the modeled interval is closed. This technique is referred to as "endpoint interpolation".

   For interior knots, the constant distance between consecutive knots $\varphi_{k_1}^{J_\varphi}$ and $\varphi_{k_1+1}^{J_\varphi}$ at *resolution level* J_φ can be computed as:

$$\Delta_{J_\varphi} = \frac{\varphi_{max} - \varphi_{min}}{2^{J_\varphi}} = \frac{\Delta\varphi}{2^{J_\varphi}}. \tag{5.16}$$

   The number of resulting basis functions is equal to the number of knot points K_φ.

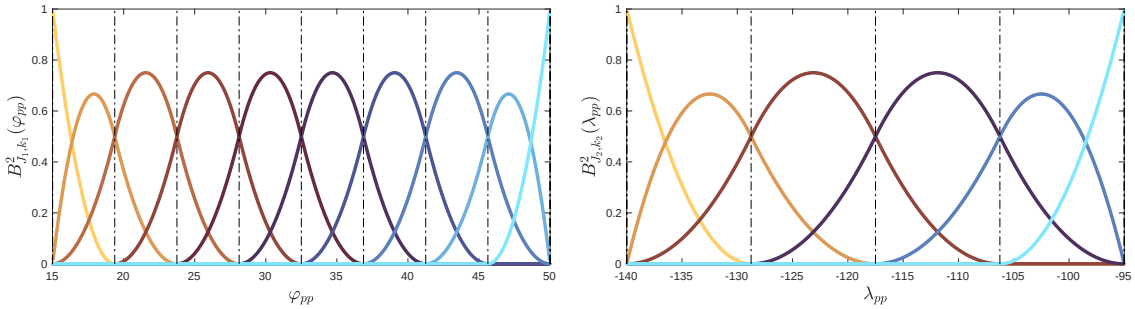
4. Compute the value of the basis functions at each latitude φ_{pp} using the normalized polynomial B-splines [88, 91, 92]:

$$\Psi_{k_1}^{J_\varphi}(\varphi_{pp}) = B_{J_\varphi,k_1}^{n}(\varphi_{pp}) \tag{5.17}$$

$$= \frac{\varphi_{pp} - \varphi_{k_1}^{J_\varphi}}{\varphi_{k_1+n}^{J_\varphi} - \varphi_{k_1}^{J_\varphi}} B_{J_\varphi,k_1}^{n-1}(\varphi_{pp}) + \frac{\varphi_{k_1+n+1}^{J_\varphi} - \varphi_{pp}}{\varphi_{k_1+n+1}^{J_\varphi} - \varphi_{k_1+1}^{J_\varphi}} B_{J_\varphi,k_1}^{n-1}(\varphi_{pp}), \tag{5.18}$$

   with

$$B_{J_\varphi,k_1}^{0}(\varphi_{pp}) = \begin{cases} 1 & \text{if } \varphi_{k_1}^{J_\varphi} \le \varphi_{pp} \le \varphi_{k_1+1}^{J_\varphi} \text{ and } \varphi_{k_1}^{J_\varphi} < \varphi_{k_1+1}^{J_\varphi}, \\ 0 & \text{otherwise} \end{cases} \tag{5.19}$$



(a) B-spline basis functions of degree $n = 2$ and resolution level $J_\varphi = 3$.

(b) B-spline basis functions of degree $n = 2$ and resolution level $J_\lambda = 2$

Figure 5.3: Polynomial B-spline basis functions of degree $n = 2$ and different resolution levels.

Fig. 5.3 shows polynomial B-spline basis functions of degree $n = 2$ and resolution levels $J = 2, 3$, chosen in Step 1. The dashed vertical gray lines show the locations of internal knot points $\varphi_2^{J_\varphi}, \ldots, \varphi_{K_\varphi}^{J_\varphi}$.

The total number of B-splines K_φ depends only the resolution level J_φ, and as a result, finer structures are modeled by increasing the number of B-splines. However, the number of B-splines must be chosen according to the average sampling interval of the input data [84,90]. Assuming a constant sampling interval between the pierce points $\Delta\varphi_{pp}$, it was found in [84] that for a spherical surface the following relationship must hold:

$$\Delta\varphi_{pp} < \frac{\varphi_{max} - \varphi_{min}}{2^{J_\varphi} + 1}, \tag{5.20}$$

which leads to the following relationship

$$J_\varphi \leq \log_2 \left(\frac{\varphi_{max} - \varphi_{min}}{\Delta\varphi_{pp}} - 1 \right). \tag{5.21}$$

Since the pierce points are not spaced at constant intervals, it is preferred to use the average pierce point spacing $\Delta\overline{\varphi}_{pp}$ instead of $\Delta\varphi_{pp}$ when determining the appropriate resolution level in eq. (5.21).

The B-splines for the longitude are similarly constructed by the analysis above.

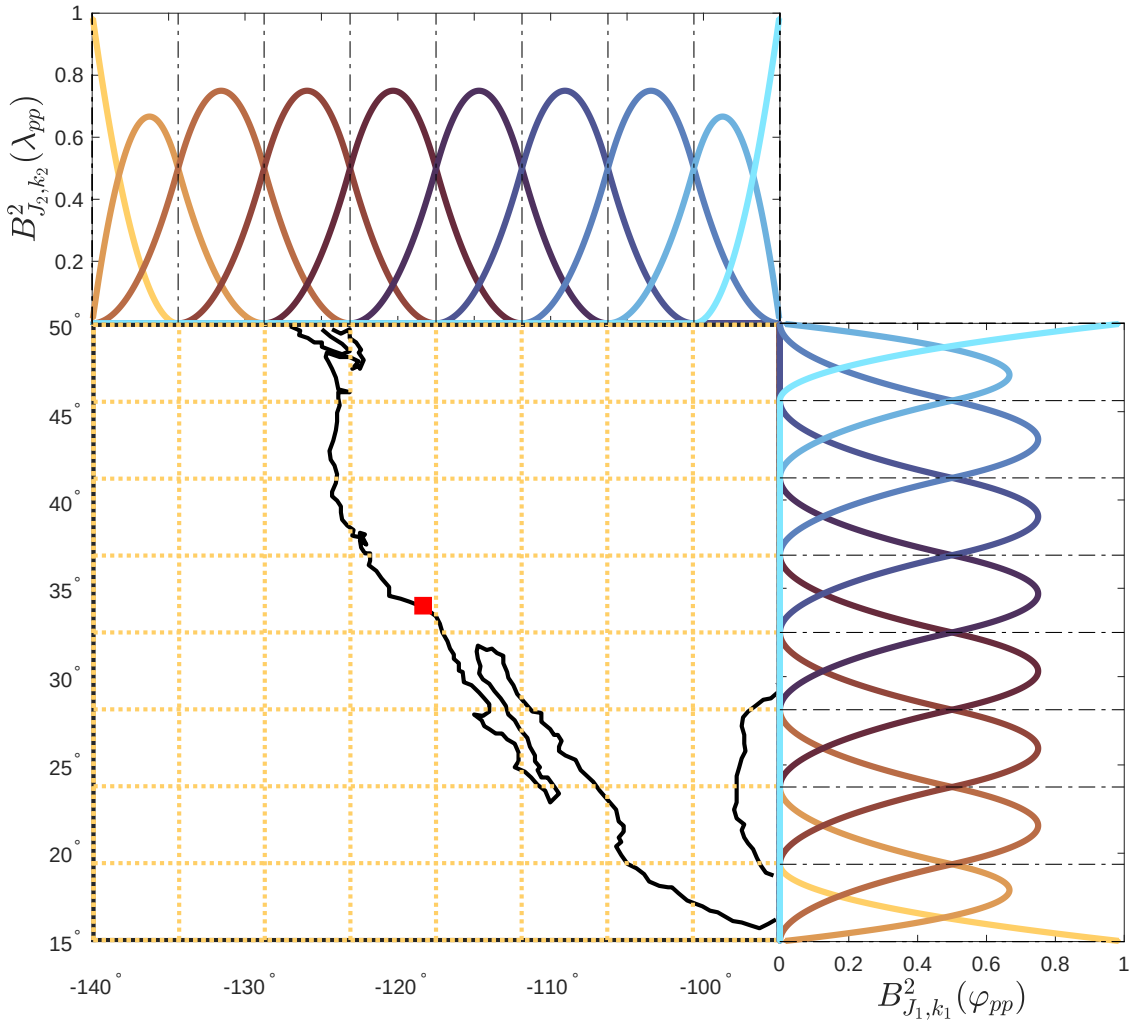


Figure 5.4: Polynomial (quadratic) B-spline basis functions covering a regional area around a user, with an elevation cutoff of 15°. The region covers an area $G = [15°, 50°] \times [-95°, -140°]$ (latitude × longitude).

Fig. 5.4 shows polynomial B-splines of degree $n = 2$ and resolution level $J_\varphi = J_\lambda = 3$, covering an area of $G = [-15°, 50°] \times [-95°, -140°]$.

5.4 Independent Estimation

This section revisits the generalized multi-GNSS and multi-frequency measurement models of Section 4.3.3, then presents a comprehensive design of the Kalman filter to estimate

the states the resulting dynamic system. For simplicity and clarity of subsequent analysis, the following set of assumptions are necessary:

Assumption 5.4.1.

1. *There are k GNSS systems, two (2) frequencies per system, and exactly ℓ usable satellites per frequency. Each usable satellite provides both code and phase measurements. In total, there are $2 * 2 * k * \ell = 4\,k\,\ell$ GNSS measurements from the receiver.*

2. *There are L additional GIM pseudo-measurements taken from points in the regional area defined by Def. 5.3.1.*

3. *There are K VTEC B-spline parameters obtained using the procedure outlined Sec. 5.3.2. These parameters are estimated as unknown states.*

Assumption 5.4.1 is mostly useful for simplifying and generalizing the subsequent theoretical analysis. In practice, no such assumption is made, i.e. it is always possible to process more or less than two frequencies per system, and there's almost always an unequal number of satellites per system.

5.4.1 GNSS Measurement Models, Revisited

Based on Assumption 5.4.1, the GNSS measurement models in eqs. (4.36) and (4.37) can be re-written to include only necessary terms for our processing. When only two frequencies are processed, the aforementioned equations for satellite s and GNSS constellation (system)

k are:

$$\rho_{1,r}^{s,sys_k} = \mathbf{1}_r^s \, \mathbf{p}_r + \widetilde{dt}_r^{sys_k} + T_{r,w}^s + I_{r,1}^s \tag{5.22}$$

$$\phi_{1,r}^{s,sys_k} = \mathbf{1}_r^s \, \mathbf{p}_r + \widetilde{dt}_r^{sys_k} + T_{r,w}^s - I_{r,1}^s + \lambda_1 \, \widetilde{N}_{r,1}^{s,sys_k} \tag{5.23}$$

$$\rho_{2,r}^{s,sys_k} = \mathbf{1}_r^s \, \mathbf{p}_r + \widetilde{dt}_r^{sys_k} - \delta b_{r,12}^{sys_k} + T_{r,w}^s + \gamma_{1,2}^{sys_k} \, I_{r,1}^s \tag{5.24}$$

$$\phi_{2,r}^{s,sys_k} = \mathbf{1}_r^s \, \mathbf{p}_r + \widetilde{dt}_r^{sys_k} - \delta b_{r,12}^{sys_k} + T_{r,w}^s - \gamma_{1,2}^{sys_k} \, I_{r,1}^s + \lambda_2 \, \widetilde{N}_{r,2}^{s,sys_k}, \tag{5.25}$$

where $\mathbf{1}_r^s \in \mathbb{R}^{1\times 3}$ is the LOS unit vector linearized around the estimated receiver position $\mathbf{p}_r$, and is defined as:

$$\mathbf{1}_r^s = \frac{\partial \rho_r^s}{\partial \mathbf{x}}\bigg|_{\mathbf{x}=\hat{\mathbf{x}}} = \frac{\hat{\mathbf{p}}_r - \hat{\mathbf{p}}^s}{\|\hat{\mathbf{p}}_r - \hat{\mathbf{p}}^s\|}, \tag{5.26}$$

and $\widetilde{N}_{r,j}^s$ for $j = 1, 2$ is as defined in eq. (4.38). The atmospheric delay error terms in eqs. (5.22)–(5.25) for frequency f_j are defined as:

$$T_{r,w}^s = MF_w(\beta_s)\,\mathrm{ZWD}, \tag{5.27}$$

$$\begin{aligned} I_r^s &= \frac{40.3}{f_j^2}\, M(z_s)\, VTEC \\ &= \frac{40.3}{f_j^2}\, M(z_s)\, \left(\mathbf{\Psi}^{J_\lambda, J_\varphi}\right)^\top \, \vec{\boldsymbol{d}}_{J_\varphi, J_\lambda}. \end{aligned} \tag{5.28}$$

where β_s is the elevation angle for satellite s, z_s is the zenith angle, and $\varphi_{pp,\ell}$ and $\lambda_{pp,\ell}$ are the geographic latitude and longitude of the corresponding pierce point. The ZWD mapping function $MF_w(\beta_s)$ was defined in eq. (3.9) of Section 3.1.3 and the ionospheric mapping functions $M(z_s)$ was defined in eq. (3.12) of 3.2. The vector of B-spline coefficients $\vec{\boldsymbol{d}}_{J_\varphi, J_\lambda}$ after eq. (5.14) can be defined as:

$$\vec{\boldsymbol{d}}_{J_\varphi, J_\lambda} = [d_1, \ldots, d_K]^\top \in \mathbb{R}^{K\times 1}, \tag{5.29}$$

where $K = K_\varphi K_\lambda$.

The measurement vector is:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y_a} \end{bmatrix} \in \mathbb{R}^{(m+L)\times 1}, \tag{5.30}$$

where

$$\mathbf{y}_i = \left[\overbrace{\rho_{r,i}^{1,sys_1}, \phi_{r,i}^{1,sys_1}, \cdots, \rho_{r,i}^{\ell,sys_1}, \phi_{r,i}^{\ell,sys_1}}^{\mathbf{y}_{i,1}\in\mathbb{R}^{2\,\ell\times 1}}, \cdots, \overbrace{\rho_{r,i}^{1,sys_k}, \phi_{r,i}^{1,sys_k}, \cdots, \rho_{r,i}^{\ell,sys_k}, \phi_{r,i}^{\ell,sys_k}}^{\mathbf{y}_{i,k}\in\mathbb{R}^{2\,\ell\times 1}} \right]^\top_{i=1,2} \in \mathbb{R}^{2\,k\,\ell\times 1}, \tag{5.31}$$

Under Assumption 5.4.1, the total number of code and phase measurements is $m = 4\,k\,\ell$, which corresponds to 2ℓ code and phase measurements from ℓ satellites, for all 2 frequencies of each of the k GNSS systems.

Code and phase measurements from one satellite alternate, with the code before phase measurements. In addition, measurements from the same frequency are grouped together. For instance, if the processing involves only GPS L1 and L2 measurements, then measurements from the L1 frequency come first, followed by those from L2. If more than one system is processed, and each system has more than one frequency, then each system is considered to have a *first* and *second* frequency. The measurements from the first frequency of each system are then grouped together. For example, if GPS L1/L2/L5 measurements are processed together with GAL E1/E5a/E5b measurements, then both L1 and E1 can be considered first frequencies, L2 and E5a second frequencies, and finally L5 and E5b third frequencies. It matters less which specific frequencies are grouped together, as long as the frequency-dependent terms in eq. (5.22)–(5.25) for each measurement are handled properly.

In other words, L2 and E5b can be grouped together just as well as the grouping of L2 and E5a.

5.4.2 GIM Pseudo-measurements

There are benefits of adding external ionospheric corrections as pseudo-observations to the measurement model. This maintains the estimate of the VTEC model in regions without pierce points and adds to the accuracy in regions where pierce points exist. The atmospheric delay pseudo-observations $\mathbf{y_a}$ are:

$$\mathbf{y_a} = \left[y_1^G, y_2^G, \cdots, y_L^G\right]^\top \in \mathbb{R}^{L\times 1}, \tag{5.32}$$

where $y_L^G = VTEC(\varphi_{pp,L}, \lambda_{pp,L})$, and $(\varphi_{pp,L}, \lambda_{pp,L})$ is a pierce point geographic coordinates. (The "G" superscript indicates that these pseudo-observations are obtained from GIM.) These pseudo-observations y_L^G are the GIM VTEC values corresponding to evenly-spaced (pseudo-) pierce points in the regional area defined in Def. 5.3.1. They are pseudo-observations, since these pierce points do not necessarily correspond to any visible satellites. Theoretically, by dividing the regional area into a finite grid, one could use any number of VTEC pseudo-observations. The pierce points must be chosen such that the observation submatrix corresponding to $\mathbf{y_a}$ does not have any all-zero columns, i.e. each B-spline parameter is activated at least once. In this work, the regional area's grid knots (shown in Fig. 5.4) were chosen as pierce points to ensure that each of the pseudo-observations corresponds to J_φ B-spline parameters. (See Section 5.4.4 for more details on the corresponding observation submatrix.)

5.4.3 State Vector

Recall from eq. (4.6) that the general PPP state vector is defined as:

$$\mathbf{x} = [\mathbf{p}, \mathbf{c}, \mathbf{n}, \mathbf{a}]^\top \in \mathbb{R}^{n \times 1}, \tag{5.33}$$

where each portion is discussed below. The total number of unknowns in the state vector is $n = \overbrace{3}^{\mathbf{p}} + \overbrace{2k}^{\mathbf{c}} + \overbrace{2k\ell}^{\mathbf{n}} + \overbrace{K+1}^{\mathbf{a}}$.

The receiver *clock vector* $\mathbf{c}$, which contains the clock error of the reference GNSS system, the clock offsets of other systems, and their respective DCBs, is defined as:

$$\mathbf{c} = \left[\widetilde{dt}_r^{\,sys_1}, \delta b_{12}^{sys_1}, \Delta t_r^{sys_2}, \delta b_{12}^{sys_2}, \cdots, \Delta t_r^{sys_k}, \delta b_{12}^{sys_k}\right]^\top \in \mathbb{R}^{(2\,k)\times 1}. \tag{5.34}$$

The reference clock error $dt_r^{sys_1}$ is that of the reference GNSS constellation, which is typically GPS due to its maturity and ubiquity in almost all GNSS receivers. If GPS is not available, the first processed GNSS constellation's clock will be used as the reference. The clock error of any other processed constellation is modeled as the reference clock error plus the constellation's constant clock offset:

$$\widetilde{dt}_r^{\,sys_i} = \widetilde{dt}_r^{\,sys_1} + \Delta t_r^{sys_i}, \quad \text{for } i = 2, \ldots, k; \tag{5.35}$$

where $\widetilde{dt}_r^{\,sys_1}$ was defined in eq. (4.24) and $\Delta t_r^{sys_i}$ was defined in eq. (2.5). The receiver DCB $\delta b_{12}^{sys_k}$ was defined in eq. (2.6).

The ambiguity vector, which contains the *float* ambiguities of all signals, is defined as:

$$\mathbf{n} = [\mathbf{n}_1, \mathbf{n}_2]^\top \in \mathbb{R}^{(2\,k\,\ell)\times 1}, \tag{5.36}$$

where

$$\mathbf{n}_i = \left[\overbrace{\widetilde{N}_{r,i}^{1,sys_1}, \widetilde{N}_{r,i}^{2,sys_1}, \cdots, \widetilde{N}_{r,i}^{\ell,sys_1}}^{\mathbf{n}_{i,1} \in \mathbb{R}^{\ell \times 1}}, \cdots, \overbrace{\widetilde{N}_{r,i}^{1,sys_k}, \widetilde{N}_{r,i}^{2,sys_k}, \cdots, \widetilde{N}_{r,i}^{\ell,sys_k}}^{\mathbf{n}_{i,k} \in \mathbb{R}^{\ell \times 1}} \right]_{i=1,2}^{\top} \in \mathbb{R}^{(k\,\ell) \times 1}, \tag{5.37}$$

Finally, the atmospheric delay vector is:

$$\mathbf{a} = [\text{ZWD}, d_1, d_2, \cdots, d_K]^{\top} \in \mathbb{R}^{(K+1) \times 1}. \tag{5.38}$$

5.4.4 Observation Matrix

The observation matrix $\mathbf{H}$, for m GNSS measurements, L GIM pseudo-measurements, and n states is:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H_p} & \mathbf{H_c} & \mathbf{H_n} & \mathbf{H_a} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{H_a^G} \end{bmatrix} \in \mathbb{R}^{(m+L) \times n}, \tag{5.39}$$

where the corresponding submatrices will be further explained below.

The matrix $\mathbf{H_p}$ comprises m rows which are the three-element LOS vectors for each measurement.

$$\mathbf{H_p} = \begin{bmatrix} \mathbf{1}_r^s \\ \mathbf{1}_r^s \\ \vdots \\ \mathbf{1}_r^s \end{bmatrix} \in \mathbb{R}^{m \times 3}, \tag{5.40}$$

where $\mathbf{1}_r^s$ was defined in eq. (5.26). The clock matrix $\mathbf{H_c}$ comprises two submatrices, which correspond to the two frequencies.

$$\mathbf{H_c} = \begin{bmatrix} \mathbf{H_{c_1}} \\ \mathbf{H_{c_2}} \end{bmatrix} \in \mathbb{R}^{m \times (2\,k)}. \tag{5.41}$$

The submatrix for the first frequency can be written as:

$$
\mathbf{H}_{\mathbf{c}_1} = \begin{array}{cl}
\begin{array}{cccc}
\overset{\left[\widetilde{dt}_r^{sys_1} \quad \delta b_{12}^{sys_1}\right]}{\longleftrightarrow} & \overset{\left[\Delta t_r^{sys_2} \quad \delta b_{12}^{sys_2}\right]}{\longleftrightarrow} & \overset{...}{\longleftrightarrow} & \overset{\left[\Delta t_r^{sys_k} \quad \delta b_{12}^{sys_k}\right]}{\longleftrightarrow}
\end{array} & \\
\left[\begin{array}{c|c|c|c}
\begin{matrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{matrix} & \mathbf{0} & \cdots & \mathbf{0} \\ \hline
\begin{matrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{matrix} & \begin{matrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{matrix} & \cdots & \mathbf{0} \\ \hline
\vdots & \mathbf{0} & \ddots & \mathbf{0} \\ \hline
\begin{matrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{matrix} & \mathbf{0} & \cdots & \begin{matrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{matrix}
\end{array}\right] & \begin{array}{l} \updownarrow \mathbf{y}_{1,1} \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array}
\end{array}
\in \mathbb{R}^{(2\,k\,\ell)\times(2\,k)}. \quad (5.42)
$$

Similarly, the submatrices for the second frequency can be written as:

$$
\mathbf{H}_{\mathbf{c}_2} = \left[\begin{array}{cc|c|c|c}
\multicolumn{2}{c|}{\overbrace{\left[\widetilde{dt}_r^{\,sys_1} \quad \delta b_{12}^{sys_1}\right]}} & \overbrace{\left[\Delta t_r^{sys_2} \quad \delta b_{12}^{sys_2}\right]} & \overbrace{\quad\cdots\quad} & \overbrace{\left[\Delta t_r^{sys_k} \quad \delta b_{12}^{sys_k}\right]} \\
\begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} & \begin{matrix} -1 \\ \vdots \\ -1 \end{matrix} & \mathbf{0} & \cdots & \mathbf{0} \\
\hline
\begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} & \begin{matrix} 1 & -1 \\ \vdots & \vdots \\ 1 & -1 \end{matrix} & \cdots & \mathbf{0} \\
\hline
\multicolumn{2}{c|}{\vdots} & \mathbf{0} & \ddots & \mathbf{0} \\
\hline
\begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} & \mathbf{0} & \cdots & \begin{matrix} 1 & -1 \\ \vdots & \vdots \\ 1 & -1 \end{matrix}
\end{array}\right] \begin{matrix} \\ \\ \updownarrow y_{2,1} \\ \\ \\ \\ \\ \end{matrix} \in \mathbb{R}^{(2\,k\,\ell)\times(2\,k)}.
$$

(5.43)

The submatrices of $\mathbf{H}_{\mathbf{c}_1}$ and $\mathbf{H}_{\mathbf{c}_2}$ are further subdivided into k^2 submatrices, each of which has the dimensions $\mathbf{H}_{\mathbf{c}_{1,k}} \in \mathbb{R}^{2\ell\times 2}$. The k diagonal submatrices correspond to the receiver clock error/offset in the first column, and the receiver DCB for that constellation in the second column.

The matrix $\mathbf{H}_{\mathbf{n}}$ corresponds to the phase ambiguities of all processed frequencies of all

processed GNSS systems.

$$\mathbf{H_n} = \begin{bmatrix} \mathbf{H_{n_1}} \\ \mathbf{H_{n_2}} \end{bmatrix} \in \mathbb{R}^{m \times (2\,k\,\ell)}, \tag{5.44}$$

where

$$\mathbf{H_{n_i}} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ \lambda_i & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \lambda_i & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix} \in \mathbb{R}^{(2\,k\,\ell)\times(2\,k\,\ell)}, \tag{5.45}$$

for $i = 1, 2$. The braces indicate that the two rows correspond to a pair measurements (code and phase) from the same signal. We note that there are alternating all zero rows and rows which have a single non-zero element. Moreover, the non-zero rows have elements in diagonal places. As a result, we should call $\mathbf{H_n}$ an *odd-even* diagonal matrix, i.e. the odd rows are all zero row vectors (because the corresponding code measurements do not have ambiguities), and the even rows would form a diagonal matrix in the absence of the odd rows.

The matrix $\mathbf{H_a}$ corresponds to the contribution of GNSS measurements to the atmo-

spheric delay parameters, and it can be written as:

$$\mathbf{H_a} = \begin{bmatrix} \mathbf{H_{a_1}} \\ \\ \mathbf{H_{a_2}} \end{bmatrix} \in \mathbb{R}^{m \times K}, \tag{5.46}$$

where

$$\mathbf{H_{a_i}} = \begin{bmatrix} \mathbf{h}^{sys_1}_{\mathbf{a}_i,1} \\ \vdots \\ \mathbf{h}^{sys_1}_{\mathbf{a}_i,\ell} \\ \vdots \\ \mathbf{h}^{sys_k}_{\mathbf{a}_i,1} \\ \vdots \\ \mathbf{h}^{sys_k}_{\mathbf{a}_i,\ell} \end{bmatrix}_{i=1,2} \in \mathbb{R}^{(2\,k\,\ell)\times K}. \tag{5.47}$$

Each submatrix $\mathbf{h}^{sys_k}_{\mathbf{a}_i,\ell}$ corresponds to the estimated atmospheric delays of the ℓ-th satellite of frequency k:

$$\mathbf{h}^{sys_k}_{\mathbf{a}_i,\ell} = \begin{bmatrix} MF_w(\beta_\ell) & \frac{40.3}{f_1^2}\,\gamma^{sys_k}_{1,i}\,M(z_\ell)\left(\mathbf{\Psi}^{J_\lambda,J_\varphi}\right)^\top \\ MF_w(\beta_\ell) & -\frac{40.3}{f_1^2}\,\gamma^{sys_k}_{1,i}\,M(z_\ell)\left(\mathbf{\Psi}^{J_\lambda,J_\varphi}\right)^\top \end{bmatrix} \in \mathbb{R}^{2\times(K+1)}, \tag{5.48}$$

where $\mathbf{\Psi}^{J_\lambda,J_\varphi} = \mathbf{\Psi}^{J_\lambda}(\lambda_{pp,\ell}) \otimes \mathbf{\Psi}^{J_\varphi}(\varphi_{pp,\ell})$ (as defined in eq. (5.14)), and $\varphi_{pp,\ell}$ and $\lambda_{pp,\ell}$ are the latitude and longitude of the ℓ-th satellite's pierce point.

The matrix $\mathbf{H}^G_\mathbf{a}$ corresponds to the contribution of pseudo-observations to the VTEC B-spline parameters. Specifically, eq. (5.32) can be written as:

$$\mathbf{y_a} = \mathbf{H}^G_\mathbf{a}\,\mathbf{a}, \tag{5.49}$$

	States	Unknowns	Total Number
Observations	–	–	$4k\ell + K$
Parameters	**p**	3	$4 + 2k + 2k\ell + L$
	c	$2k$	
	n	$2k\ell$	
	a	$(K+1)$	
Redundancy	–	–	$2k(\ell-1) - 4 + (L-K)$

Table 5.1: Redundancy of the estimation system.

where $\mathbf{a}$ is defined in eq. (5.38). Then, using the relationship in eq. (5.14), we get:

$$\mathbf{H}_{\mathbf{a}}^{G} = \begin{bmatrix} 0 & \left(\mathbf{\Psi}^{J_\lambda, J_\varphi}(\lambda_{pp,1}, \varphi_{pp,1})\right)^\top \\ \vdots & \vdots \\ 0 & \left(\mathbf{\Psi}^{J_\lambda, J_\varphi}(\lambda_{pp,L}, \varphi_{pp,L})\right)^\top \end{bmatrix} \in \mathbb{R}^{L \times (K+1)}, \tag{5.50}$$

where $\mathbf{\Psi}^{J_\lambda, J_\varphi}(\lambda_{pp,L}, \varphi_{pp,L}) \in \mathbb{R}^{1 \times K}$ is a vector of B-spline basis functions at the L-th knot point.

Table 5.1 shows that under Assumption 5.4.1, the observation matrix $\mathbf{H}$ has positive redundancy as long as measurements from more than one signal are used, i.e. $\ell > 1$. In practice, several satellites are necessary to have good geometry, among other benefits such as increased protection against the effects of faulty measurements [9].

5.4.5 Filter Design

This subsection presents the PPP problem in terms of a linear dynamic system, then describes the design of the Kalman Filter (KF) used to estimate the state.

5.4.5.1 Discrete-time System Model

Consider the state vector defined in eq. (5.33). Each element of the state vector is modeled as a discrete-time GM process. The state vector at epoch k is written as $\mathbf{x}_k$, and

is described by the following dynamical linear system:

$$\mathbf{x}_{k+1} = \mathbf{\Phi}_k \, \mathbf{x}_k + \boldsymbol{\omega}_k \tag{5.51}$$

$$\mathbf{y}_k = \mathbf{H}_k \, \mathbf{x}_k + \boldsymbol{\nu}_k, \tag{5.52}$$

where $\mathbf{y}_k \in \mathbb{R}^m$, is defined in eq. (5.30). The observation matrix $\mathbf{H}_k$ has been defined in eq. (5.39) and discussed in more detail in the previous subsection. It is computed as the partial derivative of the non-linear observation function $\boldsymbol{h}_k$ with respect to the state vector, evaluated at the state prior:

$$\mathbf{H}_k = \left. \frac{\partial \mathbf{h}_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_k^-}. \tag{5.53}$$

The non-linearity in $\boldsymbol{h}_k$ is due to the range function in eqs. (4.36) and (4.37), and it is linearized into the LOS vector in eq. (5.26). The linear system in eq. (5.51) is a GM process because: (1) it is Gaussian (assuming initial Gaussian conditions), and (2) its driving noise is white (see Section 4.6.3 in [61]).

The linear model described by eqs. (5.51) and (5.52) assumes that all systematic biases in the measurements and the system model are accounted for, either by estimating them or using appropriate models. As a result, the remaining errors only have random characteristics and can be accommodated by zero-mean white noise processes with known covariance and zero cross-correlation. The process (driving) noise and measurement noise vectors are modeled as:

$$\boldsymbol{\omega}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \tag{5.54}$$

$$\boldsymbol{\nu}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \tag{5.55}$$

where $\mathbf{Q} = E\left\langle \boldsymbol{\omega}_k \, \boldsymbol{\omega}_k^\top \right\rangle \in \mathbb{R}^{n \times n}$ and $\mathbf{R} = E\left\langle \boldsymbol{\nu}_k \, \boldsymbol{\nu}_k^\top \right\rangle \in \mathbb{R}^{m \times m}$ are the respective process and measurement noise covariance matrices, which are assumed to be positive semi-definite.

The process noise $\boldsymbol{\omega}_k$ in eqs. (5.51) and (5.54) models the unknown uncertainty introduced by unmodeled random effects during state propagation. The process noise matrix $\mathbf{Q}$ then serves to *inflate* the propagated state covariance, and thus has to be positive semi-definite. Since the state's evolution is assumed to be a Markov process with uncorrelated process noise terms, then $\mathbf{Q}$ is a diagonal matrix. Small amounts of process noise imply that the dynamics of the state are well-known and modeled accordingly by the transition matrix $\boldsymbol{\Phi}$ in eq. (5.51). New measurements will have less effect on the current state, which could lead to the filter divergence if the state dynamics model turns out to be inaccurate. To avoid filter divergence, the covariance is increased during the propagation step according to the driving white noise, modeled by the process noise matrix $\mathbf{Q}$ in eq. (5.58). The additional process noise prevents the state covariance to drop below a certain threshold, which would lead to measurement rejection. On the other hand, large amounts of process noise imply that the state dynamics modeled by $\boldsymbol{\Phi}$ are not sufficient to capture the actual state evolution. The impact of past measurements will be reduced in favor of new measurements.

5.4.5.2 Kalman Filter

The optimal recursive filter for the system defined in eqs. (5.51) and (5.52) is the KF, which detailed below[1]. The KF algorithm involves two steps: prediction (in the form of time propagation) and correction (in the form of measurement update). In the first step, the KF predicts the state using a state transition model. The KF advances the state from

[1]↑A more complete treatment of the Kalman filter and optimal linear filtering can be found in Chapter 5 of [61].

epoch k to epoch $k+1$ using the following relationship:

$$\hat{\mathbf{x}}_{k+1}^{-} = \boldsymbol{\Phi}_k \, \hat{\mathbf{x}}_k^{+}. \tag{5.56}$$

The superscript '$-$' denotes the state vector prior to the measurement update, whereas the superscript '$+$' indicates the state vector after incorporating measurements. The transition matrix $\boldsymbol{\Phi}_k \in \mathbb{R}^{n \times n}$ propagates the state forward in time with respect to the modeled dynamics of the system.

The KF uses the error covariance matrix $\mathbf{P}$ to track the state vector's estimation error. Define the error of the propagated state estimate as:

$$\begin{aligned} \mathbf{e}_{k+1} &= \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1}^{-} \\ &= (\boldsymbol{\Phi}_k \, \mathbf{x}_k + \boldsymbol{\omega}_k) - \boldsymbol{\Phi}_k \, \hat{\mathbf{x}}_k^{+} \\ &= \boldsymbol{\Phi}_k \, \mathbf{e}_k + \boldsymbol{\omega}_k. \end{aligned} \tag{5.57}$$

Then, the covariance of the propagated error is:

$$\begin{aligned} \mathbf{P}_{k+1}^{-} &= E\left\langle (\mathbf{e}_{k-1}) \, (\mathbf{e}_k)^{\top} \right\rangle \\ &= E\left\langle (\boldsymbol{\Phi}_k \, \mathbf{e}_k + \boldsymbol{\omega}_k) \, (\boldsymbol{\Phi}_k \, \mathbf{e}_k + \boldsymbol{\omega}_k)^{\top} \right\rangle \\ &= E\left\langle (\boldsymbol{\Phi}_k \, \mathbf{e}_k) \, (\boldsymbol{\Phi}_k \, \mathbf{e}_k)^{\top} \right\rangle + E\left\langle \boldsymbol{\omega}_k \, \boldsymbol{\omega}_k^{\top} \right\rangle \\ \mathbf{P}_{k+1}^{-} &= \boldsymbol{\Phi}_k \, \mathbf{P}_k^{+} \boldsymbol{\Phi}_k^{\top} + \mathbf{Q}, \end{aligned} \tag{5.58}$$

where $\mathbf{Q}$ is the process noise covariance matrix defined in eq. (5.54). Both eqs. (5.56) and (5.58) complete the state propagation step.

The measurement update uses the Kalman gain, $\mathbf{K}$. The gain determines the magnitude of the corrections applied to the state vector during the measurement update. The Kalman

gain at epoch k is computed according to:

$$\mathbf{K}_k = \mathbf{P}_k^- \, \mathbf{H}_k^\top \left(\mathbf{H}_k \, \mathbf{P}_k^- \, \mathbf{H}_k^\top + \mathbf{R} \right)^{-1} \tag{5.59}$$

After computing $\mathbf{K}_k \in \mathbb{R}^{n \times m}$, the state vector is updated using:

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left(\mathbf{y}_k - \hat{\mathbf{y}}_k \right), \tag{5.60}$$

where $\hat{\mathbf{y}}_k$ is the predicted measurement vector computed using the state prior:

$$\hat{\mathbf{y}}_k = \mathbf{H}_k \, \hat{\mathbf{x}}_k^-. \tag{5.61}$$

The term $\mathbf{R} = \mathbf{y}_k - \hat{\mathbf{y}}_k$ in eq. (5.60) is the measurement residual, whose covariance is:

$$\mathbf{S}_k = \mathbf{H}_k \, \mathbf{P}_k^- \, \mathbf{H}_k^\top + \mathbf{R}. \tag{5.62}$$

This term $\mathbf{R}$ is also referred to as the *innovation* to reflect the fact that it accounts for the new *information* brought by the measurements to the state's estimate. Indeed, eq. (5.59) shows that the optimal gain is a product of $\mathbf{H}_k$, the prior state covariance $\mathbf{P}_k^-$, and the information matrix $\mathbf{S}_k^{-1}$ corresponding to the measurement residual $\mathbf{R}$.

Similarly, the estimated state covariance is updated according to:

$$\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k \right) \mathbf{P}_k^-. \tag{5.63}$$

Note that implementing the computation in eq. (5.63) poses significant risk, since small rounding errors can cause $\mathbf{P}_k^+$ to lose its symmetric properties, which in turn can lead to catastrophic numerical instability. For that reason, this work uses the alternative Joseph's form formulation, which preserves $\mathbf{P}_k^+$ symmetry:

$$\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k \right) \mathbf{P}_k^- \left(\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k \right)^\top + \mathbf{K}_k \, \mathbf{R} \, \mathbf{K}_k^\top. \tag{5.64}$$

The downside of eq. (5.64) is the increased computation load compared to eq. (5.63), but it comes with guaranteed numerical stability, which is important when a filter is likely to run for a long time before it is reinitialized.

Table 5.2 summarizes the formulation of the KF [61]. The various formulations of the covariance update are equivalent; either one can be used, but they are suitable to different implementations and interpretations.

The covariance of the estimated VTEC is a function of the pierce point coordinates and the covariance of the estimated B-spline parameters. Using eq. (5.14), with $\mathbf{\Psi}^{J_\lambda, J_\varphi} = \mathbf{\Psi} = \mathbf{\Psi}(\varphi_{pp}, \lambda_{pp})$ and $\vec{\boldsymbol{d}}_{J_\varphi, J_\lambda} = \vec{\boldsymbol{d}}$, the variance of the VTEC at the pierce point $(\varphi_{pp}, \lambda_{pp})$ is:

$$\begin{aligned}
\operatorname{var}\big(VTEC(\varphi_{pp}, \lambda_{pp})\big) &= E\Big\langle \big(VTEC(\varphi_{pp}, \lambda_{pp}) - E\langle VTEC(\varphi_{pp}, \lambda_{pp})\rangle\big)^2 \Big\rangle && (5.65)\\
&= E\left\langle \left(\mathbf{\Psi}^\top \vec{\boldsymbol{d}} - E\left\langle \mathbf{\Psi}^\top \vec{\boldsymbol{d}} \right\rangle\right)^2 \right\rangle && (5.66)\\
&= E\left\langle \left(\mathbf{\Psi}^\top \left(\vec{\boldsymbol{d}} - E\left\langle \vec{\boldsymbol{d}} \right\rangle\right)\right)^2 \right\rangle && (5.67)\\
&= E\left\langle \mathbf{\Psi}^\top \left(\vec{\boldsymbol{d}} - E\left\langle \vec{\boldsymbol{d}} \right\rangle\right) \left(\vec{\boldsymbol{d}} - E\left\langle \vec{\boldsymbol{d}} \right\rangle\right)^\top \mathbf{\Psi} \right\rangle && (5.68)\\
&= \mathbf{\Psi}^\top E\left\langle \left(\vec{\boldsymbol{d}} - E\left\langle \vec{\boldsymbol{d}} \right\rangle\right) \left(\vec{\boldsymbol{d}} - E\left\langle \vec{\boldsymbol{d}} \right\rangle\right)^\top \right\rangle \mathbf{\Psi} && (5.69)\\
&= \mathbf{\Psi}^\top(\varphi_{pp}, \lambda_{pp})\ \mathbf{P}_{\vec{\boldsymbol{d}}, \vec{\boldsymbol{d}}}\ \mathbf{\Psi}(\varphi_{pp}, \lambda_{pp}), && (5.70)
\end{aligned}$$

where $\mathbf{P}_{\vec{\boldsymbol{d}}, \vec{\boldsymbol{d}}}$ is the marginalized covariance of the B-spline parameters from the full state covariance matrix $\mathbf{P}$.

5.4.6 Stochastic Modeling

The KF requires the knowledge of noise statistics for its initialization and tuning. This subsection describes how to determine the stochastic parameters of the KF. Specifically,

Initialization	$\hat{\mathbf{x}}_0^- = \mathrm{E}\langle \mathbf{x}_0 \rangle$ $\mathbf{P}_0^- = \mathrm{var}\left(\mathbf{x}_0^-\right)$
Kalman Gain	$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^\top \left(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^\top + \mathbf{R}\right)^{-1}$
Measurement update	$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left(\mathbf{y}_k - \hat{\mathbf{y}}_k\right)$
Covariance update	$\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right) \mathbf{P}_k^-$ $\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right) \mathbf{P}_k^- \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right)^\top + \mathbf{K}_k \mathbf{R} \mathbf{K}_k^\top$ $\mathbf{P}_k^+ = \mathbf{P}_k^- - \mathbf{K}_k \left(\mathbf{R} + \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^\top\right) \mathbf{K}_k^\top$ $\left(\mathbf{P}_k^+\right)^{-1} = \left(\mathbf{P}_k^-\right)^{-1} + \mathbf{H}_k^\top \mathbf{R}^{-1} \mathbf{H}_k$
Time propagation	$\hat{\mathbf{x}}_{k+1}^- = \mathbf{\Phi}_k \hat{\mathbf{x}}_k^+$ $\mathbf{P}_{k+1}^- = \mathbf{\Phi}_k \mathbf{P}_k^+ \mathbf{\Phi}_k^\top + \mathbf{Q}$

Table 5.2: Summary of the Kalman filter formulation.

it describes how the assumed GM model is used to determine the initial state covariance $\mathbf{P}_0$, the state transition matrix $\mathbf{\Phi}_k$, the process noise matrix $\mathbf{Q}$, and the measurement noise matrix $\mathbf{R}$.

Recall from Section 5.4.5.1 that each element of the state vector can be modeled as a scalar continuous-time GM process $x(t)$, written as:

$$\dot{x}(t) = -\frac{1}{\tau} x(t) + \omega(t), \tag{5.71}$$

where $\tau > 0$ is the time constant, and $\omega(t)$ is the white process noise with Power Spectral Density (PSD) Q. The variance of $x(t)$ is $P(t)$, and it was shown in Chapter 4 of [61] that the steady-state covariance $P(\infty)$ is:

$$P(\infty) = \sigma^2 = \frac{\tau}{2} Q. \tag{5.72}$$

When the steady-state covariance σ^2 and the time constant τ are known, the process noise PSD is computed using:

$$Q = \frac{2\,\sigma^2}{\tau}. \tag{5.73}$$

The continuous-time model for the state vector in eq. (5.33) can be written as:

$$\dot{\mathbf{x}} = \mathbf{F}\,\mathbf{x} + \mathbf{G}\,\boldsymbol{\omega}, \tag{5.74}$$

where $\boldsymbol{\omega}$ is a white noise process with PSD $\mathbf{Q}_t$. The discrete-time model in eq. (5.51) is equivalent to the continuous-time model in eq. (5.74) when sampled at discrete time instants $t_k = kT$, where $T = t_{k+1} - t_k$ is the sampling period.

The discrete-time state transition matrix $\boldsymbol{\Phi}_k$ in eq. (5.51) can be approximated by the first order Taylor series:

$$\boldsymbol{\Phi}_k = e^{\mathbf{F}\,T} \approx \mathbf{I} + \mathbf{F}\,T. \tag{5.75}$$

The above approximation only works when the sampling period is relatively small compared to the time constant of the GM process. For PPP, the sampling period is typically 1–30 seconds, which is relatively small compared to the time-constants of all states, i.e. $T \ll \tau$ for all states. (See Table 5.3 for time constants of different states of the state vector.) Consequently, the transition matrix can be further approximated by $\boldsymbol{\Phi}_k \approx \mathbf{I}$.

The discrete-time process noise matrix $\mathbf{Q}$ is usually approximated from the continuous $\mathbf{Q}_t$ by:

$$\mathbf{Q} \approx \mathbf{G}\,\mathbf{Q}_t\,\mathbf{G}^\top T. \tag{5.76}$$

Note that $\mathbf{Q}$ does not depend on the epoch k since we assume equal sampling time T between measurement update intervals, i.e. T is constant.

The process noise matrix of the state in eq. (5.33), for a sampling period T, can be

written as:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{\mathbf{p}_r} & & & & \\ & \mathbf{Q}_{\mathbf{c}} & & & \\ & & \mathbf{Q}_{\mathbf{n}} & & \\ & & & \mathbf{Q}_{\mathbf{a}} & \end{bmatrix} T \in \mathbb{R}^{n \times n}, \tag{5.77}$$

where $\mathbf{Q}_{\mathbf{p}_r} = \mathbf{0}_{3\times 3} \in \mathbb{R}^{3\times 3}$ and $\mathbf{Q}_{\mathbf{n}} = \mathbf{0}_{(2\,k\,\ell)\times(2\,k\,\ell)} \in \mathbb{R}^{(2\,k\,\ell)\times(2\,k\,\ell)}$ are the process noise submatrices of constant coordinate and ambiguity states, respectively. The process noise submatrices corresponding to scalar GM processes of the clock and atmosphere states in eqs. (5.34) and (5.38), respectively, are:

$$\mathbf{Q}_{\mathbf{c}} = 2 \begin{bmatrix} \frac{\sigma^2_{\widetilde{dt}}}{\tau_{\widetilde{dt}}} & 0 & \cdots & 0 \\ 0 & \frac{\sigma^2_{\delta b}}{\tau_{\delta b}} & & \vdots \\ \vdots & & \frac{\sigma^2_{\Delta t}}{\tau_{\Delta t}} & 0 \\ 0 & \cdots & 0 & \frac{\sigma^2_{\delta b}}{\tau_{\delta b}} \end{bmatrix} \in \mathbb{R}^{(2\,k)\times(2\,k)}, \tag{5.78}$$

$$\mathbf{Q}_{\mathbf{a}} = 2 \left[\begin{array}{c|c} \frac{\sigma^2_{T_w}}{\tau_{T_w}} & 0 \cdots 0 \\ \hline 0 \\ \vdots & \frac{\sigma^2_d}{\tau_d}\,\mathbf{I}_{K\times K} \\ 0 \end{array}\right] \in \mathbb{R}^{(K+1)\times(K+1)} \tag{5.79}$$

The initial state covariance $\mathbf{P}_0$ quantifies the expected errors when initializing the states, and it contains each state's squared initial standard deviations σ_0^2 on the main diagonal,

and it can be written as:

$$\mathbf{P}_0 = \begin{bmatrix} \mathbf{P}_{0,\mathbf{p}_r} & & & \\ & \mathbf{P}_{0,\mathbf{c}} & & \\ & & \mathbf{P}_{0,\mathbf{n}} & \\ & & & \mathbf{P}_{0,\mathbf{a}} \end{bmatrix} \in \mathbb{R}^{n \times n}, \tag{5.80}$$

where the submatrices corresponding to constant states are $\mathbf{P}_{0,p} = \sigma_{0,p}^2 \mathbf{I} \in \mathbb{R}^{3\times 3}$ and $\mathbf{P}_{0,N} = \sigma_{0,N}^2 \mathbf{I} \in \mathbb{R}^{(2\,k\,\ell)\times(2\,k\,\ell)}$. The submatrices corresponding to the clock and atmosphere states in eqs. (5.34) and (5.38), respectively, are

$$\mathbf{P}_{0,\mathbf{c}} = \begin{bmatrix} \sigma_{0,\tilde{dt}}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{0,\delta b}^2 & & \vdots \\ \vdots & & \sigma_{0,\Delta t}^2 & 0 \\ 0 & \cdots & 0 & \sigma_{0,\delta b}^2 \end{bmatrix} \in \mathbb{R}^{(2\,k)\times(2\,k)}, \tag{5.81}$$

$$\mathbf{P}_{0,\mathbf{a}} = \left[\begin{array}{c|c} \sigma_{0,T_w}^2 & 0 \cdots 0 \\ \hline 0 \\ \vdots \\ 0 \end{array} \begin{array}{c} \\ \\ \sigma_{0,d}^2 \mathbf{I}_{K\times K} \\ \\ \end{array}\right] \in \mathbb{R}^{(K+1)\times(K+1)} \tag{5.82}$$

Using the scalar GM process model in eq. (5.71), the time constant, initial variance, and steady-state variance of each state is discussed in the following list.

- Position coordinates: The position coordinates are assumed to be constant and have an initial variance $\sigma_{0,p}^2$. Since the coordinates are not changing, their time constant $\tau_{0,p} = \infty$, and their PSD is zero. The initial coordinates are assumed to be approximately known (or completely unknown), their initial variance is set to meter-level values.

- Receiver clock: The *clock* states (biases and offsets) are completely unknown initially, and their initial covariance is set to very large values. The PSD of these parameters depends on the quality of the receiver's oscillator. High-quality receivers such as survey-grade/geodetic receivers have better oscillators and are given smaller process noise PSD, compared to low-quality receivers found in consumer-grade products. The behavior of receiver clocks is largely unknown and unpredictable, hence a reasonable process noise assumes variations on the order of 100 m over 100 s.

- Receiver DCBs: The stability of DCBs over a day has been shown to be less than 0.3 ns/day (0.09 m/day) [13]. Due to this relatively small variation, the DCBs are modeled as constant, with a meter-level initial variance.

- Float ambiguities: Similar to coordinates and DCBs, the phase ambiguities are modeled as constants, with meter-level initial variance.

- Residual ZWD: The ZTD has a small variation of a few centimeters per hour [21,93]. As discussed in Section 3.1, almost all of this variation is attributable to the ZWD. Consequently, the residual ZWD is modeled as a GM process with a time constant τ_{T_w} and a steady state variance σ_{T_w}. For example, it was shown in [93] that the variance of 2–5 $\frac{mm}{\sqrt{h}}$ is applicable to most regions of the globe.

- The ionospheric B-spline parameters: The ionospheric delay has a variation of several meters per hour. Both the initial and steady-state variance values are accordingly set to meter-level values.

Table 5.3 summarizes the values for initial and steady-state variance, as well as time

State	Symbol[a]	$\tau[s]$	$\sigma_0\,[m]$	$\sigma\left[m/\sqrt{h}\right]$
Coordinates	p	∞	5	0
Clock error	$\widetilde{dt}$	100	30,000	100
Clock offset	Δt	3,600	30,000	30,000
DCB	δb	∞	3	0
Ambiguities	N	∞	20	0
Residual ZWD	T_w	3,600	0.1	0.002–0.005
B-spline parameters[b]	d	3,600	1–2	0.5–2

[a] These symbols have been modified from their counterparts in Section 5.4.3 to simplify notation.

[b] The variance values are in range delay units (meters). The range delay to VTEC conversion is: 1 m $\approx$ 6.16 TECu. The conversion assumes the GPS L1 signal from a satellite directly above the receiver at the equator.

Table 5.3: The modeled values of the time constant τ, the initial variance σ_0, and the steady-state covariance σ for all states. These three parameters are enough to characterize the continuous-time stochastic model of each state.

constants used for all states.

The measurement noise $\boldsymbol{\nu}_k$ in eqs.(5.52) and (5.55) models the uncertainty in the measurements. The values of the measurement noise covariance matrix $\mathbf{R}$ are obtained from the standard deviations of corresponding measurements. The standard deviations depend on the quality of the measurements, which in turn depends on factors such as the type of measurement (phase vs. code), satellite elevation, and signal strength (or carrier-to-noise ratio). Phase measurements are more accurate than code measurements, and are accordingly given smaller measurement noise. High-elevation satellites are given smaller measurement noise values, since they typically have better signal strength and fewer reflective surfaces.

The measurement variances σ^2_{meas} in $\mathbf{R}$ are the squared standard deviations, multiplied by the elevation-dependent weighting factor:

$$\sigma^2_{meas} = \left(\frac{1}{\sin\beta^2}\right)\sigma^2_{0,meas}, \tag{5.83}$$

Measurement type	Standard deviation
Pseudorange [m]	0.3–2
Carrier phase [m]	0.002–0.006
GIM pseudo-measurements [TECu]	2–8

Table 5.4: Typical ranges of standard deviations for the measurement noise of geodetic-grade receivers.

where σ^2_{meas} is the resulting measurement variance, $\sigma_{0,meas}$ is the measurement standard deviation, and β is the elevation of the corresponding satellite. The elevation-dependent weighting function in eq. (5.83) is used because measurements from low elevation satellites are degraded by increased noise and multipath. As a result, measurement residuals from survey-grade receivers are typically elevation-dependent. This is demonstrated by Figs. B.1 and B.2 of Appendix B, which show examples of dual-frequency code and phase measurement residuals, respectively, plotted as a function of satellite elevation. Moreover, measurements from very low elevation satellites are prone to other anomalies such as phase cycle slips; therefore they are excluded by applying an elevation mask. Other weighting functions were also considered[2].

The appropriate standard deviations of pseudorange and carrier-phase measurements were determined empirically by analyzing their corresponding PPP residuals. The standard deviation of GIM pseudo-measurements was obtained from the reported accuracy of GIM products [21]. Table 5.4 shows the typical range of standard deviations of code and phase measurements from geodetic-quality receivers and VTEC from IGS GIM products used in this work.

[2]↑For most receivers, signal strength and satellite elevation are highly correlated. Consequently, either one may be used for the purposes of modeling measurement noise [94]. However, unlike signal strength, satellite elevation is not a property of the receiver, and it is typically preferred when processing data from multiple receivers. This ensures homogeneity in noise modeling, since different receiver manufacturers might use different methods to determine signal strength values.

5.5 Experimental Results

The category of PPP approaches which estimate local tropospheric and ionospheric delays as unknowns is henceforth referred to as PPP with Local Atmospheric Modeling (LAM-PPP). In the following discussion,"LAM-PPP Spline" refers to the novel approach in which ionospheric delays are estimated as B-spline parameters. The conventional uncombined approach in the literature is referred to as "LAM-PPP Slant" to indicate that it estimates slant ionospheric delays for each used satellite.

For a fair comparison, both LAM-PPP Spline and LAM-PPP Slant approaches, use the same measurement and process noise matrices, $\mathbf{R}$ and $\mathbf{Q}$, respectively. The only difference are the values used for slant delays, which are approximated into their VTEC counterparts. For instance, the steady-state variance of slant delays $\sigma_I = 1$ m corresponds to the steady-state variance of B-spline parameters $\sigma_d = 6$ TECu. Additionally, the chosen stations all use the same receiver model (the Septentrio PolarX5), equipped with antennae from the same manufacturer.

The results presented herein are for dual-frequency GPS and GAL measurements, with a sampling interval of 15 seconds, unless otherwise stated. The archived GNSS observations were obtained as RINEX files from the GAGE Facility[3]. The external GIM corrections were acquired as from IGS as IONEX files. The tropospheric delay corrections were obtained using GPT3 model and VMF3 mapping functions [38]. Table 5.5 summarizes the details of processing individual CORS stations.

[3]↑The Geodetic Facility for the Advancement of Geoscience (https://www.unavco.org/) is operated by the EarthScope Consortium, on behalf of the National Science Foundation with support from NASA and the USGS.

Data source	Regional CORS stations
Observations	GPS: C1W + C2W, L1C + L2W GAL: C1C + C5Q, L1C + L5Q
Observation interval	15 s
Measurement stdev.	Code: 0.300 m, Phase: 0.003 m, GIM: 6 TECu
Measurement weighting	Elevation-dependent factor: $\frac{1}{\sin(elev)^2}$
Elevation mask	12°
Correction products	CODE (orbits, clocks, biases)
Ionosphere corrections	IGS Final product
Troposphere corrections	GPT3 [38]
Other corrections	Phase windup, relativistic effects, PCO/PCV, solid Earth tides, ocean loading
Cycle-slip detection	Absolute phase difference: $\lvert\phi^s_{r,1} - \phi^s_{r,2}\rvert$
B-spline parameters	Quadratic B-splines, degree $n = 2$ Resolution level: $J_\varphi = J_\lambda = 3$

Table 5.5: Processing details for individual stations.

5.5.1 Tropospheric Delay Estimates

As reference data for the tropospheric delay estimates, we use the IGS final ZTD products which are produced using other IGS final products (clocks and orbits) derived from all ACs [95]. The products have a latency of around 3 weeks, with temporal resolution of 5-min and a reported accuracy of 1.5–5 mm. Hence, this product provides suitable values for assessing the accuracy of the estimated ZTD.

Fig. 5.5 shows comparisons of estimated ZTD (ZTD_{est}), the modeled GPT3 ZTD (ZTD_{GPT3}), and the IGS Final ZTD product (ZTD_{IGS}) for two CORS stations in different latitude regions. The modeled and estimated ZTD are defined as:

$$ZTD_{GPT3} = ZHD_{GPT3} + ZWD_{GPT3}$$

$$ZTD_{est} = ZHD_{GPT3} + ZWD_{est},$$

where ZWD_{est} is the estimated wet delay defined in eq. (5.38). The modeled ZTD_{GPT3} (shown in black) does not account the temporal variations that occur throughout the day. Fig. 5.5 also shows the comparison results for station `BILL` on two different days. The difference $ZTD_{IGS} - ZTD_{est}$, whose statistics are shown in Table 5.5d, provides an assessment of ZTD_{est} accuracy. Our approach is able to accurately estimate the local tropospheric delay in different regions, and different environmental conditions.

It should be noted that experiments showed similarity of ZTD_{est} for both LAM-PPP Spline and LAM-PPP Slant, which is to be expected since they use similar models of residual ZWD.

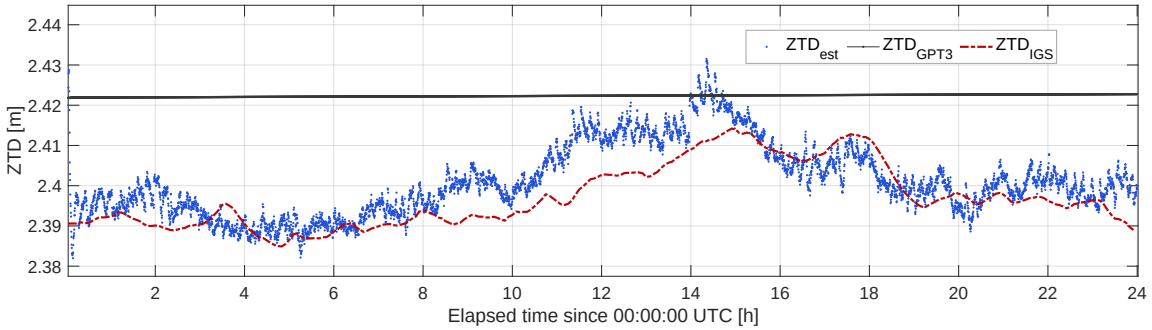
5.5.2 Ionospheric Delay Estimates

Unlike the tropospheric delay estimates, there is no equivalent of the IGS Final ZTD product which provides accurate reference values with which to assess the accuracy of our approach. Instead, in the next subsections, we will use measurements from multiple receivers to compare LAM-PPP Spline estimates to: (1) IGS GIM corrections, and (2) LAM-PPP Slant estimates.

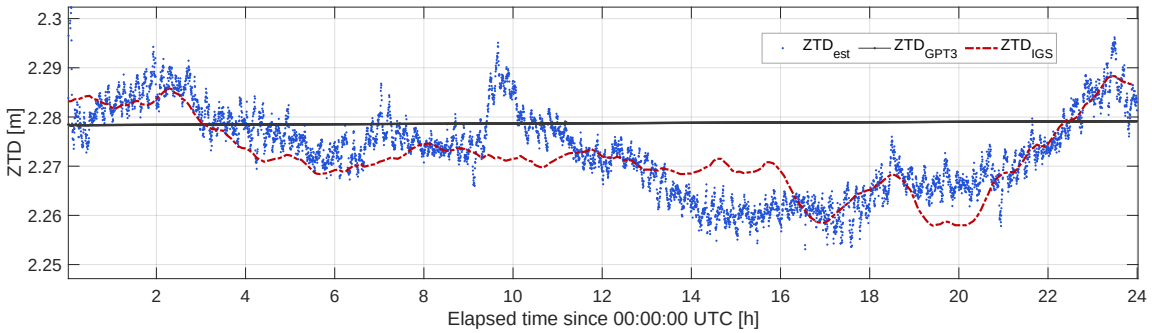
5.5.2.1 Estimated Regional VTEC Maps

The estimated B-spline parameters in eq. (5.38) can be used to construct a regional VTEC map using eq. (5.14). Such a map covers a regional area defined by Definition 5.3.1. The corresponding map of VTEC covariance can be obtained by eq. (5.70).

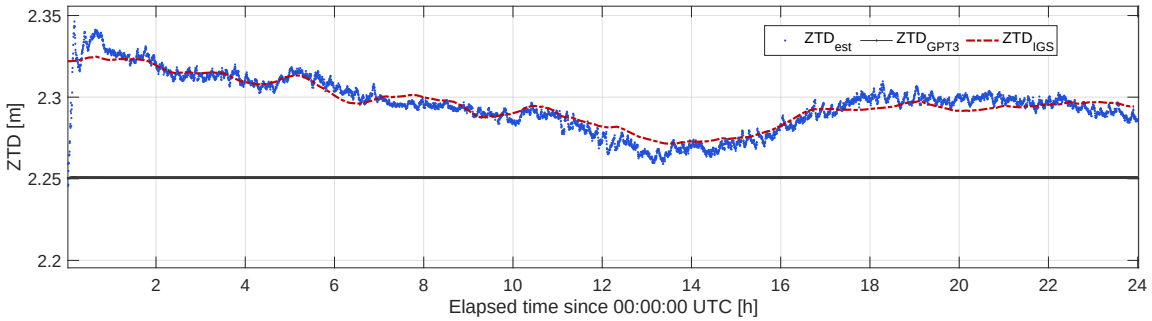
Fig. 5.6 shows an example of the estimated regional VTEC map for regions centered around receivers (green square) in California, United States and Libreville, Gabon. To give



(a) AC23(Anchorage, Alaska) in the summer.



(b) BILL(Temecula, California) in the summer.



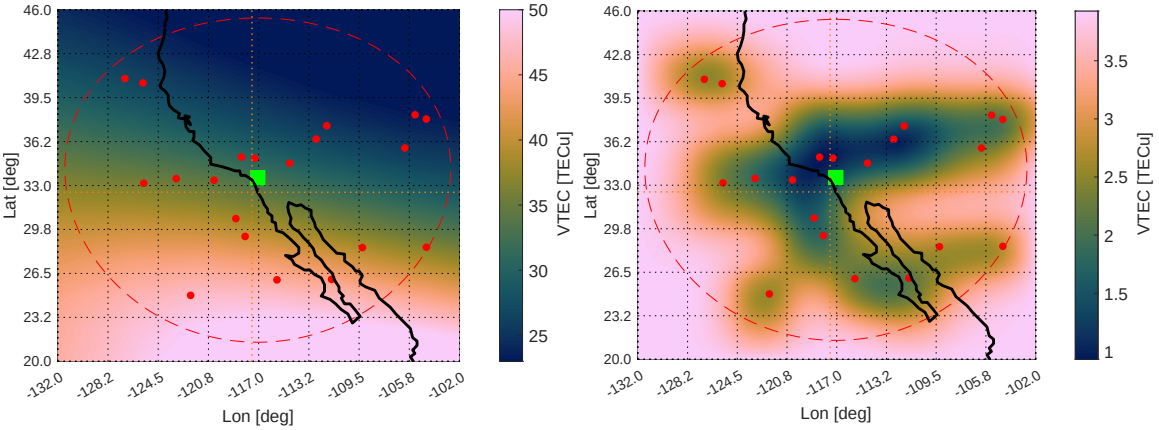
(c) BILL(Temecula, California) in the winter.

	Metric	AC23[a]	BILL[a]	BILL[b]
$ZTD_{IGS} - ZTD_{est}$	Mean [m]	0.004	0.001	−0.001
	Stdev [m]	0.006	0.006	0.006

[a] Summer
[b] Winter

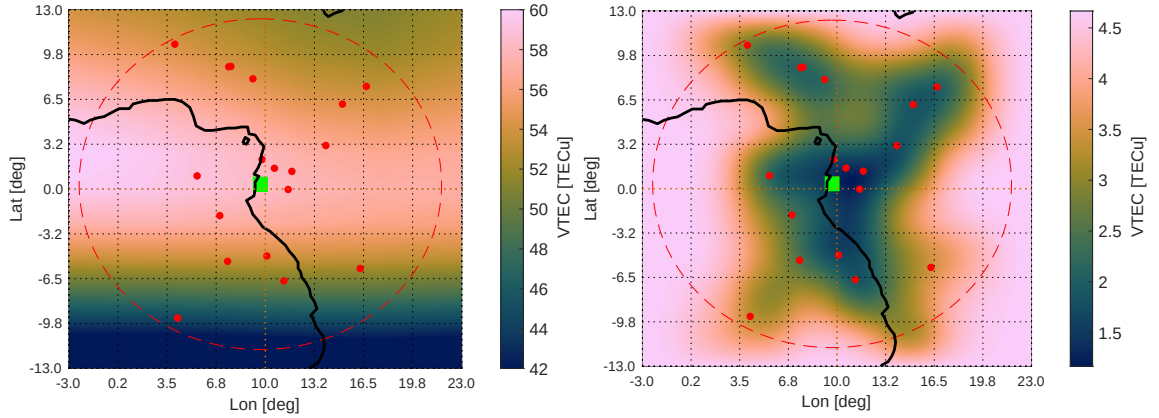
(d) Statistics for the difference between ZTD_{IGS} and ZTD_{est} for two stations in different seasons.

Figure 5.5: Comparison of estimated ZTD and IGS Final ZTD for CORS stations in (a) sub-polar and (b)–(c) mid-latitude regions.



(a) Estimated regional VTEC map for `BILL` station.

(b) Covariance of the estimated map for `BILL` station.



(c) Estimated regional VTEC map for `NKLG` station.

(d) Covariance of the estimated map for `NKLG` station.

Figure 5.6: Estimated regional VTEC maps and their corresponding covariances for two CORS reference stations in tropical and mid-latitude regions, on June 30, 2023.

an idea of the geographic size of the regional area, the black curves are used to show the coastlines of Southern California: (a)–(b), and the west coast of central Africa: (c)–(d). The red dots are the pierce points of used GPS and GAL satellites at the particular epoch. The dashed red ellipse represents the applied elevation mask, i.e. all the used pierce points are within the ellipse.

Since the maps in Figs. 5.6a and 5.6c are estimated around mid-day local time for both stations, we can compare the characteristics of the VTEC maps for the two stations. As

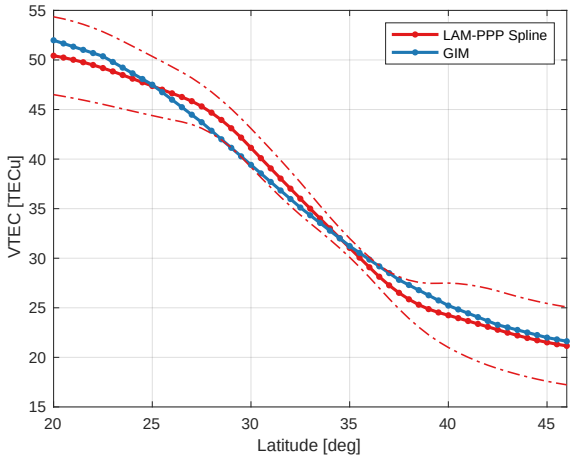
expected, the VTEC magnitude is smaller in the mid-latitude region (a) than in the tropical region (c).

Figs. 5.6b and 5.6d show the covariances of the estimated maps. Both figures show that areas without pierce points have larger covariances (i.e. larger uncertainty), compared to regions with one or more pierce points. Moreover, areas that have a lot of pierce points (e.g. the areas around the receivers) have smaller covariances, owing to the fact that the satellites are collaborating to estimate the corresponding VTEC.

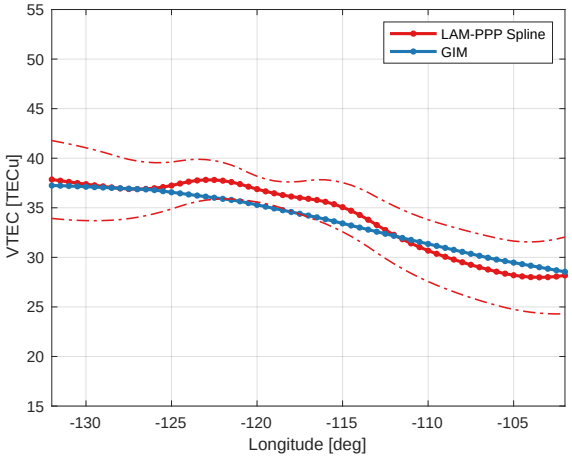
5.5.2.2 Comparison with IGS GIM Corrections

The benefits of Local Atmospheric Modeling (LAM) can be demonstrated by comparing estimates from the LAM-PPP Spline approach with theVTEC values from global GIM corrections. The comparison can be accomplished by: (1) comparing the VTEC values of multiple pierce points along a longitude or latitude line, and (2) comparing slant delays of one pierce point (one satellite) using estimates from multiple receivers in the same local area.

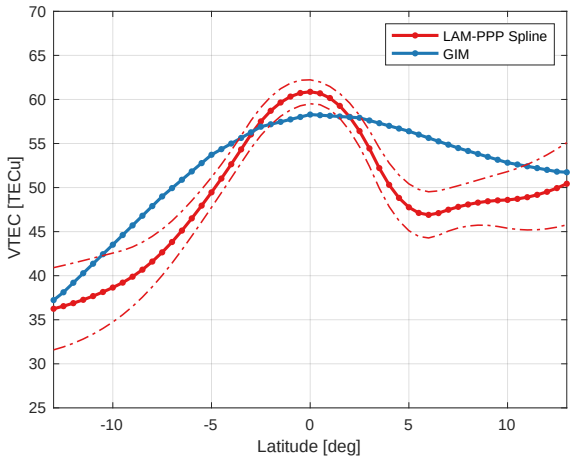
Fig. 5.7 shows comparisons of VTEC values for pierce points along the dotted orange latitude and longitude lines illustrated in Fig. 5.6. For instance, the "latitude line" for the NKLG station corresponds to multiple latitude values of the same 10° meridian, while the "longitude line" corresponds to multiple longitude values of the 0° parallel. The LAM-PPP Spline VTEC estimates are shown with the solid red line, and the envelope of their corresponding standard deviations is illustrated by dash-dotted lines. We observe that standard deviations are larger on the ends of the latitude and longitude *lines* and diminish in the middle. Additionally, the compared VTEC values are closer on the ends of the lines.



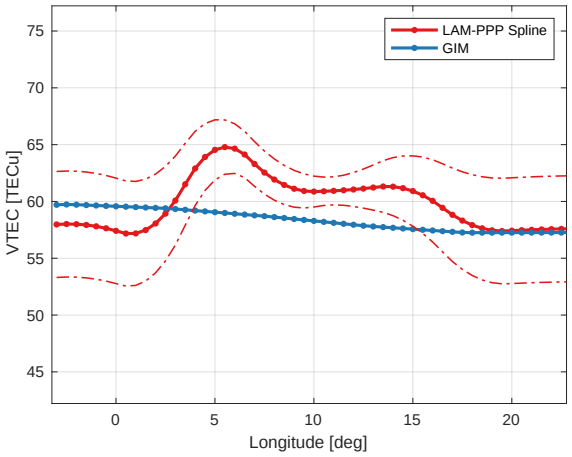
(a) VTEC values along a latitude line for `BILL`



(b) VTEC values along a longitude line for `BILL`



(c) VTEC values along a latitude line for `NKLG`

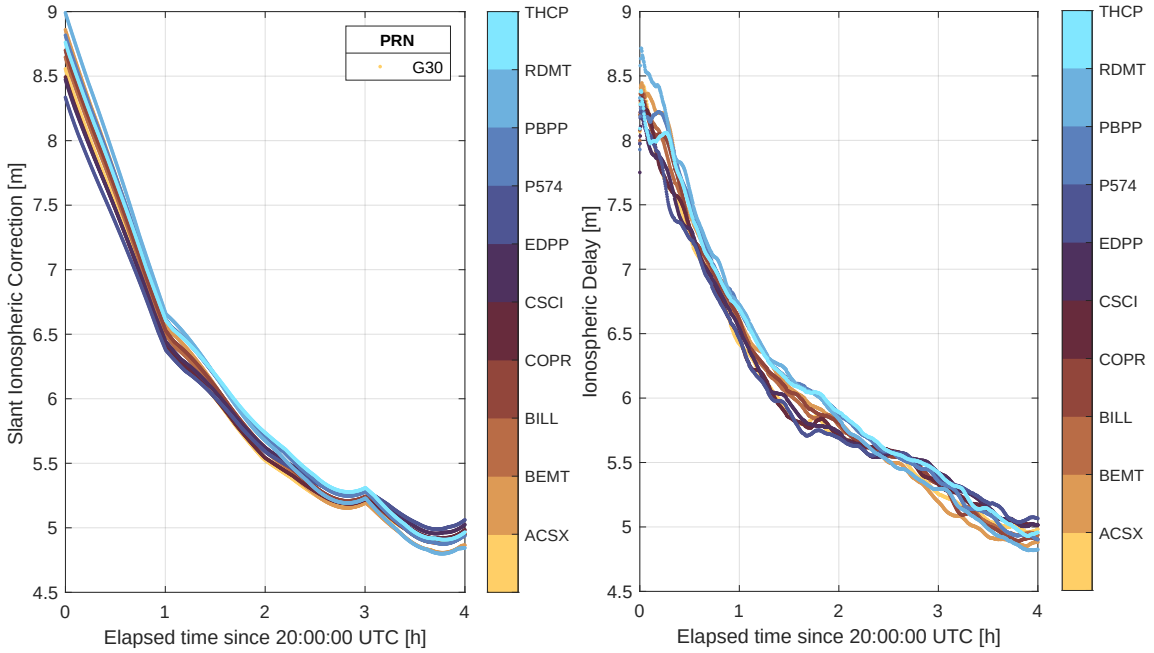


(d) VTEC values along a longitude line for `NKLG`

Figure 5.7: Estimated regional VTEC maps and their corresponding covariances for two CORS reference stations in tropical and mid-latitude regions, on June 30, 2023.

This is due to the fact that the ends of the lines correspond to outer edges of the regional maps, i.e. areas that do not have any satellite pierce points by definition.

Fig. 5.8 uses multiple receivers in the same *local* area to show both the GIM corrections and the LAM-PPP Spline estimates of the GPS L1 slant ionospheric delay for satellite G30. The receivers are from CORS stations separated by no more than 50 km. (See Fig. C.1 of Appendix C for a geographic map showing the locations of the stations.) As expected, the figure shows that the estimated slant delays for all stations are uniform in their variation



(a) GIM correction (b) LAM-PPP Spline estimates

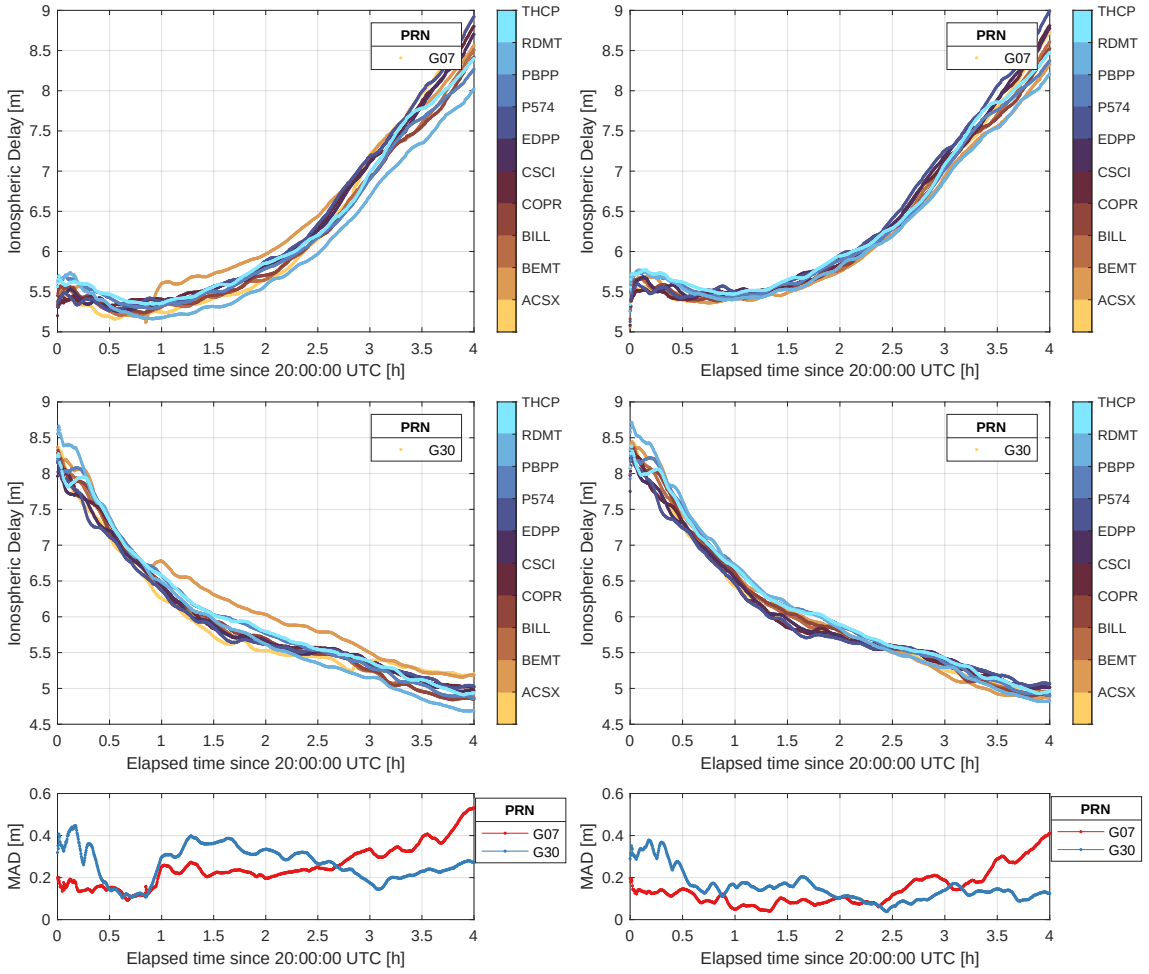
Figure 5.8: The GIM corrections and estimates of slant ionospheric delays for the L1 signal of G30 GPS satellite. Each curve represents the slant delay estimate for one station, for a total of ten stations in the same local area.

with time, i.e. they increase and decrease in a similar fashion. This is due to the fact that the stations are all located in the same local area as defined by Definition 6.1.1. (Recall that this uniform variability was also shown in the modeled corrections in Fig. 3.3.)

The artifacts introduced by spatio-temporal interpolation of GIM corrections appear as sharp changes in VTEC values around the hour marks in Fig. 5.8a. Fig. 5.8b shows that the LAM-PPP Spline approach removes these artifacts, which underlines the importance of estimating local ionospheric delays.

5.5.2.3 Comparison of Slant and Spline Approaches

The stability of estimated ionospheric delays is one of the major benefits of the LAM-PPP Spline approach over conventional LAM approaches. This is a direct consequence



(a) LAM-PPP Slant estimate

(b) LAM-PPP Spline estimate

Figure 5.9: Estimates of slant ionospheric delays for the L1 frequency of two GPS satellites (G07 and G30). The plots in the top 4 panels have 10 curves, each representing the estimate at one receiver, for a total of ten receivers in the same local area. The plots in the bottom 2 panels are the epoch-by-epoch median absolute deviation (MAD) statistics.

of satellite collaboration, i.e. the fact that each satellite's measurements contribute to estimating values of multiple B-spline parameters, which in turn results in the collaboration of multiple satellites in estimating each other's slant delays. This is in contrast to the LAM-PPP Slant approach, for which each satellite's measurements can only contribute to estimating its own slant delay state. Consequently, the LAM-PPP Slant approach does not have the same stability, as demonstrated by Fig. 5.9a in which the LAM-PPP Slant

estimate for BEMT station deviates from that of the remaining stations around the hour mark. This is an example of the instability that could happen with the LAM-PPP Slant approach, but one which is easily handled by the LAM-PPP Spline approach.

The overall convergence performance of the LAM-PPP Spline approach can be compared to that of LAM-PPP Slant approach using multiple individual receivers in the same regional area, all using the same stochastic parameters and filter settings. To that end, consider a set of N slant ionospheric delays for satellite s, $\{I_i^s\}_{i=1}^N$, obtained from N receivers in the same local area. The median absolute deviation (MAD) can be used to robustly quantify the variability of these slant delays at each epoch. The MAD is defined as:

$$\text{MAD} = \text{median}(|I_i^s - \bar{I}^s|), \tag{5.84}$$

where $\bar{I}^s = \sum_{i=1}^{N} I_i^s$ is the mean of the N slant delays for satellite s. The bottom panels of Fig. 5.9 show the epoch-by-epoch MAD of both the LAM-PPP Slant and LAM-PPP Spline approaches for two GPS satellites (G30 and G07). As expected, the MAD for LAM-PPP Spline is consistently smaller than that of LAM-PPP Slant. This consistency of the LAM-PPP Spline approach is a consequence of its ability to estimate one satellite's slant delay using measurements from more than one satellite.

5.5.3 Position Coordinates Estimates

The estimated receiver position coordinates $\hat{\mathbf{p}}_r$ in eq. (5.33) are in the ECEF frame, whose origin is the center of the Earth. To evaluate the error of the estimated receiver position, the estimated ECEF coordinates are transformed into local tangent plane coordinates whose origin is the *true* ECEF coordinates of the receiver. The resulting coordinates

are called East-North-Up/Height (ENU) coordinates[4]. The true ECEF coordinates of the receiver are obtained from various analysis centers.

For a stationary receiver, the ENU coordinates provide and intuitive definition of position errors, in a frame centered at the true position. Denote the receiver's true ECEF coordinates as $\mathbf{p}_{r,0}$, then the position error vector can be transformed from ECEF into local tangent coordinates by:

$$[dE, dN, dU]^\top = \mathbf{R}_e^t \left(\hat{\mathbf{p}}_r - \mathbf{p}_{r,0}\right), \tag{5.85}$$

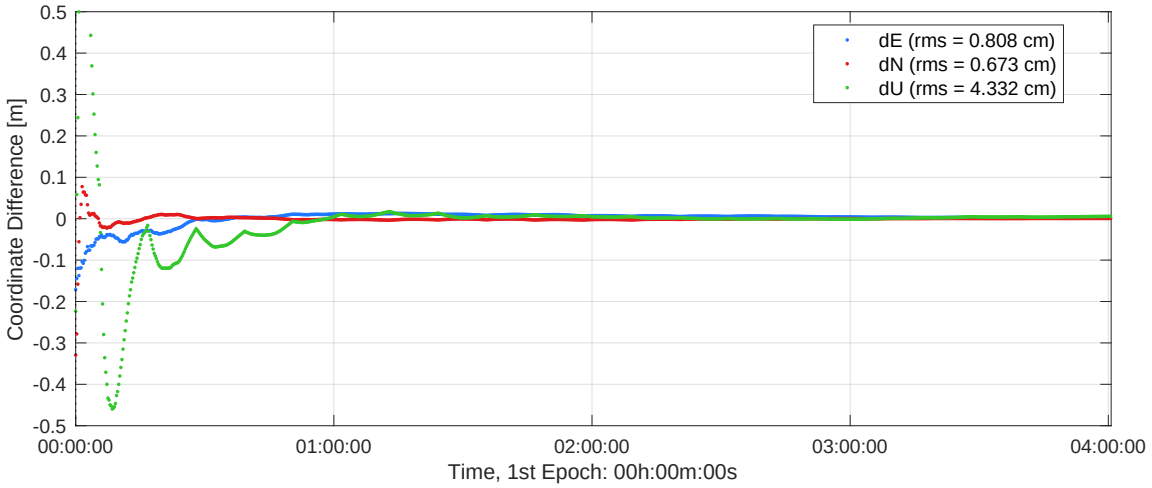
where $\mathbf{R}_e^t$ is a transformation matrix from the ECEF frame e to the local tangent plane t, defined in Section 2.5 of [61]. The vector $[dE, dN]^\top$ is the horizontal error vector, and dU is the vertical error. The 2D (horizontal) error E_{2D} and 3D error E_{3D} are then defined, respectively, as:

$$E_{2D} = \sqrt{(dE)^2 + (dN)^2}, \tag{5.86}$$

$$E_{3D} = \sqrt{(dE)^2 + (dN)^2 + (dU)^2}. \tag{5.87}$$

The convergence performances of LAM-PPP Slant and LAM-PPP Spline approaches are similar for individual receivers using exactly the same measurements. This is primarily due to two reasons: (1) both methods have similar single-epoch redundancy, as shown in Table 4.2, and (2) both methods use similar stochastic modeling and filter settings. Fig. 5.10 shows the convergence of ENU position coordinates error for the CORS station `BILL`, with observation interval of 15 seconds. The true position coordinates were obtained from the daily IGS post-processed solutions that are computed using the most accurate IGS

[4]↑The ENU local tangent plane coordinates are also referred to as local geodetic coordinates. They are formed from a plane tangent to the Earth's surface fixed to a specific location (in this case, the receiver's true geodetic coordinates).



(a) Observation period: 17:00–19:00 PT (Local time)

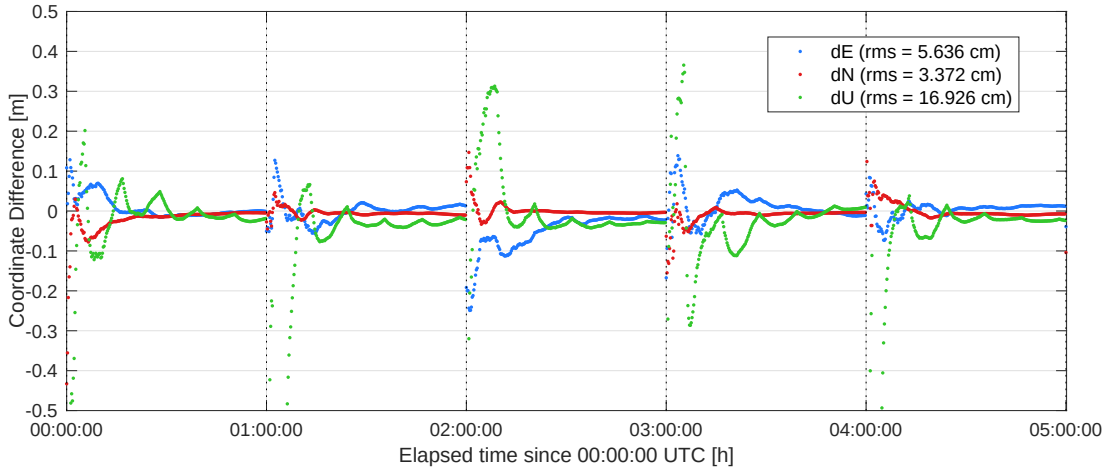
Figure 5.10: Convergence of estimated position coordinates for CORS station `BILL`, on June 30, 2023. The plots correspond to 4-hour observations period.

final products derived from all ACs. For this particular station, the coordinates converge to centimeter-level accuracy within a few minutes. For survey-grade receivers, it generally takes 5–30 minutes for coordinates to convergence to less than 10 cm, when using measurements with a sampling interval of 15 seconds.

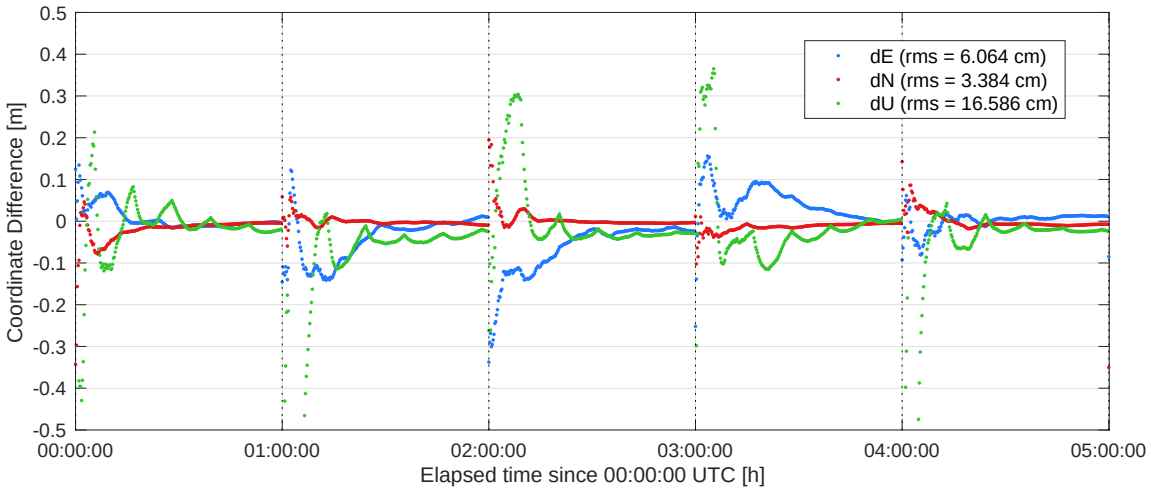
To assess the performance and convergence behavior, many convergence periods are needed in order to gather meaningful statistics. Fig. 5.10 shows that the position coordinate error converges to centimeter-level well within one hour, and therefore the Kalman Filter (KF) can be reset every one hour. The KF reset at epoch k involves re-initializing the state vector and the corresponding covariance, such that at epoch $k+1$:

$$\begin{aligned} \hat{\mathbf{x}}_{k+1}^{-} &= \hat{\mathbf{x}}_{0}^{-}, \\ \mathbf{P}_{k+1}^{-} &= \mathbf{P}_{0}^{-}, \end{aligned} \tag{5.88}$$

where $\hat{\mathbf{x}}_{0}^{-}$ and $\mathbf{P}_{0}^{-}$ are the initial state and covariance. The reset ensures that no information from previous epochs is carried over to epoch $k+1$. The epochs between two KF resets constitute one *convergence period*. Fig. 5.11 shows plots of estimated ENU coordinate



(a) LAM-PPP Spline



(b) LAM-PPP Slant

Figure 5.11: Convergence of estimated position coordinates for CORS station BILL, using both the LAM-PPP Spline and LAM-PPP Slant approaches. The vertical dotted lines indicate the Kalman Filter resets which are performed to assess the convergence performance. The plots correspond to a 5-hour observations period, resulting in 5 convergence periods since the Kalman filter is reset every hour.

error, with KF resets every one hour, for both the LAM-PPP Spline and LAM-PPP Slant approaches. It is clear that for both approaches, the convergence of position coordinates is also reset after each KF re-initialization.

The initial convergence behavior might differ for the two approaches. Since the number of estimated atmospheric delay states is different, it follows that the total number of

unknown states is also different for both methods. In most cases, the number of B-spline VTEC parameters estimated for LAM-PPP Spline is larger than the number of slant delays estimated for LAM-PPP Slant. For example, for a region $G = \{[20°, 45°] \times [-135°, -100°]\}$ defined for a user in the Los Angeles area of Southern California, there are 100 B-spline VTEC parameters to estimate for LAM-PPP Spline, resulting from choosing quadratic B-splines (degree $n = 2$) at resolution levels $J_\varphi = J_\lambda = 3$ for both latitude and longitude. On the other hand, assuming there are on average 10 visible satellites for each GNSS constellation, there are on average 20 slant delays to estimate for a dual-GNSS system of GPS+GAL, and less than 50 slant delays for full multi-GNSS. This implies that in practice, if the estimated regional VTEC map has sufficient resolution to match that of external GIM maps, the LAM-PPP Spline approach will always have more atmospheric delay unknown parameters compared to LAM-PPP Slant. However, the use of GIM pseudo-measurements means that, ultimately, both approaches have similar redundancy and hence similar convergence performance, as clearly illustrated by Fig. 5.11.

The accuracy of the coordinate estimates can also be evaluated by looking at the distributions of the 2D and 3D errors defined in eqs. (5.86) and (5.87). Fig. 5.12 shows the cumulative distributions of 2D and 3D errors, for the Spline and Slant approaches. The two curves are closer for the larger errors, indicating that both approaches have similar initial accuracy, However, once the coordinate estimates have converged, LAM-PPP Spline performs slightly better, perhaps owing to the stability brought by collaboration between satellites.

Table 5.6 shows statistics of three performance indicators. For each indicator, the table provides the mean and median convergence time (in minutes) and the mean and standard

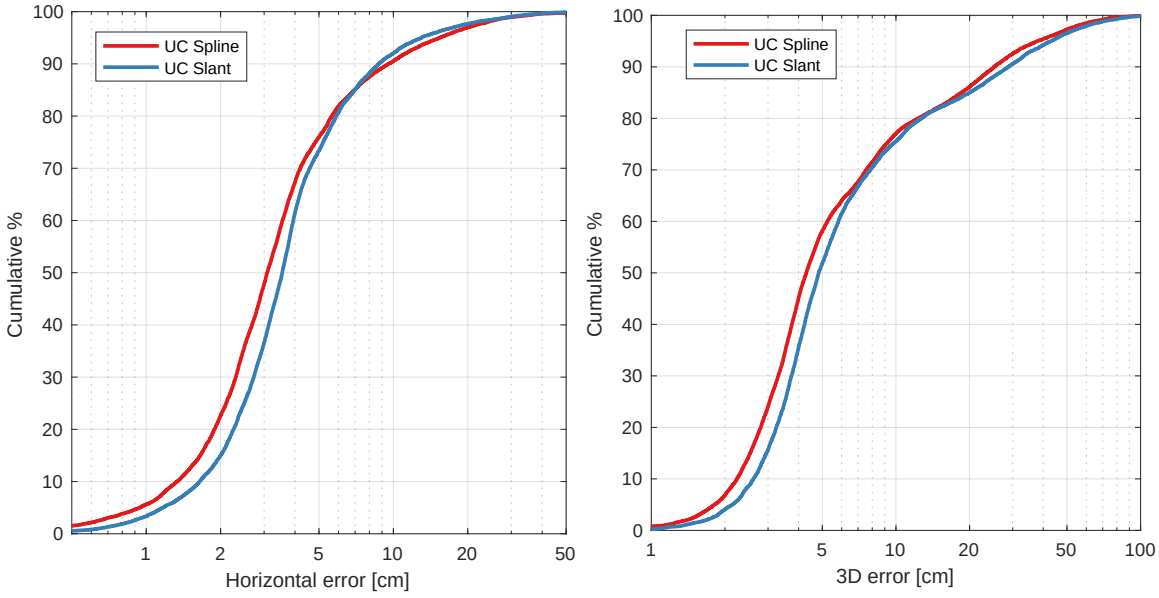


Figure 5.12: Cumulative distributions of the 2D and 3D errors for the LAM-PPP Spline and LAM-PPP Slant approaches. The plots correspond to observations in a 4-hour period: 16:00–20:00 PT (local time).

deviation of the final 2D and 3D accuracies (in centimeters). The statistics summarize the performance of the two LAM-PPP approaches for mid-latitude CORS stations. The statistics are obtained from an experiment that processes 4-hour datasets from 20 mid-latitude CORS stations in the same local area. (See Appendix Table C.1 for details of the used stations.) When processing each station's dataset, the KF is reset every hour to give four convergence periods per station. Therefore, the statistics in Table 5.6 correspond to 80 independent convergence periods. Although the Spline approach outperforms the Slant approach in almost all performance indicators, the margins are small as to be similar.

5.6 Summary

This chapter developed LAM-PPP Spline: a method that enables collaboration between satellites when estimating the local atmospheric delay. It was shown that the use of uncom-

Performance indicator	Metric	Slant	Spline
Convergence time[a] [min]	Mean	6.67	**6.34**
	Median	**5.25**	6.00
Final 2D accuracy[b] [cm]	Mean	2.96	**2.67**
	Stdev	1.05	1.06
Final 3D accuracy[c] [cm]	Mean	3.58	**3.28**
	Stdev	1.23	**1.18**

[a] The time it takes for the 2D position error to fall below 10 cm.
[b] The 2D position error (E_{2D}) after 60 minutes.
[c] The 3D position error (E_{3D}) after 60 minutes.

Table 5.6: Statistical comparison of three performance indicators for the LAM-PPP Spline and LAM-PPP Slant approaches. The statistics are derived by processing datasets from 20 mid-latitude CORS stations, each with 4 convergence periods.

bined measurements enables satellites to collaborate on estimating the local VTEC map, i.e. a measurement from one satellite contributes to the estimation of slant delays for other satellites. For instance, high elevation satellites with precise measurements contribute to the estimation of slant delays of low-elevation satellites. This is in contrast to the LAM-PPP Slant approach in which each satellite's measurements are used to estimate its own slant delay. However, even though the LAM-PPP Spline approach provides precise ionospheric delay estimates that speed up convergence, the ultimate accuracy of position estimates is still heavily dependent on the precision (noise level) of code measurements.

It was demonstrated that one of the crucial benefits of the LAM-PPP Spline approach is the representation of ionospheric delay as a function that can be parametrized by B-spline parameters. This representation gives each individual receiver the capability of generating a regional VTEC map which can easily be shared with other local receivers.

As with other current PPP approaches, the main downside of LAM-PPP Spline is the time it takes to converge to centimeter-level accuracy. This long convergence time is a big disadvantage to applications that require real-time precision. The next chapter aims to

exploit the information sharing capabilities of LAM-PPP to reduce the converge time.

Chapter 6

Collaborative PPP Using a VTEC Model

One of the primary features of Intelligent Transportation Systems (ITS) is the availability of connected agents which can easily share information with each other [96]. These agents, which can be vehicles or other devices in the transportation infrastructure, are often equipped with GNSS receivers and connectivity. Consequently, these vehicles are capable of performing the LAM-PPP approach described in Chapter 5. Recall that in LAM-PPP, each agent individually estimates the common atmospheric delay states (tropospheric and ionospheric delay), in addition to its unique states (position, clock/biases, and ambiguities). As they are connected and equipped with computation and communication capabilities, they could then share information and collaborate on estimating the common states. The primary goal of this chapter is to develop a method that uses a network of agents to reduce PPP convergence time via collaborative estimation of regional atmospheric delays.

6.1 Introduction

There has been previous attempts at collaborative estimation of local and regional ionospheric corrections using networks of receivers. Most of the approaches work on a similar principle as NRTK, described in Section 4.2.3. The most recent attempts use networks composed of reference receivers and *user* receivers. These approaches can be summarized as using some variation on the following three-step approach [7, 8]: (1) reference receivers in the network estimate the slant ionospheric delays of all satellites visible in the region, (2) the ionospheric delays are shared in the network, and finally (3) the user interpolates the slant delays from neighboring reference receivers and uses them as local external corrections instead of using GIM corrections. The interpolation of the slant delays in the last step is often accomplished using inverse-distance weighting strategies, i.e. the slant delays estimates are given less weight the farther away the user is from the reference receiver. The main drawback of such strategies is the unavoidable introduction of interpolation errors in slant delay corrections, which ultimately lead to degraded PPP performance. Furthermore, the previous chapter has shown that there are disadvantages of estimating ionospheric delay as slant delays, and that better performance is achieved by estimating the corresponding VTEC spline parameters. This chapter develops and demonstrates a distributed method for estimating regional atmospheric corrections via collaborative estimation of residual ZWD and VTEC B-spline parameters.

As already mentioned, the current approaches of collaborating on estimating ionospheric slant delays is not optimal, since each agent has a different slant delay value for the same satellite, owing to the fact that the corresponding pierce points are different. It is desirable

to parametrize the atmospheric delays in such a way that their corresponding states are the same for all agents and they can collaborate in estimating them. For ionospheric delay, the B-spline parametrization from Chapter 5 ensures that agents can collaborate on estimating the B-spline parameters. For the tropospheric delay, all agents in a *regional* area have the same residual ZWD which they can collaboratively estimate. The goal then is: (1) for each agent to compute the network average of the atmospheric delay parameters (residual ZWD and B-spline parameters) from all collaborating agents thorough communications with *only* its neighbors; then (2) for each agent to incorporate the computed neighborhood average into its individual filter. This is a distributed estimation problem in which a network of agents running individual filters are also collaborating on estimating dynamic atmospheric delays states (along with other states). It will be shown that this problem involves an agreement on linear combinations of dynamically changing quantities; and that it can be formulated into the well-known *average consensus* problem [97].

Consensus problems have been around in the literature for scenarios when multiple agents are estimating the same quantity. A typical such scenario involves a network of wireless robotic sensors tracking the position of a moving target. This work is the first to attempt formulating collaborative PPP into a consensus problem. To that end, there are two equivalent ways of conceptualizing the collaborative PPP estimation problem:

- The pierce points of each agent are mobile *sensors* estimating a set of time-varying signals (VTEC at the grid knots).
- Each agent computes a regional VTEC map and collaborates with other agents to track the average of the map.

Since the distributed collaborative estimation scheme above relies on neighbors communicating their estimates of time-varying parameters, the uncertainty of these estimates is important to the overall performance. For that reason, this chapter introduces a distributed estimation scheme which weighs each agent's estimate according to its corresponding uncertainty. There are three primary goals for this collaboration:

(1) Accurately estimate the atmospheric delay states.

(2) Enable agents joining the network to instantaneously achieve the network accuracy for atmospheric delay states.

(3) Improve the performance of the remaining unique states, since reducing the uncertainty of atmospheric states reduces uncertainty of the overall state, thereby improving the overall system performance for each agent.

Before proceeding, it is important to define the regional area whose atmospheric delays will be estimated.

Definition 6.1.1. *A local geographic area is defined to have roughly the following dimensions in mid-latitudes:* 300 *km*$\times$500 *km (North* $\times$ *East), or approximately* $2.5^\circ \times 5^\circ$ *(latitude* $\times$ *longitude).*

The above definition should not be taken as strict, but rather as a way to guide the subsequent discussion. The *local* area should also be distinguished from the *regional* area defined in Definition 5.3.1; the latter covers a much larger area. Indeed, the regional area is several times larger the local area, and consequently, receivers in the same local area will roughly have the same corresponding regional area. Fig. 6.1 shows an example of a

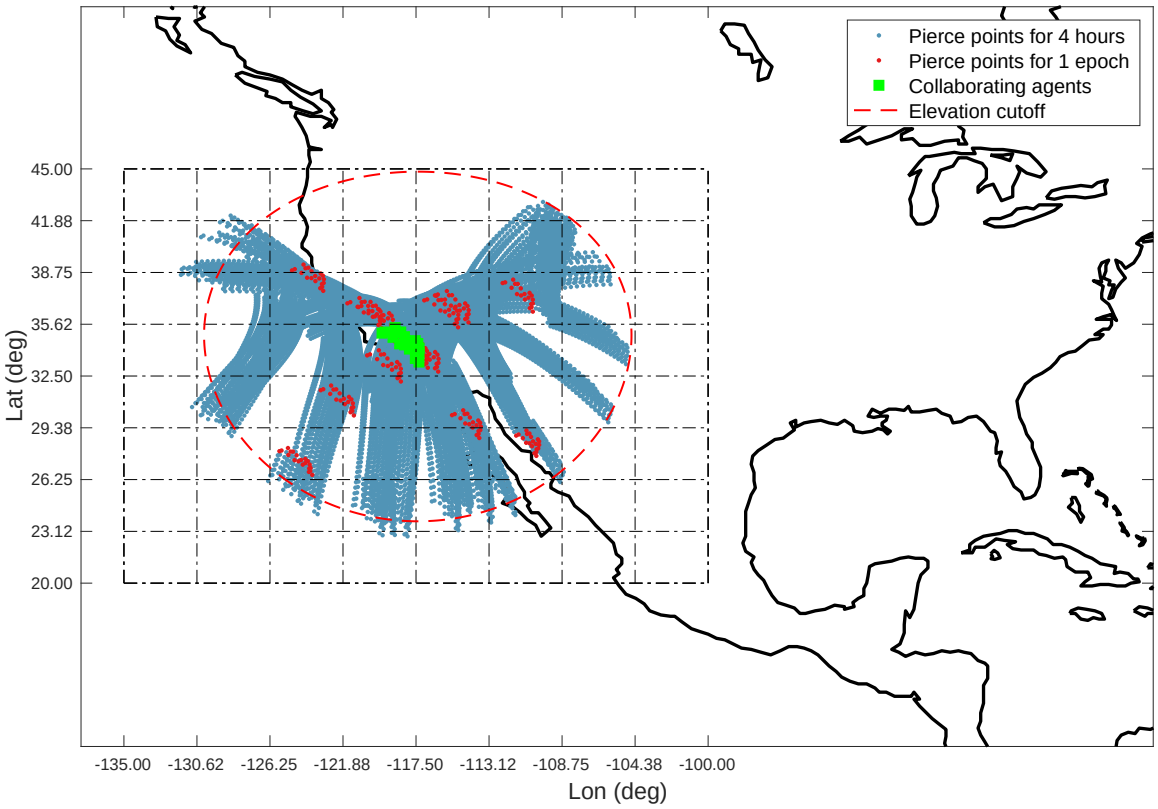


Figure 6.1: A map showing GPS+GAL ionospheric pierce points (in blue and red) from $N = 20$ collaborating agents (in green) in a local area located within a regional area (dashed grid). The solid black lines represent the coast outlines of North America. The red dashed ellipse is the satellite elevation cutoff for one agent.

Southern California local area containing 20 agents. The pierce points corresponding to all agents are shown to remain in an ellipse defined by the elevation cutoff (in red dashed line). The blue pierce points correspond to a period of four hours, whereas the red pierce points correspond to one epoch. The resulting VTEC grid (in black dash-dotted lines) covers the geographic regional area determined using Definition 5.3.1.

6.2 Graph Theory Preliminary

A pair of agents (receivers/users) are engaged in *collaborative estimation* (or simply, collaboration) if they can communicate and use each other's state estimate and information.

It is assumed that a set of collaborating agents forms a network that is represented by a connected graph, i.e., a graph in which there is a collaboration path for any arbitrary pair of agents [98]. The graph can either be directed (there is at least one pair of agents in which only one can send, but not receive information) or undirected (all agents can send and receive information). A graph forms a network when each collaboration between any pair of agents can be assigned a positive real number. A network is also referred to as a *weighted graph* [99], since each edge is assigned a weight. Reasons and implications for graph weighting are important and will be discussed throughout.

Suppose we have a network represented by an undirected connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \ldots, N\}$ represents the N agents and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ represents the agent pairs $(v, w) \in \mathcal{E}$ that are able to share information directly. Agents can also be referred to as nodes of the graph. The neighbors of agent v is a set denoted as $\aleph_v = \{w \in \mathcal{V} : (v, w) \in \mathcal{E}\}$ and its cardinality (the number of neighbors) is denoted by $|\aleph_v|$. (In this work, neighbors are chosen by geographic proximity.) The *adjacency matrix* of $\mathcal{G}$, $\mathbf{A} = \{a_{vw}\}$, is defined as

$$a_{vw} \doteq \begin{cases} 1 & \text{, if } v = w \\ 0 & \text{, if } v \neq w. \end{cases} \tag{6.1}$$

The *degree matrix* is of $\mathcal{G}$ an $N \times N$ matrix defined as $\mathbf{\Delta} = \mathbf{\Delta}(\mathcal{G}) = \{\Delta_{vw}\}$ where:

$$\Delta_{vw} \doteq \begin{cases} |\aleph_v| & \text{, if } v = w \\ 0 & \text{, if } v \neq w. \end{cases} \tag{6.2}$$

The average degree of a graph is defined as:

$$\bar{\Delta}_{\max} = \frac{1}{N} \sum_{v \in \mathcal{V}} |\aleph_v| \tag{6.3}$$

The Laplacian of graph $\mathcal{G}$ is:

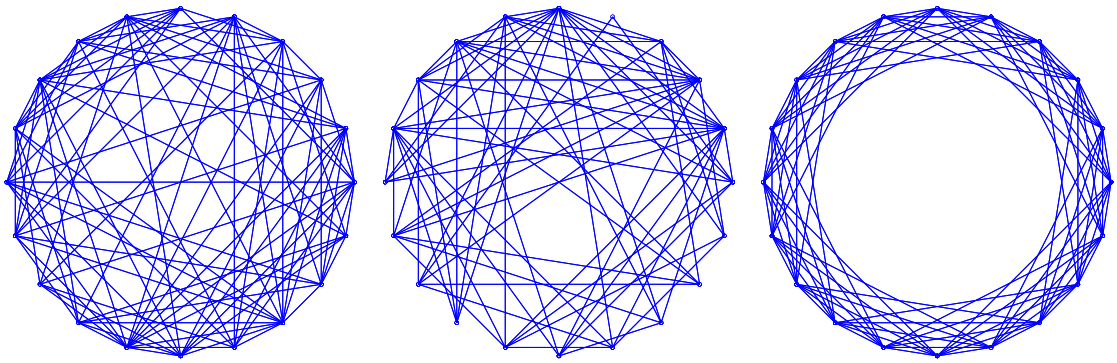
$$\mathbf{L} = \mathbf{\Delta} - \mathbf{A} \tag{6.4}$$

A *path* on the graph is a sequence of nodes $(w_0, w_1, \ldots, w_m)$ such that $w_n \in \mathcal{V}, (w_n, w_{n+1} \in \mathcal{E})$ for all n, and m is the *path length*. An undirected graph is *connected* if there is an undirected path between every pair of distinct nodes. A graph is *strongly connected* if there is a directed path connecting any two arbitrary nodes v, w of the graph. In this work, graphs are assumed to be strongly connected. The *diameter* of $\mathcal{G}$ is defined as the maximum length of the shortest path between any pair of nodes [100–102].

There are three main types of collaboration topologies: ring lattice, scale-free, and small-world. In ring networks, each node is connected to its *immediate* neighbors in a closed loop such that the corresponding graph is connected. A ring lattice network has the added property that each node v is connected to exactly $\aleph_v$ neighbors. Small-world networks have a quasi-random topology consisting of a ring lattice with additional random connections generated by a fixed probability [103, 104]. Scale-free networks have a degree distribution that follows a power law, i.e. few nodes have most of the connections and act as hubs [105]. These three types of networks are illustrated in Fig. 6.2. Table 6.1 shows the statistics of the three topologies, each with 20 agents. It was shown in [103] that small-world networks perform better than the other two types, due to the random connections between non-immediate neighbors. Evaluating the performance of different types of networks is out of scope of this work; hence, the small-world topology will be used exclusively.

Topology	Degree			Connections
	Min.	Max.	Med.	
Ring-lattice	10	10	10	100
Small-world	6	13	10	100
Scale-free	2	16	7.5	79

Table 6.1: Statistics for network topologies with 20 nodes.



(a) Small-world (b) Scale-free (c) Ring-lattice

Figure 6.2: Topologies of the three types of collaboration networks, each with 20 nodes.

6.3 Centralized Estimation

Having set up a collaborating network, this section discusses how the shared information is used to derive a network-wide estimate. The most straightforward approach is to collect estimates from all agents, compute a centralized estimate of the shared states, then send the solution back to all agents. Chapter 4 discussed two different forms of centralized collaborative estimation approaches already used for GNSS positioning. Section 4.2.3 discussed NRTK, in which receivers in a network exchange their measurements in order to eliminate common-mode errors. The exchange involves sending measurements to a central Service Provider (SP) that computes the corresponding corrections and generates virtual measurements, which are then sent back to users to perform the usual RTK. Section 4.3.4 discussed Network PPP and PPP-RTK, in which receivers in a network send their estimated full SSR corrections to a central SP, which computes appropriate corrections and interpolations, then sends them back to users to perform the usual PPP. In both of the above approaches, users send some information to a central node, then receive back improved versions of that information, which they then use to improve their own estimates.

The exchanged information is either full measurements (for NRTK), or full state estimates (for Network PPP).

There are two important remarks:

- Each agent estimates a full state vector that contains states unique to that agent (e.g. its own position, receiver clocks and biases), and states common to all agents (atmospheric delay states). This full state vector is described in eq. (5.33). Agents cannot collaborate on estimating unique states, but they can collaborate on estimating the common states. Therefore, unlike NRTK and traditional Network PPP approaches, it is sensible to share only the portion of the state vector pertinent to the estimation of common atmospheric delays. This means that agents share the atmospheric states $\mathbf{a}$ in eq. (5.33), as well as their corresponding uncertainties.

- Each agent could benefit from the knowledge of atmospheric delay estimates derived from pierce points in different areas of the map. This is especially true for the B-spline approach in Chapter 5, for which one pierce point contributes to estimating the VTEC of areas around it, as shown in Fig. 5.6b.

In the next subsections we formulate a centralized approach, which involves exchanging only the atmospheric delay states and their uncertainties. Since all agents in a network are individually estimating the same VTEC map, the centralized estimate can be interpreted as the weighted average of their estimates, as illustrated by Fig. 6.3. The solution is presented in two different parametrizations; namely the covariance form and information form.

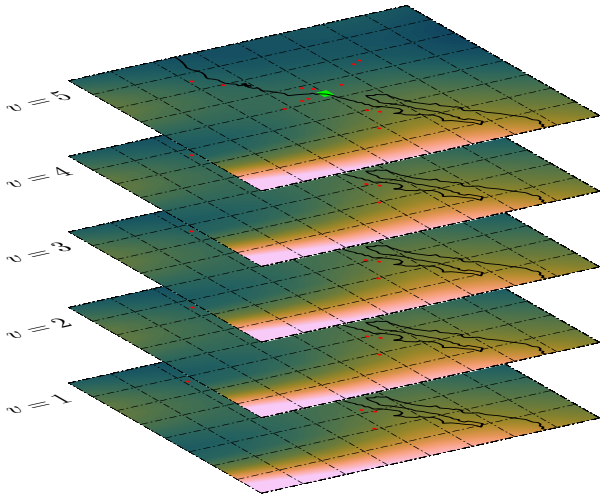


Figure 6.3: A conceptual illustration of collaborative VTEC map estimation using estimates from five agents. The centralized solution can be conceptualized as the weighted average of these maps.

6.3.1 Covariance Form

Recall from Section 5.4.3 that the full state vector of each agent v can be written as:

$$\mathbf{x}_v = [\mathbf{z}_v,\, \mathbf{a}_v]^\top \in \mathbb{R}^{n\times 1}, \tag{6.5}$$

where $\mathbf{z}_v = [\mathbf{p}_r, \mathbf{c}, \mathbf{n}]^\top \in \mathbb{R}^{(n-K-1)\times 1}$ comprises each agent v's unique states, and $\mathbf{a}_v \in \mathbb{R}^{(K+1)\times 1}$ comprises the common atmospheric delay states.

In the centralized framework, at epoch k, each agent v sends an estimate of the atmospheric delay parameters $\mathbf{a}_v$ and a covariance $\mathbf{P}_{\mathbf{a}_v} = \mathbf{P}^+_{\mathbf{a},k}$. The covariance for atmospheric delay states $\mathbf{P}^+_{\mathbf{a},k}$ is is obtained by marginalizing the full posterior covariance $\mathbf{P}^+_k$, as formulated in Table 5.2 or in eqs. (5.60) and (5.63). These estimates $\mathbf{a}_v$ can then be modeled as noisy measurements, whose model is described as:

$$\mathbf{a}_v = \tilde{\mathbf{a}} + \boldsymbol{\rho}_v, \tag{6.6}$$

where $\tilde{\mathbf{a}}$ is the mean and $\boldsymbol{\rho}_v$ is the estimation error. Each agent's estimation error $\boldsymbol{\rho}_v$ is

assumed to have a zero-mean Gaussian distribution, described as:

$$\boldsymbol{\rho}_v \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{\mathbf{a}_v}), \tag{6.7}$$

with $\mathbf{P}_{\mathbf{a}_v}$ as agent v's marginalized covariance. The estimates from multiple agents are independent (because agents are using independent measurements with independent estimators), and identically distributed (because agents have similar assumptions about distributions of their own measurement noise). Given all estimates from N agents, the concatenated vector of estimates from all N agents is written as:

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{H}}\,\mathbf{a} + \boldsymbol{\rho}, \tag{6.8}$$

where:

$$\boldsymbol{\mathcal{A}} = [\mathbf{a}_1, \ldots, \mathbf{a}_N]^\top \in \mathbb{R}^{(K+1)N \times 1}, \tag{6.9}$$

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \mathbb{1} \\ \vdots \\ \mathbb{1} \end{bmatrix} \in \mathbb{R}^{(K+1)N \times (K+1)}, \text{ with } \mathbb{1} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \in \mathbb{R}^{(K+1)\times(K+1)}, \tag{6.10}$$

$$\boldsymbol{\rho} = [\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_N] \in \mathbb{R}^{(K+1)N \times 1}. \tag{6.11}$$

and $\boldsymbol{\rho} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{\mathcal{A}})$ with $\mathbf{P}_{\mathcal{A}}$ a block diagonal matrix of $\mathbf{P}_{\mathbf{a}_v}$ covariance matrices from N agents:

$$\mathbf{P}_{\mathcal{A}} = \begin{bmatrix} \mathbf{P}_{\mathbf{a}_1} & & \\ & \ddots & \\ & & \mathbf{P}_{\mathbf{a}_N} \end{bmatrix} \in \mathbb{R}^{(K+1)N \times (K+1)N}. \tag{6.12}$$

The maximum-a-posteriori (MAP) solution to the problem in (6.8) is:

$$\begin{aligned}\hat{\mathbf{a}} &= \left(\boldsymbol{\mathcal{H}}^{\top}(\mathbf{P}_{\mathcal{A}})^{-1}\boldsymbol{\mathcal{H}}\right)^{-1}\left(\boldsymbol{\mathcal{H}}^{\top}(\mathbf{P}_{\mathcal{A}})^{-1}\,\boldsymbol{\mathcal{A}}\right) \\ &= \left(\sum_{v=1}^{N}(\mathbf{P}_{\mathbf{a}_v})^{-1}\right)^{-1}\left(\sum_{v=1}^{N}(\mathbf{P}_{\mathbf{a}_v})^{-1}\mathbf{a}_v\right).\end{aligned} \tag{6.13}$$

This is the optimal solution for the specific network consisting of N agents. If the estimates from one or more agents cannot be used by the central estimator, for one reason or another, then the solution becomes sub-optimal.

6.3.2 Information Form

In the previous section, the Gaussian distribution of each agent's estimate $\mathbf{a}_v$ is parametrized in terms of its mean $\tilde{\mathbf{a}}$ and covariance $\mathbf{P}_{\mathbf{a}_v}$, i.e. $\mathcal{N}(\mathbf{a}_v;\,\tilde{\mathbf{a}},\mathbf{P}_{\mathbf{a}_v})$. The probability density function of the above distribution can be written as:

$$p(\mathbf{a}_v) = \mathcal{N}(\mathbf{a}_v;\,\tilde{\mathbf{a}},\mathbf{P}_{\mathbf{a}_v}) \tag{6.14}$$

$$\propto \exp\left\{-\frac{1}{2}\left(\mathbf{a}_v-\tilde{\mathbf{a}}\right)^{\top}\mathbf{P}_{\mathbf{a}_v}^{-1}\left(\mathbf{a}_v-\tilde{\mathbf{a}}\right)\right\} \tag{6.15}$$

$$= \exp\left\{-\frac{1}{2}\left(\mathbf{a}_v^{\top}\mathbf{P}_{\mathbf{a}_v}^{-1}\mathbf{a}_v - 2\,\tilde{\mathbf{a}}^{\top}\mathbf{P}_{\mathbf{a}_v}^{-1}\mathbf{a}_v + \tilde{\mathbf{a}}^{\top}\mathbf{P}_{\mathbf{a}_v}^{-1}\tilde{\mathbf{a}}\right)\right\} \tag{6.16}$$

$$\propto \exp\left\{-\frac{1}{2}\,\mathbf{a}_v^{\top}\mathbf{P}_{\mathbf{a}_v}^{-1}\mathbf{a}_v + \tilde{\mathbf{a}}^{\top}\mathbf{P}_{\mathbf{a}_v}^{-1}\mathbf{a}_v\right\}, \tag{6.17}$$

where the Gaussian distribution's normalization constant is accounted for by the proportionality sign $\propto$. The covariance parametrization in eq. (6.14) involves a quadratic form, as shown by eq. (6.15). The quadratic form is expanded in eq. (6.16) to give the equivalent parametrization in eq. (6.17).

The Gaussian distribution can also be parameterized in terms of the information matrix $\mathbf{J}_{\mathbf{a}_v}$ and information vector $\mathbf{j}_{\mathbf{a}_v}$. However, unlike the marginalized covariance $\mathbf{P}_{\mathbf{a}_v}$ which can

be easily extracted from the full covariance, the marginalized information submatrix $\mathbf{J}_{\mathbf{a}_v}$ has to be carefully computed from the information matrix of the full state $\mathbf{J}_v = \mathbf{P}_v^{-1}$. The individual full-state information matrix for agent v is obtained from the posterior covariance matrix from eq. (5.63). It can be expressed in terms of submatrices corresponding to the two sub-states $\mathbf{z}_v, \mathbf{a}_v$:

$$\mathbf{J}_v = (\mathbf{P}_v)^{-1} = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{B}^\top & \mathbf{C} \end{array}\right], \tag{6.18}$$

where $\mathbf{A} \in \mathbb{R}^{(n-K-1)\times(n-K-1)}$, $\mathbf{B} \in \mathbb{R}^{(n-K-1)\times(K+1)}$, and $\mathbf{C} \in \mathbb{R}^{(K+1)\times(K+1)}$. (For brevity, the subscript v has been omitted in the submatrices.) The corresponding full-state information vector is defined as:

$$\mathbf{j}_v \triangleq \mathbf{J}_v \mathbf{x}_v = \left[\begin{array}{c} \mathbf{j}_{\mathbf{z}_v,\mathbf{a}_v} \\ \hline \mathbf{j}_{\mathbf{a}_v,\mathbf{z}_v} \end{array}\right] = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{z}_v \\ \mathbf{a}_v \end{bmatrix} \tag{6.19}$$

$$= \left[\begin{array}{c} \mathbf{A}\,\mathbf{z}_v + \mathbf{B}\,\mathbf{a}_v \\ \hline \mathbf{B}^\top \mathbf{z}_v + \mathbf{C}\,\mathbf{a}_v \end{array}\right]. \tag{6.20}$$

The joint distribution of each agent's unique states and the common atmospheric states can be written in both covariance and information form, respectively, as:

$$\begin{aligned} p(\mathbf{z}_v, \mathbf{a}_v) &= \mathcal{N}\left(\begin{bmatrix} \mathbf{z} \\ \mathbf{a} \end{bmatrix}, \begin{bmatrix} \mathbf{P}_{\mathbf{z}_v} & \mathbf{P}_{\mathbf{z}_v,\mathbf{a}_v} \\ \mathbf{P}_{\mathbf{a}_v,\mathbf{z}_v} & \mathbf{P}_{\mathbf{a}_v} \end{bmatrix}\right) \\ &= \mathcal{N}^{-1}\left(\begin{bmatrix} \mathbf{j}_{\mathbf{z}_v,\mathbf{a}_v} \\ \mathbf{j}_{\mathbf{a}_v,\mathbf{z}_v} \end{bmatrix}, \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix}\right). \end{aligned} \tag{6.21}$$

(The matrix $\mathbf{P}_{\mathbf{z}_v,\mathbf{a}_v} = (\mathbf{P}_{\mathbf{a}_v,\mathbf{z}_v})^\top$ is the cross-covariance matrix of $\mathbf{z}_v$ and $\mathbf{a}_v$.) The Schur complement is used to obtain the marginal information submatrix[1] corresponding only to subvector $\mathbf{a}_v$ [106]:

$$\mathbf{J}_{\mathbf{a}_v} = \mathbf{C} - \mathbf{B}^\top(\mathbf{A})^{-1}\mathbf{B}, \tag{6.22}$$

with $\mathbf{J}_{\mathbf{a}_v} \in \mathbb{R}^{(K+1)\times(K+1)}$. Similarly,the marginal information information vector is:

$$\begin{aligned} \mathbf{j}_{\mathbf{a}_v} &= \mathbf{j}_{\mathbf{a}_v,\mathbf{z}_v} - \mathbf{B}^\top(\mathbf{A})^{-1}\mathbf{j}_{\mathbf{z}_v,\mathbf{a}_v} \\ &= \mathbf{J}_{\mathbf{a}_v}\mathbf{a}_v. \end{aligned} \tag{6.23}$$

(The derivation of the information vector expression in eq. (6.23) is given in Appendix A.) Then the parametrization in eq. (6.17) can be written as:

$$\begin{aligned} p(\mathbf{a}_v) &\propto \exp\left\{-\frac{1}{2}\,\mathbf{a}_v^\top \mathbf{J}_{\mathbf{a}_v}\mathbf{a}_v + \mathbf{a}_v^\top \mathbf{j}_{\mathbf{a}_v}\right\} \\ &\propto \mathcal{N}^{-1}\left(\mathbf{a}_v; \mathbf{j}_{\mathbf{a}_v}, \mathbf{J}_{\mathbf{a}_v}\right). \end{aligned} \tag{6.24}$$

The parametrization in eq. (6.24) is referred to as the information form parametrization of the Gaussian distribution, in contrast to the covariance form in eq. (6.17).

The information matrix and information vector allow the formulation of the so-called "natural" parametrization of the Gaussian distribution, in which the error $\boldsymbol{\rho}$ in (6.8) is modeled as $\boldsymbol{\rho} \sim \mathcal{N}^{-1}(\mathbf{0}, \mathbf{J}_\mathbf{a})$, where $\mathbf{J}_\mathbf{a} = \left(\mathbf{P}_\mathbf{a}\right)^{-1}$. Additionally, the MAP solution in (6.13) can then be written in terms of the expressions in (6.23) and (6.22):

$$\hat{\mathbf{a}} = \left(\sum_{v=1}^{N} \mathbf{J}_{\mathbf{a}_v}\right)^{-1} \left(\sum_{v=1}^{N} \mathbf{j}_{\mathbf{a}_v}\right). \tag{6.25}$$

In other words, the MAP solution to the problem in (6.8) is the information-weighted average of estimates from all agents.

[1]↑In the unlikely case when $\mathbf{a}_v$ is a scalar, then both the information "matrix" and "vector" quantities are themselves scalars.

The equivalent centralized solutions in eqs. (6.13) and (6.25) require sending data to a central node. Therefore, these solutions are collaborative, but not fully distributed. There are two main disadvantages of relying on a central node for collaboration: (1) it could fail and stop collaboration for the entire network, and, (2) the computation complexity directly depends on the number of agents in the network, i.e. the addition of more agents to the network increases the overall requirements for communication and computation at the central node [100]. Furthermore, communications happen exclusively between the central node and each agent, and the centralized solutions do not take advantage of the fact that information can be easily broadcast from one agent to its neighbors, as is the case in ITS. The next two sections discuss two approaches of computing the centralized estimate in a distributed framework.

6.4 Average Consensus

A distributed alternative to the centralized approach uses the *average consensus algorithm* [97, 107]. In the average consensus approach, the estimates $\mathbf{a}_v$ from each agents are shared between neighbors, and in turn each agent keeps an *average* of its neighbors estimates.

At epoch k, an agent v with neighbors $w \in \aleph_v$ can implement the distributed average consensus algorithm using:

$$\mathbf{a}_v^{(k+1)} = \mathbf{a}_v^{(k)} + \epsilon \sum_{w \in \aleph_v} a_{vw} \big(\mathbf{a}_w^{(k)} - \mathbf{a}_v^{(k)} \big), \tag{6.26}$$

where $0 < \epsilon < 1/\Delta_{\max}$ and $\Delta_{\max}$ is the maximum degree of the network. The indices $k = \{1, \ldots, \kappa\}$ are the consensus iterations, with κ the maximum number of iterations.

Algorithm 1 Performing Average Consensus on atmospheric states at node v.

Input: Atmospheric states estimate, $\hat{\mathbf{a}}_v$; consensus speed, ϵ; threshold, τ; adjacency matrix, $\mathbf{A} = \{a_{vw}\}$; and maximum consensus iterations, κ.

Output: Consensus estimate $\hat{\mathbf{a}}_v$.

1: Initialize consensus proposal:
$$\mathbf{a}_v^{(0)} = \hat{\mathbf{a}}_v.$$
2: **while** $\|\hat{\mathbf{a}}_v^{(k+1)} - \hat{\mathbf{a}}_v^{(k)}\| > \tau$ and $k < \kappa$ **do**
3: Send $\mathbf{a}_v^{(k)}$ to all neighbors $w \in \aleph_v$;
4: Receive $\mathbf{a}_w^{(k)}$ from all neighbors $w \in \aleph_v$;
5: Perform consensus using :
$$\mathbf{a}_v^{(k+1)} = \mathbf{a}_v^{(k)} + \epsilon \sum_{w \in \aleph_v} a_{vw}\big(\mathbf{a}_w^{(k)} - \mathbf{a}_v^{(k)}\big),$$
6: $k = k + 1$
7: **end while**

The initial conditions $\mathbf{a}_v^{(0)} = \mathbf{a}_v$ and $\mathbf{a}_w^{(0)} = \mathbf{a}_w$ are the posterior estimates of each agent's individual filter. The average consensus algorithm is illustrated in Algorithm 1.

The main property of the average consensus algorithm is that in a strongly connected network, the consensus estimate converges asymptotically to the average of the initial conditions of the entire network, i.e. the average of all agents' estimates.

Other properties of the average consensus algorithm have been thoroughly discussed in the literature (e.g. [97, 101, 108]).

A major drawback of average consensus is the implied assumption that all agents' estimates are equally accurate. This assumption means that each agent's uncertainty, quantified by the covariance (or information matrix), is not taken into consideration when computing the consensus estimate. When there is uncertainty mismatch, the consensus estimate degrades as a result of mixing high-quality and low-quality estimates. Consequently, although the algorithm guarantees convergence to the initial conditions, the consensus estimate might not converge to the optimal centralized estimate found in eq. (6.13). Therefore, we need a

distributed algorithm that also takes into account the estimate uncertainty of each collaborating agent.

6.5 Information-weighted Consensus

The information-form centralized estimate in eq. (6.25) can be computed in a distributed manner by having agents communicate their information vectors and matrices to their neighbors. Define a matrix $\mathbf{D}$:

$$\mathbf{D} = \frac{1}{N}\left(\sum_{v=1}^{N} \mathbf{J}_{\mathbf{a}_v}\right) = \frac{1}{N}\sum_{v=1}^{N} \mathbf{D}_v, \tag{6.27}$$

where $\mathbf{D}_v = \mathbf{J}_{\mathbf{a}_v}$, and $\mathbf{D}, \mathbf{D}_v \in \mathbb{R}^{(K+1)\times(K+1)}$. The matrix $\mathbf{D}_v$ is each agent's information matrix which is communicated to the central estimator for MAP estimation in eq. (6.25). The matrix $\mathbf{D}$ is the average of information matrices across the entire network. Similarly, define a vector $\mathbf{d}$:

$$\mathbf{d} = \frac{1}{N}\left(\sum_{v=1}^{N} \mathbf{j}_{\mathbf{a}_v}\right) = \frac{1}{N}\sum_{v=1}^{N} \mathbf{d}_v, \tag{6.28}$$

where $\mathbf{d}_v = \mathbf{j}_{\mathbf{a}_v}$, and $\mathbf{d}, \mathbf{d}_v \in \mathbb{R}^{(K+1)\times 1}$. The vector $\mathbf{d}_v$ is each agent's information vector, whereas the vector $\mathbf{d}$ averages information vectors across the entire network. Then eq. (6.25) can be written as:

$$\hat{\mathbf{a}} = \mathbf{D}^{-1}\mathbf{d}. \tag{6.29}$$

Note that the dimension of $\mathbf{D}$ does not depend on the number of nodes N, i.e. as the dimension of the problem grows (more agents added to the network) the requirements for storing $\mathbf{D}$ and $\mathbf{D}_v$ do not change. The same applies to $\mathbf{d}$ and $\mathbf{d}_v$.

In this approach, each node maintains estimates of both $\mathbf{D}_v$ and $\mathbf{d}_v$. Each node can exchange information only with its neighbors, and solve for both $\mathbf{D}$ and $\mathbf{d}$ by iterating the

discrete-time difference equations:

$$\mathbf{D}_v^{(k+1)} = \mathbf{D}_v^{(k)} + \epsilon \sum_{w \in \aleph_v} a_{vw} \left(\mathbf{D}_w^{(k)} - \mathbf{D}_v^{(k)} \right), \tag{6.30}$$

$$\mathbf{d}_v^{(k+1)} = \mathbf{d}_v^{(k)} + \epsilon \sum_{w \in \aleph_v} a_{vw} \left(\mathbf{d}_w^{(k)} - \mathbf{d}_v^{(k)} \right), \tag{6.31}$$

where $0 < \epsilon < 1/\Delta_{\max}$, $k = \{1, \ldots, \kappa\}$ is the consensus iteration, and $\mathbf{D}_v^{(0)} = \mathbf{D}_v$ and $\mathbf{d}_v^{(0)} = \mathbf{d}_v$, respectively. After consensus has been reached at iteration $k = k_c$, the atmospheric delay estimate and its corresponding covariance are computed at each agent by:

$$\hat{\mathbf{a}}_v = \left(\mathbf{D}_v^{(k_c)} \right)^{-1} \mathbf{d}_v^{(k_c)}, \tag{6.32}$$

$$\mathbf{P}_{\mathbf{a}_v} = \left(N \, \mathbf{D}_v^{(k_c)} \right)^{-1}. \tag{6.33}$$

The information-weighted consensus algorithm is illustrated in Algorithm 2.

For analysis, eqs. (6.30) and (6.31) can also be written as:

$$\mathbf{D}^{(k+1)} = \mathbf{P} \, \mathbf{D}^{(k)}, \tag{6.34}$$

$$\mathbf{d}^{(k+1)} = \mathbf{P} \, \mathbf{d}^{(k)}, \tag{6.35}$$

with $\mathbf{P} = \mathbf{I} - \epsilon \, \mathbf{L}$, where $\mathbf{I} \in \mathbb{R}^{N \times N}$ is the identity matrix, and $\mathbf{L} \in \mathbb{R}^{N \times N}$ is the Laplacian matrix defined in eqn. (6.4). (Similarly, eq. (6.31) can be written the same way.) Similar to the algorithm represented by eq. (6.26), the consensus to the average of initial values is asymptotically reached. Since the initial values in eqs. (6.30) and (6.31) are the information vectors and matrices of all agents in the network, their average is the MAP solution in eq. (6.29). Hence, the distributed information-weighted consensus algorithm converges to optimal centralized estimate, i.e. $\hat{\mathbf{a}}_v = \hat{\mathbf{a}}$. After the network has converged, by only exchanging information with neighbors, each agent has access to the optimal estimate $\hat{\mathbf{a}}$.

Algorithm 2 Performing Information-Weighted Distributed Consensus on atmospheric states at node v.

Input: Atmospheric states estimate, $\hat{\mathbf{a}}_v$; marginalized information matrix $\mathbf{J}_{\mathbf{a}_v}$; consensus speed, ϵ; threshold, τ; adjacency matrix, $\mathbf{A} = \{a_{vw}\}$; and maximum consensus iterations, κ.

Output: Consensus estimate $\hat{\mathbf{a}}_v$, and corresponding information matrix $\mathbf{J}_{\mathbf{a}_v}$.

1: Initialize consensus proposals:

$$\mathbf{D}_v^{(0)} = \mathbf{J}_{\mathbf{a}_v},$$

$$\mathbf{d}_v^{(0)} = \mathbf{J}_{\mathbf{a}_v}\hat{\mathbf{a}}_v.$$

2: **while** $\|\hat{\mathbf{a}}_v^{(k+1)} - \hat{\mathbf{a}}_v^{(k)}\| > \tau$ and $k < \kappa$ **do**

3: Send $\mathbf{D}_v^{(k)}$ and $\mathbf{d}_v^{(k)}$ to all neighbors $w \in \aleph_v$;

4: Receive $\mathbf{D}_w^{(k)}$ and $\mathbf{d}_w^{(k)}$ from all neighbors $w \in \aleph_v$;

5: Perform consensus using eqs. (6.30) and (6.31):

$$\mathbf{D}_v^{(k+1)} = \mathbf{D}_v^{(k)} + \epsilon \sum_{w \in \aleph_v} a_{vw}\left(\mathbf{D}_w^{(k)} - \mathbf{D}_v^{(k)}\right),$$

$$\mathbf{d}_v^{(k+1)} = \mathbf{d}_v^{(k)} + \epsilon \sum_{w \in \aleph_v} a_{vw}\left(\mathbf{d}_w^{(k)} - \mathbf{d}_v^{(k)}\right)$$

6: Compute the distributed estimate $\hat{\mathbf{a}}_v$ by:

$$\hat{\mathbf{a}}_v = \left(\mathbf{D}_v^{(k+1)}\right)^{-1} \mathbf{d}_v^{(k+1)}$$

7: $k = k + 1$

8: **end while**

6.6 Experimental Results

This section evaluates the performance of a collaborating network of agents. The network is set up using agents in the same local area, as defined by Definition 6.1.1. Fig. C.1 in Appendix C shows a geographic map of $N = 20$ agents collaborating in the Southern California region. Fig. 6.1 shows the corresponding regional area whose VTEC is estimated.

In Chapter 5, the approximate position of each individual receiver v and the elevation mask was used to compute regional area G_v. The ionospheric B-spline parameters would then be only applicable to pierce points inside G_v. Since collaborating agents have to share their estimated B-spline parameters, it is important that all their pierce points are applicable to the collaborative regional area. Hence, all collaborating agents use the same

collaborative regional area G_{collab}, which is a superset of their individual regional areas G_v:

$$G_{collab} \supseteq G_v, \text{ for } v = 1, \ldots, N, \tag{6.36}$$

where N is the number of agents.

The processing details for each agent are the same as in Table 5.5, with the addition of the collaboration parameters discussed in the following subsections.

Some errors in this dissertation are assumed to have a Gaussian distribution. For such errors, the 68% and 95% percentiles fall within one and two standard deviations from the mean, respectively. Specifically, the 68% percentile gives an idea of the typical or average performance of the PPP approaches, while the 95% percentile provides insights into the worst-case performance, helping to assess the reliability and robustness of each approach. Hence, these percentiles are used to provide a convenient way of summarizing the spread and concentration of the estimated errors.

6.6.1 Filter Re-initialization

To assess collaborative performance and convergence behavior, the Kalman Filter (KF) is re-initialized every one hour. As shown in Fig. 5.10a, the position typically converges to cm-level within one hour. Unlike the complete KF reset described in eq. (5.88) of the previous chapter, collaborating agents have access to the converged network estimate of atmospheric delays and their corresponding covariances. Hence, after the filter reset at epoch k, each agents re-initializes its atmospheric delay states and corresponding covariance to the converged network averages $\hat{\mathbf{a}}$ and $\mathbf{P}_{\hat{\mathbf{a}}}$, respectively, such that the full state vector

and covariance at epoch $k+1$ is:

$$\begin{aligned} \hat{\mathbf{x}}_{k+1}^{-} &= [\mathbf{z}_0,\, \hat{\mathbf{a}}]^\top, \\ \mathbf{P}_{k+1}^{-} &= \begin{bmatrix} \mathbf{P}_{\mathbf{z}_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\hat{\mathbf{a}}} \end{bmatrix}, \end{aligned} \tag{6.37}$$

where $\mathbf{z}_0$ and $\mathbf{P}_{\mathbf{z}_0}$ are the initial values of each agent's unique states and covariance, respectively. The re-initialization in eq. (6.37) describes the case of an agent joining a converged network. In practice, this might correspond to a scenario of a vehicle turning on its GNSS receiver for the first time. For this reason, filter re-initialization is used to simulate agents joining a collaborating network. In the ensuing discussion, collaboration performance is evaluated for convergence periods *after* the first KF re-initialization.

The ability of agents to join the network and instantly get a very precise estimate is one of the main advantages of the collaborative estimation approach described in this chapter. To simulate the agents joining the network from a cold start, each agent's individual KF is reset and at regular intervals. To simulate joining a converged network, each vehicle's KF is re-initialized with the distributed estimate (mean and covariance) of atmospheric delay states. In this way, agents receive a precise estimate from the network instantaneously after re-initializing their filters, i.e after joining the network.

6.6.2 Effect of Collaboration on Tropospheric Delay Estimates

Fig. 5.5 in Section 5.5.1 showed that the LAM-PPP Spline approach is able to estimate the ZTD with millimeter-level accuracy, as reported in Table 5.5d. However, the figure did not readily illustrate that the ZTD estimate takes several minutes to converge to its eventual accuracy. Fig. 6.4 shows an example of residual ZWD estimates from seven collaborating

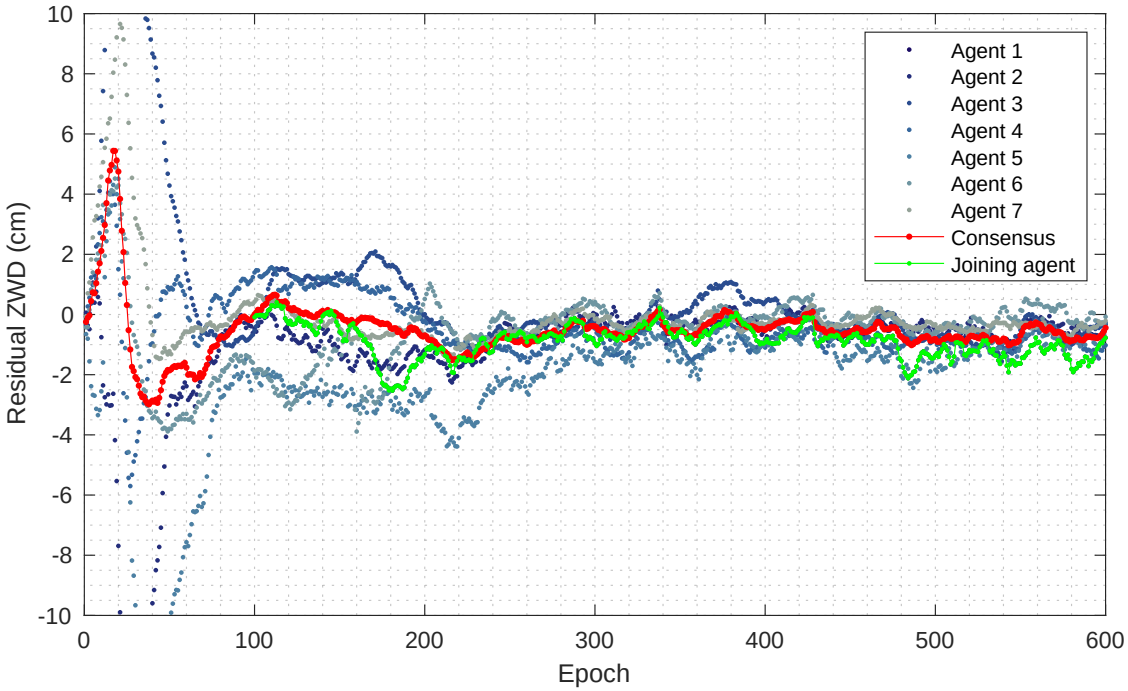


Figure 6.4: An example of residual ZWD estimates from a subset of collaborating agents. The blue lines are the KF estimates from each individual agent. The red curve is the resulting consensus estimate of residual ZWD . The green depicts an agent that joins the network and instantly obtains the current network estimate of the ZWD.

agents (in shades of blue), as well as with the consensus estimate (in red). Recall that the estimated residual ZWD indicates the accuracy of the ZTD, i.e., the smaller the residual ZWD, the more accurate the ZTD. To achieve a residual delay of approximately 1 cm, the estimates from all seven agents require roughly 80 epochs, equivalent to 15 minutes given the 15-second measurement interval. The red curve shows that the consensus estimate robustly tracks the *average* of collaborating agents, such that a new agent joining the network (in green) at epoch 100 obtains the consensus estimate instantly.

Fig. 6.5 shows the ZTD error for reference station `BILL`, estimated using three different approaches: the LAM-PPP Slant and Spline approaches (in black and blue, respectively), and the collaborative approach (in red) using a network of 7 local agents. The ZTD error

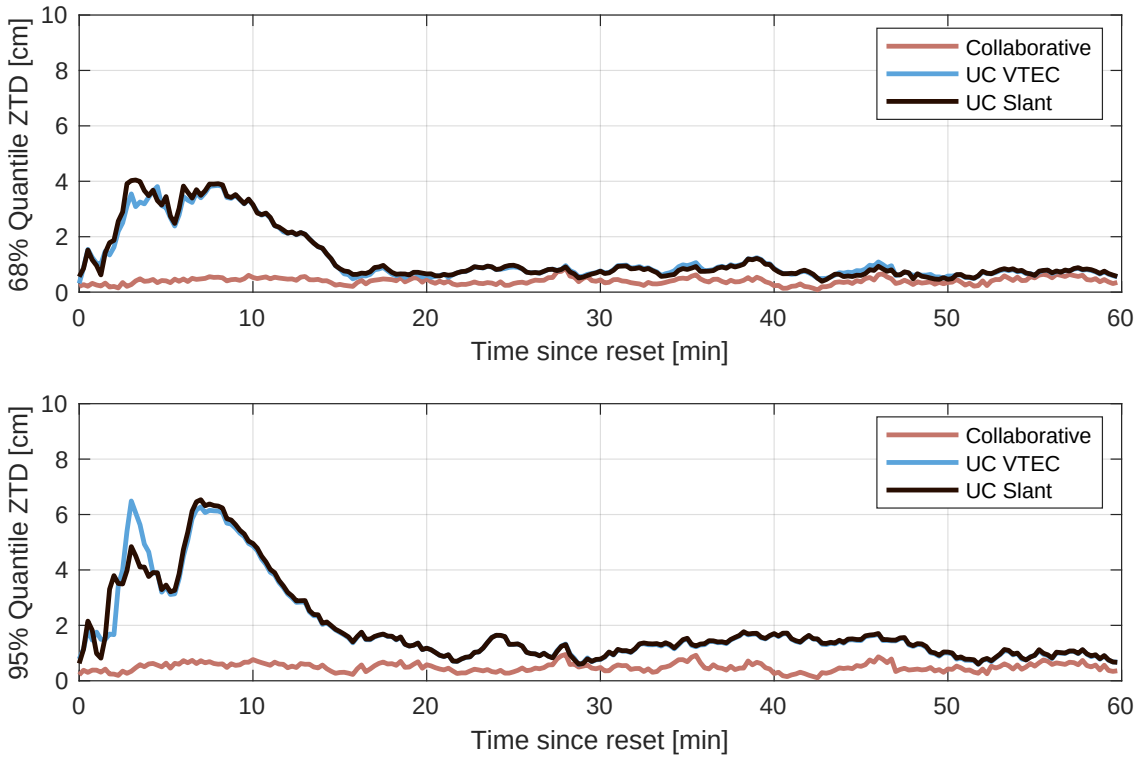
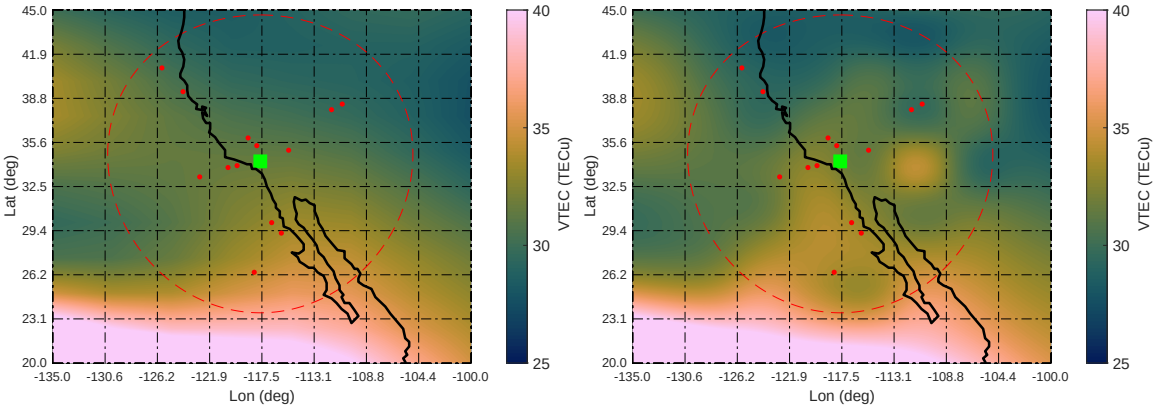


Figure 6.5: Comparison of ZTD error for IGS station `BILL`, using both individual approaches and the collaborative approach.

is computed using the IGS final ZTD product, as described in Section 5.5.1. The agents collaborate for 24 hours, resulting in 24 convergence periods when the KF is re-initialized every hour. Hence, for each epoch after re-initialization, there are 24 values of estimated ZTD error. The top and bottom panels show the 68% and 95% percentiles of the ZTD error, respectively, after every re-initialization. For instance, 10 minutes after re-initialization, 68% of the 24 convergence periods have a ZTD error of around 3 cm, and 95% have an error of 5 cm when using either of the individual approaches. As already discussed in Section 5.5.1, the individual approaches (in black and blue) have similar convergence and accuracy performance. They take roughly 10 to 20 minutes to reach an accuracy of 1 cm. The collaborative approach, on the other hand, In contrast, the collaborative approach



(a) GIM VTEC map

(b) Collaborative VTEC map

Figure 6.6: A Comparison of a regional VTEC map from: (a) GIM source, and (b) collaborative estimation. The map corresponds to 14:00 Pacific Time (local time), around the peak of ionospheric activity on that specific day. The green square is the user agent's location, and the red dots are the pierce points for used satellites at that agent.

achieves immediate convergence to millimeter-level accuracy upon each re-initialization.

6.6.3 Effect of Collaboration on Ionospheric Delay Estimates

6.6.3.1 Regional VTEC Maps

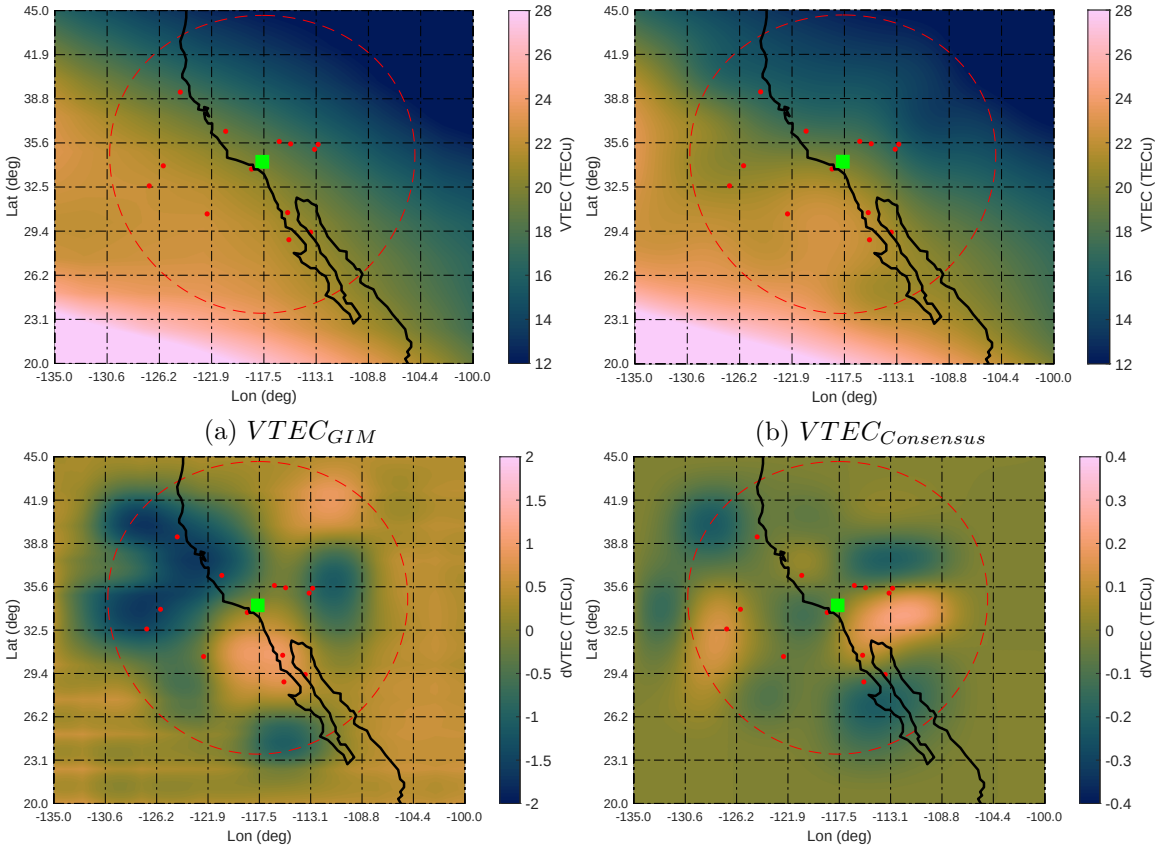
Fig. 6.6 shows a comparison of two VTEC maps: one obtained from a GIM source (Fig. 6.6a), and another obtained via collaborative estimation (Fig. 6.6b). The pierce points (in red) correspond to only one agent. The figure reveals that the collaborative approach captures more spatial details in the regional VTEC map. This suggests that the collaborative method is better equipped to model the localized variations in ionospheric delay, leading to more accurate and reliable estimates compared to global models.

Fig. 6.7 provides a visual assessment of the accuracy of regional VTEC maps derived from collaborative estimation compared to those from the GIM model and the local KF estimate of an agent. The top left panel (a) displays the VTEC map obtained from the GIM model, while the top right panel (b) presents the consensus map derived from collaborative

estimation. The bottom left panel (c) illustrates the difference between the collaborative estimate and the GIM estimate, highlighting the areas where the collaborative approach deviates from the global model. The bottom right panel (d) shows the difference between the collaborative estimate and the estimate obtained using the KF from one of the collaborating agents. The color bar on the right side of each panel indicates the VTEC or difference in VTEC values (dVTEC) in TEC units (TECu). Note the order of magnitude difference in dVTEC values when comparing the consensus map with those obtained from the GIM and KF. The consensus estimate slightly differs from the KF estimate of a collaborating agent because each agent's measurements have unique noise characteristics that influence the estimated B-spline parameters.

6.6.3.2 Instantaneous Convergence of Ionospheric Delays

The rapid convergence of position coordinates depends on the instantaneous convergence of atmospheric delays. The preceding subsection showed that the collaborative tropospheric delay estimate converges instantly after filter re-initialization. Fig. 6.8 shows a comparison of ionospheric slant delay estimates for all visible GAL satellites in a 4-hour period, estimated using the individual LAM-PPP Spline approach and the Collaborative approach. It is important to note that neither approach directly estimated the slant delays. Instead, both methods estimated a regional VTEC map, providing VTEC values at the satellite pierce points, which were subsequently used to compute the more intuitive slant delays. Fig. 6.8a illustrates that for LAM-PPP Spline, each re-initialization resets the slant delays of each visible satellite, requiring several epochs to regain the accuracy achieved in the prior convergence period. Fig. 6.8b shows that the re-initialization does not affect LAM-PPP



(a) $VTEC_{GIM}$

(b) $VTEC_{Consensus}$

(c) $VTEC_{Consensus} - VTEC_{GIM}$

(d) $VTEC_{Consensus} - VTEC_{KF}$

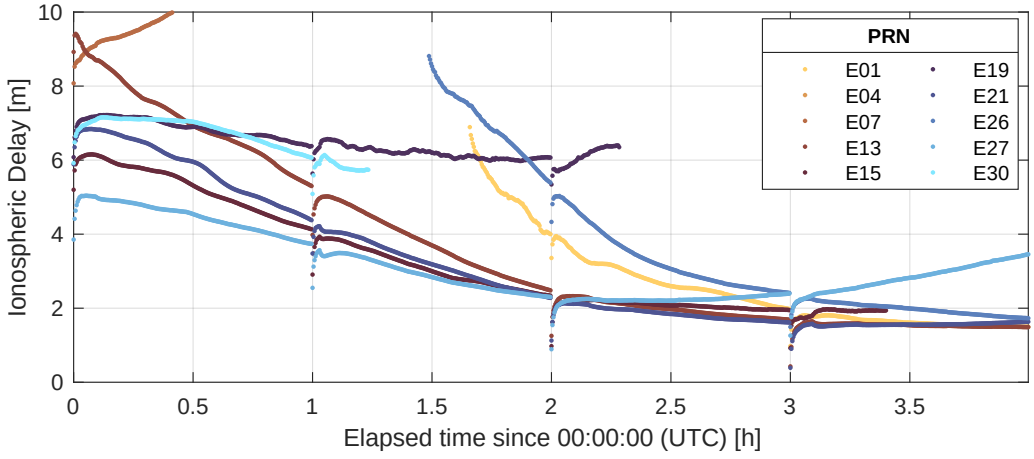
Figure 6.7: Comparison of a regional VTEC map derived from collaborative estimation compared to one from the GIM model.

Collab as much as it affects the LAM-PPP Spline approach (Fig. 6.8a). The slant delays largely maintain their accuracy.

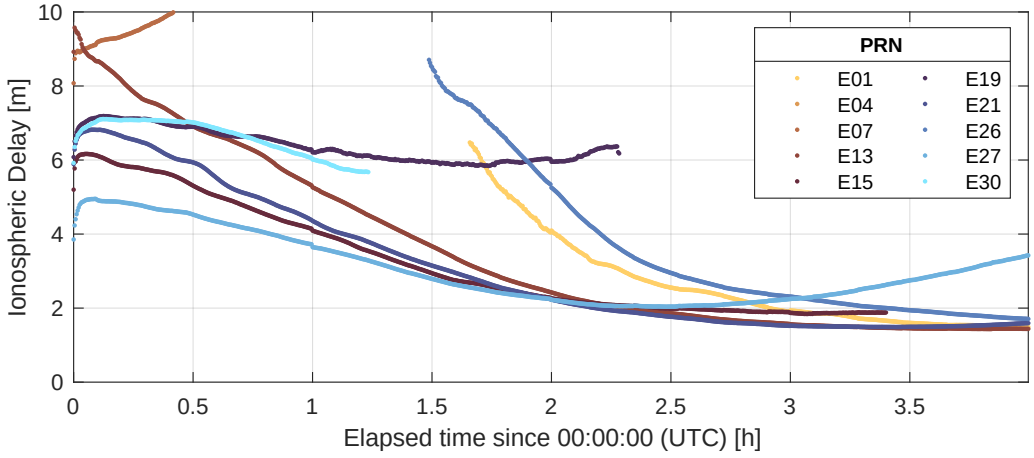
6.6.4 Effect of Collaboration on Position Convergence

The primary goal of the collaborative approach is to reduce PPP convergence time. Collaboration among agents speeds up the convergence of the estimated atmospheric delay states, which in turn speeds up the convergence of the estimated position coordinates.

To evaluate the effect of collaboration on position convergence, we set up an experiment involving a collaborating network of 20 agents located in a mid-latitude region, collaborating



(a) LAM-PPP Spline estimate

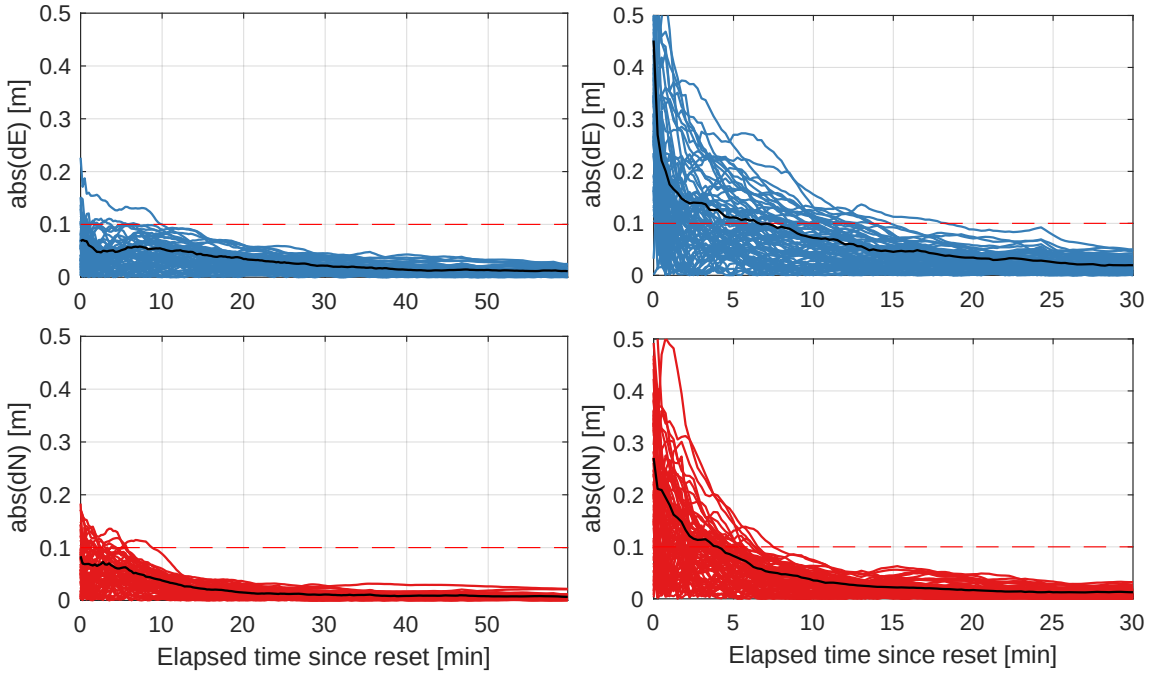


(b) Collaborative estimate

Figure 6.8: Estimates of slant ionospheric delays for the E5a frequency of visible GAL satellites. The delays are decreasing because they correspond to an afternoon period: 16:00 – 20:00 Pacific Time (local time).

for 4 hours. As a result, there are 60 convergence periods resulting from filter re-initialization every hour and retaining only the three convergence periods after the first re-initialization. Similarly, the same experiment is repeated twice, with each agent using one of the two individual approaches and similar KF settings. For a fair comparison, the receiver hardware from each agent is similar, i.e. they all used the same receiver model and similar antennae. See Table C.1 and Fig. C.1 for each agent's equipment details and location, respectively.

Fig. 6.9 illustrates the convergence of horizontal coordinates for both the collaborative



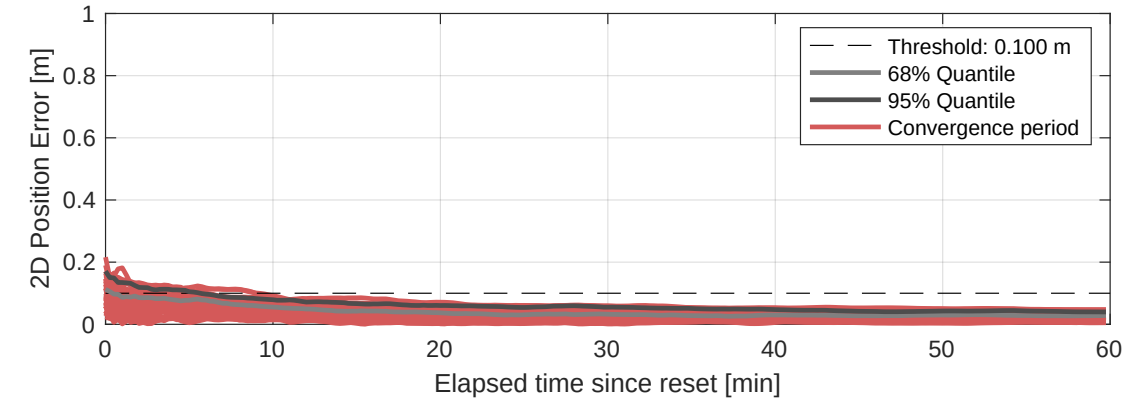
(a) Collaborative approach.

(b) LAM-PPP Spline approach.

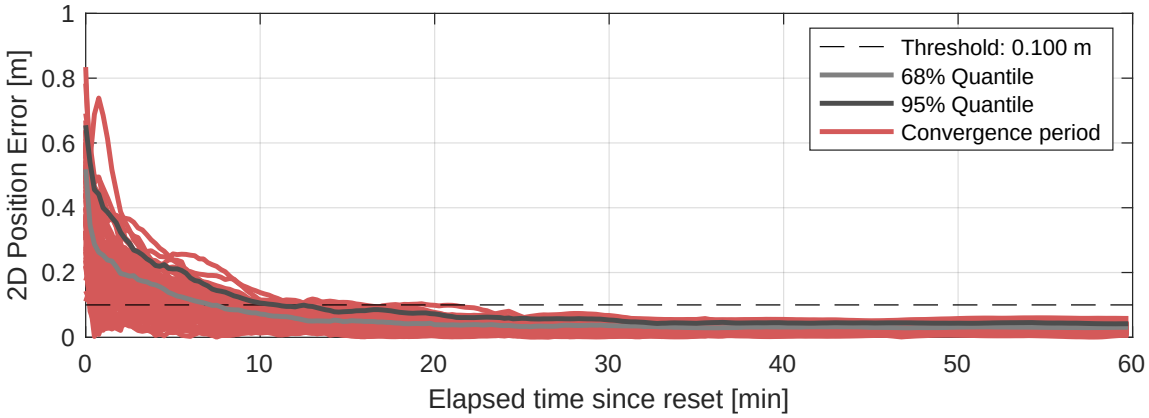
Figure 6.9: Comparison of absolute horizontal coordinates convergence using the Collaborative and LAM-PPP Spline approaches. Each plot has 60 convergence periods; 3 from each of the 20 agents. The black lines in each plot represent the 68% percentiles of convergence periods.

and LAM-PPP Spline approaches. The x-axis represents the time elapsed since the last filter re-initialization, while the y-axis shows the absolute difference between the estimated and true north (dN) or east (dE) coordinates in meters. Each colored line represents a single convergence period, and the black line indicates the 68th percentile of all convergence periods. The 68th percentile lines show that, on average, the collaborative approach achieves instant convergence of horizontal coordinates, whereas LAM-PPP Spline takes around 5 minutes. However, after 60 minutes, both approaches demonstrate comparable final accuracy.

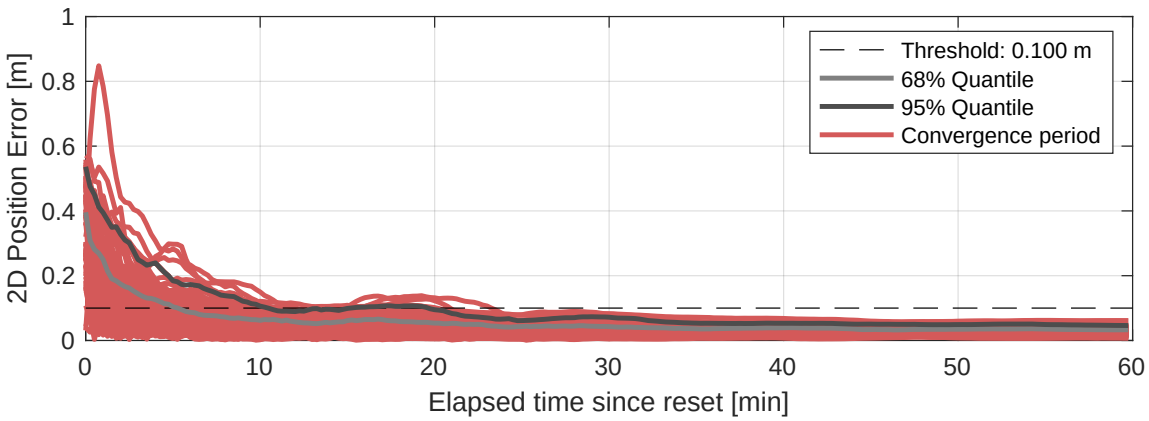
Fig. 6.10 illustrates the convergence of the 2D position error for three different PPP approaches; the collaborative approach (Collab), and the two individual LAM-PPP ap-



(a) Collab



(b) LAM-PPP Spline



(c) LAM-PPP Slant

Figure 6.10: Comparison of the convergence behavior of 2D position error estimated using the collaborative approach and both individual approaches (LAM-PPP Spline and LAM-PPP Slant). The convergence periods are shown in red.

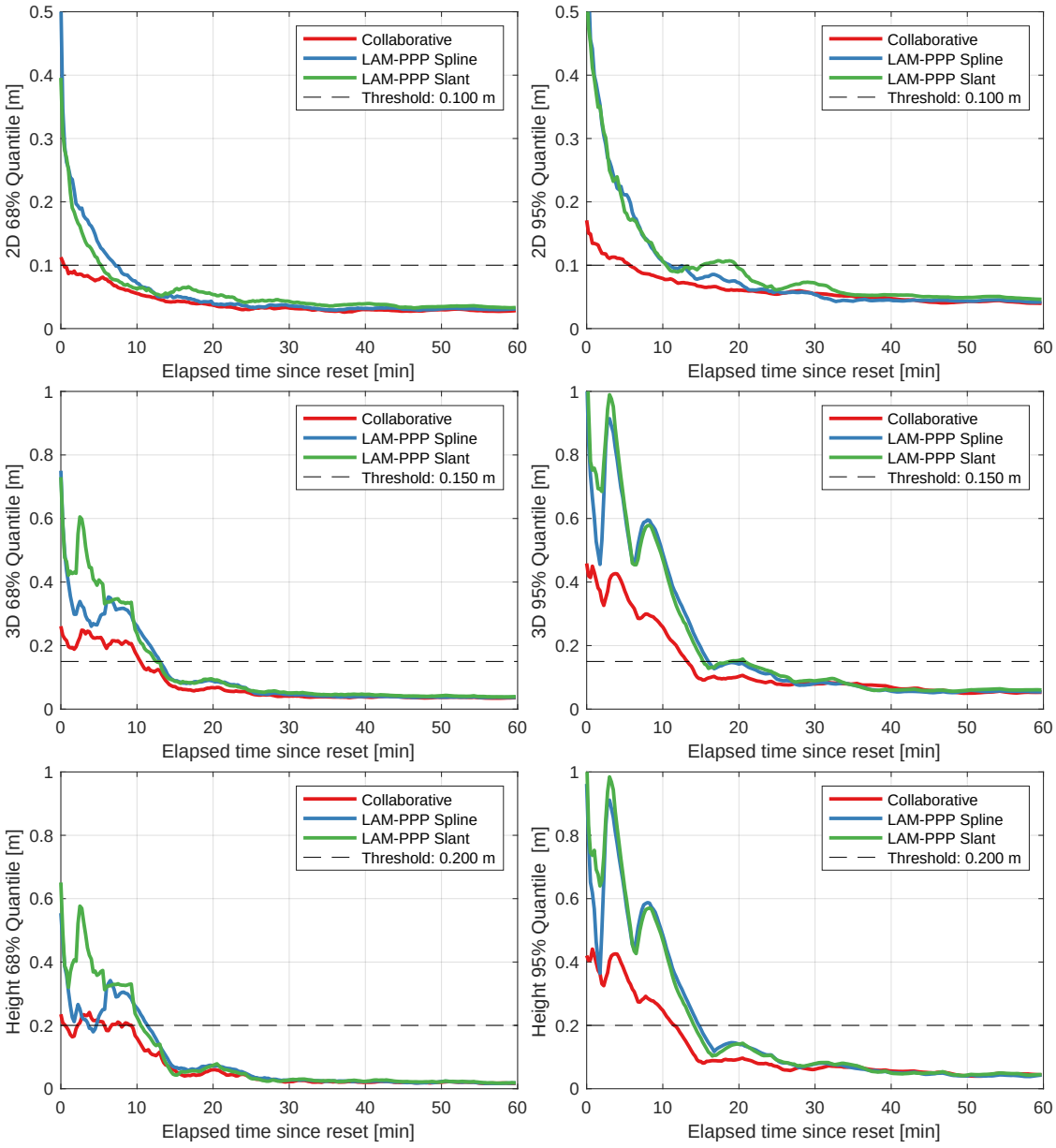
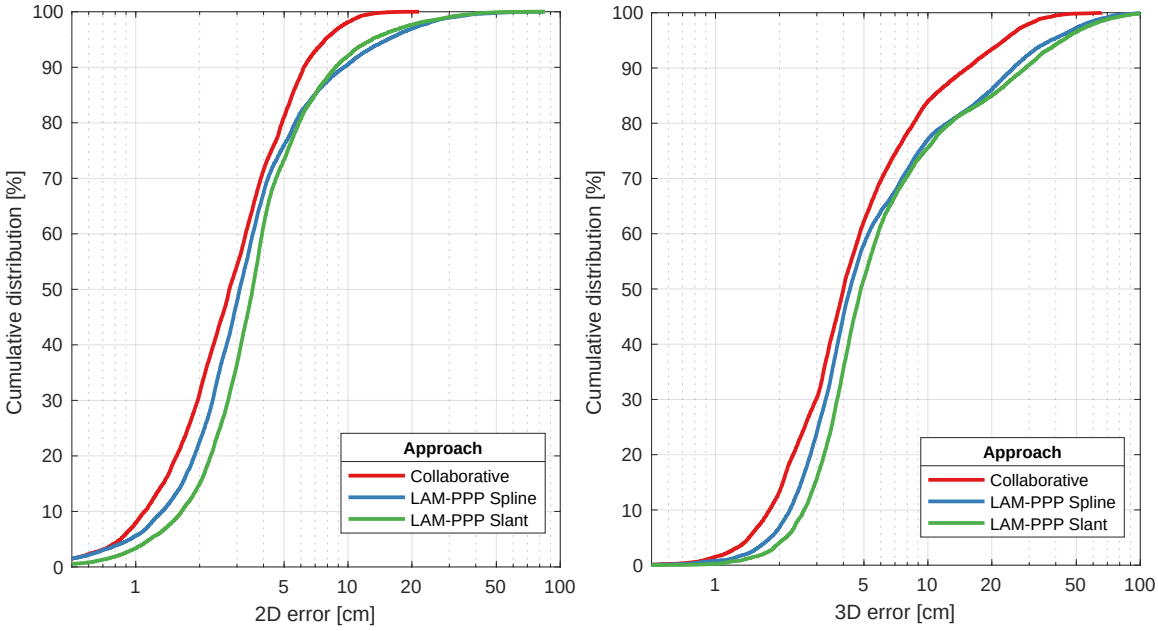


Figure 6.11: Percentile convergence of position metrics for different estimation approaches, using 20 stations. The metrics are: height error (top panels), 2D error (middle panels), and 3D error (bottom panels). The left-hand-side panels and right-hand-side panels show the 68% and 95% percentiles, respectively.

proaches: one using a B-spline model (LAM-PPP Spline) and another using a slant model (LAM-PPP Slant). The 2D position error, representing the horizontal error in meters, is plotted against the elapsed time since the last filter reset in minutes. (The 2D position error



(a) (b)

Figure 6.12: Cumulative distributions of the 2D and 3D errors for the two individual approaches and the collaborative approach. The plots correspond to observations in a 4-hour period: 16:00–20:00 PT (local time).

was defined in eq. (5.86).) The 68% and 95% quantiles are displayed for each approach, indicating the range within which the specified percentage of errors falls. The collaborative approach demonstrates significantly faster convergence than the other two methods, achieving a 2D position error below 10 cm within the first minute. In contrast, the LAM-PPP Spline and LAM-PPP Slant approaches require approximately 10 minutes to reach the same level of accuracy. This highlights the superior performance of the collaborative approach in rapidly converging to a precise position solution.

Fig. 6.11 extends the analysis of Fig. 6.10 to show the convergence statistics (68% and 95% percentiles) of not just the 2D position error, but also the 3D position error and the height error. As previously discussed in Section 5.5.3, both figures illustrate the comparable convergence behavior of the individual approaches, requiring a similar amount of time to reach 10 cm, 15 cm, and 20 cm for the 2D, 3D, and height errors, respectively. In contrast,

the collaborative approach demonstrates superior performance, converging more rapidly for each type of error. The curves for the collaborative approach are consistently lower than those of the other two approaches This indicates that the collaborative approach achieves any accuracy specification faster than the individual approaches. For example, to reach a 2D error of 10 cm, the collaborative approach requires less time than the LAM-PPP Spline and LAM-PPP Slant approaches. As anticipated, Fig. 6.11 demonstrates that the ultimate positioning accuracy after convergence is comparable across all approaches, with the collaborative approach exhibiting a marginal improvement. This is because the final accuracy is primarily determined by the precision of each agent's pseudorange measurements, which remains the same across all methods. Therefore, while collaboration accelerates convergence, it does not significantly impact the final accuracy.

Fig. 6.12 displays the Cumulative Distribution Functions (CDFs) of the 2D and 3D position errors for the three approaches. The x-axis represents the position error in centimeters, while the y-axis shows the cumulative percentage of occurrences. The collaborative approach demonstrates a slight improvement in accuracy compared to the individual approaches, particularly in the higher percentiles. This suggests that collaboration not only speeds up convergence but also leads to slightly more accurate position estimates overall. Figs. 6.11 and 6.12 highlight the superior performance of the collaborative approach in terms of both convergence speed and accuracy. The collaborative approach consistently outperforms the individual approaches across a wide range of error thresholds, indicating its effectiveness in achieving and maintaining high-precision positioning. It is clear that the Collab approach outperforms both individual approaches at nearly all percentiles. This demonstrates that not only does the Collab approach converge faster, the sharing of infor-

mation enables it to maintain the accuracy throughout the convergence period.

Table 6.2 shows statistics of three performance indicators for the three approaches. For each indicator, the table provides the mean and median convergence time (in minutes) and the mean and standard deviation of the final 2D and 3D accuracies (in centimeters).

As discussed in Section 5.5.3, the performance of the Spline and Slant approaches is similar. The collaborative approach outperforms both individual approaches in all metrics of every performance indicator. Crucially, the convergence time decreases dramatically when agents collaborate. The median convergence time for this particular network is 15 seconds, compared to 315 and 360 seconds for the Slant and Spline approaches, respectively. This is a more than 20 times reduction in median convergence time. Since the sampling interval is 15 seconds, in half of the convergence periods, the horizontal coordinates of agents joining an already converged network convergence after just one epoch. This suggests that the instantaneous convergence of atmospheric delays, is a prerequisite for achieving rapid convergence in the estimated position coordinates. While the collaborative approach also shows a slight improvement in the final 2D and 3D accuracies compared to the individual approaches, the differences are not as significant as the differences in convergence time. This suggests that the primary benefit of the collaborative approach is the faster convergence, rather than a substantial improvement in the final accuracy. Fig. 6.13 visualizes the performance indicators in Table 6.2, making it more evident that the collaborative approach drastically reduces convergence time. Overall, Table 6.2 highlights the importance of accurate and timely regional atmospheric delay corrections in PPP, especially for real-time applications where fast convergence is crucial.

Performance indicator	Metric	Slant	Spline	Collaborative
Convergence time[a] [min]	Mean	6.67	6.34	**1.40**
	Median	5.25	6.00	**0.25**
Final 2D accuracy[b] [cm]	Mean	2.96	2.67	**2.46**
	Stdev	1.05	1.06	**0.99**
Final 3D accuracy[c] [cm]	Mean	3.58	3.28	**3.06**
	Stdev	1.23	1.18	**1.17**

[a] The time it takes for the 2D position error to fall below 10 cm.
[b] The 2D position error (E_{2D}) after 60 minutes.
[c] The 3D position error (E_{3D}) after 60 minutes.

Table 6.2: Statistical comparison of three performance indicators for the individual LAM-PPP Spline and LAM-PPP Slant approaches, and the collaborative approach. The statistics are derived by processing datasets from 20 mid-latitude CORS stations, each with 3 convergence periods.

6.7 Summary

This chapter developed and demonstrated a method of receiver collaboration in estimating the regional atmospheric delays. This was accomplished by having a network of agents share their estimates of the atmospheric delay states (ZWD and VTEC B-spline parameters), without requiring them to share raw measurements. The collaborative estimation problem was formulated as an average consensus problem, and a distributed information-weighted consensus algorithm was developed to solve it. The algorithm ensures that each agent in the network is able to compute the optimal estimate by only communicating with its neighbors.

The results show that the collaborative approach significantly outperforms the individual estimation approaches in terms of convergence time. For example, it was shown that a 20 times reduction in median convergence time is possible under favorable conditions. The results also show that the collaborative approach is able to estimate a regional VTEC map with more spatial details than the GIM map.

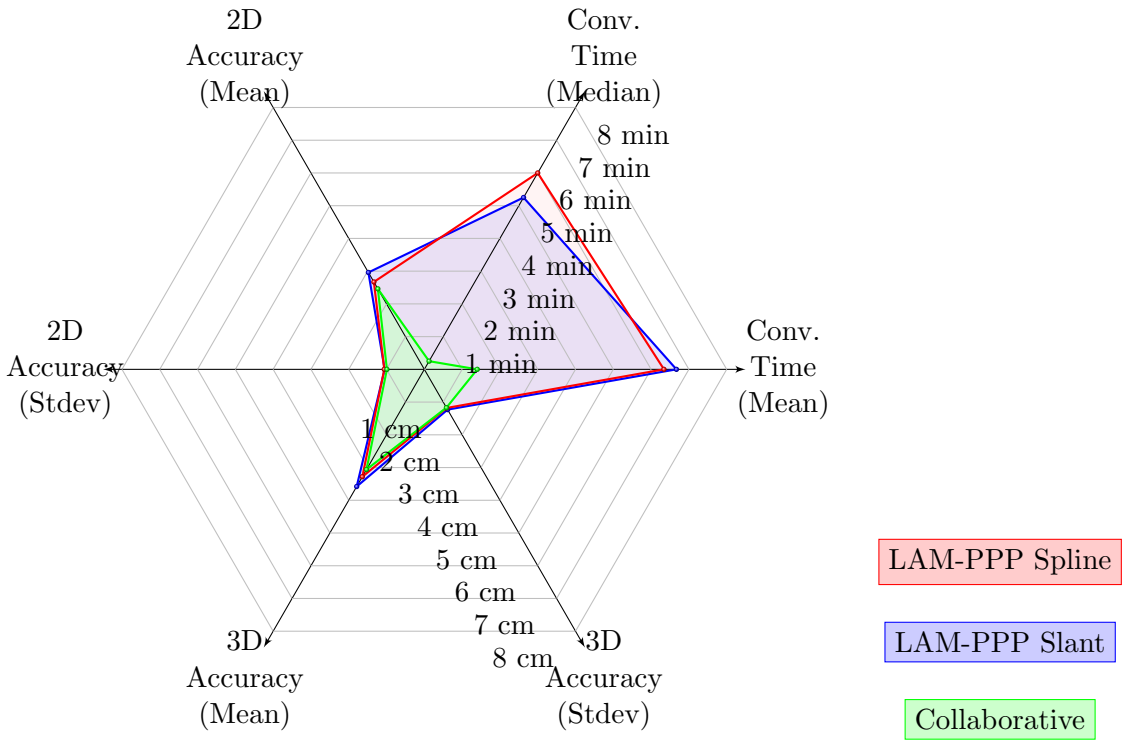


Figure 6.13: A diagram visualizing the metrics of three performance indicators for the LAM-PPP Spline, LAM-PPP Slant, and collaborative approaches.

The main advantage of the collaborative approach is that it enables agents to join the network and instantly get a very precise estimate of the atmospheric delays. This is a significant advantage for real-time applications such as ITS requiring lane-level navigation, where precise positioning is required within seconds of the receiver being switched on.

Chapter 7

Discussion and Future Directions

7.1 Conclusion

This dissertation has introduced and developed two innovative approaches aimed at reducing the convergence time of PPP using a network of GNSS receivers. The research concentrated on refining the estimation of regional atmospheric delays, which have been identified as the primary source of residual errors in PPP.

The first approach, LAM-PPP Spline, enables collaboration between satellites to estimate local atmospheric delays using observations from a single receiver. This technique leverages the expanding capabilities of multi-GNSS by utilizing uncombined measurements from all usable satellites. The ionospheric delay is modeled as a regional VTEC map, parameterized using polynomial B-splines. The B-spline parameters and the residual tropospheric delay are estimated using a Kalman Filter. Results showed that the proposed approach accurately estimates local atmospheric delays, thereby improving the overall PPP solution.

Furthermore, the dissertation proposes a collaborative PPP approach that utilizes estimated regional atmospheric corrections from a network of local agents. This approach combines corrections from global models with estimates from local collaborating agents to derive more accurate regional corrections. The resulting collaborative estimation problem is formulated as an average consensus problem, and a distributed information-weighted consensus algorithm is developed to solve it. The algorithm allows agents to share only their atmospheric delay estimates and uncertainties with their neighbors. This distributed approach eliminates the need for a central processing center and enables agents to instantaneously achieve network accuracy upon joining the network, thereby significantly improving PPP convergence time. The collaborative PPP approach is shown to achieve instantaneous convergence for some users under favorable conditions, demonstrating its potential for real-time precise positioning applications.

The proposed approaches promise to significantly improve the performance of PPP, making it more viable for real-time ITS applications that require high accuracy and rapid convergence.

7.2 Future Directions

7.2.1 Collaborative Ambiguity Resolution

Ambiguity Resolution (AR) is crucial for achieving instantaneous convergence in PPP. The residual atmospheric delay problem addressed in this dissertation has been a major obstacle to resolving ambiguities, and the solutions presented here could significantly enhance AR. However, challenges remain for successful full ambiguity resolution. In addition

to residual atmospheric delays, the lack of accurate phase bias estimates poses a significant challenge for the uncombined measurement models used in this work. Ambiguities for newly initialized signals, such as those from low-elevation satellites, are often difficult to resolve into integers. One potential solution is Partial Ambiguity Resolution (PAR), which focuses on resolving a subset of ambiguities rather than the entire set. This technique has gained attention in recent literature [109–111] A promising avenue for future research would be to explore PAR using collaborative LAM.

7.2.2 Kinematic Collaborative PPP

A straightforward extension of this work would be to adapt it to kinematic scenarios. The long convergence time of traditional PPP has typically hindered its use in kinematic applications. However, the rapid convergence achieved in this research could make PPP a viable option for real-time kinematic positioning. This could involve incorporating a Position-Velocity-Acceleration (PVA) model to account for vehicle dynamics and refining the stochastic modeling to better reflect the dynamic nature of the system.

7.2.3 Integration with INS

Integrating the proposed approaches with an Inertial Navigation System (INS) could bring several benefits. INS provides very accurate short-term position and velocity estimates, which can help bridge gaps in GNSS signals during periods of obstruction or signal loss. Combining INS with the precise atmospheric delay estimates from the proposed approaches can lead to a more robust and reliable navigation solution, especially in challenging environments where GNSS signals may be degraded. Additionally, the integration of INS

can aid in faster ambiguity resolution, further improving the convergence time and accuracy of PPP.

7.2.4 Collaboration Using Low-cost Receivers

One of the main advantages of both the proposed LAM-PPP Spline and collaborative approaches over traditional approaches is their potential to better utilize GNSS observations from low-cost receivers, such as premium smartphones and automotive-grade receivers [112]. These receivers often have a limited channels to track all available frequency bands and may lack advanced multipath mitigation capabilities, posing challenges for precise atmospheric modeling. However, the widespread availability of these devices presents a unique opportunity to leverage their observations for enhanced atmospheric modeling. Recent studies have demonstrated the potential of crowd-sourced smartphone observations for precise tropospheric delay monitoring [113]. Future research could focus on developing algorithms specifically tailored to incorporate observations from low-cost receivers, addressing their limitations and harnessing their potential for improving PPP performance.

7.2.5 Adaptive Regional VTEC Maps

The collaborative regional VTEC map generated in this dissertation has a predefined resolution and size. The designer must choose beforehand the geographic area the map covers and the resolution of the resulting map. This choice is straightforward when all agents in the network are static or operating in a network with predefined behavior. However, for ad-hoc networks with dynamic agent arrivals and departures, the regional VTEC map must adapt to the network's changing characteristics. Future research could focus on developing

collaborative approaches that can adapt to changing network dynamics.

7.3 Publications Resulting from PhD Study

The following is a list of publications resulting from this PhD study:

Published

- Jean-Bernard Uwineza and Jay A. Farrell. *"Cooperative Estimation of Local Atmospheric Delay in Precise Point Positioning with GNSS."* IEEE Conference on Control Technology and Applications (IEEE CCTA). Trieste, Italy. August 22-25, 2022.

- Jean-Bernard Uwineza and Jay A. Farrell. *"RAIM and Failure Mode Slope: Effects of Increased Number of Measurements and Number of Faults."* Sensors (Basel, Switzerland) 23.10 (2023).

- Jean-Bernard Uwineza and Jay A. Farrell. *"Supplementary Materials: RAIM and Failure Mode Slope: Effects of Increased Number of Measurements and Number of Faults."* Sensors (Basel, Switzerland) 23.10 (2023).

- Wang Hu, Jean-Bernard Uwineza, Jay A. Farrell. *"Outlier Accommodation for Multi-GNSS Precise Point Positioning using Risk-Averse Performance-Specified Approach."* American Control Conference (ACC), Toronto, Canada. July 8-12, 2024.

Working Papers

- Jean-Bernard Uwineza and Jay A. Farrell. *"Collaborative PPP Estimation for Transportation Applications."* In Preparation.

- Felipe O. Silva, Jean-Bernard Uwineza, Farzana R. Rahman, Zeyi Jiang, Wang Hu, Gustavo S. Carvalho, Jay A. Farrell. *"Dual-antenna GNSS-aided INS Stationary Alignment with Sensor Parameter Estimation."* In Preparation.

- Ashim Neupane, Wang Hu, Jean-Bernard Uwineza, and Jay A. Farrell. *"A Tutorial on Robust Regression Methods in Linear State Estimation."* In Preparation.

Other Publications

- Jean-Bernard Uwineza, Farzana Rahman, Felipe O. Silva, Wang Hu, and Jay A. Farrell. (2019). *"Characterizing GNSS Multipath at Different Antenna Mounting Positions on Vehicles."* Sirius XM Technical Report. 2019. Online: https://escholarship.org/content/qt9mj3459s/qt9mj3459s.pdf

- Farzana Rahman, Jean-Bernard Uwineza, and Jay A. Farrell. (2020). *"ECEF Position Accuracy and Reliability: Inertial Navigation with GNSS Precise Point Positioning (PPP)."* Sirius XM Technical Report. 2020. Online: https://escholarship.org/content/qt4km0d6pz/qt4km0d6pz.pdf

- Jean-Bernard Uwineza (2021). *"Towards Adaptive Eco-Approach and Departure For Actuated Signalized Corridors."* Technical Report. 2021. Online: https://escholarship.org/content/qt9jk0v98t/qt9jk0v98t.pdf

Bibliography

[1] I. Fernandez-Hernandez, T. Senni, D. Calle, S. Cancela, G. A. Vecchione, and G. Seco-Granados, "Analysis of high-accuracy satellite messages for road applications," *IEEE Intelligent Transportation Systems Magazine*, vol. 12, no. 3, pp. 92–108, 2020.

[2] "Report on Road User Needs and Requirements: Outcome of the EUSPA User Consultation Platform. Issue/ Revision 3.0," European Union Ageny for the Space Programme, Tech. Rep., September 2021. [Online]. Available: https://www.gsc-europa.eu/sites/default/files/sites/all/files/Report_on_User_Needs_and_Requirements_Road.pdf

[3] C. J. Hegarty and E. Chatre, "Evolution of the global navigation satellite system (GNSS)," *Proc. of the IEEE*, vol. 96, no. 12, pp. 1902–1917, 2008.

[4] S. Choy, S. Bisnath, and C. Rizos, "Uncovering common misconceptions in GNSS Precise Point Positioning and its future prospect," *GPS Solutions*, vol. 21, no. 1, pp. 13–22, 2017.

[5] S. Banville, P. Collins, W. Zhang, and R. B. Langley, "Global and regional ionospheric corrections for faster PPP convergence," *NAVIGATION: Journal of the Institute of Navigation*, vol. 61, no. 2, pp. 115–124, 2014.

[6] X. An, R. Ziebold, and C. Lass, "From RTK to PPP-RTK: towards real-time kinematic precise point positioning to support autonomous driving of inland waterway vessels," *GPS Solutions*, pp. 1–16, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s10291-023-01428-2

[7] Z. Yan and X. Zhang, "Assessment of the performance of GPS/Galileo PPP-RTK convergence using ionospheric corrections from networks with different scales," *Earth, Planets and Space*, vol. 74, no. 1, p. 47, 2022.

[8] P. Hou and B. Zhang, "Decentralized GNSS PPP-RTK," *Journal of Geodesy*, vol. 97, no. 7, pp. 1–18, jul 2023. [Online]. Available: https://link.springer.com/article/10.1007/s00190-023-01761-x

[9] J.-B. Uwineza and J. A. Farrell, "RAIM and Failure Mode Slope: Effects of Increased Number of Measurements and Number of Faults," *Sensors*, vol. 23, no. 10, p. 4947, 2023.

[10] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas (Second Edition)*. Princeton University Press, 2009.

[11] M. Håkansson, A. B. Jensen, M. Horemuz, and G. Hedling, "Review of code and phase biases in multi-GNSS positioning," *GPS Solutions*, vol. 21, no. 3, pp. 849–860, jul 2017. [Online]. Available: https://link.springer.com/article/10.1007/s10291-016-0572-7

[12] C. H. Yinger, W. A. Feess, R. D. Esposti, A. Chasko, B. Cosentino, D. Syse, H. Air, F. Base, M. B. Wheaton, and S. M. C. Czut, "GPS Satellite Interfrequency Biases," in *Proceedings of the 55th Annual Meeting of The Institute of Navigation*, June 1999, pp. 347–354.

[13] O. Montenbruck, A. Hauschild, and P. Steigenberger, "Differential Code Bias Estimation using Multi-GNSS Observations and Global Ionosphere Maps," *Navigation, Journal of the Institute of Navigation*, vol. 61, no. 3, pp. 191–201, sep 2014.

[14] S. Beer, L. Wanninger, and A. Heßelbarth, "Estimation of absolute GNSS satellite antenna group delay variations based on those of absolute receiver antenna group delays," *GPS Solutions*, vol. 25, no. 3, pp. 1–10, 2021. [Online]. Available: https://doi.org/10.1007/s10291-021-01137-8

[15] A. Hauschild, O. Montenbruck, J. M. Sleewaegen, L. Huisman, and P. J. Teunissen, "Characterization of Compass M-1 signals," *GPS Solutions*, vol. 16, no. 1, pp. 117–126, 2012.

[16] S. Beer, L. Wanninger, and A. Heßelbarth, "Galileo and GLONASS group delay variations," *GPS Solutions*, vol. 24, no. 1, pp. 1–8, jan 2020. [Online]. Available: https://link.springer.com/article/10.1007/s10291-019-0939-7

[17] R. Zhou, Z. Hu, Q. Zhao, P. Li, W. Wang, C. He, C. Cai, and Z. Pan, "Elevation-dependent pseudorange variation characteristics analysis for the new-generation BeiDou satellite navigation system," *GPS Solutions*, vol. 22, no. 3, pp. 1–11, jul 2018. [Online]. Available: https://link.springer.com/article/10.1007/s10291-018-0726-x

[18] G. Wübbena, M. Schmitz, and A. Warneke, "Geo++ absolute multi frequency GNSS antenna calibration," in *EUREF AC Workshop*, 2019. [Online]. Available: http://www.geopp.com/pdf/gpp_cal125_euref19_p.pdf

[19] T. Kersten and S. Schön, "GNSS group delay variations - Potential for improving GNSS based time and frequency transfer?" in *43rd Annual Precise Time and Time Interval Systems and Applications Meeting*, 2011, pp. 255–269.

[20] J.-T. Wu, S. C. Wu, G. A. Hajj, W. I. Bertiger, and S. M. Lichten, "Effects of antenna orientation on GPS carrier phase," *Astrodynamics 1991*, pp. 1647–1660, 1992.

[21] J. Kouba, "A guide to using International GNSS Service (IGS) products," 2009. [Online]. Available: http://acc.igs.org/UsingIGSProductsVer21.pdf

[22] B. Luzum and G. Petit, "The IERS Conventions (2010): Reference systems and new models," *Proceedings of the International Astronomical Union*, vol. 10, no. H16, pp. 227–228, 2012.

[23] G. Petit and B. Luzum, "IERS conventions (2010)," *IERS Technical Note*, vol. 36, p. 180, 01 2010. [Online]. Available: https://www.iers.org/SharedDocs/Publikationen/EN/IERS/Publications/tn/TechnNote36/tn36.pdf

[24] H. Dragert, T. James, and A. Lambert, "Ocean loading corrections for continuous GPS: A case study at the Canadian coastal site Holberg," *Geophysical research letters*, vol. 27, no. 14, pp. 2045–2048, 2000.

[25] L. Petrov and J.-P. Boy, "Study of the atmospheric pressure loading signal in very long baseline interferometry observations," *Journal of Geophysical Research: Solid Earth*, vol. 109, no. B3, 2004.

[26] C. Chao, "New tropospheric range corrections with seasonal adjustment," *Technical Report 32-1526 Volume VI, Jet Propulsion Laboratory*, vol. 77, no. 4, p. 67, 1971.

[27] J. Böhm, J. Kouba, and H. Schuh, "Forecast Vienna Mapping Functions 1 for real-time analysis of space geodetic observations," *Journal of Geodesy*, vol. 83, pp. 397–401, 2009.

[28] J. Askne and H. Nordius, "Estimation of tropospheric delay for microwaves from surface weather data," *Radio science*, vol. 22, no. 03, pp. 379–386, 1987.

[29] T. Nilsson, J. Böhm, D. Wijava, A. Tresch, V. Nafisi, and H. Schuh, "Path delays in the neutral atmosphere," in *Atmospheric Effects in Space Geodesy*. Springer, 2013, pp. 73–136.

[30] J. Böhm and H. Schuh, *Atmospheric effects in space geodesy*. Springer, 2013, vol. 5.

[31] J. Saastamoinen, "Atmospheric correction for the troposphere and stratosphere in radio ranging satellites," *The Use of Artificial Satellites for Geodesy*, vol. 15, pp. 247–251, 1972.

[32] J. Davis, T. Herring, I. Shapiro, A. Rogers, and G. Elgered, "Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length," *Radio science*, vol. 20, no. 6, pp. 1593–1607, 1985.

[33] H. S. Hopfield, "Tropospheric effect on electromagnetically measured range: Prediction from surface weather data," *Radio science*, vol. 6, no. 3, pp. 357–367, 1971.

[34] ——, "Two-quartic tropospheric refractivity profile for correcting satellite data," *Journal of Geophysical Research*, vol. 74, no. 18, pp. 4487–4499, 1969.

[35] R. Leandro, M. Santos, and R. Langley, "UNB neutral atmosphere models: development and performance," in *Proceedings of the 2006 National Technical Meeting of the Institute of Navigation*, 2006, pp. 564–573.

[36] J. Böhm, R. Heinkelmann, and H. Schuh, "Short note: a global model of pressure and temperature for geodetic applications," *Journal of Geodesy*, vol. 81, pp. 679–683, 2007.

[37] J. Böhm, G. Möller, M. Schindelegger, G. Pain, and R. Weber, "Development of an improved empirical model for slant delays in the troposphere (GPT2w)," *GPS solutions*, vol. 19, pp. 433–441, 2015.

[38] D. Landskron and J. Böhm, "VMF3/GPT3: refined discrete and empirical troposphere mapping functions," *Journal of Geodesy*, vol. 92, pp. 349–360, 2018.

[39] T. Herring, "Modeling atmospheric delays in the analysis of space geodetic data," *Publications on Geodesy, Proceedings of Refraction of Transatmospheric Signals in Geodesy*, vol. 36, pp. 157–164, 1992.

[40] C. Chao, "A model for tropospheric calibration from daily surface and radiosonde balloon measurements," *Tech. Report. 391-350, Jet Propulsion Laboratory*, 1972.

[41] A. Niell, "Global mapping functions for the atmosphere delay at radio wavelengths," *Journal of Geophysical Research*, vol. 101, no. B2, pp. 3227–3246, 1996.

[42] A. E. Niell, "Improved atmospheric mapping functions for VLBI and GPS," *Earth Planets Space*, vol. 52, pp. 699–702, 2000.

[43] W. Li, Y. Yuan, J. Ou, H. Li, and Z. Li, "A new global zenith tropospheric delay model IGGtrop for GNSS applications," *Chinese Science Bulletin*, vol. 57, pp. 2132–2139, 2012.

[44] W. Li, Y. Yuan, J. Ou, and Y. He, "IGGtrop_SH and IGGtrop_rH: Two improved empirical tropospheric delay models based on vertical reduction functions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5276–5288, 2018.

[45] M. Abdelazeem, "AFRC-Trop: New real-time zenith tropospheric delay model over Africa," *Journal of Surveying Engineering*, vol. 147, no. 2, 2021.

[46] Z. Du, Q. Zhao, W. Yao, and Y. Yao, "Improved GPT2w (IGPT2w) model for site specific zenith tropospheric delay estimation in China," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 198, 2020.

[47] G. Chen and T. Herring, "Effects of atmospheric azimuthal asymmetry on the analysis of space geodetic data," *Journal of Geophysical Research: Solid Earth*, vol. 102, no. B9, pp. 20 489–20 502, 1997.

[48] D. Landskron, "Modeling tropospheric delays for space geodetic techniques," Ph.D. dissertation, Technische Universität Wien, 2017.

[49] N. Putri, D. Landskron, and J. Böhm, "Assessing the performance of vienna mapping functions 3 for GNSS stations in Indonesia using precise point positioning," *Advances in Geosciences*, vol. 50, pp. 77–86, 2020.

[50] P. Kintner, T. Humphreys, and J. Hinks, "GNSS and ionospheric scintillation," *Inside GNSS*, vol. 4, no. 4, pp. 22–30, 2009.

[51] Y. Jiao, Y. T. Morton, S. Taylor, and W. Pelgrum, "Characterization of high-latitude ionospheric scintillation of GPS signals," *Radio Science*, vol. 48, no. 6, pp. 698–708, 2013.

[52] M. Hernández-Pajares, J. M. Juan, J. Sanz, R. Orus, A. Garcia-Rigo, J. Feltens, A. Komjathy, S. C. Schaer, and A. Krankowski, "The IGS VTEC maps: A reliable source of ionospheric information since 1998," *Journal of Geodesy*, vol. 83, pp. 263–275, Feb 2009.

[53] A. Chulliat, W. Brown, P. Alken, C. Beggan, M. Nair, G. Cox, A. Woods, S. Macmillan, B. Meyer, and M. Paniccia, "The US/UK World Magnetic Model for 2020-2025: Technical Report," 2020. [Online]. Available: https://doi.org/10.25923/ytk1-yx35

[54] E. Astafyeva, I. Zakharenkova, and M. Förster, "Ionospheric response to the 2015 St. Patrick's Day storm: A global multi-instrumental overview," *Journal of Geophysical Research: Space Physics*, vol. 120, no. 10, pp. 9023–9037, 2015.

[55] "Global Positioning System Standard Positioning Service Performance Standard, 5th Edition," US Department of Defense (DoD), Washington DC, Tech. Rep., 2020. [Online]. Available: https://www.gps.gov/technical/ps/2020-SPS-performance-standard.pdf

[56] C. J. Hegarty, "The global positioning system (GPS)," *Springer Handbook of Global Navigation Satellite Systems*, pp. 197–218, 2017.

[57] O. Montenbruck, P. Steigenberger, and A. Hauschild, "Comparing the 'Big 4'-A User's View on GNSS Performance," in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. IEEE, 2020, pp. 407–418.

[58] G. T. Kremer, R. M. Kalafus, P. V. Loomis, and J. C. Reynolds, "The effect of selective availability on differential GPS corrections," *Navigation*, vol. 37, no. 1, pp. 39–52, 1990.

[59] D. Odijk and L. Wanninger, "Differential Positioning," in *Springer Handbook of Global Navigation Satellite Systems*, P. J. Teunissen and O. Montenbruck, Eds. Springer, 2017, pp. 753–780. [Online]. Available: https://doi.org/10.1007/978-3-319-42928-1_26

[60] H. Vo and J. Foster, "A quantitative study of ionospheric density gradients at mid-latitudes," *Journal of Geophysical Research: Space Physics*, vol. 106, no. A10, pp. 21 555–21 563, 2001.

[61] J. A. Farrell, *Aided Navigation: GPS with High Rate Sensors*. McGraw-Hill, Inc., 2008.

[62] P. Teunissen, R. Odolinski, and D. Odijk, "Instantaneous BeiDou + GPS RTK positioning with high cut-off elevation angles," *Journal of Geodesy*, vol. 88, pp. 335–350, 2014.

[63] P. J. Teunissen, "Least-squares estimation of the integer GPS ambiguities," in *Invited lecture, section IV theory and methodology, IAG general meeting, Beijing, China*, 1993, pp. 1–16.

[64] ——, "A new method for fast carrier phase ambiguity estimation," in *Proc. of the IEEE Position, Location and Navigation Symposium (PLANS)*. IEEE, 1994, pp. 562–573.

[65] V. Wegener and L. Wanninger, "Communication options for Network RTK/SAPOS realization," in *Joint 2nd Workshop on Positioning, Navigation and Communication 2005 (WPNC) & 1st Ultra-Wideband Expert Talk 2005 (UET)*, 2005.

[66] X. Chen, T. Allison, W. Cao, K. Ferguson, S. Grunig, V. Gomez, A. Kipka, J. Kohler, H. Landau, R. Leandro *et al.*, "Trimble RTX, an innovative new approach for network RTK," in *Proc. of the 24th international Technical Meeting of the Satellite Division of the institute of navigation (ION GNSS)*, 2011, pp. 2214–2219.

[67] U. Vollath, A. Buecherl, H. Landau, C. Pagels, and B. Wagner, "Multi-base RTK positioning using virtual reference stations," in *Proc. of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS)*, 2000, pp. 123–131.

[68] H. Euler, C. Keenan, and B. Zebhause, "Study of a simplified approach in utilizing information from permanent reference station arrays," in *Proc. of the 14th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS)*, 2001, pp. 379–391.

[69] G. Wübbena, M. Schmitz, and A. Bagge, "PPP-RTK: Precise Point Positioning using state-space representation in RTK networks," in *Proceedings of the 18th ION international technical meeting of the satellite division (ION GNSS)*, 2005, pp. 2584–2594.

[70] W. Hu, A. Neupane, and J. A. Farrell, "Using PPP Information to Implement a Global Real-Time Virtual Network DGNSS Approach," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 10, pp. 10 337–10 349, 2022.

[71] P. Collins, J. Henton, Y. Mireault, P. Héroux, M. Schmidt, H. Dragert, and S. Bisnath, "Precise point positioning for real-time determination of co-seismic crustal motion," in *Proceedings of the 22nd International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS)*, 2009, pp. 2479–2488.

[72] G. Guerova, J. Jones, J. Douša, G. Dick, S. de Haan, E. Pottiaux, O. Bock, R. Pacione, G. Elgered, H. Vedel *et al.*, "Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe," *Atmospheric Measurement Techniques*, vol. 9, no. 11, pp. 5385–5406, 2016.

[73] RTCM SC-104 SSR Working Group and IGS Real-Time Working Group, "IGS State Space Representation (SSR) Format," 2020. [Online]. Available: https://files.igs.org/pub/data/format/igs_ssr_v1.pdf

[74] G. Wübbena, C. Perschke, J. Wübbena, T. Wübbena, and M. Schmitz, "GNSS SSR real-time correction: The Geo++ SSRZ format and applications," in *Proceedings of the 33rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+)*, 2020, pp. 55–98.

[75] Y. Gao and X. Shen, "A new method for carrier-phase-based precise point positioning," *Navigation*, vol. 49, no. 2, pp. 109–116, 2002.

[76] M. Hernández-Pajares, J. Juan, J. Sanz, and R. Orús, "Second-order ionospheric term in GPS: Implementation and impact on geodetic estimates," *Journal of geophysical research: Solid Earth*, vol. 112, no. B8, 2007.

[77] S. Bisnath and Y. Gao, "Current state of precise point positioning and future prospects and limitations," in *Observing Our Changing Earth*. Springer, 2009, pp. 615–623.

[78] J. Kouba and P. Héroux, "Precise point positioning using IGS orbit and clock products," *GPS Solutions*, vol. 5, no. 2, pp. 12–28, 2001.

[79] Y. Xiang, Y. Gao, and Y. Li, "Reducing convergence time of precise point positioning with ionospheric constraints and receiver differential code bias modeling," *Journal of Geodesy*, vol. 94, pp. 1–13, 2020.

[80] N. Naciri and S. Bisnath, "RTK-Quality Positioning With Global Precise Point Positioning Corrections," *Navigation*, vol. 70, no. 3, 2023.

[81] D. Odijk, "Fast precise GPS positioning in the presence of ionospheric delays," Ph.D. dissertation, 2004.

[82] X. Li, J. Huang, X. Li, H. Lyu, B. Wang, Y. Xiong, and W. Xie, "Multi-constellation GNSS PPP instantaneous ambiguity resolution with precise atmospheric corrections augmentation," *GPS Solutions*, vol. 25, no. 3, pp. 1–13, 2021. [Online]. Available: https://doi.org/10.1007/s10291-021-01123-0

[83] R. J. Blakely, *Spherical Harmonic Analysis*. Cambridge University Press, 1995, p. 100–127.

[84] M. Schmidt, D. Dettmering, M. Mößmer, Y. Wang, and J. Zhang, "Comparison of spherical harmonic and B spline models for the vertical total electron content," *Radio Science*, vol. 46, no. 06, pp. 1–8, 2011.

[85] S. Farzaneh and E. Forootan, "A least squares solution to regionalize VTEC estimates for positioning applications," *Remote Sensing*, vol. 12, no. 21, 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/21/3545

[86] I. Cherniak, A. Krankowski, and I. Zakharenkova, "ROTI Maps: a new IGS ionospheric product characterizing the ionospheric irregularities occurrence," *GPS Solutions*, vol. 22, pp. 1–12, 7 2018. [Online]. Available: https://link.springer.com/article/10.1007/s10291-018-0730-1

[87] Y. Jin and K. Oksavik, "GPS scintillations and losses of signal lock at high latitudes during the 2015 St. Patrick's Day storm," *Journal of Geophysical Research: Space Physics*, vol. 123, no. 9, pp. 7943–7957, 2018.

[88] C. De Boor, *A practical guide to splines*. Springer-Verlag, New York, 1978, vol. 27.

[89] M. Schmidt, D. Dettmering, and F. Seitz, "Using B-spline expansions for ionosphere modeling," in *Handbook of Geomathematics (Second Edition)*. Springer, 2015, pp. 939–983.

[90] A. Goss, M. Schmidt, E. Erdogan, B. Görres, and F. Seitz, "High-resolution vertical total electron content maps based on multi-scale B-spline representations," *Annales Geophysicae*, vol. 37, no. 4, pp. 699–717, aug 2019.

[91] L. L. Schumaker and C. Traas, "Fitting scattered data on spherelike surfaces using tensor products of trigonometric and polynomial splines," *Numerische Mathematik*, vol. 60, no. 1, pp. 133–144, 1991.

[92] C. De Boor, "Package for calculating with B-splines," *SIAM Journal on Numerical Analysis*, vol. 14, no. 3, pp. 441–472, 1977.

[93] T. Hadas, F. N. Teferle, K. Kazmierski, P. Hordyniec, and J. Bosy, "Optimum stochastic modeling for GNSS tropospheric delay estimation in real-time," *GPS solutions*, vol. 21, pp. 1069–1081, 2017.

[94] X. Luo, M. Mayer, B. Heck, and J. L. Awange, "A realistic and easy-to-implement weighting model for GPS phase observations," *IEEE Transactions on Geoscience and Remote sensing*, vol. 52, no. 10, pp. 6110–6118, 2014.

[95] S. H. Byun and Y. E. Bar-Sever, "A new type of troposphere zenith path delay product of the international GNSS service," *Journal of Geodesy*, vol. 83, pp. 1–7, 2009.

[96] N. Williams and M. Barth, "A qualitative analysis of vehicle positioning requirements for connected vehicle applications," *IEEE Intelligent Transportation Systems Magazine*, vol. 13, no. 1, pp. 225–242, 2020.

[97] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and Cooperation in Networked Multi-agent Systems," *Proc. of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[98] B. Bollobás, *Modern Graph Theory*, ser. Graduate Texts in Mathematics. Springer New York, 1998, vol. 184.

[99] G. Strang, *Linear algebra and its applications, Fourth Edition*. Brooks/Cole, 2012.

[100] R. Tron and R. Vidal, "Distributed computer vision algorithms," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 32–45, 2011.

[101] W. Ren, R. W. Beard, and E. M. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control Systems Magazine*, vol. 27, no. 2, pp. 71–82, 2007.

[102] W. Ren and R. W. Beard, *Distributed consensus in multi-vehicle cooperative control*. Springer-Verlag London, 2008, vol. 27, no. 2.

[103] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[104] A. Barrat and M. Weigt, "On the properties of small-world network models," *The European Physical Journal B: Condensed Matter and Complex Systems*, vol. 13, no. 3, pp. 547–560, 2000.

[105] A.-L. Barabási and E. Bonabeau, "Scale-free networks," *Scientific American*, vol. 288, no. 5, pp. 60–69, 2003.

[106] R. M. Eustice, H. Singh, and J. J. Leonard, "Exactly sparse delayed-state filters for view-based SLAM," *IEEE T. on Robotics*, vol. 22, no. 6, pp. 1100–1114, 2006.

[107] R. Olfati-Saber and R. M. Murray, "Consensus protocols for networks of dynamic agents," in *Proceedings of the 2003 American Control Conference*, vol. 2. IEEE, 2003, pp. 951–956.

[108] S. S. Kia, B. Van Scoy, J. Cortes, R. A. Freeman, K. M. Lynch, and S. Martinez, "Tutorial on dynamic average consensus: The problem, its applications, and the algorithms," *IEEE Control Systems Magazine*, vol. 39, no. 3, pp. 40–72, 2019.

[109] S. Banville, E. Hassen, P. Lamothe, J. Farinaccio, B. Donahue, Y. Mireault, M. A. Goudarzi, P. Collins, R. Ghoddousi-Fard, and O. Kamali, "Enabling ambiguity resolution in CSRS-PPP," *NAVIGATION: Journal of the Institute of Navigation*, vol. 68, no. 2, pp. 433–451, 2021.

[110] J. Geng and C. Shi, "Rapid initialization of real-time PPP by resolving undifferenced GPS and GLONASS ambiguities simultaneously," *Journal of Geodesy*, vol. 91, pp. 361–374, 2017.

[111] X. Li, J. Huang, X. Li, H. Lyu, B. Wang, Y. Xiong, and W. Xie, "Multi-constellation GNSS PPP instantaneous ambiguity resolution with precise atmospheric corrections augmentation," *GPS Solutions*, vol. 25, no. 3, p. 107, 2021.

[112] J.-B. Uwineza, F. Rahman, F. Silva, W. Hu, and J. A. Farrell, "Characterizing GNSS Multipath at Different Antenna Mounting Positions on Vehicles," 2019. [Online]. Available: https://escholarship.org/uc/item/9mj3459s

[113] Y. Pan, G. Kłopotek, L. Crocetti, R. Weinacker, T. Sturn, L. See, G. Dick, G. Möller, M. Rothacher, I. McCallum, V. Navarro, and B. Soja, "Determination of High-Precision Tropospheric Delays Using Crowdsourced Smartphone GNSS Data," pp. 1–22, 2024. [Online]. Available: https://egusphere.copernicus.org/preprints/2024/egusphere-2024-66/

Appendix A

Derivation of the Information Vector

As in the information matrix in eqn. (6.22), the information vector is also obtained by Schur complement, which leads to

$$\mathbf{j}_{\mathbf{a}_v} = \mathbf{J}_{\mathbf{a}_v}\mathbf{a}_v. \tag{A.1}$$

The above expression is proven below.

Proof. The Schur complement of $\mathbf{j}^+_{v_z}$ in $\mathbf{j}^+_v$ is obtained by using (6.18) and (6.20). Specifically,

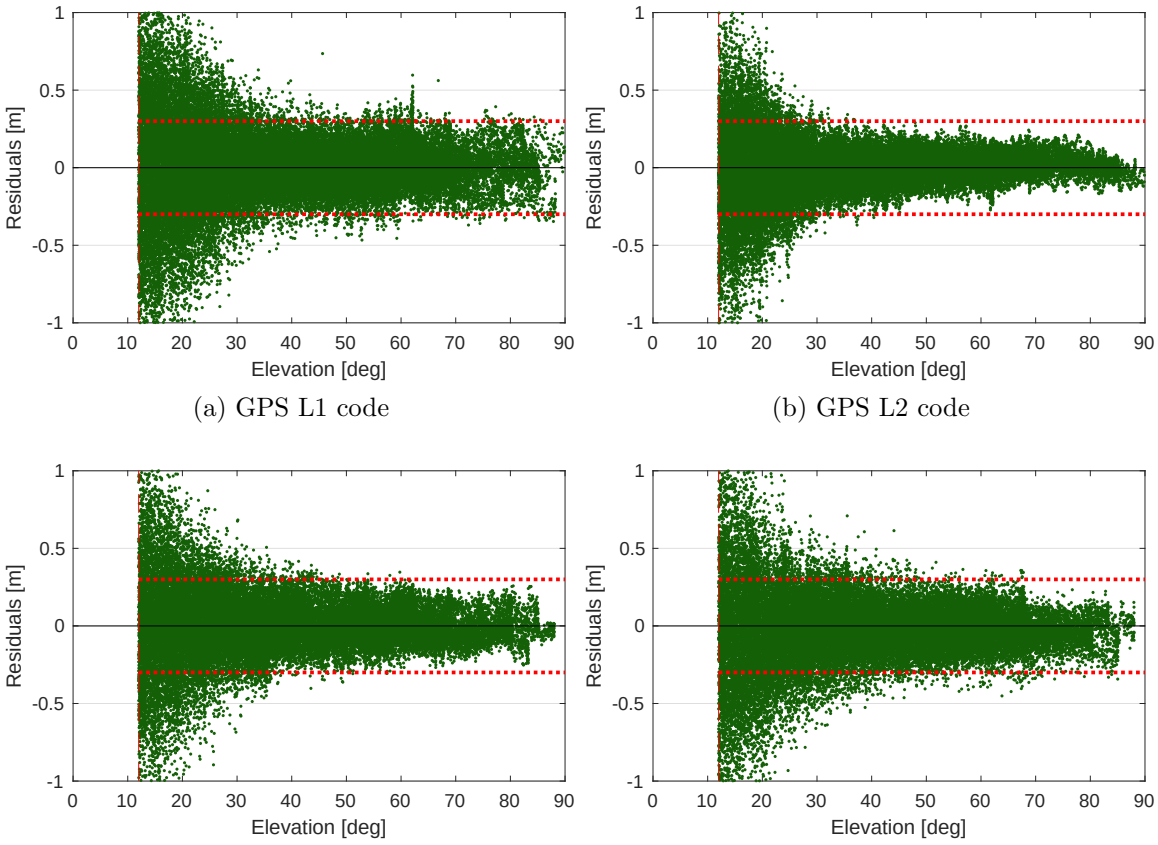
$$\begin{aligned}
\mathbf{j}_{\mathbf{a}_v} &= \mathbf{j}_{\mathbf{a}_v,\mathbf{z}_v} - \mathbf{B}^\top(\mathbf{A})^{-1}\mathbf{j}_{\mathbf{z}_v,\mathbf{a}_v} \\
&= \Big(\mathbf{B}^\top\mathbf{z}_v + \mathbf{C}\,\mathbf{a}_v\Big) - \mathbf{B}^\top\mathbf{A}^{-1}\left(\mathbf{A}\,\mathbf{z}_v + \mathbf{B}\,\mathbf{a}_v\right) \\
&= \mathbf{B}^\top\mathbf{z}_v + \mathbf{C}\,\mathbf{a}_v - \mathbf{B}^\top\mathbf{A}^{-1}\mathbf{A}\,\mathbf{z}_v - \mathbf{B}^\top\mathbf{A}^{-1}\mathbf{B}\,\mathbf{a}_v \\
&= \mathbf{B}^\top\mathbf{z}_v - \mathbf{B}^\top\mathbf{z}_v + \mathbf{C}\,\mathbf{a}_v - \mathbf{B}^\top\mathbf{A}^{-1}\mathbf{B}\,\mathbf{a}_v \\
&= \mathbf{C}\,\mathbf{a}_v - \mathbf{B}^\top\mathbf{A}^{-1}\mathbf{B}\,\mathbf{a}_v \\
&= \Big(\mathbf{C} - \mathbf{B}^\top\mathbf{A}^{-1}\mathbf{B}\Big)\,\mathbf{a}_v \\
&= \mathbf{J}_{\mathbf{a}_v}\mathbf{a}_v.
\end{aligned}$$

Thus, we have obtained the expression in (6.23). □

Appendix B

Code and Phase Measurement Residuals

This appendix shows plots of typical code and phase measurement residuals. The receiver used is SEPT POLARX5, equipped with a TRM59800.99 antenna. The measurement sampling interval is 15 seconds.



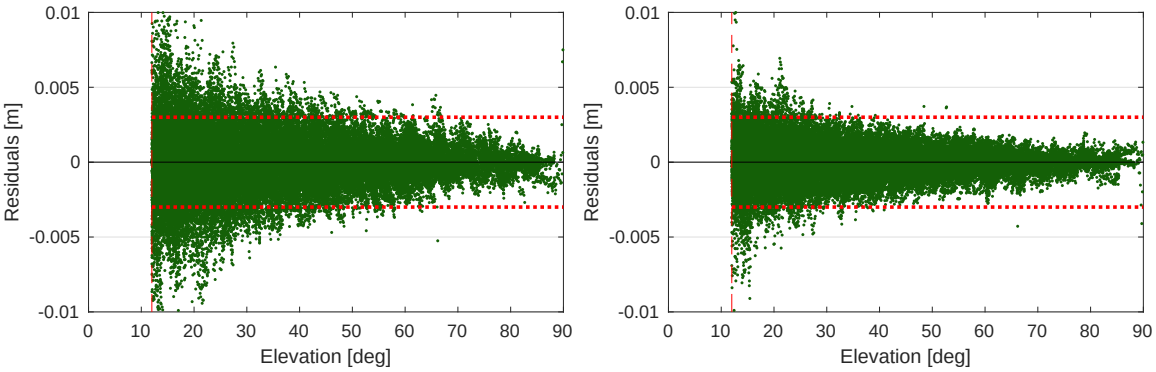
(a) GPS L1 code

(b) GPS L2 code

(c) GAL E1 code

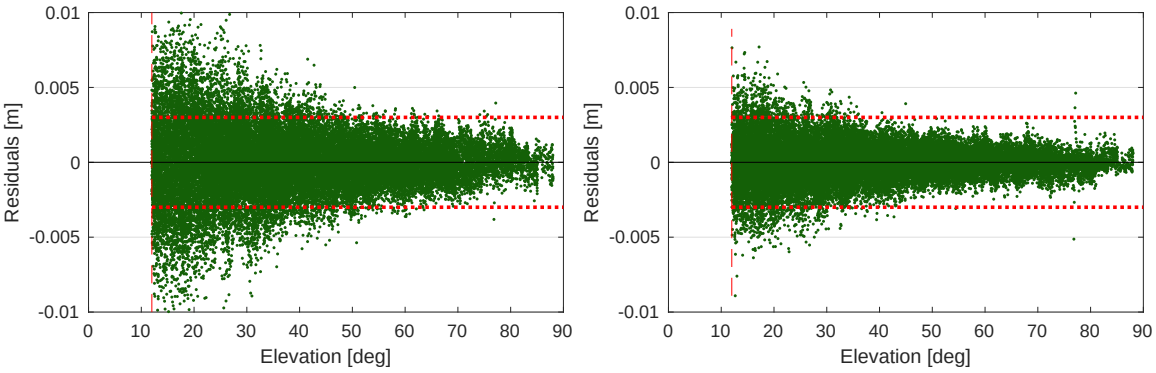
(d) GAL E5a code

Figure B.1: Pseudorange residuals as a function of satellite elevation for GPS and GAL measurements. The horizontal red lines shows the code standard deviation (0.03 m) used in the measurement covariance matrix $\mathbf{R}$. The vertical dashed line shows the applied elevation mask.



(a) GPS L1 phase

(b) GPS L2 phase



(c) GAL E1 phase

(d) GAL E5a phase

Figure B.2: Phase residuals as a function of satellite elevation for GPS and GAL measurements. The horizontal dotted red lines show the phase standard deviation (0.003 m) used in the measurement covariance matrix $\mathbf{R}$. The vertical dashed line shows the applied elevation mask.

Appendix C

Collaborating Reference Stations

Table C.1 shows details of twenty (20) CORS reference stations used throughout this text. Note that to facilitate fair comparison, all of the stations use the same receiver model and antennae from the same manufacturer.

Fig. C.1 is a geographic map showing the location of the reference stations.

Abbr.	Location	Height [m]	Receiver[a]	IGS Antenna Name[b]
AR27	Camp Pendleton, CA	68.000	POLARX5	TRM59800.00
AZU1	Azusa, CA	144.752	POLARX5	TRM59800.00
BVPP	Taft, CA	123.058	POLARX5	TRM59800.99 w/ TW605[c]
ECFS	Al Cariso, CA	801.420	POLARX5	TRM59800.99 w/ TW605
MAT2	Riverside, CA	398.321	POLARX5	TRM59800.80
MLFP	Riverside, CA	472.981	POLARX5	TRM59800.99 w/ TW605
P470	Victorville, CA	991.384	POLARX5	TRM59800.99 w/ TW605
P471	San Juan Capistrano, CA	174.755	POLARX5	TRM59800.80
P517	Angeles Nat'l. Forest, CA	1,959.161	POLARX5	TRM59800.99 w/ TW605
P522	Taft, CA	991.087	POLARX5	TRM59800.00
P553	Lebec, CA	1,335.835	POLARX5	TRM59800.99 w/ TW605
P558	Tehachapi, CA	1,411.439	POLARX5	TRM59800.00
P567	Bakersfield, CA	706.313	POLARX5	TRM59800.80
P574	Upland, CA	2,873.768	POLARX5	TRM59800.99 w/ TW605
P582	El Mirage, CA	848.025	POLARX5	ASH701945B
P612	San Bernardino, CA	531.718	POLARX5	TRM59800.99 w/ TW605
PBPP	Pearblossom, CA	901.959	POLARX5	TRM59800.99 w/ TW605
RSTP	Rosamond, CA	712.738	POLARX5	TRM59800.99 w/ TW605
SKYB	Santa Clarita, CA	520.229	POLARX5	TRM59800.99 w/ TW605
THCP	Tehachapi, CA	1,652.403	POLARX5	TRM59800.99 w/ TW605

[a] All stations use the Septentrio PolaRx5, a multi-GNSS geodetic-grade reference receiver.
[b] All antennas have a Dorne Margolin antenna element with choke rings.
[c] Trimble choke ring antenna (TRM59800.99) with Tallysman Low Noise Amplifier (TW605).

Table C.1: Collaborative reference stations throughout Southern California.

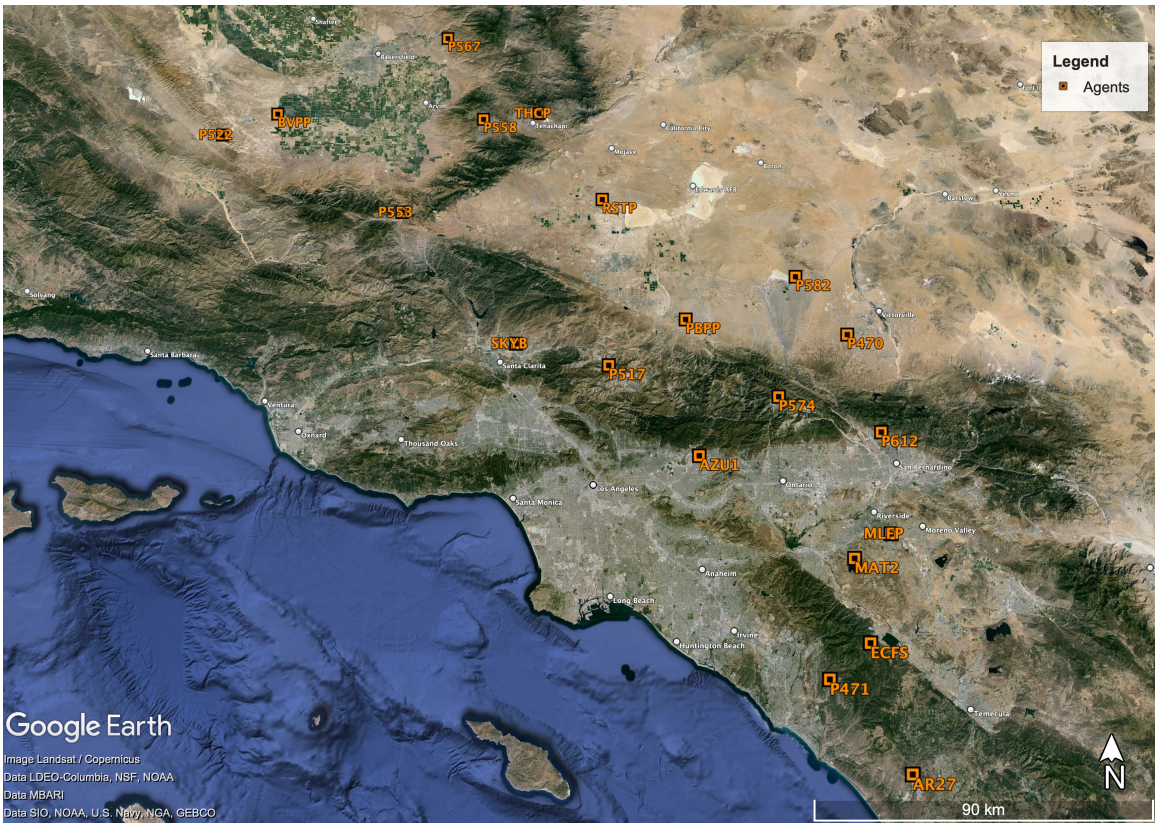


Figure C.1: Geographic map of reference stations throughout Southern California.