ORIGINAL ARCHIVAL COPY

RELATIVE RECEIVER AUTONOMOUS INTEGRITY MONITORING FOR FUTURE GNSS-BASED AIRCRAFT NAVIGATION

BY

LIVIO RAFAEL GRATTON

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mechanical and Aerospace Engineering
in the Graduate College of the
Illinois Institute of Technology

Approved _____
Adviser

Chicago, Illinois
May 2011

ACKNOWLEDGMENT

I would like to thank my advisor, Professor Boris Pervan for giving me the chance to be part of his research group, for his support, patience and open-mindedness throughout these years; and, particularly in the last six months, for the time and dedication devoted towards the completion of my dissertation, amounting to a personal effort more expected from a friend than from an academic advisor. His precise, practical and thorough way of thinking, along with his considerate way of transmitting his ideas, has made this an exciting knowledge-absorbing process. I would also like to thank my defense and dissertation committee, including Professors John Kallend, Sudhakar Nair, Xiaoping Qian and Matthew Spenko.

I want to thank all the past and current members of the IIT Navigation Laboratory, particularly Moon Heo, for his friendship and for introducing me to the basic principles or Relative RAIM, and Fang-Cheng Chan, Samer Khanafseh, Mathieu Joerger and Steven Langel for the innumerable hours of, sometimes heated but always intellectually challenging, technical discussions.

I want to thank my parents: my mother for her constant love and unbreakable faith in me, and my father for, among so many other things, instilling in me the desire for excellence in my work and the awareness of the beauty in the laws of physics.

I want to thank the people without whose support, material and otherwise, I would have not been able to accomplish anything: Federico, Lora, Jerónimo, Giampaolo and Ana María, Moche, Juan Pablo, Corinna, Vera, Tato, Maite, Tom Hoban, Marti, and Sophia.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYSMS

AL	Alert Limit
ARAIM	Absolute RAIM
BR	Broadcast Radius
C/A	Course Acquisition
CCD	Code Carrier Divergence
CCDM	CCD Monitor
CDF	Cumulative Distribution Function
CDMA	Code-Division Multiple Access
CS	Control Segment
CT	Coasting Time
DH	Decision Height
DOF	Degrees Of Freedom
EGNOS	European Geostationary Overly Service
ERAIM	Extended RAIM
FAA	Federal Aviation Administration
FC	Failure during Coasting
FDMS	Frequency-Division Multiple Access
FF	Fault Free
FMS	Failure Mode Slope
GBAS	Ground Based Augmentation System
GEAS	GNSS Evolutionary Architecture Study
GEO	GEosynchronous Orbit

GF	Ground Facility
GGM	Ground Gradient Monitor
GIC	GNSS Integrity Chanel
GIVE	Grid Ionospheric Vertical Error
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GSM	Geometry Screening Monitor
HMI	Hazardously Misleading Information
ICME	Interplanetary Coronal Mass Ejection
IF	Ionospheric delay-Free
IIT	Illinois Institute of Technology
ILS	Instrument Landing System
IP	Ionospheric pierce Point
IS	Ionospheric Storm
LAAS	Local Area Augmentation System
LPV	Localized Performance with Vertical guidance
LSR	Least Squares Residual
MCS	Master Control Station
MEO	Medium Earth Orbit
NAVLAB	Navigation Laboratory
NC	Normal Coasting
P/Y	Precision/encrYpted
PCM	Position Comparison Method

PD	Position Domain
PDF	Probability Density Function
PH	Potentially Hazardous
PL	Protection Level
PPS	Precise Positioning System
PS	Parity Space
RAIM	Receiver Autonomous Integrity Monitoring
RCM	Range Comparison Method
RD	Range Domain
RRAIM	Relative RAIM
SA	Selective Availability
SAS	Standard Accuracy Service
SBAS	Space Based Augmentation System
SIF	Static Ionospheric Front
SPS	Standard Positioning System
SS	Solution Separation
SV	Space Vehicle
TEC	Total Electron Content
TTA	Time To Alert
UDRE	User Differential Range Error
WAAS	Wide Area Augmentation System

ABSTRACT

The Global Positioning System (GPS) has enabled reliable, safe, and practical aircraft positioning for en-route and non-precision phases of flight for more than a decade. Intense research is currently devoted to extending the use of Global Navigation Satellite Systems (GNSS), including GPS, to precision approach and landing operations.

In this context, this work is focused on the development, analysis, and verification of the concept of Relative Receiver Autonomous Integrity Monitoring (RRAIM) and its potential applications to precision approach navigation. RRAIM fault detection algorithms are developed, and associated mathematical bounds on position error are derived. These are investigated as possible solutions to some current key challenges in precision approach navigation, discussed below.

Augmentation systems serving large areas (like the Wide Area Augmentation System (WAAS) covering the North American continent) allow certain precision approach operations within the covered region. More and better satellites, with dual frequency capabilities, are expected to be in orbit in the mid-term future, which will potentially allow WAAS-like capabilities worldwide with a sparse ground station network. Two of the main challenges in achieving this goal are (1) ensuring that navigation fault detection functions are fast enough to alert worldwide users of hazardously misleading information, and (2) minimizing situations in which navigation is unavailable because the user's local satellite geometry is insufficient for safe position estimation.

Local augmentation systems (to be implemented at individual airports, like the Local Area Augmentation System or LAAS) have the potential to allow precision

approach and landing operations by providing precise corrections to user-satellite range measurements. An exception to these capabilities arises during ionospheric storms (caused by solar activity), when hazardous situations can exist with residual range errors several orders of magnitudes higher than nominal. Until dual frequency civil GPS signals are available, the ability to provide integrity during ionospheric storms, without excessive loss of availability, will be a major challenge.

For all users, with or without augmentation, some situations cause short duration losses of satellites in view. Two examples are aircraft banking during turns and ionospheric scintillation. The loss of range signals can translate into gaps in good satellite geometry, and the resulting challenge is to ensure navigation continuity by bridging these gaps, while simultaneously maintaining high integrity.

It is shown that the RRAIM methods developed in this research can be applied to mitigate each of these obstacles to safe and reliable precision aircraft navigation.

CHAPTER 1

INTRODUCTION

The Global Positioning System (GPS) has enabled reliable, safe and practical aircraft positioning for the en-route and non-precision phases of flight for more than a decade. Intense research work has been devoted to extending the use of satellite based navigation to precision approach and landing operations, much of it through some type of augmentation.

Augmentation systems serving large areas (like the Wide Area Augmentation System (WAAS) covering the North American continent) allow certain precision approach operations within the covered region. More and better satellites with dual frequency capabilities are expected to be in orbit in the mid-term future, which will potentially allow WAAS-like capabilities worldwide with a sparse ground station network. Two of the main challenges in achieving this goal are:

- Ensuring that the navigation system is fast enough to alert users of misleading information, and;
- Minimizing situations in which the system is unavailable for certain users because the satellite geometry does not support requirements for safe position estimation.

Local augmentation systems (for each airport, like the Local Area Augmentation System or LAAS) have the potential to allow precision approach and landing operations by providing corrections to user-satellite range measurements. An exception to these capabilities arises during ionospheric storms caused by solar activity; when hazardous

situations can exist with residual range errors being several orders of magnitudes bigger than the nominal. Until dual frequency civil GPS signals are available, this introduces another challenging problem:

- Providing integrity during ionospheric storms, without excessive loss of availability.

For all users, with or without augmentation, some situations cause short duration loss of satellites in view. Two examples are aircraft banking during turns and ionospheric scintillation. This type of occurrence requires:

- Coasting through gaps of poor satellite geometry to ensure navigation continuity.

Solutions to the four fundamental issues listed in the bulleted paragraphs are provided in this work by proposing, analyzing, developing and verifying several novel implementations based on carrier phase time-differential Relative Receiver Autonomous Integrity Monitoring (RRAIM) applications.

This chapter should give the reader enough background to understand the meaning and relevance of the contributions of the dissertation. Many concepts outlined in Chapter 1 are repeated with more detail and mathematical derivations in subsequent chapters.

1.1 Global Navigation Satellite Systems Background

1.1.1 History of GNSS. Navigation using radio signals began in World War II with systems based on ground transmitting antennas, like the British Gee and the American Loran, used primarily for 2D positioning of ships. In the 1970s Omega (operated by the

US and 6 international partners) provided continuous worldwide coverage for 2D positioning with an accuracy of 2 to 4 kilometers. Both Loran and Omega had differential implementations, in which ground stations would send corrections to nearby users, greatly improving accuracy. When the Sputnik 1 satellite started transmitting a signal in 1957, its orbit could be deduced from one ground station (with the station's position known) by analyzing the pattern of the signal's Doppler shift. It soon became obvious that if satellite orbits where known instead, then the user position on earth could hypothetically be determined from satellite signals. The US Navy quickly developed a system commonly known as Transit. Transit had 4 to 7 nearly polar low orbit satellites, which allowed stationary or slow moving users to get a 2D position fix after several minutes of data processing, with an expected precision of approximately 25 m. Its main use was for the Navy's submarines to reset their inertial navigation equipment. A stationary user could get a 3D position fix with 5 m accuracy after several days of accumulating measurements. Transit transmitted at two different frequencies, which allowed elimination of the ionospheric delay, a concept currently very useful for Global Navigation Satellite Systems (GNSS)-based navigation. The USSR developed its own system called Tsikada, which, in contrast to Transit decommissioned in 1996, is still operating [Mis01].

The success of Transit triggered the development of several systems. In the case of the US, the efforts were finally merged into GPS. GPS (described in detail in Chapter 2) nominally has 24 mid-orbit satellites or Space Vehicles (SVs). It became fully operational in 1997, and since then any user on the Earth's surface who has at least 4 satellites in view can estimate his/her position at all times. The current rms position

accuracy for a standalone user is roughly 5 m horizontal, and 7.5 m vertical on average [Mis01]. In 1996 the USSR had deployed a similar 24 SV constellation named GLONASS. It decayed after the breakup of the Soviet Union, but several satellite launches from the Russian Federation have reinvigorated the system with 22 SVs already in orbit [IAC11]. The European Union is developing its own GNSS system called Galileo with several prototype SVs already in orbit and is planned to be operational in 2013. All of the satellite systems described thus far are passive ranging systems. That is, they do not require any transmission of information from the user to the SV. China is also developing its own navigation system named Compass which has 5 SVs already in orbit. The Chinese government claims it will be operational in 2020 when its 35 satellites will have been launched. Contrary to the systems mentioned so far, Compass user receivers have transponders that communicate with the system's SVs [Heg08].

1.1.2 GNSS Based Navigation. GNSS-based positioning is accomplished using a technique called trilateration. It determines the user's position by finding the point in space where three imaginary spheres intersect. Each sphere is centered at the location of its respective satellite, and has a radius equal to the corresponding satellite-user measured distance. The user can compute the SV position at each moment with orbital ephemeris parameters broadcast from each satellite. In general, the clocks at the receivers are not as precise as the atomic clocks in the satellites. For this reason four satellites are needed to estimate four unknowns: three position coordinates, and the bias between GPS and user time. Each satellite broadcasts a ranging signal, and the SV-user distance is computed by

subtracting the time of reception of the signal from the time of emission at the SV, and multiplying it by the speed of light.

The work in this dissertation is related to aviation navigation systems, in which safety of life is the main concern. When an aircraft is performing an operation (for example a precision landing), many requirements have to be met to allow the operation to be executed safely. These can be separated into four basic categories:

Accuracy. The error in the position estimate has to be below a certain value a significant portion of the time (for example 95% of the time).

Continuity. The risk of aborting an operation once it is initiated has to be small.

Integrity. The risk of the position estimate error being bigger than a tolerable bound for the operation without the user being alerted of this situation has to be minimal.

Availability. The system is available when the accuracy, continuity and integrity requirements are met before the operation is initiated.

In general, when studying the feasibility of an implementation, or how good it is compared to other implementations, performance is evaluated by the percentage of time the system is available.

1.1.3 GNSS Performance. Of the basic GNSS in existence, only the GPS constellation, run by the US government, is fully operational. It has a service available free of charge to civil users worldwide, specified to provide horizontal and vertical positioning accuracy of 13 and 22 m respectively 95% of the time [Heg08]. The dominant error source is the signal delay as it travels through the ionosphere. The actual performance of the system is much better than the specification mentioned above. When

Galileo becomes fully operational (planned for 2013) it is claimed by its designers that it will provide a service equal to or better than GPS. Currently GLONASS provides positioning inside the Russian Federation territory and in a few years should be able to do so worldwide. These services and error margins refer to civil applications available to any user.

1.1.4 Future Development of GNSS. The most modern GPS SVs (and all space vehicles under development), as well as all Galileo satellites, will broadcast signals to civil users at two frequencies. This allows the user to eliminate the ionospheric delay by combining the measurements from the two different frequencies. The standalone civil user will then be able to achieve a positioning accuracy that before was only available with augmentation systems (discussed in section 1.3). Other planned improvements are increased robustness in signal tracking, better monitoring of satellite clock and ephemeris errors, and communication between satellites, allowing real time uploads from the ground to any satellite.

All of these improvements, added to the much larger number of satellites visible in the sky (from different constellations) open the door for many opportunities for safer and more precise navigation. To plan ahead, analyze the feasibility of different implementations, and be ready to take advantage of these future improvements when they become operational, the US Federal Aviation Administration (FAA) initiated the GNSS Evolutionary Architecture Study (GEAS). A considerable part of the work in this dissertation is in contribution to the GEAS effort.

1.2 Measurements and Error Sources

There are two basic types of measurements of the user-to-satellite distance: the code phase and the carrier phase measurements. They can be combined in different ways to reduce or eliminate errors.

The code signal is a unique message that is modulated on top of a carrier signal that is broadcast from each satellite. The receiver knows the unique code signal from each satellite, and reproduces the code signals of all satellites in view. By seeing how much phase difference there is between the broadcast and reproduced messages, the receiver can compute the time between broadcast and reception. The corresponding user-to-satellite range is obtained by multiplying this time difference by the speed of light. The carrier phase measurement tracks the phase difference between the satellite's broadcast carrier signal and a reproduction generated in the receiver.

The code for each SV has a unique sequence, making it unambiguous. The carrier however is a continuous sine wave, and when computing the phase shift between the SV signal and the one reproduced at the receiver, the number of whole cycles between the two is unknown. This integer is commonly referred to as the cycle ambiguity, and its determination is the major obstacle in using the precise carrier phase measurement.

These measurements differ from the actual satellite-receiver distance because of various nominal errors or because of faults. Mathematical models that bound the nominal errors can be achieved by analyzing data. These models typically take the form of a statistical distribution, usually a bounding Gaussian or Chi Square. In contrast, measurement errors caused by failures are very difficult to model (for lack of sufficient faulted data, or knowledge of their causes), and monitors are implemented to detect them.

For measurement faults, the frequency, or probability of occurrence is modeled. Following is a brief description of the possible error sources. More details are provided in chapters 2-5.

Satellite Errors: caused by errors in the SV clock (deviations from the GPS time), or errors in the broadcast ephemeris parameters (that do not accurately describe the actual satellite orbit). The nominal bounds on these errors are well known, and are taken into account by the user when computing the bounds on its position estimation error. The ground segment of the GNSS is constantly monitoring these two elements for potential failures, but sometimes the user might not be notified fast enough to meet the required Time To Alert (TTA) for a certain operation. Satellite failures are considered rare (three or less per year for GPS) [DoD08], but for some operations with stringent integrity risk requirements, additional monitoring is required at the user level.

Another potential satellite failure is code-carrier divergence, in which the code signal and its carrier become unsynchronized. This failure has some of the same characteristics as an ionospheric delay (introduced in the next paragraph), and so monitors built to detect it can also help detect ionospheric related events (details of this in chapters 2 and 3).

Atmospheric Errors: caused by the delay suffered by the signal as it goes through the atmosphere. The delay occurs mainly in the ionospheric and the tropospheric regions. The tropospheric delay error can be reduced by models based on the time of day, humidity, and location. Residual errors can be bounded with confidence. In contrast, the ionosphere is highly unpredictable, making this error source dominant for standalone users. However, it is a dispersive medium (the magnitude of the delay depends on the

frequency of the signal), which allows the elimination of the ionosphere-caused errors if the system has two different synchronized signals with different frequencies. Two facts that will become relevant later on are 1) that once the ionospheric delay is eliminated, other error sources become dominant; and 2) that the ionosphere delays the code phase but advances the carrier's by the same magnitude.

Receiver-end errors: when the signal reaches the receiver, it might be affected by reflections of itself on nearby objects, an effect called multipath. It will also have errors caused by the receiver noise as it travels from the antenna to the processing unit. Because of the way each signal is tracked, these errors are significantly lower for the carrier measurement than for the code measurement.

1.3 Aviation Augmentation Systems

Augmentation systems enable better accuracy and hazard mitigation by providing users with corrections and bounds on errors and monitors to diminish the probability of unknown failures. There are two basic types of aviation augmentation systems, those that provide information to nearby users (for example in the vicinity of an airport), and those that provide information to users over a wide area (for example the North American continent, or Europe). For historical reasons, the denomination of each one is related to the way the information is transmitted rather than the extension of their coverage. Thus the former are called Ground Based Augmentation Systems (GBAS) and the latter Space Based Augmentation Systems (SBAS).

1.3.1 Ground Based Augmentation Systems. Each GBAS will have a Ground Facility (GF) at an airport, with a group of receiving antennas, a computation center and a broadcasting VHF antenna. One main advantage over the predominant landing systems on major airports is that one set of equipment can service all runways, as opposed to two per runway (as Instrument Landing Systems or ILS requires).

The location of a static receiver antenna can be precisely known, allowing the ground facility to estimate errors in the range measurements, and transmit corrections to nearby users to account for them. The correction is generated as the difference between the expected range (computed) and the measured range. In general, satellite and atmospheric errors will only have a small residual error after the corrections are applied by the user, and this error is proportional to the distance between the ground antenna and the aircraft. Receiver-end errors can't be reduced by augmentation.

The FAA's version of GBAS is called the Local Area Augmentation System, and it is in an advanced prototype stage. It is targeted to support Category I precision approaches and in the future Category III. However under ionospheric storm conditions the integrity requirements for these operations are more difficult to meet, which could translate into some loss of availability. This problem will cease when dual frequency is available, but unfortunately this will not happen in the near term, as will be explained in Chapter 2. It is highly desirable then to have a transition solution to the ionospheric storm scenario.

1.3.2 Spaced Based Augmentation Systems. SBAS consists of a network of ground stations (spread over a desired coverage area) that collect measurements from a core

constellation (for example GPS), a processing center, an upload antenna, and a group of GEosynchronous Orbit satellites (GEOs). The GEOs receive corrections and error bounds for the core constellation satellite signals via the upload antenna, and rebroadcast them to users over a wide area.

Since ranges to many different ground stations are processed, the corrections are very effective for errors that are not dependent on specific user location, like satellite errors. Under nominal conditions a relatively accurate description of the ionospheric delay can be mapped for the covered area. During ionospheric storms, the storm will most likely be detected, but the system will not always be able to provide useful corrections for the ionospheric delay.

The FAA's SBAS implementation is called the Wide Area Augmentation System, and it has been operational since 2003. Its core constellation is GPS, and it supports a precision approach operation called LPV-200, which is similar (but with less stringent requirements, see Appendix D) than Category I.

The European Union has a similar system called EGNOS, operational since 2009. It can be used over Europe with GPS as the core constellation until Galileo is ready. China's Compass includes an SBAS segment, and India and Japan are also developing SBAS systems that will augment the existing core constellations of their respective countries. Australia is developing a system that is a mixture of GBAS and SBAS, as the corrections come from a widely spaced network of antennas, but they are transmitted to the user from local antennas instead of a GEO [Heg08].

1.4 Receiver Autonomous Integrity Monitoring (RAIM)

All methods of monitoring threats are based on checking the consistency of redundant information. For standalone users, this means having more ranging sources than states.

The methods briefly described in this section (and in more detail in subsequent chapters) can be used to detect failures without considering the cause, or to detect a specific threat, taking advantage of the knowledge of the behavior of the potential failure source. The work in this dissertation has applications of both types.

1.4.1 RAIM Concept. Receiver Autonomous Integrity Monitoring (RAIM), is so named because the airborne receiver performs self-contained fault detection. In RAIM each GPS measurement is compared to the consensus of all available GPS measurements. In this way, RAIM detects the presence of a faulty satellite within the current set of in-view satellites. In some circumstances, RAIM can also reliably isolate which satellite is faulty or inconsistent with the other satellites in-view. The failure detection and/or isolation can be done with different implementations that will be described in detail in Chapter 2, and expanded into novel concepts and algorithms in chapters 3 through 5.

1.4.2 History of RAIM. RAIM was initially developed thinking mainly of a standalone user basing its position estimation on code measurements (or carrier smoothed code, a derived measurement explained in Chapter 2), during the en-route and non-precision approach phases of flight. It was also generally assumed that the likelihood of simultaneous failures (i.e. on more than one satellite) was negligible.

The concept of RAIM was initially introduced by Lee in 1986 [Lee86] who proposed two methods: the Range Comparison Method (RCM) and the Position Comparison Method (PCM). In 1988, Parkinson and Axelrad showed that the Least-Squares Residual (LSR) could be used as a test statistic, and that it is Chi Square distributed with n-4 degrees of freedom (n being the number of available satellites) [Par88]. In 1998 Brown matured the details for the computation of the monitor's threshold and introduced the concept of the horizontal protection level [Bro98].

In 1988 Sturza [Stu88] introduced the Parity Space (PS) method, further developed by others later, including van Graas and Farrell in 1993 [vGr93] and Kelly in 1998 [Kel98].

These implementations (LSR and PS) relate the possible impact of failures on each satellite on the position estimate error and on the monitor residual, allowing one to quantify the probability of an undetected hazard. Both methods are of the same nature, and achieve identical results, but provide different insight when extending the purpose of the monitor to include isolation rather than detection alone.

In 1988, Brown and McBurney introduced the Solution Separation (SS) algorithm [Bro88], which compares the position solution using all available ranges with solutions eliminating one satellite at a time. As opposed to the LSR and PS implementations described above, this method realizes the detection in the position domain. In 1995 Brenner developed a method using SS to integrate GPS and inertial information through the use of a bank of Kalman filters generating bounds on the position estimation errors [Bre96]. In 2006, Hwang and Brown introduced the concept of Novel Integrity-

Optimized RAIM, which optimizes the LSR algorithm based on integrity rather than accuracy.

1.4.3 Applications and Limitations of RAIM. In 1992, the FAA included RAIM as a required capability of avionics for all non-precision approach phases of flight [RTC06], thus making RAIM a very practical research topic. However, the quality of code phase measurements and the number of satellites currently available do not allow for the extension of existing RAIM techniques to more demanding aviation applications.

RAIM is currently used to provide supplemental navigation in the en route and terminal area phases of flight, and is also used to support lateral guidance during the approach phase of flight for certain specific type of approaches. At present, RAIM cannot support vertical navigation [GEA10]. The extension of RAIM applications to other phases of flight is highly desirable, since the user can ensure safety autonomously with minimal or no augmentation.

1.5 Relative RAIM (RRAIM)

As is obvious from the limitations of RAIM implementations stated in subsection 1.4.3, there are two main areas for improvement to expand RAIM to phases of flight with more stringent requirements: signal quality and number of ranging sources. The current (and planned) improvement and expansion in satellite constellations provides more ranging sources, and in most cases better signal quality. This has triggered a new wave of research in RAIM applications after some years of waning interest by the navigation community.

One way to reduce dramatically the measurement error magnitudes is to use the carrier phase measurement rather than the code. To do this the uncertainty of the cycle ambiguity has to be dealt with. The main idea is to use time differential measurements, eliminating the cycle ambiguity in the process. However, this introduces new complications that have to be considered and addressed properly.

1.5.1 Concept, Potential and Challenges of RRAIM. RRAIM is a RAIM application that uses differential measurements. The measurements are differenced in time, which, as the aircraft is moving, also translates to a difference (or baseline) in space. The vector of differences in measurements is processed to provide risk mitigation for the period between the current and the initial measurements. By taking time differential measurements, many errors are eliminated-most importantly the carrier phase cycle ambiguity. Another major advantage is that the remaining errors can be modeled more precisely. For example, during an ionospheric storm, the magnitude of the delay is unknown, but changes in delay over shorter distances can be modeled with greater confidence.

The reduction of the dominant satellite and atmospheric measurement errors through time differencing causes receiver-end errors to play a much more dominant role. As mentioned previously, carrier phase measurements have much smaller receiver-end errors than code measurements, creating enormous potential to reduce the overall measurement error level. Carrier phase measurements have been used for decades for geodetic applications. However, for navigation, the user position is needed in real time, making its use more challenging. RRAIM applications introduce new ways of combining

code and carrier measurements, which requires a detailed analysis of the correlation between all the elements of information used for position estimation and hazardous error monitoring.

1.5.2 Prior work. In 1994 Pervan, Cohen and Parkinson introduced the idea of using the precise carrier phase measurements for RAIM [Per94][Per96a]. In that implementation additional measurements are provided by pseudolites (antennas at the ground facility transmitting GPS-like signals). Since these ranging signals are near the user, the corresponding geometry changes quickly, allowing for the estimation of carrier phase cycle ambiguities in real time.

In 2004, Heo, Pervan et al. introduced a method using carrier phase RAIM to detect orbit estimation errors in the GPS broadcast ephemeris [Heo04]. In this case the cycle ambiguity is eliminated rather than estimated, by using time differential measurements from the baseline formed by the aircraft itself as it moves during its approach. Thus, this is the first RRAIM application per se.

The RRAIM concept was subsequently much more fully developed at the Illinois Institute of Technology (IIT) NAVigation LABoratory (NAVLAB) as part of the GEAS research activity. In January 2007, the concept and algorithms of a Range Domain (RD) RRAIM implementation were presented [Per07]. In January 2008 the concept and algorithms of a Position Domain (PD) RRAIM implementation were introduced. Also in January 2008 the dynamic allocation of the integrity risk between the failure hypothesis and the fault free hypothesis was introduced [Per08]. These concepts are discussed in this dissertation. The PD implementation concept was also explored by Lee later in 2008, but

using SS instead of a LSR, with similar results [Lee08]. In 2007 Lee introduced an optimized method that reallocates the integrity budget between satellites to improve availability [Lee07].

In 2006 Angus generalized the LRS algorithm from one failure to multiple failures [Ang06]. In 2007, Blanch, Ene, et al. developed a multiple hypothesis RAIM algorithm for vertical guidance [Ene07] [Bla07], which are particularly suitable for managing the risk of failures in more than one SV using the SS implementation. It develops a dynamic allocation technique to minimize the protection levels, and also shows a preliminary comparison between the SS and LSR methodologies. Even though the algorithms for simultaneous satellite failures are introduced in the work mentioned in this paragraph, the great majority of research in the area currently relies on very small prior probability of failures per satellite to generate the position bounds. In practice this is often equivalent to assuming single satellite failure. In the future, with multiple constellations in the sky and a great increase in the number of SVs consideration of multiple SV failures might become unavoidable.

1.6 Contributions

- Two new methods are introduced for integrity monitoring during aircraft precision approaches:
    - The Relative RAIM method, with two different implementations: Range Domain and Position Domain; and
    - The Extended RAIM method.

These implementations use carrier phase RAIM and support worldwide SBAS.

- Formulas are provided that bound the user position errors within the allowed integrity risk. The necessary measurements and their respective covariances are precisely defined.

- New algorithms are developed to implement the proposed monitors. These include PD and RD applications, with LS and SS residuals, using GBAS augmentation or traditional RAIM for a reference position. Conservative assumptions are used when necessary to reduce computation time.

- The feasibility of the proposed implementations is proven through sensitivity analysis. Performance with respect to constellation and available augmentation systems is studied, which can be easily translated into requirements for the future ground and space segment installations and functions.

- A novel monitor is introduced which mitigates the static ionospheric storm threat for GBAS. It uses carrier phase relative RAIM. The computational burden is reduced through theoretical analysis and practical assumptions.

- The feasibility of the proposed implementation is proven through simulation covering the entire ionospheric front threat space for selected cities.

CHAPTER 2

GNSS-BASED NAVIGATION AND MONIORING SYSTEMS

In this chapter, the principles of GNSS-based navigation and the status of the different worldwide implementations are explained. Many of the concepts introduced in Chapter 1 are developed here in more detail. The chapter also includes a brief description of relevant aircraft navigation operations, as well as the principles by which the safety of life of these operations is ensured. Within this context, the reader will get an idea of what problems are considered to be already solved, and which are still a challenge; and the role of RAIM, and more specifically RRAIM, in ensuring safety. The primary purpose of this chapter is to provide background and foundational material for the following chapters.

2.1 Global Navigation Satellite Systems

A GNSS can be divided into three segments: 1) A space segment, consisting of a constellation of satellites as ranging sources and transmitting information necessary for positioning. This information includes time and satellite-position estimation parameters as well as certain error bounds. 2) A ground segment that controls the satellites and the information transmitted to the users by the SVs. 3) A segment comprised of all the users of the system. In this section current GNSS constellations and those under development are described. GPS is described more thoroughly in a separate subsection as it is the one used in most of the work, and the only constellation that is currently fully operational. Augmentation systems like WAAS or EGNOS (that can be considered part of a GNSS), are described in the specific augmentation systems sections 2.3 and 2.4.

2.1.1 Global Positioning System. GPS is operated by the US Air Force. The first satellite was launched in 1978, and the system achieved initial operating capability in 1993 and full operational capability in 1995 [Mis01].

The nominal GPS constellation has 24 satellites at a 20000 km altitude (Medium Earth Orbit or MEO), with an inclination of 55 degrees. The satellites are distributed in 6 orbital planes with four satellites in each plane. They are unevenly spaced in such a way as to mitigate the impact on the users position fixes in case of any satellite failing to broadcast a signal. This constellation provides a minimum of 4 satellites in view at all times for any point on the earth's surface with an unblocked view of the sky.

So far 66 SVs have been launched, with 30 being operational at this moment [Heg08][USC10]. Different GPS SV models (Block I, Block II, Block IIA, Block IIR, Block IIR-M and Block IIF) manufactured by different companies (Rockwell International, Lockheed Martin, Boeing) have consistently improved the service provided by GPS with each SV generation, generally surpassing the specified requirements and predicted lifespan by large margins.

GPS provides two services: the Standard Positioning Service (SPS) for civilian users, and the Precise Positioning Service (PPS), restricted to the military forces of the US and its allies. The US government has committed to providing the SPS service worldwide free of charge since 1994. Furthermore, the US has guaranteed a 6 year warning in case the provision of service was terminated [Heg08]. Initially the accuracy of the SPS service was intentionally degraded by a pseudorandom dithering of the SV clock that could only be removed by PPS receivers equipped with the generating algorithm and cryptographic key (this process was called selective availability or SA). In May 2000, SA

was ceased, and in 2007 the US announced that the capability to activate SA would not be procured in the design of future satellites [Heg08].

GPS satellites use common frequencies to broadcast their codes using a technique known as Code-Division Multiple Access (CDMA). All GPS satellites broadcast navigation signals in two frequencies: L1 (1575.42 MHz) and L2 (1227.6 MHz). L1 transmits a Course/Acquisition (C/A) code signal with a 1.023 MHz chipping rate generated using a Gold code [Par96] repeated every millisecond. The C/A code is used for SPS. The Gold code is different for each SV, allowing the receiver to identify different satellite sources in the received signal. Both L1 and L2 transmit an identical week long Precision (P) code, denominated Y code when it is encrypted before broadcast. The P/Y code is used for PPS and has a 10.23 MHz chipping rate. Both the C/A and P/Y code signals are modulated with navigation data at 50 bps that includes ephemeris parameters, satellite clock corrections, health information for the satellite, and an almanac with basic information for the rest of the satellites in the constellation.

Block IIR-M satellites introduced new military signals, and a civil use signal in the L2 band named L2C. However the L2 frequency is not in a bandwidth protected for radio navigation, and so in some areas of the world it could potentially suffer interference from radars, and fixed or mobile services. Consequently the International Civil Aviation Organization does not include L2C as a usable signal in its Standards And Recommended Practices.

Block IIF satellites (1 launched of 12 planned by 2012) have an additional broadcast signal at 1176.45 MHz. Both the frequency and the signal are named L5. The signal has a 10.23 MHz chipping rate, and includes a dataless signal component

(allowing for more robust tracking), and the data includes forward data correction (allowing the receiver to detect certain transmission errors and correct them without retransmission). L5 (as is the case for L1) is within the 960-1215 MHz band protected worldwide for Distance Measuring Equipment in civil aviation. It is designed with a minimum receiver power level more than three times higher than the C/A signal (-154.9 dBW compared to -160 dBW for C/A at L1) [Heg08].

Block III satellites (built by Lockheed Martin, and to be launched starting in 2014) will include another signal named L1C which is intended to allow greater interoperability with the Galileo constellation.

The C/A, L1C and L2C signals are created employing a Binary Offset Carrier modulation, having a multiple peak autocorrelation function that produces a narrower primary peak, making tracking at the receiver more robust.

The GPS satellites are controlled and monitored by the GPS Control Segment (CS), which has a Master Control Station (MCS) at Schriever Air Force Base (Colorado) and a global set of monitoring stations and data upload antennas. Since 2005 each satellite is continuously in view of at least 2 monitor stations. The monitor stations continuously measure ranges and clock information from the SVs, which are transmitted to the MCS and fed into a Kalman filter for orbit and satellite clock estimation. The MCS generates 15 input parameters for the SV position pseudo-keplerian estimation formula (the six keplerian elements, three rates, and six magnitudes of sinusoidal like waves) [ICD00]. These parameters are uploaded once or more a day to the satellites, in the form of 12 overlapping sets that are valid for 4 hours each. The satellite rebroadcasts these parameters to all users in view, automatically shifting to the corresponding set every two

hours, and the aircraft uses these inputs to generate its estimate of the SVs positions which is needed for trilateration.

Historical GPS performance shows that the specifications of the system are met with a considerable margin [FAA11]. This will be described in more detail in the measurements section 2.2.

2.1.2 Other Constellations. GLONASS is a GNSS operated by the Russian Federation. Its nominal configuration is a Walker 24/3/1 constellation, i.e. 24 satellites in three orbital planes with regular spacing. The inclination is 64.8° and the altitude is 19100 km. Contrary to GPS, GLONASS has a Frequency-Division Multiple Access (FDMA) design, broadcasting at frequencies that are different for each satellite. The future generation (GLONASS-K) will broadcast at an additional frequency. In 2007 the Russian Federation reiterated the offer made by the USSR in 1988 to provide free of charge the GLONASS Standard Accuracy Service (SAS), intended for civil use worldwide. For older satellites, the SAS message is only broadcast at one frequency. Due to a temporary decline in the system's operability in the past, the fact that FDMA designs are more complicated to implement in receivers than CDMA, and reference frame and time coordination issues, there is a limited number of GLONASS capable equipment compared to GPS receivers. A CDMA implementation, added to the current FDMA, is being considered for the GLONASS-K satellites. The GLONASS constellation is not considered further in this work.

The European Union's GNSS system, Galileo, is planned as a Walker 27/3/2 constellation, with an inclination of 56° and an altitude of approximately 26000 km. It

will not be described here as it will not be operational before 2013. It is designed to be compatible with GPS (GPS receivers need minimal adjustments to also use Galileo signals). The quality of the signals and error bounds are, by design, improved compared to GPS. These include an open service which will be available to users worldwide free of charge. In this work, when the Galileo constellation is used in simulations, GPS error models will be utilized.

China is developing its own navigation system named Compass, which will have 30 MEO satellites. It is also intended to have an open service available to all users worldwide for free. The details of the signal structure have not been released by the Chinese government. Therefore, Compass is not considered in this work.

2.2 GNSS Based Navigation

2.2.1 Positioning Algorithm. A user needs the vector of ranges to the satellites in view ($\mathbf{y}$) to determine its 3-D position vector $\mathbf{x}^*$ using trilateration. A measured estimate of the range to SV i is given by:

$$\hat{\mathbf{y}}_i^* = |\mathbf{r}_i - \mathbf{x}^*| + \tau + \delta\mathbf{y}_i^* \tag{2.1}$$

where the first term on the right side is the difference between the ith satellite position vector $(\mathbf{r}_i)$ and $\mathbf{x}^*$. The second term (τ) is the difference between receiver clock time and GPS system time, multiplied by the speed of light. The source of this term is the user receiver, and thus it is a bias common to measurements from the same user to all satellites. The term $\delta\mathbf{y}_i^*$ includes all remaining measurement error sources such as

satellite clock offset from GPS time, ionospheric delay, tropospheric delay, multipath error, and receiver noise.

Using the Taylor series expansion of the norm of a vector Eq. (2.1) can be linearized as:

$$\hat{\mathbf{y}}_i^* = \begin{bmatrix} -\hat{\mathbf{e}}_i^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{r}}_i - \hat{\mathbf{x}}^* \\ \tau \end{bmatrix} + \delta \mathbf{y}_i^* \tag{2.2}$$

where:

$\hat{\mathbf{e}}_i \equiv \dfrac{\hat{\mathbf{r}}_i - \hat{\mathbf{x}}^*}{|\hat{\mathbf{r}}_i - \hat{\mathbf{x}}^*|}$ is the estimate of the line of sight unit-vector,

$\hat{\mathbf{x}}^*$ is the estimate of the user position vector,

$\hat{\mathbf{r}}_i$ is the estimate of the SV position vector and a

T superscript means transposed.

Element $\delta \mathbf{y}_i^*$ now includes some errors due to the linearization, which in all applications treated in this work can be considered negligible. [Note: this involves some iteration, using $\hat{\mathbf{y}}_i^*$ to estimate a better $\hat{\mathbf{x}}^*$ until it converges. This is discussed more after Eq. (2.5) is introduced].

Stacking the element or row components of Eq. (2.2) and redefining our vector estimate as:

$$\hat{\mathbf{y}} = \hat{\mathbf{y}}^* - \begin{bmatrix} -\mathbf{e}_1^{\mathrm{T}} \\ \vdots \\ -\mathbf{e}_n^{\mathrm{T}} \end{bmatrix} [\hat{\mathbf{r}}] , \tag{2.3}$$

(where $\hat{\mathbf{r}}$ is our estimate of SV position vectors), we can relate the state vector and the estimates by:

$$\hat{\mathbf{y}} = \begin{bmatrix} -\mathbf{e}_1^{\mathrm{T}} & 1 \\ \vdots & \\ -\mathbf{e}_n^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}^* \\ \tau \end{bmatrix} + \delta\mathbf{y} = \mathbf{H}\mathbf{x} + \delta\mathbf{y} \tag{2.4}$$

where we have defined our state vector $\mathbf{x}$ and our observation matrix $\mathbf{H}$. Now $\delta\mathbf{y}$ also includes errors in the estimate $\hat{\mathbf{r}}$.

Assuming we know the line of sight vectors needed to construct $\mathbf{H}$, we can obtain a least squares estimate of the user position (and the receiver clock bias nuisance parameter) from:

$$\hat{\mathbf{x}} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{z} \tag{2.5}$$

where the range vector estimate has been renamed ($\mathbf{z} = \hat{\mathbf{y}}$) to conform to the standard notation used within the navigation community for measurement vectors.

When a prior estimate of user position used to construct the matrix $\mathbf{H}$ is in error by a large amount, iteration might be necessary until the change of the least squares estimate is sufficiently small. However, the matrix $\mathbf{H}$ is constructed from line-of-sight unit vectors and the satellite distance from the user is very large, so it is not very sensitive to prior position estimate error.

If the measurement error variances are not equal across all satellites, a diagonal weighting matrix may be introduced to give a weighted least squares estimate that is more sensitive to more reliable SVs [Gel74]:

$$\begin{aligned}\hat{\mathbf{x}} &= (\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{z} \\ &= \mathbf{Q}\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{z} = \mathbf{S}\mathbf{z}\end{aligned} \tag{2.6}$$

The elements composing $\mathbf{z}$ and $\mathbf{W}$ will be clearly defined for each application in the following chapters. The covariance matrix $\mathbf{Q}$ and the weighted pseudoinverse $\mathbf{S}$ have also been defined to simplify notation in subsequent formulas.

The equations defined above represent the basic single point solution in GPS positioning. In general, more precise measurements and good satellite geometry (well separated and redundant SVs) will result in a more accurate user position estimate.

2.2.2 Notation. In subsection 2.2.1 some notation was introduced that will be used throughout the dissertation. The following is a summary to clarify it to the reader:

A bold capital letter represents a matrix, for example: $\mathbf{H}$

A bold lower case letter represents a vector, for example: $\mathbf{e}$

A 'hat' means 'an estimate of', for example $\hat{\mathbf{y}}$

A lower case 'delta' means an error in the estimate, where the identifying symbol with no hat or delta represents the actual value, giving, for example:

$$\hat{\mathbf{y}} = \mathbf{y} + \delta\mathbf{y} \tag{2.7}$$

To adapt to the common notation within the navigation research community, an exception to the notation in Eq. (2.7) is made when what is being estimated (through a measurement) is a user-satellite distance (or differential distance). In these cases the estimate vector will not be represented with a hat, but with a letter (for example 'z') such that:

$$\mathbf{z} = \mathbf{r} + \delta\mathbf{z} \tag{2.8}$$

A lower case italic symbol represents a scalar parameter. Some examples: satellite number 'i', a system specification (in this example a probability) 'P_{HMIreq}', an element from row i and column j of matrix **H**, '$H_{(i,j)}$'.

Operators will always be non-italic, for example superscript T (for transpose) or P for a probability of occurrence: $\mathrm{P}(a > b)$

Superscript asterisks '*' are used to name intermediate definitions during a derivation.

The absence of a time subscript implies the term refers to 'current time'.

2.2.3 Available Measurements and Error Sources. The basic ranging measurements available to users are described in this subsection, as well as ways of combining them (derived measurements) that are later applied in the implementations in subsequent chapters. The basic measurements are the code phase and the carrier phase. The two most important derived measurements are the carrier smoothed code measurement and the ionospheric delay-free dual frequency measurement. Also in this subsection the error sources are presented for both basic and derived measurements, grouping them as satellite based errors, atmospheric generated errors, and receiver-end errors, with a coarse description of their magnitudes. The concept of differential measurements is also introduced.

2.2.3.1 Code and Carrier Phase Measurements. The code measurement is based on the required phase shift to align the C/A code replica generated at the receiver with the

original code generated at the satellite. Ideally, multiplying this phase angle by the code chip length gives the SV-user range. However, since there are several error sources and a receiver clock bias, these measurements are commonly referred to as pseudoranges. For a certain SV and a given time (both subscripts are avoided in the following formulas), the pseudorange can be modeled as:

$$\rho^* = r + \delta\rho^* = r + c\left[\frac{\tau}{c} - v_{SV\tau}\right] + v_I + v_T + v_{MN} \tag{2.9}$$

where:

r is the range between the satellite (at time of transmission) and receiver (at reception time),

c is the speed of light

$v_{SV\tau}$ is the difference between GPS time and SV clock time at time of transmission, and

v_I, v_T and, v_{MN}, are respectively the ionospheric delay, the tropospheric delay and the error from multipath and receiver noise.

Since the distance from a satellite to a user on the Earth's surface varies from approximately 20000 to 26000 km, the transit time is between 70 and 90 ms. The Gold code broadcast by each SV repeats every millisecond, but the whole millisecond ambiguity is easily resolved with a rough estimate of user position (within hundreds of kilometers).

Receiver clocks are usually basic quartz oscillators and tend to drift, but are kept within a threshold difference from GPS time. Manufacturers have different ways of

ensuring that, but the details are not important here, since the receiver clock offset is part of our state vector.

The other error elements in the pseudorange will be described shortly, after the carrier phase measurement is introduced.

The carrier phase measurement is the phase shift required to align the receiver generated carrier signal at time of reception with the signal broadcast from the SV at time of emission. It can be modeled as:

$$\phi = r + \delta\phi = r + c\left[\frac{\tau}{c} - v_{SV\tau}\right] + N\lambda - v_I + v_T + v_{\phi MN}, \tag{2.10}$$

were

r, $c\left[\frac{\tau}{c} - v_{SV\tau}\right]$, v_I, v_T have the same values as in Eq. (2.9) for the same satellite and time (the difference in sign in the ionospheric delay term will be addressed shortly),

N is the number of whole cycles between SV and receiver generated signals; this integer ambiguity or cycle ambiguity is unknown,

λ is the carrier signal's wavelength, and

$v_{\phi MN}$ is the error caused by noise and multipath on the carrier phase measurement. A 'ϕ' is added to the subscript because the magnitude of this error is different for code and carrier measurements.

The satellite clocks diverge from GPS time, but a significant part of that error can be predicted by the CS. In the information uploaded to the SVs, three clock correction parameters are included. These are inputs for the user in a quadratic SV clock error

model. All SV clock errors mentioned in this dissertation refer to residual errors remaining after the quadratic correction is applied.

When the code and/or carrier phase measurements are put into the form defined in Eq. (2.3), and additional term v_e has to be introduced on the right hand side of equations (2.9) and (2.10) to account for errors in the SV position estimation. The simplified orbital model introduced in subsection 2.1.1 is accurate to within a few meters during its time of validity.

The nature and behavior of the satellite ephemeris and clock errors is sufficiently known, and it can be bounded with confidence with Gaussian distributions when there is no failure present. A rough descriptive characterization of the CS generated errors (SV clock and ephemeris together) on measured pseudorange and carrier phase would be as a Normal distribution with zero mean and a standard deviation of 3 m [Mis01].

The effects on ranging signals from going through the Earth's atmosphere can be circumscribed to two regions: the ionosphere, extending from about 50 km altitude to 1000 km (including most of the mesosphere, the thermosphere and the exosphere), and the troposphere, that extends from the surface to approximately 14 km altitude.

The ionosphere is a region of ionized gases, and the ionospheric delay of the signal is proportional to the number of free electrons it encounters in its path. This is normally measured as the Total Electron Content (*TEC*), an integral of the electron density in an imaginary 1 m^2 cross section cylinder from user to SV.

The signal delay is also proportional to the frequency (f) of the signal (i.e. the ionosphere is a dispersive medium). For the carrier phase signal and the code signal it is respectively given by [Mis01]:

$$v_{I\phi} = -\frac{40.3\ TEC}{f^2}, \qquad v_I = \frac{40.3\ TEC}{f^2} \tag{2.11}$$

which translate into a delay for the code, and an advance for the carrier. This explains the notation for this term in equations (2.9) and (2.10); the derivation of this property can be found in [Mis01]. The dispersive quality of the ionosphere can be used to remove the delay from the measurements, as will be shown shortly.

The ionization is the effect of solar ultraviolet radiation, so the *TEC* varies substantially with latitude, between day and night, and the period within the solar cycle. These effects can be modeled to an extent, and GPS satellites broadcast eight parameters that are processed by the user to generate inputs for a formula known as the Klobuchar model [Par96]. This reduces the effect of ionospheric delay by about 50% on average. There are also rapid local variations that cannot currently be modeled, particularly in the equatorial and polar regions, causing rapid fluctuations in the carrier phase (scintillation) or in signal amplitude (fading).

A useful simplification of the ionosphere is to view it as a spherical shell whose surface is separated from the earth's by 350 kilometers (on average the altitude of the highest free electron density). This is known as the thin shell model. The intersection of the range vector and the 'shell' is referred to as the Ionospheric pierce Point (IP). If a signal passes through the ionosphere at a slant angle the *TEC* will be bigger than if it does so vertically. In general the effect of a slant angle through an idealized rectangular slab can be quantified by a factor $OF_I^* = \frac{1}{\cos\zeta}$, where ζ is the zenith angle. For our thin

shell model, representing the effect is slightly more complicated; in this case the Obliquity Factor (OF_I) for the ionospheric delay is [Mis01]:

$$OF_I = \sqrt{1 - \left(\frac{R_E \sin \zeta}{R_E + 350\,km} \right)^2} \tag{2.12}$$

where R_E is the radius of the earth. It is easy to verify that this value ranges from 1 (for a satellite with a 90 deg elevation) to about 3 (for a satellite with a 5 deg elevation). Under normal conditions, the zenith delay will vary from 1 m to 10 m depending on the factors mentioned in the previous paragraphs.

Interplanetary Coronal Mass Ejections (ICME) at the sun's corona can produce an Ionospheric Storm (IS) when it reaches the earth's atmosphere. This increases the ionospheric delays by several orders of magnitude, and it also increases the fluctuations of *TEC* as the separation between two IPs increases. This can potentially cause extremely hazardous conditions for navigation. Chapter 3 is dedicated to minimizing this danger.

The troposphere is a non-dispersive medium, and all signals, code or carrier, of all frequencies will have the same delay for the same path. The magnitude of the tropospheric delay is proportional to the amount of dry gases and water vapor the signal encounters from the satellite to the user. Many complicated models based on weather conditions exist to estimate the delay and are used for surveying and similar applications. However, the inputs to these models (accurate meteorological data) are rarely available to GNSS users, and so they are seldom used for navigation. Similarly the obliquity factor is sometimes computed differently for the dry delay (caused by gases, and amounting for approximately 90% of the total delay) and the wet delay (caused by water vapor mostly

below a 4 km altitude). That approach is not practical for navigation. An example of a simplified model used to compute the Obliquity Factor for the Tropospheric delay (OF_T) is [Mis01]:

$$OF_T = \frac{1}{\sqrt{1-\left(\frac{\cos\xi}{1.001}\right)^2}} \tag{2.13}$$

where ξ is the elevation angle of the satellite.

Since there is no generalized model for the tropospheric delay used by GPS, the description and bounds of it will be given in detail for each application in subsequent chapters. Unless stated otherwise, Eq. (2.13) is used to generate the obliquity factor. Since the altitude of the troposphere is at the same level of the user, the effect of a slant trajectory will be bigger than the case of the ionosphere. It is easy to verify that the values of OF_T vary from 1 (for an SV at zenith), to almost 10 (for a 5 deg elevation).

The unpredictable nature of the atmosphere's behavior (particularly the ionosphere) makes it necessary to model the errors separately under nominal conditions and under anomalous conditions. A rough descriptive characterization of the ranging error magnitudes due to atmospheric delays (ionospheric and tropospheric together) would be as a Normal distribution with zero mean and a standard deviation of 5 m.

The final group of error sources is affected by code structure, signal power and receiver design. A simplified description can divide them into multipath and receiver noise. Multipath refers to components of the tracked signal that reach the antenna through indirect paths. This degrades the measurement of the direct signal, which is the one needed to obtain the SV-user range. Multipath can be caused by surrounding structures

from where the signals can reflect or by the ground. Some antennas are specifically designed to de-weight waves coming from low or negative elevations, where the secondary signals are more likely to be coming from. Receiver noise refers to contributions from signals that are not being tracked, and cannot be filtered completely, interference from other GPS signals, and the noise introduced by the receiver equipment itself.

The details of receiver design, software and hardware, are not the focus of this work. However, the level of noise in its receivers affects how a system performs. Accordingly, parts of the work presented in Chapter 3 will evaluate what level of noise would be acceptable to meet a system's demands, allowing comparison to what is actually achievable with the current technology.

Receiver-end errors affect the code and the carrier phase measurements differently, being about two orders of magnitude smaller for the carrier. The reason will be explained in a conceptual way, while a more rigorous mathematical explanation can be found in Chapter 6 of [Mis01]. The receiver-end noise affects the capacity of the receiver to precisely correlate the phases of the received signal and the replica. This results in an error in phase shift estimation. The wavelength of the carrier phase is tens of centimeters (for example 19 cm for L1). The unit of the Gold code used by GPS to generate the navigation message has a wavelength of approximately 300 m. Given this difference in wavelength magnitude, the same phase shift error caused by the same noise, when multiplied by the wavelength, will produce much bigger errors in the code estimation. We can summarize the receiver noise and multipath effect on measurement errors as a

normal distribution with 0 mean, and a standard deviation of roughly 1 m for the code and 0.01 m for the carrier.

As can be seen in the summary in table 2.1, the noise is not the dominant error source for the standalone user. This is why code measurements are suitable for an aircraft not using augmentation, as the benefit of using carrier phase is not worth the complications of using the ambiguous measurement. However, when augmentation is introduced, common errors can be eliminated, reducing SV and atmospheric related errors. Receiver-end errors cannot be reduced, as they are mostly uncorrelated between receivers or sites; and in those cases they can become dominant, making the use of precise carrier phase measurements worth the effort, or even a necessity, for precision approach operations.

Table 2.1 Approximate Contributions of Different Sources to Range Error (standalone)

Source	Satellite	Atmosphere	Receiver		Total	
Standard Deviation	Code or Carrier	Code or Carrier	Code	Carrier	Code	Carrier
	3 m	5 m	1 m	0.01 m	5.91 m	5.83 m

2.2.3.2 Carrier Smoothed Code Measurement. It was mentioned in the previous subsection that code measurements are suitable for standalone users in non-precision phases of flight. However, the most widely used range estimate in aircraft navigation is the carrier-smoothed code measurement. It takes advantage of the low noise level of the

carrier phase measurement without need of resolving the cycle ambiguity. It is defined for the current epoch as:

$$\rho = \frac{1}{M}\rho^* + \frac{(M-1)}{M}\left[\rho_{-1} + (\phi - \phi_{-1})\right] \tag{2.14}$$

where M is a scalar used to define the filter gains. In this formula, the subscript '-1' means 'from the immediately previous epoch' with respect to current time [Note: this is consistent with the notation introduced in subsection 2.2.2 where no time subscript implies current time]. The smoothed measurement ρ consists of a weighted sum of two estimates of the range: the current code measurement and a second estimate, based on the smoothed measurements from the previous epoch, to which the difference of the (current minus prior) carrier phase measurements is added. The filter (known as the Hatch filter) is explained conceptually in the following paragraphs. A mathematical development of the concepts introduced can be found in [Mis01].

When the filter starts it has to use a code measurement $\rho_0 = \rho_0^*$, with its corresponding (comparatively high) noise level, as it is the only available information. For subsequent epochs, as the filter applies Eq. (2.14) recursively, two effects take place simultaneously. The first one is a reduction of the receiver noise effect in the smoothed measurement from the code measurement level, to a level roughly corresponding to a carrier phase measurement. This happens because the filter is estimating the initial code measurement ρ_0^* at each epoch using the time differential carrier phase measurement and averaging it with all the estimates of ρ_0^* obtained in previous epochs. This is not obvious from Eq. (2.14) but can be easily shown [Mis01]. A larger value of M gives more weight

to the carrier phase measurements, maximizing the advantage of this averaging process. However, a second effect is the growth of an error in the last term in Eq. (2.14) generated by the different sign of the ionospheric delay for code and carrier phase measurements introduced in Eq. (2.11). This error increases with time, asymptotically approaching a maximum value [And00]. This maximum value is proportional to the M chosen for the smoothing filter in Eq. (2.14). M is then selected considering these two competing effects. In many current applications it is set at $M = 200$, equivalent to a time constant of 100 seconds for one sample every half a second. Errors caused by the satellite and atmosphere are not reduced by this filter and then it is obvious from Table 2.1 that the gain from this filter is not significant for standalone users. From now on, unless stated otherwise, when referring to the code measurement, it means the carrier-smoothed code defined in Eq. (2.14).

2.2.3.3 Ionospheric Delay-Free Error Measurement. Another useful derived ranging estimate is the dual frequency Ionospheric delay-Free (IF) error measurement ρ^{IF}. Consider two frequencies (in this example L1 and L5) available and using equations (2.11) and (2.14):

$$\rho_{L1} = \rho^{IF} + v_{I_{L1}} \quad , \qquad \rho_{L5} = \rho^{IF} + v_{I_{L5}} = \rho^{IF} + \left(\frac{f_{L1}}{f_{L5}}\right)^2 v_{I_{L1}} \tag{2.15}$$

With some basic algebra we can construct a measurement without the ionospheric delay as:

$$\rho^{IF} = \frac{f_{L1}^2}{f_{L1}^2 - f_{L5}^2}\rho_{L1} - \frac{f_{L5}^2}{f_{L1}^2 - f_{L5}^2}\rho_{L5} \approx 2.26\rho_{L1} - 1.26\rho_{L5} \tag{2.16}$$

Although this measurement does not have the ionospheric delay in it, a price is paid through an increase in other errors present. If we assume the receiver-end error sources are independent and identically distributed for L1 and L5 we can say that their contribution to the ranging error will be $\sqrt{2.26^2 + 1.26^2} \approx 2.77$ times bigger than for a standard pseudorange. An analogous derivation can be done for the carrier phase measurement, with similar advantages, in which the resulting cycle ambiguity becomes a linear combination of the integer for each individual frequency, but will still be invariant with time.

In general, if dual frequency is available, the IF measurements should be used, as the ionospheric delay is the dominant error. The symbol ρ will be used in certain chapters instead of ρ^{IF} (or ϕ for ϕ^{IF}) to simplify the notation. These cases will be explicitly identified.

A concept that is fundamental to the applications derived in this work is that of differential measurements. A *single difference* measurement is understood as the difference between measurements from two different receivers to the same SV at the same epoch. It is used to eliminate SV clock errors (completely correlated) and atmospheric and ephemeris errors (highly correlated if receivers are close to each other). A *double difference in time* measurement is the difference between two single difference measurements, one at the current time and the other at a reference time. It is most useful to eliminate the cycle ambiguity for carrier phase measurements. While the major error sources are cancelled out, the contributions of the multipath and noise are increased (as

was explained for the dual frequency IF measurement). The differential measurements concept will be expanded further in subsequent chapters of the dissertation.

2.2.4 Navigation Systems and Failure Monitoring. A navigation system (for example GPS) has a set of specifications that are met by the ground and/or space segments of the system provider (for example the US Air Force) before the user utilizes any information. For example, the GPS SPS Performance Standards guarantees measurement accuracy of 6 m or better 95% of the time [DoD08].

There is a second set of specifications, to be met by the user to allow a certain operation to be conducted. These can be summarized into three categories:

- *Accuracy requirements*: the distribution of the magnitude of the position estimate error should be bounded by a given Gaussian distribution. This requirement is usually stated as a threshold and a probability for a point of that distribution, and it can be divided into vertical and horizontal errors. For example, for an LPV-200 operation the vertical accuracy has to be less than 4 m 95% of the time. The accuracy requirement is not the main focus of this work, because it is assumed that if the other requirements (explained in the following paragraphs) are met, the accuracy requirement is met, as its bounding distribution is less conservative than the one required for integrity. However, the two requirements aim at different things. The accuracy requirement is related to nominal conditions, while as we will see soon, the integrity requirement is aimed at anomalous conditions.

- *Continuity*: If a pilot has to abort an approach after it was initiated, it is called a break in continuity. The probability of this occurrence is the continuity risk (P_c). As this

is an undesirable situation, there is a maximum tolerable probability, defined as the required continuity risk: P_{cReq}. Before starting an operation, the continuity risk is estimated by the user and if

$$\hat{P}_c < P_{cReq} \tag{2.17}$$

then the operation proceeds. For a break in continuity to happen, a monitor would have to trigger an alarm after the approach was initiated. Let's schematically think of a system operating n monitors for different threats. In principle the user could produce an estimate of the continuity risk from each monitor 'i' and generate: $\hat{P}_c = \sum_{i=1}^{n} \hat{P}_{c,i}$. However, this is not the way the continuity specifications are typically used, as will be explained in the next paragraph.

Since there are many monitors operating at the same time, P_{cReq} can be considered as a 'budget' that is distributed amongst them, such that $P_{cReq} = \sum_{i=1}^{n} P_{cReq,i}$ (see Appendix C for an allocation example). The alarm might trigger because there is a Potentially Hazardous (PH) condition, or it might be a Fault Free (FF) or false alarm. Because failures are considered rare, most of the allocation is given to FF alarms. Each monitor will generate a test statistic (α) that it will compare to a threshold (T). If $\alpha < T$ it is assumed that a FF condition is present. Setting the threshold is done in such a way that the continuity risk allocation is met for a given monitor i:

$$\mathrm{P}(\alpha_i > T_i|\mathrm{FF})\mathrm{P}(\mathrm{FF}) + \mathrm{P}(\alpha_i > T_i|\mathrm{PH})\mathrm{P}(\mathrm{PH}) \approx \mathrm{P}(\alpha_i > T_i|\mathrm{FF}) < P_{cReq,i} \tag{2.18}$$

- *Integrity requirements*: the user generates a bound on its position estimate error in real time, known as a Protection Level (PL). When the PL does not bound the

actual error, and the user is unaware of it, it is said that integrity has been lost. The probability of this occurrence is called integrity risk (P_{in}). Similarly to the continuity requirement, the system will have a tolerable P_{in} (P_{inReq}) that can be allocated to the n different monitors such that $P_{inReq} = \sum_{i=1}^{n} P_{inReq,i}$. For each phase of flight during an operation, the system will also have a maximum tolerable position estimate error magnitude, commonly defined as the Alert Limit (AL). The AL will have more stringent values as the aircraft gets closer to landing. The first priority of any analysis of monitor performance is that the integrity requirement is met. The starting point formula to achieve this is:

$$\mathrm{P}\left(\left|\delta x^{*}\right| > PL,\ \alpha_i < T \middle| \mathrm{FF}\right)\mathrm{P}(\mathrm{FF}) + \mathrm{P}\left(\left|\delta x^{*}\right| > PL,\ \alpha_i < T \middle| \mathrm{PH}\right)\mathrm{P}(\mathrm{PH}) = P_{in,i} \qquad (2.19)$$

were PL, P(FF), P(PH) are usually given by or derived from the system specifications (although the meaning of PH requires more explanation, which will be found in subsequent chapters), and T_i is derived directly from the continuity allocation for the monitor. In contrast to Eq. (2.18), in Eq. (2.19) we are interested in $\mathrm{P}(\alpha_i < T_i)$.

There are different ways to ensure a monitor is complying with integrity. One option is to compute $P_{in,i}$ from Eq. (2.19) and verify that:

$$P_{in,i} < P_{inReq,i} \qquad (2.20)$$

another is to replace $P_{in,i}$ with $P_{inReq,i}$ in Eq. (2.19), derive the corresponding PL, and then ensure that:

$$PL < AL \qquad (2.21)$$

- *Availability*: availability is defined as the percentage of time the system is available to initiate an operation. As integrity and continuity requirements are in place to ensure safety of life, algorithms for candidate architectures are constructed to ensure both are met, and then the effectiveness of the system is measured in availability.

2.2.5 System evaluation. In this work a system is considered available, for a given location and time, if the integrity requirement is met. (Remember continuity compliance is ensured by setting the monitors' thresholds appropriately. In this work it was also assumed that if the integrity requirement is met, the accuracy requirement will also be met). The result (available or unavailable) will depend basically on four things: measurement quality, satellite geometry, prior probability of faults and the effectiveness of the monitor implementation in detecting potential failures.

In most cases presented in this dissertation, simulations are run for 24 hours. The GPS ground track repeats itself every day, making a day's sample equivalent to long term performance. Depending on the application, different measures of availability are used:

- Specific sites: when more insight is required into the causes of unavailability, a detailed evaluation is done for specific sites, usually major airport locations.

- Global availability: a simulation is run for a grid (for example 5 deg spacing) over the whole globe. The total number of available cases (for all epochs and all locations) is divided by the total number of samples, giving a fraction (usually expressed as a percentage).

- Latitude evaluation: a longitude is picked, and availability is evaluated for different latitudes (for example 5 deg spacing), getting an availability percentage.

- Average availability: the availability for each site is computed individually, and then all the availabilities are averaged.

- Coverage: the availability is computed individually for each point in a worldwide grid (for example 5 deg per 5 deg). Then the percentage of earth surface with an availability above a set value is considered 'covered'. The result is then a percentage of the earth's surface. In this dissertation *weighted* coverage is usually used, where each point in the grid is multiplied by the cosine of its latitude, thus giving an actual area covered (for higher latitudes, smaller distances are subtended for a given longitudinal spacing). A latitude grid spanning -70 to 70 deg is used, as a very small number of airports are located in the polar regions.

2.3 Ground Based Augmentation Systems: LAAS

GBAS is a local augmentation architecture that serves one airport. Its name has a historical origin, coming from the way the information is transmitted to the users, via a VHF ground antenna.

The underlying principle of GBAS systems is that receivers that are close to each other will have a strong correlation on the main sources of error. Since the airport receiver's locations are precisely known, corrections can be generated and transmitted to nearby users. Then each user will apply these corrections to its own measurements. A GBAS equipped facility does not need the ILS implementation that is currently used at most airports suited for commercial airplane landings. ILS has proven to be a very safe system, since its first use in 1929. However a GBAS implementation has several advantages over ILS, some of which are: one set of equipment can service all runways at

the same airport (rather than two per runway which is the case for ILS); the cost of the GBAS equipment is significantly less than for ILS equipment, especially considering maintenance and replacement costs over time; it has the potential to support curved landing approaches, which can reduce fuel costs for airplanes; and it uses the latest technology available making more precise user positioning possible.

Several countries are developing GBAS equipped airports, including Australia, France, Germany, and Brazil, with several others planned for the near future. The US through the FAA has installed several LAAS prototypes for testing in different cities. Some details of GBAS will be explained in the next section.

2.3.1 Local Area Augmentation System. LAAS installations include a GF with 3 or more reference receivers within the airport's property, a data processing unit, which handles the measurements and generates the corrections, and a VHF transmitting antenna.

The GF will compute an estimate of each reference antenna-SV range based on the known antennae locations and the satellite positions obtained from the ephemeris broadcast. It will then compare this range estimate with the measured distance in each case. All the discrepancies between computed and measured estimates are consolidated into a correction.

The corrections for each satellite are transmitted to the users via a VHF Data Broadcast. The required coverage distance from each runway is 20 nautical miles. The frequency of transmission (108-117.975 MHz) is used for ILS and omnidirectional radio range systems, thus commercial aircraft are already capable of processing the message.

As one set of corrections is enough for all runways in an airport, bandwidth overcrowding is not an issue.

A generally unrecognized benefit of having multiple antennas in a GF is that differential carrier phase measurements across the antenna baselines can be used to detect and isolate certain signal-in-space failures and anomalies that are hazardous to LAAS. Two major classes of anomalies that fall into this category are satellite ephemeris failures and ionospheric storms. The last one is the topic of Chapter 3.

2.3.2 Ionospheric Delay and LAAS. The ionospheric delay is removed by the correction generated at the GF, except for a residual error due to the fact that the signal path is not identical for reference antenna and user antenna. During days of normal ionospheric activity, the GF broadcasts a conservative bound on the standard deviation of the spatial ionospheric gradient (σ_{vig}) to LAAS users. Under these normal circumstances, navigation integrity is ensured by incorporating σ_{vig} into the computation of position domain protection levels. Typically σ_{vig} varies from 1 mm/km to 4 mm/km [Pul09], and to obtain the standard deviation of the user's remaining ionospheric delay after the LGF's correction is applied, σ_{vig} is multiplied by the distance between the ground and user antennas.

Anomalies exhibiting abrupt changes in the ionospheric gradient have been observed during ionospheric storms in October and November 2003. Usually a big ionospheric delay would not be of concern, because if it is big for both the reference antenna and the user, the correction will mitigate its effect. However if the ionospheric storm front (described in detail in Chapter 3) is between the airport and the aircraft, the

ionospheric delay gradient can be several orders of magnitude larger than the nominal value, and σ_{vig} will not bound it. This creates an integrity breach.

Once dual frequency is available for reference receivers and aircraft, all ionospheric related errors can be eliminated. However LAAS will be ready to operate years before L5 is universally available (2019 for a 24 SV GPS L5 constellation, potentially before (2014/1015) if Galileo SVs are considered) [Gak11], and thus a solution is needed for the ionospheric front integrity threat during this transition. The details of the ionospheric storm threat model and mitigation techniques can be found in Chapter 3.

2.4 Space Based Augmentation Systems

A space based augmentation system provides corrections for satellites of the core constellation over a large area. Currently the core constellation for all SBAS systems is GPS. In the future they could include Galileo and GLONASS and/or Compass. The corrections are based on dual frequency measurements from an array of ground stations distributed over the coverage area. All measurements are sent to a master station that processes them and generates corrections for satellite clock and ephemeris errors. It also creates a model that describes the current ionospheric delay and bounds for all remaining errors. All of this information is uploaded to one or more GEOs that will broadcast it over the whole coverage area (this way of transmitting the information is the origin of the name SBAS).

The SBAS advantage over ILS and GBAS implementations is that it can cover a much larger area (continent sized) without an installation at each airport. However, it

cannot support the ILS precision approach operations (Categories I, II or III). Instead it can support oceanic/en route, terminal and non-precision approach phases of flight, as well as an operation not defined for ILS called LPV-200, similar to Category I, but less stringent (see Appendix D).

2.4.1 Existing SBAS. Currently several SBAS implementations are either operating or under development. WAAS will be explained in a separate subsection. All current generation SBAS GEO satellites broadcast in the L1 frequency, with a signal that resembles the L1 C/A GPS signal but with a data rate of 250 bps instead of 50 bps. The European Geostationary Navigation Overlay Service (EGNOS) is composed of three geostationary satellites: two Inmarsat-3 satellites, one over the eastern part of the Atlantic, the other over the Indian Ocean, and the ESA Artemis satellite above Africa. Unlike the GPS satellites, these three SVs do not have signal generators on board. A transponder re-transmits signals up-linked to the satellites from the ground. The ground segment consists of 34 ranging stations, four master control centers and six up-link stations. This configuration is highly redundant to guarantee continuity of service. The EGNOS open service has been available since October 2009. EGNOS positioning data through satellite signals is freely available in Europe to anyone equipped with an EGNOS-enabled GPS receiver, but no specifications are guaranteed as of today. Safety of life and commercial services are expected to be operational by the end of 2011 [ESA10]. The Japanese Multi-functional Satellite Augmentation System (MSAS) has a ground network in Japan, plus two stations in Canberra and Honolulu. It utilizes two GEOs. It

has supported safety of life operations since 2007. India is developing the GPS Aided Geo Augmented Navigation (GAGAN) system, expected to be operational by 2014.

There are other systems under development which are hybrids between a GBAS and SBAS implementations. These include the Australian Regional Augmentation System (GRAS), which has an SBAS-like ground station structure, but uses a VHF link to transmit the information to users, and the Japanese Quasi-Zenith Satellite System (QZSS) which includes inclined orbits at geosynchronous altitude with the intention of enhancing GPS in 'urban canyon' scenarios (i.e. to lessen blockages by buildings) and the Indian Regional Navigation Satellite System (IRNSS) with a combination of GEOs and inclined orbit SV's for a total of seven satellites to enhance the operational core constellations available (GPS, GLONASS, etc.). The Chinese Beidu/Compass system also has an SBAS component [Heg08].

2.4.2 Wide Area Augmentation System. WAAS is a GBAS that serves the North American continent. It has been commissioned for aircraft navigation in 2003, and can provide guidance down to an altitude of 200 feet for Localized Performance with Vertical guidance (LPV) operations. LPV operations are similar to (ILS) Category I, but have requirements derived to maximize the capabilities of WAAS.

It currently has 38 reference stations, most of them in the US, including stations in Alaska and Hawaii, and several stations in Canada, Mexico and Puerto Rico. It also has three master stations, two GEO SVs (a third SV is still broadcasting, but does not respond to commands since 2007), four ground stations (for uploads) and two operational control centers. Each station has three antennas, a precise atomic clock, and dual frequency

receivers, which provide redundant information at the master stations to generate reliable corrections and robust failure detection.

WAAS supplies two different sets of corrections: a set of corrections for GPS parameters (SV position and clock) and a set of ionospheric parameters. The first set of corrections is independent of the user's position. The second set of corrections is area specific. WAAS supplies correction parameters for a number of points (organized in a grid pattern) across the WAAS service area. The user receiver computes ionospheric corrections for the received GPS signals based on algorithms which use the appropriate grid points for where the user and each satellite are located [Han96] [RTC06]. The combination of the two sets of corrections allows for significantly increased user position accuracy and confidence anywhere in the WAAS service area. See for example Table 2.2 with the values for a user equipped with GPS only, compared to a user with equipment certified to execute WAAS LPV operations [Eld08].

WAAS generates two confidence bounds: the User Differential Range Error (UDRE) for the satellite errors, and the Grid Ionospheric Vertical Error (GIVE) for the

Table 2.2. GPS V.S. WAAS-LPV Accuracy

	GPS Specified	GPS Actual	WAAS LPV Specified	WAAS LPV Actual
Horizontal 95%	36 m	2.74 m	16 m	1.08 m
Vertical 95%	77 m	3.89 m	20 m	1.26 m

ionospheric errors, in both cases to account for residual errors after the WAAS corrections have been applied. For single frequency users, the GIVE is dominant when computing position error bounds, as the ionospheric delay is measured by WAAS in a relatively sparse manner compared to potential user locations. GIVE has to assume the worst possible user position, so the bound is significantly inflated. For dual frequency users, the ionospheric delay is eliminated, so the UDRE becomes dominant. The second case is the one of interest for this dissertation, as will become clear in chapters 4 and 5.

When UDRE does not bound the actual remaining range error WAAS must either correct the situation or notify users not to use the satellite, within the specified time to alert. The TTA becomes more stringent as the user approaches the runway. Meeting the TTA requirement has proven to be one of the most challenging aspects of WAAS applications.

2.5 Receiver Autonomous Integrity Monitoring

RAIM allows a user to autonomously verify, with a required high probability, that there are no undetected ranging errors big enough to cause a hazardous error in the computed user position. This is done by generating a test statistic that measures the consistency of redundant measurements in generating a position solution.

There are different RAIM implementations. Much of the work in the following chapters in this dissertation has its foundations in two classical concepts: the Failure Mode Slope (FMS) and solution separation. For both cases and throughout the dissertation, it is assumed that only one SV fails at a time. The reason why this

assumption is made will be described in each subsequent chapter; as it is different for each type of threat.

The two basic RAIM concepts are explained in this section, as well as how the threshold for each monitor is derived, leaving the generation of the position estimate error bound for chapters 3, 4 and 5.

2.5.1 Failure Mode Slope Implementation. The FMS implementation is based on the LSR concept introduced in 1.4.2. The procedure to use it is as follows:

- The user generates a position fix $\hat{\mathbf{x}}$ using a weighted least squares computation as defined in Eq. (2.6).

- Using the observation matrix as defined in Eq. (2.4), it then generates a vector of user-SV range estimates based on $\hat{\mathbf{x}}$:

$$\hat{\mathbf{y}} = \mathbf{H}\hat{\mathbf{x}} \tag{2.22}$$

- It then generates a weighted vector of residuals between the measurements and the estimates from Eq. (2.22), where each element is divided by the standard deviation σ_i of the corresponding measurement i:

$$r_i = \frac{z_i - \hat{y}_i}{\sigma_i} \tag{2.23}$$

The test statistic of the FMS RAIM is the norm of vector $\mathbf{r}$:

$$r = \| \mathbf{r} \| \tag{2.24}$$

- Finally it compares the test statistic to a pre-defined threshold T.

Note that if there is no redundancy in ranging sources, the residual will always be zero, thus giving the monitor no detection capabilities. Therefore a minimum of five satellites is required; this applies to all RAIM implementations.

It was shown in [Par88] that under fault free conditions, and assuming all ranging sources are Gaussian, r is Chi Square distributed with n-4 Degrees Of Freedom (DOF), where n is the number of satellites used to obtain the position estimate in Eq. (2.6). The threshold T has to satisfy the continuity requirement, and so it is computed from:

$$P_{cReq} = 1 - \frac{\gamma\left(\frac{n-4}{2}, \frac{T}{2}\right)}{\Gamma\left(\frac{n-4}{2}\right)} \tag{2.25}$$

where γ and Γ are the lower incomplete Gamma function and the Gamma function respectively, and the second term (i.e. the fraction) on the right hand side is the Cumulative Distribution Function (CDF) of a Chi Square distribution with n DOF at point T.

This implementation is very effective in detecting failures, as when a failure occurs, the effect of the failure (for satellite i) on the position error magnitude $\left\|\delta\mathbf{x}^i\right\|$ is linearly related to the effect in the residual r. Since the effect of random ranging error sources in r is independent of their effect on the position error $\boldsymbol{\delta}\mathbf{x}$ when computing $\hat{\mathbf{x}}$ [Per96b], the detection capabilities of the monitor can be easily determined for each geometry. The failure will act as a bias (non-zero mean) in the Gaussian position error distribution, and as a non-centrality parameter in the Chi Square residual distribution. By

considering the entire range of possible failure magnitudes f for each satellite i, a line can be drawn with a slope defined as the *FMS* [Per96b]:

$$FMS_i \equiv \frac{\left\|\delta\mathbf{x}^i\right\|}{\left\|\mathbf{r}^i\right\|} = \frac{\left\|\mathbf{S}\mathbf{f}^i\right\|}{\left\|(\mathbf{I}-\mathbf{HS})\mathbf{f}^i\right\|} = \frac{\left\|\mathbf{S}_{:,i}\right\|}{\left\|(\mathbf{I}-\mathbf{HS})_{:,i}\right\|}\frac{f}{f} = \frac{\left\|\mathbf{S}_{:,i}\right\|}{\left\|(\mathbf{I}-\mathbf{HS})_{:,i}\right\|} \tag{2.26}$$

Superscript i means the hypothesis is a failure in SV i. The failure vector $\mathbf{f}^i$ is a vector of zeros except for element i which is the size of the failure in the range measurement to SV i. We can observe in (2.26) that the *FMS* only depends on the satellite considered, the current geometry, and the weighting matrix used in pseudoinverse **S**. It is assumed in Eq. (2.26) that the same measurements are used in generating the position fix and the monitor's residual.

The concept of the FMS can be more easily visualized in Fig. 3.1, representing a residual-vertical error magnitude $(r - |\delta\mathbf{x}_v|)$ plane (the selection of the vertical error rather than the 3-D error simplifies the example, and the reason for that choice will become obvious in later chapters). Four *FMS* slopes, each one for a different satellite at the same epoch are represented. The plot also shows threshold T and the Vertical Alert Limit (*VAL*) on each corresponding axis.

For FF cases, the constant probability density contours will be centered at the intersection of the mean values for each axis: (n-4, σ_{x_v}), and the density of the joint Probability Density Function (PDF) is schematically represented by the blue colored ellipse. The probability that the vertical error is larger than the maximum tolerable error $\mathrm{P}(|\delta x_v| > VAL)$, is very small, and the probability of the monitor sounding the alarm $\mathrm{P}(r > T)$ is also small. However, when a failure is introduced, the center point of the

distribution moves along the *FMS* line, from Eq. (2.26): $\|\mathbf{S}_{,t}\|f = FMS_t \times \|(\mathbf{I} - \mathbf{HS})_{,t}\|f$ and the probabilities change.

Four potential situations exist, represented in the plot by the four quadrants of the *T-VAL* lines intersection:

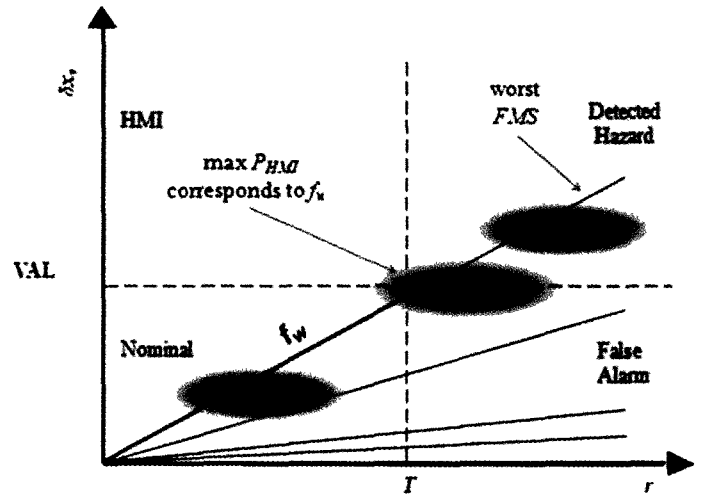


Figure 2.1. FMS Representation

- Nominal: $|\delta x_v| < VAL$ and $r < T$ (lower left quadrant),
- Detected hazard: $|\delta x_v| > VAL$ and $r > T$ (upper right quadrant),
- False Alarm: $|\delta x_v| < VAL$ and $r > T$ (lower right quadrant), and
- Hazardous Misleading Information (HMI): $|\delta x_v| > VAL$ and $r < T$ (upper left quadrant)

Another blue ellipse schematically represents the constant probability contours for the 'detected hazard' situation.

To obtain the total probability of HMI (equivalent to the integrity risk) we would have to integrate through all satellites and possible failure magnitudes:

$$P(HMI) = \sum_{i=1}^{n} \int_{f=-\infty}^{\infty} \mathrm{P}\left(|\delta x_v| > VAL \ , r < T \middle| f \right) \mathrm{p}(f)\, df\, P_{f^i} \tag{2.27}$$

where P_{f^i} is the probability that the failure occurred in SV i and $\mathrm{p}(f)$ is the PDF of f. In general, it will be assumed that the probability of a failure is the same for all satellites; furthermore, the whole probability $P_f = \sum_{i=1}^{n} P_{f^i}$ will be assigned to the satellite with the worst (biggest) *FMS*. This is conservative, and practical from a computational point of view. We can then write a simplified version of (2.27):

$$P(HMI) = \int_{f=-\infty}^{\infty} \mathrm{P}\left(|\delta x_v| > VAL \ , r < T \middle| f \right) \mathrm{p}(f)\, df\, \mathrm{P}_f \tag{2.28}$$

The PDF of failure magnitude is usually unknown, so the aim is to identify the worst failure magnitude f_w, and assign it a probability of 1 (conditional on the existence of a fault). Then (2.27) is further simplified to:

$$P(HMI) = \mathrm{P}\left(|\delta x_v| > VAL \ , r < T \middle| f_w \right) \mathrm{P}_f \tag{2.29}$$

P(HMI) can now be schematically represented by the portion of the red ellipse inside the upper left quadrant in Fig.3.1. Notice that f_w is not the biggest possible magnitude for the fault, but the one that maximizes *P(HMI)* .

The specifics of how variations of Eq. (2.27) are used in each case to ensure Eq. (2.19) is met will be developed in detail in the following chapters. For the implementations introduced in this dissertation, things will be significantly more complicated; as to expand and improve the FMS application, different measurements for the numerator and the denominator will be used in Eq. (2.25), which introduces new challenges that need to be overcome. Another assumption in the classical implementation is that the nominal measurement errors can be modeled as Gaussian. The consideration of some error types that cannot be modeled as Gaussian is also addressed later in this dissertation.

2.5.2 Solution Separation Implementation. The SS implementation is based on comparing the full solution estimate ($\hat{\mathbf{x}}$ from Eq. 2.6) with a subset solution that avoids using satellite *i*:

$$\begin{aligned}\hat{\mathbf{x}}_i &= (\mathbf{H}^{\mathrm{T}}\mathbf{W}_i\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{W}_i\mathbf{z} \\ &= \mathbf{Q}_i\mathbf{H}^{\mathrm{T}}\mathbf{W}_i\mathbf{z} = \mathbf{S}_i\mathbf{z}\end{aligned} \tag{2.30}$$

where $\mathbf{W}_i$ is the same as $\mathbf{W}$ but with the ith row and column set to zero. Since $\mathbf{W}$ is the inverse of the covariance matrix of the measurements, this is equivalent to ignoring measurement *i* in the solution by giving it infinite error variance.

The test statistic vector (we add a subscript 'ss' to distinguish it from the FMS residual) is:

$$\mathbf{r}_{ss,\,i} = \hat{\mathbf{x}} - \hat{\mathbf{x}}_i \tag{2.31}$$

The idea is that if there is a failure in SV i, the residual will be big, because we have a 'good' estimate ($\hat{\mathbf{x}}_i$) that is not using the failed satellite.

The SS method generates a residual for each satellite in view, and it also has to generate a threshold for comparison for each SV geometry and potentially failed satellite. In this sense it is more complicated than the FMS. However, once the monitor test passes, establishing a corresponding position bound is much more straightforward because there is no need to search for the worst failure size f_w, as will be explained in chapters 4 and 5.

For the FF case, the residual will be generated exclusively by nominal errors in the measurements (i.e. if they were perfect the full and the subset solutions would be identical). Thus the covariance describing the behavior of $\mathbf{r}_{ss,i}$ in each case is given by:

$$\begin{aligned} \mathrm{E}\left[\mathbf{r}_{ss,i}^2\right] &= \mathrm{E}\left[(\hat{\mathbf{x}} - \hat{\mathbf{x}}_i)(\hat{\mathbf{x}} - \hat{\mathbf{x}}_i)^{\mathrm{T}}\right] \\ &= (\mathbf{S} - \mathbf{S}_i)\mathrm{E}\left[\delta\mathbf{z}\delta\mathbf{z}^T\right](\mathbf{S} - \mathbf{S}_i)^{\mathrm{T}} \\ &= (\mathbf{S} - \mathbf{S}_i)\mathbf{W}^{-1}(\mathbf{S} - \mathbf{S}_i)^{\mathrm{T}} \end{aligned} \tag{2.32}$$

Again we will focus on the vertical error. The threshold in each case will be given by:

$$T_i = k_c\sqrt{\mathrm{E}\left[\mathbf{r}_{ss,i}^2\right]_{(3,3)}} \tag{2.33}$$

where the continuity multiplier k_c is such that:

$$\frac{P_{cRec}}{2n} = \frac{1}{\sqrt{2\pi}}\int_{k_c}^{\infty} e^{\left(\frac{-a^2}{2}\right)} da \tag{2.34}$$

to account for a false alarm in any of the n monitors and the two sides of each Gaussian distribution for the vertical error.

Each monitor producing a test statistic smaller than T_i can be translated into a protected position estimate error bound that complies with the integrity risk requirement. As for the case of the FMS implementation, this is left for the following chapters.

2.6 GNSS Evolutionary Architecture Study

The current process of modernization of GNSS will improve navigation user capabilities in many ways. There will be additional civil use signals, which will be more powerful and easier to acquire by receivers. They will also facilitate interoperability between different systems (GPS, Galileo and GLONASS for example), by transmitting inter-system clock corrections. Two of the predicted improvements are particularly relevant to this work: the additional protected civil signal, and the ability to use a larger number of satellites in view at all times. This will translate into more precise ranging (as seen in 2.2.3.3, ionospheric delay is eliminated using dual frequency measurements) and greater redundancy. Ranging precision and redundancy are naturally two key elements affecting the efficiency of RAIM implementations.

The FAA has organized a GNSS Evolutionary Architecture Study group to explore the different possibilities for aircraft navigation utilizing the new satellite signals available in the mid-term future. The objective of GEAS is to develop new navigation architectures for aviation to provide worldwide coverage for aircraft precision approach, initially LPV-200 operations, with minimal ground infrastructure. In response, the GEAS has focused its investigations on three different architectures, one of which is based on RRAIM. These were first described in [GEA08], and are summarized briefly below.

The first architecture most nearly resembles the existing Wide Area Augmentation System (WAAS), which is already capable of achieving LPV-200 performance, but not globally. This SBAS-equivalent concept is called the GNSS Integrity Channel (GIC) architecture. It is WAAS-like in that it will use sparsely placed reference stations to generate ranging corrections and perform integrity monitoring. However, ionospheric corrections are not required because L5 GPS signals are assumed to be available to airborne users. This means that a global network of GIC ground stations can be widely spaced, requiring perhaps only 20 stations worldwide [GEA08]. Based on WAAS experience, one challenge for this architecture is that it may be difficult to meet aircraft TTA requirements (6.2 seconds for LPV 200) with an integrated global system.

The second concept is called the Absolute RAIM (ARAIM) architecture. This approach is essentially a traditional RAIM architecture that uses carrier-smoothed code measurements for both positioning and fault detection. In this concept, the integrity burden is placed almost entirely at the aircraft. A simplified version of the GIC fault detection function is needed only to ensure that the prior probability of undetected multiple, simultaneous satellite faults is kept low. The ARAIM implementation depends heavily on redundancy, so it is very demanding on the satellite constellation. Initial results suggest that achieving good availability for worldwide LPV-200 with ARAIM requires constellations of 30 or more satellites [GEA08]. The choice of this architecture would therefore presume a suitably expanded GPS constellation or a combined-constellation GNSS. Chapter 5 develops a methodology to extend the use of ARAIM to otherwise unavailable geometries with a carrier phase RRAIM architecture.

The final GEAS concept is called the 'RRAIM' architecture. Like the previous two concepts, this system uses carrier-smoothed code for positioning. However, fault detection is performed using a combination of GIC ground based monitoring and a carrier phase RRAIM function. This architecture represents a practical intermediate solution between the GIC and ARAIM architectures. It assumes the same integrity monitoring capabilities of the 'GIC' architecture described above, but eliminates the resulting TTA concern by providing integrity for the latest segment of flight with a carrier phase RRAIM function. Preliminary analysis in [GEA08] showed that a RRAIM implementation could potentially enable LPV-200 operations with worldwide coverage using a 27 SV constellation, which is larger than the present nominal *specified* GPS constellation, but is consistent with the actual GPS constellation in orbit. This implementation is the topic of Chapter 4.

2.7 Relative RAIM

RRAIM is a novel RAIM application that is based on two features: the use of time differenced measurements in an innovative form and the use of carrier phase measurements, which have smaller receiver-end errors than code measurements.

The need for cycle ambiguity estimation can be eliminated by differencing measurements in time, creating spatial baselines associated with the user's translation over the time-difference interval. Each implementation introduces different challenges, and according to what measurements are available and what threat is to be mitigated, the differential measurements can be used solely for failure detection, or for both failure detection and navigation. The advantages and tradeoffs in each case are presented in each

corresponding chapter, but a summary of the architectures developed in this work is provided here:

RRAIM LSR aircraft baseline failure monitoring: The time differential carrier phase measurements are not used for positioning, but only for failure detection. This allows existing systems like LAAS to retain their current positioning algorithms, using RRAIM to detect specific threats like a static ionospheric front.

RRAIM LSR Range Domain: The user computes an initial reference position using carrier smoothed code, with integrity assured by augmentation from an SBAS source. The aircraft then uses the time-differential carrier phase measurements to navigate the latest path portion (from the initial position to the current time) and to detect failures in that period. Available information is merged in the RD before the position is estimated.

RRAIM LS Position Domain: similar to the implementation above, but the ranges are merged in the PD. PD implementations have the advantage that all satellites can be used for the initial reference position, but have the disadvantage of introducing new error terms due to uncertainty in the initial position.

RRAIM SS PD: similar to the LS PD implementation, but the monitor uses a solution separation technique rather than a least squares residual.

The last three of the methods described above use stored information which is known to be fault-free (from the SBAS augmentation) to compute the initial reference position.

ERAIM SS PD: this implementation is similar to the *RRAIM SS PD* detection technique, but the stored information used to obtain the initial position does not need SBAS information.

CHAPTER 3

IONOSPHERIC STORM THREAT MITIGATION

When a severe ionospheric storm occurred in 2003, its impact on navigation safety was clearly identified by WAAS, and became a source of concern for GBAS. At that point the FAA's LAAS program had been under intense development for several years. Conservative error bounds were being used to account for the ionospheric delay under normal error conditions, but the effects of extreme conditions during an IS had not been considered. When the bulk of the work included in this chapter was executed (2004-2006), it contributed in reducing the concern within the LAAS program by showing a simple effective solution was possible to mitigate the newly identified threat. Compared to other proposed monitors, the RRAIM implementation developed in this work is more focused on analysis, rather than presenting an ad hoc solution that covers all threats seen in the past, and so it is less sensitive to variations on the threat model. One negative aspect is that the user will need the broadcast from the GBAS installation of an additional parameter (message Type 6 for LAAS) with information on the ground carrier phase measurements.

The work presented in this chapter is based on a carrier phase time-differenced measurement RRAIM monitor with a FMS detection structure. The main challenge in this particular application is that the measurements used for positioning and the ones used for failure detection are different, making the practical advantages of the FMS monitor more difficult to capitalize.

Overcoming the challenge mentioned in the previous paragraph and providing a potential solution to the static ionospheric front threat are a specific contribution from the work in this chapter; but in a more general sense, it was the first in-depth development of a carrier phase relative RAIM application using carrier phase measurements, paving the way to its use in many other areas, for example the applications presented in chapters 4 and 5 of this dissertation.

3.1 Ionospheric Storms

Ionospheric storms are caused by solar activity when its effects interact with the earth's magnetosphere. When an IS takes place the impact on GNSS based navigation can be devastating as the ionospheric delay of signals can increase by several orders of magnitude. In 2003 during a peak of solar activity, an ionospheric storm occurred with unexpected negative effects. WAAS lost availability but had no integrity issues, as the errors were detected by its monitors; however, it raised questions about LAAS integrity if a similar situation occurred. This concern motivated this research. Properly describing this threat is extremely difficult, and most of the work in this chapter was based on threat models derived at Stanford University (SU) (described in subsection 3.1.2). Because of the difficulties in modeling the ionospheric delay during storms, the threat model was repeatedly updated as more data was analyzed [Luo03] [Luo04]. It is important then that the analysis and simulation tools developed are flexible to changes in the threat model, or can cover all possible threats. This is an aspect in which the RRAIM solution provided in this dissertation is comparatively strong. Of the different hazardous situations created by ISs, some can be detected by monitors already implemented for other threats. However, a

'static' ionospheric storm is less likely to be detected by any of those monitors (as will become clear in section 3.3), that is why the RRAIM solution provided focuses on a static front threat, which is the most problematic.

3.1.1 Nature of Ionospheric Storms. ISs are extreme events of space weather that can have an adverse effect on technological systems, including electric power blackouts due to induced currents, damage to satellites, exposure by humans in space or high altitudes, loss of HF communications, disruption of UHF satellite links due to scintillation, errors in VLF navigation systems and, of concern to this work, abnormally large delays in GNSS satellite signals.

The sun has a magnetic activity cycle that lasts approximately 22 years. During this period the magnetic polarity is reversed twice. From the Earth's received-radiation perspective it can be considered an 11 year-long cycle with clear maximum and minimum activity periods. During maximum activity, two types of events producing the negative effects described in the previous paragraph happen with greater frequency (for a classification of events, their magnitudes, effects and frequency of occurrence, see [NOA10]). These solar events affecting the Earth can be divided into solar flares and interplanetary coronal mass ejections (where interplanetary means they are in the solar system's plane). The former travel at speeds close to the speed of light, and have significant negative effects on communications and navigation; but usually not on GNSS differential measurements [Cer06]. The particles accelerated by ICME's hit the Earth's magnetosphere after one to five days, potentially causing a geomagnetic storm.

Geomagnetic storms cause high energy inputs in the upper atmosphere resulting in enhanced electric fields and currents and energetic particle precipitation. This changes the rates of production and loss of ionization, generating significant variations in the number of ionized particles. It also causes abrupt changes in atmospheric temperatures, leading to a thermal storm with rapidly moving particles between upper and lower atmosphere layers, which can all be summarized as an IS [Buo99]. The total electron content values will depart several orders of magnitude from the nominal and, more important to this work, the variation in *TEC* between IPs can also be significantly abnormal.

There are models describing the effect of ISs based on local time, universal time, onset time and magnitude of the geomagnetic storm as well as seasonal variations among other parameters. They can reproduce some phenomena, but usually post factum and not to the level of detail useful for navigation, even in the unlikely event that all inputs were available in a timely manner. However, there is enough information available from satellites and ground stations to know an IS storm is likely to happen or is taking place (more on this in subsection 3.3.4).

3.1.2 Ionospheric Storm Threat Model. As was briefly described in the previous section ISs are very difficult to model, however, for GBAS applications a simplified model might be sufficient. Because of the differential and local nature of GBAS, the threat model does not require an understanding of the causes, long term dynamics, or large area characteristics of the ionospheric storm.

The large magnitudes of the delay are not a concern, as the bulk of the error is removed by the ground corrections. We are only concerned with the residual errors. SU developed a model that describes the IS by its effect (delay) on the signals. It is described as a disturbed surface of the thin shell sphere with delay values that are significantly different from the rest of the surface that is undisturbed by the storm. IPs inside the disturbed surface will have similar delay values; IPs outside of that surface will also have similar values if they are close to each other, even if they are very different from the values of the storm-affected area. The transition between disturbed and undisturbed surfaces is defined as an ionospheric front which is modeled as a linear structure and can be defined by a few parameters. The parameter ranges are bounded based on observed data. Defining this model and its bounds was a difficult task, involving raw and post processed data from the Continuously Operating Reference Station network subjected to an exhaustive automated and manual analysis. This was done for all days for which WAAS availability was affected by IS conditions. A description of the challenges encountered in identifying the front structure and parameter values can be found in [Pul09]. The complex nature of the task led to the model evolving with time as it accommodated new data and better analysis techniques. Consequently different threat space models were used to generate results in [Gra05a] [Gra05b] and [Gra06]; however all results presented in this dissertation were regenerated using the current threat model [Pul09] which is presented in Appendix A.

The parameters used to describe the front are: the front width (w), the front gradient (k) and the front's velocity v_f. The front height (h) is derived from w and k. A sketch of the model is shown in Figure 3.1. Another parameter derived from empirical

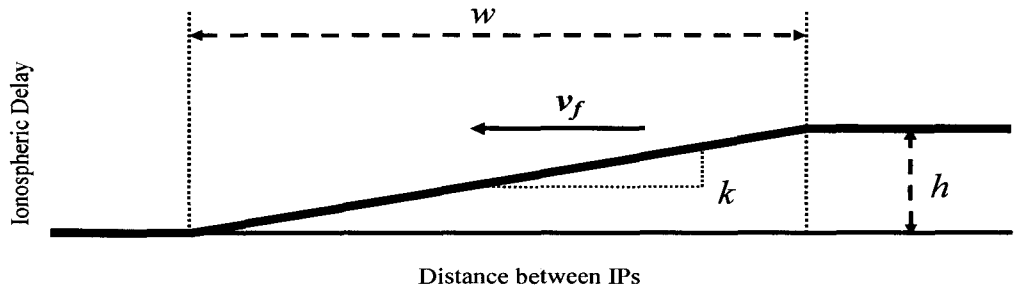


Figure 3.1. Ionospheric Front Model

data is a maximum plausible delay h_{max}. This bound implies discarding combinations of k and w for which:

$$h_{max} < k \times w. \tag{3.1}$$

The bounds on k on the SU threat model refer to slant gradient. This is convenient because the observed data used to generate the bounds can be translated into a threat model in a direct way.

One contribution of the work presented in this chapter is that it gives insight on what shapes of ionospheric fronts are more troublesome for the integrity of a GBAS system. This is relevant when studying the effectiveness of the RRAIM monitor acting jointly with other monitors, as they can each cover different parts of the threat space. It is convenient then to analyze the fronts considering their w-k combination.

The threat model does not consider the existence of fronts with w < 25 km plausible. However, it could be that only part of the physical front stands between the ground facility and the user, making the effect on the measurements equivalent to a front thinner than 25 km (see Figure. 3.7, where the front extends past the ground facility, but

w is measured only between the ground antenna and the aircraft). Accordingly in all the simulations in this work, the threat space covers $0 < w < 200$ km.

The assumption that only one satellite at a time is affected by a front is based on the fact that even though the IS can affect a large part of the sky in view, the actual front (layer between affected and non-affected parts of the sky) is very small as seen from the ground, so usually each satellite will lie on either the storm affected or the none affected part of the sky, with the exception being an SV right between the two. As will be seen in the following section, other monitors being developed consider the possibility of simultaneous failures. However, a more comprehensive RRAIM monitor that considers multiple SVs affected by the front is left for future work.

3.2 Mitigation Techniques

The simplest solution for GBAS to account for ISs is to augment the value used in position error bounding to account for the residual delay. This can be done by increasing the value (broadcast to the users from the GF) of the standard deviation of the nominal vertical ionospheric gradient σ_{vig}. However, this is not conservative enough for certain scenarios.

Additional mitigation can come from monitors specifically implemented to detect ionospheric fronts or even from monitors existing for other purposes. At the present moment the GBAS architectures under development base ionospheric storm mitigation for Cat I approaches on the combined work of three different monitors: the Code-Carrier Divergence monitor (CCDM), the Ground Gradient Monitor (GGM) and the Geometry Screening Monitor (GSM). Current research for mitigation during Cat III approaches

includes airborne monitors. The SU-developed GSM covers the threat space mitigated by the RRAIM monitor developed in this chapter.

At the end of this section, options that could provide an early alarm of a storm are described, as it might be convenient to activate the monitors only when an ionospheric storm threat is present, which would reduce availability loss considerably.

Research in this area initially assumed that one satellite at a time is being affected by an ionospheric front (see subsection 3.2.1). Later on, cases with two affected satellites were also considered. The following three monitor implementations, presented as context, have been evaluated considering two satellites at a time. Currently some research is dedicated to three or more affected satellites scenarios to take into account the 'bubble effect' of ionospheric delay [ICA10].

3.2.1 Code Carrier Divergence Monitor. As was explained in subsection 2.2.3.2 the carrier-smoothed code is widely used as the basic range measurement for positioning. The threat of an SV failure causing the code and carrier measurements to diverge excessively resulted in the implementation of a CCDM [Sim06]. This is advantageous to ionospheric front threat detection, because as was described in subsection 2.2.3, the ionosphere naturally causes a divergence between the code and the carrier ranges: an advance on the carrier, and a delay on the code. This effect, under nominal circumstances is considered when setting the threshold for the CCDM, but during a storm this divergence could get larger, making the monitor trigger an alarm.

For the case of LAAS, the CCDM has a divergence rate estimator and a detection test that compares the value of the estimator to a threshold. The input to the rate estimator

is the raw code-minus-carrier measurement. If the ground and aircraft are using the same filter with the same time constant and starting time, the error in code minus carrier will be accounted for in the correction broadcast by the ground. However, if the filter start times are different, the transient responses to CCD can be different for each monitor.

In [Sim06] it is shown that the CCDM is very effective in detecting divergence rates down to a few centimeters per second. The monitor assumes a nominal standard deviation for the CCD of $\sigma_d = 0.0039\,\frac{m}{s}$, which after being multiplied by a factor k_{ffa} to ensure an allowable probability of false alarm of the system, gives a monitor threshold of:

$$T = k_{ffa} \times \sigma_d = 5.83 \times \sigma_d \approx 0.023\,\frac{m}{s} \tag{3.2}$$

A moving front with a certain speed and gradient will be perceived to a static antenna (and the CCDM) as a divergence rate of magnitude: speed × gradient. Based on the model described in subsection 3.1.2, Figure 3.2 shows the divergence rates (in m/s and as seen from the monitor), caused by ionospheric fronts with different speeds and gradients (y and x axes respectively). To make it easier to read for our purposes, all divergence rates smaller than the CCDM threshold have been converted to a value of 0 in the plot. Without analyzing in detail the effect of noise, it is nevertheless obvious from the plot that for big values of speed and gradient, the CCDM will sound the alarm. On the other hand for small values of either one, the observed gradient is smaller than the threshold and the alarm will not trigger. In the case of small gradients, it can be argued

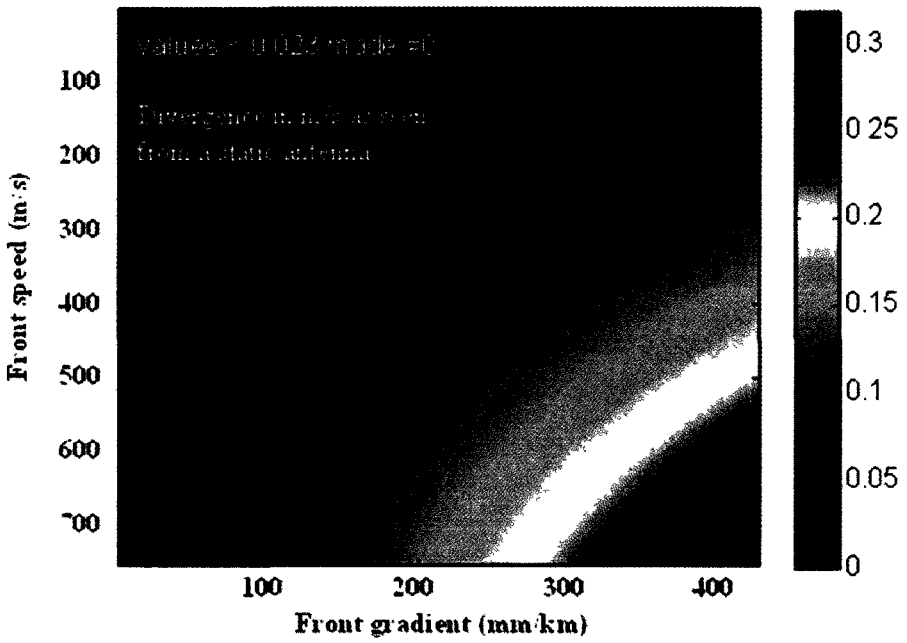


Figure 3.2. Potential Ionospheric Front Gradients Observed by the CCDM

that they are less likely to cause a hazardous position error. However for the small velocity, very big errors cannot be detected by the CCDM regardless of the front shape.

3.2.2 Ground Gradient Monitoring. Ground gradient monitors detect anomalous ionospheric delay values by comparing the ranges to the same satellite from two antennas forming a spatial baseline on the ground. Their detection capabilities depend on baseline length and measurement accuracy. Gratton and Pervan introduced the concepts and preliminary results in [Gra04] and [Gra05a], which are described below.

Consider the basic carrier phase measurement ϕ. A difference is taken between measurements from the two antennas forming each baseline to detect the ionospheric gradient. A double difference is then taken from a different satellite to remove the

receiver clocks biases, and finally a triple difference is taken from a reference epoch to remove the carrier phase cycle ambiguity. This measurement is defined as ϕ_{TD}. The standard deviation of this triple difference will be:

$$\sigma_{TD} = \sqrt{2}\sigma_{DD} \tag{3.3}$$

where σ_{DD} is the double difference standard deviation. The difference from a second satellite can be taken because it is assumed that measurements to only one satellite will be affected by the front. The test statistic for this monitor will then be:

$$\mu = \frac{\phi_{TD} - \phi_C}{\sigma_{TD}} \tag{3.4}$$

where ϕ_C is the computed triple difference using the broadcast ephemerides and known baseline vector. For this preliminary analysis ϕ_C is assumed fault free and with a negligible error. Under normal conditions, assuming Gaussian errors in ϕ_{TD}, μ is Chi Square distributed with one *DOF*.

Considering the linear model adopted for the ionospheric front, the differential delay introduced by it can always be represented by a bias in the measurement from one antenna to one satellite, and thus will be a bias (b) in the triple difference measurement. Accordingly it will be also a bias in μ equivalent to

$$b_\mu = \frac{b}{\sigma_{TD}} \ . \tag{3.5}$$

For this to be true, the front does not have to be present at the moment the reference measurement is taken, otherwise b will be eliminated when doing the triple difference. This will be addressed shortly.

The effectiveness of this monitor can be measured by computing the minimal detectable ionosphere gradient (k_{MDG}) between antennas, that will flag an alarm within the required probability of missed-detection P_{MDreq}.

If T_{GM} is the monitor's threshold and λ is the non-centrality parameter such that:

$$P_{MDreq} = \mathrm{P}\left(\mu^2 < T_{GM} \middle| \lambda\right) \tag{3.6}$$

then, with L being the distance between the antennas, we can compute k_{MDG} from equations (3.5) and (3.6) as:

$$k_{MDG} = \sqrt{\lambda}\,\frac{\sigma_{TD}}{L} \tag{3.7}$$

To get a rough idea of the effectiveness of this monitor, a test using real data and a baseline of 100 m was performed (See [Gra04] for antenna specifications). The ionospheric front was simulated by introducing different bias magnitudes in the measurements from one antenna after rise-time. Specifications were arbitrarily chosen as: $P_{ffa} = 1.5 \times 10^{-4}$ and $P_{MD} = 10^{-4}$, resulting in values of $T_{GM} = 3.791$ and $\lambda = 8.053$. The simulations resulted in μ being always bigger than the threshold for values of $b > 1.5$ cm. Figure 3.3 is reproduced from [Gra04] showing the borderline case where the monitor stops being reliable in detecting the bias (when some points in the plot are below the detection threshold, in red in the plot, for $b = 1.2$ cm). For this example the monitor can detect ionospheric fronts with a gradient of 150 mm/km or more (150mm/km×100m=1.5cm).

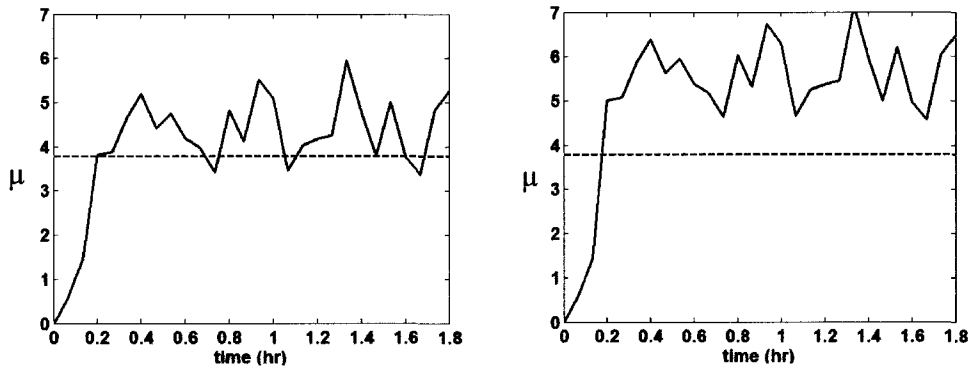


Figure 3.3. Ionospheric Front Bias Detection

The gradients being studied at this moment range from 0 to 425 mm per km. Thus, in order to cover the whole range of threats, the baselines would have to be 425 m long (for the same antenna performance).

As mentioned when introducing the monitor, the obvious flaw of this monitor is that if the front was present when the satellite rises, the bias caused by the gradient will be eliminated when doing the triple difference to remove the cycle ambiguity, and it will go undetected. A possible solution is to take the reference measurement from the lapse in which the rising satellite is below the elevation mask, or to elevate the mask if there is an ionospheric storm warning. Because of w size restrictions in the threat model, if there is enough time from the moment the reference measurement is taken, the signal at reference time could not have been affected by a front affecting the current signal. The noise in the measurement will be worse because of the low elevation of the satellites, and the standard deviation of the triple difference measurement will be higher. Then as the satellite rises, the monitor's performance will improve as the SV elevation gets higher.

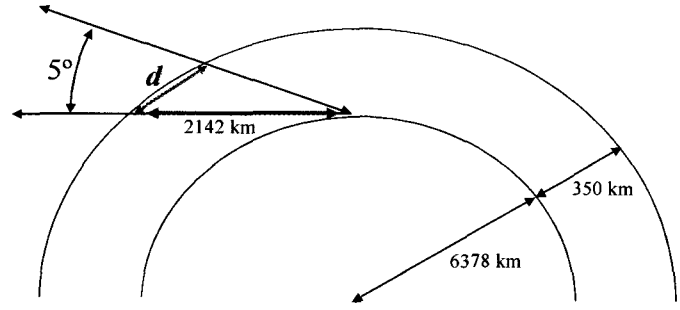


Figure 3.4. Mask Augmentation Analysis

Doing a conservative analysis, assuming the satellite rises in a direction perpendicular to the horizon, the distance the satellite signal's projection travels before getting over the mask is bigger than 183 km ('*d*' in Figure 3.4). If its signal goes through an ionospheric front when the satellite becomes available for navigation, it is possible to find an unaffected reference measurement for all fronts with $w < d$. If the mask is 7.5 deg, then the distance will be bigger than 280 km. This solution does not seem plausible at the moment, as the best standard deviation obtained with the current antenna technology for measurements below the 5 deg mask is larger than 10 mm, making the minimal gradient detectable larger than 400 mm/km.

A different monitor concept also uses baselines between the existing antennas, but its test statistic is based on integer ambiguity estimation, avoiding the triple difference. The test statistic for this monitor is the difference between the estimate of the cycle ambiguity c_{Amb} and the closest integer n:

$$\mu = \hat{c}_{Amb} - n = \mathrm{N}(0, \sigma_{DD}) \tag{3.8}$$

The distribution of μ is normal, and is represented by the black curve in Figure 3.5. In the event of a storm causing a bias b in the measurement, the distribution of μ will be:

$$\mu = \mathrm{N}(b, \sigma) \tag{3.9}$$

Given a threshold T is set to satisfy the continuity specification for the system, there is a magnitude for b, b_{min} for which:

$$P_{MDreq} = P\left(\mu < T \,\middle|\, b_{\min}\right) \tag{3.10}$$

This case is represented in Figure 3.5 by the blue curve showing the distribution, and the green line showing $b_{\min}$. As b increases, the area under the right tail past $-T$ (for the n+1 integer) will be big enough to make the P_{MD} larger than the allowed integrity risk. This determines a maximum detectable bias $b_{\max}$. This is represented by the red distribution and the pink bias magnitude in Figure 3.5.

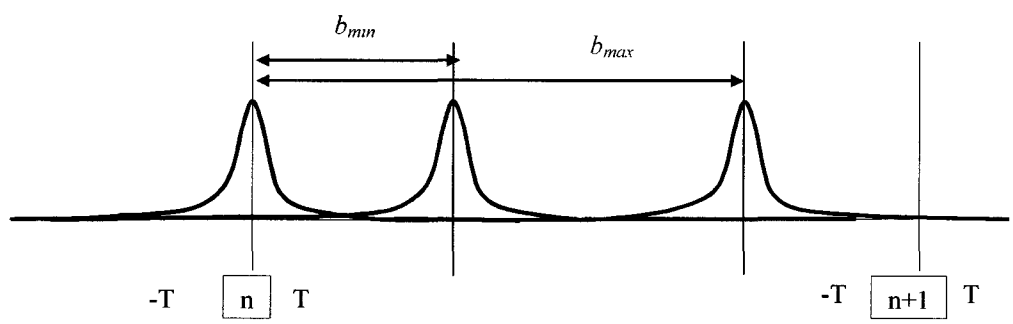


Figure 3.5. Cycle Ambiguity Baseline Monitor Test Statistic

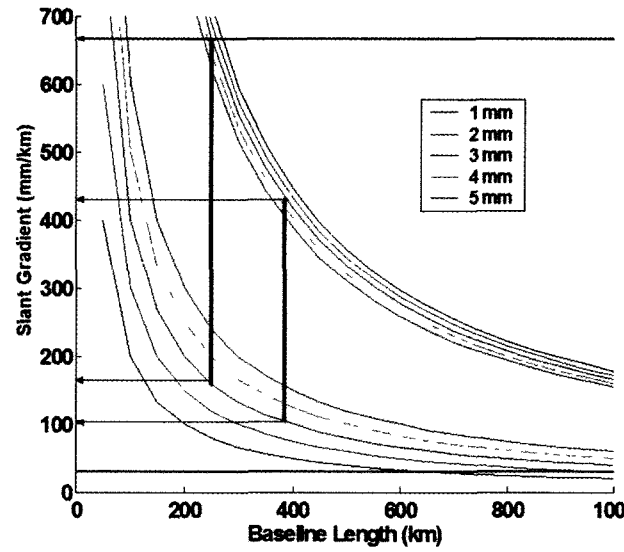


Figure 3.6. Maximum and Minimum Detectable Gradients for Cycle Ambiguity Baseline Monitor

Figure 3.6 shows the detectable slant delays with respect to L, with a different pair of curves corresponding to different values of measurement error standard deviation σ_{DD}. The top and bottom horizontal lines represent the maximum and minimum slant gradients for the threat space. The formulas corresponding to minimum and maximum detectable gradients are:

$$k_{MAX} = \frac{0.19\,\text{cm} - (k_{ffa} + k_{MD})\sigma_{DD}}{L} \tag{3.11}$$

$$k_{MIN} = \frac{(k_{ffa} + k_{MD})\sigma_{DD}}{L} \tag{3.12}$$

with the k's on the right hand side being the multipliers corresponding to the fault free alarm and missed-detection specifications (not to be confused with gradient values on the left hand side).

For each combination of L and σ_{DD}, a range of detectable gradients can be extracted from the plot. For example, the left vertical line shows that for a baseline of 200 m, all gradients bigger than 200 mm/km can be detected with a σ_{DD} of 3 mm. The second vertical line shows that for a baseline length of 400 m, any gradient between 100 and 425 mm/km can be detected with the same σ_{DD}. Double difference standard deviations smaller than 3 mm are not credible at this moment, so assuming a best case scenario of $\sigma_{DD} = 3$ mm and unrestricted siting possibilities for the baselines, we could potentially build a pair of orthogonal baselines of 1 km to detect all gradients between 50 mm/km and 200 mm/km, and a pair of orthogonal baselines of less than 400 m to detect any gradient larger than 200 mm/km. However, the necessary values of measurement standard deviation and the restrictions on antenna siting at airports are not minor obstacles to consider. Another factor to take into account is the fact that the measurement noise for this implementation would be increased by the use of at least two baselines, as the front direction cannot be assumed to be in the direction of a certain baseline.

The work in 3.2.2 was expanded in 2010 by Khanafseh, et al., who updated it with current antenna performance capabilities [Kha10].

3.2.3 Geometry screening monitoring. The two methods described in subsections 3.2.1 and 3.2.2 mitigate certain ionospheric storm threats. One way to deal with the remaining uncovered threat space is by geometry screening, a method studied in depth at SU [Pul09]. Each GBAS GF will examine at a certain specified time interval (for example every 5 minutes, and when an SV rises or sets) all possible SV subset

geometries in sight. This is required as the GF does not know exactly which SVs are being tracked by individual users. The GF will make unavailable those SVs for which a hazardous condition due to an ionospheric front is not guaranteed to be detected by the CCD or baseline monitors. To do so, at all Decision Height (DH) points (the last position in which a pilot can abort a landing operation) of all runways for the airport, simulations for all plausible geometries (4+ SVs) are executed for all possible front gradients and orientations. For each geometry, the worst case potential position error is computed, and this is compared to a threshold (according to [Shi08], 28.8 m is a safe value). In parallel, the broadcast σ_{vig} values used by the aircraft to compute position bounds are inflated in such a way that the users would generate protection levels higher than the allowable alert limits for all hazardous subset geometries. This will result in the aircraft not initiating an approach. According to the studies made so far, the availability loss due to the implementation of this method is sustainable within CONUS, particularly if the monitors are only implemented during ISs, with the alarm coming from WAAS. [Pul09]

3.2.4 Airborne Monitors. Category II and III approaches will require the implementation of airborne monitors. Current research anticipates the use of the following monitors: a CCDM at the aircraft, a consistency check between two position estimates that use two different filtering times for the carrier smoothed pseudorange (30 s and 100 s), and possibly geometry screening based on potential undetected ionospheric threats. These monitors are not discussed further in this dissertation, but a description of them and their performance can be found in [Mur07].

3.2.5 Early Warning. The ionospheric storm condition is a relatively rare event. The monitors implemented to mitigate its effect on GBAS GFs (the geometry screening presented before, or the RRAIM method following) will produce availability loss by discarding poor geometries when the monitor would not be able to detect a failure 'if there was one'.

It would be desirable to have an early warning system that would allow these monitors to be turned on only when there is a storm threat, rather than on a permanent basis. SU suggests using WAAS warnings [Pul09]. However, this is only possible for GFs in the North American continent, unless future studies show that if a storm hits the earth's atmosphere anywhere, its effects will always impact the WAAS covered region to a degree that makes it detectable.

Another option is to rely on the satellites in space with instruments specifically deployed to study solar activity. ACE, in a halo type orbit around the Earth-Sun L1 libration point, can provide accurate information determining if an ICME will affect the earth's ionosphere with at least an hour warning, while SOHO can provide a less accurate early warning of CMEs, but without knowledge of whether it will impact the atmosphere. The Solar Terrestrial Relations Observatory (STEREO) can also provide the same early warnings as SOHO but with fewer false alarms [NAS10].

3.3 Static Front Threat Mitigation Using RRAIM

The monitor described in this section has the advantages of the FMS implementation introduced in subsection 2.5.1. However, as will become clear shortly, it requires the use of carrier phase measurements for efficient front detection. In contrast to

the monitor presented in subsection 2.5.1, the measurements used for positioning (numerator of Eq. (2.26)) and for fault detection (denominator in Eq. (2.26)) are now different. In the classical RAIM implementation the measurements are carrier smoothed code measurements for both. In the RRAIM implementation that follows, time/space differenced carrier phase measurements are used for failure detection. Although the effect of the noise on the position error is still independent from the effect of the noise on the residual, in this case the same ionospheric front corrupting the measurements will have a different 'failure magnitude' on the numerator and denominator of Eq. (2.26) [Note: I only refer to the definition of FMS_l in Eq. (2.26); the terms after the left-most equal sign are not valid anymore]. The convenient linear relationship between position error and residual value is lost, complicating the evaluation of a system's integrity. This problem is solved through analysis and the introduction of innovative practical algorithms.

The following is a schematic explanation of how the RRAIM monitor would work. The details of the algorithm will be described in the next section.

Before executing a precision approach within a GBAS, an aircraft evaluates the monitor's capability of detecting a Static Ionospheric Front (SIF) if should one be present. This is done by validating that the integrity requirement is met using the current SV geometry and the error models for the measurements; no actual measurements are involved in this evaluation (this will be described in detail in section 3.3). If the integrity requirements are not met, the approach is not initiated. In this case, the detection function is said to be 'unavailable'.

When the aircraft reaches the ground VHF antenna Broadcast Radius (BR), it will store a set of differential carrier phase measurements. After that, at each point during the

approach, the monitor will generate a RRAIM residual using the stored carrier phase measurements and the current carrier phase measurement vector.

As long as the residual is below the threshold for the monitor, the user can assume there is no hazardous SIF present and can compute its position as usual using the carrier smoothed code measurement vector $\boldsymbol{\rho}$ in place of '$\mathbf{z}$' in Eq. (2.6).

3.3.1 RRAIM Monitor Implementation. The monitor needs to store a reference measurement to generate its test statistic (the RRAIM residual). In general it is desirable to have taken the reference measurement as far from the current aircraft location as possible, because with larger distances the effect of the SIF on the residual will be more significant, while the effect of the measurement noise will remain (in a statistical sense) the same, making detection more likely. So the reference measurement vector $\boldsymbol{\varphi}_0^{**}$ is taken at the BR distance. The user will also store a measurement vector $\boldsymbol{\varphi}_{G0}^{**}$ from the GBAS GF, which is transmitted through the VHF antenna. With these reference measurements and the two corresponding current time measurements, the time double difference carrier phase measurement vector is formed:

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}^{*} - \boldsymbol{\varphi}_G^{*} = \left(\boldsymbol{\varphi}^{**} - \boldsymbol{\varphi}_G^{**}\right) - \left(\boldsymbol{\varphi}_0^{**} - \boldsymbol{\varphi}_{G0}^{**}\right) \tag{3.13}$$

For GBAS, a typical BR is about 45 km. If the user is close to the BR point, the baseline formed with the aircraft translation will be small, but the integrity requirements are less stringent, as the aircraft is still far from the runway. When it reaches the DH point, the integrity requirements are more stringent, but they are also more likely to be

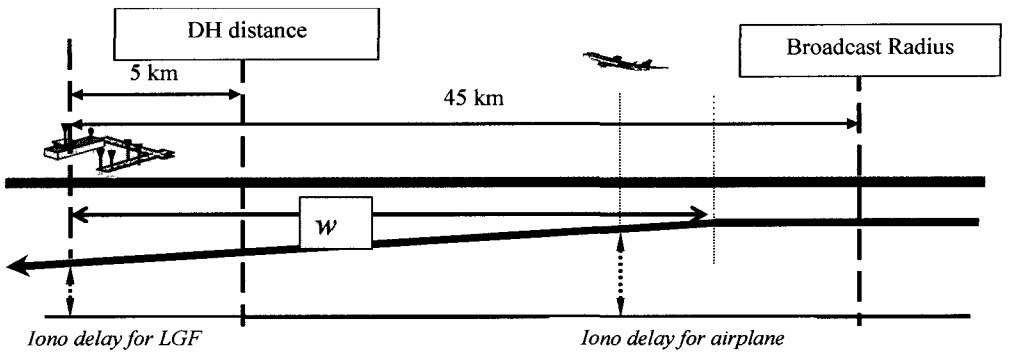


Figure 3.7. Approach with Ionospheric Front Present

met by the monitor because the user-GF baseline will be 40 km long. For this analysis the DH point is considered to be 5 km from the GF's VHF antenna; and it is the closest point to the runway in which a monitor's alarm would be useful. It must be stated that for the user to be able to compute the differential measurements, the GF would need to broadcast its carrier phase measurements and corrections (in the case of LAAS, a message Type 6). Figure 3.7 shows a scheme of an approach, with a representation of the delay caused by the ionospheric front at each point between the airport and the user in red on the lower part of the figure.

Under normal ionospheric conditions, the error in each component of measurement vector φ can be modeled as Gaussian:

$$\delta\phi_i = N(0, \sigma^2_{\phi,i}) \tag{3.14}$$

where:

$$\sigma^2_{\phi,i} = 4\sigma^2_{\phi\, n+MP,i} + \left(d_i \sigma_{vig}\right)^2 \tag{3.15}$$

where $\sigma_{\phi\ n+MP,\ i}$ is the standard deviation of the noise and multipath error for the carrier phase measurement to SV*i*, which has been multiplied twice by $\sqrt{2}$ to account for the double difference operation, and d_t is the distance between pierce points of the current and BR aircraft measurements.

An ionospheric front will affect ϕ_t in the form of a failure of size $f_{\phi,\ t}$, and its impact on the test statistic r will be:

$$r^t = \left\| \left(\mathbf{I} - \mathbf{HS}_\phi\right)_{\cdot,\ t} \right\| f_{\phi,\ t} \tag{3.16}$$

where $\mathbf{S}_\phi$ is the pseudoinverse of $\mathbf{H}$ as defined in Eq. (2.6), weighted with matrix $\mathbf{W}_\phi$ with its elements given by:

$$W_{\phi\,(i,i)} = \sigma_{\phi,i}^{-2}, \qquad W_{\phi\,(i,j)} = 0 \tag{3.17}$$

$f_{\phi,\ t}$ is an element of vector $\mathbf{f}_\phi$ affecting the corresponding element ϕ_t of vector $\boldsymbol{\varphi}$, and is given by:

$$f_{\phi,\ i} = k \times d_i \tag{3.18}$$

where k is the slant delay. The other elements of vector $\mathbf{f}_\phi$ are zero as the front is assumed to affect one SV at a time.

In this work, a simplified version of the error model will be used, in which d_t is replaced by the distance d_t, travelled by the aircraft between taking the reference and current measurements:

$$f_{\phi,\ t} = k \times d_t \tag{3.19}$$

The vertical position error caused by the ionospheric front will be:

$$\delta x_{I,v} = S_{\rho\,(3,:)} f_{\rho,\,i} \tag{3.20}$$

where $\mathbf{S}_\rho$ is the pseudoinverse of $\mathbf{H}$ as defined in Eq. (2.6), weighted with matrix $\mathbf{W}_\rho$ whose elements are given by:

$$W_{\rho\,(i,i)} = \sigma_{\rho,i}^{-2}, \qquad W_{\rho\,(i,j)} = 0 \tag{3.21}$$

with $\sigma_{\rho,i}^2$ being the variance from the error model used by the GBAS for carrier smoothed code and f_ρ is the magnitude of the bias introduced in the code measurement by the ionospheric front.

The accumulated delay on the Hatch filter (from Eq. (2.14), see derivation in Appendix B) is given by:

$$\begin{aligned} f_{\rho,\,i} = & \left(\frac{M-1}{M}\right)^q \left(-f_{\phi^{**}0,\,i}\right) + \sum_{j=1}^{q} \left(\frac{M-1}{M}\right)^j \left[f_{\phi^{**}(q+1-j),i} - f_{\phi^{**}(q-j),i}\right] \\ & + \sum_{j=1}^{q} \frac{1}{M} \left(\frac{M-1}{M}\right)^{j-1} \left[-f_{\phi^{**}(q+1-j),\,i}\right] \end{aligned} \tag{3.22}$$

were $q \times \Delta t$ is the time since the user entered the front, and Δt is the interval between measurements.

The inputs in this formula are the delays caused by the front on the user's undifferenced (but after applying the correction transmitted from the GF) carrier phase measurement φ_l^{**} at each epoch l. Thus, if we assume the higher end of the front is further from the ground facility (Fig. 3.7), the inputs to Eq. (3.22) are given by:

$$f_{\phi^{**}\,l,\,i} = k\left(w - v_a \times \Delta t \times l\right) \tag{3.23}$$

where v_a is the aircraft speed. Like Eq. (3.19), Eq. (3.23) also uses a simplified error model that assumes the ionospheric gradient refers to distance travelled by the aircraft.

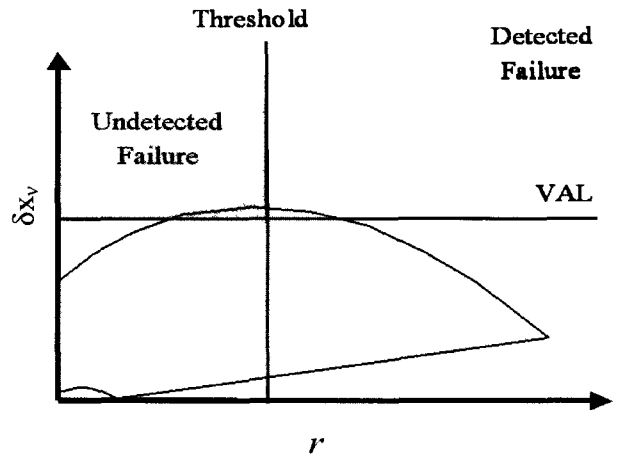


Figure 3.8. Carrier Phase RAIM Failure Space

Using the FMS RAIM method, we apply Eq. (2.26) to obtain:

$$FMS_i \equiv \frac{\left\|\delta\mathbf{x}^i\right\|}{\left\|\mathbf{r}^i\right\|} = \frac{\left\|\mathbf{S}_\rho \mathbf{f}_\rho^i\right\|}{\left\|\left(\mathbf{I} - \mathbf{HS}_\phi\right)\mathbf{f}_\phi^i\right\|} = \frac{\left\|\mathbf{S}_{\rho,:,i}\right\|}{\left\|\left(\mathbf{I} - \mathbf{HS}_\phi\right)_{:,i}\right\|}\frac{f_\rho}{f_\phi} \tag{3.24}$$

However, the effect of a failure magnitude on the position error and the RAIM residual cannot be described as a line as was the case for Eq. (2.26). In this case f_ρ has a 'memory' of past ionospheric delay values from the filtering (Eq. (3.22)), while f_ϕ is an 'instantaneous' value (Eq. (3.19)). Figure 3.8 shows schematically what the threat space to be evaluated looks like in this case as the different values of w and k are combined. All points on the irregular light blue shape in the figure have to be evaluated for each satellite.

This process is prohibitively time consuming; the computational time needs to be reduced to make this an effective tool. To do so, we will look at the continuous version of Eq. (3.22) [Gra05a]:

$$f_\rho = k\left[d_{ga} + 2\,v_a\,\tau\left(1 - e^{-\frac{w-d_{ga}}{v_a\,\tau}}\right)\right] \tag{3.25}$$

where d_{ga} is the current user distance from the ground antenna.

Using equations (3.19) and (3.25) we define:

$$\alpha_w \equiv \frac{f_\rho}{f_\phi} = \frac{\left[d_{ga} + 2\,v_a\tau\left(1 - e^{-\frac{w-d_{ga}}{v_a\tau}}\right)\right]}{d_t} \tag{3.26}$$

For a given w, Eq. (3.24) can be rewritten as:

$$\begin{aligned} FMS_{w,t} &\equiv \frac{\left\|\delta\mathbf{x}^t\right\|}{\left\|\mathbf{r}^t\right\|} = \alpha_w \frac{\left\|\mathbf{S}_{\rho,,t}\right\|}{\left\|\left(\mathbf{I} - \mathbf{HS}_\phi\right)_{,t}\right\|}\frac{f_\phi}{f_\phi} = \alpha_w \frac{\left\|\mathbf{S}_{\rho,,t}\right\|}{\left\|\left(\mathbf{I} - \mathbf{HS}_\phi\right)_{,t}\right\|}\frac{d_t k}{d_t k} \\ &= \alpha_w \frac{\left\|\mathbf{S}_{\rho,,t}\right\|}{\left\|\left(\mathbf{I} - \mathbf{HS}_\phi\right)_{,t}\right\|}\frac{k}{k} = \alpha_w \frac{\left\|\mathbf{S}_{\rho,,t}\right\|}{\left\|\left(\mathbf{I} - \mathbf{HS}_\phi\right)_{,t}\right\|} \end{aligned} \tag{3.27}$$

were all the intermediate steps have been explicitly shown because it will help visualize concepts expressed later on in this chapter.

The advantages of the FMS form have been recovered, where, for a given user position, the failure magnitude that moves along the line with slope *FMS* is now the front gradient k. The importance of this will become obvious in subsection 3.3.2.

To verify integrity is ensured, following the reasoning in equations (2.27)-(2.29) the user will verify that:

$$P_{MDreq} > P_{MD,f(w,k)} = \mathrm{P}\left(\left|\delta x_v\right| > VAL \ , r < T \middle| f(w,k)\right) \tag{3.28}$$

for all *w-k* combinations in the threat space, before initiating an approach. In this and subsequent formulas, '$f(w,k)$' represents a front with width w and gradient k. Unlike f_ρ and f_ϕ which represent the effect of f on measurements, i.e. specific numbers, $f(w,k)$ expresses the assumption that a certain front shape is present..

3.3.2 Computation Time Issues. The condition in Eq.(3.28) needs to be evaluated at all points of the threat space, i.e. the area covered by the filled shape in Figure. 3.8. The advantage of covering all front shapes by fixing the w and looking at different magnitudes of k using Eq. (3.27) is that it allows knowledge of $P_{MD,f(w,k)}$ without actually computing $\mathrm{P}\left(\left|\delta x_v\right| > VAL \ , r < T \middle| f(w,k)\right)$ for many $f(w,k)$. This is accomplished in three ways:

1) By deducing there is no value of k in the threat model that can cause a $P_{MDreq} < P_{MD,f(w,k)}$ for that value of w;

2) By finding an 'equivalent' front shape w_{eq}, k_{eq}, where $P_{MD,f(w_{eq},k_{eq})}$ was previously computed, and for which $P_{MD,f(w_{eq},k_{eq})} = P_{MD,f(w,k)}$; or

3) By finding a 'conservative' front shape w_{co}, k_{co}, for which $P_{MD,f(w_{co},k_{co})}$ was previously computed, and for which $P_{MD,f(w_{co},k_{co})} > P_{MD,f(w,k)}$.

The effect of applying these three concepts can be schematically summarized in Figure. 3.9, where the threat space to be evaluated is reduced for each satellite compared to Figure. 3.8.

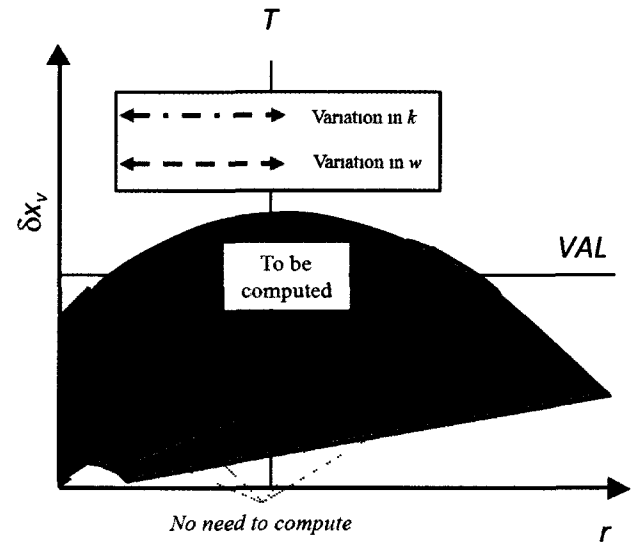


Figure 3.9. Carrier Phase RAIM Failure Space to Evaluate

The time saved by avoiding the actual computation of Eq. (3.28) for many $f(w,k)$ is not only useful for availability simulations, but, as the user would actually be performing this analysis before initiating an approach, it is also relevant to the flight software needed to implement the monitor. Accordingly the three concepts introduced at the end of the previous page will be explained in some detail, although the reader does not need to go through the rest of the subsection to understand results section 3.5.

The range of w to be computed is derived first. On the lower end, for any width thinner than 5 km, there will be no detection (the front will be entirely located between the airport and the DH point) so the detection capabilities cannot 'get worse'. As the

position error will only be smaller as w decreases, it is conservative to assign the following value:

$$P_{MD,f(w<5\,\mathrm{km},k)} = P_{MD,f(w=5\,\mathrm{km},k)} \tag{3.29}$$

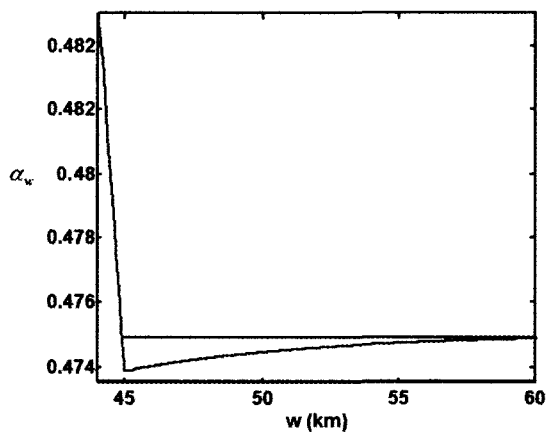


Figure 3.10. Values of α_w

To save computation time on the higher end of values of w, let's look at how α_w changes with w. The blue curve in Figure 3.10 shows the values for 43 km < w < 60 km. For 0 < w < 5 km, $\alpha_w = \infty$; after that the value decreases rapidly until w = 45 km, then increasing again, until at w = 200 km it takes the same value than for w = 44.9 km (represented by the black horizontal line). The variation in α_w over the range 43 km < w < 200 km is 2×10^{-2}. Since a lower slope reduces the P_{MD}, it is conservative to assign the following value:

$$\alpha_{w=44.9\,\mathrm{km}<w<200\,\mathrm{km}} = \alpha_{w=200\,\mathrm{km}} \tag{3.30}$$

It can be seen from equations (3.26) and (3.27) that if α_w is fixed, equal values of f_ϕ will produce equal values of P_{MD}. So once the availability has been computed for all values of k with w = 200 km, the availability for 44.9km < w < 200km can be obtained with minimal computation by looking at the already known values on the right hand side of

$$P_{MD,f(w,k)} = P_{MD,f(w=200\text{km},k_{equiv})} \tag{3.31}$$

where:

$$k_{equiv} = \frac{k(w - d_{ga})}{200\ km - d_{ga}} \tag{3.32}$$

is the gradient that will produce the same f_ϕ when combined with w = 200 km than the f_ϕ corresponding to the evaluated w-k pair.

So for each geometry, and each satellite, only the P_{MD} for 5 km < w < 44.9 km, and w = 200 km need to be computed, the rest can be deduced or bounded from these values.

Reducing the k space to be evaluated is more complex, as it will change with each geometry and satellite. The P_{MD} is the outcome of a joint probability (Eq. (3.28)); but if it is assumed conservatively that any failure size will cause $\delta x_v > VAL$, the Non-Central Chi Square CDF of the residual alone will determine the P_{MD}. To find the highest value of k that can be hazardous (any larger value would trigger the alarm within the required P_{MD} regardless of the position error produced) the limiting formula:

$$P_{MDreq} = \mathrm{P}\left(D_{dof} < T, q_{high}^{dof}\right) \tag{3.33}$$

can be used to find the non-centrality parameter q_{high}^{dof} that causes a $CDF = P_{MDreq}$ for a Non-Central Chi Square distributed random variable D_{dof} with dof DOF. The Non-centrality parameter affecting the RAIM residual distribution is given by the square of the denominator in Eq. (3.24). Using the value q_{high}^{dof} obtained using Eq. (3.33) the maximum value of k to be evaluated is given by:

$$k_{\max}(\mathbf{H}, w) = \frac{\sqrt{q_{high}^{dof}}}{\left\| \mathbf{I} - \mathbf{HS}_{\phi} \right\|_{(:,\,i)} (w - d_{ga})} \tag{3.34}$$

Similar reasoning can be used to obtain the minimum k to be evaluated by assuming there is no detection. The bias μ_{low} that gives a $CDF = P_{MDreq}$ for a Gaussian distributed $D_g = \mathrm{N}(0,\sigma_v)$ can be computed:

$$P_{MDreq} = P\left(D_g > VAL \,|\mu_{low}\right) \tag{3.35}$$

with

$$\sigma_v = \sqrt{\left(\mathbf{H}^T \mathbf{W}_{\rho} \mathbf{H}\right)_{3,3}} \tag{3.36}$$

In this case, the bias in the P_{MD} evaluation comes from the numerator in (3.27), (no detection is assumed); then the k minimum (k_{min}) to be evaluated is obtained as:

$$k_{\min}(\mathbf{H}, w) = \frac{\mu_{low}}{\alpha_{(w)} H_{3,i} (w - X_{DH})} \tag{3.37}$$

The main advantage of finding these limits relies in finding satellites for which no computation needs to be done at all for a certain w and $\mathbf{H}$. Figure 3.11 shows an example of this with two different slopes. It can be seen that for similar values of k_{min} and $k_{max,}$ for

the satellite with the smaller *FMS*, $k_{min} > k_{max}$. This means no value of k can produce a hazardous situation, and no evaluation of P_{MD} is necessary for that satellite.

For those satellites with kmin < kmax it must be determined if there is a value of k for which PMD > PMDreq. If that is the case, the geometry is unavailable. As was explained at the beginning of this chapter it is useful to know what combinations of w-k values make a geometry unavailable to evaluate what cases might be mitigated by other monitors. For a given w, finding which values of k make a geometry unavailable does not demand excessive computational time; once a hazardous value of k is detected.

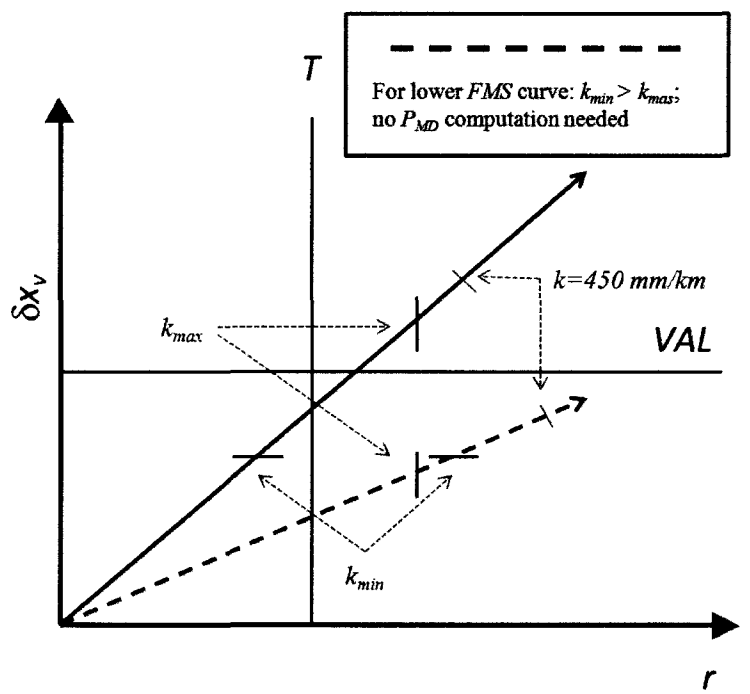


Figure 3.11. Search Limits

What consumes most of the search time is being able to discard the existence of a k value that will produce a $P_{MD} > P_{MDreq}$. The key is to find the biggest search interval (Δk) between evaluated k values that will guarantee finding any value of k that causes unavailability (detection not guaranteed by the monitor). A detailed method to minimize computational time, while guaranteeing detection, is presented in Appendix F.

3.4 Results

To evaluate the effectiveness of the RRAIM monitor introduced in section 3.3, a simulation is executed for two different scenarios. In one case specifications that resemble the requirements of a Category I approach are used, while for the second case, a more stringent set of requirements, similar to a Category III approach, is used (see Appendix A). The availability of integrity is verified for each epoch with an interval of 4 minutes for ten representative airport locations.

At each epoch it is verified if $P_{MD,f(w,k)} < P_{MDreq}$, for all w-k combinations in the threat space, using Eq. (3.28) and the algorithms described in subsections 3.3.1 and 3.3.2. This simulation is based on covariance analysis, which is the same type of analysis an aircraft will do before initiating an approach to verify integrity availability. The outcomes of these evaluations are used in three different ways:

Overall availability. If for any $f(w,k), P_{MD,f(w,k)} < P_{MDreq}$, then the system is declared unavailable. The aircraft would not begin the approach under these circumstances. By dividing the available epochs by the total number of epochs, an overall

availability can be obtained for each site evaluated. These results assume the RRAIM monitor is the only monitor providing integrity against static ionospheric fronts.

Overall availability acting with other monitors. The overall availability is computed using only part of the threat space, as the integrity risk from the other part of the threat space is assumed mitigated by another monitor (CCDM or GGM).

Hazardous front shape identification. The availability with respect to each specific *w-k* combination is computed (it is equivalent to computing the availability assuming each *w-k* combination was the only front shape possible in the threat model). This way it can be identified which fronts are difficult to detect and cause unavailability, and which are not problematic.

Before showing the results, it is appropriate to explain the constellation used. It introduces a new way of evaluating the availability by looking at the effect of the monitor using the real GPS constellation, with the actual outages. The period considered was the whole year 2004. For that period, the Notice Advisory to Navigation Users (NANUS) were consulted [USC10], establishing all cases in which a satellite was added or taken out of the GPS constellation. For each lapse of time without changes in the number of SV's, a set of broadcast ephemerides close to the middle point of that period was used to generate the satellite positions for that lapse of time. This has the advantage of giving a real sense of the impact of the monitor implementation with the actual satellites in space and their outages. It also has the added value, that for segments of time (with unchanged constellation configuration) bigger than a sidereal day, only the availability for the first 24 hours need to be computed as the results will be repeated in a daily basis.

3.4.1 Availability Results. Table 3.1 shows the results of the overall availability for the RRAIM monitor compared to the availability of the RRAIM monitor acting jointly with a GGM monitor. In this case the GGM monitor is assumed to detect any front with a $k >$ 150 mm/km and a $w < 24.5$ km. The value of minimum detectable k is considered reasonable given the preliminary analysis presented in subsection 3.2.2. The constraint on the size of w arises from the possibility of the front lying between the ground facility and the aircraft. In that case the antenna baselines at the airport will not be able to detect any gradient, as the front is not on top of them. However, as was explained in subsection 3.1.2, a front of a certain w (larger than 25 km) can have the effect of an equivalent front with $w < 25$ km if only part of the front lies between the ground antenna and the user (Fig. 3.7). This implies that any front evaluated with $w < 25$ km is necessarily 'on top' of the ground facility. By using a maximum detectable w (by the GGM) of 24.5 km we are ensuring that a length of front of at least 0.5 km is observable at the ground facility.

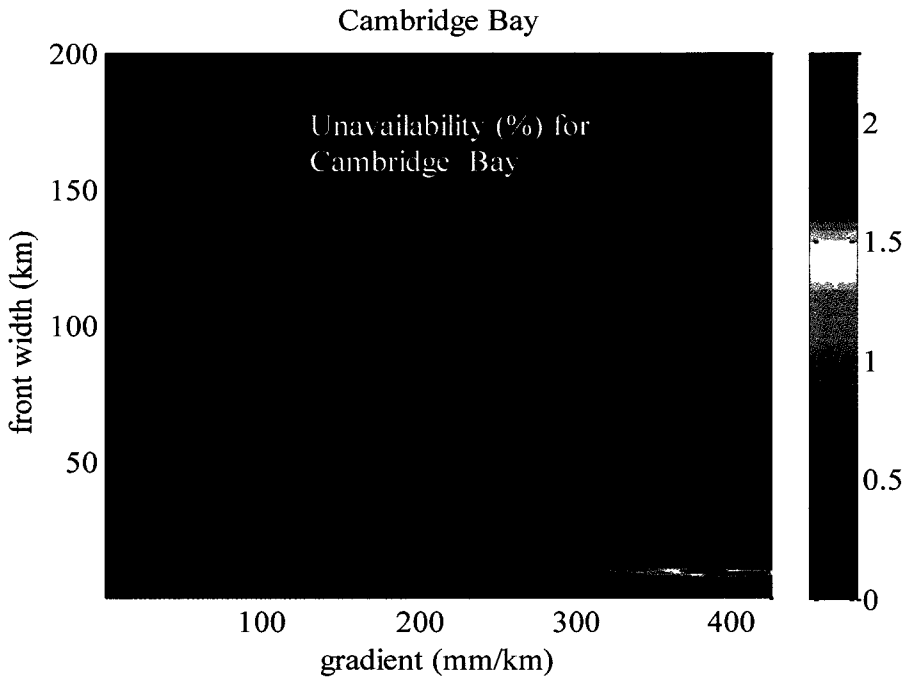
Results for the ten locations evaluated show and average availability of 98.91% for the 'Category I' simulation, and 97.76% for the 'Category III' simulation when using the RRAIM monitor alone. These values show a significant loss of system availability when monitoring an ionospheric static front with the RRAIM monitor alone. However if a GGM monitor like the one described in this section is functioning, the average availability for the same ten sites increases to 99.96% and 99.93% respectively for the 'Category I' and 'Category III' simulations. In this case the worst case values are 99.78% and 99.67% respectively. The loss of availability is still considerable. However, the availability loss is tolerable if the RRAIM monitor is only activated during ionospheric storm alarms, as described in subsection 3.2.5.

Table 3.1. Static Ionospheric Front Threat-Mitigation Availability (Percentage %)

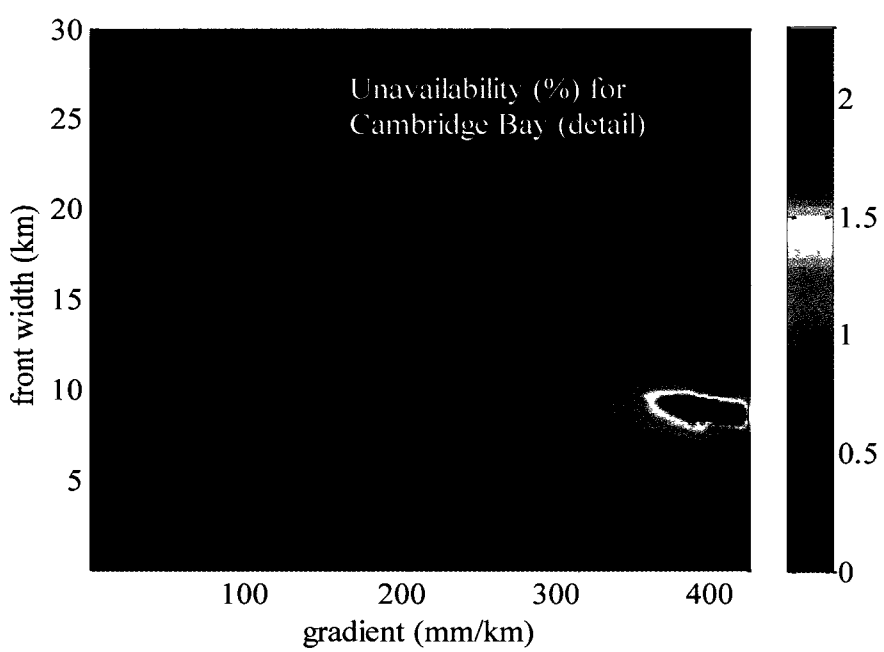
Location		Camb. Bay (Canada)	London (UK)	Los Angeles	Panama (Panama)	Lisbon (Portugal)	Seattle	Miami	New York	Chicago	Dallas
Cat I	RRAIM	97.11	99.43	99.06	98.70	99.39	99.65	98.30	98.44	99.26	99.83
	RRAIM + GGM	99.78	99.99	99.98	99.99	100	100	99.96	100	99.98	99.99
Cat III	RRAIM	93.30	98.10	97.33	98.32	98.66	99.29	96.61	97.85	98.58	99.56
	RRAIM + GGM	99.67	99.94	99.86	99.99	100	100	99. 94	99.99	99.97	99.97

3.4.2 Hazardous Fronts Identification. The improvement shown in subsection 3.4.1 in system availability from using the GGM could be quantified thanks to the ability to identify the shapes of the fronts that can't be reliably detected by the RRAIM monitor in each case. An example is shown in Fig. 3.12 that gives further insight into what *w-k* combinations are the most problematic. The plot for Cambridge Bay was chosen, as it is the site with the worst results in all four measures shown in Table 3.1. The general shape is the same for all ten sites evaluated; only changing the magnitude of the availability loss. Looking at the results it can be seen that the front shapes that can potentially go undetected while being a threat to position estimation, are the ones with k values larger than 100 mm/km (and mainly larger than 150 mm/km) and a w smaller than 25 km. This explains why the results using a RRAIM and a GGM simultaneously improve significantly with respect to either monitor alone.

It is worth noting that the analysis of fronts with $w < 25$ km (which cause most of the unavailability) correspond to a very specific case in which a front larger than 25 km w can only be located in a position extending from the ground facility and away from the user, as physical fronts with $w < 25$ km are not considered in the threat space. The analysis has been conservative, assuming this particular front is static, standing 'on top' of the ground facility, without taking any credit for the fact that as the static front was forming, changes in ionospheric delay could trigger the CCDM. For larger fronts (bigger w) this assumption is correct, as the front could be unobservable from the ground (lying between the airport and the aircraft, but no part of it on over the airport). But a $w < 25$ km implies the front must be observable from the ground. So it could be that for some cases that cause unavailability for the RRAIM monitor, CCDM will detect the front as it forms.



a. Complete Threat Space



b. Reduced Threat Space ($w < 30$ km)

Figure 3.12. Unavailability Caused by RRAIM Implementation for Each $f(w,k)$

3.5 Chapter Conclusions

In this chapter, the LSR RAIM implementation was expanded to a novel application called Relative RAIM (RRAIM) using time-differential carrier phase measurements.

The RRAIM implementation was applied to solve a specific threat within the context of GBAS: the static ionospheric front. It requires the broadcast of a carrier phase measurement and/or correction from the ground. The analysis in this chapter assumes the front affects one satellite at a time. Multiple satellite failures are the focus of future work.

In this implementation the RRAIM monitor is used for failure detection using the baseline formed by the aircraft as it approaches the runway. The measurements used for failure detection (time differenced carrier phase) are different form the ones used for positioning; creating some challenges that were resolved through analysis and practical implementations in the algorithms.

This work provided the foundations for more universal RRAIM applications (two of which are developed in chapters 4 and 5).

For the specific static ionospheric front detection application, the method of analysis developed provides insight into what ionospheric front shapes are more hazardous during an approach, which simplifies the analysis of total mitigation of the threat space using different monitors.

Simulation results for selected airport locations show that the threat can be mitigated with a tolerable availability loss if it is used in conjunction with an ionospheric storm external-alarm system and a ground gradient monitor at the GBAS ground facility.

CHAPTER IV

GEAS RELATIVE RAIM ARCHITECTURE

The concept of Relative Receiver Autonomous Integrity Monitoring using time differential carrier phase measurements was presented in section 2.7. In section 2.6, the GEAS activity was introduced with a brief description of its candidate navigation architectures. In this chapter, the concepts, algorithms, and error bounds for the GEAS 'RRAIM' architecture are developed in detail. The implementations presented solve the main problem of the GIC based architecture to provide worldwide LPV-200 coverage, which is meeting the time to alert requirement using a sparse worldwide network of ground stations.

The methods described in this chapter for fault detection and performance evaluation are valid for any error model and a wide variety of navigation operations. However given the importance of the GEAS study, the results shown are based on an LPV-200 operation [see Appendix D] and use the GEAS error models [see Appendix A].

In the RRAIM navigation architecture, an initial position estimate is obtained using data whose integrity is validated by the GIC. Because there is a delay between the epoch when corrections are computed by the GIC and the moment the user receives them, the aircraft will first use stored carrier-smoothed code measurements, corresponding to the time of generation of the last received GIC correction, to obtain a 'reference' position $\mathbf{x}_0$. The integrity of the reference position is therefore ensured by the GIC. A relative vector is then computed using time-differential carrier phase measurements and added to $\mathbf{x}_0$ to determine the current position. The integrity of this relative vector is provided by a

RRAIM test. The duration of the differential time interval is referred to as the carrier 'Coasting Time' (*CT*).

As noted earlier, the principal advantage of the RRAIM architecture is that it allows for significant relaxation of the GIC TTA requirement. This is true because the RRAIM function ensures navigation integrity after the latest available GIC correction. In the airborne realization of the concept, there are two interesting implications to consider: the effect of increasing *CT* on nominal differential ranging measurement errors (some error sources will increase as the coasting time gets larger), and the potential advantages of looking for the 'best' reference epoch rather than using the latest available one (reference position error depends heavily on satellite geometry).

The focus of this chapter is to provide a detailed mathematical development of the GEAS RRAIM concept and to derive the associated algorithms for positioning, fault detection, and position-domain protection level bounding. The RRAIM fault detection function is based on the weighted least squares residual, developed with two fundamentally different approaches for carrier phase coasting. The first one was introduced in the last paragraph of the previous page, and will be referred to as the 'position domain' implementation. An alternative 'range domain' implementation is also presented in this chapter, which differs in the way the available measurements at the reference and current positions are merged, as will become clear in section 4.1. Each of these two implementations has its own advantages, as will be shown later on. The range domain implementation is more straightforward to derive, and its position estimation error is independent of changes in SV geometry during coasting. The position domain

implementation allows the use of satellites visible at the reference time, even if tracking is lost during the coasting time.

All the position error bounds derived in this work correspond to the case in which the monitors do not trigger an alarm (which is the only case that an integrity threat can be present). In case of an alarm, a failure is assumed to have occurred, and the approach is aborted with no isolation being attempted. The monitor's threshold is computed to guarantee that continuity requirements are met as explained in subsection 2.5.1. The rare nature of the failures being monitored implicitly guarantees that real alarms will not violate the continuity requirements of the system.

As explained in section 2.2, the aircraft must evaluate, before starting an operation, the capabilities of its monitors to avoid HMI events. The necessary condition to start an operation can be written as

$$P_{HMIreq} = \mathrm{P}\left\{\left(\left|\delta\mathbf{x}_p\right| > AL\right) \cap \left(r < T\right)\right\} \tag{4.1}$$

where

$\delta\mathbf{x}_p$ is the error in the user position estimation ,

AL is the Alert Limit, (minimum position error magnitude considered hazardous),

r is the residual generated by the RRAIM monitor, and

T is the monitor threshold.

An equivalent way to implement inequality (4.1) is to derive a Protection Level (PL) such that

$$P_{HMIreq} = \mathrm{P}\left\{\left(\left|\delta\mathbf{x}_p\right| > PL\right) \cap \left(r < T\right)\right\} \tag{4.2}$$

and verify that

$$PL < AL \tag{4.3}$$

Potential contributors to P_{HMI}, or equivalently PL, can include fault-free errors (due for example, to an 'unlucky' combination of the nominal errors from all ranging sources) or measurement faults. In this work, it is assumed that the GIC provides a fault detection and removal rate sufficient to ensure that the probability that a user is exposed to multiple, simultaneous failed ranging sources is negligible.

The RRAIM test is responsible for the integrity from the last GIC correction time to the current time, which we have defined as the coasting time. During the CT interval, two situations (mentioned above) are possible: there is a Fault during Coasting (FC), or that there is fault free, or Normal Coasting (NC).

These two events are mutually exclusive and exhaustive. Because the random parts of $\delta\mathbf{x}_p$ and r are independent, Eq. (4.2) can be written as:

$$\begin{aligned} &\mathrm{P}\left(|\delta\mathbf{x}_p| > PL \middle| FC\right)\mathrm{P}\left(r < T \middle| FC\right)\mathrm{P}(FC) \\ &+ \mathrm{P}\left(|\delta\mathbf{x}_p| > PL \middle| NC\right)\mathrm{P}\left(r < T \middle| NC\right)\mathrm{P}(NC) = P_{HMIreq} \end{aligned} \tag{4.4}$$

Ideally we would want to compute one PL that will satisfy Eq. (4.4). However, this is difficult in practice because it typically requires an iterative process that can be very time consuming. A more conservative but practical approach will be to find two PL values for each hypothesis separately. To do this, the overall risk requirement P_{HMIreq} can be sub-allocated into separate components for the two hypotheses, $P_{HMIreq(FC)}$ and $P_{HMIreq(NC)}$, such that $P_{HMIreq} = P_{HMIreq(FC)} + P_{HMIreq(NC)}$. Then the protection level under the NC hypothesis is defined by

$$\begin{aligned} &\mathrm{P}\left(|\,\delta\,\mathbf{x}_p\,| > PL_{NC} \middle|\; NC \right) \mathrm{P}\left(r < T \middle| NC\right) \mathrm{P}(NC) \\ &\approx \mathrm{P}\left(|\,\delta\,\mathbf{x}_p\,| > PL_{NC} \middle|\; NC \right) = P_{HMIreq(NC)} \end{aligned} \tag{4.5}$$

where it has been conservatively assumed that $\mathrm{P}\left(r < T \,\middle|NC\right) \mathrm{P}(NC) \approx 1$. Similarly, for the FC hypothesis,

$$\mathrm{P}\left(|\,\delta\mathbf{x}_p\,| > PL_{FC} \middle| FC\right) \mathrm{P}\left(r < T \middle| FC\right) \mathrm{P}(FC) = P_{HMIreq(FC)} \tag{4.6}$$

Inequality (4.3) can then be re-expressed as:

$$PL = \max\left(PL_{NC}, PL_{FC}\right) < AL \tag{4.7}$$

Note that an incorrect allocation of P_{HMIreq} between $P_{HMIreq(NC)}$ and $P_{HMIreq(FC)}$ is not an integrity risk, but it impacts the availability of the system by conservatively making either PL_{NC} or PL_{FC} larger than needed (for the correct allocation $PL_{NC} = PL_{FC}$).

In the following development, we will provide a conservative and practical way to compute the PL. The right-hand sides of equations (4.5)-(4.7) are derived from system integrity requirements, and the detection threshold T on the left-hand side of Eq. (4.6) is derived from the system continuity risk requirement. The necessary intermediate steps toward computing the PL are to statistically describe the RRAIM residual r and the position error $\delta\mathbf{x}_p$.

The RRAIM architecture uses two sets of user-satellite ranges (to be described in detail shortly): $\mathbf{z}_\phi$, which is the vector of time-differenced carrier phase measurements between two epochs of interest, and $\mathbf{z}_c$, which is the vector of carrier smoothed code

measurements at a single time. All measurements in this chapter are the ionospheric delay-free measurements introduced in subsection 2.2.3.3.

Based on these measurements, two different RRAIM architecture implementations will be developed. The first one adds $\mathbf{z}_{c0}$ ($\mathbf{z}_c$ at time '0') and $\mathbf{z}_\phi$ (time difference carrier ranges between the current time and time 0) in the range domain to produce a current ranging measurement. This measurement is then used to obtain the current user position and clock bias state vector estimate $\hat{\mathbf{x}}$. The second implementation uses $\mathbf{z}_{c0}$ to obtain $\hat{\mathbf{x}}_0$, an estimate of the state at time 0, and then uses $\mathbf{z}_\phi$ to obtain a relative state estimate $\Delta\hat{\mathbf{x}}$. These two state estimates will then be added together to obtain $\hat{\mathbf{x}}$, the state estimate at the current time. In this case, the information is combined in the position domain. Because of satellite motion, and the possibility of satellite acquisition or loss during the coasting interval, we will see that the two methods do not provide equivalent performance.

4.1 Range Domain Implementation

4.1.1 Range Domain Implementation Measurements. For each SV the user will have a GIC-corrected carrier smoothed code measurement (from here on referred to as simply the 'code measurement') at a past epoch '0'. The corrected measurement for SV i can be expanded as:

$$z_{C0}^{*i} = r_0^i + \tau_0 + v_{C0}^{*i} + b_0^i = \mathbf{e}_0^{i\,\mathrm{T}}\left(\mathbf{x}_{SV0}^i - \mathbf{x}_{p0}\right) + \tau_0 + v_{C0}^{*i} + b_0^i \qquad (4.8)$$

where:

r_0^i is the actual distance between the user and SV,

$\mathbf{e}_0^l$ is the user-SV line of sight unit vector,

$\mathbf{x}_{SV0}^l$ is the true SV position,

$\mathbf{x}_{p0}$ is the true user position,

τ_0 is the receiver clock bias (expressed in units of distance),

v_{C0}^{*l} represents the sum of SV clock error ($v_{sv\tau 0}^l$) after the GIC corrections have been applied, residual tropospheric delay (v_{trop0}^l) after model-based correction, and multipath and receiver noise for code ($v_{Cmpn\ 0}^l$), and

b_0^l is included to account for two additional potential sources of measurement error: (1) a satellite measurement fault small enough to go undetected by the GIC, and (2) nominal errors that cannot be easily modeled or bounded by Gaussian distributions, and will instead be characterized by bounds on their potential magnitudes.

Note that even if there is a potential small fault in b_0^l, it is a fault from the point of view of the GIC, which is responsible for detecting it. In other words, only a failure occurring during carrier coasting distinguishes between NC and FC events. Nevertheless the effect of b_0^l must be accounted for. To avoid allocating our tolerable integrity risk between the GIC and RRAIM detection functions, we allocate P_{HMIreq} entirely to the RRAIM monitoring and introduce a bound β^l on the magnitude of b_0^l such that the

probability that $\left|b_0^i\right| > \beta^i$ is negligibly small compared with P_{HMIreq}, which may be mathematically expressed as:

$$\sum_i \mathrm{P}\left(\left| b_0^i \right| > \beta^i\right) << P_{HMIreq} \tag{4.9}$$

The choice of bound β^i will depend primarily on the integrity monitor capabilities of the GIC.

We will now re-express our measurement in Eq. (4.8) as

$$z_{C0}^i = z_{C0}^{*i} - \mathbf{e}_0^{i\,\mathrm{T}} \hat{\mathbf{x}}_{SV0}^i = -\mathbf{e}_0^{i\,\mathrm{T}} \mathbf{x}_{p0} + \tau_0 + v_{C0}^i + b_0^i \tag{4.10}$$

where $\hat{\mathbf{x}}_{SV0}^i$ is the ephemeris generated (and GIC corrected) position of the satellite. The error in the term $\mathbf{e}_0^{i\,\mathrm{T}} \hat{\mathbf{x}}_{SV0}^i$ is

$$v_{eph0}^i = \mathbf{e}_{C0}^{i\,\mathrm{T}} \delta \mathbf{x}_{SV0}^i, \tag{4.11}$$

and, for compactness of notation, it is now included in v_{C0}^i:

$$v_{C0}^i = v_{Cmpn0}^i + v_{sv\tau 0}^i + v_{trop0}^i + v_{eph0}^i \tag{4.12}$$

Note that terms in Eq. (4.12) are the errors after all GIC corrections have been applied.

The user will also have a stored carrier phase measurement from epoch 0, for any SV i:

$$z_{\phi 0}^{**i} = l_0^i + \tau_0 + N^i + v_{\phi 0}^{*i} + f_0^{*i} = \mathbf{e}_0^{i\,\mathrm{T}}\left(\mathbf{x}_{SV0}^i - \mathbf{x}_{p0}\right) + \tau_0 + N^i + v_{\phi 0}^{*i} + f_0^{*i} \tag{4.13}$$

where:

N^i is a bias that includes the carrier phase cycle ambiguity (expressed in units of distance),

$\nu_{\phi 0}^{*i}$ represents the sum of SV clock error ($\nu_{sv\tau 0}^{i}$) after the GIC corrections have been applied, residual tropospheric delay (ν_{trop0}^{i}) after model-based correction, and multipath and receiver noise for the carrier phase measurement ($\nu_{\phi mpn0}^{i}$), and

f_0^{*i} is a potential failure affecting the measurement.

As we did for the code measurement, we re-express our carrier phase measurement at 0 as

$$z_{\phi 0}^{*\,i} = z_{\phi 0}^{**\,i} - \mathbf{e}_0^{i\,\mathrm{T}} \hat{\mathbf{x}}_{SV0}^{\,i} = -\mathbf{e}_0^{i\,\mathrm{T}} \mathbf{x}_{p0} + \tau_0 + N^i + \nu_{\phi 0}^i + f_0^{*i} \tag{4.14}$$

where, as in equations (4.10) and (4.12), the ephemeris error is now implicitly included in $\nu_{\phi 0}$. We then do the same for the current time to obtain

$$z_{\phi}^{*\,i} = -\mathbf{e}^{i\,\mathrm{T}} \mathbf{x}_p + \tau + N^i + \nu_{\phi}^i + f^{*i} \tag{4.15}$$

where

$$\nu_{\phi}^i = \nu_{\phi mpn}^i + \nu_{\phi sv\tau}^i + \nu_{\phi trop}^i + \nu_{\phi eph}^i \tag{4.16}$$

The time-differential carrier phase measurement that will be used in the RRAIM system is

$$z_{\phi}^i = z_{\phi}^{*i} - z_{\phi 0}^{*i} = -\mathbf{e}^{i\,\mathrm{T}} \mathbf{x}_p + \mathbf{e}_r^{i\,\mathrm{T}} \mathbf{x}_{p0} + \Delta\tau + \Delta\nu_{\phi}^i + f^i \tag{4.17}$$

where

$$\Delta\tau = \tau - \tau_r; \quad \Delta\nu_{\phi}^i = \nu_{\phi}^i - \nu_{\phi 0}^i; \quad f^i = f^{*i} - f_0^{*i}. \tag{4.18}$$

We can now define our RD basic measurement at the current time from equations. (4.10) and (4.17) as:

$$\begin{aligned} z_{RD}^{l} &= z_{C0}^{l} + z_{\phi}^{l} = -\mathbf{e}^{l\,\mathrm{T}}\mathbf{x}_p + \tau + v_{C0}^{l} + \Delta v_{\phi}^{l} + b_0^{l} + f^{l} \\ &= \mathbf{h}^{l\,\mathrm{T}}\mathbf{x} + v_{C0}^{l} + \Delta v_{\phi}^{l} + b_0^{l} + f^{l} \end{aligned} \tag{4.19}$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_p & \tau \end{bmatrix}^{\mathrm{T}}; \qquad \mathbf{h}^{l} = \begin{bmatrix} -\mathbf{e}^{l\,\mathrm{T}} & 1 \end{bmatrix}^{\mathrm{T}}. \tag{4.20}$$

The state vector (position and time) estimate will then be obtained from

$$\hat{\mathbf{x}}_{RD} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{R}_{RD}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}_{RD}^{-1}\mathbf{z}_{RD}, \tag{4.21}$$

where

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}^{1} & \cdots & \mathbf{h}^{n} \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{z}_{RD} = \begin{bmatrix} z_{RD}^{1} & \cdots & z_{RD}^{n} \end{bmatrix}^{\mathrm{T}} \tag{4.22}$$

and n is the number of SVs whose signals have been continuously tracked by the user between time 0 and the current time. $\mathbf{R}_{RD}$ is the covariance matrix of the errors in Eq. (4.19) that can be bounded by Gaussian distributions: $v_{C0}^{l} + \Delta v_{\phi}^{l}$ $(i = 1, \ldots, n)$. These will be defined in detail in the next subsection. Distributions of b_0^{l} are unlikely to be available, so the Gaussian weighting is used here. Nevertheless, the effects of b_0^{l} on position error must be accounted for, and this will be addressed shortly as well.

4.1.2 Range Domain Implementation Covariance Matrices. In this section we will explicitly define the elements of the measurement error covariance matrix $\mathbf{R}_{RD}$ introduced in Eq. (4.21).

Assuming all error components are independent and that the receiver errors are not correlated in time; the diagonal elements of $\mathbf{R}_{RD}$, from equations (4.12), (4.16), (4.18) and (4.19) can be expressed as:

$$\begin{aligned} R_{RD(i,i)} &= \mathrm{E}\left[\left(\nu_{C0}^{i} + \Delta\nu_{\phi}^{i}\right)^2\right] \\ &= \left(\sigma_{Cmpn0}^{i}\right)^2 + \left(\sigma_{\Delta\phi mpn}^{i}\right)^2 + \left(\sigma_{sv\tau}^{i}\right)^2 + \left(\sigma_{trop}^{i}\right)^2 + \left(\sigma_{eph}^{i}\right)^2 - 2\mathrm{E}[\nu_{Cnpm0}^{i}\nu_{\phi npm0}^{i}] \quad (4.23) \\ &\approx \left(\sigma_{Cmpn0}^{i}\right)^2 + \left(\sigma_{\Delta\phi mpn}^{i}\right)^2 + \left(\sigma_{sv\tau}^{i}\right)^2 + \left(\sigma_{trop}^{i}\right)^2 + \left(\sigma_{eph}^{i}\right)^2 \end{aligned}$$

where the last term has been conservatively eliminated assuming a positive correlation between the multipath and noise terms for code and carrier. Assuming errors from different ranging sources are independent:

$$R_{RD(i,j)} = 0 \quad (4.24)$$

Representative values for the standard deviations in Eq. (4.23) can be found in [GEA10], some of which are summarized in Appendix A.

4.1.3 Range Domain Implementation Estimation Errors. From Eq. (4.21) and using the notation introduced in Eq. (2.6):

$$\delta\mathbf{x}_{RD} = \mathbf{Q}_{RD}\mathbf{H}^{\mathrm{T}}\mathbf{R}_{RD}^{-1}\delta\mathbf{z}_{RD} \quad (4.25)$$

Now defining the vectors:

$$\mathbf{b}_V = \left[b_0^1 \cdots b_0^n\right]^{\mathrm{T}}, \mathbf{v}_{C0} = \left[\nu_{C0}^1 \cdots \nu_{C0}^n\right]^{\mathrm{T}}, \Delta\mathbf{v}_{\phi} = \left[\Delta\nu_{\phi}^1 \cdots \Delta\nu_{\phi}^n\right]^{\mathrm{T}}, \mathbf{f} = \left[f^1 \cdots f^n\right]^{\mathrm{T}} \quad (4.26)$$

and using Eq. (4.19); Eq. (4.25) can be expanded as:

$$\delta\mathbf{x}_{RD} = \mathbf{Q}_{RD}\mathbf{H}^{\mathrm{T}}\mathbf{R}_{RD}^{-1}\left[\mathbf{v}_{C0} + \Delta\mathbf{v}_{\phi} + \mathbf{b}_V + \mathbf{f}\right] = \mathbf{S}_{RD}\left[\mathbf{v}_{C0} + \Delta\mathbf{v}_{\phi} + \mathbf{b}_V + \mathbf{f}\right] \quad (4.27)$$

After introducing the PD implementation, the position error bounds will be defined for both methods.

4.2 Position Domain Implementation

4.2.1 Position Domain Implementation Measurements. From equations (4.10) and (4.20), we may write

$$z_{C0}^{l} = \mathbf{h}_0^{l\,\mathrm{T}}\mathbf{x}_0 + v_{C0}^{l} + b_0^{l} \tag{4.28}$$

and the initial position and clock bias estimate can then be obtained as:

$$\hat{\mathbf{x}}_0 = \left(\mathbf{H}_0^{\mathrm{T}}\mathbf{R}_{C0}^{-1}\mathbf{H}_0\right)^{-1}\mathbf{H}_0^{\mathrm{T}}\mathbf{R}_{C0}^{-1}\mathbf{z}_{C0} \tag{4.29}$$

where

$$\mathbf{z}_{C0} = \begin{bmatrix} z_{C0}^{1} & \cdots & z_{C0}^{m} \end{bmatrix}^{\mathrm{T}} \tag{4.30}$$

and m is the number of SVs available at time 0, and the rows of $\mathbf{H}_0$ are $\mathbf{h}_0^{l\,\mathrm{T}}$ for each satellite. $\mathbf{R}_{C0}$ is the covariance matrix of the Gaussian errors (defined in Eq. (4.12)) in Eq. (4.28).

From equations (4.17), (4.18) and (4.20):

$$z_{\phi}^{i} = \mathbf{h}^{i\,\mathrm{T}}\mathbf{x} - \mathbf{h}_0^{i\,\mathrm{T}}\mathbf{x}_0 + \Delta\mathbf{v}_{\phi}^{i} + f^{i} = \mathbf{h}^{i\,\mathrm{T}}\Delta\mathbf{x} - \Delta\mathbf{h}^{i\,\mathrm{T}}\mathbf{x}_0 + \Delta\mathbf{v}_{\phi}^{i} + f^{i} \tag{4.31}$$

where

$$\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_0, \qquad \Delta\mathbf{h}^{i} = \mathbf{h}^{i} - \mathbf{h}_0^{i} \tag{4.32}$$

Using Eq. (4.31) and the initial state estimate from Eq. (4.29), we now define the time-differenced carrier phase measurement for the PD implementation:

$$z^i_{PD\phi} = z^i_\phi + \Delta\mathbf{h}^{i\,\mathrm{T}}\hat{\mathbf{x}}_0 = \mathbf{h}^{i\,\mathrm{T}}\Delta\mathbf{x} + \Delta v^i_{PD\phi} + f^{\,i} \tag{4.33}$$

where the error in the initial state estimate $\hat{\mathbf{x}}_0$ contributes through satellite Line Of Sight (LOS) changes as:

$$\Delta v^i_{\Delta LOS} = \Delta\mathbf{h}^{i\,\mathrm{T}}\delta\mathbf{x}_0 \ , \tag{4.34}$$

and is included in $\Delta v^i_{PD\phi}$ in Eq. (4.33):

$$\Delta v^i_{PD\phi} = \Delta v^i_\phi + \Delta v^i_{\Delta LOS} \tag{4.35}$$

4.2.2 Position Domain Implementation Covariance Matrices. In this section we will explicitly define the elements of the measurement error covariance matrices $\mathbf{R}_{C0}$ and $\mathbf{R}_{PD\phi}$. These are needed for the PD positioning algorithms defined above, as well as for RRAIM residual generation.

To obtain the elements of $\mathbf{R}_{C0}$ in Eq. (4.29) we assume that all ranging sources and component sources have independent errors. From Eq. (4.12):

$$R_{C0(i,i)} = \mathrm{E}\left[{v^i_{C0}}^2\right] = \left(\sigma^i_{C0}\right)^2 \ , \quad R_{C0(i,j)} = 0 \tag{4.36}$$

$$\left(\sigma^i_{C0}\right)^2 = \left(\sigma^i_{Cmpn0}\right)^2 + \left(\sigma^i_{C\tau 0}\right)^2 + \left(\sigma^i_{CTropo0}\right)^2 + \left(\sigma^i_{Ceph0}\right)^2 \tag{4.37}$$

For the position domain implementation, from equations (4.16) (4.18) and (4.31) to (4.35):

$$R_{PD\phi(i,i)} = \mathrm{E}\left[\left(\Delta v^i_\phi + v^i_{\Delta LOSg}\right)^2\right] = \left(\sigma^i_{\Delta\phi}\right)^2 + \mathrm{E}\left[\left(v^i_{\Delta LOSg}\right)^2\right] = \left(\sigma^i_{\Delta\phi}\right)^2 + \Delta\mathbf{h}^{i\,\mathrm{T}}\mathbf{Q}_{C0}\Delta\mathbf{h}^i \tag{4.38}$$

where 'g' in the $v^{i}_{\Delta LOSg}$ subscript means that only the Gaussian components of the error term are considered, and

$$\left(\sigma^{i}_{\Delta\phi}\right)^2 = \left(\sigma^{i}_{\Delta\phi mpn}\right)^2 + \left(\sigma^{i}_{\Delta svr}\right)^2 + \left(\sigma^{i}_{\Delta trop}\right)^2 + \left(\sigma^{i}_{\Delta eph}\right)^2 \tag{4.39}$$

Note that in contrast to the values in Eq. (4.23) the terms in equations (4.38) and (4.39), with the exception of $\left(\sigma^{i}_{\Delta\phi mpn}\right)^2$, are a function of the *CT*. Finally, $\mathbf{Q}_{C0}$ in Eq. (4.38) is the covariance matrix of the initial state estimate error due to the Gaussian measurement error components at time 0.

For the non-diagonal elements $R_{PD\phi(i,j)}$ we can assume error sources for different SVs are independent, except for the $\mathbf{v}_{\Delta LOS}$ term, as the error in $\hat{\mathbf{x}}_0$ affects all SVs:

$$R_{PD\phi(i,j)} = \Delta\mathbf{h}^{i\,\mathrm{T}}\mathbf{Q}_{C0}\Delta\mathbf{h}^{j} \tag{4.40}$$

4.2.3 Position Domain Implementation Estimation Errors. We can now estimate the relative position vector as:

$$\Delta\hat{\mathbf{x}} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}_{PD\phi}\mathbf{H}\right)^{-1}\mathbf{H}^{T}\mathbf{R}^{-1}_{PD\phi}\mathbf{z}_{PD\phi} \tag{4.41}$$

where

$$\mathbf{z}_{PD\phi} = \left[z^{1}_{PD\phi} \quad \cdots \quad z^{n}_{PD\phi}\right]^{\mathrm{T}}, \tag{4.42}$$

and n is the number of SVs continuously tracked between time 0 and the current time, and the rows of $\mathbf{H}$ are $\mathbf{h}^{i\,\mathrm{T}}$ for each satellite i.

Using the results of equations (4.29) and (4.41), we obtain the PD implementation state vector estimate as:

$$\hat{\mathbf{x}}_{PD} = \hat{\mathbf{x}}_0 + \Delta\hat{\mathbf{x}}. \tag{4.43}$$

Note that $\hat{\mathbf{x}}_{PD}$ and $\hat{\mathbf{x}}_{RD}$ are two different estimates of the same state true vector $\mathbf{x}$. The most notable source of difference is that in the RD implementation SVs that are not present at the current time are not used at all, while in the PD implementation all SVs present at epoch 0 are used to obtain $\hat{\mathbf{x}}_0$. A second point of difference is that the error terms $v^{i}_{\Delta LOS}$ do not appear in the RD implementation.

From equations (4.10), (4.29), (4.30), (4.33)-(4.35) and (4.41)-(4.43):

$$\begin{aligned}
\delta\mathbf{x}_{PD} &= \delta\mathbf{x}_0 + \delta\Delta\mathbf{x} = \mathbf{S}_{C0}\delta\mathbf{z}_{C0} + \mathbf{S}_{PD\phi}\,\delta\mathbf{z}_{PD\phi} \\
&= \left(\mathbf{S}_{C0} + \mathbf{S}_{PD\phi}\Delta\mathbf{H}\,\mathbf{S}_{C0}\right)\mathbf{v}_{C0} + \mathbf{S}_{PD\phi}\Delta\mathbf{v}_{PD\phi} \\
&\quad + \left(\mathbf{S}_{C0} + \mathbf{S}_{PD\phi}\Delta\mathbf{H}\,\mathbf{S}_{C0}\right)\mathbf{b}_V + \mathbf{S}_{PD\phi}\mathbf{f} \\
&= \mathbf{B}\mathbf{v}_{C0} + \mathbf{S}_{PD\phi}\Delta\mathbf{v}_{PD\phi} + \mathbf{B}\mathbf{b}_V + \mathbf{S}_{PD\phi}\mathbf{f}
\end{aligned} \tag{4.44}$$

where $\mathbf{S}_{C0} = \mathbf{Q}_{C0}\mathbf{H}^{\mathrm{T}}\mathbf{R}_{C0}^{-1}$, $\mathbf{S}_{PD\phi} = \mathbf{Q}_{PD\phi}\mathbf{H}^{\mathrm{T}}\mathbf{R}_{PD\phi}^{-1}$, and $\Delta\mathbf{H} = -\left[\Delta\mathbf{h}^1 \cdots \Delta\mathbf{h}^n\right]$. The matrix $\mathbf{B}$ has been introduced for the sole purpose of simplifying the notation in the following section, and its definition is obvious from equation (4.44).

4.2.4 Protection Levels. Now the algorithms to obtain four protection levels are developed, satisfying equations (4.5) and (4.6) for the NC and FC hypotheses and each of the two implementations presented: $PL_{NC(RD)}$, $PL_{NC(PD)}$, $PL_{FC(RD)}$ and $PL_{FC(PD)}$.

The starting point to generate the protection levels are the position error formulas (4.27) and (4.44). Each of these formulas has terms originated by errors that can be

modeled as Gaussian, terms originated by non-Gaussian errors $\mathbf{b}_V$, and for the FC cases, a term caused by failure f.

The distribution of $\mathbf{b}_V$ is unknown, but it is assumed known that the probability of any $|b_0^l| > \beta^l$ is negligible (Eq. (4.9)). Assuming further that each b_0^l is independent of any other error source from the same (or different) satellite, we can conservatively bound the effect of non-Gaussian errors in the RD case by

$$\mathbf{S}_{RD}\mathbf{b}_V \leq |\mathbf{S}_{RD}||\mathbf{b}_V| \leq |\mathbf{S}_{RD}|\boldsymbol{\beta}_V \tag{4.45}$$

where

$$\boldsymbol{\beta}_V = \begin{bmatrix}\beta^1 & \cdots & \beta^n\end{bmatrix}^{\mathrm{T}} \tag{4.46}$$

and the notation $|\bullet|$ denotes element-wise absolute value operations for a vector or a matrix (not a determinant). Similarly, for the PD case (4.44):

$$\mathbf{B}\mathbf{b}_V \leq |\mathbf{B}||\mathbf{b}_V| \leq |\mathbf{B}|\boldsymbol{\beta}_V \tag{4.47}$$

To compute the NC protection levels we must then add the effect of Gaussian errors terms. For a hypothetical zero mean Gaussian random variable y with standard deviation σ, we first compute an integrity factor k_{int} such that

$$\mathrm{P}\left(|y| > k_{int}\sigma\right) = P_{HMIreq(NC)}\,. \tag{4.48}$$

Then we compute the position error standard deviation for the Gaussian error terms. For the RD implementation from equations (2.6) and (4.23)-(4.24):

$$\mathrm{E}\left[\delta\mathbf{x}_{RDg}\delta\mathbf{x}_{RDg}^{\mathrm{T}}\right] = \mathbf{Q}_{RD} \tag{4.49}$$

We will write the final results in terms of the Vertical Protection Level (VPL); from Eq's (4.27), (4.45), (4.48) and (4.49):

$$VPL_{NC(RD)} = k_{int}\sqrt{Q_{RD(3,3)}} + \left(\left|\mathbf{S}_{RD}\right|\boldsymbol{\beta}_V\right)_{(3,1)} \tag{4.50}$$

Computing the position error standard deviation for the Gaussian errors terms in the PD implementation is slightly more complicated. For the PD implementation, from Eq. (4.44):

$$\begin{aligned}
\mathbf{Q}_{PD} &\equiv \mathrm{E}\left[\delta\mathbf{x}_{PDg}\delta\mathbf{x}_{PDg}^{\mathrm{T}}\right] \\
&= \mathrm{E}\left[\left(\delta\mathbf{x}_{0g} + \delta\Delta\mathbf{x}_g\right)\left(\delta\mathbf{x}_{0g} + \delta\Delta\mathbf{x}_g\right)^{\mathrm{T}}\right] \\
&= \mathrm{E}\left[\left(\left(\mathbf{S}_{C0} + \mathbf{S}_{PD\phi}\Delta\mathbf{H}\,\mathbf{S}_{C0}\right)\mathbf{v}_{C0g} + \mathbf{S}_{PD\phi}\Delta\mathbf{v}_{PD\phi g}\right)\right. \\
&\qquad \left.\left(\left(\mathbf{S}_{C0} + \mathbf{S}_{PD\phi}\Delta\mathbf{H}\,\mathbf{S}_{C0}\right)v_{C0g} + \mathbf{S}_{PD\phi}\Delta\mathbf{v}_{PD\phi g}\right)^{\mathrm{T}}\right] \\
&= \mathbf{Q}_{C0} + \mathbf{S}_{C0}\mathrm{E}\left[\mathbf{v}_{C0g}\Delta\mathbf{v}_{PD\phi g}^{\mathrm{T}}\right]\mathbf{S}_{PD\phi}^{\mathrm{T}} + \mathbf{S}_{PD\phi}\mathrm{E}\left[\Delta\mathbf{v}_{PD\phi g}\mathbf{v}_{C0g}^{\mathrm{T}}\right]\mathbf{S}_{C0}^{\mathrm{T}} + \mathbf{Q}_{PD\phi}
\end{aligned} \tag{4.51}$$

We assume that the *change* in carrier phase measurement errors over the CT is uncorrelated from the initial code phase errors at time 0. Therefore, using equations (4.33) through (4.35), we know that

$$\begin{aligned}
\mathrm{E}\left[\Delta\mathbf{v}_{PD\phi g}v_{C0g}^{\mathrm{T}}\right] &= \mathrm{E}\left[\left(\Delta\mathbf{H}\delta\mathbf{x}_{0g}\right)\mathbf{v}_{C0g}^{\mathrm{T}}\right] = \mathrm{E}\left[\Delta\mathbf{H}\mathbf{Q}_{C0}\mathbf{H}_0^{\mathrm{T}}\mathbf{R}_{C0}^{-1}\mathbf{v}_{C0g}\mathbf{v}_{C0g}^{\mathrm{T}}\right] \\
&= \Delta\mathbf{H}\mathbf{Q}_{C0}\mathbf{H}_0^{\mathrm{T}}\mathbf{R}_{C0}^{-1}\,\mathrm{E}\left[\mathbf{v}_{C0g}\mathbf{v}_{C0g}^{\mathrm{T}}\right] = \Delta\mathbf{H}\mathbf{Q}_{C0}\mathbf{H}_0^{\mathrm{T}}
\end{aligned} \tag{4.52}$$

Substituting this result into Eq. (4.51), we obtain the covariance matrix for the Gaussian position error for the PD implementation:

$$\mathbf{Q}_{PD} = \mathbf{Q}_{C0} + \mathbf{Q}_{C0}\Delta\mathbf{H}^{\mathrm{T}}\mathbf{S}_{PD\phi}^{\mathrm{T}} + \mathbf{S}_{PD\phi}\Delta\mathbf{H}\,\mathbf{Q}_{C0} + \mathbf{Q}_{PD\phi} + \mathbf{S}_{PD\phi}\Delta\mathbf{H}\,\mathbf{Q}_{C0}\Delta\mathbf{H}^{\mathrm{T}}\mathbf{S}_{PD\phi}^{\mathrm{T}} \tag{4.53}$$

Using the results from equations (4.47) and (4.53) the resulting NC VPL equation is:

$$VPL_{NC(PD)} = k_{int}\sqrt{Q_{PD(3,3)}} + \left(|\mathbf{B}|\boldsymbol{\beta}_V\right)_{(3,1)} \tag{4.54}$$

The RRAIM test statistic (in both implementations) will be computed to detect a potential failure [in equation (4.19) or (4.33)] occurring during the coasting period:

$$r^2 = \left(\mathbf{z}_{PD\phi} - \mathbf{H}\Delta\hat{\mathbf{x}}\right)^{\mathrm{T}} \mathbf{R}_{PD\phi}^{-1}\left(\mathbf{z}_{PD\phi} - \mathbf{H}\Delta\hat{\mathbf{x}}\right) \tag{4.55}$$

If all the errors in $\mathbf{z}_{PD\phi}$ (i.e., $\Delta v_{PD\phi}^{l} = \Delta v_{\phi}^{l} + v_{\Delta LOS}^{l}$) were Gaussian, r^2 would be Chi Square distributed with n-4 DOF. However, the non-Gaussian term b_0^l, defined in Eq. (4.10), can lead to a non-Gaussian contribution through $v_{\Delta LOS}^{l}$. The effect is accounted for later in the chapter, where it will be shown that the actual distribution of r^2 can be bounded by a non-central Chi Square distribution.

To generate the FC protection levels, we need to identify the worst failure size and SV combination that could go undetected (meaning no monitor alarm, $r < T$). Doing this precisely is a time consuming iterative process. A conservative but more practical approach is used here instead. For each SV i, a fault magnitude f_+^l is found such that

$$\mathrm{P}\left(r < T \middle| f_+^l\right) P(FC) = P_{HMIreq(FC)} \tag{4.56}$$

with

$$r^2 = \left(\mathbf{z}_{PD\phi}^{+} - \mathbf{H}\Delta\hat{\mathbf{x}}^{+}\right)^{\mathrm{T}} \mathbf{R}_{z_{PD\phi}}^{-1}\left(\mathbf{z}_{PD\phi}^{+} - \mathbf{H}\Delta\hat{\mathbf{x}}^{+}\right), \tag{4.57}$$

$$\mathbf{z}_{PD\phi}^{+} = \mathbf{z}_{PD\phi} + \mathbf{f}_V^l, \tag{4.58}$$

where $\mathbf{f}_V^l$ is a vector, that has as its only non-zero element the value of f_+^l for that satellite:

$$\mathbf{f}_V^l = \begin{bmatrix} 0 & \cdots & f_+^l & \cdots & 0 \end{bmatrix}^{\mathrm{T}} \tag{4.59}$$

and $\Delta\hat{\mathbf{x}}^+$ is the estimate $\Delta\hat{\mathbf{x}}$ obtained using $\mathbf{z}_{PD\phi}^+$.

The resulting VPL must account for the impact in the position domain of the various fault vectors $\mathbf{f}_V^l$ and also the nominal measurement errors. Furthermore, a fault on the worst case SV, (i.e., the satellite fault causing the worst integrity impact) is used to define the FC protection levels. The method for obtaining the associated worst-case fault magnitude f_+^l will be discussed shortly. The protection levels for the RD and PD implementations are, respectively,

$$VPL_{FC(RD)} = \max_l \left(\mathbf{S}_{RD}\,\mathbf{f}_V^l\right)_{(3,1)} + k_{intFC}\sqrt{Q_{RD(3,3)}} + \left(|\mathbf{S}_{RD}|\boldsymbol{\beta}_V\right)_{(3,1)} \tag{4.60}$$

$$VPL_{FC(PD)} = \max_l \left(\mathbf{S}_{PD\phi}\,\mathbf{f}_V^l\right)_{(3,1)} + k_{intFC}\sqrt{Q_{PD(3,3)}} + \left(|\mathbf{B}|\boldsymbol{\beta}_V\right)_{(3,1)} \tag{4.61}$$

where k_{intFC} is a multiplier such that, for a zero mean Gaussian distributed random variable y with standard deviation σ:

$$\mathrm{P}\left(|y| > k_{intFDC}\,\sigma\right) = \frac{P_{HMIreq(FC)}}{P(FC)} \tag{4.62}$$

The RRAIM detection threshold T is set such that the fault-free alarm probability meets the allocated system continuity requirement (P_{Creq}) and is obtained for a fault free Non-Central Chi Square distribution:

$$P\left(r > T|\lambda_{NC}\right) = P_{Creq} \tag{4.63}$$

The non-centrality parameter, λ_{NC}, is obtained using equation (4.57), together with (4.33) through (4.35) and (4.44):

$$\lambda_{NC} = (\mathbf{VCb}_V)^{\mathrm{T}} \mathbf{R}_{z_{PD\phi}}^{-1} \mathbf{VCb}_V \tag{4.64}$$

where

$$\mathbf{V} = \mathbf{I} - \mathbf{HS}_{PD\phi} \tag{4.65}$$

where **I** is the identity matrix of the corresponding size; and

$$\mathbf{C} = \Delta\mathbf{H}\,\mathbf{S}_{C0} \tag{4.66}$$

Although $\mathbf{b}_V$ is unknown, we can obtain an upper bound on λ_{NC},

$$\lambda_{NC} \le \mu \parallel \boldsymbol{\beta}_V \parallel^2 \tag{4.67}$$

where μ is the maximum eigenvalue of $(\mathbf{VC})^{\mathrm{T}} \mathbf{R}_{PD\phi}^{-1} \mathbf{VC}$ and $\parallel \boldsymbol{\beta}_V \parallel$ is the magnitude of the bounding vector $\boldsymbol{\beta}_V$.

Note that for short *CT*, the geometry change effect $\Delta\mathbf{H}$ is generally small, and therefore μ will also be small. In these cases, depending on the magnitude of $\boldsymbol{\beta}_V$, it may be acceptable to use a (central) Chi Square distribution to compute *T*. [The results shown in section 4.3 compute *T* using a central Chi Square distribution].

Under the fault hypothesis (FC), with a fault on arbitrary satellite *i*, the RRAIM residual is Non-Central Chi Square distributed with non-centrality parameter:

$$\lambda_{FC}^{i} = (\mathbf{Cb}_V + \mathbf{f}_V^i)^{\mathrm{T}} \mathbf{V}^{\mathrm{T}} \mathbf{R}_{PD\phi}^{-1} \mathbf{V}(\mathbf{Cb}_V + \mathbf{f}_V^i). \tag{4.68}$$

To determine the worst-case fault for satellite *i*, we choose λ_{FC}^{i} (same for all satellites) such that

$$P\left(r < T \middle| \lambda_{FC}^{i}\right) = \frac{P_{HMIreq(FC)}}{P(FC)}, \tag{4.69}$$

and then find $\mathbf{f}_V^i$ with the largest magnitude that satisfies Eq. (4.69). For convenience in notation, we define a matrix

$$\mathbf{F} = \mathbf{R}_{PD\phi}^{-1/2}\mathbf{V}, \tag{4.70}$$

then from Eq.(4.68), we know that

$$\sqrt{\lambda_{FC}^i} = \left\| \mathbf{F}(\mathbf{C}\mathbf{b}_V + \mathbf{f}_V^i) \right\|. \tag{4.71}$$

From equations (4.64) and (4.67), we also know that the largest possible magnitude for $\mathbf{FCb}_V$ is $\sqrt{\lambda_{NC}}$, where λ_{NC} is assigned the bounding value on the right-hand side of Eq. (4.67). Therefore, the vector $\mathbf{f}_V^i$ with the largest magnitude that is consistent with Eq. (4.71) must satisfy

$$\sqrt{\lambda_{FC}^i} = \left\| \mathbf{F}\mathbf{f}_V^i \right\| - \sqrt{\lambda_{NC}}. \tag{4.72}$$

Recall from Eq. (4.59) that $\mathbf{f}_V^i$ describes a fault on satellite i with magnitude f_+^i. Substituting Eq. (4.59) into Eq. (4.72), the worst-case fault magnitude for satellite i is then

$$f_+^i = \frac{\sqrt{\lambda_{FC}^i} + \sqrt{\lambda_{NC}}}{\| \mathbf{F}_{\cdot,i} \|}, \tag{4.73}$$

where $\mathbf{F}_{\cdot,i}$ is the i-th column of the matrix $\mathbf{F}$. The result in Eq. (4.73), when reincorporated into Eq. (4.59), provides the input vector $\mathbf{f}_V^i$ for the protection level equations (4.60) and (4.61).

4.3 Results

4.3.1 Sample Case. An example of results for a sample location (Chicago) and the nominal GPS 24 SV GPS constellation is presented next. The specifications used, inputs to determine measurement error standard deviations, and values for the biases can be found in [GEA10], with a summary in Appendix A. Figure 4.1 shows VPL_{RD} and VPL_{PD} at each epoch for one whole day. The result using traditional RAIM (ARAIM architecture) is also plotted to show the significant improvement in availability when using an RRAIM implementation. The overall availability is obtained dividing available/total epochs, were 'available' corresponds to a VPL < 35m. For the RD implementation, the coasting time was one minute, which also served as the minimum allowable coasting time for the PD implementation.

Figure 4.1 shows VPL values are very similar for both implementations. However there is a slight availability gain from using the PD implementation. This gain is due to the fact that the PD implementation allows reaching back in time to find a better reference geometry. However, the acceptable CT is limited by the rapid growth of differential SV clock and tropospheric delay errors. Reference times were searched up to one hour before the current epoch, but the smallest VPL_{PD} was never found to be one using a reference epoch more than 3 minutes old. This can be seen in Figure 4.2, where the CT that gives the smallest VPL_{PD} for each epoch is shown.

This preliminary analysis is relevant because it allowed determination of how far back in time to search for a reference user position for the PD implementation. This is important because the coverage simulations are very computationally intensive.

4.3.2 Coverage Results. Coverage values (see subsection 2.2.5) were generated with LPV-200 operation requirements (see Appendix D) for the PD and RD implementations. In accordance to what was expressed in subsection 4.3.1, the search for a reference epoch was extended to six minutes prior to the current epoch, evaluated at intervals of 30 seconds (i.e. up to 12 reference epochs are used). Availability was computed for a whole day with a sample interval of four minutes. The constellations used are the six agreed upon by the GEAS panel for simulations [GEA08]: three nominal constellations of 24, 27 and 30 SVs, and three constellations constructed from removing one satellite from each of the nominal constellations (24-1, 27-1 and 30-1 SVs).

Two criteria are used to consider a site covered: 99.5% and 99.9% availability. The *CT* for the RD implementation is 30s. The initial purpose of this implementation was to relax the TTA requirement. Selecting a 30s *CT* is equivalent to relaxing the *TTA* to that time interval, so it is a conservative number (recall that the TTA for WAAS is close to 6 s). Accordingly the minimum *CT* for the PD implementation is also 30 s.

Three coverage results are presented for each constellation and each of the two criteria (99.5 and 99.9%): the coverage using the RD implementation, coverage using the PD with a *CT* of 30 s, and coverage using the PD implementation using the *CT* (up to 6 min long) producing the lowest VPL_{PD}.

The results in Table 4.2 agree with what was expected. Coverage is slightly better for the PD implementation even when using only a *CT=30s*, as for some cases the number of satellites in the reference geometry is higher than for the RD implementation. The gain in coverage is always small (< 1%). When the PD implementation is allowed to search back in time for a better reference geometry, then performance improves more,

sometimes by more than 3% with respect to the RD implementation. In particular for the case of the 30-1 SV constellation, 100% coverage is achieved even with the more stringent 99.9% availability criteria.

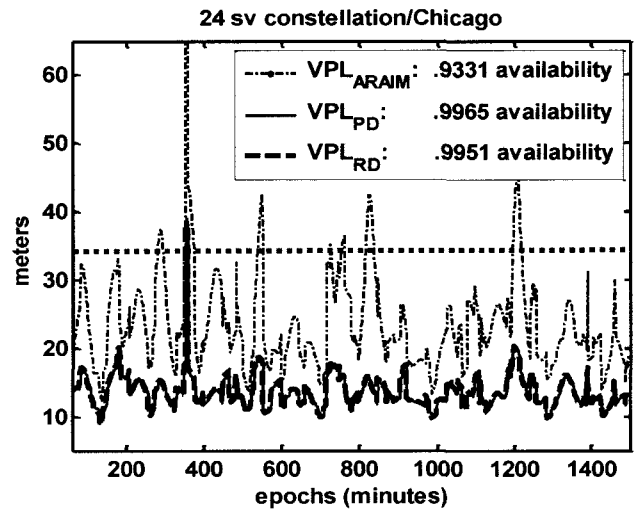


Figure 4.1. VPL with PD and RD Implementation

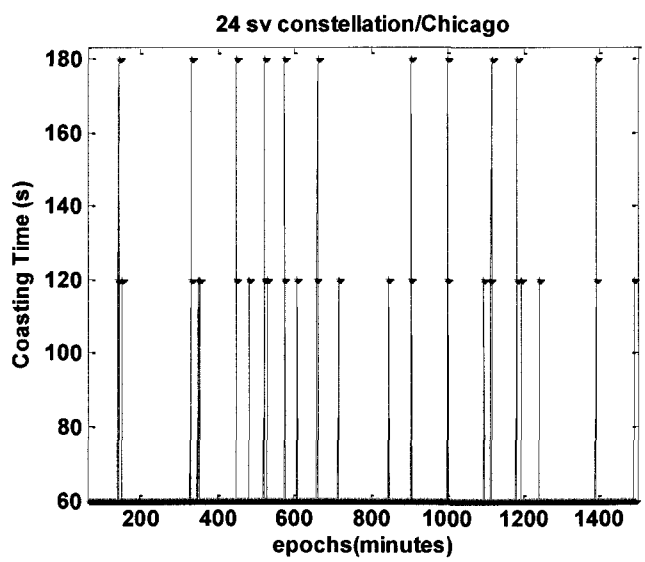


Figure 4.2. *CT* for Smallest VPL_{PD} at Each Epoch

Table 4.1 Coverage for Range Domain and Position Domain Implementations

Avail. criteria	Constellation:	24-1	24	27-1	27	30-1	30
99.5%	$RRAIM_{RD}$	78.87%	98.98%	95.56%	100%	100%	100%
	$RRAIM_{PD(30\,s)}$	79.28%	99.06%	95.93%	100%	100%	100%
	$RRAIM_{PD}$	80.75%	99.59%	97.43%	100%	100%	100%
99.9%	$RRAIM_{RD}$	67.52%	96.87%	87.96%	100%	99.70%	100%
	$RRAIM_{PD(30\,s)}$	67.88%	97.15%	89.09%	100%	99.73%	100%
	$RRAIM_{PD}$	69.96%	97.63%	91.15%	100%	100%	100%

4.4 Chapter Conclusions

The use of Relative Receiver Autonomous Integrity Monitoring (RRAIM) for aircraft precision approach navigation was investigated in this chapter. In the concept investigated, the responsibility for detecting hazardously misleading information is divided between the RRAIM-equipped user and the navigation system provider, whose space and ground systems are the basis for a GNSS Integrity Channel (GIC). The GIC performs ranging source integrity screening, generates corrections, and then broadcasts this information to users worldwide via a space based communication channel. During the correction processing and communication interval, there will be a latent period during which the user must rely on past GIC information. This latency (also called coasting time) can be significantly longer than the TTA for SBAS systems like WAAS, depending

on the implementation. The user is continuously positioning in real time, and integrity against threats occurring during the latency period can be provided by RRAIM.

In this paper two versions of RRAIM were studied: a Range Domain (RD) and a Position Domain (PD) implementation. In both cases the user stores past carrier smoothed code and carrier phase measurements, and selects from those a reference epoch for which it has already received the GIC corrections. In both cases the user has three sets of measurements to use: the stored code measurements with corrections from the reference epoch (whose integrity is ensured by the GIC), a stored carrier phase measurement from the same reference epoch, and the current carrier phase measurement.

In the RD implementation the user creates a set of projected measurements by adding time-differential carrier phase measurements (current minus reference) to the reference GIC-corrected code measurements. The user position is obtained directly using these projected measurements. This implementation and corresponding covariance analysis is relatively straightforward. However, it requires use of only those ranging sources that are continuously available between the current and reference epochs.

The PD implementation generates an initial user position at the reference epoch using the corrected code measurements, and then adds to it a differential position vector, generated with the differential carrier phase measurements. The PD implementation allows for the use of all ranging sources available at the reference time to generate the initial position. The disadvantage is that the covariance analysis is significantly more complicated, and the corresponding bounding protection levels are more difficult to define. This is true because the carrier differential position error includes the effects of changes in user-SV lines of sight. These geometry changes introduce a correlation

between the differential (current minus reference) carrier phase position error and the initial code-based reference position error. When *CTs* of several minutes are considered, these effects cannot be neglected, and have therefore been carefully analyzed and modeled in this chapter.

The RRAIM detection function described in this chapter is the same for both the RD and the PD implementations, as it needs only detect hazardous measurement errors during the *CT* (because for the initial position integrity is provided by the GIC). When general error models for the corrected code measurement errors are considered, including potential unknown biases and Gaussian errors, the geometry change effects introduce a non-centrality parameter in the fault free Chi Square distribution of the RRAIM residual. How to account for this effect is carefully considered in this work.

In summary, this chapter provides the general formulas and derivations for positioning, fault detection, and protection level generation to meet a given set of integrity and continuity requirements. The mathematical justification of assumptions and models is provided, along with practical algorithms that pave the way toward real time implementation. An example of *VPL* results using the two algorithms for one site was provided. Worldwide coverage results were also provided for both implementations.

The RRAIM architectures described in this chapter have the advantage of not being excessively demanding on satellite constellation size, and they also allow the GIC segment to significantly relax requirements on time to alert. They provide an improvement in worldwide navigation using methods that are straightforward enough to not pose serious obstacles in certification or installation on the user's end. The actual implementation of this architecture will depend on strategic non-technical decisions, like

the future size of the constellations, the ability to install ground stations world-wide, and/or the willingness to use satellites not operated by the user's own country.

CHAPTER 5

EXTENDED RRAIM (ERAIM)

As described in previous chapters, the availability of RAIM to provide integrity is often limited by satellite geometry and signal quality. The algorithm described in this chapter allows recovery of fault detection availability when the current SV geometry does not support integrity using ARAIM. The concept introduced here is called Extended RAIM (ERAIM). It consists of finding a prior position with a better geometry, and projecting the user's position to the current time with a carrier phase RRAIM implementation. In contrast to the implementation presented in Chapter 4, in this case the integrity for both the initial position and the coasting period is provided by the ERAIM monitor. This fact makes it a very attractive architecture, as it needs no augmentation; as a tradeoff, the burden of also detecting potential failures in the reference position makes both the analysis of and demands on the monitor more challenging.

As seen in section 2.6, the GEAS panel developed new navigation architectures for aviation to provide worldwide coverage for aircraft precision approach, initially LPV-200 operations, with minimal ground infrastructure. One of those architectures is based on ARAIM, and it proved insufficient to provide suitable coverage results. Another architecture, developed extensively in Chapter 4, used the GIC to provide integrity on the initial position and RRAIM to provide integrity during coasting, thus significantly relaxing the TTA requirement. For most geometries, the 'ERAIM architecture' presented uses the simple ARAIM test for integrity. However when the current geometry does not support ARAIM-based integrity, the user 'looks back' to a past epoch where it did, and

projects that initial position to the current position using RRAIM. Thus, the specific purpose of RRAIM in this case is not to relax TTA (which is essentially instantaneous for ARAIM anyway), but to potentially coast through gaps in ARAIM availability.

To test the concept, candidate ARAIM and RRAIM detection functions based on the FMS technique were developed and a resulting methodology to generate VPL_{ERAIM} was defined. The initial availability results using this approach showed some, but not especially significant, gains in availability relative to ARAIM alone (these analyses and results are not presented in this dissertation). The choice of the LSR residual-based detection method made it necessary to use very conservative assumptions to be able to generate VPL_{ERAIM} without prohibitively large computation times. This approach was suitable for the applications in chapters 3 and 4 because it only involved failure detection during the coasting period. In the current case the integrity of the initial reference position must also be ensured. In response a solution separation methodology (see subsection 2.5.2) is pursued here because it is far more direct to implement with ERAIM, without the conservative assumptions that were needed to make the residual-based approach computationally tractable.

The algorithms to implement ERAIM are presented in this chapter, with the proof that for a given error model a specified level of continuity risk can be met and the position error bounds meet the integrity specifications also. These derivations are valid in a general sense, but given the relevance of GEAS for the future of navigation, examples of availability improvement from traditional RAIM are shown for an LPV-200 approach using the GEAS error models (see Appendix A and Appendix D). Some results using a different error model, of particular interest to the US Navy are also shown.

The efficiency of all RRAIM applications, of which ERAIM is a novel variation, is diminished with longer *CT*s; consequently, ERAIM is particularly attractive to mitigate short duration threats to continuity. An example considered in this work is the loss of ranging sources due to aircraft banking. Another potential application for ERAIM is to mitigate the effects of loss of satellite lock due to scintillation in equatorial regions. This second application is left for future work.

5.1 ERAIM

The ERAIM concept is illustrated in Figure 5.1. At a current time t_k, if the *VPL* for ARAIM ($VPL_{ARAIM(t_k)}$), exceeds the *VAL*, the aircraft can search back in time through stored measurements and satellite geometries by an interval *CT* such that $t_0 = t_k - CT$, to find an epoch for which $VPL_{ARAIM(t_0)} < VAL$. The punctual position fix at t_k is then generated using time differenced carrier measurements across *CT*. Fault detection for the current position fix is performed using a SS RAIM test. The residual errors in the time differenced carrier phase measurements are much smaller than those for the code-based initial position; consequently the resulting ERAIM protection level can potentially be smaller than the punctual ARAIM protection level: $VPL_{ERAIM(t_k)} < VPL_{ARAIM(t_k)}$.

5.1.1 ERAIM Positioning Algorithm. When the current ARAIM solution does not meet the system requirements, the ERAIM navigation solution for the current position is

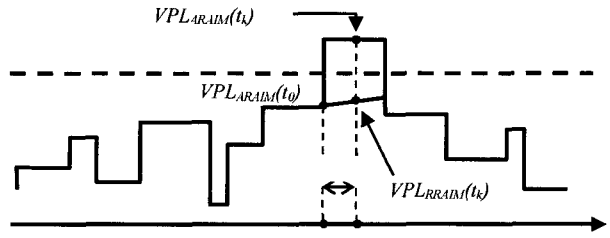


Figure 5.1. ERAIM Concept

based on a previous ARAIM-approved position estimate $\hat{\mathbf{x}}_0$ projected forward to the current time using a relative carrier-phase position estimate $\Delta\hat{\mathbf{x}}$:

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}_0 + \Delta\hat{\mathbf{x}} \tag{5.1}$$

where, as in previous chapters and all subsequent formulas, no-subscript is short notation for current time t_k and Δ is an equivalent of a '$_{k\text{-}0}$' subscript (meaning coasting time $t_k - t_0$).

The reference position is computed as:

$$\hat{\mathbf{x}}_0 = \mathbf{S}_0\mathbf{z}_0 \tag{5.2}$$

with:

$$\mathbf{S}_0 = \left(\mathbf{G}_0^{\mathrm{T}}\mathbf{W}_0\mathbf{G}_0\right)^{-1}\mathbf{G}_0^{\mathrm{T}}\mathbf{W}_0 \tag{5.3}$$

where:

$$\mathbf{W}_0^{-1} = \begin{bmatrix} \sigma_{1,0}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{n,0}^2 \end{bmatrix} \tag{5.4}$$

and each standard deviation $\sigma_{i,0}$ is a function of the satellite based errors (represented in Eq. (5.5) by the standard deviation of the User Range Accuracy or URA), the airborne multipath and receiver noise error model, and the tropospheric delay model for SV *i*:

$$\sigma_{i,0}^2 = \sigma_{URA,i,0}^2 + \sigma_{MP+N,i,0}^2 + \sigma_{tropo,i,0}^2 \tag{5.5}$$

(see Appendix A for representative values), and $\mathbf{z_0}$ is the vector of carrier smoothed code measurements, from now on referred to as 'code' measurements.

The relative position estimate $\Delta\hat{\mathbf{x}}$ from epoch 0 to the current epoch *k* is computed using relative carrier-phase measurements from all satellites continuously tracked between epochs 0 and *k*:

$$\Delta\hat{\mathbf{x}} = \mathbf{S}_\phi \boldsymbol{\varphi} \tag{5.6}$$

$$\mathbf{S}_\phi = \left(\mathbf{G}^{\mathrm{T}}\mathbf{W}_{\phi c}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{W}_{\phi c} \tag{5.7}$$

and the compensated time-differenced carrier-phase measurement is formed by the raw carrier-phase measurements $\boldsymbol{\varphi}^*$ at current epoch, and those at the reference position epoch, $\boldsymbol{\varphi}_0^*$, plus geometry change compensation:

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}^* - \boldsymbol{\varphi}_0^* - \Delta\mathbf{G}\hat{\mathbf{x}}_0 \tag{5.8}$$

$\Delta\mathbf{G}$ is the geometry change matrix:

$$\Delta\mathbf{G} = \mathbf{G} - \mathbf{G}_0 \tag{5.9}$$

The weighting matrix used in the projection matrix $\mathbf{S}_\phi$ in Eq. (5.7) is defined as:

$$\mathbf{W}_{\phi c}^{-1} = \mathbf{W}_\phi^{-1} + \Delta\mathbf{G}\,\mathrm{cov}\left[\delta\mathbf{x}_0\right]\Delta\mathbf{G}^{\mathrm{T}} \tag{5.10}$$

with

$$\mathbf{W}_{\phi}^{-1} = \begin{bmatrix} \sigma_{\phi,1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{\phi,n}^2 \end{bmatrix} \tag{5.11}$$

where $\sigma_{\phi,i}^2$ is the variance of the ith relative carrier-phase measurement coasting error. The second term on the right-hand side of Eq. (5.10) projects the error of the estimated position at time 0 through the change in the geometry matrix from 0 to k.

The carrier-phase measurement coasting error of the ith satellite over the CT is due to three error sources: 1) satellite clock and ephemeris coasting error, 2) tropospheric coasting error, and 3) double difference carrier-phase measurement noise-plus-multipath error. Each individual coasting error is treated as independent from the others. Therefore:

$$\sigma_{\phi,i}^2 = \sigma_{\Delta clock+\Delta ephem,i}^2 + \sigma_{\Delta trop,i}^2 + \sigma_{\Delta MP+\Delta N\phi,i}^2 \tag{5.12}$$

[Note: $\hat{\mathbf{x}}_0$ (the reference position estimate) always refers to a solution using carrier smoothed code measurement (an ARAIM implementation), while the current estimate $\hat{\mathbf{x}}$ could be the ARAIM implementation solution, or the ERAIM solution if the ARAIM solution does not meet the requirements (the case we are interested in). To simplify the notation the indication of what type of solution is used has been eliminated, as in either case Eq. (5.1) is valid: if the ARAIM solution is used, $0 = k$, and $\Delta\hat{\mathbf{x}}_{k-k}$ is a vector of zeros. This note also applies to the subset solutions $\hat{\mathbf{x}}_l$ and $\hat{\mathbf{x}}_{l,0}$ introduced in equations (5.13) and after.]

5.1.2 ERAIM Detection Algorithms. Several possibilities were initially explored as failure detection schemes. The LSR FMS implementation needs to account for the

possibility of a failure on the initial position and/or during coasting separately. That implies that to compute the P_{MD}; either all combinations of failure magnitudes are considered (unacceptable computation-time wise), or conservative assumptions have to be made which result in an unacceptable availability loss. Consequently the solution separation methodology was adopted.

The position estimate error for a full-set solution (using all available satellites) at the current epoch can be bounded using solution separation fault detection.

Applying Eq. (2.7) to the position estimate, and considering that there is only one true position regardless of how we estimate it, we can describe the full set solution estimate error at current time as:

$$\delta\mathbf{x} = \hat{\mathbf{x}} - \hat{\mathbf{x}}_l + \delta\mathbf{x}_l \tag{5.13}$$

where $\hat{\mathbf{x}}_l$ is the lth subset position estimate (i.e., estimate with satellite l removed) and $\delta\mathbf{x}_l$ are the estimate errors for the lth subset estimate.

From equations (5.1) and (5.13):

$$\delta\mathbf{x} = \left(\hat{\mathbf{x}}_0 + \Delta\hat{\mathbf{x}}\right) - \left(\hat{\mathbf{x}}_{l,0} + \Delta\hat{\mathbf{x}}_l\right) + \delta\mathbf{x}_{l,0} + \delta\Delta\mathbf{x}_l \tag{5.14}$$

where the lth subset ERAIM position estimate is computed as:

$$\hat{\mathbf{x}}_l = \hat{\mathbf{x}}_{l,0} + \Delta\hat{\mathbf{x}}_l = \mathbf{S}_{l,0}\mathbf{z}_0 + \mathbf{S}_{\phi,l}\boldsymbol{\varphi}\,, \tag{5.15}$$

$$\mathbf{S}_{l,0} = \left(\mathbf{G}_{l,0}^{\mathrm{T}}\mathbf{W}_0\mathbf{G}_{l,0}\right)^{-1}\mathbf{G}_{l,0}^{\mathrm{T}}\mathbf{W}_0 \tag{5.16}$$

where $\mathbf{G}_{l,0}$ is $\mathbf{G}_0$ with row l replaced by zeros, and

$$\mathbf{S}_{\phi,l} = \left(\mathbf{G}_l^{\mathrm{T}}\mathbf{W}_{\phi c}\mathbf{G}_l\right)^{-1}\mathbf{G}_l^{\mathrm{T}}\mathbf{W}_{\phi c} \tag{5.17}$$

where $\mathbf{G}_l$ is $\mathbf{G}$ with row l replaced by zeros.

The test statistic d_l for detection of a failure (in the user vertical direction) caused by satellite l is defined as:

$$d_l \equiv \left| \left(\hat{\mathbf{x}}_0 + \Delta\hat{\mathbf{x}}\right)_{(3)} - \left(\hat{\mathbf{x}}_{l,0} + \Delta\hat{\mathbf{x}}_l\right)_{(3)} \right| \tag{5.18}$$

where the subscript '3' within the parenthesis indicates the third component of the position vector, which corresponds to the vertical position state in an East, North, Up coordinate system.

The error bounds developed in this chapter account for bounded code measurement errors that cannot be modeled as Gaussian. Considering that possibility, the corresponding detection threshold for the lth test statistic (D_l) is

$$D_l = k_{ffd,l} \times \sigma_{dV,l} + \sum_{\iota=1}^{n} \left| \mathbf{F}_0 - \mathbf{F}_{\iota,0} \right|_{(3,\ )} \times b_{nom} \tag{5.19}$$

where

$$\mathbf{F}_0 \equiv \left[\mathbf{I} - \mathbf{S}_\phi \Delta\mathbf{G}\ \right] \mathbf{S}_0, \quad \mathbf{F}_{\iota,0} \equiv \left[\mathbf{I} - \mathbf{S}_{\phi,\iota} \Delta\mathbf{G}_\iota\ \right] \mathbf{S}_{\iota,0} \tag{5.20}$$

b_{nom} is an upper bound on non-Gaussian errors at time zero under normal error conditions; with the same criteria explained in subsection 4.1.1. [Note: in Eq. (5.19), as well as the following formulas (5.23) and (5.26) $|\bullet|$ means element-wise absolute value and not a norm.] The fault-free standard deviation of the Gaussian contribution to the test statistic is represented by:

$$\sigma_{dV,l} \equiv \sqrt{dP_{l(3,3)}} \tag{5.21}$$

where

$$d\mathbf{P}_l = \left[\mathbf{F}_0 - \mathbf{F}_{l,0}\right] \mathbf{W}_0^{-1} \left[\mathbf{F}_0 - \mathbf{F}_{l,0}\right]^{\mathrm{T}} + \left[\mathbf{S}_\phi - \mathbf{S}_{\phi,l}\right] \mathbf{W}_\phi^{-1} \left[\mathbf{S}_\phi - \mathbf{S}_{\phi,l}\right]^{\mathrm{T}} \tag{5.22}$$

The fault-free detection multipliers $k_{ffd,l}$ are selected (for $l = 1, 2, ..., n$) to ensure that the sum of the fault-free alarm probabilities for all n tests is lower than the continuity risk requirement. In the simplest implementation, this risk is allocated equally across all n tests. This uniform allocation is used to generate the results in section 5.2.

5.1.3 VPL equations for ERAIM. Following what is done in the ARAIM case, the ERAIM protection levels are derived for the fault-free full-set (VPL_{ff}) and, under the fault hypotheses (VPL_l), for all the subsets. For the fault-free hypothesis:

$$VPL_{ff} = k_{md,ff} \times \sigma_{v,ff} + \sum_{i=1}^{n} \left|\mathbf{F}_{i,0}\right|_{(3,:)} \times b_{\max} \tag{5.23}$$

where $b_{\max}$ is an upper bound on non-Gaussian errors at time zero under worst-case error conditions, $k_{md,ff}$ is selected to meet the fault-free integrity allocation, and

$$\sigma_{v,ff} \equiv \sqrt{P_{ff\,(3,3)}} \tag{5.24}$$

with

$$\mathbf{P}_{ff} = \left(\mathbf{I} - \mathbf{S}_{\Delta\phi}\Delta\mathbf{G}\right)\left(\mathbf{G}_0\mathbf{W}_0\mathbf{G}_0^{\mathrm{T}}\right)^{-1}\left(\mathbf{I} - \mathbf{S}_{\phi}\Delta\mathbf{G}\right)^{\mathrm{T}} + \left(\mathbf{G}\mathbf{W}_{\phi}\mathbf{G}^{\mathrm{T}}\right)^{-1} \tag{5.25}$$

The error contribution due to geometry change (over the coasting interval) is captured in the first term in Eq. (5.25) (as is the case in Eq. (5.28) for the subset solution).

Under the fault hypotheses, VPL_l is derived as:

$$VPL_l = D_l + k_{md,l} \times \sigma_{v,l} + \sum_{i=1}^{n} \left|\mathbf{F}_{i,0}\right|_{(3,:)} \times b_{\max} \tag{5.26}$$

where $k_{md,l}$ is selected to meet the integrity allocation for a fault on the lth satellite, and

$$\sigma_{v,l} \equiv \sqrt{P_{l\,(3,3)}} \tag{5.27}$$

$$\mathbf{P}_l = \left(\mathbf{I} - \mathbf{S}_{\phi,l}\Delta\mathbf{G}_l\right)\left(\mathbf{G}_{l,0}\mathbf{W}_0\mathbf{G}_{l,0}^{\mathrm{T}}\right)^{-1}\left(\mathbf{I} - \mathbf{S}_{\phi,l}\Delta\mathbf{G}_l\right)^{\mathrm{T}} + \left(\mathbf{G}_l\mathbf{W}_\phi\mathbf{G}_l^{\mathrm{T}}\right)^{-1} \tag{5.28}$$

The system will be available when:

$$VAL > \max\left(VPL_{ff}, \max_{l=1 \text{ to } n}(VPL_l)\right) \tag{5.29}$$

5.2 ERAIM Applications

5.2.1 LPV-200 ARAIM improvement on coverage and availability. The first application where the algorithms presented in section 5.1 were used was in improving availability for the ARAIM architecture for LPV-200 and LPV-250 approaches within the GEAS. Results produced at IITs NAVLAB showed substantial improvement in coverage and average availability, but not enough to envision a worldwide implementation without substantial GBAS or SBAS augmentation (as is the aim of ARAIM) [GEA10]. This analysis is still in process, studying how improvement in certain variables (user clock error, less conservative error models, etc.) could obtain the desired autonomous monitoring. The results are not presented in this dissertation.

5.2.2 Aircraft Banking. When a user makes a turn, the roll angle might cause a loss of ranging sources (SVs). This is of special interest for GNSS-based navigation because one of its advantages over ILS guided operations is that it allows for curved approaches. Curved approaches for future navigation systems have not been defined, but it would be important to evaluate the effect of banking on continuity and availability. This analysis is also relevant for military operations, as they include and allow sharper turns. A general analysis of this topic is not within the scope of this paper. However, it will be shown that

ERAIM could be very effective in neutralizing the negative effects of banking. This is illustrated in this section through an extensive simulation of specific examples. If an aircraft is using ARAIM to provide integrity during an LPV-200 approach, a geometry diminished by satellites lost during a turn might not meet the operation's requirements. With the same algorithms detailed in section 5.1, the user could search back in time for an epoch when the operation specifications were met, and coast to the current position using differential carrier phase measurements. To avoid prohibitive simulation times, the effect of banking was simulated in a straightforward way, by introducing different elevation masks, assuming in each case that all SVs below that mask were not visible in any direction at the given epoch being evaluated (see Appendix E regarding the validity of this simplified model). Note that in this simulation, for each epoch being evaluated, banking affects the current geometry, but not past geometries, as the satellite loss is considered to be a short term effect. The minimum *CT* will be the duration of the aircraft's turn or 'bank' time. In this work the *CT* used is one minute (see Appendix E).

For each set of error model and system specifications, worldwide average availability, and coverage for 95%, 99% and 99.5% availability criteria, were computed using ARAIM for different masks (5 deg, 15 deg and in one case 30 deg). The difference from the 5 deg mask-results to the results using higher masks represents the availability lost due to banking.

Then the same results were generated, but this time assuming the airplane also has an ERAIM implementation. Comparing the 15 deg and 30 deg results between the ARAIM implementation and the ARAIM+ERAIM implementation allows quantifying how effective ERAIM is in neutralizing the negative effect of banking.

Table 5.1 shows results for an LPV-200 approach, but with a more benign situation, in which there are two constellations available, and the maximum bias is b_{max} = 0.75 m (instead of b_{max}= 1.25 m used on previous simulations). For the 99% and 99.5% coverage criteria cases, the simulated banking effect introduced (by increasing the SV elevation mask) reduces the coverage significantly, but the 100% coverage is recovered by the ERAIM implementation.

A second example is presented in Table 5.2. In this case, the current GPS constellation is used (30 SV as of July 15th 2009) but a less stringent VAL is applied to define when the system is available (compared to VAL= 35 m for an LPV-200 approach). Results show how the effects of banking are neutralized for a 15 deg mask bringing the average availability back to 99% from 70%, and the 95% coverage back to 97% (from 0%). The results for an extreme 30 deg mask are also introduced (representing a sharp maneuver of a military aircraft for example), showing how the ERAIM implementation recovers the average availability back to a 95% level (from 58%).

The values presented in tables 5.1 and 5.2 are not valuable in themselves because curved approaches have not been defined, and because of the way the banking effects are simulated is just a simplified representation; however the big gains in availability, which result in changes of 30 to 40% in average availability and coverage results, are a clear indication of the power of ERAIM in avoiding loss of continuity in cases of short duration loss of ranging sources.

Table 5.1 ERAIM Neutralization of Effect of Banking: dual constellation- b_{max}=0.75 m

	Average avail.		95%		99%		99.5%	
Mask	ARAIM	ERAIM	ARAIM	ERAIM	ARAIM	ERAIM	ARAIM	ERAIM
5 deg	100%	100%	100%	100%	100%	100%	100%	100%
15 deg	100%	100%	100%	100%	**86%**	**100%**	**76%**	**100%**

Table 5.2 ERAIM Neutralization of Effect of Banking: current constellation- VAL=50 m

	Average avail.		95%		99%		99.5%	
Mask	ARAIM	ERAIM	ARAIM	ERAIM	ARAIM	ERAIM	ARAIM	ERAIM
5 deg	99%	100%	98%	100%	52%	86%	38%	78%
15 deg	**70%**	**99%**	**0%**	**97%**	0%	53%	0%	35%
30 deg	**58%**	**95%**	0%	57%	0%	7%	0%	3%

5.3 Chapter Conclusions

The use of RRAIM for aircraft precision approach navigation was investigated in this chapter under a novel implementation called Extended RRAIM (ERAIM). Contrary to the application developed in Chapter 4, for ERAIM, the responsibility for detecting hazardously misleading information resides entirely on the aircraft, including both the reference position and the coasting segment. In this application the aircraft will provide integrity using ARAIM, but when the geometry does not support the operation in process, the aircraft will attempt to find a previous geometry and project from it to its current position using time differential carrier phase measurements. Failure detection is achieved

with a RRAIM test on the whole position solution, including both the reference and the coasting measurements, using a SS methodology. SS has shown to be more efficient for this implementation, as it does not require the conservative assumptions needed to make the LSR approach tractable form a computation time perspective. It is also better suited for producing bounds on the vertical position estimate error, as the LSR includes errors in all three directions.

The algorithms to produce a *VPL* that bounds the error in the position estimate are provided, with the definitions of measurement and error covariance matrices needed for its computation. These include the effect of bounded errors that cannot be modeled as Gaussian.

An example application of ERAIM is shown, to avoid loss of continuity in case of satellite loss due to aircraft banking during turns. Several cases are found in which this negative effect is almost completely neutralized by implementing an ERAIM test to provide integrity.

CHAPTER 6

CONCLUSIONS

The work presented in this dissertation has the general aim of extending the use of satellite based aircraft navigation to precision approach and landing operations. Solutions to fundamental navigation issues (integrity and continuity risk mitigation during precision approach operations) are presented, by proposing, analyzing, developing, and verifying several novel implementations based on carrier phase time-differential Relative Receiver Autonomous Integrity Monitoring (RRAIM). These solutions contribute to different areas of airborne navigation.

6.1 Summary of Work and Findings

The two RRAIM methods introduced in Chapter 4 support a worldwide SBAS:

RRAIM Least Square Range Domain: The user computes an initial reference position using carrier smoothed code, with integrity assured by augmentation from an SBAS source. The aircraft then uses the time-differential carrier phase measurements to navigate the latest path portion (from the initial position to the current time) and to detect failures in that period. Available information is merged in the range domain before the position is estimated.

RRAIM LS Position Domain: similar to the implementation above, but the ranges are merged in the Position Domain (PD). PD implementations have the advantage that all satellites can be used for the initial reference position, but have the disadvantage of introducing new error terms due to uncertainty in the initial position.

Either of these two implementations (1) ensures that users are instantaneously alerted of hazardously misleading information, and (2) lowers the number of situations in which the system is unavailable because the local satellite geometry is insufficient for safe position estimation.

Other RRAIM implementations can support GBAS operations, like the *RRAIM LS aircraft baseline failure monitoring* method introduced in Chapter 3. In this case the time differential carrier phase measurements are not used for positioning, but to provide integrity during ionospheric storms, without excessive loss of availability.

Finally some RRAIM contributions, like the *ERAIM Solution Separation (SS) PD* method introduced in Chapter 5, are relevant for all users, with or without augmentation. This implementation uses a SS PD detection technique, and the stored information used to obtain the initial position does not need SBAS information. It allows the users to ensure continuity by bridging gaps in good satellite geometry while simultaneously maintaining high integrity

RRAIM is based on two features: the use of time differenced measurements in an innovative form and the use of carrier phase measurements, which have smaller receiver-end errors than code measurements. The need for cycle ambiguity estimation can be eliminated by differencing measurements in time, creating spatial baselines associated with the user's translation over the time-difference interval.

For the static ionospheric front detection application (*RRAIM LS aircraft baseline failure monitoring*), the method of analysis developed provides insight into what ionospheric front shapes are most hazardous during an aircraft approach, which simplifies the analysis of the total mitigation of the threat space using different monitors. Simulation

results for selected airport locations show that the threat can be mitigated with a tolerable availability loss if it is used in conjunction with an ionospheric storm external-alarm system and a ground gradient monitor at the GBAS ground facility.

For the *RRAIM LS PD* and *RD* implementations, the responsibility for detecting hazardously misleading information is divided between the RRAIM-equipped user and the navigation system provider (GIC). The GIC performs ranging source integrity screening, and then broadcasts information to users worldwide. During the processing and communication interval, there will be a latent period (also called coasting time) which can be significantly longer than the required time-to-alert for SBAS systems like WAAS. The user is continuously positioning in real time, and integrity against threats occurring during the latency period can be provided by either of the two RRAIM implementations mentioned, as is shown in Chapter 4.

After achieving the goal of lowering the time to alert, it became obvious that the use of these RRAIM methods could be extended to recovering availability for times with a poor SV geometry. The PD implementation is better suited for this purpose, as it allows the use of all ranging sources available at the reference time to generate the initial position. The disadvantage is that the covariance analysis is significantly more complicated, and the corresponding bounding protection levels are more difficult to define. It also introduces an extra error term that projects errors in the initial position estimation into the current position estimate error. Nevertheless, the results in Chapter 4 show that the PD implementation is favored in that tradeoff, making it the implementation of choice.

A user providing its own integrity autonomously with ARAIM can have excessive breaks in continuity for various reasons, including rapid changes to poor SV geometries that do not support the integrity requirements. Some of these gaps in continuity might be recovered by using the RRAIM method introduced in Chapter 5; Extended RRAIM (ERAIM).

ERAIM is similar to *RRAIM LS PD* in that the user searches back for a good SV geometry. However, for ERAIM, failure detection is achieved with a RRAIM test on the whole position solution, including both the reference and the coasting measurements. For ERAIM, a SS methodology for fault detection is more suitable than a LS residual approach.

An example application of ERAIM is shown in Chapter 5, where ERAIM is implemented to avoid loss of continuity in case of satellite tracking loss due to aircraft banking during turns. Several scenarios are found in which this negative effect is almost completely neutralized by implementing an ERAIM test to provide integrity.

6.2 Contributions

This work introduces *new concepts* for integrity monitoring using RRAIM:

- RRAIM Least Square Range Domain
- RRAIM LS Position Domain
- RRAIM LS aircraft baseline failure monitoring
- ERAIM Solution Separation Position Domain

Algorithms: For each of the new concepts proposed, implementation methodology is developed in detail, for use at the aircraft and/or ground facility as well as for

simulations to evaluate the efficiency of the implementation for a given set of requirements or threat model.

Error Bounds: For each implementation the derivation of conservative error bounds is presented. These include the protection levels for the hypotheses of fault free and faulted situations. The criteria for determining a system's availability of continuity and availability of integrity are clearly defined for each case. The protection levels include a means to account for the effect of potential (bounded) errors that cannot be modeled as Gaussian.

Measurements and Covariance Matrices: The meaning of measurements and their covariance matrices are defined for each algorithm, with particular care on considering the role of the reference position estimate error as it is projected to the current epoch in the position domain implementations.

Applications: Applications are developed using the new RRAIM concepts in the context of current GBAS and future SBAS configurations, to provide integrity against concrete threats.

Results: Simulation results are generated to evaluate the effectiveness of the developed implementations in detecting potential threats. The results obtained range from individual performance metrics for a given SV geometry and user location to worldwide coverage.

6.3 Future Work

A number of suggestions for future work are given to expand and complete the foundations for RRAIM applications laid out in this dissertation:

- Consider the probability of multiple simultaneous SV failures (assumed negligibly small in the present analysis). Where appropriate, expand the detection capabilities of the monitors presented in this work, to detection of simultaneous SV faults.
- Quantify the benefits of reducing residual errors that would make the coasting error smaller. This would allow going further back in time looking for a strong SV geometry to generate a reference position. These error reductions could come from more precise clocks in the next generation of GNSS satellites and/or less conservative tropospheric error models.
- Investigate curved approaches for future operations, and establish the roll angles during these types of approaches. Then do a more precise analysis of signal loss due to aircraft banking and how they are mitigated by ERAIM.
- Study the relationship between the Least Squares Residual and the Solution Separation residuals for RRAIM, determining which one is more suitable for different types of problems.

APPENDIX A

ERROR MODELS

A.1 Ionospheric Front Threat Space

Table A.1. Ionospheric Front Threat Space

Elevation	Speed	Width (w)	Slant Slope (k)	Max.delay (h_{max})
5-15 deg	0-750 m/s	25-200 km	0-375 mm/km	50 m
15-65 deg			0-(375+el-15) mm/km	
65-90 deg			0-425 mm/km	

A.2 Specification for Static Ionospheric Front Threat Mitigation Simulations

Category I scenario:

P_{creq}: 1-10^{-6}/approach

P_{MDreq}: 10^{-4}

σ_{vig}: 2 mm/km

Double Difference standard deviation (σ_{DD}): 1 mm/km

VAL=10 m

Category III scenario

P_{creq}: 1-10^{-8}/approach

P_{MDreq}: 10^{-6}

σ_{vig}: 4 mm/km

Double Difference standard deviation (σ_{DD}): 1 mm/km

VAL: 10 m

A.3 Error model for GEAS 'RRAIM' Architecture and ERAIM Simulations

The details of the GEAS error models can be found in [GEA08] here is a brief description:

$P_{HMIreq} = P_{HMIreq(NC)} + P_{HMIreq(FC)} = 4.35\times10^{-8} + 4.35\times10^{-8}$

A-priori failure rate: 10^{-4} per SV per hour

$P_{Creq} = 4\times10^{-6}$ per approach

$0.11\ \text{m}^2 < \sigma^2_{Cmpn} < 1.37\ \text{m}^2$ (function of elevation [GEA08])

$\sigma^2_{C\tau} + \sigma^2_{Ceph} = 0.5625\ \text{m}^2$

$0.01\ \text{m}^2 < \sigma^2_{CTropo} < 1.5\ \text{m}^2$ (function of elevation in [GEA08])

$\sigma^2_{\Delta\phi mpn} = 0.0016\ \text{m}^2$

$$\sigma^2_{\Delta\phi\, sv\tau} + \sigma^2_{\Delta\phi\, eph} = \left(8.5\times10^{-4}\ \frac{\text{m}}{\text{s}}\right)^2 \times (CT)^2$$

$0 < \sigma^2_{\Delta\phi trop} < 0.408\ \text{m}^2$ (function of elevation and CT in [GEA08]))

$\beta_{\max} = b_{\max} = 1.125\ \text{m}$

$\beta_{\text{nom}} = b_{\text{nom}} = 0.1\ \text{m}$

APPENDIX B

TRANSIENT RESPONSE OF A HATCH FILTER TO AN IONOSPHERIC FRONT

The formula for the accumulated effect of the ionospheric delay on the carrier smoothed code measurement was introduced in Eq. (3.22). The derivation for the first three epochs is presented in this appendix. Defining $f_0 \equiv -f_{\phi^{**}(0)}$, from Eq. (2.14):

$$f_1 = \left(\frac{M-1}{M}\right)^1 - f_{\phi^{**}(0)} + \sum_{J=1}^{1}\left(\frac{M-1}{M}\right)^J \left[f_{\phi^{**}(1+1-J)} - f_{\phi^{**}(1-1)}\right]$$
$$+ \sum_{J=1}^{1}\left(\frac{1}{M}\right)\left(\frac{M-1}{M}\right)^{J-1}\left(-f_{\phi^{**}(1+1-J)}\right)$$

$$f_2 = \left(\frac{1}{M}\right)\left(-f_{\phi^{**}(2)}\right) + \left(\frac{M-1}{M}\right)\left[\left(\frac{1}{M}\right)\left(-f_{\phi^{**}(1)}\right)\right.$$
$$\left. + \left(\frac{M-1}{M}\right)\left(f_{(0)} + \left[f_{\phi^{**}(1)} - f_{\phi^{**}(0)}\right]\right) + \left[f_{\phi^{**}(2)} - f_{\phi^{**}(1)}\right]\right]$$
$$= \left(\frac{M-1}{M}\right)^2 f_{(0)} + \left(\frac{M-1}{M}\right)^2 \left(\left[f_{\phi^{**}(1)} - f_{\phi^{**}(0)}\right]\right) + \left(\frac{M-1}{M}\right)^1$$
$$\times \left[f_{\phi^{**}(2)} - f_{\phi^{**}(1)}\right] + \left(\frac{M-1}{M}\right)\left(\frac{1}{M}\right)\left(-f_{\phi^{**}(1)}\right) + \left(\frac{1}{M}\right)\left(-f_{\phi^{**}(2)}\right)$$
$$= \left(\frac{M-1}{M}\right)^2 f_{(0)} + \sum_{J=1}^{2}\left(\frac{M-1}{M}\right)^J \left(\left[f_{\phi^{**}(2+1-J)} - f_{\phi^{**}(2-J)}\right]\right)$$
$$+ \sum_{J=1}^{2}\left(\frac{1}{M}\right)\left(\frac{M-1}{M}\right)^{J-1}\left(-f_{\phi^{**}(2+1-J)}\right)$$

$$
\begin{aligned}
f_3 &= \left(\frac{1}{M}\right)\left(-f_{\phi^{**}(3)}\right) + \left(\frac{M-1}{M}\right) \\
&\quad \times \left[\left(\frac{M-1}{M}\right)^2 f_{(0)} + \left(\frac{M-1}{M}\right)^2 \left(\left[f_{\phi^{**}(1)} - f_{\phi^{**}(0)}\right]\right) + \left(\frac{M-1}{M}\right)^1 \left[f_{\phi^{**}(2)} - f_{\phi^{**}(1)}\right]\right. \\
&\quad \left. + \left(\frac{M-1}{M}\right)\left(\frac{1}{M}\right)\left(-f_{\phi^{**}(1)}\right) + \left(\frac{1}{M}\right)\left(-f_{\phi^{**}(2)}\right) + f_{\phi^{**}(3)} - f_{\phi^{**}(2)}\right] \\
&= \left(\frac{M-1}{M}\right)^3 f_{(0)} + \left(\frac{M-1}{M}\right)^3 \left(\left[f_{\phi^{**}(1)} - f_{\phi^{**}(0)}\right]\right) \\
&\quad + \left(\frac{M-1}{M}\right)^2 \left[f_{\phi^{**}(2)} - f_{\phi^{**}(1)}\right] + \left(\frac{M-1}{M}\right)\left(f_{\phi^{**}(3)} - f_{\phi^{**}(2)}\right) + \\
&\quad + \left(\frac{M-1}{M}\right)^2 \left(\frac{1}{M}\right)\left(-f_{\phi^{**}(1)}\right) + \left(\frac{M-1}{M}\right)\left(\frac{1}{M}\right)\left(-f_{\phi^{**}(2)}\right) + \left(\frac{1}{M}\right)\left(-f_{\phi^{**}(3)}\right) \\
&= \left(\frac{M-1}{M}\right)^3 f_{(0)} + \sum_{J=1}^{3} \left(\frac{M-1}{M}\right)^J \left[f_{\phi^{**}(3+1-J)} - f_{\phi^{**}(3-J)}\right] \\
&\quad + \sum_{J=1}^{3} \left(\frac{M-1}{M}\right)^{J-1} \left(\frac{1}{M}\right)\left(-f_{\phi^{**}(3+1-J)}\right)
\end{aligned}
$$

APPENDIX C

ALLOCATION EXAMPLE

Figure C.1 shows an example of allocation from [RTC04].

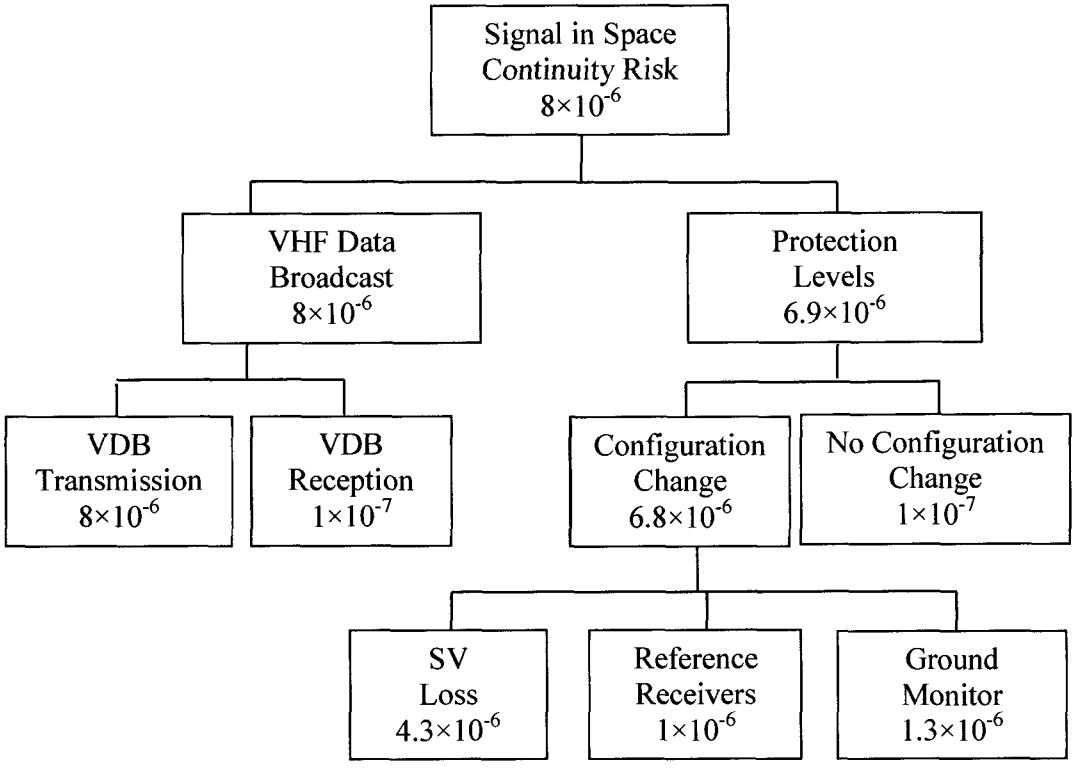


Figure C.1. Category I Continuity Allocation for LAAS

APPENDIX D

SPECIFICATIONS FOR DIFFERENT FLIGHT OPERATIONS

Table D.1 Standards for Aviation Integrity for Different Operations

	En Route (RNAV-2)	Terminal (RNAV-1)	LNAV/ VNAV	(*LPV*) LPV 200	Cat I	Cat II	Cat III
TTA	15 s	15 s	10 s	6.2 s	6 s	2 s	2 s
HAL	2 nm	1 nm	556 m	40 m	40 m	17.3 m	15.5 m
VAL	-	-	50 m	(*50 m*) 35 m	10 m	5.3 m	5.3 m
P_{HMIreq}	10^{-7}/hour	10^{-7}/hour	2×10^{-7} /approach	2×10^{-7} /approach	2×10^{-7} /approach	2×10^{-9} /approach	2×10^{-9} /approach
P_{Creq}	-	-	-	8×10^{-6}/15 s	4×10^{-6}/15 s	2×10^{-6}/15 s	2×10^{-6}/15 s (vertical) 2×10^{-6}/30 s (horizontal)
Accuracy (H) (95%)	0.4 nm	0.4 nm	220 m	16 m	16 m	6.9 m	6 m
Accuracy (V) (95%)	-	-	20 m	(*20 m*) 4 m	7.6 m	2 m	2 m

Specification from [Eld08] [ICA09].

APPENDIX E

VALIDITY OF AIRCRAFT BANKING MODEL

As curved approaches are not defined for civil aviation, in subsection 5.2.2, additional elevation masks of 15 deg and 30 deg were used to simulate loss of SV ranging signals do to aircraft banking during turns. A *CT* of 60 s was used, equivalent to assuming prior geometries are unavailable to compute the reference position (Eq. 5.2) during that time, as the same SVs than at the current geometry might have been already blocked. This appendix provides the information to evaluate how representative these three numbers are.

A study of more than 1500 approaches of a Bombardier CRJ-200 aircraft gave the following relevant information [Chi10]:

- Within all cases and phases of flight the maximum bank angle was 46 deg.
- The mean maximum bank angle was 28 deg.
- The maximum bank angle for the final approach phase from 500 ft to 200 ft was 12 deg.
- The mean maximum bank angle for the same phase was 3.5 deg.
- The expected azimuth change rate with autopilot is 3 deg/s.

Considering that the roll angles for larger aircraft are expected to be smaller than for this regional jet case, the 15 deg SV elevation mask, representing civil aviation approaches, seems to be conservative when considering outages during the 500 ft to 200 ft phase of flight. Considering the given azimuth rate, a 180 deg turn would take one minute, the 60 s *CT* is representative.

Regarding the 30 deg elevation mask, from analyzing aircraft carrier landing operations (for example Case I and Case III described in [Lan09]), it becomes clear that it is useful to have a coarse idea of the power of ERAIM, but a more precise model is

required. Military operations include maneuvers with 45 deg angles, and holding patterns in circles at lower roll angles that can potentially last several minutes, making a one minute *CT* less representative of the geometries that might actually be available to generate the reference position.

APPENDIX F

IONOSPHERIC GRADIENT SEARCH INTERVAL DERIVATION

As was shown in subsection 3.3.2, for SIF threat mitigation, for each SV with a *FMS* corresponding to a certain geometry and front width, the P_{MD} has to be evaluated for a segment from k_{min} to k_{max}. It is difficult to find a maximum value for P_{MD} analytically; and when the MATLAB functions for that purpose were used, they proved to be unreliable for this specific application, particularly for the small P_{MDreq} values for Category III approaches. P_{MD} has to be evaluated numerically for the whole k_{min} to k_{max} segment, and to do that, a search interval (Δk) between each two evaluated points has to be chosen. Our purpose is to avoid having a Δk that is too large, as to risk "missing" a value of k for which $P_{MD} > P_{MDreq}$, which could potentially make available geometries that really should be discarded. On the other hand, as Δk is made smaller, more points need to be computed (think of a limit case in which only one value of k produces a $P_{\mathrm{MD}} > P_{MDreq}$, requiring a Δk=0), so from a computation-time point of view is desirable to make is as large as possible

From Eq. (3.28), and considering the effects of the noise on the position error (to simplify analysis only the vertical error is considered), and on the residual are independent; it can be said that the P_{MD} is the product of two probabilities:

$$P_{MD,f(w,k)} = \mathrm{P}\left(\left|\delta x_v\right| > VAL \;\; \left|f(w,k)\right.\right) \times \mathrm{P}\left(r < T \left| f(w,k)\right.\right) \tag{F.1}$$

To make notation more concise, for the reminder of the appendix, it is always assumed there is an ionospheric front present of width w, and the P_{MD} is a function of k. Eq. (F.1) is rewritten as:

$$P_{MD} = \mathrm{P}\left(\left|\delta x_v\right| > VAL\right) \times \mathrm{P}\left(r < T\right) \tag{F.2}$$

where the two factors on the right of the equation, and consequently the result on the left, are functions of k.

If there exists a range of values of k ($\bar{k}_{HMI}$) for which $P_{MD} > P_{MDreq}$, another range of values of k ($\bar{k}_{search}$) will exist that will make $P_{MD} > P_{MDsearch}$. As long as $P_{MDsearch} < P_{MDreq}$, $\bar{k}_{search} \subset \bar{k}_{HMI}$. The lower the value of $P_{MDsearch}$ the bigger the search interval can be (the idea is to look for $P_{MD} > P_{MDsearch}$ with a larger search interval, i.e. fewer points to evaluate, to find/discard cases of $P_{MD} > P_{MDreq}$). [Note: setting $P_{MDsearch}$ excessively low should be avoided as it will produce situations in which $\bar{k}_{search}$ exists even though $\bar{k}_{HMI}$ does not.]

As the CDF of $P(\delta x_v > VAL)$ is a monotonically increasing function with k, and the CDF of $P(r < T)$ is a monotonically decreasing function with k, there can be a maximum of two values of k that give the same P_{MD}. This allows us to state that if we choose a $P_{MDsearch} < P_{MDreq}$ (see Fig. F1), no point with $P_{MD} > P_{MDreq}$ exists in the search segment (i.e. between k_{min} and k_{max}) if no point with $P_{MD} > P_{MDsearch}$ is found in our evaluation, and the distance between individual evaluated points complies with:

$$\Delta k < 2 \times \min|k_{req} - k_{search}| < \text{length}\left(\bar{k}_{search}\right) \tag{F.3}$$

where k_{req} and k_{search} are the values of k that give a P_{MDreq} and $P_{MDsearch}$ respectively. [Note: the 'min' in Eq. (F.3) refers to the minimum of the two possible values, as there is a potential set of k_{req}- k_{search} on each side of the maximum value of P_{MD} inside the search segment. In Figure F.1, only the values on the right of that maximum are shown. 'length $\left(\bar{k}_{search}\right)$' is the maximum minus the minimum value within $\bar{k}_{search}$. In Figure F.1, only

the maximum value is shown.]. The second inequality in Eq (F.3) is always true; the first inequality is our condition to choose a suitable search interval Δk. As we don't know k_{req} or k_{search}, the purpose of this analysis is to find a conservative value for Δk.

Again using the monotonic properties of the individual probabilities that produce P_{MD} (Eq. F.2), it can be stated that the rate of change in the P_{MD} curve can never be bigger than any of the two individual rates (for $\mathrm{P}(|\delta x_v| > VAL)$ and $\mathrm{P}(r < T)$) at any given point, as the two individual rates are always in opposite directions.

The idea is to identify the point with the highest rate of change in each of the individual curves, select the higher of these two, and at that point compute the Δk necessary to produce a change in $P(\delta x_v > VAL)$ or $P(r < T)$ (depending what curve we are looking at) that is equal to $\Delta P = P_{MDreq} - P_{MDsearch}$ Given the property stated in the previous paragraph, the Δk found this way complies with the condition in Eq. (F.3).

The biggest rate of change of a Gaussian CDF (or the maximum PDF) is at the $P = 0.5$ point. We assume that for a non-central Chi^2 distribution the non-centrality parameter in big enough as to make the maximum change also close to the $P = 0.5$ point; for cases where this assumption does not stand, it is not difficult to find the maximum slope point. [Note: $P(r < T)$ is actually given by 1-CDF of a Chi Squared distribution. It is the CDF of $P(r < T)$ that we are referring to when we talk about a "Chi Square" or "non- central Chi Square" distribution].

To find Δk, $P(\delta x_v > VAL)$ and $P(r < T)$ are evaluated both at k_{min} and k_{max}, depending on the result, the procedure is the following:

If for both curves one P_{MD} value is lower than 0.5, and the other one is higher (this is the case shown in Figure F.1), then Δk must be evaluated in each curve in the vicinity of were $P_{MD} = 0.5$. To do so, four values of k are found: k_+^{Chi}, k_-^{Chi}, k_+^{Gauss} and k_-^{Gauss} such that:

$$\mathrm{P}(r < T | k_+^{Chi}) = 0.5 + \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.4}$$

$$\mathrm{P}(r < T | k_-^{Chi}) = 0.5 - \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.5}$$

$$P(\delta x_v > VAL | k_+^{Chi}) = 0.5 + \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.6}$$

$$P(\delta x_v > VAL | k_-^{Chi}) = 0.5 - \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.7}$$

These points can be observed in Figure F.1, which is out of scale to allow visualization. All values of "k" should be read in the x axis. Now the value of the search interval is determined as:

$$\Delta k = \min\left(\left(k_+^{Chi} - k_-^{Chi}\right), \left(k_+^{Gauss} - k_-^{Gauss}\right)\right) \tag{F.8}$$

If all four points (for $P(\delta x_v > VAL)$ and $P(r < T)$ evaluated both at k_{min} and $k_{max,}$) give a $P > 0.5$ (imagine the green curve shifted to the right, and the red curve shifted to the left in Figure F.1), then the inputs to Eq. (F.8) are evaluated from:

$$\mathrm{P}(r < T | k_+^{Chi}) = P_{k_{\max}} + \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.9}$$

$$\mathrm{P}(r < T | k_-^{Chi}) = P_{k_{\max}} - \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.10}$$

$$P(\delta x_v > VAL \,|\, k_+^{Chi}) = P_{k_{\min}} + \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.11}$$

$$P(\delta x_v > VAL \,|\, k_-^{Chi}) = P_{k_{\min}} - \frac{P_{MDreq} - P_{MDsearch}}{2} \; ; \tag{F.12}$$

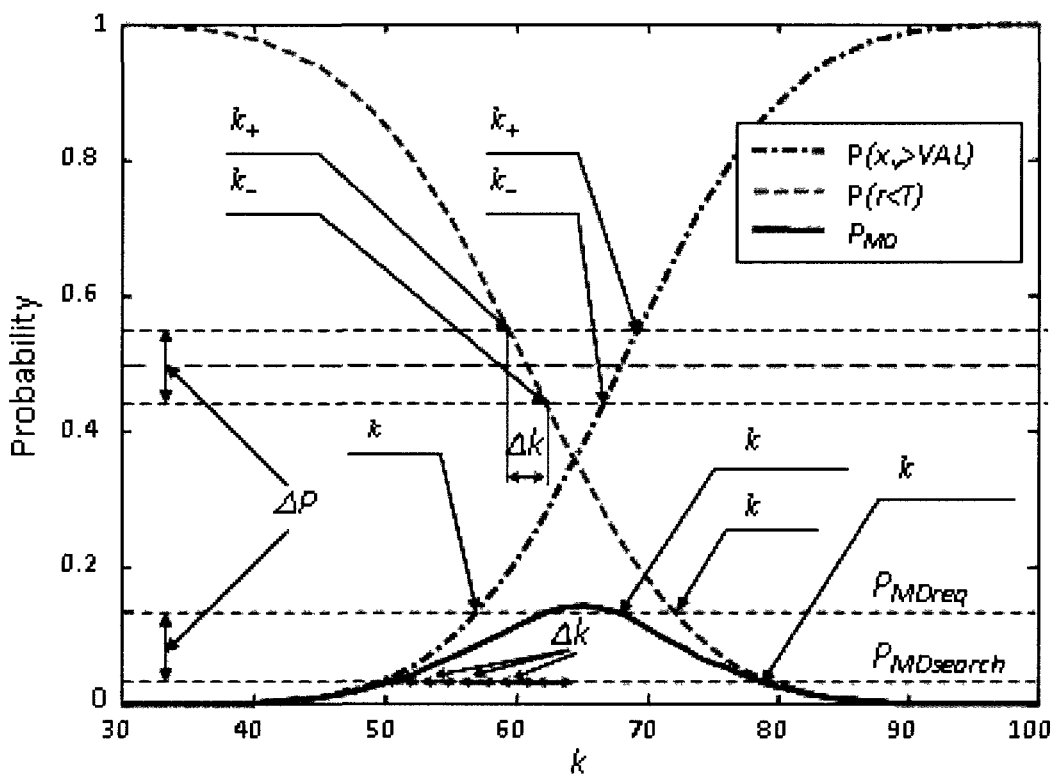


Figure F.1. Search Interval Computation

if all four points give a $P < 0.5$ (imagine shifting the green curve to the left and the red curve to the right in Figure 3.1), then Δk is evaluated at k_{min} for the Chi^2 distribution, and at k_{max} for the Gaussian. For this case, the inputs to Eq. (F.8) have to satisfy:

$$\mathrm{P}(r < T \,|\, k_+^{Chi}) = P_{k_{\min}} + \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.13}$$

$$\mathrm{P}(r < T | k_{-}^{Chi}) = P_{k_{\min}} - \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.14}$$

$$P(\delta x_v > VAL | k_{+}^{Chi}) = P_{k_{\max}} + \frac{P_{MDreq} - P_{MDsearch}}{2} \tag{F.15}$$

$$P(\delta x_v > VAL | k_{-}^{Chi}) = P_{k_{\max}} - \frac{P_{MDreq} - P_{MDsearch}}{2} \; ; \tag{F.16}$$

If the results are a mix of these cases, the rules in the paragraphs above are applied to each individual curve accordingly.

This procedure allows discarding the possibility of the existence of a $k_{min} < k_{HMI} < k_{max}$ that produces a $P_{MD} > P_{MDreq}$, in a reasonable time. If it is found such $\bar{k}_{HMI}$ exists, finding its lower and upper limits is not time consuming.

BIBLIOGRAPHY

[And00] Andreacchi, J, "Characterizing the Effect of Ionospheric Divergence on Differential GPS Navigation", *MSc Thesis.* Illinois Institute of Technology, Chicago, IL. (2000).

[Ang06] Angus, J.E., "RAIM with Multiple Faults." *NAVIGATION Journal of the Institute of Navigation. Vol. 53, No. 4,* Winter 2006. (2006).

[Bla07] Blanch, J., A. Ene, T. Walter, P. Enge, "An Optimized Multiple Hypothesis RAIM Algorithm for Vertical Guidance." *Proceeding of ION GNSS 20th International Technical Meeting of the Satellite Division,* Fort Worth, TX. (2007).

[Bre96] Brenner, M., "Integated GPS/Inertial Fault Detection Availability." *NAVIGATION Journal of the Institute of Navigation. Vol. 43, No. 2, Summer 1996.* (1996).

[Bro88] Brown, R.G., "Self-Contained GPS Integrity Check Using Maximum Solution Separation." *NAVIGATION Journal of the Institute of Navigation. Vol. 35, No. 1, Spring 1988.* (1988).

[Bro98] Brown, R.G., "GPS RAIM: Calculations of Threshold and Protection Radius Using Chi-Square Methods- A Geometric Aprroach." *GPS Monograph Series, Vol. V, The Institute of Navigation.* (1988).

[Buo99] Buosanto, M.J., "Ionospheric Storms- A review." *Space Science Reviews, 88, Kluwer Academic Publishers.* (1999).

[Cer06] Cerruti, A., "Observed GPS and WAAS Signal-to-Noise Degradation Due to Solar Radio Bursts," *Proceedings of the 19th International Technical Meeting of the Satellite Division of The Institute of Navigation*, Fort Worth, TX. (2006)

[Chi10] Chiecchio, J. "Report on Flight Banking Angles", Flight Operational Quality Assurance Program, Atlantic Southeast Airlines, Atlanta, GA. (2010).

[DoD08] USA Department of Defense, *Global Positioning System Standard Positioning System Performance Standard.* Washington, DC. (2008).

[Eld08] Eldredge, L., "Wide Area Augmentation System (WAAS) and Local Area Augmentation System (LAAS) Update" (Slides presentation to the Civil Global Positioning System Service Interface Committee (CGSIC)) (2008).

[Ene07] Ene, A., J. Blanch, J.D. Powell, "Fault Detection and Elimination for Galileo-GPS Vertical Guidance." *Proceedings of the 2007 National Technical Meeting of The Institute of Navigation,* San Diego CA, (2007).

[ESA10] European Space Agency, *ESA Official Webpage (www.esa.int)*

[FAA11] Federal Aviation Administration, 'Archive List of GPS Report Cards' *(http://www.nstb.tc.faa.gov/DisplayGPSReportCardArchive.htm)*

[Gak11] Gakstatter, Eric, 'A Closer Look at L5: The Future of High-Precision GNSS', webinar sponsored by GPS World and Geospatial Solutions, March 17th 2011.

[GEA08] GEAS Panel, *Phase I Report of the GNSS Evolutionary Architecture Study.* (2008).

[GEA10] GEAS Panel, *Phase II Report of the GNSS Evolutionary Architecture Study.* (2010).

[Gel74] Gelb, A. (editor). *Applied Optimal Estimation.* Cambridge, MA: The MIT Press. (1974).

[Gra04] Gratton, L. R., S. Khanafseh, B. Pervan, S. Pullen, J. Warburton, "Experimental Observations and Ephemeris Monitor Applications of LAAS Carrier Phase IMLA Measurements." *Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute of Navigation,* Long Beach, CA. (2004).

[Gra05a] Gratton, L. R., F. Chan, B. Pervan, "Algorithms for Airborne Ionospheric Front Detection in LAAS Using Carrier Phase and INS Measurements." *Proceedings of the 2005 National Technical Meeting of The Institute of Navigation,* San Diego, CA. (2005).

[Gra05b] Gratton, L. R., B. Pervan, "Airborne and Ground Monitors for Ionospheric Front Detection in LAAS Using Carrier Phase." *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation,* Long Beach, CA. (2005).

[Gra06] Gratton, L. R., B. Pervan, "Carrier Phase Airborne and Ground Monitors for Ionospheric Front Detection for Category III LAAS." *Proceedings of the 19th International Technical Meeting of the Satellite Division of The Institute of Navigation,* Fort Worth, TX. (2006).

[Han96] Hansen, Andrew J., T. Walter, P. Enge, "Ionospheric Correction Using Tomography," *Proceedings of the 10th International Technical Meeting of the Satellite Division of The Institute of Navigation*, Kansas City, MO. (1996)

[Heo94] Heo, M., B. Pervan, S. Pullen, J. Gautier, P. Enge, and D. Gebre-Egziabher, "Robust Airborne Navigation Algorithm for SRGPS." *Proceeding of the IEEE Position, Location, and Navigation Symposium,* Monterey, CA. (2004).

[Heg08] Hegarty, C., E. Chatre "Evolution of the Global Navigation Satellite System (GNSS)." *Proceedings of the IEEE.* (2008).

[Hwa06] Hwang, M., R. G. Brown, "RAIM FDE Revisited: A New Breakthrough in Availability Performance With NIORAIM (Novel Integrity-Optimized RAIM)." *NAVIGATION Journal of the Institute of Navigation , Vol. 53, No. 1.* Spring 2006. (2006).

[IAC11] Information Analytical Centre, "GLONASS status" *(http://www.glonass-center.ru/en/GLONASS/)*. Russian Federal Space Agency. (2011).

[ICA09] ICAO NSP Cat II/III Subgroup Group. *GAST D Time To Alert Requirements and Guidance Material.* Working Paper. (2009).

[ICA10] ICAO NSP WGW/WP14, *Validation of Ionospheric Anomaly Mitigation for GAST D Meeting*. Montreal, Canada. (2010).

[ICD00] Interface Control Document GPS-200, Revision C, Arinc Research Corporation (2000).

[Kel98] Kelly, R., "The Linear Model, RNP, and the Near Optimum Fault Detection and Exclusion Algorithm." *GPS Monograph Series, Vol. V, The Institute of Navigation. January 1998.* (1988).

[Lan09] Langel, S., S. Khanafseh, C. Fang-Cheng, B. Pervan, "Cycle Ambiguity Reacquisition in UAV Applications using a Novel GPS/INS Integration Algorithm," *Proceedings of the 2009 International Technical Meeting of The Institute of Navigation*, Anaheim, CA. (2009).

[Lee86] Lee, Y., "Analysis of Range and Position Comparison Methods as a means to provide GPS Integrity in the User Receiver." *Proceedings of the Institute of Navigation ION 42nd Annual Meeting*. Seattle, WA. (1986).

[Lee07] Lee, Y., "Two New RAIM Methods Based on the Optimally Weighted Average Solution (OWAS) Concept" *NAVIGATION Journal of the Institute of Navigation*. Vol. 54, No. 4, Winter 2007. (2007).

[Lee08] Lee, Y., M. McLaughlin, "Optimization of Position Domain Relative RAIM" *Proceedings of the 21st International Technical Meeting of the Satellite Division of The Institute of Navigation,* Savannah, GA. (2008).

[Luo03] Luo, M., S. Pullen, J. Dennis, H. Konno, G. Xie, T. Walter, P. Enge, S. Datta-Barua, T. Dehel, "LAAS Ionosphere Spatial Gradient Threat Model and Impact of LGF and Airborne Monitoring," *Proceedings of the 16th International Technical Meeting of the Satellite Division of The Institute of Navigation*, Portland, OR. (2003)

[Luo04] Luo, M, S. Pullen, A. Ene, D. Qiu, T. Walter, P. Enge, "Ionosphere Threat to LAAS: Updated Model, User Impact, and Mitigations," *Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute of Navigation*, Long Beach, CA. (2004)

[Mis01] Misra, P., P. Enge. *Global Positioning Systems*. Lincoln, MA, Ganga-Jamuna Press. (2001).

[Mur07] Murphy, T., M. Harris, "More Ionosphere Anomaly Mitigation Considerations for Category II/III GBAS" *Proceedings of the 20th International Technical Meeting of the Satellite Division of The Institute of Navigation,* Fort Worth, TX. (2007).

[NAS10] NASA. "STEREO Mission Status" *(stereo-ssc.nascom.nasa.gov/status.shtml)*. Stereo Science Center. National Aeronautics and Space Administration. (2010).

[NOA10] NOAA. "NOAAscales/index.html#GeomagneticStorms" *Space Weather Prediction Center (www.swpc.noaa.gov/)*. National Weather Center. National Oceanic and Atmospheric Administration. (2001).

[Par88] Parkinson, B., P. Axelrad, "Autonomous GPS Integrity Monitoring Using the Pseudorange Residual." *NAVIGATION Journal of the Institute of Navigation*. Vol. 35, No. 2, Summer 1988. (1988).

[Par96] Parkinson, B., J. Spilker, P. Axelrad, P. Enge (editors) *The Global Positioning System: Theory and Applications*. AIAA Progress in Aeronautics and Astronautics Volume 163-4, Washington DC., (1996).

[Per94] Pervan, B., C. E. Cohen, B. Parkinson, "Integrity in Cycle Ambiguity Resolution for GPS-Based Precision Landing." *Third International Conference on Differential Satellite Navigation Systems*. London, UK. (1994).

[Per96a] Pervan, B., C. E. Cohen, B. Parkinson, "Parity Space Methods for Autonomous Fault Detection and Exclusion Using GPS Carrier Phase." *IEEE Position, Location and Navigation Symposium*. Atlanta, GA. (1996).

[Per96b] Pervan, B. "Navigation Integrity for Aircraft Precision Landing Using the Global Positioning System." *PhD Dissertation*. Stanford, CA: Stanford University. (1996).

[Per98] Pervan, B., S. Pullen, J. Christie, "A Multiple Hypothesis Approach to Satellite Navigation Integrity." *NAVIGATION Journal of the Institute of Navigation*. Vol. 45, No. 1, Spring 1998. (1998).

[Per07] Pervan, B., "Relative RAIM" (Slides Presentation). *GEAS Meeting #3. January 11th 2007*. Arlington, VA (2007).

[Per08] Pervan, B., L. Gratton, M. Joerger, "RRAIM Phase I Results and Data Validation/Phase II Work Plans" (Slides Presentation). *GEAS Meeting #10. January 30th 2008*. San Diego, CA (2008).

[Pul09] Pullen, S., Y. S. Park, P. Enge "Impact and mitigation of ionospheric anomalies on ground-based augmentation of GNSS." *Radio Science, vol 44.* (2009).

[RTC04] RTCA, *Minimum Aviation System Performance Standards for Local Area Augmentation Systems (LAAS) DO-245A*. Washington, DC. (2004).

[RTC06] RTCA, *Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment DO-229D.* Washington, DC. (2006).

[Shi08] Shively, C. A., R. Niles "Safety Concepts for Mitigation of Ionospheric Anomaly Errors in GBAS." *Proceedings of the 2008 National Technical Meeting of The Institute of Navigation*. San Diego, CA. (2006).

[Sim06] Simili, D., B. Pervan "Code-Carrier Divergence Monitoring for the GPS Local Area Augmentation System." *IEEE Aerospace Electronic Systems Magazine*. (2006).

[Stu88] Sturza, M., "Navigation System Integrity Monitoring Using Redundant Measurements." *NAVIGATION Journal of the Institute of Navigation*. Vol. 35, No. 2, Winter 1988-89. (1988).

[USC10] United States Coast Guard, *Coast Guard Navigation Center (www.cgnc.org)*.

[vGr93] vanGraas, F., J. Farrell, "Baseline Fault Detection and Exclusion Algorithms" *Proceedings of the 49th Annual Meeting of The Institute of Navigation*. Cambridge, MA. (1993).