Analysis of Iridium-Augmented GPS for Floating Carrier Phase Positioning

MATHIEU JOERGER, LIVIO GRATTON, and BORIS PERVAN
Illinois Institute of Technology, Chicago, Illinois, 60616

CLARK E. COHEN
Coherent Navigation, Inc., San Mateo, California, 94404

Received December 2009; Revised March 2010

ABSTRACT: *Carrier-phase ranging measurements from Global Positioning System (GPS) and low-Earth-orbiting Iridium telecommunication satellites are integrated in a precision navigation system named iGPS. The basic goal of the system is to enhance GPS positioning and timing performance, especially under jamming. In addition, large satellite geometry variations generated by fast-moving Iridium spacecraft enable rapid estimation of floating cycle ambiguities. Augmentation of GPS with Iridium satellites also guarantees signal redundancy, which enables Receiver Autonomous Integrity Monitoring (RAIM). In this work, parametric models are developed for iGPS measurement error sources and for wide-area corrections from an assumed network of ground reference stations. A fixed-interval positioning and cycle ambiguity estimation algorithm is derived and a residual-based carrier-phase RAIM detection method is investigated for integrity against single-satellite step and ramp-type faults of all magnitudes and start-times. Predicted overall performance is quantified for various ground, space, and user segment configurations.*

INTRODUCTION

The iGPS navigation and communication system combines ranging measurements from the Global Positioning System (GPS) and low Earth orbit (LEO) Iridium telecommunication satellites. iGPS is a Boeing initiative now under contract from the U.S. Navy aimed at enhancing GPS timing and positioning performance, especially under jamming. In addition, while many details of the iGPS system are not public, assumptions are made herein indicating that iGPS opens the possibility for rapid, robust, and accurate floating carrier-phase positioning over wide areas (without need for local reference stations). The system's promise for real-time high-accuracy positioning performance makes it a potential navigation solution for demanding precision applications such as autonomous terrestrial and aerial transportation. The ability of iGPS to harness the exacting precision of carrier phase emanates from the concept of geometric diversity. The algorithms and analyses developed in this paper are founded on two core principles that exploit the addition of fast-moving LEO Iridium space vehicles (SVs): large changes in satellite geometry for rapid cycle ambiguity estimation and satellite redundancy for fault-detection.

The peer review of this paper was independently handled by Professor John Crassidis, Associate Editor for the *AIAA Journal of Guidance, Control, and Dynamics*.

Historical Background

Observations from the fast-moving LEO spacecraft Sputnik were at the origin of the first satellite radio-navigation system, known as the Navy Navigation Satellite System or Transit, which became operational in 1964 [1]. The Transit constellation was comprised of 4-7 LEO SVs in nearly circular, polar orbits, which broadcast radio-frequency signals with encoded orbital parameters and time corrections. Users could determine their position by tracking the apparent compression and stretching of the carrier wavelength due to spacecraft motion over 10–20 min passes. Each location in sight of the satellite observed a unique Doppler shift curve (defined as the difference between signal frequencies at the transmitter and at the receiver). As a result, *Doppler-based position* fixes were achievable several times a day (at 100 min intervals at mid-latitudes) with better than 70 meters of accuracy, which met the requirements originally intended for slow moving military vessels and

submarines [2]. It was often used in conjunction with inertial navigation systems, which were employed to correct for the added uncertainty due to user motion and to bridge gaps between infrequent position updates.

In the 1990s, Transit was superseded in both military and civilian applications by GPS, which directly utilizes range instead of range rate. Codes modulated on GPS signals provide instantaneous and absolute measurement of the travel time between the satellite transmitter and user receiver. In addition, the GPS medium Earth orbit constellation ensures that at least four SVs are continuously visible anywhere on Earth. This enables real-time determination by *trilateration* of the user's receiver clock deviation and three-dimensional position within about 10–20 m of accuracy [3]. The ultimate in GPS performance is obtained using measurements of the signal's carrier phase. Its tracking error is lower than the code's by two to three orders of magnitude, but it requires that an unknown constant cycle ambiguity be determined (receivers can only track the carrier phase modulus 2π). An efficient solution for this problem is to exploit the bias observability provided by satellite motion. Unfortunately, the large amount of time for GPS spacecraft to achieve significant changes in line of sight (LOS) precludes its use in most real-time applications.

In contrast, angular variations from LEO satellites quickly become substantial. Therefore in this work, the geometric diversity of GPS ranging sources is enhanced using additional carrier phase measurements from fast moving Iridium satellites. In fact, carrier phase observations are equal to integrated Doppler shift, so that the underlying concepts of utilizing spacecraft motion to resolve cycle ambiguities and of Transit's Doppler positioning are equivalent. Combined with GPS, real-time unambiguous carrier-phase based trilateration is possible without restriction on the user's motion. In the longer term future, integration of GPS and Iridium signals with additional ranging sources from other global navigation satellite systems (GNSS) such as GLONASS and Galileo will lead to unprecedented levels of navigation performance.

Related Work

The Integrity Beacon Landing System (IBLS) devised in the early 1990s was an explicit implementation of this principle for aircraft precision approach and landing [4, 5]. GPS signal transmitters serving as pseudo-satellites ('pseudolites') placed on the ground along the airplane's trajectory provided additional ranging sources and a large geometry change as the receiver's downward-looking antenna flew over the installation. The efficiency of IBLS was demonstrated in 1994 as it enabled 110 successful automatic landings of a Boeing 737 [4] and in 2003 for the high angle-of-attack Extremely Short Takeoff and Landing (ESTOL) of the Navy/Boeing/EADS X-31 aircraft [6]. However, pseudolite placement constraints and off-field operations and maintenance access cost has dampened wider application of the system.

By 2000, Rabinowitz et al. designed a receiver capable of tracking carrier-phase measurements from GPS and from GlobalStar (another LEO telecommunication constellation) [7]. Using GlobalStar satellites' large range variations, precise cycle ambiguity resolution and positioning was achieved within 5 min. Numerous practical issues relative to the synchronization of GPS and GlobalStar data (without modification of the SV payload) had to be overcome to obtain experimental validation results. Such considerations are outside the scope of this paper, but Rabinowitz's work offers a proof of concept for the precision applications of iGPS.

Measurement Error Models

The treatment of measurement errors plays a central part in the design of the iGPS navigation system. Error sources include uncertainties in satellite clocks and positions, signal propagation delays in the ionosphere and troposphere, user receiver noise, and multipath (unwanted signal reflections reaching the user antenna). The ionosphere, a non-uniform upper layer of the atmosphere, is the largest source of measurement error. It is a dispersive medium where refraction varies with signal frequency. Dual-frequency implementations effectively eliminate the ionospheric error thus improving the positioning accuracy, but will not be widely available for civilian applications before 2025 [8].

Instead, differential corrections can help mitigate satellite-dependent and spatially-correlated atmospheric errors. In differential GPS, measurements collected at ground reference stations are compared with the known distance between these stations and the satellites. The resulting correction accuracy varies with user-to-ground-station separation distance. In the aforementioned pseudolite and GlobalStar-augmented GPS research, the short baseline-distance from the differential reference station to the user (1–5 km) was instrumental in achieving high performance. In Rabinowitz's work in particular, residual measurement errors over short baselines could be modeled reliably enough to allow for integer cycle ambiguities to be fixed.

iGPS however aims at servicing wide areas with as little ground infrastructure as possible, and therefore relies on long-range, lower-accuracy

corrections similar to the ones generated by the Wide Area Augmentation System (WAAS). The latter has been operational since 2003 and produces a 95% positioning accuracy better than 10 m for single-frequency code-phase GPS users across the United States [9]. When using long-range corrections, the unpredictability of atmospheric effects makes it difficult to capture residual errors with high levels of confidence.

Hence, in this work, a conservative approach is adopted for the derivation of new parametric measurement error models. They account for the instantaneous uncertainty at signal acquisition (absolute measurement error) as well as variations over the signal tracking duration (relative error with respect to initialization). Unlike existing GPS measurement models used in WAAS and in the Local Area Augmentation System (LAAS), they deal with large drifts in ranging error for LEO satellite signals moving across wide sections of the atmosphere. Published data and experimental results help establish an initial knowledge of the measurement error probability distributions. They also show that the dynamics of the errors can be reliably modeled over short time periods [10].

Algorithms and Analysis

Thus, two conflicting considerations are shaping the carrier-phase iGPS estimation and detection processes that are devised in this paper: ranging measurements must be tracked for as long as possible to draw maximum benefit from changes in satellite geometry, but as this filtering duration increases, the robustness of the measurement error model decreases. In response, an upper limit on filter duration is set to ensure the model's validity, and a fixed-interval filtering algorithm is carried out for the simultaneous estimation of user position and real-valued (floating) carrier-phase cycle ambiguities. It is compatible with real-time implementations provided that sufficient memory is allocated to the storage of a finite number of past measurements and LOS coefficients.

In addition, Iridium and GPS code and carrier-phase observations collected within the filtering interval are all vulnerable to rare-event integrity threats such as user equipment and satellite failures. In this regard, the augmentation of GPS with Iridium offers an advantage in guaranteeing redundant measurements, which enables Receiver Autonomous Integrity Monitoring (RAIM) [11]. Indeed, if five or more satellites are available, the consistency of the over-determined position solution is verifiable. The accuracy of carrier-phase observations further allows for an extremely tight detection threshold while still ensuring a very low false-alarm probability. To protect the system against faults that may affect successive measurements, an innovative least-squares residual-based detection routine is developed. It is similar to an instantaneous or 'snapshot' RAIM algorithm [12], but is applied to batches of code and carrier phase measurements, whose time-correlated error distributions are bounded.

Potential applications for iGPS are investigated, including ground and aerial transportation. Specifically, the ability of iGPS to meet a predefined vertical alert limit (VAL) [13] is investigated here. Since transportation involves safety-of-life operations, special emphasis is placed on integrity. The most stringent known requirements proposed for such operations specify that no more than one undetected hazardous navigation system failure is allowed in a billion approaches [13].

Performance evaluations are structured around these example requirements. They measure the impact of Iridium's near-polar orbits on positioning and monitoring for a series of satellite geometries, thus providing detailed insights into observability mechanisms. Finally, the multidimensionality of the algorithm and the multiplicity of system parameters make the design of the navigation architecture particularly complicated. A sensitivity analysis is proposed that compares the relative influence of individual system parameters on the overall end-user output. The methodology singles out system components likely to bring about substantial performance improvement and establishes recommendations on possible orientations for future design iterations.

Outline of the Paper

The second section of this paper introduces a conceptual architecture for the iGPS space, ground, and user segments. In the third section, measurement error models applicable to short time periods are derived. They are employed in a positioning and cycle-ambiguity estimation process, which serves as basis for the residual-based RAIM detection algorithm, as developed in the fourth section. Next, the benchmark aircraft precision approach mission and its requirements are presented, which set the framework for the simulations. In the sixth section, fault-free (FF) integrity is evaluated by covariance analysis, and residual-based detection is tested for a set of canonical step and ramp-type single-satellite faults (SSF) of all magnitudes and of multiple start-times. In the last section, a sensitivity analysis of the combined FF and SSF performance brings dominant system parameters to the foreground, investigates alternative system configurations, and assesses the potential of iGPS to provide nationwide high-integrity positioning in a near-term and longer-term future.

IGPS SYSTEM DESCRIPTION

As part of this work, a nominal navigation system configuration is proposed. It was iteratively refined as a result of simulated performance sensitivity analyses (final section of the paper), so that iGPS could fulfill the intended civilian application requirements.

GPS and Iridium Satellite Constellations

In this work, a baseline GPS constellation is considered [14]. 24 satellites follow near-circular orbits at about 20,000 km altitude. They are arranged in six orbital planes of four spacecraft each, with 55 deg inclination angles, and with an orbital period, T_{GPS}, of 1/2 sidereal day. One distinctive feature of GPS satellites is that they are equipped with highly-stable atomic cesium and rubidium clocks (long-term stability on the order of 10^{-13} [15]), which are essential to the system's precise synchronization on a common time-reference for direct transit time measurements.

The Iridium satellite constellation's primary function is to provide telecommunication capabilities to users worldwide, particularly in remote places where other communication means are unavailable. Messages are exchanged between users and satellites and satellite-to-satellite cross-links enable uninterrupted communications so that any two points on the globe are connected [16]. Telecommunication satellite clocks do not meet the atomic standard, but it has been demonstrated that their clock drift can be effectively modeled and corrected using GPS measurements at the user receiver [17]. Continuous global coverage is realized using 66 spacecraft orbiting at an altitude of 780 km, which is much lower than the 20,000 km GPS orbit altitude. As a result, a spacecraft spends on average 10 minutes in view of a given location on Earth's surface, and circles the earth in a period, T_{IRI}, of 100 min 28 s [18]. As illustrated in Figure 1, the 66 satellites are distributed among six planes in near-circular orbits at 86.4 deg inclination. A 31.6 deg angle separates each co-rotating orbital plane, and the remaining 22 deg angle separates the two planes at the seam of the constellation, where spacecraft are counter-rotating [19]. Each plane contains 11 regularly spaced satellites, whose positions are offset from SVs in the adjacent co-rotating plane by one-half of the satellite spacing.

As a consequence of the constellation design, the satellite density is much higher near the poles than at lower latitudes; for example, the average number of Iridium SVs visible at any instant is 2.2 in Chicago and 1.8 in Miami. In addition, the spacecraft trajectories generate larger North-South

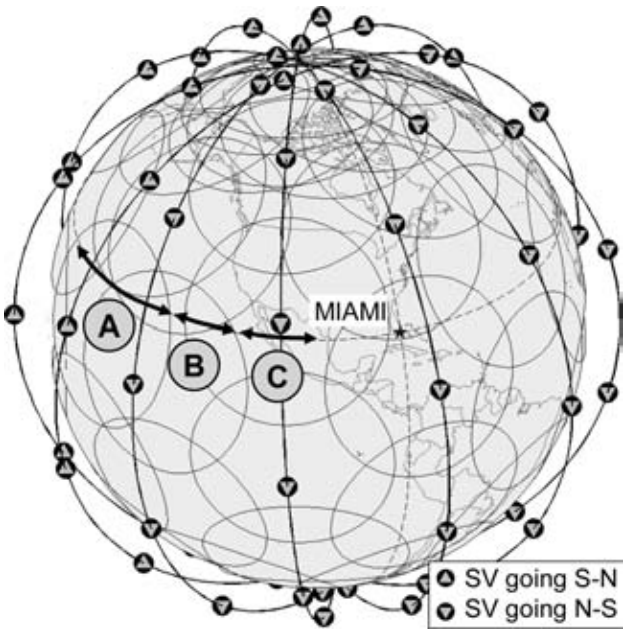

Fig. 1–Iridium Satellite Coverage

LOS variations relative to a ground observer than East-West. Accordingly, the horizontal positioning performance is heterogeneous and higher precision is generally achieved for the North coordinate.

In this work, the possibility of expanding the Iridium space segment was considered with a modified 88 Iridium satellite constellation. In this case, eight orbital planes are separated by a 23 deg angle, leaving 19 deg at the seam of the constellation. The resulting increased average spacecraft numbers over Chicago and Miami are respectively 3 and 2.4.

Finally, nominal GPS and Iridium orbits are pictured together in Figure 2(a). The constellations' orbital planes are quasi-stationary in an Earth-centered inertial (ECI) frame, whose origin is the center of the Earth and whose axes are fixed with respect to the stars. Satellites travel along the orbits, while the Earth rotates about its North-South axis. Figure 2(b) shows (from the point of view of a user at the Miami location) the large difference in accumulated angular variations between GPS and LEO satellites over a 10 minute period. In parallel, for the same location and duration, an azimuth-elevation sky plot underscores again the difference in spacecraft motion, and the North-South directionality of Iridium satellites.

Assumed Ground Segment

The GPS ground-based Operational Control Segment (OCS) makes satellite position and time synchronization information available to users. Spacecraft dynamics are modeled using observations from six ground monitoring stations spread around

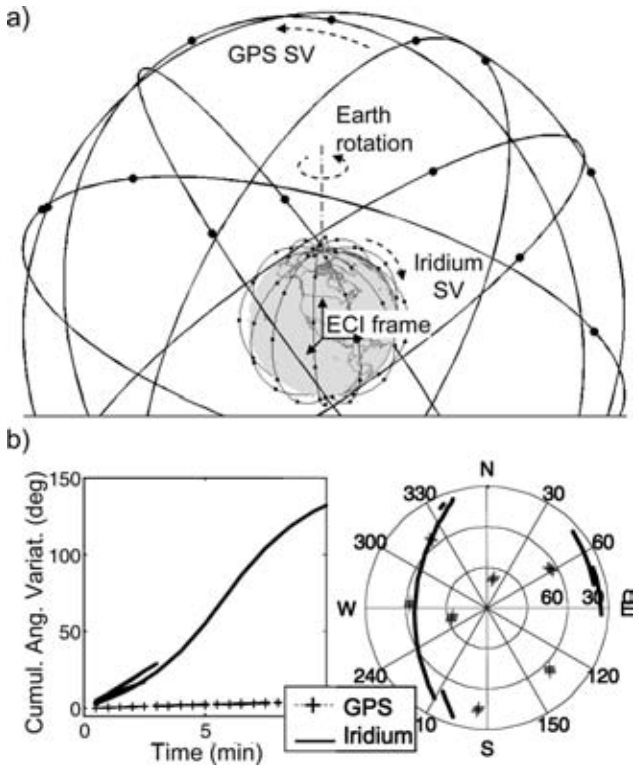

Fig. 2–Joint GPS and Iridium Constellations, (a) Orbits, (b) from a User's Perspective

the world. Orbit ephemeris parameters are computed at a master control station, uploaded to the satellites, and broadcast to users as part of the navigation message modulated on the GPS signal. The six monitoring stations are equipped with atomic clocks to establish satellite clock corrections also transmitted in the navigation message. Additional functions fulfilled by the OCS include monitoring satellite health and commanding occasional SV station-keeping maneuvers. A similar architecture is assumed for Iridium satellites. Although they cannot be continuously tracked by ground stations, they are visible several times a day.

Precision navigation requires that additional information be transmitted, in particular to correct for errors due to refraction in the ionosphere. Unlike clock errors that have a similar impact for all ground stations within a satellite's footprint, the effects of ephemeris errors and, to a greater extent, of ionospheric disturbances vary with receiver location. Indeed, the LOS to the receiver determines the section of the atmosphere crossed by the signal. Two differential approaches to mitigate ionospheric effects have been the subject of extensive studies over the past decade. The first approach, used in LAAS, aims at providing corrections at selected airport locations, based on measurements from ground receivers experiencing errors that are very similar to the aircraft's within a limited broadcast radius. The second approach, implemented in WAAS, consists of sampling the ionosphere over wide areas using observations from a network of ground stations. In WAAS, 25 wide area reference stations (WRS) spread across the United States collect dual-frequency measurements, which are used at a wide area master station (WMS) to compute ionospheric delays for a 5 deg $\times$ 5 deg latitude-longitude grid of locations [14]. The WAAS ionospheric delay estimates (as well as additional satellite ephemeris and clock corrections) are broadcast via geostationary satellites to users who can interpolate corrections at the location of interest. The precision of these estimates decreases in coastal areas due to depleted ground station coverage.

In this work, the conceptual iGPS ground segment consists of a network of ground reference stations (illustrated in Figure 3), whose density determines the accuracy of ionospheric corrections. In a first attempt to determine the overall system performance, iGPS ground stations are assumed colocated with the WRS, whose correction accuracy has been documented over the past five years [9]. WAAS-like ionospheric delay estimates are assumed to be derived at a master station (following the algorithm given in reference [20]) and broadcast to the user via Iridium communication channels. Moreover, Iridium's communication capability expands the potential to transmit data, which is limited both for GPS and for WAAS (50 bits per second (bps) for the GPS navigation message, 250 bps for WAAS corrections). This feature is exploited in the third section of this paper by assuming precise Iridium clock and ephemeris data (assuming more numerous, more frequently updated orbital and clock parameters).

User Segment

The user segment is composed of all GPS/Iridium receivers. Receivers are equipped with low-cost quartz oscillator clocks that are unstable over long durations ($10^{-6} - 10^{-9}$ over a day [15]). The deviation from GPS time introduces a nuisance parameter that can be solved for if four or more satellites are available.

The iGPS concept described in this work assumes civilian users, who can collect single-frequency L-band code and carrier ranging observations (centered at 1575 MHz for GPS (L_1) and at 1624 MHz for Iridium). Users also have access to navigation messages for each constellation and measurement corrections. In the perspective of GPS III, a modernization of GPS which plans to provide L_1 and L_5 (1176 MHz) signals to civilians by 2025 [8], dual-frequency GPS measurements are considered in the sensitivity analysis for

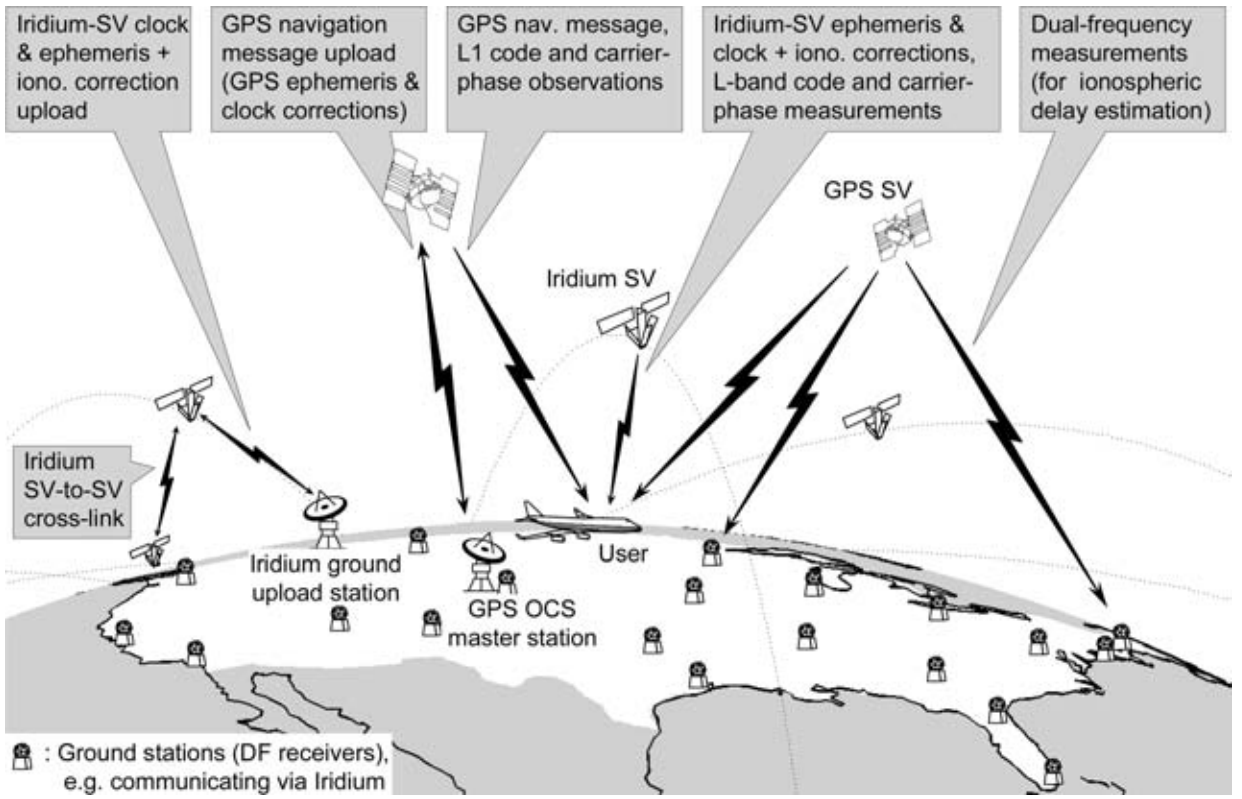

Fig. 3–Conceptual Overview of the Assumed iGPS Architecture

longer-term future implementations. In addition, Iridium satellites are equipped with Ka-band (19.5 GHz) transmitters [19]. Dual-frequency Iridium signals are therefore simulated as well, although Ka signals might be attenuated by heavy rain.

Finally as mentioned in the introduction, user equipment is assumed to include the necessary computational and memory resources to process current and past-time observations collected within a fixed filtering period, T_F, which is limited to ensure the validity of the measurement error models.

NOMINAL MEASUREMENT ERROR MODELS

Wide area differential corrections are insufficient to make residual errors negligible with respect to carrier phase tracking errors. Therefore, a conservative approach is adopted for the derivation of new parametric error models under nominal fault-free conditions. In the absence of actual Iridium measurements and of detailed information on the iGPS system (which is not publicly available), further assumptions are made on the accuracy of Iridium clock and orbit ephemeris corrections. These assumptions can in turn be considered as intermediary ground and space system requirements necessary to achieve the targeted end-user performance.

Residual Satellite Orbit Ephemeris and Clock Errors

Individual GPS satellite clocks, in spite of their high stability and corrections provided by the OCS, exhibit a slow but significant drift with respect to true GPS system time. During the short filtering interval, the remaining ranging error can be overbounded for a satellite, s, by:

- an undetermined clock bias, ${}^{s}b_C$, at the time the satellite first comes in sight, which is constant over T_F,
- plus a ramp over time with an unknown but constant gradient, ${}^{s}g_C$, accounting for linear variations from the initial value over T_F.

Based on several years of GPS data [21, 22], the initial uncertainty on ${}^{s}g_C$ is assumed normally distributed with zero mean and standard deviation, σ_{CG}, of $4{\cdot}10^{-4}$ m/s. The following notation is used in the rest of the paper:

$$ {}^{s}g_C \sim \mathrm{N}(0, \sigma_{CG}). $$

The same gradient model is applied to Iridium satellites. The satellite clock bias is modeled as ${}^{s}b_C \sim \mathrm{N}(0,\sigma_{CB})$, where σ_{CB-GPS} is approximately 1.5 m for GPS [15]. The corresponding value for Iridium is addressed later when combined with the ephemeris bias.

Another primary source of error stems from the orbit ephemeris parameters computed by the ground segment. In reference [23], several years of broadcast GPS ephemeris data were compared to decimeter-level precision post-processed satellite (truth) positions: from 1997 to 2003, daily root-mean-square ranging errors due to orbit parameter errors remain around 1.1 m. In addition, Reference [24] investigates the sensitivity over 24 hour periods of computed satellite positions to individual ephemeris parameter errors. It shows that the most sensitive parameter is the inclination angle, which causes satellite position deviations to vary periodically with the orbital period, T_{GPS}. This is further corroborated in the 24-hour broadcast ephemeris error plots of References [22, 23].

For short filtering durations, T_F (10 minutes or less) relative to T_{GPS}, changes in orbit errors are linear [10]. In this work, a worst slope approximation is used to evaluate the ephemeris gradient, ${}^s g_E$, assuming periodic variations with frequency $2\pi/T_{GPS}$ and normally distributed amplitude m_{E-GPS} (with zero mean and standard deviation σ_{E-GPS} of 1.1 m [23]):

$$ {}^s g_{E-GPS} = {}^s m_{E-GPS} \frac{2\pi}{T_{GPS}}, $$

so that ${}^s g_{\mathrm{E-GPS}} \sim \mathrm{N}(0,\sigma_{EG-GPS})$ with $\sigma_{EG-GPS} = \sigma_{E-GPS} 2\pi/T_{GPS}$. In addition, the GPS ephemeris bias is modeled as ${}^s b_{E-GPS} \sim \mathrm{N}(0,\sigma_{EB-GPS})$ (with $\sigma_{EB-GPS} = \sigma_{E-GPS}$). Since Iridium benefits from higher communication data rates, more numerous and more frequently updated orbital parameters can be exploited. The proposed orbit error model for Iridium is similar to GPS, with $\sigma_{EG-IRI} = \sigma_{E-IRI} 2\pi/T_{IRI}$. A lower sigma value is assumed ($\sigma_{E-IRI} = 0.1$ m). The value chosen is consistent with that realistically achievable in near-real-time using GPS receivers onboard the LEO spacecraft [25, 26]. Decimeter-level performance is also assumed for residual Iridium satellite clock errors and can be considered as a correction accuracy requirement on the iGPS ground and space segments.

The combined ephemeris and clock bias is modeled as ${}^s b_{EC} \sim \mathrm{N}(0,\sigma_{ECB})$ [15, 22], and different values are allocated to σ_{ECB} for GPS and Iridium models (respectively, $\sqrt{\sigma^2_{CB-GPS} + \sigma^2_{EB-GPS}}$ and 0.1 m). Finally, the combined ephemeris and clock gradient is defined as:

$$ {}^s g_{EC} = {}^s g_C + {}^s g_E, $$

so that ${}^s g_{EC} \sim \mathrm{N}(0, \sqrt{\sigma^2_{CG} + \sigma^2_{EG}})$. Altogether at epoch k of the filtering interval, for a satellite, s, that has been visible over a period Δt_k (from filter initiation at t_0 to the sample time of interest t_k), GPS and Iridium SV-related errors are expressed as:

$$ {}^s \varepsilon_{SAT,k} = {}^s b_{EC} + \Delta t_k \cdot {}^s g_{EC}. \quad (1) $$

Residual Ionospheric Error

The ionosphere is a layer of the atmosphere extending from an altitude of 50 km to 1000 km above Earth. It is composed of electrons and charged atoms and molecules of gases that have been ionized by solar ultraviolet radiation. The resulting non-uniform density of electrons causes changes in the satellite signal propagation speed that vary with geomagnetic latitude, time of day, season, and level of activity in the 11-year long solar cycle.

The ionosphere is the largest source of uncertainty in SV ranging observations. It generates a delay in code measurements and an advance of equal magnitude in carrier-phase data, which are proportional to the total electron content in the path of the signal, and to the inverse square of the carrier's frequency. As mentioned in the introduction, this frequency-dependence is exploited in dual-frequency architectures to eliminate ionospheric disturbances. Approximately 50% of the error for single-frequency users can be removed using Klobuchar's empirical model, whose parameter values are broadcast by GPS [27].

The residual ionospheric error model implemented in this work hinges on three major assumptions. First, consider an Earth-centered sun-fixed (ECSF) frame whose x-axis is pointing toward the sun and whose z-axis is Earth's axis of rotation. Under anomaly-free conditions, the ionosphere's slow dynamics in the mid-latitude temperate zones justifies that it be assumed constant over the short mission duration in an ECSF frame [10, 28, 29]. In Figure 4, the varying thickness of the egg-shaped grey area surrounding Earth represents the non-uniform electron density in the ionosphere, and is fixed in ECSF.

Second, the peak electron density occurs between 250 km and 400 km above Earth's surface. A spherical thin shell approximation is typically adopted to localize the effect of the ionosphere. An ionospheric pierce point (IPP) is defined as the intersection

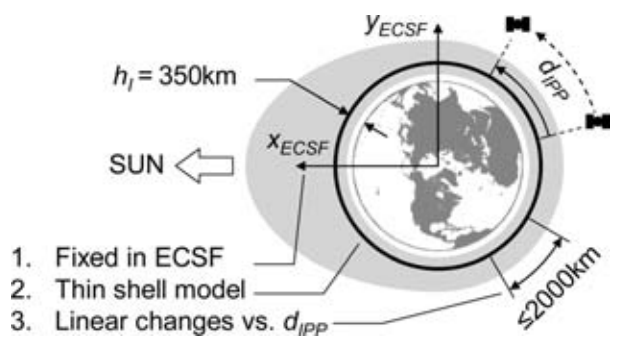

Fig. 4–Three Assumptions for the Ionospheric Error Model

between the satellite-user LOS and the thin shell at an altitude, h_I, of 350 km. IPP displacement in ECSF coordinates is due to the relative motions of the SV, of the user close to Earth's surface, and of the Earth itself. In most precision applications (including aircraft final approach with a relatively slow vehicle velocity of 70 m/s), GPS IPPs move mostly westward in ECSF (especially for high-elevation satellites) because the rotation of the Earth is the dominant factor (surface velocity larger than 200 m/s at latitudes lower than 65 deg). Iridium IPPs in contrast move along a North-South axis due to fast satellite motion. The effect of Earth's rotation is highlighted in Figure 5: the IPP displacement relative to some initial time is plotted over 10 minutes, for a user at a Miami location, in Earth-Centered-Earth-Fixed (ECEF) and ECSF reference frames.

Third, dual-frequency GPS data as well as extensive LAAS and WAAS-motivated research reported in Appendix I suggest that the vertical ionospheric delay varies linearly with IPP separation distances (actually 'great circle distances' or GCD) of up to 2000 km (it levels off for larger distances) [30–33]. The distribution of the corresponding slope can be bounded by a Gaussian model [30, 31].

The equivalent delay or advance is therefore modeled as an initial vertical ionospheric bias, b_{VI}, associated with a ramp, whose slope over IPP displacement, d_{IPP} (in an ECSF frame), is the vertical ionospheric gradient, g_{VI}. A single gradient per SV accounts for ramps along one direction only, which means that the model assumes pierce points trace approximately straight paths along the great circle, with little or no lateral motion (as seen in Figure 5). An obliquity factor, c_{OI}, adjusts this error for the fact that the LOS does not pierce the ionosphere perpendicularly, but with a slant angle function of the satellite elevation angle, ${}^s\theta$. As a result,

$$ {}^s\varepsilon_{I,k} = {}^s c_{OI,k} \cdot \left({}^s b_{VI} + d_{IPP,k} \cdot {}^s g_{VI} \right) \qquad (2) $$

with

$$ {}^s c_{OI,k} = 1 \Big/ \sqrt{1 - \left[R_E \cos({}^s\theta_k) / (R_E + h_I) \right]^2}, $$

where R_E is the radius of the Earth (6378 km).

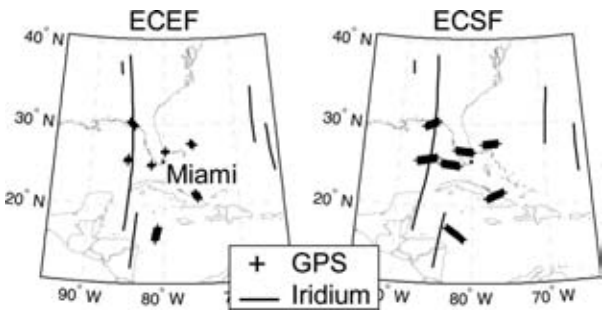

Fig. 5–IPP Displacement

Initial uncertainties on the bias and gradient are modeled as:

$$ {}^s b_{VI} \sim \mathrm{N}(0, \sigma_{VIB}) \ \text{ and } \ {}^s g_{VI} \sim \mathrm{N}(0, \sigma_{VIG}). $$

Values allocated to these parameters are justified in Appendix I using published data [9, 10, 15, 30–34], experimental results, and a covariance analysis of the WAAS algorithms [20]. Cases with and without corrections from the iGPS network of ground reference stations are considered. Hence, values of 1 mm/km and 4 mm/km for the corrected and uncorrected σ_{VIG}, respectively, and of 1–2 m and 5 m for the corrected and uncorrected σ_{VIB} are implemented.

Finally, the maximum GCD traveled by an Iridium IPP when occasionally crossing the sky with near-zero azimuth amounts to 3300 km. With a bounding σ_{VIG} value, linearity for distances beyond the suggested 2000 km limit makes the model overly conservative (in the sense that the ranging error due to ionosphere is over-bounded), but has small consequences on the overall performance results. Therefore, Eq. (2) is used to model ionospheric errors affecting both GPS and Iridium signals. A second order spatial correlation model may be considered for future work. As mentioned in Appendix I, further analysis of the ionosphere using experimental data is in progress.

Residual Tropospheric Error

Signal refraction in the troposphere, the lower part of Earth's atmosphere, delays the transmission of SV measurements. The troposphere is made of electrically neutral gases not uniform in composition, including dry gases whose behavior is largely predictable, and water vapor, which is generally treated as random but represents a much smaller fraction of the error. The majority of the delay can therefore be removed by troposphere modeling [35].

The residual uncertainty is modeled as a zenith tropospheric delay, b_{ZT} (i.e., associated with a hypothetical signal coming from 90 deg elevation), which is constant over the time interval T_F. In addition, user motion causes variations relative to this initial value, which are captured by a LAAS-like residual tropospheric error model [13] expressed as a function of the local air refractivity index, Δn, so that the total zenith tropospheric error is:

$$ \varepsilon_{ZT,k} = b_{ZT} + 10^{-6} h_0 \left(1 - e^{-\Delta h_k / h_0} \right) \cdot \Delta n. $$

Here, Δh_k designates the difference in height that the user (e.g., aircraft) experiences from the start

Table 1—Summary of Error Parameter Values

Parameter Description (σ: standard deviations, a.f.i.: at filter initialization)	Nominal
σ_{CG} : residual satellite clock gradient a.f.i.	$4{\cdot}10^{-4}$ m/s
σ_{CB-GPS} : residual GPS satellite clock bias a.f.i.	1.5 m
σ_{E-GPS} : amplitude of GPS ephem. error variations a.f.i.	1.1 m
σ_{E-IRI} : amplitude of Iridium ephem. error variations a.f.i.	0.1 m
$\sigma_{ECB-IRI}$: residual Iridium SV clock and ephemeris bias a.f.i.	0.1 m
σ_{ZTB} : residual zenith tropospheric bias a.f.i.	0.12 m
$\sigma_{\Delta n}$: residual refractivity index a.f.i. (unit-less)	30
σ_{VIB} : vertical ionospheric bias a.f.i. (corrected)[a]	1 m
σ_{VIG} : vertical ionospheric gradient a.f.i. (uncorrected)[a]	4 mm/km
$\sigma_{RN-\rho}$: code-phase receiver noise[b]	0.3 m
$\sigma_{RN-\phi}$: carrier-phase receiver noise[b]	0.003 m
$\sigma_{M-\rho}$: code-phase multipath noise[b]	1 m
$\sigma_{M-\phi}$: carrier-phase multipath noise[b]	0.01 m
$T_{M,GPS}$: GPS multipath time constant	1 min
$T_{M,IRI}$: Iridium multipath time constant	1 s
T_K : sampling interval	30 s

[a] for dual-frequency at f_1 and f_2 (GPS: L_1/L_5, Iridium: L/Ka): the terms in b_{VI} and g_{VI} are eliminated
[b] for dual-frequency (f_1, f_2), these terms are multiplied by $([f_1^2/(f_1^2 - f_2^2)]^2 + [f_2^2/(f_1^2 - f_2^2)]^2)^{1/2}$

of the filtering interval to epoch k. A fixed value of 15 km is assigned to the tropospheric scale height, h_0. Notations are simplified as follows:

$$\varepsilon_{ZT,k} = b_{ZT} + c_{T,k} \cdot \Delta n,$$

where

$$c_{T,k} = 10^{-6} h_0 \left(1 - e^{-\Delta h_k / h_0}\right).$$

An obliquity factor, ${}^s c_{OT,k}$, is applied [13] (${}^s c_{OT,k} = (0.002 + \sin^2({}^s\theta_k))^{-1/2}$) because a larger segment of the signal's path travels through the troposphere at lower elevations:

$${}^s\varepsilon_{T,k} = {}^s c_{OT,k}\left(b_{ZT} + c_{T,k} \cdot \Delta n\right). \quad (3)$$

The parameters b_{ZT} and Δn are not satellite-dependent because they characterize the environment surrounding the airplane (Eq. (3) applies to GPS and Iridium signals). They are modeled as random constants over the time interval T_F such that $b_{ZT} \sim \mathrm{N}(0,\sigma_{ZTB})$ [14] and $\Delta n \sim \mathrm{N}(0,\sigma_{\Delta n})$ [13], where the nominal standard deviations are given in Table 1.

Receiver Noise and Multipath

Receiver noise and multipath errors depend on the signal structure, signal to noise ratio, antenna design, and receiver electronics. A signal can typically be tracked to within approximately 1% of a cycle, which explains the difference of two orders of magnitude for the receiver measurement noise of GPS code (300 m chip length) and carrier-phase (19 cm wavelength). The code and carrier-phase receiver noise (${}^s\varepsilon_{RN\text{-}\rho,k}$ and ${}^s\varepsilon_{RN\text{-}\phi,k}$) for both GPS and Iridium are modeled as normally distributed white sequences with standard deviations, $\sigma_{RN\text{-}\rho}$ and $\sigma_{RN\text{-}\phi}$, respectively [36].

In addition, multipath error caused by unwanted signal reflections reaching the user receiver depends on the satellite geometry, on the environment surrounding the antenna, and on the antenna technology. Signals reaching the antenna at elevations lower than five degrees are not considered (which eliminates low-elevation reflections and accounts for minor changes in user dynamics). To account for the time-correlation it introduces, multipath is modeled as a first-order Gauss-Markov Process (GMP) with time constant, T_M, standard deviation, $\sigma_{M\text{-}\phi/\rho}$ [36], and driving noise, $v_{M,k}$:

$${}^s\varepsilon_{M,k+1} = e^{-T_K/T_M} \cdot {}^s\varepsilon_{M,k} + v_{M,k}$$

with

$$v_{M-\phi/\rho,k} \sim \mathrm{N}\left(0, \sigma_{M-\phi/\rho}\sqrt{1 - e^{-2T_K/T_M}}\right).$$

Large azimuth-elevation variations generate fast changes in the directions of signal reflections for Iridium. The multipath time-constant for Iridium, $T_{M,IRI}$, was therefore computed by multiplying the angular rate ratio between GPS and Iridium satellites (approximately 0.017) with the time constant for GPS, $T_{M,GPS}$ (conservatively estimated at 60 s in a dynamic environment [37]). Current receivers are capable of tracking signals with high update rates (higher than 1 Hz). Within the filtering interval, T_F, there is limited benefit in using all the available measurements both because of their correlation in time, and because the contribution of geometry change to the estimation process far outweighs that of redundant measurement averaging. Therefore past measurements within T_F are considered at regular intervals, T_K, of 30 s, which greatly decreases the computational burden. Finally, the combined receiver noise and multipath (RNM) is expressed as:

$${}^s\varepsilon_{RNM,k} = {}^s\varepsilon_{RN,k} + {}^s\varepsilon_{M,k} \quad (4)$$

In summary, the complete linearized GPS and Iridium carrier-phase measurement equation for a

satellite, s, at epoch, k, can be written as (using Eq. (1–4)):

$$
\begin{aligned}
{}^{s}\phi_k = {}^{s}\mathbf{g}_k^T\mathbf{u}_k + {}^{s}N + {}^{s}b_{EC} + \Delta t_k \cdot {}^{s}g_{EC} \\
+ {}^{s}c_{OT,k}\left(b_{ZT} + c_{T,k}\cdot\Delta n\right) \\
- {}^{s}c_{OI,k}\left({}^{s}b_{VI} + {}^{s}d_{IPP,k}\cdot {}^{s}g_{VI}\right) + {}^{s}\varepsilon_{RNM-\phi,k}
\end{aligned} \quad (5)
$$

where $\mathbf{u}_k = [x_E\ x_N\ x_U\ \tau]_k^T$ is the user position (in a local East-North-Up or ENU coordinate frame) and receiver clock deviation, and N is the cycle ambiguity bias. The vector, ${}^{s}\mathbf{g}_k$, is a function of the unit LOS vector between user and satellite (noted ${}^{s}\mathbf{e}_k$) such that ${}^{s}\mathbf{g}_k^T = [-{}^{s}\mathbf{e}_k^T\ 1]$. The equation for the code-phase measurement, ${}^{s}\rho_k$, is identical except for the absence of ${}^{s}N$, a positive sign on the ionospheric error, and the receiver noise, ${}^{s}\varepsilon_{RNM\text{-}\rho,k}$, which replaces ${}^{s}\varepsilon_{RNM\text{-}\phi,k}$.

The assertion that error models are conservative is only true if the Gaussian models over-bound the cumulative distribution functions of each error sources' ranging errors [38]. The next phase of this research will involve establishing probability distributions for the error parameters, and verifying the fidelity of the dynamic models to experimental data. Alternatively, error parameter values may be considered as requirements that ground corrections should meet in order to achieve the desired system performance.

Nominal values for all error parameters are listed in Table 1. They constitute the *prior knowledge* acquired from experimental observations of physical phenomena, and are a crucial input to the estimation algorithm.

ESTIMATION AND DETECTION ALGORITHM DERIVATION

Position and Cycle Ambiguity Estimation

The satellite clock and ephemeris as well as the ionosphere and the troposphere error models derived previously all assume that measurements are collected over a short, limited duration. In order not to exceed this period of validity, optimal position and cycle ambiguity estimation is performed using a fixed-interval filtering algorithm. Continuous real-time operation requires that measurements and LOS coefficients be stored over the interval T_F: at each new epoch, t_k, incoming data is updated and the oldest data at $t_{k-1} - T_F$ can be erased from memory. Current-time optimal state estimates are obtained from iteratively feeding the stored finite sequence of observations (from $t_k - T_F$ to t_k) into a Kalman filter (KF). The KF also provides an indication of the estimation uncertainty in the form of a state covariance matrix, which serves as the basis for the upcoming analysis. State augmentation is used to integrate the dynamics of the multipath GMP [39]. Practical implementation of the KF also necessitates that rows and columns for all vectors and matrices (including covariance matrices) be added and removed as satellites come in and out of sight (which is frequent for Iridium).

In addition, measurement faults, whether they affect recent or older observations within the interval, are just as likely to generate hazardous current-time positioning information. In anticipation of the RAIM-type residual-based fault detection introduced in the next subsection, a fixed-interval smoothing (instead of filtering) process is used; it can be efficiently realized using, for example, a forward-backward iterative smoother, or a Rauch-Tung-Striebel algorithm [40]. Although smoothing is computationally more intensive than filtering, the computation time is negligible with respect to the specified one-second time-to-alarm limit [13]. Finally, in applications where timing and computational load are not of primary concern, measurements can be processed as a batch, which is the method presented below for clarity in exposition. Batch processing produces results identical to the KF for the current time as well as optimal estimates for past epochs that are used for residual generation.

Consider first the vector of carrier-phase observations for a satellite, s, in view between epochs k_O and k_F:

$$
{}^{s}\boldsymbol{\varphi} = \left[\,{}^{s}\phi_{k_O}\cdots{}^{s}\phi_{k_F}\right]^T.
$$

These epochs are generally the first and last of the smoothing interval for GPS signals, but not for Iridium satellites, whose passes are often shorter than T_F. Let $\mathbf{0}_{n\times m}$ be an $n \times m$ matrix of zeros. State coefficients defined in Eq. (5) are arranged in matrices that are needed in later steps, so that for spacecraft s:

$$
{}^{s}\mathbf{G} = \begin{bmatrix} {}^{s}\mathbf{g}_{k_O}^T & & \mathbf{0}_{1\times 4} \\ & \ddots & \\ \mathbf{0}_{1\times 4} & & {}^{s}\mathbf{g}_{k_F}^T \end{bmatrix},
$$

$$
{}^{s}\Delta\mathbf{t} = \left[0\ \ \Delta t_{k_1}\ \ \cdots\ \ \Delta t_{k_F}\right]^T,
$$

$$
{}^{s}\mathbf{c}_{OT} = \left[\,{}^{s}c_{OT,k_O}\ \ \cdots\ \ {}^{s}c_{OT,k_F}\right]^T,
$$

$$
{}^{s}\mathbf{c}_{T} = \left[0\ \ {}^{s}c_{OT,k_1}\ \ c_{T,k_1}\ \ \cdots\ \ {}^{s}c_{OT,k_F}\ \ \cdot\ c_{T,k_F}\right]^T,
$$

$$
{}^{s}\mathbf{c}_{OI} = \left[\,{}^{s}c_{OI,k_O}\ \ \cdots\ \ {}^{s}c_{OI,k_F}\right]^T \text{ and}
$$

$$
{}^{s}\mathbf{c}_{I} = \left[0\ \ {}^{s}c_{OI,k_1}\cdot{}^{s}d_{IPP,k_1}\ \ \cdots\ \ {}^{s}c_{OI,k_F}\cdot{}^{s}d_{IPP,k_F}\right]^T.
$$

Carrier-phase observations for all n_S Iridium and GPS satellites are then stacked together:

$$
\boldsymbol{\varphi} = \left[\,{}^{1}\boldsymbol{\varphi}^T\ \ \cdots\ {}^{n_S}\boldsymbol{\varphi}^T\right]^T,
$$

and

$$\boldsymbol{\varphi} = \mathbf{H}_{\varphi}\mathbf{x} + \boldsymbol{v}_{\varphi}, \tag{6}$$

where $\boldsymbol{v}_{\varphi}$ designates the carrier-phase measurement noise vector and $\mathbf{H}_{\varphi}$ will be defined below. The state vector is

$$\mathbf{x} = \left[\mathbf{u}_{k_O}^T \ \cdots \ \mathbf{u}_{k_F}^T \ \mathbf{N}^T \ \mathbf{b}_{EC}^T \ \mathbf{g}_{EC}^T \ b_{ZT} \ \Delta n \ \mathbf{b}_{VI}^T \ \mathbf{g}_{VI}^T\right]^T,$$

where k_O and k_F (subscripts of $\mathbf{u}$) are now the first and last epochs of the smoothing interval. Bold face characters for parameters other than $\mathbf{u}_k$ designate vectors of states for all satellites, for example:

$$\mathbf{N} = \left[\ {}^1N \ \cdots \ {}^{n_S}N\right]^T,$$

The dynamics of the user position and clock deviation vector, $\mathbf{u}_k$, are unknown; $\mathbf{u}_k$ is therefore allocated different states at each time step, as opposed to the other parameters that are modeled as constants over interval T_F.

The observation matrix, $\mathbf{H}_{\varphi}$, is constructed by blocks:

$$\mathbf{H}_{\varphi} = [\mathbf{G} \ \mathbf{B}_N \ \mathbf{B}_{ECB} \ \mathbf{B}_{ECG} \ \mathbf{B}_{ZTB} \ \mathbf{B}_{\Delta n} \ \mathbf{B}_{VIB} \ \mathbf{B}_{VIG}];$$

each block corresponds to a state parameter, and contains coefficients for all spacecraft, for the entire sequence of measurements. Let $n_K(s)$ be the number of samples for satellite s (which generally differs for Iridium SVs), and $\mathbf{1}_n$ be an $n \times 1$ column-vector of 1′s:

$$\mathbf{G} = \left[\ {}^1\mathbf{G}^T \ \cdots \ {}^{n_S}\mathbf{G}^T\right]^T,$$

$$\mathbf{B}_N = \mathbf{B}_{ECB} = \begin{bmatrix} \mathbf{1}_{n_K(1)} & & 0 \\ & \ddots & \\ 0 & & \mathbf{1}_{n_K(n_S)} \end{bmatrix},$$

$$\mathbf{B}_{ECG} = \begin{bmatrix} {}^1\Delta\mathbf{t} & & 0 \\ & \ddots & \\ 0 & & {}^{n_S}\Delta\mathbf{t} \end{bmatrix},$$

$$\mathbf{B}_{ZTB} = \begin{bmatrix} {}^1\mathbf{c}_{OT} \\ \vdots \\ {}^{n_S}\mathbf{c}_{OT} \end{bmatrix}, \quad \mathbf{B}_{\Delta n} = \begin{bmatrix} {}^1\mathbf{c}_T \\ \vdots \\ {}^{n_S}\mathbf{c}_T \end{bmatrix}$$

$$\mathbf{B}_{VIB} = -\begin{bmatrix} {}^1\mathbf{c}_{OI} & & 0 \\ & \ddots & \\ 0 & & {}^{n_S}\mathbf{c}_{OI} \end{bmatrix},$$

$$\text{and} \quad \mathbf{B}_{VIG} = -\begin{bmatrix} {}^1\mathbf{c}_I & & 0 \\ & \ddots & \\ 0 & & {}^{n_S}\mathbf{c}_I \end{bmatrix}.$$

A measurement equation similar to (6) is established for the code-phase observation vector, $\boldsymbol{\rho}$. In this case, the sign on the ionospheric coefficients, $\mathbf{B}_{VIB}$ and $\mathbf{B}_{VIG}$, is positive. Also, the columns of ones in $\mathbf{B}_N$ corresponding to the cycle ambiguity vector, $\mathbf{N}$, are replaced by zeros; this explains why state vectors, $\mathbf{N}$ and $\mathbf{b}_{EC}$, have to be distinguished, even though columns of $\mathbf{B}_N$ and $\mathbf{B}_{ECB}$ are linearly dependent for carrier-phase measurements. Since no prior knowledge is assumed for $\mathbf{N}$, the system's performance sensitivity to $\mathbf{b}_{EC}$ (investigated later) reflects the influence of code-phase measurements. In addition, since the integer nature of the constant cycle ambiguities, $\mathbf{N}$, is not exploited, potential nuisance parameters on carrier-phase data that are constant over time (such as inter-frequency biases for dual-frequency signals) can be lumped together with $\mathbf{N}$ without affecting the results.

The complete sequence of code and carrier-phase signals for all satellites over the smoothing interval are included into a batch measurement vector:

$$\mathbf{z}_B = \left[\boldsymbol{\varphi}^T \ \ \boldsymbol{\rho}^T\right]^T,$$

and

$$\mathbf{z}_B = \mathbf{H}_B\mathbf{x} + \mathbf{v}_B. \tag{7}$$

The measurement noise vector, $\mathbf{v}_B$, is utilized to introduce the time-correlated noise due to multipath modeled as a GMP. Its covariance, $\mathbf{V}_B$, is block diagonal, each block corresponding to observations from the same SV over time. Within each block, the time-correlation between two measurements originating from the same satellite s at sample times t_i and t_j is modeled as $\sigma^2_{M\text{-}\rho/\phi} \cdot e^{-\Delta t_{ij}/T_M}$, where $\Delta t_{ij} = |t_i - t_j|$. The quantities $\sigma^2_{RN-\rho}$ and $\sigma^2_{RN\text{-}\phi}$ are also added to the diagonal elements to account for code-phase and carrier-phase uncorrelated receiver noise, respectively.

Finally as mentioned in the prior section, valuable information is gained from the study of the physical phenomena causing measurement disturbances. This prior knowledge of the error parameters is expressed in terms of bounding values on their probability distributions. It can be included as a vector of pseudo-measurements, $\mathbf{z}_P$, that provide direct *a priori* observations to the corresponding states. Let n_S, n_E, and n_K be respectively the numbers of available satellites, of error states and of samples over T_F; in addition, let $\mathbf{I}_n$ be an $n \times n$ identity matrix. $\mathbf{H}_P$ is defined as $[\mathbf{0}_{n_E \times (4n_K + n_S)} \ \mathbf{I}_{n_E}]$, so that:

$$\mathbf{z}_P = \mathbf{H}_P\mathbf{x} + \mathbf{v}_P. \tag{8}$$

The covariance matrix, $\mathbf{V}_P$, of the pseudo-measurement noise vector, $\mathbf{v}_P$, is diagonal, with values of

the initial conditions on the error states, $\mathbf{b}_{EC}$, $\mathbf{g}_{EC}$, b_{ZT}, Δn, $\mathbf{b}_{VI}$, and $\mathbf{g}_{VI}$. In other words, the diagonal vector of $\mathbf{V}_P$ is:

$$\left[\mathbf{1}_{n_S}^T \cdot \sigma_{ECB}^2 \; \mathbf{1}_{n_S}^T \cdot \sigma_{ECG}^2 \; \sigma_{ZTB}^2 \; \sigma_{\Delta n}^2 \; \mathbf{1}_{n_S}^T \cdot \sigma_{VIB}^2 \; \mathbf{1}_{n_S}^T \cdot \sigma_{VIG}^2\right].$$

The elements of $\mathbf{z}_P$ are the mean values of the error states (zeros in this case). The vector $\mathbf{z}_P$ can be added to the system by direct augmentation of $\mathbf{z}_B$, in which case $\mathbf{z} = [\mathbf{z}_B^T \; \mathbf{z}_P^T]^T$, $\mathbf{H} = [\mathbf{H}_B^T \; \mathbf{H}_P^T]^T$, and the covariance matrix, $\mathbf{V}$, of the measurement noise, $\mathbf{v}$, is block diagonal with $\mathbf{V}_B$ and $\mathbf{V}_P$ on its diagonal. It can also be incorporated using the equivalent reduced-order form proposed in Appendix II. An alternative derivation based on the method of Lagrange multipliers and without introducing the concept of pseudo-measurements is available in Reference [40]. The total measurement equation becomes

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v}. \tag{9}$$

The weighted least squares state estimate is defined as $\hat{\mathbf{x}} = \mathbf{S}\,\mathbf{z}$, where $\mathbf{S}$ is the weighted pseudo-inverse of $\mathbf{H}$:

$$\mathbf{S} = \left(\mathbf{H}^T\mathbf{V}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{V}^{-1}, \tag{10}$$

and the state covariance matrix is:

$$\mathbf{P}_{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{V}^{-1}\mathbf{H}\right)^{-1}.$$

The diagonal element of $\mathbf{P}_{\mathbf{x}}$ corresponding to the current-time vertical position covariance (noted σ_U^2) is used in the next sections to determine the positioning performance under fault-free conditions. The focus is on the Up-coordinate, both because of the tighter requirements in this direction (see the next section) and because of the generally poorer satellite geometry due to the absence of open LOS below the horizon.

RAIM-Type Single-Satellite Fault (SSF) Detection Algorithm

State estimation is based on a history of observations, all of which are vulnerable to equipment faults (satellite clock excessive acceleration, corrupted ephemeris parameter, user receiver cycle slip, etc. [13]) or unusual atmospheric phenomena. To protect the system against abnormal events potentially affecting successive observations, a RAIM-type detection algorithm is derived based on the batch least-squares residual.

The residual-based RAIM methodology is articulated around two dimensions. First, let $\mathbf{s}_{U,k_F}^T$ be the row of $\mathbf{S}$ corresponding to the vertical position at the last epoch of the filtering interval k_F. The corresponding positioning error due to a measurement fault vector, $\mathbf{f}$ (of the same dimension as $\mathbf{z}$), whose non-zero elements introduce deviations from normal FF conditions, is such that:

$$\delta x_U \sim \mathrm{N}\left(\mathbf{s}_{U,k_F}^T\mathbf{f}, \sigma_U\right).$$

The system is said to produce hazardous information if a failure causes a vertical position error that exceeds a vertical alert limit, *VAL*:

$$|\delta x_U| > VAL;$$

Second, a failure may be detected using a residual vector, $\mathbf{r}$, defined as the difference between a possibly faulty measurement $\mathbf{z}$ and the best current estimate of this measurement $\hat{\mathbf{z}}$, so that:

$$\mathbf{r} = \mathbf{z} - \hat{\mathbf{z}}.$$

The residual vector corresponding to a failure, $\mathbf{f}$, can be written as [12] (see Appendix II for a reduced-order form that includes prior knowledge on the error states):

$$\mathbf{r} = (\mathbf{I} - \mathbf{H}\mathbf{S})\mathbf{f}.$$

The norm of $\mathbf{r}$ weighted by the measurement noise matrix, $\mathbf{V}$, is noted $\|\mathbf{r}\|_W$. $\|\mathbf{r}\|_W$ is chi-square distributed with $n_Z - (4n_K + n_S)$ degrees of freedom (n_Z is the dimension of $\mathbf{z}_B$) and non-centrality parameter, $\sqrt{\mathbf{f}^T\mathbf{V}^{-1}(\mathbf{I} - \mathbf{H}\mathbf{S})\mathbf{f}}$ [41]. A detection threshold, R_C, is set in compliance with a continuity requirement ($2 \cdot 10^{-6}$) to limit the probability of false alarms [42]. As a result, a measurement failure is undetected if:

$$\|\mathbf{r}\|_W < R_C.$$

The influence of an SSF on both of these dimensions can be represented on a plot of δx_U versus $\|\mathbf{r}\|_W$ (Figure 6). The upper left quadrant delimited by *VAL* and R_C corresponds to the missed-detection (MD) area, where failures are both hazardous and undetected. The probability of missed detection, P_{MD}, is defined as a joint probability:

$$P_{MD} = P\left(|\delta x_U| > VAL, \|\mathbf{r}\|_W < R_C\right).$$

Therefore, within the MD area, P_{MD} is the product of the cumulative probability distribution functions of $|\delta x_U|$ and $\|\mathbf{r}\|_W$. The normal and chi-square distributions of δx_U and $\|\mathbf{r}\|_W$, respectively, explain the ovoid shape of the isoprobability contours partially depicted in Figure 6 for an example failure mode $\mathbf{f}$.

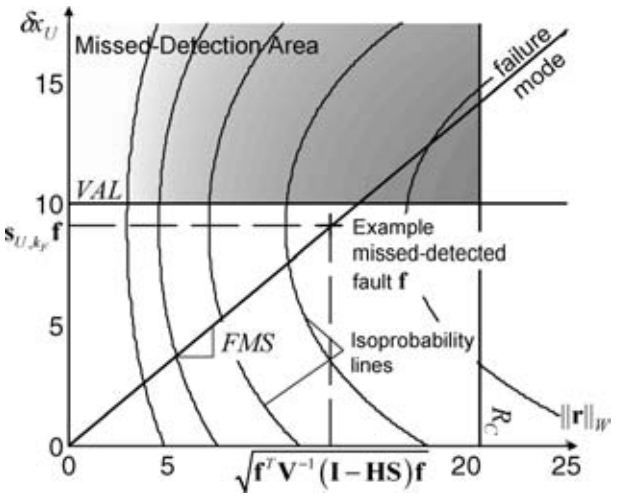

Fig. 6–Failure Mode Plot

The failure mode slope, *FMS*, defined as the ratio of $|\delta x_U|$ over $\|\mathbf{r}\|_W$, is a useful concept that is independent of the fault's magnitude; the resulting failure mode curve for all magnitudes is a line of slope *FMS* passing through the origin. On the failure mode plot of Figure 6, as the magnitude is increased from zero (fault-free case) to some larger value, the point moves along the line from the origin towards the right of the plot. The detection performance of the integrity threat search algorithm can therefore be evaluated by finding the steepest *FMS* for a set of failures, and then varying magnitudes along the corresponding line in search of the highest missed-detection probability, P_{MD} (corresponding to the worst-case fault).

A set of canonical failure modes is injected into the system for evaluation. At this stage of the research, simulated faults are limited to satellite failures because they are the only types of faults for which the failure rate is reliably known. Simulated faults of arbitrary magnitude, of numerous starting and stopping times, are easily constructed using blocks of truncated triangular matrices. For example, for a satellite s available over three epochs, the set of failures is:

$$^{s}\mathbf{f} = \begin{bmatrix} 0 & 0 & 3 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 0 & 2 \end{bmatrix}.$$

The first two columns represent step faults, the next two are descending ramps, and the last one is an ascending ramp. For efficiency, all redundant failure modes were eliminated (e.g., ascending and descending ramps over the entire duration, T_F, have identical effects). Similar matrices are generated for each satellite over the entire smoothing duration and applied to code and carrier-phase measurements individually as well as simultaneously, one SV at a time. The number of failures hence simulated exceeds 1,500 for a ten-minute smoothing interval. Future work includes testing for failures with multiple breakpoints and affecting multiple satellites at a time.

FRAMEWORK FOR THE PERFORMANCE ANALYSIS

Example Mission Definition and Requirements

A preliminary performance analysis is structured around a benchmark application of an aircraft precision approach. As illustrated in Figure 7, an airplane equipped with a GPS/Iridium receiver is following a simplified straight-in trajectory: it is flying at a constant speed of 70 m/s with a 3 deg glideslope angle towards the runway until touchdown (TD) where lateral and vertical requirements apply (symbolized by a rectangle). As mentioned earlier, the focus in this work is on the vertical position coordinate. A *VAL* requirement of 10 m is assumed for the purpose of this analysis to apply from a 200 ft altitude to TD. The filtering interval (which equals 5 min in the example of Figure 7) is simulated for position estimation at TD. Protection levels (PL) are defined below.

In this work, system performance is measured in terms of availability of a high-integrity vertical position solution. The integrity risk requirement, or probability of hazardous misleading information (HMI), noted P_{HMI}, is defined as the limit probability of any information sent by iGPS resulting in out-of-specification position error without timely warning [13]. It is subdivided here into two hypotheses, so that the integrity budget allocated to normal fault-free conditions (FF) is αP_{HMI}, and the fraction for rare-event single-satellite failures (SSF) is $(1 - \alpha)P_{HMI}$. The coefficient α ranges between 0 and 1 and is selected to maximize the

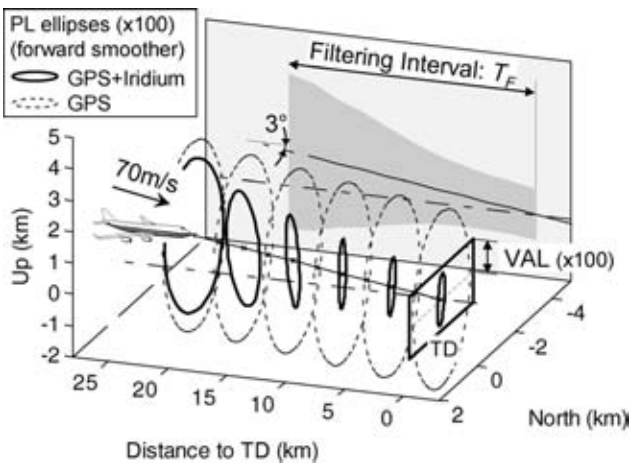

Fig. 7–Final Approach Simulation Description (Case s in Fig. 9)

Table 2—Summary of Simulation Parameters

Parameter Description	Nominal
P_{HMI} : integrity risk	10^{-9}
α : FF integrity risk coefficient	10^{-5}
Continuity risk	$2{\cdot}10^{-6}$
FR : failure rate	10^{-4}/hr
VAL : vertical alert limit	10 m
T_{AV} : availability simulation period	3 days
T_F : filtering period	10 min[a]
Location : near-worst case (25.5deg North, −81.1deg East)	Miami
GPS constellation	24 SVs
Iridium constellation	66 SVs
Signals : single-frequency (SF) or dual-frequency (DF)	SF

[a] 5min is used in the FF-availability analysis (first subsection of Sec. VI)

combined FF and SSF performance. Cases of multiple simultaneous measurement faults (considered independent events, hence having an even lower probability of occurrence) are neglected in this phase of the analysis, but will be addressed in the future.

Under normal conditions, the vertical protection level, *VPL*, a statistical over-bound on the positioning error in the Up-direction, is defined as a function of the standard deviation of the vertical position coordinate, σ_u:

$$VPL = \kappa_{FF} \cdot \sigma_U,$$

where the probability multiplier, κ_{FF}, corresponding to αP_{HMI} is a confidence-level coefficient (it is in fact the value for which the cumulative distribution function of a normal distribution equals $1 - \alpha P_{HMI}/2$). Nominal simulation parameters are summarized in Table 2. An approach or a geometry is deemed available under FF conditions if and only if

$$VPL < VAL. \tag{11}$$

Rare-event faults such as equipment and satellite failures (whose rate *FR* is assumed to be 10^{-4}/hr [43]) become significant threats when aiming at ensuring an integrity risk, $(1 - \alpha)P_{HMI}$, on the order of 10^{-9} [13]. The SSF-availability criterion, established here using a RAIM methodology, specifies that:

$$P_{MD} < (1 - \alpha)P_{HMI}/(FR \cdot T_F). \tag{12}$$

When testing the detection algorithm, if the probability of missed-detection, P_{MD}, for any one of the simulated faults violates Eq. (12), the approach is considered SSF-unavailable. The method presented in the previous section aims at directly finding the fault causing the highest P_{MD}.

Equations (11) and (12) are the expressions of FF and SSF binary criteria that either validate or nullify availability for an approach. In the following sections, approaches starting at regular intervals are simulated for sequences of satellite-user geometries, over a period T_{AV} defined below. Ultimately, the percentage of available approaches is the measure of iGPS FF and SSF performance. It is evaluated for a nominal iGPS configuration, conservatively selected to produce reliable availability results, and to exhibit performance variations that can be analyzed in the next section. This nominal navigation architecture is also used in the final section as a reference for comparisons when evaluating the sensitivity of individual system parameters.

Period for the Availability Analysis

When combining information from GPS and Iridium constellations, the duration over which availability simulations are carried out (T_{AV}) should enable sampling of a complete set of satellite geometries. For GPS, T_{AV} is one sidereal day (the Earth's rotation period in an ECI frame, which equals $2T_{GPS}$). It corresponds to the time-period the constellation needs to completely repeat itself with respect to the Earth.

The orbital period for Iridium, T_{IRI}, is 6028 s. The combined GPS/Iridium constellation repeatability period with respect to the Earth can be easily calculated if secular variations due to Earth's oblateness are neglected. Secular effects prove to be very small both for GPS and Iridium due respectively to the high-altitude and the high-inclination of their orbits. Hence, it takes 1,507 sidereal days (more than 4 years) for the geometry between the earth, GPS, and Iridium satellites to completely repeat itself. Simulating the algorithms over 1,507 days is computationally too intensive.

Fortunately, an approximated duration representative of a large number of geometries can be utilized. In fact, Iridium satellites circle the earth exactly 43 times in 3 solar days and 4 seconds. Also, the remainder of the closest integer number of intervals, T_{IRI}, within $n \cdot 2T_{GPS}$, where n is a series of consecutive integers, exhibits a 3 to 4 day cyclic trend. Finally, Figure 8 demonstrates that the computed cumulative FF-availability averages out after a few days, for the nominal system configuration. Indeed, the maximum deviation relative to the accumulated value after 3 days does not exceed 0.03% at the Miami location over one month of simulation. Concurrently, it is important that the interval between simulated approaches be selected short enough as illustrated with the 30 s

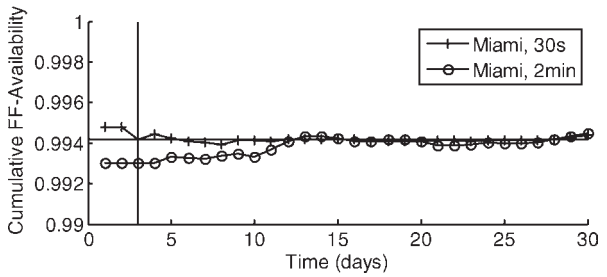

Fig. 8–Determination of the Period of the Availability Analysis (T_{AV})

and 2 min interval curves at the Miami location. In view of these results, approaches are simulated every 30 s over a period, T_{AV}, of 3 days.

FF AND SSF ALGORITHM ANALYSIS

FF Availability Analysis

Under FF conditions, the recursive KF or forward smoother that results in an optimal position estimate at TD is illustrated in Figure 7 for the aircraft precision approach mission. The covariance ellipses represent the protection levels (PLs) in the North and Up coordinates at different time steps. Position estimate standard deviations at each epoch are multiplied by κ_{FF} (to get PLs), and inflated by a constant scaling factor for the clarity of the plot. Their shape and size change with geometry (mainly because of Iridium satellite motion). A gray area for the corresponding vertical position covariance envelope along the aircraft trajectory is projected in the background. This example illustrates the incremental improvement in positioning accuracy within the filtering interval, T_F, and the dramatic enhancement of GPS/Iridium over non-augmented GPS, for which the position error barely changes. In this case, the FF integrity requirements are met, meaning that the ellipse at TD is fully contained inside the specified rectangle of alert limits.

The estimation algorithm performance is further analyzed to understand the impact of satellite geometries on carrier phase positioning. Cycle ambiguity estimation for mobile users requires SV redundancy as well as variations in user-to-satellite lines of sight, both of which are provided by augmenting GPS with Iridium [44]. In general, position and cycle ambiguity estimates improve as the change in LOS angle increases [5, 17]. The example geometry depicted in Figure 2(b) underscores the sharp contrast between an Iridium satellite whose angular variation over 10 minutes exceeds 130 deg, and GPS satellites whose LOS rotation barely reaches 5 deg.

A similar argument on observability helps explain the influence of range variations on individual position coordinates. Approaches are simulated at regular 30-second intervals over the period T_{AV} for a nominal configuration. Nominal parameters summarized in Tables 1 and 2 describe system conditions for a single-frequency aircraft user attempting a precision approach in Miami (near-worst location for the conterminous United States, or CONUS) and receiving WAAS-like ionospheric corrections. In this FF analysis, in order to investigate cases of FF unavailable geometries, the filtering period, T_F, is set to 5 min.

In Figure 9(a), a 15-hour period is extracted out of the total 72-hour simulation period, T_{AV}, in order to better visualize the variations in vertical protection level (VPL). Solid gray vertical lines indicate cases where the *VAL* is exceeded. The average number of Iridium satellites in view over T_F (i.e., the sum of visible Iridium SVs at each epoch divided by the number of epochs) for each simulated approach is presented beneath, over the same 15-hour period. In the following discussion, parallels are established between these two plots, which shed light on underlying estimation mechanisms. These observations are consistently repeated during T_{AV}, and have been established for a wide range of system parameters. The first finding is that GPS satellite measurements have a lesser influence on positioning. Neither the number of GPS satellites nor the vertical dilution of precision (not shown for conciseness) influence VPL results as significantly as Iridium spacecraft, which have a major impact.

In this regard, three regions are identified in Figure 1 to distinguish different areas of Iridium satellite visibility at Miami's latitude. Region A is located at the seam of the constellation, where the orbital plane separation-angle is smaller, so that the number of satellites in view at any instant is expected to be higher than elsewhere. Regions of Type B are located in-between co-rotating orbital planes and benefit from the coverage of satellites from both planes. Finally, regions of Type C designate areas around orbital planes, covered by only one or two satellites from a single plane at a time. Moving along a parallel at Miami's latitude, regions of Types B and C succeed each other, presenting respectively 1–3 and 1–2 visible satellites at a time.

Over time, the Earth rotates about its axis while SVs move in their orbital planes, which remain quasi-stationary in an ECI-frame. So, regions of Types B and C in Figure 1 correspond to intervals of time in Figure 9(a), which presents an alternating of Type B periods over which about 2.2 satellites are visible (on average over T_F) and Type C periods of about 1.3 satellites. The number of satellites increases every half sidereal day during Type A intervals. This increase is more pronounced on

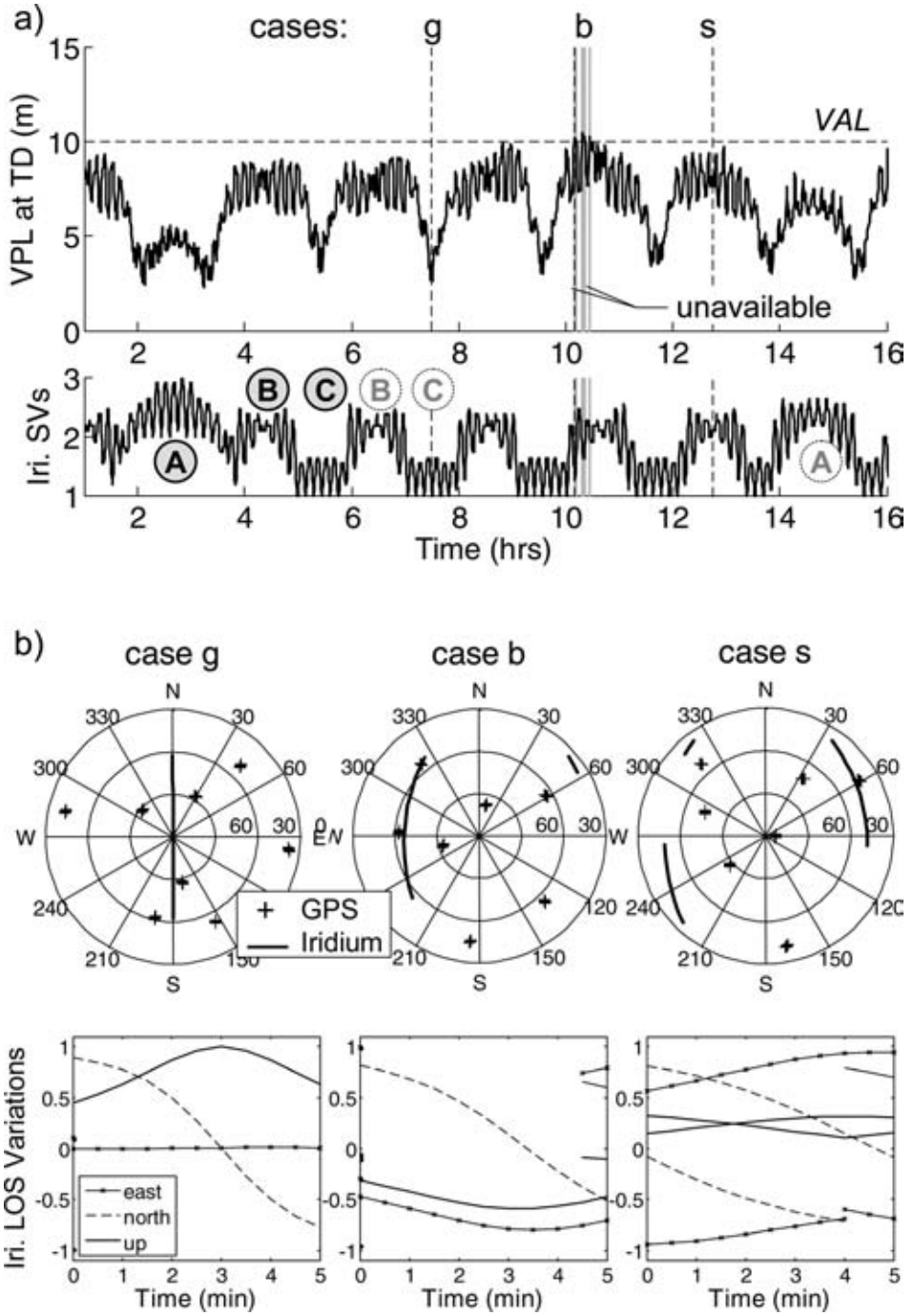

Fig. 9–FF Availability Analysis: a) VPL at TD and Average Number of Visible Iridium Satellites vs. Time; b) Azimuth-Elevation Sky Plots and Range Variations in the Position-Domain

one end of the seam than on the other (around the 15-hour point) because of the less-than-90 deg orbital plane inclination. Finally, users at Miami's latitude all experience patterns similar to Figure 9(a), which are shifted in time depending on their longitude.

The first obvious parallel between the two curves in Figure 9(a) is that during intervals of Type A, numerous Iridium measurements logically produce lower VPLs. A close look at the high-frequency variations of both curves shows that peaks in VPLs, especially within phases of Types A and B, correspond to valleys in average number of visible Iridium satellites over T_F (this again, regardless of GPS geometry). Another strong observation is that all unavailability cases occur around the beginning or the end of Type B intervals. This is verified over the 3-day T_{AV} period for all 50 unavailable approaches (out of 8620 total simulated cases). Now what seems paradoxical is that sharp drops

in VPLs are achieved during intervals of Type C, where the average number of Iridium satellites is the lowest.

To further examine this point, three characteristic cases are investigated: a 'good' case (noted g), a 'bad' case (b), and a 'standard' case (s). For each case, Figure 9(b) introduces azimuth-elevation sky plots of the SV trajectories over the filtering period, T_F, and Iridium satellite LOS variations along each of the three local position coordinates. Case (g), the best of the three cases, occurs in the middle of a Type C interval: one LEO satellite is traveling directly overhead the user, so that the variation in LOS coefficients corresponding to the vertical coordinate is the largest. Case (b), the worst of the three, is at the beginning of a Type B phase, and a single Iridium SV is visible over the largest part of the interval. Two mechanisms have been identified that explain the poor performance. First, the amount of angular variation with respect to the vertical direction is decreased relative to the previous case. Second, an observability issue appears that prevents estimation: unlike the two other cases, the LOS variation profiles in the East and Up coordinates have become difficult to distinguish so that the corresponding states remain undetermined, hence causing VPL to exceed the *VAL*. Finally in the standard case (s), in the middle of a Type B interval, additional satellites from the adjacent orbital plane come into sight, slightly augmenting the cumulated amount of angular variations and above all, resolving the East-Up ambiguity.

In summary, excellent performance is obtained when a satellite crosses the sky directly overhead the user, where angular variations with respect to the vertical axis are the most substantial. As time passes, the user's location drifts away from the orbital plane due to Earth rotation, causing vertical LOS variations to decrease with Iridium satellite elevation. On top of that, for a few spacecraft trajectories, occasional observability issues arise from the difficulty to distinguish variations in the East and Up directions. Additional low-elevation satellites come into sight as the user location approaches the adjacent plane, which solves the observability problem, and provides sufficient accumulated geometric diversity to meet the *VAL* requirement.

Practical lessons learned from this exercise are, for example, that if Iridium satellites were added for navigation purposes, the constellation should be rearranged to include extra orbital planes with tighter separation angles, rather than launching more SVs in the existing planes. Results for an 88-Iridium-satellite constellation (with eight orbital planes instead of six) are presented in the next section. Another costless and straightforward way to improve FF performance is to increase the filtering period, T_F, in order to include extra range variations. An upper limit of 10 minutes is imposed in order to ensure the validity of the error models. Using an interval, T_F, of 10 minutes, 100% availability is achieved over T_{AV} at the Miami location, under FF conditions.

Undetected Failures Analysis

The RAIM-type single satellite fault detection process is tested against measurement steps and single-breakpoint ramps of all magnitudes and of multiple starting times, for simulated approaches repeated every 30 seconds over the 3 day T_{AV}-period. A small value is assigned to coefficient α in order to maximize the fraction of the integrity risk allocated to SSF (see Table 2). The 10 minute smoothing period, T_F, enables 100% FF availability but does not prevent SSF unavailable approaches, using the nominal parameter values listed in Tables 1 and 2.

Fault-modes causing unavailability can be identified. First, results show that step-type faults are systematically detected. They generate huge residuals, even for low failure magnitudes, which translate into a gentle slope on a failure mode plot (i.e., Figure 6), and reaffirm the ability of the system to estimate constant biases. Then, after inspection, all 96 unavailable approaches (out of 8620 simulated cases) are due to ascending or descending ramp-type faults on Iridium carrier-phase measurements. Carrier-phase observations are the most sensitive because they carry by far the most weight in the estimation algorithm as opposed to code-phase measurements, which if corrupted but undetected, are typically not hazardous to the final position estimate. For the same reason, faults on Iridium signals have a more dramatic impact on the vertical position estimate error than faults on GPS measurements.

Further examination shows that the overwhelming majority of undetected faults affect satellites visible for short periods (for only part of the filtering interval), and the corresponding ramps stretch over the entire time the spacecraft is in view. In general, ramps whose starting or ending points occur sometime within the satellite pass are easily detected. Indeed, it was verified that elements of the residual corresponding to measurements collected just before and just after the breakpoint exhibit sharp variations, hence inflating $\|\mathbf{r}\|_W$. Rare exceptions, where faults that include a breakpoint go undetected, are addressed below.

Uninterrupted ramps over the entire pass constitute the main cause of unavailability. Analysis of the residuals suggests that these ramps might be indistinguishable from fault-free behaviors. Faults

causing missed-detection were therefore compared to the signature profiles of individual terms in the measurement equation, and turn out to match the satellite range variations. Another clue that confirms this idea is simply that the longest undetected failures do not stretch over more than 6 minutes, which is about half the maximum duration of an Iridium satellite pass (e.g., from rise to maximum elevation), and beyond which its range variation is no longer ramp-like (see the range variation curve in Figure 2). Even more revealing are the aforementioned exceptions, where the undetected fault includes a breakpoint. In all of these rare cases, the failure-mode is a ramp stretching over most of the pass, and a constant over one or two epochs. Now all undetected cases apply to satellites whose passes are approximately truncated in half by the filtering interval and whose peak satellite elevation (where range variations level off) occurs either at the beginning or at the end of the filtering period, so that the fault's breakpoint matches to the change of slope in range variation. Here again, the fault is masked by an FF behavior. Interestingly, the same ramp-type geometric sensitivity in range that is providing observability for cycle ambiguity bias estimation is also sometimes making ramp-type failures undetectable.

In some applications, operational system requirements include a limit on the period of navigation service outage, i.e., the maximum period of time over which high-integrity positioning is unavailable. As illustrated in the previous sub-section, cases of poor geometries are isolated, and do not last long. Accordingly, SSF availability results show that outages last on average 2.4 min, with a maximum of 4 min, meaning that users would never have to wait more than 4 min to recover the required performance.

Besides, in addition to verifying SSF availability (by finding the fault magnitude that generates the highest P_{MD}), the boundaries of the integrity threat space can be identified by determining the minimum and maximum fault magnitudes for which P_{MD} violates Eq. (12). Simulations indicate that the slopes of undetected hazardous ramp-like failures range from 7.7 to 33 mm/s over 2–6 min periods. Whether physical phenomena actually exist that cause such faults and how likely they are to actually occur is as yet unknown, but would be a useful subject for future work.

The analysis results described above provide a preliminary evaluation of the system's detection performance for a variety of faults that are encountered in actual operating environments. For example, steps model receiver cycle slips and ramps account for certain types of excessive satellite clock deviations and erroneous ephemeris parameters. However, it is understood that ultimate validation of fault detection performance must be verified against a comprehensive set of all real fault modes, as well as unusual ionospheric disturbances.

Now that the root causes for loss of availability are understood, the performance sensitivity to individual system parameters can be investigated.

COMBINED FF-SSF AVAILABILITY SENSITIVITY ANALYSIS

Parameter values for the nominal configuration (Tables 1 and 2) are used here as a reference for comparisons, and were conservatively selected to describe a system architecture implementable in the short term. Results are presented in terms of 'combined availability,' which is only granted for an approach if both the FF and SSF criteria are satisfied. As mentioned earlier, the SSF criterion is the driving factor for loss of availability. The nominal combined availability performance is 98.89% at the Miami location, which is a near-worst-case location for CONUS.

Parameters of the Measurement Error Source Models

The performance sensitivity to individual error model parameters is investigated for realistic ranges of values in Figure 10. More precisely, the combined FF-SSF availability at the Miami location, for a fixed smoothing interval, T_F, of 10 minutes, is plotted for each parameter's nominal standard deviation, σ_{NOM} (listed in Table 1; e.g., for the vertical ionospheric bias, b_{VI}, σ_{NOM} is σ_{VIB}), inflated by a scaling factor given on the x-axis (ranging between 0.2 and 1.8). As expected for all parameters, values lower than σ_{NOM} produce better results than the nominal case, and conversely availability decreases for higher values.

Three parameters stand out as being the most influential. First, the GPS ephemeris and clock

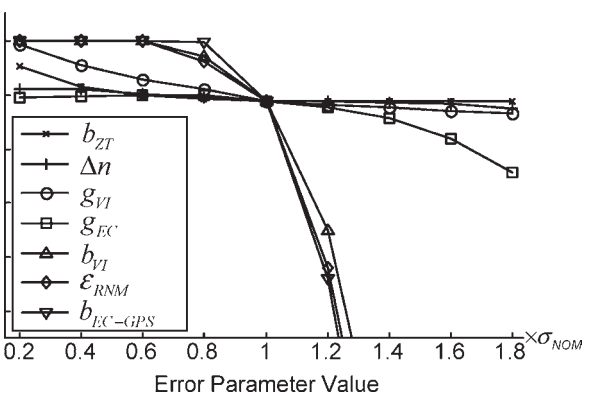

Fig. 10–Sensitivity to Parameters of the Measurement Error Source Models at Worst-Case Miami Location

bias parameter, $\sigma_{ECB-GPS}$, generates the largest performance variations. As noticed earlier, the impact of $\sigma_{ECB-GPS}$ reveals the importance of GPS code-phase measurements to the overall system performance. The nominal 1.86 m parameter value for $\sigma_{ECB-GPS}$, representative of GPS ephemeris and clock errors after OCS correction, was selected based on multiple years of data [22, 23]. Further data analysis is required before taking credit for WAAS-like ground corrections. Segments of the curve label, b_{EC-GPS}, to the left-hand side of the nominal point show the improvement such corrections could bring about. The corresponding parameter for Iridium, $\sigma_{ECB-IRI}$, only produces minute variations, and was not included for clarity of the plot.

Receiver noise and multipath are the second largest cause of performance variations. Values for σ_{RN} and σ_M depend on user receiver technology, and may vary with satellite elevation depending on the antenna. Thus the corresponding result in Figure 10 describes the sensitivity of iGPS performance to user equipment. The nominal values selected in Table 1 are typical of aircraft equipment.

Third, availability performance is very sensitive to the standard deviation of the vertical ionospheric bias, σ_{VIB}. Assuming WAAS-like ground corrections under non-anomalous ionosphere conditions at mid-latitudes, a 1-meter nominal value for σ_{VIB} was selected [9]. The need for ionospheric corrections determines in large part the scale of the ground infrastructure, which motivates further analysis below.

Finally, in view of the remaining results, biases (b_{VI} and b_{EC}) have a more significant impact than gradients (g_{VI} and g_{EC}) and than the troposphere parameters, b_{ZT} and Δn. Obviously, σ_{ZTB} is relatively small, and the accumulated error for the gradient-terms, $c_{OI}d_{IPP}g_{VI}$, $\Delta t \cdot g_{EC}$ and $c_{OT}c_T\Delta n$, over the short smoothing interval is not nearly as large as the bias-terms, b_{EC} and $c_{OI}b_{VI}$.

Locations and System Configurations

Combined FF and SSF availability (for the nominal configuration) is presented for a 5 deg $\times$ 5 deg and a 4 deg $\times$ 4 deg latitude-longitude grid of locations, respectively, over CONUS and over Europe in Figure 11. As expected, results improve at higher latitudes, as the density of Iridium satellites increases. The map shows that the availability performance can drop below 99% at 25 degrees of latitude but reaches 100% for all locations at latitudes higher than 40 degrees.

Since the performance is driven by Iridium SV motion, and since a large part of the variations in longitude averages out over the 3-day simulation period, availability is plotted versus latitude in Figure 12, for an example longitude of -80 deg. Six different system configurations are considered, including the nominal case. The bottom curve presents the worst performance, obtained without ionospheric correction, assuming a σ_{VIB} of 5 m. This is evidence that single-frequency iGPS without corrections from a sizeable network of ground stations (e.g., WAAS-like) is not sufficient to enable applications that require high accuracy and integrity, such as aircraft precision approach. Still, the

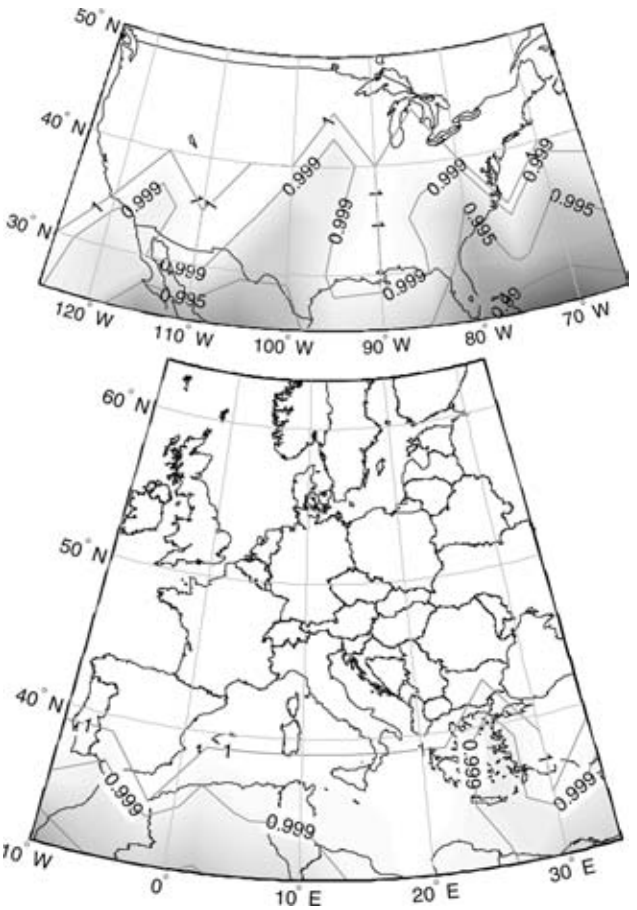

Fig. 11–Combined FF-SSF Availability Maps for the Nominal Configuration

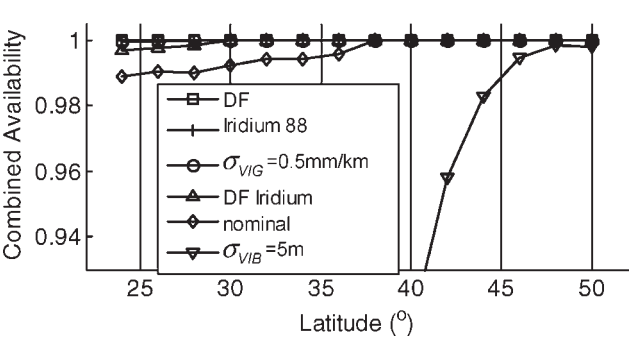

Fig. 12–Sensitivity to System Configurations

system produces 99% combined availability for latitudes higher than 45 deg.

The addition of ionospheric corrections for the vertical ionospheric bias, b_{VI}, is decisive as underscored with the nominal case. Furthermore, the existing WAAS ionospheric delay estimation algorithm presented in Reference [20] not only computes corrections for the bias, b_{VI}, but for the gradient, g_{VI}, as well. Significant enhancement would be obtained if these corrections were also broadcast, increasing the availability at 24 deg latitude, −80 deg longitude, from 98.9% for the nominal case to 99.9%. The curve labeled 'DF Iridium' shows that substantial improvement could also be observed if the second Iridium broadcast frequency (Ka-band) could be reliably tracked by the user.

More promising results are obtained for future long-term evolutions. First, extending the Iridium constellation by adding two additional orbital planes, as suggested earlier, produces maximum availability. Then, modernization of GPS (scheduled over the next two decades) will provide civilians with dual-frequency (DF) signals, which are free of ionospheric error. There is also potential to add DF capability to the next generation of Iridium satellites. Thus, ranging measurements would only require corrections for satellite orbit and clock related errors. This considerably reduces the need for densely spaced ground stations. Results generated for the dual-frequency GPS and Iridium implementation are excellent at all latitudes.

Potential Near-Future iGPS Performance

To refine the sensitivity analysis, the emphasis is placed on two of the most influential parameters: the smoothing interval, T_F, and the standard deviation of the residual vertical ionospheric bias, $\sigma_{VIB} \cdot T_F$ is limited to a maximum of 10 minutes to ensure robustness of the error models. σ_{VIB} depends on the accuracy of ionospheric corrections, which is determined by the ground segment. Thus, this preliminary analysis investigates what investment in ground infrastructure is needed to achieve high-integrity positioning.

The system's potential near-future performance is evaluated for a single-frequency GPS and Iridium architecture. As a reminder, the nominal configuration assumes that users are provided with GPS ephemeris and clock data from the OCS as well as precise Iridium satellite orbit information. Moreover, the uncorrected instantaneous ionospheric error is such that σ_{VIB} ranges between 5 m and 10 m. After correction from a WAAS-like network of 25 reference stations spread across the U.S., this number drops to 0.5–1.5 m. In fact, a one-sigma root-mean-square value of 0.51 m was computed using quarterly 95% ionospheric error indeces for all locations and all GPS satellites tabulated in the WAAS performance analysis reports [9] from spring 2002 to spring 2008.

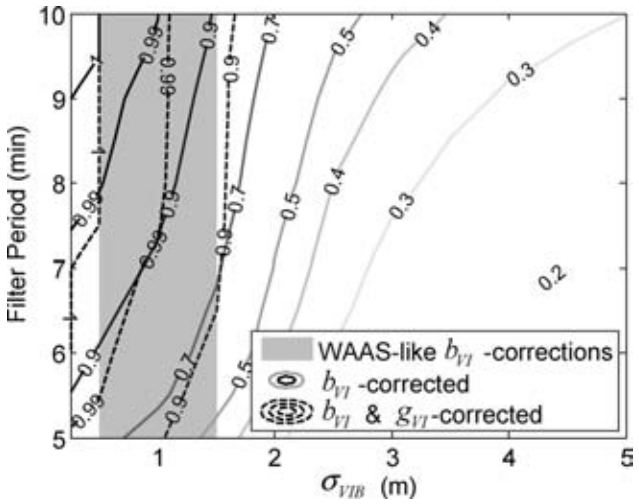

Fig. 13–Sensitivity to Filtering Period and Ionospheric Corrections

Combined FF-SSF availability results for the Miami location are given in Figure 13 versus σ_{VIB} and T_F. Contours of constant availability show that the performance sensitivity to filtering period decreases at low σ_{VIB} values, especially in the grey-shaded area corresponding to WAAS-like ionospheric corrections. The 100%-availability domain ranges from σ_{VIB} of 0.5 m and lower, and filtering periods, T_F, longer than 9 min. Further enlargement of this range is immediately achievable using corrections for the gradient, g_{VI} (dashed contours), transmitted here via Iridium communication channels.

Finally, the superposition of dashed and solid lines in Figure 13 highlights the multidimensionality of the problem. In the next step of this navigation system design process, similar analyses as well as further error model validation will be carried out for the satellite ephemeris and clock parameters, whose impact on the performance results has been illustrated earlier in Figure 10. Ultimately, quantifying the influence on the overall end-user performance of parameters such as ionospheric or satellite-related corrections helps establish and prioritize recommendations on individual system components.

CONCLUSION

This paper investigates the potential for Iridium-augmented GPS to enable rapid, robust, and accurate navigation at continental scales. Iridium's large satellite motion over short periods of time opens the

possibility to perform floating carrier-phase based positioning in real-time applications over extended geographic areas. The resulting bias observability is effectively exploited in the optimal position and cycle ambiguity estimation algorithm developed in this work. The detection of faults affecting sequences of time-correlated code and carrier phase signals in the presence of bounded measurement errors is achieved using a batch least-squares residual-based RAIM procedure also derived in this paper.

A methodology is established to evaluate the availability of high-integrity positioning solutions under fault-free and single-satellite fault conditions. Single-frequency iGPS without ionospheric ground corrections was shown to offer practical utility at latitudes of 40 deg and above. Below 40 deg of latitude, the influence of nominal ionospheric errors on more challenging geometries was shown to be largely mitigated using ground corrections broadcast by a network of ground reference stations.

Early performance analysis results over CONUS and Europe show promising results for navigation with high accuracy and integrity. However, a detailed integrity analysis, including the effects of actual satellite fault modes and atmospheric anomalies, is still needed to prove sufficiency for high-integrity applications. In particular, the system's robustness under stormy ionospheric conditions has yet to be evaluated. Future evolutions including dual-frequency architectures yield an even more decisive impact for Iridium-augmented GPS, as they may relax the requirements on ground infrastructure while extending the availability of high-integrity carrier-phase positioning.

ACKNOWLEDGMENTS

This work was sponsored in part by The Boeing Co., Phantom Works Division. In particular the authors gratefully acknowledge the efforts Dr. David Whelan and Mr. Bart Ferrell of Boeing for their early recognition of the significance of embarking on this research and their subsequent support.

References

1. Guier, W. and Weiffenbach, G., "Genesis of Satellite Navigation," *John Hopkins APL Technical Digest*, Vol. 19, No. 1, 1998, pp. 14–17.
2. Danchik, R., "An Overview of Transit Development," *John Hopkins APL Technical Digest*, Vol. 19, No. 1, 1998, pp. 18–26.
3. NSTB/WAAS T&E Team, "Global Positioning System (GPS) Standard Positioning Service (SPS) Performance Analysis Report," Report Nos. 26–62, William J. Hughes Technical Center, Atlantic City, NJ, 1999–2008.
4. Cohen, C., Cobb, S., Lawrence, D., Pervan, B., Powell, D., Parkinson, B., Aubrey, G., Loewe, W., Ormiston, D., MCNally, D., Kaufmann, D., Wullschleger, V., and Swider, R., "Autolanding a 737 Using GPS Integrity Beacons," *NAVIGATION*, Vol. 42, No. 3, 1995, pp. 467–486.
5. Pervan, B., "Navigation Integrity for Aircraft Precision Landing Using the Global Positioning System," Ph.D. Thesis, Stanford University, Stanford, CA, 1996.
6. Gutt, G., Fischer, S., Sheen, J., and Lawrence, D., "Precision Autoland Guidance of the X-31 Aircraft Using IBLS – The Integrity Beacon Landing System," *AIAA Guidance, Navigation, and Control Conference*, Providence, RI, 2004.
7. Rabinowitz, M., Parkinson, B., Cohen, C., O'Connor, M., Lawrence, D., "A System Using LEO Telecommunication Satellites for Rapid Acquisition of Integer Cycle Ambiguities," *Proceedings of ION/IEEE Position Location and Navigation Symposium*, Palm Springs, CA, 1998, pp. 137–145.
8. Van Dierendonck, A. J., "GNSS User Assessment of the Plans and the Benefits of GNSS Modernized Signals and Services," *Proceedings of The Institute of Navigation 61st Annual Meeting*, Cambridge, MA, 2005, pp. 201–209.
9. NSTB/WAAS T&E Team, "Wide-Area Augmentation System Performance Analysis Report," Report Nos. 5–23, William J. Hughes Technical Center, Atlantic City, NJ, 2003–2008.
10. Olynik, M., Petovello, M. G., Cannon, M. E., and Lachapelle, G., "Temporal Variability of GPS Error Sources and Their Effect on Relative Positioning Accuracy," *Proceedings of the 2002 National Technical Meeting of The Institute of Navigation*, San Diego, CA, January 2002, pp. 877–888.
11. Brown, R., "A Baseline RAIM Scheme and a Note on the Equivalence of Three RAIM Methods," *NAVIGATION*, Vol. 39, No. 4, 1992, pp. 127–137.
12. Brown, R., Parkinson, B., and spilker, J., "Receiver Autonomous Integrity Monitoring," *Global Positioning System: Theory and Applications, Volume 2,* AIAA Progress in Aeronautics and Astronautics, Washington, DC, Vol. 163, 1996, pp. 143–166.
13. RTCA Special Committee 159, "Minimum Aviation System Performance Standards for the Local Area Augmentation System (LAAS)," Document No. RTCA/DO-245, 2004.
14. RTCA Special Committee 159, "Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment," Document No. RTCA/DO-229C, 2001.
15. Misra, P. and Enge, P., *Global Positioning System: Signals, Measurments, and Performance*, Ganga-Jamuna Press, Lincoln, MA, Chap. 1–4, 2001, pp. 123–174.
16. Fossa, C., Raines, R., Gunsch, G., and Temple, M., "An Overview of the Iridium Low Earth Orbit (LEO) Satellite System," *Proceedings of the IEEE 1998 National Aerospace and Electronics Conference*, 1998, pp. 152–159.
17. Rabinowitz, M., "A Differential Carrier-Phase Navigation System Combining GPS with Low Earth Orbit Satellites for Rapid Resolution of Integer Cycle

Ambiguities," Ph.D. Thesis, Stanford University, Stanford, CA, 2000.
18. Kidder, S. and Vonder Haar, T., "A Satellite Constellation to Observe the Spectral Radiance Shell of Earth," *Conference on Satellite Meteorology and Oceanography*, 2004, pp. 1–5.
19. International Civil Aviation Organization (ICAO), "Manual for ICAO Aeronautical Mobile Satellite (Route) Service," Part 2-Iridium, draft edition, URL: http://www.icao.int/anb/panels/acp/repository.cfm [cited September 2007].
20. Walter, T., Hansen, A., Blanch, J., Enge, P., Mannucci, T., Pi, X., Sparks, LaL.rry, Iijima, B., El-Arini, B., Lejeune, R., Hagen, M., Altshuler, E., Fries, R., and Chu, A., "Robust Detection of Ionospheric Irregularities," *NAVIGATION*, Vol. 48, No. 2, Summer 2001, pp. 89–100.
21. van Graas, F., personal communication, 2007.
22. Parkinson, B. and Spilker, J., *Global Positioning System: Theory and Applications, Volume 1*, AIAA Progress in Aeronautics and Astronautics, Washington, DC, Volume 163, 1996.
23. Warren, D. and Raquet, J., "Broadcast vs. Precise GPS Ephemerides: a Historical Perspective," *GPS Solutions*, Vol. 7, 2003, pp. 151–156.
24. Gratton, L., "Orbit Ephemeris Monitors for Category I Local Area Augmentation of GPS," MS Thesis, Illinois Institute of Technology, Chicago, IL, 2003.
25. Bae, T. S., "Near-Real-Time Precise Orbit Determination of Low Earth Orbit Satellites Using an Optimal GPS Triple-Differencing Technique," Ph.D. Thesis, Ohio State University, Columbus, OH, 2006.
26. Bisnath, S. and Langley, R., "Precise Orbit Determination of Low Earth Orbiters with GPS Point Positioning," *Proceedings of the National Technical Meeting of The Institute of Navigation*, Long Beach, CA, 2001, pp. 725–733.
27. Klobuchar, J., "Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users," *IEEE Transactions on Aerospace Electronic Systems*, Vol. AES-23, No. 3, 1987, pp. 325–331.
28. Cohen, C. E., Pervan, B., and Parkinson, B. W., "Estimation of Absolute Ionospheric Delay Exclusively through Single-Frequency GPS Measurements," *Proceedings of the 5th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1992)*, Albuquerque, NM, September 1992, pp. 325–330.
29. Christie, J. R., Ko, P. Y., Hansen, A., Dai, D., Pullen, S., Pervan, B. S., and Parkinson, B. W., "The Effects of Local of Local Ionospheric Decorrelation on LAAS: Theory and Experimental Results," *Proceedings of the 1999 National Technical Meeting of The Institute of Navigation*, San Diego, CA, January 1999, pp. 769–777.
30. Hansen, A., Blanch, J., Walter, T., and Enge, P., "Ionospheric Correlation Analysis for WAAS: Quiet and Stormy," *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, Salt Lake City, UT, September 2000, pp. 634–642.
31. Hansen, A., Peterson, E., Walter, T., and Enge, P., "Correlation Structure of Ionospheric Estimation and Correction for WAAS," *Proceedings of the 2000 National Technical Meeting of The Institute of Navigation*, Anaheim, CA, January 2000, pp. 454–463.
32. Blanch, J., "Using Kriging to Bound Satellite Ranging Errors Due to the Ionosphere," Ph.D. Thesis, Stanford University, Stanford, CA, 2003.
33. Datta-Barua, S., Walter, T., Pullen, S., Luo, M., Blanch, J., and Enge, P., "Using WAAS Ionospheric Data to Estimate LAAS Short Baseline Gradients," *Proceedings of the 2002 National Technical Meeting of The Institute of Navigation*, San Diego, CA, January 2002, pp. 523–530.
34. Lee, J., Pullen, S., Datta-Barua, S., and Enge, P., "Assessment of Nominal Ionosphere Spatial Decorrelation for GPS-Based Aircraft Landing Systems," *Proceedings of the IEEE/ION Position Location and Navigation Symposium*, San Diego, CA, 2006.
35. McGraw, G., Murphy, T., Brenner, M., Pullen, S., and Van Dierendonck, A., "Development of the LAAS Accuracy Models," *Proceedings of The Institute of Navigation GPS Conference*, Salt Lake City, UT, 2000, pp. 1212–1223.
36. Murphy, T., Harris, M., Booth, J., Geren, P., Pankaskie, T., Clark, B., Burns, J., and Urda, T., "Results from the Program for the Investigation of Airborne Multipath Errors," *Proceedings of the 2005 National Technical Meeting of The Institute of Navigation*, San Diego, CA, January 2005, pp. 153–169.
37. Khanafseh, S., "GPS Navigation Algorithms for Autonomous Airborne Refueling of Unmanned Air Vehicles," Ph.D. Thesis, Illinois Institute of Technology, Chicago, IL, 2008.
38. DeCleene, B., "Defining Pseudorange Integrity - Overbounding," *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, Salt Lake City, UT, September 2000, pp. 1916–1924.
39. Gelb, A., (editor), "Optimal Linear Filtering", *Applied Optimal Estimation*, The MIT Press, Cambridge, MA, 1974, pp. 102–155.
40. Crassidis, J. and Junkins, J., *Optimal Estimation of Dynamic Systems*, Chapman & Hall/CRC, Boca Raton, FL, 2004, pp. 123–174.
41. Walter, T., and Enge, P., "Weighted RAIM for Precision Approach," *Proceedings of the 8th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1995)*, Palm Springs, CA, September 1995, pp. 1995–2004.
42. Sturza, M., "Navigation System Integrity Monitoring Using Redundant Measurements," *NAVIGATION*, Vol. 35, No. 4, 1988, pp. 69–87.
43. Federal Aviation Administration, "Category I Local Area Augmentation System Ground Facility," FAA-E-2937A, 2002.
44. Hwang, P., "Kinematic GPS for Differential Positioning: Resolving Integer Ambiguities on the Fly," *NAVIGATION*, Vol. 38, No. 1, 1991, pp. 1–15.
45. Kolb, P. F., Chen, X., and Vollath, U., "A New Method to Model the Ionosphere Across Local Area Networks," *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2005)*, Long Beach, CA, September 2005, pp. 705–711.

APPENDIX I IONOSPHERIC ERROR MODEL ASSUMPTIONS

Extensive analysis of ionospheric data is in progress and has been the subject of significant prior research. Assumptions made in this paper are based on existing literature and a limited set of new experimental data.

Literature Review

As mentioned in the fourth section, the ionospheric delay varies greatly with sun exposure. Under normal conditions, a σ_{VIB} of 10–20 m is considered [15] and empirical ionosphere modeling (e.g., using the Klobuchar model, whose parameter values are broadcast by GPS) helps decrease this number by approximately 50% [27]. In addition, an important amount of work motivated by LAAS and WAAS aims at determining the vertical ionospheric gradient, g_{VI}, under anomaly-free conditions. Researchers agree on a σ_{VIG} of 1 mm/km for quiet days [34, 45] and 4 mm/km for active (but nonstormy) ionospheric days. The results in [29–34] strongly suggest that the linearity and bounding Gaussian assumptions are valid for IPP separation distances of up to 2000 km.

Dual-Frequency GPS Data Processing

Dual-frequency GPS data (L_1 and L_2) were collected in Chicago on one winter and one summer day (12/01/06 and 07/10/07). Measurements were processed using a 'time-step method' [34], which exploits the dispersive nature of the ionosphere, and the fact that most unwanted quantities are constant biases over time (cycle ambiguities, interfrequency biases, and part of the ionospheric delay). The noise of the carrier-phase double-differenced measurement (used in this method) is small with respect to the effect of the vertical ionospheric gradient, g_{VI}. Experimental data confirm that the probability distribution of the gradient, g_{VI} (after de-trending using Klobuchar's model), can be bounded by a Gaussian with a 1 mm/km σ_{VIG} value during quiet days. A line-fitting algorithm is employed to verify the linearity assumption; the coefficient of recession remains close to 1 for large periods of time (up to an hour). Further data processing is ongoing, but the initial results support the assumptions made in this paper.

WAAS Data Analysis and Simulation

Dual-frequency measurements from redundant receivers are collected at 25 wide area reference stations (WRS) and processed at the Wide area Master Stations (WMS) to compute ionospheric corrections for a grid of locations over CONUS. According to [20] and [32], for each grid point (IGP) location, a planar fit is applied to all WRS ionospheric delay measurements contained within a certain radius. The resulting IGP vertical delay estimates (IVDE) are sent to the user.

In our work, IVDEs were recorded for the aforementioned summer and winter days, for a grid of 40 IGP locations over CONUS. The time interval between b_{VI} corrections (2–5 min) does not allow for efficient g_{VI} corrections (using values of b_{VI} over time). An estimate of large-scale variations of the gradient, g_{VI}, can be computed by differencing the IVDEs at different IGPs. A maximum of 1.5 mm/km was observed.

In addition to the bias, b_{VI}, corrections for the gradient, g_{VI}, are computed at the WMS (the slopes of the plane fit), but are not broadcast due to the low data transmission rate (250 bps) and because they are not needed in aviation applications currently serviced by WAAS. A covariance analysis replicated from [20] establishes that a σ_{VIG} after correction of 0.5 mm/km is achievable. Furthermore, the additional data bandwidth enabled by Iridium provides ample margin for conveying new aiding data to the user. Finally, a one meter value is assigned to the corrected σ_{VIB} based on numerous WAAS Performance Analysis Reports [9] documenting the daily and quarterly 95% ionospheric error for multiple locations since 2002.

APPENDIX II REDUCED-ORDER WEIGHTED RESIDUAL EQUATION WITH PRIOR KNOWLEDGE

Pseudo-measurements (Eq. (8)) can be incorporated to the batch observation (Eq. (7)) by augmentation:

$$\mathbf{z} = \begin{bmatrix} \mathbf{H}_B \\ \mathbf{H}_P \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{v}_B \\ \mathbf{v}_P \end{bmatrix}, \tag{A.1}$$

in order to get an equation of the same form as Eq. (9):

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{v}.$$

$\mathbf{H}_P$ is subdivided to distinguish states with no prior knowledge ($\mathbf{u}$ and $\mathbf{N}$) from the error states whose initial uncertainty can be bounded: $\mathbf{H}_P = [\mathbf{0}_{n_E \times (4n_K + n_S)} \ \mathbf{I}_{n_E}]$. The same subdivision is performed on $\mathbf{H}_B$, so that $\mathbf{H}_B = [\mathbf{H}_{uN} \ \mathbf{H}_E]$. The weighted least squares state covariance matrix is:

$$\mathbf{P}_{\mathbf{x}} = \left(\mathbf{H}^T \mathbf{V}^{-1} \mathbf{H}\right)^{-1}, \text{ where } \mathbf{V} = \begin{bmatrix} \mathbf{V}_B & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_P \end{bmatrix}$$

$$\mathbf{P_x} = \begin{bmatrix} \mathbf{H}_{uN}^T \mathbf{V}_B^{-1} \mathbf{H}_{uN} & \mathbf{H}_{uN}^T \mathbf{V}_B^{-1} \mathbf{H}_E \\ \mathbf{H}_E^T \mathbf{V}_B^{-1} \mathbf{H}_{uN} & \mathbf{H}_E^T \mathbf{V}_B^{-1} \mathbf{H}_E + \mathbf{V}_P^{-1} \end{bmatrix}^{-1}$$

which can be written as

$$\mathbf{P_x} = \left(\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_P^{-1} \end{bmatrix} + \mathbf{H}_B^T \mathbf{V}_B^{-1} \mathbf{H}_B \right)^{-1}. \tag{A.2}$$

The last expression is actually a KF measurement update covariance equation, in which the first term of the addition is the information matrix at mission initialization, the inverse of which is referred to as the *a priori* state estimate covariance matrix in [40].

The following notation is then defined:

$$\mathbf{P_x} = \begin{bmatrix} \mathbf{P}_{uN} & \mathbf{P}_{uNE} \\ \mathbf{P}_{uNE}^T & \mathbf{P}_E \end{bmatrix}.$$

The augmented residual ($\mathbf{r} = [\mathbf{r}_B^T \ \mathbf{r}_P^T]$) subject to a failure, $\mathbf{f}$, is:

$$\begin{bmatrix} \mathbf{r}_B \\ \mathbf{r}_P \end{bmatrix} = (\mathbf{I} - \mathbf{HS}) \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}.$$

With the above notations, $\mathbf{HS}$ is subdivided into four blocks. Based on Eq. (10) describing $\mathbf{S}$, computations of $\mathbf{r}_B$ and $\mathbf{r}_P$ result in:

$$\mathbf{r}_B = \mathbf{f} - \Big(\mathbf{H}_{uN} \mathbf{P}_{uN} \mathbf{H}_{uN}^T + \mathbf{H}_E \mathbf{P}_{uNE}^T \mathbf{H}_{uN}^T + \mathbf{H}_{uN} \mathbf{P}_{uNE} \mathbf{H}_E^T + \mathbf{H}_E \mathbf{P}_E \mathbf{H}_E^T \Big) \mathbf{V}_B^{-1} \mathbf{f}$$

which reduces to:

$$\mathbf{r}_B = \Big(\mathbf{I} - \mathbf{H}_B \mathbf{P_x} \mathbf{H}_B^T \mathbf{V}_B^{-1} \Big) \mathbf{f} \tag{A.3}$$

and

$$\mathbf{r}_P = -\mathbf{H}_P \mathbf{P_x} \mathbf{H}_B^T \mathbf{V}_B^{-1} \mathbf{f}. \tag{A.4}$$

Finally, the weighted norm of the residual $\|\mathbf{r}\|_W$ can be expressed as:

$$\|\mathbf{r}\|_W^2 = \mathbf{r}^T \mathbf{V}^{-1} \mathbf{r} = \mathbf{r}_B^T \mathbf{V}_B^{-1} \mathbf{r}_B + \mathbf{r}_P^T \mathbf{V}_P^{-1} \mathbf{r}_P. \tag{A.5}$$

If Eqs. (A.2) and (A.5) are implemented rather than the augmented system (A.1), the dimension of the residual vector, $\mathbf{r}_B$, decreases by an average of 50 elements (for the nominal configuration) with respect to $\mathbf{r}$, and brings about substantial computational gain when determining $\|\mathbf{r}\|_W$.