Carrier Phase Ionospheric Gradient Ground Monitor for GBAS with Experimental Validation

Samer Khanafseh, Fan Yang and Boris Pervan, Illinois Institute of Technology
Sam Pullen, Stanford University
John Warburton, William J. Hughes FAA Technical Center

ABSTRACT

This paper describes a Ground Based Augmentation System (GBAS) ground-based monitor capable of instantly detecting anomalous ionospheric gradients at the time of satellite acquisition. The monitor utilizes differential carrier phase measurements across multiple reference station baselines as the basis for detection. Performance analysis shows that the monitor is highly sensitive to the quality of the carrier phase measurements. Therefore, data collected from a GBAS prototype ground facility is used to quantify the measurement quality and validate the concept monitor.

INTRODUCTION

The Federal Aviation Administration (FAA) has sponsored two programs to augment GPS for civil aviation applications: the Wide Area Augmentation System (WAAS), which is referred to internationally as a Space Based Augmentation System (SBAS), and the Local Area Augmentation System (LAAS), commonly referred to as a Ground Based Augmentation System (GBAS). GBAS is a safety-critical system that is intended to support all close-proximity flight procedures including landing, departure, and surface operations at an airport. The main advantages of GBAS, relative to the existing Instrument Landing System (ILS), are its ability to support multiple types of approaches and its ability to support operations at several runways simultaneously with a single Ground Facility (GF). Because of its versatility, GBAS is envisioned to replace ILS, which is currently being used for final approach and landing phases of flight. Each GF will include multiple, spatially separated GPS receivers utilizing Low Multipath Antennas (LMAs). The primary reasons for the use of multiple reference receivers and antennas at the GF are to provide a means for detection and isolation of a failed receiver and also to allow for a net reduction in ranging error by averaging measurements for a given satellite. However, a generally unrecognized benefit of such antenna separation is that differential carrier phase measurements across the baselines can be used to detect and isolate certain signal-in-space (SIS) failures and anomalies that are hazardous to GBAS. Examples of such SIS failures are ephemeris anomalies and ionospheric fronts. In this work, we will focus on ionospheric front detection and mitigation.

The ionosphere is a region of ionized gases that extend from 50 km to about 1000 km above the Earth’s surface. When the GPS radio signal travels through this region a change occurs in its speed and direction (usually referred to as refraction), which has a direct impact on the calculation of user-to-satellite range. The state of the ionosphere changes over a given day and night and is also highly influenced by solar activity, such as solar storms and geomagnetic disturbances. Such fronts have been observed in 2000 and 2003 and are detailed in [1 – 4]. Because of the dispersive nature of the ionosphere, GPS code and carrier measurements are affected differently: the ionosphere delays the code phase signal and advances the carrier phase signal by the same amount. This phenomenon is commonly referred to as code-carrier divergence. More details about the hazardous impact of sharp ionospheric fronts on GBAS navigation is discussed in [4].

At the GF, existing monitors use code-carrier divergence to help detect ionospheric fronts. Another type of monitor which uses triple difference carrier phase measurements for fault detection is described in [5]. The major disadvantage of these monitors is that they require a time history of measurements and cannot detect ionospheric fronts present at satellite acquisition. In other words, if a newly acquired satellite is affected by an ionospheric front, such a monitor will not be able to detect it. Subsequent monitor performance may also be affected due to the presence of the front at the commencement of
monitoring. In this case, it is necessary for ground-based fault detection to rely on instantaneous GF measurements directly.

At first, this paper will tackle the issue of detecting ionospheric fronts at satellite acquisition. We will begin by considering a monitor capable of instantaneously detecting ionospheric fronts using double difference carrier phase measurements [6]. The performance of this monitor is quantified as a function of ionospheric front gradient, baseline length and carrier phase measurement quality. Due to the high sensitivity of monitor performance on carrier phase measurement quality, data validation is presented and discussed. Finally, improvements to the monitor performance via calibration of the LMA antenna phase center variation are discussed.

IONOSPHERIC FRONT MONITOR CONCEPT

The concept explained herein is based on the work in [6]. A very effective method toward detecting ionospheric fronts instantaneously is to use double difference carrier phase measurements between the known LMA baselines to directly observe the effect of the ionospheric anomalies. This method utilizes differences in baseline lengths to cover different ionospheric front gradient magnitudes. Under nominal conditions (fault free), a simplified version of the single difference (two reference station antennas constructing a baseline) carrier phase measurement can be written as,

\[ \Delta \phi = e^T x_b + \Delta \tau + \lambda \Delta n + \Delta \beta + \epsilon_{\phi p} \]  

(1)

where,

- \( \Delta \phi \): the single difference carrier phase measurement vector
- \( e \): user-satellite unit line of sight vector
- \( x_b \): baseline vector between the two antennas
- \( \Delta \tau \): differential receiver clock bias
- \( \lambda \): carrier phase wavelength
- \( \Delta n \): single difference ambiguity vector
- \( \Delta \beta \): differential ionospheric error between the antennas (baseline).
- \( \epsilon_{\phi p} \): the single difference carrier phase measurement noise vector

Under nominal ionospheric conditions, if the baseline length is relatively short (less than a kilometer), \( \Delta \beta \) will be on the order of millimeters and usually can be neglected. If an ionospheric front exists, the ionospheric front gradient will be significant and might cause an erroneous position estimate due to substantial \( \Delta \beta \). For simplicity, the ionospheric front is modeled as a slant gradient \( \alpha \) connecting two zones. Note that since \( \alpha \) is expressed as a slant gradient, there is no need to apply an obliquity factor to the measurements. Furthermore, it is assumed that the baseline between the antennas in the GBAS GF is aligned with the runway for the landing approach. Therefore, the ionospheric error is the product of the baseline length with the effective ionospheric slant gradient. In this case, Equation 1 becomes:

\[ \Delta \phi = e^T x_b + \Delta \tau + \lambda \Delta n + \alpha \left| x_b \right| + \epsilon_{\phi p} \]  

(2)

Pullen, et al. [4] found that in CONUS, \( \alpha \) could be as large as 425 mm/km. They also stated that much higher values are expected in more active regions near the equator. In order to be conservative, we will assume that values of up to 2000 mm/km can occur. For the case of single-frequency GBAS Approach Service Type D (GAST-D) CAT III, the current Standards and Recommended Practices (SARPs) requirement limits the responsibility of the ground system monitor to detection of gradients above 300 mm/km [7]. The reason for this is that an extensive analysis of the airborne monitoring system provided in GAST-D showed that only fronts which are larger than 300 mm/km and are undetected by the GF monitors are potentially hazardous [7].

It is also assumed that measurements to only one satellite will be affected by the front. That is based on the fact that, taking the maximum values for the front characteristics in [8] produces a zone that is only a small part of the sky as seen from the GF (at 350 km altitude), and a satellite far away from the satellite being monitored can always be found as to guarantee it is out of the affected zone. In addition, the double difference can be taken from a satellite that has already been validated as fault-free. As a result, differencing the carrier phase measurement of another fault-free satellite from Equation 2 (known as double difference measurement) eliminates the receiver clock bias and retains the gradient as shown in Equation 3.

\[ \Delta \phi = \Delta e^T x_b + \lambda \Delta n + \alpha \left| x_b \right| + \epsilon_{\phi p} \]  

(3)

Since the line-of-sight to each satellite is computed from the ephemeris and the baseline vector between the GF antennas is known, the first term on the right hand side of Equation 3 can be removed from the measurements:

\[ \Delta \phi - \Delta e^T x_b = \lambda \Delta n + \alpha \left| x_b \right| + \epsilon_{\phi p} \]  

(4)

Separating the ionospheric front error from the ambiguities in Equation 4 is quite challenging. However, it is known that the ambiguities can only take on integer values. Therefore, if the measurement noise is ignored, a necessary condition for the detection of an ionospheric front is that the value of \( \alpha \left| x_b \right| \) is not equal to an integer.
multiple of $\lambda$. This fact forms the basis of the instantaneous ionospheric front detection monitor.

Now let us consider the case with measurement noise. The carrier phase measurement noise is assumed to be bounded by a Gaussian distribution with zero mean and a standard deviation $\sigma_{\phi}$. Therefore, the noise term in Equation 4 will be normally distributed with zero mean and a standard deviation of $\sigma_{\phi}^2 = 2\sigma_{\phi}^2$. Bounded by a Gaussian distribution with zero mean and a standard deviation $\sigma_{\phi}$.

Based on the concept described earlier, the test statistic $s$ is defined as

$$s = \Delta^2 \phi - \Delta^2 \phi^T x - \lambda \text{round}\left(\frac{\Delta^2 \phi - \Delta^2 \phi^T x}{\lambda}\right). \quad (5)$$

Under fault free conditions, this test statistic has a zero mean and a standard deviation of $\sigma_{\phi}^2$. Given an acceptable false alarm probability under fault free conditions, $P_{\text{ffd}}$, a detection threshold can be defined as:

$$T = k_{\text{ffd}} \sigma_{\phi}, \quad (6)$$

where $k_{\text{ffd}}$ is the false alarm multiplier computed from the inverse of the standard normal cumulative distribution function ($\Phi$) as

$$k_{\text{ffd}} = -\Phi^{-1}\left(\frac{P_{\text{ffd}}}{2}\right) = -\sqrt{2} \text{erf}^{-1}\left(P_{\text{ffd}} - 1\right) \quad (7)$$

The probability of missed detection $P_{\text{md}}$ of this monitor must also comply with the integrity risk requirement. Therefore, given that there is a fault (ionospheric front), the probability that the test statistic will be less than the threshold must meet the required $P_{\text{md}}$. Using a similar approach to Equations 6 and 7, the MDE becomes

$$MDE = T + k_{\text{md}} \sigma_{\phi} = \left(k_{\text{ffd}} + k_{\text{md}}\right) \sigma_{\phi}, \quad (8)$$

where $k_{\text{md}}$ is the missed detection multiplier and is computed as

$$k_{\text{md}} = -\Phi^{-1}\left(P_{\text{md}}\right) = -\sqrt{2} \text{erf}^{-1}\left(2P_{\text{md}} - 1\right). \quad (9)$$

Due to the existence of a nearby integer with its own MDE buffers, there will also be a maximum detectable error. This case is better illustrated in Figure 2. The origin point (corresponding to point 0) represents the integer value of $\lambda \text{round}\left(\frac{\Delta^2 \phi - \Delta^2 \phi^T x}{\lambda}\right)$. The minimum detectable $\alpha|x_b|$ about integer 0 is given by $(k_{\text{ffd}} \sigma_{\phi} + k_{\text{md}} \sigma_{\phi})$. Gradients larger than this value are detectable while satisfying the required availability and integrity requirement allocations. However, there is also an upper bound to the detectable gradient because of the presence of the adjacent integer and the undetectable region around it. In other words, if $\alpha|x_b|$ is significant, it might falls close to another integer where it cannot be detected with the required probability. This illustration can be expressed in a mathematical form that defines the undetected ionospheric slant gradient band as

$$\lambda n - \left(k_{\text{ffd}} + k_{\text{md}}\right) \sigma_{\phi} < \alpha < \lambda n + \left(k_{\text{ffd}} + k_{\text{md}}\right) \sigma_{\phi} \quad (10)$$

Figure 1: Detection principles used to define the threshold and MDE

Figure 2: Illustration of the undetectable lanes around each integer
The effectiveness of this monitor is fundamentally based on the widths of the detection bands, which in turn are directly related to the baseline lengths, $k_{n_d}$ and $k_{n_d}$, and the quality (standard deviation) of the differential carrier phase measurement error. As an example, suppose that $P_{md}$ and $P_{fd}$ are $10^{-4}$. This value for $P_{md}$ was chosen as a conservative number given an availability requirement of 99.9% given that the GBAS integrity risk requirement is $10^{-9}$, setting $P_{md}$ equal to $10^{-4}$ is based on a prior probability of an ionospheric front occurring. Using these values in equations (7) and (9), we conclude that $k_{n_d}$ and $k_{n_d}$ are 3.7 and 3.9, respectively. Assuming that $\sigma_{\Delta \phi}$ is 3 mm, it is possible to plot regions of non-detectable slant gradient as a function of baseline length. These regions are shown as shaded lanes in Figure 3. Other values of $\sigma_{\Delta \phi}$ have also been used to generate similar plots in Figures 4 – 6.

If $\sigma_{\Delta \phi}$ increases to 6 mm (Figure 4), the range of detectable slant gradients using a 100 m baseline decreases to 470 – 1450 mm/km. By adding another 300 m baseline, the monitor will be able to detect gradients in the range of 170 – 1750 mm/km. Figure 5 shows the case when $\sigma_{\Delta \phi}$ is 10 mm. In this case, the undetectable lanes get even thicker leaving small detection ranges. Therefore, multiple carefully chosen baselines are necessary even to detect an ionospheric gradient ranging from 190 – 1100 mm/km. If the measurement quality becomes so poor that $\sigma_{\Delta \phi}$ is 12 mm or more, the monitor is no longer feasible for instantaneous ionospheric front detection. This situation is depicted in Figure 6, where no detection space is left when the measurement noise reaches 13 mm ($1\sigma$).

---

Figure 3: Detection performance for the instantaneous monitor with $\sigma_{\Delta \phi} = 3$ mm

Figure 4: Detection performance for the instantaneous monitor with $\sigma_{\Delta \phi} = 6$ mm

Figure 5: Detection performance for the instantaneous monitor with $\sigma_{\Delta \phi} = 10$ mm
Figure 6: Detection performance for the instantaneous monitor with $\sigma_{\Delta}\phi = 12$ mm and $\sigma_{\Delta}\phi = 13$ mm (in the sub figure)

Even though the results shown in Figures 3 – 6 were obtained for specific values of $P_{fkd}$ and $P_{mud}$, they can be generalized to other values of $P_{fkd}$ and $P_{mud}$ by appropriately modifying the value of $\sigma_{\Delta}\phi$ sigma for each plot. For example, Figure 3 (which was generated with $\sigma_{\Delta}\phi$ set equal to 3 mm) is in fact for a case where $(k_{fkd} + k_{mud}) \sigma_{\Delta}\phi = 22.8$ mm. Therefore, if different values for $k_{fkd}$ and $k_{mud}$ are required, the figure will correspond to a value of $\sigma_{\Delta}\phi$ that is equal to $\frac{22.8}{k_{fkd} + k_{mud}}$. Figures 4 – 6 can be reinterpreted similarly.

In summary, the monitor feasibility is sensitive to the quality of the carrier phase measurements. The uses envisioned here for the GBAS LMAs pose a new challenge because such precision for carrier phase measurements is not required for the nominal fault-free GBAS functionality and was not a specific requirement in the design of these antennas. The exceptional theoretical thermal noise and multipath performance of the LMA suggests that such carrier phase precision should be achievable in principle, but the behavior of the antennas in detail at the millimeter level must be experimentally analyzed. Of particular importance in this regard are the potential variations in phase patterns between antennas, which will be discussed next.

**EXPERIMENTAL VALIDATION**

To examine the performance of the LMA antennas, measurements were collected from four LMA antennas at the LAAS Test Prototype (LTP) facility at the William J. Hughes FAA Technical Center in Atlantic City, New Jersey. The LMA is an array antenna that covers elevation angles of 5-90 degrees; it is resistant to terrestrial interference sources and meets the signal-to-noise ratio and multipath rejection requirements at low elevations. Although the antenna phase center variations are bounded by a few centimeters, in the previous section we showed that the measurement accuracy required by the instantaneous ionospheric front monitor must be within millimeters. In this paper, results will be shown for an example antenna pair first two antennas.

Figure 7 shows an example of the double difference carrier phase residual for one satellite (PRN 31). This residual is computed as

$$r = (\Delta\phi - \Delta e^T \mathbf{x}_s - \lambda \text{round}\left(\frac{\Delta\phi - \Delta e^T \mathbf{x}_s}{\lambda}\right)$$

Figure 7: Double difference carrier phase residual for one satellite (PRN 31).

As shown in Figure 7, non-random residual errors are much larger and occur at lower frequencies than what can be caused by multipath and thermal noise variations. This behavior is not unique to this particular satellite. Figure 8 shows a composite of all visible satellites over a 24 hour period in which several satellites display similar wavy patterns in their residuals. Furthermore, the residual wavy pattern shows day-to-day repeatability indicating that the source of the effect is a systematic difference in the phase patterns between antennas. This pattern cannot be ignored because of its effect on increasing the standard deviation of the double difference residual.

Figure 8: Double difference carrier phase residual for all visible satellites.
Since this monitor will be used in high integrity applications, simply computing an estimate of the standard deviation ($\sigma_{\Delta\phi}$) of the residual errors is not enough to ensure integrity. Instead, the tails of the distribution must be appropriately overbounded by a Gaussian distribution with a standard deviation that is inflated relative to this estimated value. A common method of constructing an overbound is to use the folded cumulative distribution function (CDF). The folded CDF is a CDF where one minus the cumulative probability is plotted on a y-axis (in log scale) for cumulative probabilities greater than one half. In Figure 9, the CDF for all satellite residuals is shown. Alongside the residual folded CDF with a zero mean and various sigma values (dashed curves). The figure shows that the residual CDF cannot be overbounded by a standard deviation of up to 6 mm, which is not a promising result for the front monitor.

Removal of the undesired systematic effects is necessary to achieve the levels of residual carrier phase error needed for the monitor applications under consideration. In this regard, the differential phase pattern behavior was captured as a function of satellite azimuth and elevation by fitting a full day’s worth of LTP satellite data to a spherical harmonic model. In previous work [5], similar models have been used to calibrate the Integrated Multipath Limiting Antennas (IMLA).

**Spherical Harmonics Model**

The spherical harmonics function, which is a Legendre trigonometric polynomial series of order $n$ and degree $m$ (Equation 12), is used for this application.

\[
U(\theta, \phi) = \sum_{n=1}^{\infty} \sum_{k} J_n P_n(c \theta) + \sum_{n=1}^{\infty} \sum_{m=1}^{n} \left[ C_{nm} \cos m \phi + S_{nm} \sin m \phi \right] P_m(c \theta)
\]

where,

- $J_n$, $C_{nm}$ and $S_{nm}$: spherical harmonic coefficients,
- $\theta$: zenith angle
- $\phi$: azimuth angle
- $P_m$: Legendre polynomial of order $n$, degree $m$, and is defined as
  \[
  P_m(x) = 2^{-m/2} \sum_{k=0}^{\lfloor m/2 \rfloor} \frac{(-1)^k}{k!(n-k)!(n-m-2k)!} x^{n-2k}
  \]

The model used to calibrate the antennas uses the double difference carrier phase residual. Therefore, after writing the spherical harmonics model in Equation 12 in a linear vector form [5], the residual can be written as

\[
r = \sum_{j=1}^{2} \sum_{l=1}^{n} \Delta P_{\lambda l} \left( \theta^j, \theta^l, \phi^j, \phi^l \right) : \Delta K_{\lambda l} + v_{\lambda|\phi}
\]

where for the two satellites used in the double difference computation ($j$ and $l$),

\[
\begin{align*}
\sum_{j=1}^{2} \sum_{l=1}^{n} \Delta P_{\lambda l} \left( \theta^j, \theta^l, \phi^j, \phi^l \right) : \Delta K &= \sum_{j=1}^{2} \Delta J_{\lambda j} \left[ P_{\lambda j}(c \theta^j) - P_{\lambda j}(c \theta^l) \right] + \\
&\sum_{j=1}^{2} \Delta C_{\lambda j} \left[ \cos m \phi^j - \cos m \phi^l \right] \left[ P_{\lambda j}(c \theta^j) - P_{\lambda j}(c \theta^l) \right] + \\
&\sum_{j=1}^{2} \Delta S_{\lambda j} \left[ \sin m \phi^j - \sin m \phi^l \right] \left[ P_{\lambda j}(c \theta^j) - P_{\lambda j}(c \theta^l) \right]
\end{align*}
\]

$\Delta K_{\lambda l}$ in Equation 13 can be estimated using a Kalman filter for all satellites in view using a 24 hour data set. The results of this filtering process are shown in Figure 10. The processed satellite passes (azimuth and elevation angles) are shown in the sky plot in Figure 10-a. In the calibration process, an 8th order model (k=8) was used because it was concluded in [5] that an 8th order model is sufficient to reduce the wave amplitude. Using this estimated $\Delta K_{\lambda l}$, a calibration map for all azimuth and elevation angles is generated and the correction magnitude is represented in the contour plot in Figure 10-b. This calibration is valid for the specific baseline under consideration and a different calibration map must be created for other baselines. Since no data was available in the northern section of the sky map (as illustrated by the sky plot in Figure 10-a), calibration at this area is not representative and was not included in the contour plot of Figure 10-b. Notice that the sky maps are not symmetric, which again suggests that the cause is individual phase center variations of these antennas.

Figure 9: Folded CDF for the residuals in Figure 8 with three overbounding Gaussian curves.
The corrected residuals using the spherical harmonics model are shown in Figure 11. The experimental results show that using a one-time calibration, the systematic wave patterns have been removed. Also, Figure 12 shows the CDF of the corrected residuals with three overbounding Gaussian CDFs. The figure shows that if a standard deviation of 3 mm is used, the tails are not overbounded. The standard deviation must be inflated to 4.5 mm to overbound the residual CDF out to a probability of $10^{-4}$. It should be noted that the remaining points that are not overbounded by this distribution were caused by outliers that were observed in the data. In order to be conservative, a standard deviation of 6 mm is used as an overbound. As a result, this monitor is feasible for detecting ionospheric slant gradients as shown in Figure 4. Using that figure, different baseline combinations can be chosen to cover the desired range of slant gradients to be detected.

The calibration map and the estimated coefficients are antenna and site dependent. If this calibration is performed using different antennas or is performed at any other airport or facility, new coefficients must be estimated. Nevertheless, in principle, this type of calibration should be valid for all times given that the antennas and siting locations have not changed. In future work, the authors will validate this conjecture by using the calibration map that is estimated here to correct data that was collected from the same site using the same antennas several months later. The source of outliers that are observed in Figures 8 and 11 and high tail activity (Figures 9 and 12) in the LTP data will also be investigated. Data from other installations at busy airports will also be considered. Furthermore, the monitor concept will be generalized to detect ionospheric front anomalies at all times (instead of only at the time of acquisition) and to potentially extend its capabilities to detect other anomalies like ephemeris failures.
CONCLUSIONS

This paper describes in detail an instantaneous monitor for detecting ionospheric front anomalies using double differential carrier phase measurements. The monitor utilized GBAS ground antenna baselines for detection. It was shown that certain combinations of baseline lengths can extend the range of detectable ionospheric slant gradients. However, in order for such a monitor to be feasible, the carrier phase measurement noise must be less than 10 mm. In order to validate the feasibility of this monitor, carrier phase performance of typical GBAS LMA antennas has been analyzed. An assessment of differential phase pattern variations with respect to azimuth and elevation has been presented. This paper defines the mathematical model developed for differential phase variation, explains its use in precise antenna calibration, and demonstrates that the resulting LMA double difference carrier phase errors can be overbounded by a standard deviation of 6 mm.

REFERENCES


