Formulation of a time-varying maximum allowable error for ground-based augmentation systems

Jason H. Rife, Tufts University & R. Eric Phelts, Stanford University

>> Accepted Article <<

CITATION:
J. Rife and R.E Phelts (2008). Formulation of a time-varying maximum allowable error for ground-based GPS augmentation systems. *IEEE Transactions on Aerospace and Electronic Systems*, 44(2), April 2008, pp. 548-560. First published under the same title in the *Proc. ION National Technical Meeting*, 2006, Monterey, CA. DOI: 10.1109/TAES.2008.4560206.

CORRESPONDING AUTHOR:
Jason Rife
jason.rife@tufts.edu

FINANCIAL SUPPORT:
Federal Aviation Administration Satellite Navigation LAAS Program Office (AND-710)

Formulation of a Time-Varying Maximum Allowable Error for Ground-Based Augmentation Systems

Jason Rife and R. Eric Phelts, *Stanford University*

BIOGRAPHY

Jason Rife is a Research Associate in the Stanford GPS laboratory. After receiving his B.S. in Mechanical Engineering from Cornell University (1996), he worked at the commercial engine division of Pratt and Whitney. He resumed his studies at Stanford and earned M.S. (1999) and Ph.D. (2004) degrees in Mechanical Engineering. His research focuses on integrity assurance for the Local Area Augmentation System (LAAS) and the Joint Precision Approach and Landing System (JPALS).

R. Eric Phelts is a Research Associate in the Department of Aeronautics and Astronautics at Stanford University. He received his B.S. in Mechanical Engineering from Georgia Institute of Technology in 1995, and his M.S. and Ph.D. in Mechanical Engineering from Stanford in 1997 and 2001, respectively. His research involves multipath mitigation techniques and satellite signal anomalies.

ABSTRACT

Ground-Based Augmentation Systems (GBAS), such as the Federal Aviation Administration's Local Area Augmentation System (LAAS), ensure integrity by enabling users to compute a conservative navigation error bound, called the Protection Level (PL). Although the PL covers both fault-free operations and certain faulted operations (reference-receiver faults or ephemeris faults), the PL does not explicitly account for all known fault modes. These remaining fault modes, including signal deformation, code-carrier divergence, and excessive acceleration, are covered implicitly through a range-domain bound known as the Maximum-allowable ERror in Range (MERR).

This paper introduces a new formulation for MERR that accounts for several limitations of earlier methods. In contrast with earlier formulations, which defined MERR as a static value, the new method defines MERR as a function of time. This time-varying MERR bound can then be compared to the expected time-evolution of the navigation error associated with each fault mode. The GBAS can be proven to maintain integrity during a fault event so long as the error magnitude does not exceed the MERR.

The time-varying MERR formulation addresses three limitations associated with earlier static formulations. Specifically, the new method (1) resolves ambiguity in the timing of bound application, (2) removes arbitrary restrictions on the performance of GBAS fault monitors, and (3) establishes a clear relationship between the MERR and the GBAS Time-to-Alert (TTA) requirement. In particular, TTA strongly impacts MERR magnitude, as illustrated by an application of the time-varying MERR to the signal-deformation fault mode.

INTRODUCTION

A primary responsibility of GBAS is to broadcast parameters that allow a user to compute a conservative error bound, called the Protection Level (PL). System availability is evaluated in real-time by comparing the PL to the alert limit, an operational criterion which describes the maximum navigation error tolerable during a precision approach. In general, the broadcast alert limit has a constant value, whereas the PL varies in time depending on the quality of the satellite geometry viewed by the GBAS user. To account for these geometry effects, the user receiver computes PL specifically for the set of satellites employed in its navigation solution.

To ensure integrity, the user PL must exceed worst-case undetected errors under both nominal and faulted conditions. Accordingly, the GBAS Local Ground Facility (LGF) both excludes unhealthy satellites and provides a sigma parameter that describes the errors for each usable satellite. LGF exclusions occur only when ground monitors, dedicated to individual fault modes such as signal deformation, code-carrier divergence, and excessive clock acceleration, detect an anomaly. The broadcast sigma parameter must protect users when anomalies are too small to detect with monitoring. Thus, as shown in Figure 1, GBAS integrity is achieved as long

as either ground monitoring detects the anomaly or the PL provides a bound for the resulting navigation error.

In practice, separate PL values are defined for each fault mode based on probability allotments from the GBAS integrity risk tree [1]. In an ideal case, allotments would be assigned dynamically such that the PL would be identical for all fault scenarios. Practical development considerations, however, have motivated the use of a static integrity tree and distinct PLs for each fault mode. Furthermore, to reduce user-receiver complexity, the user is required only to evaluate PL expressions for three scenarios: the fault-free (H0) case, the single reference-receiver fault (H1) case, and the ephemeris-fault case. The LGF assumes responsibility for other fault scenarios.

In taking responsibility for these fault modes, the LGF cannot explicitly determine the user PL, as the LGF lacks information about the satellite set employed in the user navigation solution. Consequently, the LGF must assume a worst-case user geometry. This assumption permits the definition of a range-domain bound on the largest permissible fault-induced error. This bound is commonly referred to as the Maximum-allowable Error in Range, or MERR.

Several previous authors have provided a framework for determining and applying MERR. Existing MERR formulations, however, place constraints on ground monitor design and provide no connection to the GBAS Time-to-Alert specification. Resolving these issues would significantly benefit the development and validation of GBAS ground facilities.

This paper introduces a new time-varying MERR concept. The time-varying MERR is developed in three steps. The first section of this paper reviews existing MERR methods to highlight their strengths and liabilities. The second section derives the time-varying MERR to address these liabilities. Finally, the last section of the paper illustrates the application of the time-varying MERR concept through an example involving the signal-deformation fault scenario.

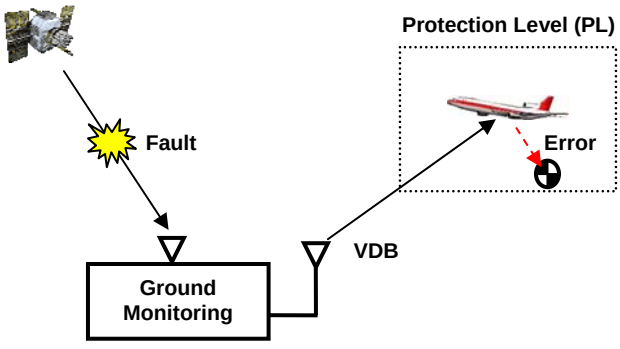

Figure 1. MERR Concept

During a fault scenario, GBAS integrity is ensured by (1) ground monitoring and (2) by the navigation-error confidence bound known as the protection level (PL).

PREVIOUS MERR FORMULATIONS

This section reviews past developments of the MERR to highlight the key concepts, milestones, and limitations of the existing MERR formulation.

The MERR first appeared in a paper by Shively under the name MERPROS (Maximum Error in Pseudorange for One-Statellite) [2]. Shively's paper laid the groundwork for converting the position-domain PL bound into a geometry-free bound on the worst-case range-domain error for a faulted satellite. This bound, the MERPROS, was derived based on two key assumptions regarding the form of the ranging error and the role of fault monitoring. The first assumption treated the ranging error for the faulted satellite (satellite k) as a random error with an abnormally wide error distribution. Using the Gaussian overbound concept [3], the widened error distribution, $p(\mathcal{E}_{fault})$, was modeled as a normal distribution, $\mathcal{N}$, with zero mean and with standard deviation scaled by a multiplicative factor, F, relative to the fault-free standard deviation, σ_{ff}.

$$p\left(\mathcal{E}_{fault}\right) = \mathcal{N}\left(0, F\sigma_{ff,k}\right) \tag{1}$$

The second assumption was to apply no direct credit for ground monitoring in the MERPROS derivation. Rather, monitor missed-detection probabilities were introduced subsequently in the integrity and continuity discussion of the MERPROS [2].

In parallel with the development of the MERPROS (soon abbreviated to the shortened form, MERR), other research explored the role of ground monitoring in greater detail. This research noted that ground monitoring provides full system integrity for fault anomalies larger than a cutoff level, called the Minimum Detectable Event (MDE) [4]. Central to the derivation of the MDE is the assumption that the monitor statistic is composed of a wholly deterministic bias, induced by the fault, and an additional random error, associated with fault-free monitor noise. Under this assumption, the Gaussian overbound for the monitor statistic, m_{fault}, is parameterized by a standard deviation for fault-free monitor operations, σ_{mon}, and a mean equal to the fault-induced bias term, η.

$$p\left(m_{fault}\right) = \mathcal{N}\left(\eta, \sigma_{mon,k}\right) \tag{2}$$

Because LGF monitoring assures GBAS integrity above the MDE, the MERR need only apply below the MDE. Based on this observation, it was proposed that the MERR would ensure GBAS integrity as long as it exceeded the ranging error at the MDE. This proposed integrity test, comparing the MDE error and the MERR, is illustrated in Figure 2.

The figure consists of two plots, which show the ranging error and missed-detection probabilities for a hypothetical fault event. The MDE time is defined in the lower plot, at the moment when the missed detection risk equals the allowed integrity risk. The ranging error at the MDE point is evaluated in the upper plot. As long as this ranging error is less than MERR (derived using the MERPROS methodology), then the proposed test suggests GBAS integrity is validated for the fault event.

In fact, as Zaugg later showed, this proposed test does not validate GBAS integrity in all cases [5]. Zaugg recognized that the point of greatest integrity risk actually occurs below the MDE point. To illustrate this fact, Zaugg plotted the combined risks associated with monitor missed detection (P_{md}) and with navigation error exceeding PL given a missed detection event (P_{pl}). The point of highest total integrity risk ($P_{md} \cdot P_{pl}$) occurs at a point inside the MDE, as illustrated in Figure 3. In the figure, the Maximum-risk Point is labeled MP.

Based on his observations, Zaugg modified the MERR formulation to incorporate the monitor missed-detection probability directly. This revised MERR expression is here referred to as $MERR_{MP}$, where the subscript denotes the Maximum-risk Point, MP. Values of $MERR_{MP}$ are lower than values of the original MERPROS, indicating that Zaugg's refined $MERR_{MP}$ provides enhanced conservatism.

In addition to resulting in a stricter bound, the $MERR_{MP}$ derivation differs from the MERPROS derivation in two respects. First, the $MERR_{MP}$ derivation treats the fault as a deterministic bias rather than a random error. Second, the $MERR_{MP}$ formulation assumes a particular form for the monitor time response under faulted conditions, in order to incorporate the missed-detection risk directly into the MERR derivation.

The first assumption replaces the distribution for the faulted-satellite ranging error, E_{fault}, previously described by (1), with a different Gaussian overbound. This alternate distribution treats the fault-induced error as a deterministic bias, E, summed with a random error. The random error, with standard deviation, σ_{ff}, describes the fault-free noise in the pseudorange measurement.

$$p\left(\mathcal{E}_{fault}\right) = \mathcal{N}\left(E, \sigma_{ff,k}\right) \tag{3}$$

This deterministic treatment of the fault-induced error better matches the monitor statistic description, (2), used in deriving the MDE.

A second key aspect of the $MERR_{MP}$ derivation was the assumption of proportionality between the fault-induced biases in the ranging error, E, and the monitor statistic, η. Assuming some relationship between these two quantities is necessary to embed the missed-detection risk in the $MERR_{MP}$ bound. As a consequence of this assumption, however, $MERR_{MP}$ is only valid for certain monitor designs with proportional time response for the monitor statistic and the error transient such that $E(t) \propto \eta(t)$. This assumed proportionality also obscures the relationship

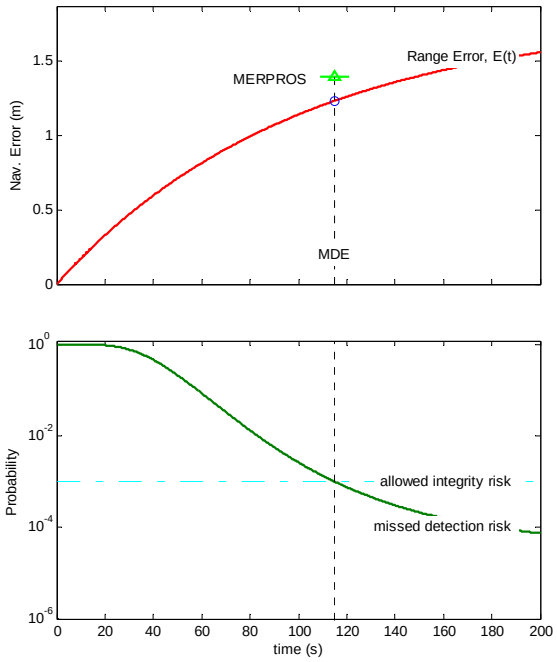

Figure 2. Single-Point MERR Concept Based on Minimum Detectable Error (MDE)

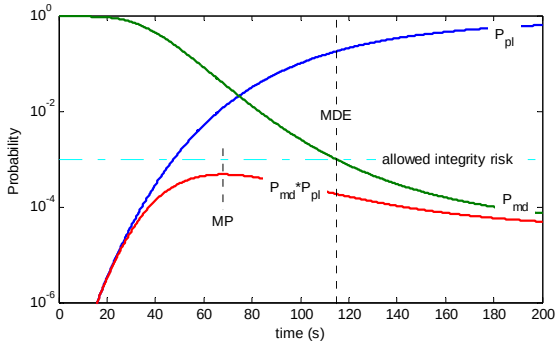

Figure 3. Maximum-risk Point

between MERR and the GBAS Time-to-Alert requirement. This ambiguity led Zaugg to conclude that Time-to-Alert was not a significant restriction in applying the MERR bound.

In practice, the $MERR_{MP}$ concept is applied to establish GBAS integrity by confirming that $MERR_{MP}$ exceeds the MDE error for each relevant system fault. The $MERR_{MP}$ approach has been generally accepted and widely applied to validating GBAS integrity for signal deformation, code-carrier divergence, and excess acceleration. Because the method treats fault-case errors as pure biases, however, it is less useful for analysis of fault modes that induce elevated levels of random error, such as the Low Power or Radio-Frequency Interference fault modes.

Despite the wide acceptance of Zaugg's MERR formulation, the methodology leaves open several unresolved issues. These open issues include (1) the absence of a precise comparison criterion for $MERR_{MP}$, (2) the restriction to monitor designs with filter transients proportional to the ranging-error transient, and (3) the lack of a quantitative relationship between $MERR_{MP}$ and Time-to-Alert.

The first open issue results from the fact that Zaugg did not derive a new integrity test along with his derivation of $MERR_{MP}$. Designers have continued to compare the MDE error to $MERR_{MP}$ as a test of integrity, but it is not clear that this is the correct comparison. Figure 4 illustrates the consequences of this assumption. The error at the MDE point, $E(t_{MDE})$, may exceed $MERR_{MP}$ even though the error at the MP point does not. In this case, the $MERR_{MP}$ fails to bound the MDE error, so the conventional MERR test fails to confirm GBAS integrity. Arguably, however, the $MERR_{MP}$ should be compared not to the error at the MDE point but to the error at the MP point, which is well bounded. Resolving this ambiguity requires a more precise definition of the MERR comparison test.

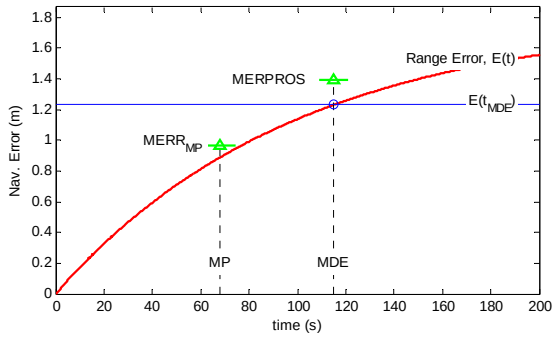

Figure 4. Comparison of $MERR_{MP}$ and Fault-Induced Ranging Error, $E(t)$

The second unresolved issue of $MERR_{MP}$ is the assumed proportionality between the offsets in monitor statistic, η, and navigation error, E. For proportionality to hold, both the monitor and user-ranging filters need to respond in a similar manner to a fault input. In practice, neither are the forms of the monitor and ranging filters similar, nor are the inputs to the individual filters necessarily proportional. Figure 5 offers two examples. The left column of Figure 5 illustrates signal-deformation transients when the monitor and range-smoothing filters are both Linear Time-Invariant (LTI), with time constants of 50 seconds for the monitor filter and 100 seconds for the ranging filter (also known as the carrier-smoothing filter). Both filters experience step inputs during the signal-deformation fault, but because of the difference in time constants, the monitor time response (lower left) is not proportional to the response of the ranging-error filter (upper left). The right column of Figure 5 illustrates a Code-Carrier Divergence (CCD) example that strays still farther from proportionality. The ranging filter is again taken to be a first-order LTI filter with a 100-s time constant. In the illustrated example, this filter experiences a restart ten seconds after the onset of CCD. By comparison, the CCD monitor filter will likely perform as a second-order cascade of two filters with individual time constants of 40 seconds, as proposed by Pervan [6]. This monitor filter response increases as the range error response decays, giving a clear indication that the proportionality assumption is invalid. A more extensive MERR theory is required to validate integrity for these types of monitor responses, which lack proportionality.

A third unresolved $MERR_{MP}$ issue is the ambiguous relationship between $MERR_{MP}$ and Time-to-Alert (TTA). Since the $MERR_{MP}$ test is not defined for a specific instant in time, it is not possible to determine at what moment a MERR violation occurs or how much time passes before the LGF monitors detect and exclude an offending satellite.

These timing issues will be increasingly critical for Category III and for alternative GBAS implementations. Currently, for Category I, the time required to transmit a monitor warning message to an airborne user is less than six seconds. As the required TTA is six seconds, the Category I system is expected to meet the TTA requirement with margin. By contrast, for Category III, the TTA is reduced to only two seconds. The LGF monitor must detect threats very quickly to ensure integrity in this case. In fact, if the transmission time remained at six seconds, the Category III monitor would need to detect threats four seconds before they became hazardous in order to meet the two second TTA requirement. The Federal Aviation Administration's proposed Local Airport Monitor (LAM) system would also experience transmission times longer than the TTA requirement [7]. LAM would provide an alternative means of achieving Category I GBAS capabilities by

rebroadcasting WAAS monitor messages. The associated transmission time of ten seconds exceeds the Category I TTA requirement by four seconds. As a consequence, LAM integrity requires that WAAS monitors detect threats four seconds before those threats become hazardous. A reformulated MERR should take this time differential into account for the purposes of integrity validation.

This paper introduces a time-varying MERR concept that addresses all three unresolved issues associated with existing MERR methodologies. In order to establish integrity using this new MERR formulation, the error transient is compared to MERR as a function of time. As long as MERR for a particular fault mode everywhere exceeds the error transient, the integrity of the GBAS is ensured. The fact that MERR is compared to the error transient at every moment in time resolves the first open issue, regarding the ambiguous timing of the $MERR_{MP}$ test. Also, because the new MERR formulation is evaluated as a function of time, no restrictions on the monitor statistic and error transients are required. This flexibility resolves the second open issue by removing the proportionality constraint and permitting validation of arbitrary monitor designs. Finally, with the time-varying MERR, TTA requirements enter directly into the MERR derivation. This quantitative connection between MERR and TTA resolves the third open issue associated with existing methodologies.

DERIVATION OF TIME-VARYING MERR

This section develops a mathematical formulation for the time-varying MERR. If the fault-induced ranging error exceeds the MERR bound at any time, GBAS integrity cannot be ensured.

For a system anomaly to result in a hazardous error, two simultaneous failures must occur. This chain of events was illustrated in Figure 1. First, ground monitoring must fail to detect the anomaly, and second, the PL must fail to bound the resulting navigation error. The MERR will be defined to reflect the conditional risks associated with this chain of events.

The Vertical Protection Level

A key step in quantifying the MERR bound is defining the protection level. The protection level used in deriving the MERR is in fact a Vertical Protection Level (VPL). The Lateral Protection Level (LPL) is neglected, both because the geometric diversity of the satellite

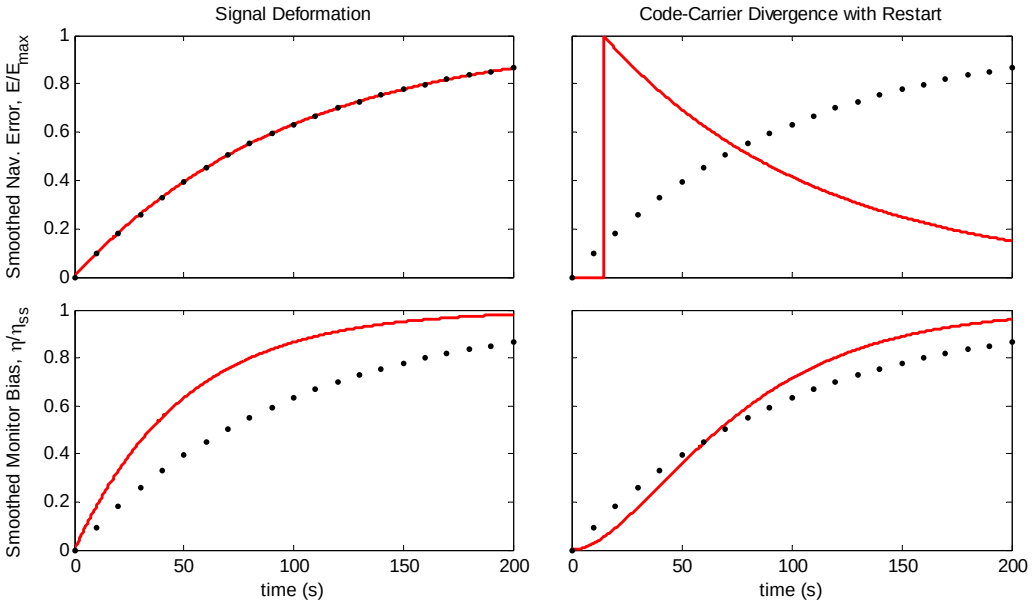

Figure 5. Comparison of Ranging Error Transient (*E*) and Monitor Statistic Transient (η) for Two Scenarios.

The left column represents transients associated with signal deformation. The right column represents transients associated with a ranging-filter restart which occurs during a code-carrier divergence event. Curves are normalized to a maximum value of one. A dotted reference curve is drawn on each plot. The dotted reference represents the normalized step response of a first-order linear time-invariant (LTI) filter with a 100-s time constant. In neither example is the ranging error transient, E, proportional to the fault-induced monitor bias, η.

constellation is poorest in the vertical direction, causing larger vertical errors than horizontal, and because the consequences of navigation errors are most severe in the vertical direction.

Although the airborne user evaluates multiple VPL expressions, only the fault-free expression (VPL_{H0}) is used in deriving MERR. The fault-free VPL is generally larger than the alternative VPL expressions for the faulted reference-receiver case (VPL_{H1}) and for the ephemeris fault case (VPL_e). Moreover, since the user evaluates VPL to be the largest of these three expressions, the user VPL will always be at least as large as VPL_{H0}.

The VPL_{H0} expression is a confidence level for the fault-free error distribution. The confidence level is based on an allotted integrity risk, P_{ffmd}. Assuming a Gaussian overbound for the position error, the allotted risk probability is encoded in a sigma-scaling factor for the fault-free missed detection event: K_{ffmd}. The scaling factor and risk probability are related by a Q-function representing the cumulative distribution for a Gaussian distribution with zero mean and unit variance.

$$K_{ffmd} = -Q^{-1}\left(\tfrac{1}{2} P_{ffmd}\right). \quad (4)$$

The VPL bound is simply the K-factor multiplied by the standard deviation of the vertical positioning error, $\sigma_{v,p}$.

$$\mathrm{VPL}_{H0} = K_{ffmd}\,\sigma_{v,p} \quad (5)$$

Assuming Gaussian overbounds, the vertical component of the positioning error is derived from the standard deviations ($\sigma_{ff,i}$) for each of the N satellite ranging errors, evaluated under fault-free conditions. These errors are projected into the vertical direction by the sensitivity weights, $S_{v,i}$.

$$\sigma_{v,p} = \sqrt{\sum_{i=1}^{N} S_{v,i}^2 \sigma_{ff,i}^2} \quad (6)$$

The Navigation Error

As illustrated in Figure 6, the protection level must bound navigation errors that ground monitors fail to detect. In the event of a monitor missed detection, an integrity violation occurs at the instant when the vertical component of the navigation error exceeds the VPL. Figure 6 provides a hypothetical example in which the total vertical component of the navigation error (red) exceeds the VPL (green dashed line).

As shown in the lower plot of Figure 6, the model for the total error signal can be decomposed into two component signals: a random signal with a probability distribution matching the fault-free condition and a deterministic, fault-induced error. The random component of the error signal is strongly correlated in time. This time correlation is modeled using the concept of an exposure window.

The Exposure Window

GBAS integrity is defined over an exposure window that reflects the duration of approach and landing operations. For Category I, the exposure window is 150 s, corresponding to the duration of an approach. For Category III operations, the exposure window is 15 s (in the vertical direction) or 30 s (in the horizontal direction), corresponding to the duration of landing and rollout. An integrity violation occurs if the navigation error exceeds the protection level at any instant during the exposure window.

The duration of the exposure window determines the prior probability that a hazardous anomaly occurs during the approach or landing operation. System anomalies occur rarely; however, an operation with a longer exposure window introduces a somewhat higher prior probability for a fault. Since the start time of the exposure window is arbitrary, the probability of occurrence of any instant in the fault transient is identical to the prior probability for that fault.

In addition to determining the prior probability of a hazardous anomaly, the exposure window also provides a means to assess the time-correlation of the random component of the navigation error. Correlation reflects slow variation of certain error sources inherent in the raw

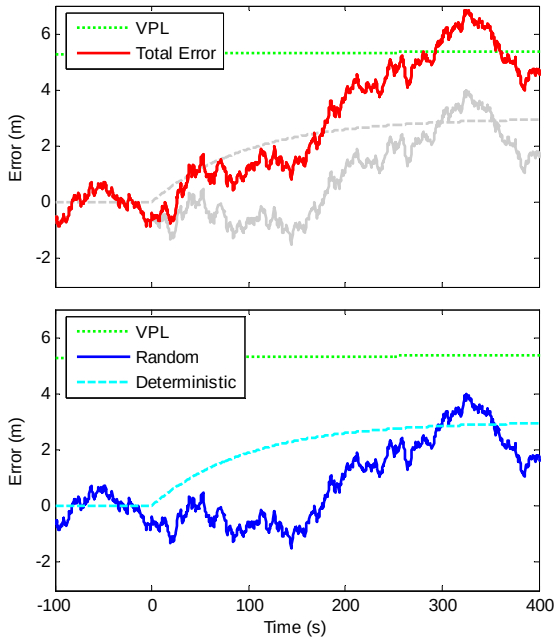

Figure 6. Instant of VPL Violation

measurement (multipath, ephemeris, ionosphere and troposphere, for instance) and, more significantly, the carrier-smoothing applied in the user receiver. Nominal carrier smoothing acts as a first-order filter with a 100-s time constant. Correlation time can be defined as twice the filter time constant (i.e. 200 s). Because the correlation time exceeds the exposure window duration (150 s, maximum), the random signal remains roughly constant throughout this window.

Because the signal is highly correlated over the entire exposure window, risk need only be assessed at the single, riskiest instant within the window. To prove integrity for any arbitrary exposure window, it is thus sufficient to show that the error transient never, at any instant, results in a risk probability greater than the allotted integrity risk.

Deriving MERR for a Fault Scenario

This section derives the time-varying MERR based on the instantaneous probability that the vertical navigation error exceeds the VPL. The vertical navigation error, $\mathcal{E}_{v,p}$, equals the weighted sum of the error associated with the faulted satellite, $\mathcal{E}_{fault}$, and the errors associated with the other unfaulted satellites, $\mathcal{E}_i$.

$$\mathcal{E}_{v,p} = \sum_{i=1}^{N} S_{v,i}\mathcal{E}_i = S_{v,k}\mathcal{E}_{fault}(t) + \sum_{i\neq k} S_{v,i}\mathcal{E}_i \qquad (7)$$

The faulted-satellite error, $\mathcal{E}_{fault}$,, is the sum of the fault-induced bias, E, and the random signal, E'. For the faulted satellite, k, the random signal is a zero-mean Gaussian with standard deviation equal to the fault-free case ($\sigma_{ff,k}$).

$$\mathcal{E}_{fault}(t) = E(t) + E'(t), \quad p(E') = \mathcal{N}(0, \sigma_{ff,k}) \qquad (8)$$

Since the random component of the faulted ranging signal is identical to that for the fault-free case, the probability distribution for the vertical positioning error is

$$p(\mathcal{E}_{v,p}) = \mathcal{N}(S_{v,k}E(t), \sigma_{v,p}). \qquad (9)$$

Based on this overbounding distribution, the probability of a vertical positioning error exceeding the protection level, P_{pl}, is given by the sum of two Gaussian Q-functions:

$$P_{pl} = Q\left(\frac{-VPL - S_{v,k}E(t)}{\sigma_{v,p}}\right) + Q\left(\frac{-VPL + S_{v,k}E(t)}{\sigma_{v,p}}\right) \qquad (10)$$

These Q-functions represent the cumulative probability in the tails of the Gaussian error distribution outside the VPL, as illustrated by the gray region of Figure 7. In practice, the total probability in one of the distribution tails is a negligible fraction of the probability in the other. After neglecting the smaller tail probability, the dominant Q-function can be inverted to give the following equation for K_{pl}.

$$K_{pl}(t) = -Q^{-1}(P_{pl}(t)) = \frac{VPL - |S_{v,k}E(t)|}{\sigma_{v,p}} \qquad (11)$$

The allowed protection level risk, P_{pl}, is a function of time that depends on the monitor missed-detection probability, as discussed in the following section. Given P_{pl}, and hence K_{pl}, can be evaluated, it is possible to rearrange (11) to solve for the fault-induced ranging error:

$$|E(t)| = \frac{VPL - K_{pl}(t)\sigma_{v,p}}{|S_{v,k}|} = \frac{[K_{ffmd} - K_{pl}(t)]\sigma_{v,p}}{|S_{v,k}|}. \qquad (12)$$

This equation for the ranging error depends on satellite geometry through the vertical positioning sigma, $\sigma_{v,p}$, and through the faulted-satellite weighting factor, $S_{v,k}$. Invoking equation (6) for $\sigma_{v,p}$, the ratio of these geometry-dependent terms may be lower bounded as follows.

$$\frac{\sigma_{v,p}}{|S_{v,k}|} = \frac{\sqrt{\sum_{i=1}^{N} S_{v,i}^2\sigma_{ff,i}^2}}{|S_{v,k}\sigma_{ff,k}|}\sigma_{ff,k} \geq \sigma_{ff,k}. \qquad (13)$$

This inequality expresses the fact that the two-norm of a vector is larger than any element of that vector. The worst constellation geometry (with the lowest maximum error) is the case dominated by the faulted satellite. Assuming this worst-case constellation, geometry is removed from (12), resulting in the following equation for Maximum-allowable Error in Range (MERR):

$$MERR(t) = [K_{ffmd} - K_{pl}(t)]\sigma_{min}. \qquad (14)$$

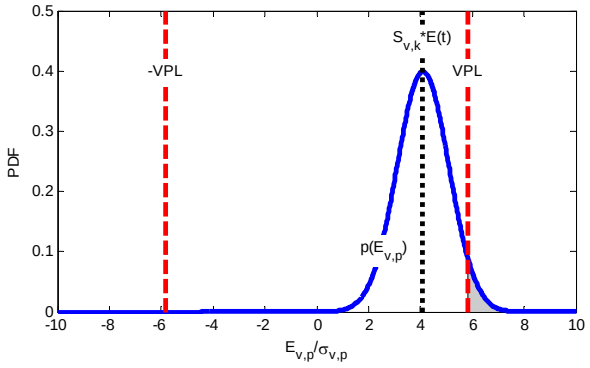

Figure 7. Probability of Error Exceeding Protection Level (Shaded Gray)

In the above equation, σ_{min} is defined to be the minimum fault-free sigma (σ_{ff}) over all elevations. For GBAS, this minimum fault-free sigma occurs at a 90° elevation. Since the MERR bound is derived to apply at this worst-case elevation for the worst-possible geometry, it applies conservatively to all other satellite geometries as well.

Instantaneous Integrity Risk

The time-varying MERR derived in the previous section varies as a function of the protection level risk, $P_{pl}(t)$. This section relates the protection level risk to the total risk allotment, P_a, determined from the GBAS fault tree.

The risk allotment takes into account three simultaneous events: first, a fault event must occur; second, ground monitoring must fail to detect the fault; and third, the VPL must fail to bound the fault-induced error given that ground monitoring has failed. These three events are each described by a conditional probability: P_f, the prior probability of the fault, P_{md}, the probability of a missed detection given the existence of the fault, and P_{pl}, the probability that the error exceeds protection level given a missed-detection. For integrity, the product of these conditional risks may equal but not exceed the risk allotment.

$$P_{pl}(t)P_{md}(t)P_f = P_a \,. \tag{15}$$

The maximum allowable protection level risk is determined by division:

$$P_{pl}(t) = \frac{P_a}{P_{md}(t)P_f} \,. \tag{16}$$

Time-variations in $P_{pl}(t)$, and hence in MERR(t), result from changes in the missed-detection probability, $P_{md}(t)$. Specifically, P_{pl} and MERR increase as P_{md} decreases.

The integrity allotment, P_a, and the fault prior, P_f, in (16) are static quantities. For Category I GBAS, P_a for each fault mode ranges between 10^{-11} and 10^{-9} depending on the severity of the fault and P_f is set to a standard value of $4.2 \cdot 10^{-6}$ per satellite per approach for all fault modes [8].

Monitor Missed Detection Probability

This section derives the monitor missed-detection risk, P_{md}, upon which the value of P_{pl}, and hence of the time-varying MERR, depends. The risk of a monitor missed-detection event decreases as the monitor statistic increases. In order to express this relationship, the monitor statistic signal is modeled, like the ranging error, as the combination of a random signal associated with fault-free noise, η', and a deterministic signal induced by the fault, η. The combined monitor statistic for a faulted satellite channel is m_{fault}:

$$m_{fault}(t) = \eta(t) + \eta(t)', \quad p(\eta') = \mathcal{N}\left(0, \sigma_{mon,k}\right) \tag{17}$$

A Gaussian overbound, with standard deviation σ_{mon}, is assumed to describe the value of the random signal component at any instant in time.

Fault monitoring at the Local Ground Facility (LGF) attempts to detect faults by comparing the monitor statistic to a threshold. The threshold is set sufficiently high to minimize false alarms caused by fault-free noise. When the monitor statistic exceeds the threshold, the LGF sends a warning message that excludes the affected satellite from the user navigation solution.

The LGF warning must reach the user in a timely fashion. The required time between the onset of a hazardous condition and the arrival of the warning message at the user is called Time-to-Alert (TTA). The actual worst-case Time-to-Transmit (TTT) the warning message may be longer or shorter than the required TTA. If transmission time is longer than the specified TTA, the monitor must make up the difference by triggering early. If the transmission time is shorter than the allowed alert time, then the monitor may trigger late, after the fault becomes hazardous. In either case, the time difference is referred to as the Relative Detection Time (RDT).

$$RDT = TTA - TTT \tag{18}$$

The RDT is illustrated in Figure 8.

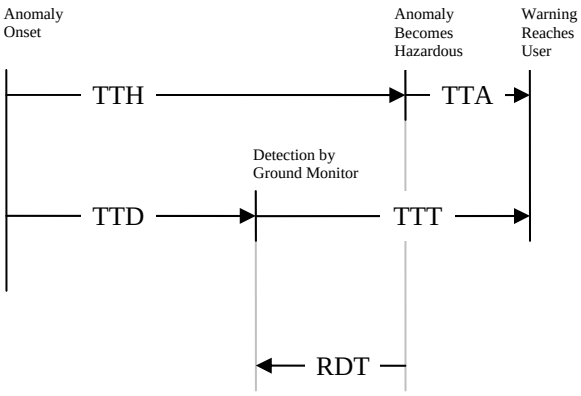

Figure 8. Relative Detection Time

In practice, the Time-to-Alert (TTA) requirement may not match the actual maximum Time-to-Transmit (TTT) warnings from the ground facility to the airborne user. The time difference, called the Relative Detection Time (RDT), alters the allowed Time-to-Detect (TTD) such that it is either shorter or longer than the duration between the anomaly onset and the moment at which the anomaly becomes hazardous, a duration labeled the Time-to-Hazard (TTH).

The RDT is defined to be negative if the monitor must trigger in advance of a protection level violation. For Category I operations, the RDT is zero seconds, because the maximum transmission time is six seconds, equal to the specified alert time requirement. For Category III operations, a six-second transmission time would be longer than the allowed 2 second TTA. In this case, the RDT would be -4 seconds, implying that a monitor would need to detect the anomaly 4 seconds before the anomaly becomes hazardous.

The RDT enters directly into the formula for the missed-detection probability, $P_{md}(t)$. The P_{md} is the probability that the monitor statistic lies inside the monitor thresholds, as illustrated by the gray region in Figure 9. Integrating the Gaussian distribution for the monitor statistic, m_{fault}, given by (17), gives the following formula for P_{md} in terms of two Gaussian Q-functions. The time at which P_{md} is evaluated must be shifted relative to PL evaluation by the RDT.

$$P_{md}(t - RDT) = Q\left(\frac{T - \eta(t)}{\sigma_{mon}}\right) - Q\left(\frac{-T - \eta(t)}{\sigma_{mon}}\right). \qquad (19)$$

This equation assumes the random components of the monitor statistic, m', and the vertical positioning error, $S_{v,k}E'$, are independent. If the random variables were strongly correlated, their dependence would need to be taken into account in the P_{md} equation, (19), in order to satisfy the conditional probability relationship of (16). The assumption of independence is justified for the following reasons.

- The errors that generate the monitor statistic noise, m', only represent a fraction of the total ranging error introduced by the ground station receivers (i.e. $\sigma_{ground,mon,i} < \sigma_{ground,i}$).
- The ranging error introduced by the ground signal is a small fraction of the total ranging error, which is dominated by airborne receiver noise ($\sigma_{air,i}$), ionosphere decorrelation noise ($\sigma_{iono,i}$) and troposphere decorrelation noise ($\sigma_{tr,i}$), such that $\sigma_{ground,i} < \sigma_{ff,i}$ where

$$\sigma_{ff,i} = \sqrt{\sigma^2_{ground,i} + \sigma^2_{air,i} + \sigma^2_{iono,i} + \sigma^2_{tr,i}} \qquad (20)$$

- Although MERR is defined for a worst-case geometry, the faulted-satellite ranging error is generally only a fraction of the total vertical navigation error, $\mathcal{E}_{v,p}$ (i.e. $\sigma_{ff,i} < \sigma_{v,p}$).

Because the errors that result in monitor noise have little impact on vertical positioning error ($\sigma_{ground,mon,i} << \sigma_{v,p}$), it is not expected that the random component of the monitor statistic, m', would display a measurable correlation with the random vertical navigation error, E'. Moreover, since any minor correlation would improve (reduce) the missed-detection probability for large errors, it is conservative to assume independence.

Time-Varying MERR Summary

The time-varying MERR has been derived as a function of the deterministic component of the monitor statistic transient, $\eta(t)$. So long as MERR(t) exceeds the fault-induced ranging bias, $E(t)$, for all time, the GBAS is ensured to maintain full integrity throughout the duration of a fault scenario.

Verifying integrity for any fault mode thus requires a fault-specific model of the time transients for $\eta(t)$ and $E(t)$. Given these curves, MERR is evaluated by computing $P_{md}(t)$ from $\eta(t)$ according to (19). Next, $P_{pl}(t)$ is derived from $P_{md}(t)$ according to (16) and converted into $K_{pl}(t)$ through (11). Finally, MERR(t) is evaluated by inserting the expression for $K_{pl}(t)$ into (14).

Integrity is verified by ensuring that the expression $E(t)$ is less than MERR(t) for all time. Figure 10 illustrates this comparison by extending the example shown earlier in Figure 2 and Figure 4. For the fault scenario illustrated, GBAS integrity is ensured since the MERR bound (green) is everywhere larger than the ranging bias (red).

A significant feature of the MERR(t) curve, as illustrated in the figure, is that the curve asymptotes and grows toward infinity at a particular time. The asymptote occurs where the missed detection probability, P_{md}, falls below the total allowed risk given the fault prior (P_a / P_f). At this point, previously identified as the Minimum Detectable Event (MDE) point, the monitor takes full responsibility for system integrity, so that the PL need not bound the faulted navigation error. In effect, MERR becomes infinite beyond the MDE point.

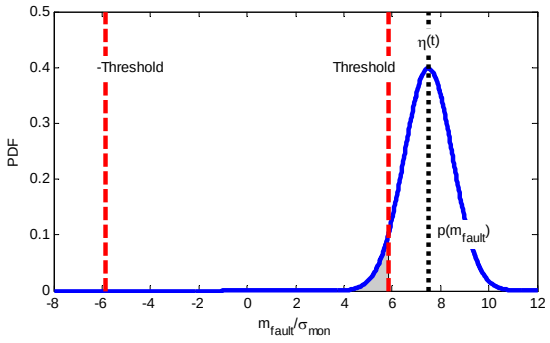

Figure 9. Probability of Missed-Detection by LGF Monitor (Shaded Gray)

Another feature of the time-varying MERR curve illustrated in Figure 10 is the relationship between MERR and $MERR_{MP}$. By design, these values are equal at the MP point when RDT is zero and when the $\eta(t)$ and $E(t)$ curves are proportional. Because the time-varying MERR curve is compared directly to the $E(t)$ curve, the time-varying MERR is slightly less conservative than the $MERR_{MP}$, which has generally been compared to the error at the MDE point. In practice, this benefit of the time-varying MERR is small unless RDT is non-zero. The signal-deformation example in the following section underscores the significance of RDT.

APPLICATION OF TIME-VARYING MERR TO SIGNAL-DEFORMATION MONITORING

In contrast with earlier MERR formulations, the time-varying MERR requires a model for the transients of both the monitor statistic and the ranging error. These transients must be evaluated for all threats in the relevant threat space. To provide an example, this section applies the time-varying MERR to the particular case of the signal-deformation fault.

Signal-Deformation Background

Signal deformation occurs when a GPS satellite emits a pathological signal, sometimes called an evil waveform. Evil waveforms distort the correlation peaks tracked by GPS receivers. These distortions may result in ranging biases as large as tens of meters for worst-case threats. Though these errors are large, they would ideally cancel in a differential correction scheme as long as the evil waveforms impacted all receivers in a like manner. Unfortunately, the diversity of user receiver designs results in substantial differences in signal-deformation errors for different receivers. Consequently, differential corrections do not necessarily eliminate large signal-deformation errors.

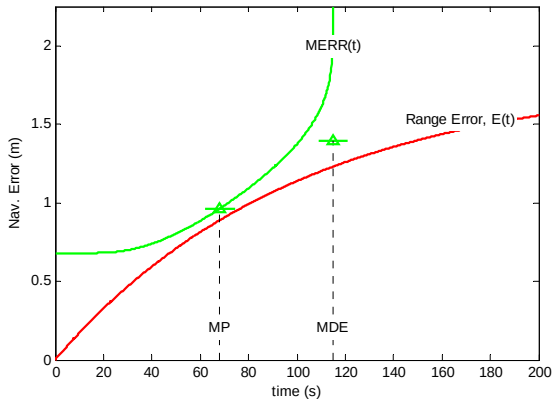

Figure 10. Comparison of MERR(*t*) to Fault-Induced Ranging Error, *E*(*t*)

To date only one signal-deformation event has ever been detected. This fault was observed for satellite SV19 in October of 1993. It is not known when the fault first occurred, and it is possible that the fault was present from the time the satellite was launched. Following the signal-deformation fault, the SV19 signal was recovered by switching over from the A-side electronics to the B-side. In the absence of data describing the signal-deformation transient, the accepted threat model treats the fault aggressively, as a step function.

Signal-Deformation Transients

The signal-deformation event model introduces a step error into the raw ranging measurement. This step is smoothed in the user receiver before being introduced into the navigation solution. Consequently, the model for the fault-induced ranging bias, $E(t)$, is the step response of the smoothing filter. The response for the baseline carrier-smoothing filter is that of a linear time-invariant (LTI), first-order filter with a time constant of 100 seconds [1]. Assuming this baseline filter for the current example, the fault-induced bias, $E(t)$, has the following form. It is convenient to divide the bias by its steady-state value, E_{ss}, so that the normalized bias, $f_E(t)$, is only a function of time and the carrier-smoothing time constant, τ_{CS}.

$$\frac{E(t)}{E_{ss}} = f_E(t) = \left(1 - \exp(-t/\tau_{CS})\right) \tag{21}$$

The signal-deformation model also assumes that the raw monitor statistic, which quantifies distortion in the code correlation peak, experiences a step input. The monitor filter smooths the raw data before applying a detection threshold. Thus the monitor statistic transient, like the range error transient, is the step response of a low-pass filter. To provide rapid detection, the time constant of the monitor filter may be as little as 50 seconds. For a first-order LTI monitor with a time constant, τ_{mon}, of 50 seconds, the step response, $\eta(t)$, can be written in a normalized form, $f_\eta(t)$, as follows.

$$\frac{\eta(t)}{\eta_{ss}} = f_\eta(t) = \left(1 - \exp(-t/\tau_{mon})\right) \tag{22}$$

Here, the monitor transient is normalized by its steady-state value, η_{ss}. Even without specifying steady-state values, it is possible to observe that the normalized transients, $f_E(t)$ and $f_\eta(t)$, are not proportional because their time constants differ. These two normalized functions were plotted in the left column of Figure 5.

Signal Deformation in Steady-State

Extensive simulations would be required to determine the specific steady-state values, E_{ss} and η_{ss}, for each threat in the space of signal deformations. These steady-state

biases, equal the raw monitor and range biases caused by signal deformation, are functions of correlator-peak shape. The simulations required to determine correlator-peak shape for each evil waveform in the signal-deformation threat space are beyond the scope of the present paper [9]. However, it is sufficient to note that these simulations can be used to map individual threats into the space of positive E_{ss} and η_{ss}. Fortunately, MERR can be mapped into this same space to yield a bound as a function of η_{ss}. GBAS integrity can be verified for each evil waveform in the signal-deformation threat model if MERR exceeds $E_{ss}(\eta_{ss})$ for each threat.

Mapping MERR onto Threat Space

The following procedure describes how to convert MERR(t) from a function of time to a function of the steady-state monitoring statistic, η_{ss}. For any given value of η_{ss} the MERR transient (14) can be evaluated using the monitor time response, (22), and intermediate equations for $P_{md,}$ P_{pl}, and K_{pl}. This value of MERR is a function of time, parameterized by the steady-state monitor statistic value: MERR(t; η_{ss}). To establish integrity, this MERR expression must equal or exceed the fault-induced error for the entire duration of the fault transient.

$$E(t) \le MERR(t;\eta_{ss}) \tag{23}$$

Dividing by the normalized transient for the ranging error, f_E, given by (21), the above inequality is rewritten as

$$E_{ss} \le \frac{MERR(t;\eta_{ss})}{f_E(t)}. \tag{24}$$

Since the MERR curve must be a bound for all time, the largest value of E_{ss} for which integrity can be established is the smallest value of right-hand side of (24) over all time. This minimum value depends on the parameterized monitor statistic, η_{ss}. The resulting bound on E_{ss} is called $\mathrm{MERR}_{ss}(\eta_{ss})$.

$$MERR_{ss}(\eta_{ss}) = \min_t \left(\frac{MERR(t;\eta)}{f_E(t)} \right) \tag{25}$$

This equation effectively maps MERR into the two-dimension space characterized by positive values of E_{ss} and η_{ss}. Any threat with a steady-state error value, $E_{ss}(\eta_{ss})$, that is below $\mathrm{MERR}_{ss}(\eta_{ss})$ is bounded with integrity.

Comparison of Time-Varying and Static MERR

In addition to providing a means to evaluate the signal-deformation integrity of arbitrary GBAS implementations, the MERR mapping provides a means to compare the performance of the time-varying MERR to previously proposed MERR implementations, such as MERR_{MP}. In evaluating MERR, a larger value is more desirable in order to bound larger errors. In present work on Category I GBAS, signal-deformation errors for certain receiver designs, particularly double-delta designs, exceed MERR [9]. A larger value of MERR might ensure integrity for the entire user space.

The static MERR derived by Zaugg (MERR_{MP}) can be converted to the space of positive E_{ss} and η_{ss} values by a process similar to that derived for the time-varying MERR, as summarized by (25). In applying (25) to the static MERR, MERR_{MP} is treated as infinite after the MDE time. Before the MDE time, MERR_{MP} has a constant value. This constant is $4.36\sigma_{min}$, as derived by Zaugg for the signal-deformation fault mode [5].

Both the static and time-varying MERR expressions are proportional to σ_{min}. This sigma term is the standard deviation of the fault-free ranging error, $\sigma_{ff,i}$, evaluated at 90° elevation. As described by (20), the ranging error consists of contributions from ground receiver noise, airborne noise, ionosphere noise and troposphere noise. For the purposes of this paper, the ground error was evaluated using the GAD-C3 curve, the airborne error was evaluated using the AAD-A curve, and the ionosphere error was evaluated at 6 km using the a vertical-ionosphere gradient sigma, σ_{vig}, equal 4 mm/km [10],[11]. The troposphere sigma was neglected. The resulting value of σ_{min} was 0.25 m.

The ratio of the allotted integrity risk to the fault prior for the time-varying MERR was assigned the same value used in [5]: $P_a/P_f = 10^{-3}$. The RDT was set to zero.

Using these parameters, the values of the time-varying and static MERR curves are plotted against η_{ss} in Figure 11. As shown in the bottom plot, the static MERR (green circles) is approximately equal to the time-varying MERR (black line) for this example. The small difference between the two curves is detailed in the top part of Figure 11, which plots the difference, $\Delta\mathrm{MERR}_{ss}$, defined as the value of the time-varying MERR curve minus the MERR_{MP} curve. The difference between the two curves results, primarily, from the assumed proportionality between the monitor statistic and ranging error biases in the formulation of MERR_{MP}. This proportionality is only approximate for the current example, since the time-constant for the monitor filter, at 50 seconds, is shorter than the time constant for the pseudorange smoothing filter, at 100 seconds. The step responses for the two filters are nearly linear (and hence proportional) only when the steady-state monitor statistic, η_{ss}, is large, beyond the right edge of the $\Delta\mathrm{MERR}_{ss}$ plot. Through much of the η_{ss} range, the MERR_{MP} curve is overconservative by up to 0.2 m relative to the time-varying MERR curve. At the low end of η_{ss}, where the

proportionality assumption breaks down, $MERR_{MP}$ is not a conservative bound, and $\Delta MERR_{ss}$ is negative. It can be concluded from Figure 11 that the static MERR is a reasonable approximation of the time-varying MERR for signal-deformation analysis, but that the static MERR may be invalid over certain regions and slightly overconservative over others.

The proportionality between the monitor statistic and ranging-error biases breaks down further if the RDT is not equal to zero. The time-varying MERR bound changes for non-zero RDT as illustrated in Figure 12. $MERR_{MP}$ is not shown in this figure, since the static MERR does not change with RDT. As is evident from comparing Figure 11 to Figure 12, the $MERR_{MP}$ formulation provides a poor approximation of the true MERR bound for cases when RDT is not equal to zero.

As shown in Figure 12, the MERR bound becomes much larger when the maximum transmission time is shorter than the required Time-to-Alert (i.e. when RDT is positive). The MERR bound becomes smaller when the transmission time is longer than the alert time (i.e. when RDT is negative). The difference is substantial, especially as the value of the steady-state monitor parameter, η_{ss}, becomes large. This result may introduce some benefit for Category I operations, if the maximum transmission time can be shown to be less than the six-second TTA. By contrast, this result will make integrity more difficult to establish for Category III operations, unless the maximum transmission time can be made shorter than the required two-second TTA. This effect also has a significant impact on the proposed Local Airport Monitor (LAM), which would rebroadcast WAAS signals in a LAAS format to enable Category I approach. The transmission time for LAM is 10 seconds, as compared to a TTA requirement of 6 seconds. As shown by Figure 12, the MERR curve is significantly depressed for the LAM case, with RDT equal -4 seconds.

CONCLUSION

This paper developed a novel method of deriving a range-domain bound to cover GBAS navigation errors in a fault scenario. This bound, called the time-varying MERR (Maximum-allowable Error in Range), replaces earlier versions of the MERR bound that did not take temporal effects into account. Specifically, the time-varying MERR formulation provides (1) a precise comparison criterion to validate GBAS integrity, (2) a relaxation of the requirements on monitor-filter design, and (3) a means to incorporate the Time-to-Alert requirement directly in the MERR bound.

The time-varying MERR concept was applied to develop a bound for the signal-deformation fault mode. For this example, it was shown that the older, static MERR formulation provided a reasonable approximation of the time-varying MERR for the baseline case. However, the static MERR approximation broke down when the Time-to-Alert requirement was considered. In fact, it is the difference between the alert time and the message transmission time, labeled the Relative Detection Time (RDT), which impacts the MERR bound. If the transmission time is shorter than the alert time, as may be the case for Category I GBAS, the time-varying MERR bound is significantly more generous than the static approximation. If the alert time is shorter than the transmission time, as may be the case for Category III GBAS, the time-varying MERR bound is significantly stricter than the static approximation.

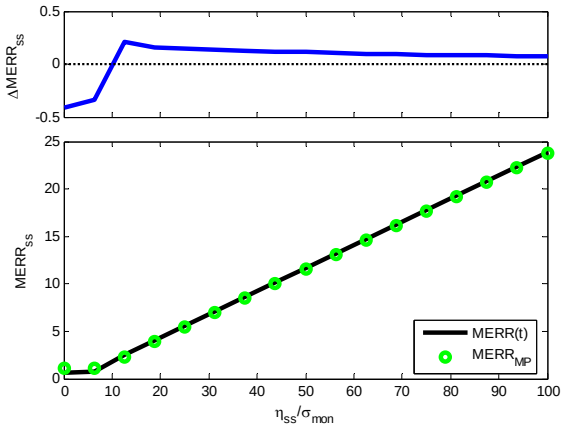

Figure 11. Comparison of MERR(*t*) to Fault-Induced Ranging Error, *E*(*t*)

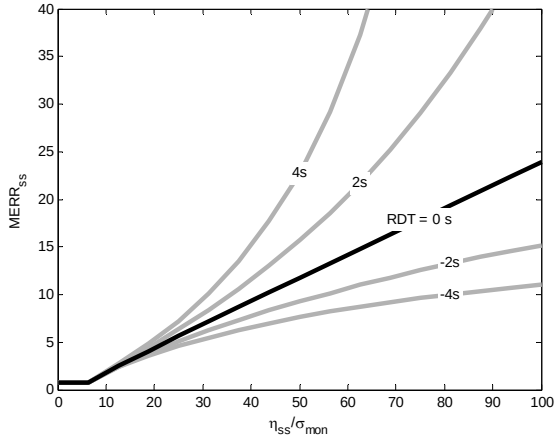

Figure 12. Time-Varying MERR Applied to Different Relative Detection Times (RDT)

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Federal Aviation Administration Satellite Navigation LAAS Program Office (AND-710) for supporting this research. The opinions discussed here are those of the authors and do not necessarily represent those of the FAA or other affiliated agencies.

REFERENCES

[1] RTCA Inc, *Minimum Aviation System Performance Standards for the Local Area Augmentation System (LAAS)*. RTCA/DO-245, 1998.

[2] C. Shively, *Derivation of Acceptable Error Limits for Satellite Signal Faults in LAAS*, Proceedings of the Institute of Navigation's ION-GPS 1999, pp. 761-770.

[3] B. DeCleene, *Defining Pseudorange Integrity – Overbounding*, Proceedings of the Institute of Navigation's ION-GPS 2000, pp. 1916-1924.

[4] S. Pullen, T. Walter, and P. Enge, *System Overview, Recent Developments and Future Outlook for WAAS and LAAS*, Proceedings of Tokyo University of Mercantile Marine GPS Symposium, November, 2002.

[5] T. Zaugg, *A New Evaluation of Maximum Allowable Errors and Missed Detection Probabilities for LAAS Ranging Source Monitors*, Proceedings of the Institute of Navigation's Annual Meeting 2002, pp. 187 - 194.

[6] B. Pervan. Personal Communication. September, 2005.

[7] J. Rife and E. Phelts. *Time-Varying MERR Applied to SDM*, Unpublished presentation from the Local Airport Monitor (LAM) Technical Interchange Meeting at the Federal Aviation Administration Technical Center, August, 2005.

[8] FAA, *Specification FAA-E-2937A: Category I Local Area Augmentation System Ground Facility*, 2002.

[9] R.E. Phelts, T. Walter, P. Enge, *Toward Real-Time SQM for WAAS: Improved Detection Techniques*, Proceedings of the Institute of Navigation's ION-GNSS 2003, pp. 2739-2749.

[10] G. McGraw, T. Murphy, M. Brenner, S. Pullen and A.J. Van Dierendonck, *Development of LAAS Accuracy Models*, Proceedings of ION GPS 2000, pp. 1212 - 1223.

[11] C. Shively and T. Hsiao, *Availability Enhancements for Cat IIIB LAAS*, NAVIGATION, Journal of the Institute of Navigation, Vol. 51, No. 1, Spring 2004.