# Integrity Lessons from the WAAS Integrity Performance Panel (WIPP)

Todd Walter, Per Enge, *Stanford University*,
Bruce DeCleene, *FAA*

**ABSTRACT**

The Wide Area Augmentation System (WAAS) is unlike any previous navigation system fielded by the FAA. Historically the FAA has implemented relatively simple and distributed systems. Each only affects a small portion of the airspace and each is maintained independently of the others. WAAS, in contrast, is a complex and centralized, system that provides guidance to the whole airspace. Consequently, the certification for WAAS must proceed very cautiously. WAAS is being pursued because its potential benefits are significant. It will provide guidance throughout the national airspace. It will enable approaches with vertical guidance to every runway end in the United States without requiring local navigational aids. It will enable advanced procedures such as curved approaches and departures. Eventually it will allow greater capacity through smaller separation standards. These and other benefits motivate the effort to create and certify this new type of system. Although the analysis becomes much more difficult, the system must maintain the same or higher level of safety than the existing infrastructure.

Another difference with WAAS is that it is inherently a non-stationary system. It relies on satellites that are constantly in motion and that may change their characteristics. Additionally, the propagation of the satellite signals varies with local conditions. Thus, the system has differing properties over time and space. However, the system requirements apply to each individual approach. In particular, the integrity requirement, that the confidence bound fails to contain the true error in fewer than one in ten million approaches, must apply to all users under all foreseeable operational conditions. To ensure that the integrity requirement would be met, the FAA formed the WAAS Integrity Performance Panel (WIPP). The role of the WIPP is to independently assess the safety of WAAS and to recommend system improvements. To accomplish these tasks, the WIPP had to determine how to interpret the integrity requirement for WAAS, develop algorithms to meet this requirement, and ultimately validate them.

**INTRODUCTION**

The Wide Area Augmentation System (WAAS) monitors the Global Positioning System (GPS) and provides both differential corrections to improve the accuracy and associated confidence bounds to ensure the integrity. WAAS utilizes a network of precisely surveyed reference receivers, located throughout the United States. The information gathered from these WAAS reference Stations (WRSs) monitors GPS and its propagation environment in real-time. However, WAAS designers must be aware of the limitations of its monitoring. The observables are corrupted by noise and biases causing certain fault modes to be difficult to detect. Because it is a safety-of-life system, WAAS must place rigorous bounds on the probability that it is in error, even under faulted conditions.

In late 1999, concerns arose over the WAAS design and the process by which WAAS was to be proven safe. In response, the FAA created the WAAS Integrity Performance Panel (WIPP). The WIPP is a body of GPS and system safety experts chartered to assess the system engineering and safety design of WAAS and recommend any required changes. The WIPP consists of members from government (FAA, JPL), industry (Raytheon, Zeta, MITRE), and academia (Stanford University). They first convened in early 2000 to address the integrity and certification of WAAS.

Primarily the WIPP quantified the degree to which WAAS mitigated the system vulnerabilities. Over the next two years, the WIPP changed the design of several system components where the system could not satisfactorily demonstrate the required level of integrity. As each threat was addressed, the WIPP built upon what it had learned.

Some of the main lessons that emerged from the WIPP are:

- The aviation integrity requirement of $10^{-7}$ per approach applies in principle to each and every approach. It is not an ensemble average over all conditions.

- For events where fault modes or rare events are not known, validated threat models are essential both to describe what the system protects against and to quantitatively assess how effectively it provides such protection.
- The system design must be shown to be safe against all fault modes and external threats, addressing the potential for latent faults just beneath the system's ability to detect them. Conventional non-aviation differential systems presume no failures exist until consistency checks fail.
- Analysis must take place primarily in the range or correction domain as opposed to the position domain.
- The small numbers associated with integrity analysis are not intuitive. Careful analysis must take priority over anecdotal evidence.

These lessons will be described in greater detail. Of these lessons, the need for threat models is the most important and was the most lacking. Threat models describe events or conditions that may cause harm to the user. In this case, harm is referred to as Hazardously Misleading Information (HMI). It is defined as a true error that is larger than the guaranteed Protection Level (PL). WAAS provides differential corrections that are applied to the received pseudoranges from GPS. At the same time, confidence bounds are also supplied to the user. These bounds are used, with the geometry of satellites about the user, to calculate the PL. In order to use the calculated position for navigation, the PL must be small enough to support the operation. The user only has real-time access to the PL, not the true error. Thus, HMI arises if the user has been told that the error in position is small enough to support the operation, but in fact, it is not.

The threat models must describe all the known conditions that could cause the true errors to exceed the predicted confidence bounds. Having a comprehensive list is essential to achieving the required level of safety and it also drives the system design. Additionally, restricting the scope of the threats is necessary for practical reasons. It is not possible to create a system that can protect against every conceivable threat. Fortunately, many such threats are either unphysical or extremely improbable. Restricting threats to those that are sufficiently likely is necessary for creating a practical system.

## INTEGRITY REQUIREMENT

The integrity requirement for precision approach guidance (APV-I through Category I) is $1\text{-}2\text{x}10^{-7}$ per approach [1]. There is a general understanding that this probabilistic requirement applies individually to every approach. This definition is further refined in the WAAS specification [2] as applying at every location and time in the service volume. Since WAAS provides service to a large number of runways, it is not acceptable for one airport to have less integrity simply because a different aircraft hundreds of miles away has margin against the requirement. Similarly, with the non-stationary characteristics arising from effects such as the orbiting satellites, it is not appropriate for operations to continue during an hour interval when the integrity requirement is not met, just because it is exceeded for the rest of the day. Generally, this can be restated as meaning that the probability of Hazardously Misleading Information (HMI) must be at or below $1\text{x}10^{-7}$ for an approach at the worst time and location in the service volume for which the service is claimed to be available. Despite this apparent understanding, a more detailed discussion of the interpretation is instructive.

The integrity requirement is that the Positioning Error (PE) must be no greater than the confidence bound, known as the Protection Level (PL), beyond the specified probability. Confusion may result because the requirement is probabilistic, yet at the worst time and place, the errors appear deterministic. Instead, the requirement should be viewed as applying to a hypothetical collection of users under essentially identical conditions. The collection of users, referred to as the ensemble, must be hypothetical in this case because satellite navigation systems and their associated errors are inherently non-stationary. Any true ensemble would average over too many different conditions, combining users with high and low risk. Thus, we must imagine an ensemble of users, for each point in space and time, whose errors follow probability distributions specific to that point.

Of course, there can only be one actual user at a given point in space and time. That user will experience a specific set of errors that combine to create the position error. These errors are comprised of both deterministic and stochastic components. The distinction is that if we could replicate the conditions and environment for the user, the deterministic components would be completely repeatable. Thus, these errors would be common mode; all users in our ensemble would suffer them to the same degree. On the other hand, stochastic errors such as thermal noise would differ for each user in our ensemble. Overall, these components combine to form a range of possible errors whose magnitudes have differing probability. When we look at a very large number (approaching infinity) of hypothetical users in the ensemble, some will have errors that exceed the

protection level while most will not. The fraction of users that exceed the PL can be used to determine the probability of an integrity failure under those conditions.

The difficult aspect of applying this philosophy is defining equivalent user conditions and then determining the error distributions. A circular definition is that user conditions can be called equivalent if they carry the same level of risk. A more practical approach is to exploit prior knowledge of the error sources. For example, if it were known that an error source only has a definite temperature dependency, then the ensembles should be formed over all users in narrow temperature ranges. The error distributions and probability of exceeding the PL would be calculated for each ensemble, and the integrity requirement would have to be met for the most difficult case for which availability is claimed. Unfortunately, true error sources usually have multiple dependencies and these dependencies are different between the various error sources. Thus, the ensembles may need to be formed over narrow ranges of numerous parameters. However, great care must be taken because if certain dependencies are not properly recognized, the ensembles may unknowingly average over different risk levels.

The restatement of the requirement that it applies to the worst time and location is misleading because it is acceptable to average against certain conditions. Some events may be sufficiently rare to ignore altogether. If, under similar conditions, the *a priori* likelihood is well below $1x10^{-7}$ per approach (considering exposure time to the failure), then there may not be any need to provide additional protection. The worst time and place should not be viewed as when and where this unlikely event occurs. The event need only be considered if it is sufficiently likely to occur, if when and where it is most likely to occur can be predicted ahead of time, or if it is strongly correlated with an observable. Even if the event is not sufficiently rare to be ignored, its *a priori* probability may be utilized provided the event remains unpredictable and immeasurable. Thus, the conditions where the event is present may be averaged with otherwise similar conditions without the event. Taking advantage of such *a priori* probabilities must be approached very cautiously on a case-by-case basis.

The goal is to ensure that all users are exposed to risk at no greater than the specified rate of $10^{-7}$ per approach. Thus, ensembles that cannot be correlated in some way with user behavior do not make sense. For example, users may tend to fly to the same airport at the same time of day or during a certain season. Therefore, an ensemble of all users with a specific geometry at a certain location and certain time of day, but theoretically infinitely extended forwards and backwards over adjacent days is reasonable. On the other hand, an ensemble of all users whose thermal noise consists of five-sigma errors aligned in the worst possible direction is neither realistic nor practical. The latter example attempts to combine rare and random events into a unifying ensemble that cannot be made to correspond to user behavior. In general, conditions leading to high risk that are both rare and random can be averaged with lower risk conditions. The requirement for rarity seeks to assure that users do not receive multiple exposures to the high-risk condition, while the requirement for randomness seeks to avoid a predictable violation of the integrity requirement. However, the correlation with conceivable user behavior must be the determining factor when deciding whether or not to average the risk. Similarly, a correlation with a system observable should be exploited to protect the user when performance goes out of tolerance.

Deciding how to define the ensembles provides the necessary information for determining the error distributions. Components will largely be divided into noise-like contributions, with some spread in their values, and bias-like contributions whose values are seen as fixed although probably unknown. Although many of these error sources may be deterministic, practically they may need to be described in stochastic terms. Many error sources fall into this category including ionosphere, troposphere, and multipath. If we knew enough about the surrounding environments, we could predict their effects for each user. However, because it is usually not practical to obtain this information it may be acceptable to view these effects as unpredictable as long as their effects cannot be correlated with user behavior.

Knowledge of the error characteristics is very important in evaluating system design. While impossible to know fully, many important characteristics such as dependencies may be recognized. This knowledge allows proper determination of the error distributions. After defining the individual distributions, the correlations between them must be established. Many deterministic error sources will affect multiple ranging sources simultaneously. Correlated deterministic errors may add together coherently for a specific user. Such effects require larger increases in the protection level than if the errors were uncorrelated. If these effects are not recognized and treated appropriately, the integrity requirement will not be met and the user will suffer excessive risk. Although the form of the protection level equations given in [1] and [2] suggest that all error sources are independent, zero-mean, and gaussian, this is not the case under all operating conditions. Each error source must be carefully analyzed, both individually and

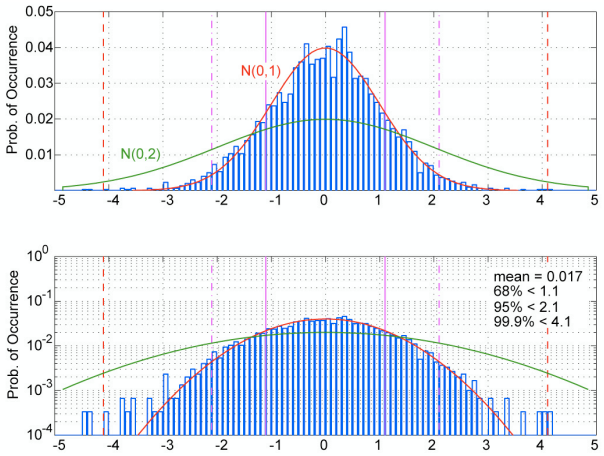


**Figure 1.** *Simulated probability distribution composed of a mixed gaussian: 2900 points with unity variance and 100 points with a variance of four. The top and bottom graphs are the same data displayed on a linear scale (top) and log-scale (bottom).*

in relation to the other sources. Only then can the appropriate confidence bounds be determined.

## ERROR MODELING

Each individual error source has some probability distribution associated with it. This distribution describes the likelihood of encountering a certain error value. Ideally, smaller errors are more likely than larger errors. Generally, this is true for most error sources. The central region of most error sources can be well described by a gaussian distribution. That is, most errors are clustered about a mean (usually near zero) and the likelihood of being farther away from the mean falls off according to the well-known model. This is often a consequence of the central-limit-theorem that states that distributions tend to approach gaussian as more independent random variables are combined.

Unfortunately, the tails of the observed distributions rarely look gaussian. Two competing effects tend to modify their behavior. The first is clipping, because there are many cross-comparisons and reasonability checks, the larger errors tend to be removed. Thus, for a truly gaussian process, outlier removal would lead to fewer large errors than would otherwise be expected. The second effect is mixing. The error sources are rarely stationary. Thus, some of the time the error might be gaussian with a certain mean and sigma and at other times have a different distribution. Such mixing may result from a change in the nominal conditions or from the introduction of a fault mode. Mixing generally leads to broader tails or large errors being more likely than otherwise expected.

The mixing causes additional problems. If the error processes were stationary, it would be possible to collect as large a data set as practical and then conservatively extrapolate the tail behavior using a gaussian or other model. However, because the distribution changes over time, it is more difficult to predict future performance based on past behavior. Furthermore, mixing leads to more complicated distributions whose tails are more difficult to extrapolate. With enough mixing, it can be very difficult to characterize the underlying distributions at all. Figure 1 is an example of a mixed distribution. The majority of the data points are selected from a zero-mean gaussian with unity variance. A few of the points are selected from a zero-mean gaussian with a variance of four. This plot contains some very typical features of the real data we collect. The majority of the data conforms very well to a gaussian model, while the tails usually contain outliers. Sampling issues are usually significant as it is very difficult to obtain large amounts of independent data. Thus, just by looking at the graph it is difficult to determine the actual distribution.

The central-limit-theorem causes error distributions to approach gaussian as several independent sources are combined. Certainly, the main body of collected data tends to be gaussian in appearance. The tails are more difficult to discern. A generalized mixed gaussian description is appropriate. Here, the errors can be described as gaussian where the mean and variance are also drawn from some joint probability distribution.

$$\square\square \; N(\square,\square)$$
$$\square,\square \; \square \; pr(\square\,,\square)$$

At any given instant, the error is gaussian, but its mean and variance have some uncertainty. By understanding the extent of the possible means and variances we can overbound the worst-case. Additional information ideally allows us to partition the space and distinguish when larger bounds are needed versus when smaller ones can be provided.

Nominally, we expect the distribution to be zero-mean and have some well-defined variance. Some small fraction of the time the error may still be zero-mean, but have a larger variance as depicted in Figure 1. During a fault mode the mean may grow in magnitude, but the variance may stay roughly the same as nominal (of course other variances are possible). Restricting the error

distribution to this class distribution allows the analysis to become tractable.

Of course, it is impossible to truly know the real distribution, particularly to $10^{-7}$ confidence. The use of a model like this must be accepted by a body of experts such as the WIPP who can assert that it is valid based on physical knowledge of the system design, supporting data, and simulation. This combination is essential for describing the tail behavior. A physical understanding of the error process is essential to describing expected behavior. Data must be collected in sufficient quantity and under many conditions. The physical knowledge must be exploited to determine what the worst-case conditions are and how data should be reduced. For example, severe ionospheric behavior is correlated with solar events and magnetic disturbances. Data must be collected during some of the most extreme operating conditions. Finally, simulation may be used to confirm that the models constructed are consistent with the observations.

Physical knowledge of the system is essential. Any information on the physical processes behind the error source can be used to separate mixtures and create better-defined distributions. For example, multipath can be related to the surrounding environment. Large reflections tend to occur at lower elevation angles. Partitioning data by elevation angles may reduce mixing. Changes to multipath can be related to changes in satellite position and to changes in the environment. Excessive multipath can sometimes be related to specific reflectors. Additionally, the magnitude of multipath errors can be bounded, by limiting the number of reflectors and strength of the reflected signals.

Data is also essential. The data must be sufficient to support assumptions or validate system performance to the degree to which the safety of the system relies on that data. It is not sufficient to collect a day or two of randomly selected data, but many days collected under extreme conditions. Examples include tropospheric data from many different climates, ionospheric data from the worst times in the 11-year solar cycle, multipath data from the most cluttered environments etc. Rare events are unlikely to be captured in small data sets. Large data sets taken over long time-periods are more likely to capture postulated events. Having data containing these events provides better insight into their effect.

## THREAT MODELS

Threat models describe the anticipated events that the system must protect the user against and conditions during which it must provide reliably safe confidence bounds. The threat model must describe the specific nature of the threat, its magnitude and its likelihood. Together, the various threat models must be comprehensive in describing all reasonable conditions under which the system might have difficulty protecting the user. Ultimately they form a major part of the basis for determining if the system design meets its integrity requirement. Each individual threat must be fully mitigated to within its allocation. Only when it can be shown that each threat has been sufficiently addressed can the system be deemed safe.

WAAS is being developed primarily to address existing threats to GPS. However, it runs the risk of introducing threats in absence of any GPS fault. By necessity, it is a complex system of hardware and software. Included in any threat model must be self-induced errors. Some of these errors are universal to any design while others are specific to the implementation. For example, the software design assurance of WAAS reference receivers was based on market availability of equipment, so reference receivers software faults were a unique threat that had to be mitigated through system integrity monitoring. The following is a high level list of generic threats. While it is not comprehensive, it does include the most significant categories either for magnitude of effect or likelihood. There are numerous other threats that have a smaller effect, are less likely, or are implementation specific.

**High-Level Threat List**

- Satellite
  - Clock/ephemeris error
  - Signal deformation
  - Code carrier incoherency
- Ionosphere
  - Local non-planar behavior
    - Well-sampled
    - Undersampled
- Troposphere
- Receiver
  - Multipath
  - Thermal noise
  - Antenna bias
  - Survey errors
  - Receiver errors
- Master station
  - SV clock/ephemeris estimate errors
  - Ionospheric estimation errors

    - SV Tgd estimate errors
    - Receiver IFB estimate errors
    - WRS clock estimate errors
    - Communication errors
    - Broadcast errors
- User errors

The following sections provide greater detail for each threat, although the true details depend on implementation and must be decided by the service provider.

**SV Clock/Ephemeris Estimate Errors**
Satellites suffer from nominal ephemeris and clock errors when there are no faults in the GPS system [3]. Additionally, the broadcast GPS clock and ephemeris information may contain significant errors in the event of a GPS system fault or erroneous upload. Such faults may create jumps, ramps, or higher order errors in the GPS clock, ephemeris, or both [4] [5]. Such faults may be created by changes in state of the satellite orbit or clock, or simply due to the broadcasting of erroneous information. Either the user or the system may also experience incorrectly decoded ephemeris information.

The UDRE must be sufficient to overbound the residual errors in the corrected satellite clock and ephemeris.

**Signal Deformations**
ICAO has adopted a threat model to describe the possible signal distortions that may occur on the GPS L1 CA code [1]. These distortions will lead to biases that depend upon the correlator spacing and bandwidth of the observing receivers. Such biases would be transparent to a network of identically configured receivers [6].

The UDRE must be sufficient to overbound unobservable errors caused by signal deformation. Unobservable errors are those that cannot be detected to the required level of integrity.

**Code-Carrier Incoherency**
A postulated threat is that a satellite may fail to maintain the coherency between the broadcast code and carrier. This fault mode occurs on the satellite and is unrelated to incoherence caused by the ionosphere. This threat causes either a step or a rate of change between the code and carrier broadcast from the satellite. This threat has never been observed, but nevertheless must be protected against as a postulated satellite failure.

The UDRE must be sufficient to overbound unobservable errors caused by incoherency. Unobservable errors are those that cannot be detected to the required level of integrity.

**Ionosphere and Ionospheric Estimation Errors**
The majority of the time, mid-latitude ionosphere is easily estimated and bounded using a simple local planar fit. However, periods of disturbance occasionally occur where simple confidence bounds fall significantly short of bounding the true error [7]. Additionally, in other regions of the world, in particular equatorial regions, the ionosphere often cannot be adequately described by this simple model [8]. Some of these disturbances can occur over very short baselines causing them to be difficult to describe even with higher order models. Gradients larger than three meters of vertical delay over a ten-kilometer baseline have been observed, even at mid-latitude [9]. These worst-case gradients are a threat to both SBAS and GBAS.

The broadcast ionospheric grid format specified in the MOPS may also limit accuracy and integrity. The simple two-dimensional model and assumed obliquity factor may not always provide an accurate conversion between slant and vertical ionosphere. There will also be instances where the five-degree grid is too coarse to adequately describe the surrounding ionosphere.

There are times and locations where the ionosphere is very difficult to model. This problem may be compounded by poor observability [10]. Ionospheric Pierce Point (IPP) placement may be such that it fails to sample important ionospheric structures. This may result from the intrinsic layout of the reference stations and satellites, or from data loss through station, satellite, or communication outages. As a result, certain ionospheric features that invalidate the assumed model can escape detection.

Finally, because the ionosphere is not a static medium there may be large temporal gradients in addition to spatial gradients. Rates of change as large as four vertical meters per minute have been observed at mid latitudes [9].

The GIVE must account for inadequacies of the assumed ionospheric model, restrictions of the grid, and limitations of observability. The GIVE must be sufficient to protect against the worst possible ionospheric disturbance that may be present in that region given the IPP distribution. Additionally, since each ionospheric correction does not time out until after ten minutes, the GIVE and the Old But Active Data (OBAD) terms [11] must protect against any changes in the ionosphere that can occur over that time scale. Because the physics of the ionosphere are incompletely understood, the most practical ionospheric

threat models are heavily data driven and contain a large amount of conservatism.

**Tropospheric Errors**

Tropospheric errors are typically small compared to ionospheric errors or satellite faults. Historical observations were used to formulate a model and analyze deviations from that model [12]. A very conservative bound was applied to the distribution of those deviations. The model and bound are described in the MOPS and SARPS [11][1]. These errors may affect the user both directly through their local troposphere, and indirectly through errors at the reference stations that may propagate into satellite clock and ephemeris estimates. The user protects against the direct effect using specified formulas.

The master station must ensure that the UDRE adequately protects against the propagated tropospheric errors and their effect on satellite clock and ephemeris estimates. Of particular concern are the statistical properties of these error sources. These errors may be correlated for long periods, and will produce correlated errors across all satellites at a reference station and each receiver at the reference station.

**Multipath and Thermal Noise**

Multipath is the most significant measurement error source. It limits the ability to estimate the satellite and ionospheric errors. It depends upon the environment surrounding the antenna and the satellite trajectories. While many receiver tracking techniques can limit its magnitude, its period can be tens minutes or greater [13]. Additionally, it contains a periodic component that repeats over a sidereal day. Thus, severe multipath may be seen repeatedly for several days or longer.

Since all measurements that form the corrections and the UDREs and GIVEs are affected by multipath, great care must be used to bound not only its maximum extent but its other statistical characteristics as well (non-gaussian, non-white, periodic, etc.). There is potential for correlation between measurements and between antennas at a single reference site. Additionally the local environment may change either due to meteorological conditions (snow, rain, ice), or physical changes (new objects or structures placed nearby).

If carrier smoothing is used to mitigate multipath then robust cycle slip detection is essential. Half integer cycle slips have been observed on different brands of receivers. In one case, several half cycle slips were observed in the same direction each several minutes apart resulting in a several meter error. Cycle slip detection must be able to reliably catch unfortunate combinations of L1 and L2 half and full integer cycle slips in order to achieve an unbiased result.

**Antenna Bias**

Look-angle dependent biases in the code phase on both L1 and L2 are present on GPS antennas. These biases may be several tens of centimeters. In the case of at least one antenna, they did not become smaller at higher elevation angle. These biases are observable in an anechoic chamber, but extremely difficult to observe in operation. They may result from intrinsic antenna design as well as manufacturing variation.

While the particular orientation of each antenna and bias is random, it is also static. Therefore, there may exist some points in the service volume where the biases tend to add together coherently consistently. Thus, these locations will experience this effect day after day. To protect these regions, the biases should be treated pessimistically as though they are all nearly worst-case and coherent. Calibration may be applied, although individual variation, the difficulty of maintaining proper orientation, and the possibility of temporal changes, hamper its practicality.

**Survey Errors**

Errors in the surveyed coordinates of the antenna code phase center can affect users in the same manner as antenna biases. However, survey errors tend to be much smaller in magnitude and cancel between L1 and L2.

These errors can typically be lumped in with antenna bias protection terms and mitigated in the same manner.

**Receiver Errors**

The receivers themselves can introduce errors through false lock or other mechanisms including hardware failure (GPS receiver, antenna, atomic frequency standard). These events have been observed to be rare and uncorrelated.

These may be mitigated through the use of redundant and independent receivers, antennas, and clocks, at the same reference station. However, the UDRE and GIVE must still protect against small errors may exist up to the size of the detection threshold.

**Tgd and IFB Estimation Errors**

For internal use, the correction algorithms often need to know the hardware differential delay between the L1 and L2 frequencies. These are referred to as Tgd for the bias on the satellite and IFB for the InterFrequency Bias in the reference station receivers. These values are typically estimated in tandem with the ionospheric delay estimation

| PRN | *EL* | *AZ* | $\sigma_i$ | $S_{3i}$ | $S_{3i}$ without PRN 8 |
|---|---|---|---|---|---|
| 2 | 45.8° | -32.3° | 2.34 m | 0.595 | 0.451 |
| 5 | 11.2° | -76.8° | 10.1 m | 0.258 | 0.437 |
| 6 | 36.6° | 48.4° | 2.32 m | 0.162 | 2.005 |
| 8 | 9.98° | 73.0° | 3.74 m | 1.000 | - |
| 9 | 61.4° | 28.5° | 2.03 m | -1.928 | -3.087 |
| 15 | 32.8° | 151.0° | 6.89 m | -0.015 | 0.174 |
| 21 | 42.3° | -136.0° | 4.83 m | 0.066 | -0.003 |
| 122 | 40.6° | 120.1° | 6.19 m | -0.139 | 0.022 |

***Table 1.*** *Satellite elevation and azimuth angles, confidence bounds and projection matrix values both for the all-in-view solution and without PRN 8.*

[14]. Although these values are nominally constant, there are some conditions under which they may change their value. One method is component switching, if a new receiver or antenna is used to replace an old one, or if different components or paths are made active on a satellite. Another means is through thermal variation either at the reference station or on the satellite as it goes through its eclipse season. Finally, component aging may also induce a slow variation

The estimation process may have difficulty in distinguishing changes in these values from changes in the ionosphere. The steady state bias value and step changes may be readily observable, but slow changes comparable to the ionosphere may be particularly difficult to distinguish. Ionospheric disturbances that don't follow the assumed model of the ionosphere may also corrupt the bias estimates. The UDREs and GIVEs must bound the uncertainty that may result from such estimation errors.

**Receiver Clock Estimate Errors**

Similarly, the satellite correction algorithm must estimate and remove the time offsets between the reference station receivers. These differences are nominally linear over long times for atomic frequency standards. However, component replacement or failure may invalidate that model.

Nominally, these differences are easily separated, however, reference station clock failures and/or satellite ephemeris errors may make this task more difficult. The UDRE must protect against errors that may propagate into the satellite clock and ephemeris correction due to these errors. Particular attention must be paid to correlations that may result from this type of misestimation

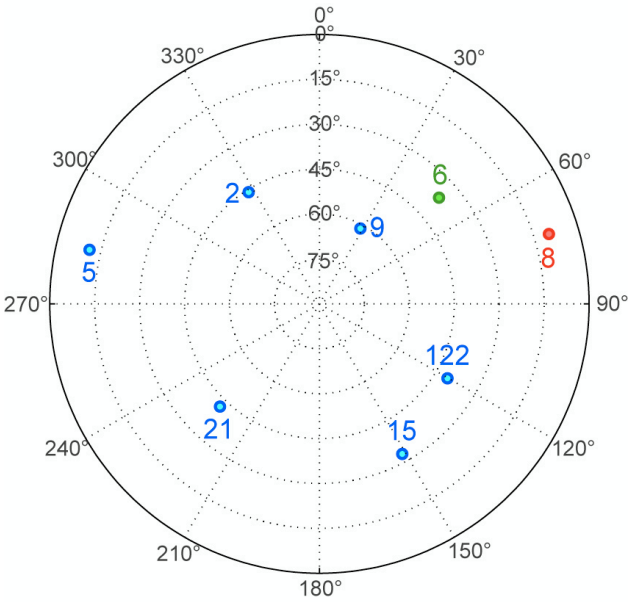


***Figure 2.*** *Satellite elevation and azimuth values for a standard skyplot. PRN 8 is a low elevation satellite that if not included in the solution dramatically changes the influence of PRN 6.*

**RANGE DOMAIN VS. POSITION DOMAIN**

The HMI requirement is specified in the position domain, yet WAAS broadcasts values in the range/correction domain. The users combine the corrections and confidences with their geometry to form the position solution and protection level. Exactly which corrections and satellites are used is known only to the user. Therefore, how the position error depends on the residual errors is known only to the users. WAAS cannot monitor solely in the position domain and fully protect its users. A combination of position domain and range/correction domain monitoring is most efficient.

To see this effect we can look at a specific user geometry. This example was created using Stanford's Matlab Algorithm Availability Simulation Tool (MAAST) [15] which can be used to simulate WAAS performance. The user has eight satellites in view as shown in Table 1. Figure 2 shows the elevations and azimuths of the satellites along with their PRN values. Table 1 also shows the PRN, elevation, azimuth, and one sigma confidence bound ($\sigma_i$). In addition, the fifth column shows the dependence of the vertical error to a pseudorange error on that satellite, $S_{3i}$. $\mathbf{S}$ is the projection matrix and is defined as $\mathbf{S} = (\mathbf{G}^T\mathbf{W}\mathbf{G})^{-1}\mathbf{G}^T\mathbf{W}$, where $\mathbf{G}$ is the geometry matrix and $\mathbf{W}$ is the weighting matrix, see Appendix J of [11]. This term multiplies the error on the pseudorange to determine the contribution to the vertical

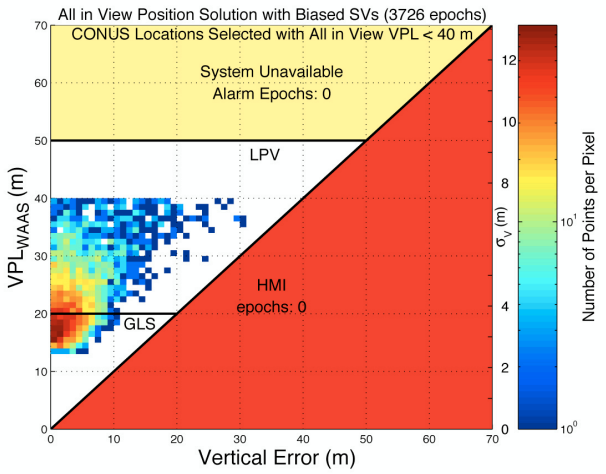


***Figure 3.*** *The triangle plot for all-in-view solutions including one biased satellite in each is shown. Here each bias is deweighted by the other satellites. No obvious problems are evident in this chart.*

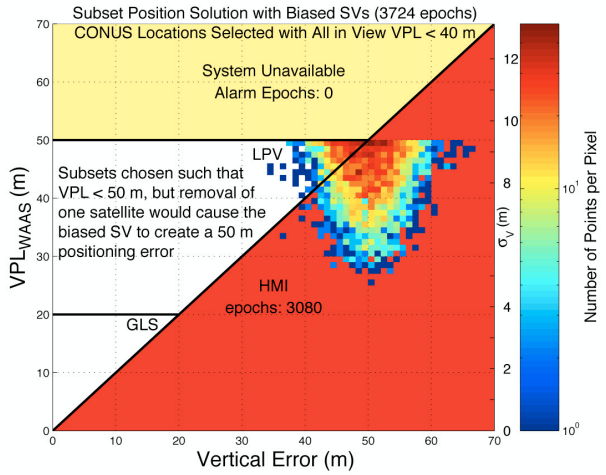


***Figure 4.*** *The triangle plot for the subset solutions that expose each biased satellite is shown. Here the biases are exposed as being hazardous for the user. This demonstrates the importance of checking each subset or in the range domain.*

error. Thus a 1 m ranging error on PRN 2 would create a positive 59.5 cm vertical error for the user with this combination of satellites and weights. The final column in Table 1 shows the projection matrix values if PRN 8, a low elevation satellite, is not included in the position solution.

With the all-in-view solution, the user has a VPL of 33.3 m (HPL = 20.4 m). When PRN 8 is dropped, the VPL increases to 48.6 m (HPL = 20.5 m). Both values are below the 50 m Vertical Alert Limit (VAL) for LPV [16]. Either solution could be used for vertical guidance. Notice that the vertical error dependency changes dramatically with the loss of PRN 8. In particular, PRN 6, which had little influence over the all-in-view solution, now has a very strong impact on this subset solution. Also notice that the other values change as well. PRNs 2, 21, and 122 lose influence while PRNs 5, 6, 9 and 15 become more important. More surprisingly, the influences of PRNs 15, 21, and 122 change sign; therefore, what was a positive error for the all-in-view solution becomes a negative error for this particular subset.

The changes in the $S_{3i}$ values with subset or superset position solutions limit the ability to verify performance exclusively in the position domain. For example, if PRN 6 had a 25 m bias on its pseudorange, it would lead to a vertical error of greater than 50 m with PRN 8 missing, but just over 4 m for the all-in-view solution. A position domain check with all satellites would not be concerned with a 4 m bias compared to a 33.3 m VPL. Thus, one would be inclined to think that all was well. However, the user unfortunate enough to lose PRN 8 would suffer a 50 m bias, large enough to cause harm. A 25 m bias would be more than a ten-sigma error in the range domain and thus would be easily detectable. Therefore, it is the combination of range and position domain checks that protect users with different combinations of satellites. It may be possible to work exclusively in the position domain by using subset solutions, however that approach is numerically much more intensive and significantly more complex when considering a wide area system that must consider users throughout the service volume.

There is nothing unique about this particular geometry. To investigate how position errors can hide for one combination of satellites and be exposed for another, we set MAAST to look for subset solutions that had very different $S_{3i}$ values in its subset solutions. We restricted the search to geometries that had VPLs below 40 m for all-in-view and then only investigated subsets with VPLs below 50 m. Of the 3726 geometries investigated, only two did not change $S_{3i}$ values by more than 40%.

To better illustrate the effect, the remaining 3724 geometries had biases placed on the satellite with the largest change. Each bias was chosen such that it would lead to a 50 m positioning bias in the subset solution (a 25 m bias on PRN 6 in the example above). Each pseudorange was also assigned a zero-mean gaussian error with a standard deviation of one half of its one-sigma confidence bound (column four of Table 1). The broadcast WAAS confidence bounds are approximately

three times larger than the nominal no-fault values (this inflation is necessary to protect against fault modes). We then calculated position errors and VPLs for both the all-in-view and subset solutions. The results are plotted in standard triangle charts, Figure 3 and 4.

Figure 3 is similar in appearance to a nominal triangle chart except the VPLs are clipped at 40 m due to our selection process and the position errors are worse than normal due to the injected error on the single satellite. However, the position errors are all below the VPL and the aggregate is not obviously biased. An observer might be inclined to declare that the system is functioning safely based on this chart. However, Figure 4 shows that with the same errors and biases, but a slightly different geometry, this is not true. The subset solution removes satellites that were masking the bias for each case. The result is an obviously faulted triangle chart. Thus, a triangle chart without obvious faults, Figure 3, is no guarantee of a safe system, as evidenced by Figure 4.

This simulation was pessimistic in its construction since the minimum, unacceptable error was placed on the most sensitive satellite. On the other hand, the geometries were chosen at random and do not have any unique subset characteristics. We were surprised we could create such radically different triangle charts from the same pseudorange errors and one biased satellite simply by looking at subsets. We were aware that we could corrupt two or more pseudoranges to create arbitrarily large errors in the subset solution and zero error for the all-in-view, but had not realized how well single errors could be hidden. The lesson is that it is not sufficient to observe a particular set of position solutions. The most effective method is to combine position domain monitoring with range domain monitoring.

## SMALL NUMBERS AND INTUITION

The integrity requirement of $10^{-7}$ is an incredibly small number. In fact, it has to be; there have been more than $2.5 \times 10^{7}$ landings per year in the US each year for the last 10 years [17]. Granted only a small fraction of these are instrument landings in poor visibility, however a larger value could have a noticeable effect on the overall accident rate. Furthermore, air traffic is expected to increase over the coming years. To reduce the total number of accidents while increasing the number of flights requires lowering the risk per operation. Satisfying and exceeding the WAAS integrity requirement is part of that overall strategy.

It is hard to imagine the exceedingly small probability of one part in ten million. By design, no individual will sample anything approaching that number of approaches. At most, an individual will sample of order tens of thousands of approaches, typically far less. Additionally that individual will likely mostly experience nominal conditions, and rarely the unusual events, such as ionospheric disturbances, where WAAS still has to meet $10^{-7}$. Thus, personal experience is only sensitive to $10^{-4}$ at best. Stand-alone GPS already has this level of integrity, so the design issues for WAAS can be counter-intuitive. It is because so many flight operations take place under such a variety of conditions that WAAS needs to extend integrity to $10^{-7}$. The greater populace samples the system every year in a more thorough way than any individual will in a lifetime.

WAAS is specifically in place to protect against rare events, events that one will infrequently encounter. As a result, the situations that WAAS is designed to protect against run counter to our intuition. It is tempting to say that an event such as the signal deformation that occurred on SVN 19 back in 1993 [18] is sufficiently unlikely to occur again that we do not need to worry about it. However, it did occur, and there is no basis to assume it will not occur again. We can estimate the likelihood of reoccurrence from the single observed event. The Block II satellites so far have a cumulative lifetime of roughly 280 years (through January 2003). One event over 280 satellite years is a likelihood of occurrence of approximately $4 \times 10^{-7}$ per satellite per hour. This is an exceedingly small number and one that many people may easily dismiss. However, a user has on average eight satellites in their solution for a combined probability of experiencing this effect of $3.2 \times 10^{-6}$ per hour. This is more than an order of magnitude over the entire integrity budget. In addition, if there were no capability to detect and mitigate the condition should it occur, then all subsequent users would be exposed to unacceptably high risk. Meanwhile, the system allocation to this effect is much less than 1% of the total, or below $10^{-9}$ per hour. Thus, it is not possible to dismiss this event out of hand. In addition, this calculated value of reoccurrence is too small. It is possible that other events have escaped attention or that we have been fortunate to date. In fact, a recent paper indicates that some form of signal deformation may be present on other satellites [19].

Other events with far greater probability are sometimes dismissed because they are perceived to be remote. The possibility that eight satellite biases all line up to have the same sign may seem to be extremely unlikely however if either sign is equally likely and they are all uncorrelated, the probability is twice (either all positive or all negative)

$2^{-8}$, or one in 128. This is roughly 1%, rare, but not nearly improbable enough to dismiss it. Additionally the likelihood that the most significant biases line up is even greater. Thus, one must work to reduce the likelihood of biases, or ensure they are all positively correlated, as random biases have a non-negligible chance of aligning.

By necessity, WAAS must work with very small numbers, probabilities of $10^{-7}$ and below. These probabilities are outside of personal experience and intuition. Events that seem unlikely must have an upper bound calculated for them. They should not simply be dismissed out of hand. Unless one does the calculation they may not be able to distinguish between probabilities of $10^{-5}$ and a $10^{-7}$.

## CONCLUSIONS

Augmentations systems for aviation are very different from conventional differential GPS. They are supplementing and ultimately replacing proven navigational aids whose safety has been demonstrated over many years of operational experience. Consequently their safety must be proven before they are put into service. Over the course of documenting the proof of safety, the WIPP learned many important lessons. Chief among these was the use of threat models. Threat models define our fault modes, how they manifest themselves and how likely they are. They describe what we must protect against. A well-defined threat model permits a quantitative assessment of the mitigation strategy. The quantitative assessment as opposed to a qualitative assessment is essential to establishing $10^{-7}$ integrity.

Another key lesson is the application of the $10^{-7}$ integrity requirement to each approach that changed how we viewed certain error sources. Rather than averaging over conditions with different risk levels, we must overbound the conditions describing the worst allowable situation. *A priori* probabilities may be used only for events that are infrequent, unpredictable, and unobservable. For example, ionospheric storms occur with certainty, therefore the system must provide at least $10^{-7}$ integrity while ionospheric disturbances are present. However, the onset time, exactly when the mid-latitude ionosphere will transition from a period of quiet to a disturbed state, is both rare and random. Thus, we may apply an *a priori* to that brief period of time when the ionosphere may be disturbed, but we haven't yet detected it. This lesson affects how we view all of our *a priori* failure rates and probability distributions.

## ACKNOWLEDGMENTS

The work for this paper was supported by the FAA Satellite Product Team under research grant 95-G-005. The authors gratefully acknowledge the contributions from the other WIPP members.

## REFERENCES

[1] ICAO Standard and Recommended Procedures (SARPS) Annex 10

[2] FAA-E-2892C WAAS Specification

[3] Jefferson, D. C. and Bar-Sever, Y. E., "Accuracy and Consistency of GPS Broadcast Ephemeris Data," in Proceeding of ION GPS-2000, Salt Lake City, UT, September 2000.

[4] Hansen, A., Walter, T., Lawrence, D., and Enge, P., "GPS Satellite Clock Event of SV#27 and Its Impact on Augmented Navigation Systems," in proceedings of ION GPS-98, Nashville, TN, September 1998.

[5] Shank, C. M. and Lavrakas, J. "GPS Integrity: An MCS Perspective," in Proceeding of ION GPS-1993, Salt Lake City, UT, September 1993.

[6] Phelts, R. E., "Multicorrelator Techniques for Robust Mitigation of Threats to GPS Signal Quality," Stanford University Thesis, June, 2001. http://waas.stanford.edu/~wwu/papers/gps/PDF/ericthesis.pdf

[7] Walter, T., et al. "Robust Detection of Ionospheric Irregularities," in NAVIGATION, Journal of the Institute of Navigation, vol. 48, no. 2, Summer 2001.

[8] Klobuchar, J., et al., "Ionospheric Issues for a SBAS in the Equatorial Region," in proceedings of the 10th International Ionospheric Effects Symposium, Alexandria, VA, May, 2002.

[9] Datta-Barua, S., Walter, T., Pullen, S., Luo, M., and Enge, P., "Using WAAS Ionospheric Data to Estimate LAAS Short Baseline Gradients," in Proceeding of ION NTM, San Diego, CA, January, 2002.

[10] Sparks, L., et al., "The WAAS Ionospheric Threat Model," in Proceedings of the International Beacon Satellite Symposium, Boston, June, 2001.

[11] WAAS Minimum Operational Performance Specification (MOPS), RTCA document DO-229C

[12] Collins, J. P. and Langley, R. B., "The residual tropospheric propagation delay: How bad can it get?," in proceedings of ION GPS-98, Nashville, TN, September 1998.

[13] Shallberg, K., Shloss, P., Altshuler, E., and Tahmazyan, L., "WAAS Measurement Processing, Reducing the Effects of Multipath," in Proceeding of ION GPS-2001, Salt Lake City, UT, September 2001.

[14] Hansen, A., "Tomographic Estimation of the Ionosphere Using Terrestrial GPS Sensors," Stanford University Thesis, March, 2002.

[15] Jan, S., Chan, W., Walter, T., and Enge, P., "Matlab Simulation Toolset for SBAS Availability Analysis," in Proceeding of ION GPS-2001, Salt Lake City, UT, September 2001.

[16] Cabler, H. and DeCleene, B., "LPV: New, Improved WAAS Instrument Approach," in Proceedings of ION GPS-2002, Portland, OR, September 2002.

[17] Statistics gathered from NTSB web page, http://www.ntsb.gov/aviation includes parts 121 and 135, general aviation and military.

[18] Enge, P., Phelts, E., and Mitelman, A., "Detecting Anomalous Signals from GPS Satellites," Global Navigation Satellite System Panel meeting, Toulouse October 18-29 1999, working paper 19.

[19] Brenner, M., Kline, P., and Reuter. R., "Performance of a Prototype Local Area Augmentation System (LAAS) Ground Installation," in Proceedings of ION GPS-2002, Portland, OR, September 2002.