Overbounding The SBAS Integrity Equation

Pieter Bastiaan Ober, *Delft University of Technology*
Rick Farnworth, *Eurocontrol Experimental Centre*
Edward Breeuwer, *Eurocontrol Experimental Centre*
Durk van Willigen, *Reelektronika BV*

BIOGRAPHY

Pieter Bastiaan Ober's areas of experience include the influence of multipath on GPS positioning, carrier phase differential GPS, ambiguity resolution and integrity monitoring, especially for aviation applications. He is currently active as a Ph.D. student doing research on integrity design and analysis of integrated navigation systems for safety critical applications.

Edward Breeuwer obtained his MSc in Electrical Engineering from Delft University of Technology in 1992 and was awarded a PhD from the same university in 1998 for his work on integrated navigation systems. Since October 1997 he has been working as a consultant for EUROCONTROL at their Experimental Centre in the GNSS Programme Office, where he is involved in various activities focusing on the implementation of satellite navigation in civil aviation.

Rick Farnworth graduated with a BSc in Electronic Engineering from the University of Wales in 1988 and was awarded a PhD in 1992 for his work on LORAN-C coverage prediction modelling. He then joined the UK's CAA's National Air Traffic Services to work on R&D projects relating to the application of satellite navigation systems in civil aviation. Since February 1996 he has been working for EUROCONTROL at their experimental centre in the GNSS Project Office where he is responsible for various R&D projects related to satellite navigation.

Dr. Durk van Willigen retired as professor at Delft University in The Netherlands where he headed a research group in Electronic Navigation Systems from 1989-1999. He is president of Reelektronika b.v., a consultancy in navigation, and is involved in many studies on navigation for the European Commission and in the implementation of Eurofix on the European Loran-C stations. In 1996, the International Loran Association awarded him the Medal of Merit. He is fellow of the Royal Institute of Navigation since 1999, and is the recipient of the Thurlow Award of the United States' Institute of Navigation in 2000.

ABSTRACT

The paper shows how the general integrity concept of an SBAS differs from that of traditional ground-based navaids and shows that this implies a need for detailed knowledge of the tails of the error distributions. Because such detailed knowledge cannot be obtained from measurement data, the assumption of gaussian distributions cannot be validated by data alone. The search for a satisfying overbounding concept to enable the use of the SBAS integrity equation with non-gaussian errors has lead to some interesting results. However, as this paper shows, the validation problem has not been satisfactory solved with the current proposals.

1. INTRODUCTION

For certification of GNSS based navigation systems for aviation, it is necessary to guarantee that the user is informed on his position with sufficient integrity. The probability that the navigation system supplies so called hazardously misleading information should be proven to remain extremely small.

The small probabilities involved make it hard to formulate how system integrity can be validated. Theoretical studies are almost exclusively based on the assumption of (white) Gaussian distributions for the signal deviations. Although this leads to useful insights, the deviations in real life never completely follow these theoretical assumptions. This leads to the important question whether navigation systems based on these assumptions can be certified.

The problem of trying to guarantee that (differential) GNSS-based systems offer sufficient integrity is known as the 'overbounding problem' because practical solutions are necessarily conservative ('bounding'). Work in this area has mainly been performed for ground based augmentation systems. However, lately the need for a better understanding in the SBAS field has also become understood.

The paper first shows how the general concept of integrity monitoring for a differential GNSS system differs from that of traditional ground-based navaids. In particular, it is shown why a much better understanding of the error behaviour of the signals is necessary. The paper then discusses the SBAS integrity concept. It gives a structured overview of the overbounding problem and briefly reviews the status of the solutions that are currently being proposed and investigated and their merits and shortcomings.

2. MONITORING CONCEPTS

There are substantial differences between the monitoring of traditional ground-based radionavigation aids and differential GNSS. These differences have important consequences for the verification of their performance.

In ground based radio-navigation systems, monitoring has traditionally been performed by separate monitoring systems that check the correctness of the transmitted signals. Examples of such systems are numerous, and include DME, VOR, ILS and MLS. Whenever a large signal deviation is detected, the monitor will cause the signal transmission to be interrupted, thus providing integrity. Because the monitor position is known, the measured value can be checked and compared to the known value a perfect signal would have. In case of large deviations action can be taken, which usually means that the signal in space is completely removed to avoid use of the degraded signal. This implies that integrity is relatively easily established *without the necessity of detailed knowledge on the transmissions error distributions*: the transmitter is switched off whenever the sum of the transmission error and the monitor noise exceed a threshold value. There is no relation between the error distribution of the transmitted signals and system integrity – the only uncertainty in the assessment of the transmitted error lies in the measurement uncertainty that is introduced by the monitor. This uncertainty should be kept sufficiently small to keep the false alarm rate within the continuity requirements.

2.1 GNSS Monitoring

For a number of reasons, the traditional monitoring philosophy does not work very well with global systems such as the current Global Navigation Satellite Systems (GNSS's) .

First of all, the global coverage makes timely detection less likely, as a widespread monitoring system would be required to check signals permanently. Furthermore, to invalidate the satellite broadcast, the satellite needs to be in view of an uplink station. The lack of sufficiently dense monitoring and uplink networks make the response to failures of systems such as GPS too slow to rely on for safety critical operations. Another problem is the occurrence of local signal distortions (such as locally extreme atmospheric delays, multipath and interference) that can affect users while remaining undetected by the monitoring network. Finally, GNSS-s serve many different kinds of users that all have different requirements. Providing a signal with better integrity for safety-related use might adversely affect accuracy and availability and thus be sub-optimal for other users.

To overcome most of the problems mentioned above, satellite based augmentation systems such as WAAS and EGNOS are under development. Both include monitoring networks that are an order of magnitude denser than those of GPS or GLONASS, to guarantee signal integrity over large areas. This solves the 'signal in view' problem for the satellites. To cope with the uplink problem, both WAAS and EGNOS use a datalink that employs geostationary satellites that notify the user of failures; since both systems are civil, they cannot change the satellite broadcast of the military GPS.

2.2 Differential GNSS

With the whole infrastructure for it in place, it makes sense not only to detect large signal deviations and transmit integrity information. The monitoring network can also be used to estimate the 'normal' deviations in the GNSS signals (such as atmospheric delays, satellite clock and ephemeris errors) that correlate with the deviations the user will experience. This information can be supplied to the user in the form of differential corrections to increase accuracy.

In a differential system, it no longer suffices for the signal monitor to look at signal deviations as such. As some of the signal deviations will be compensated for by the differential corrections, a check on the ground has become much harder to perform. In some way or another, the ground system now has to distinguish between differentially 'correctable' and 'uncorrectable' deviations. Only in case of large 'uncorrectable' deviations, an integrity warning is to be issued to preserve as much availability as possible.

Practical systems, especially when they need to serve multiple classes of users such as an SBAS, only flag obviously erroneous measurements. For all other measurements, they compute a quality metric and send this along with the corrections. This enables the user to apply this information to decide whether his particular requirements are met by the system. This way of operation has two important implications:

- Detailed knowledge of the transmitted GNSS error is required to enable the user to compute the integrity of his position obtained with the system
- A failing monitoring infrastructure can comprise integrity when erroneous corrections or correction quality information is broadcast. To deal with receiver failures, a consistency check among multiple redundant receivers will be required.

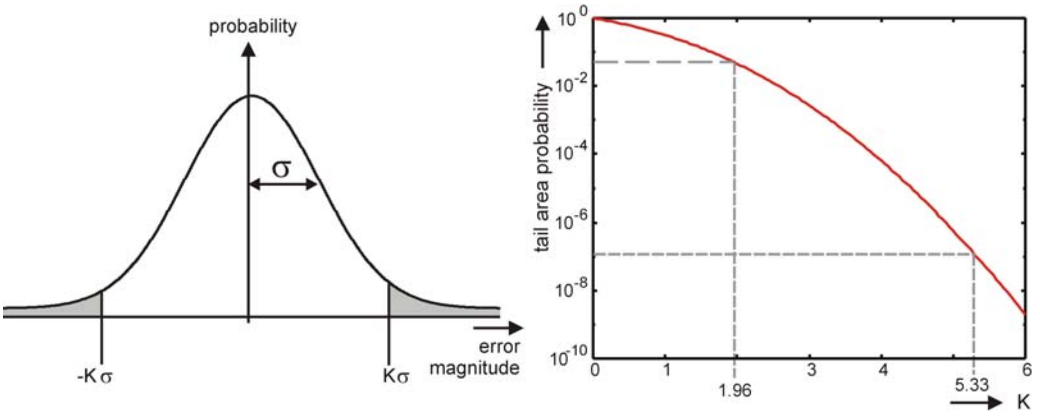

Figure 1. The zero mean gaussian distribution is characterised by a single parameter: the standard deviation σ (or its variance σ^2). It is therefore practical to express the location of special points on the distribution as a scale-factor (*K*) times σ. On the right hand side, the probability that the magnitude of the error exceeds K (grey area) is given as a function of the scale-factor *K*.

The focus in the remainder of this paper will lie on the first problem: the requirement to know the error distributions of the corrected GNSS ranges to enable the user to make an assessment of the integrity the system provides.

3. SBAS INFORMATION FLOW

In an SBAS environment, the user will contain three types of information:

- satellite geometry information
- ranging information
- measurement quality related information

The *satellite geometry information* comprises the ephemeris data of the ranging satellites from which the position of the satellites as a function of time can be derived, and contains the satellite ephemeris data and the corrections to the satellite positions. The *ranging information* consists of three different contributions: the ranges, the clock and ephemeris corrections and the ionospheric corrections. *Measurement quality information* is provided in the form of variances that are related to the two types of corrections: the UDRE for the ephemeris and clock corrections and the UIVE for the ionospheric corrections.

The measurement flows from the ranging sources and the SBAS messages are asynchronous. Therefore, the SBAS information needs to be synchronised with the range measurements. For the ionospheric correction, rather than a time-synchronisation, an ionospheric path correction is applied in which the received values are all transformed to relate to the user's position. For the scope of this paper, it will be assumed that these synchronisations have been performed and all parameter values relate to the same space-time point.

The user equipment translates the differentially corrected ranges into a position solution by a weighted least squares algorithm. To determine the optimal position solution, low noise measurements will be more heavily weighted than more noisy measurements. To determine the weights, the quality information that is provided by the UDRE and UIVE are used: high variances imply small weights to compensate for the high measurement uncertainty.

The UDRE and UIVE are combined with the local noise variance to obtain the variances of the differentially corrected ranges:

$$\sigma_i^2 = UDRE + UIVE + \sigma_{local}^2$$

in which:

σ_i^2: variance of the position error distribution
UDRE: variance of residual clock and ephemeris errors in corrected range to the i^{th} ranging satellite
UIVE: variance of residual ionospheric errors in corrected range to the i^{th} ranging satellite
σ_{local}^2: variance of the local receiver noise and multipath

This information is fed to the SBAS integrity equation that will be discussed in the next section. Note that no ranging information whatsoever is used to assess the integrity of the system. The integrity of the system does not depend on the information of these sources.

4. THE SBAS INTEGRITY EQUATION

The SBAS integrity equation [Walter97] describes the position error distribution that is obtained by using differentially corrected measurements of which the validity has been checked by the ground network in the absence of failures that are local to the user. It gives the user a simple means to assess the externally provided accuracy and integrity performance. The equation is of the following form:

$$e_{pos} \sim N(0, \sigma_{pos}^2)$$

which means that the positioning error e_{pos} has a gaussian distribution with mean 0 and variance σ_{pos}^2.

The variance is a function of:

- geometry
- variances σ_i^2 of the corrected ranges

A zero mean gaussian distribution is fully characterised by its variance or the square root of the variance, the standard deviation. Because of that, specific points on the distribution are often represented as a multiple of this standard deviation. The relationship between this multiple and the energy (probability content) in the tails of the distribution beyond this multiple are shown in Figure 1. In the figure, two such special points are indicated: the one that specifies the accuracy, and the one that specifies the protection level that serves as a measure of integrity.

For any phase of flight, [SARPS] specifies the required accuracy at the 95% level, which corresponds to a total tail energy of 5% and a *K*-factor of 1.96:

$$\text{Accuracy} = 1.96 \cdot \sigma$$

Integrity is specified in terms of the protection level, which is related to the probability that the alert limit may be exceeded. The exact probability – and therewith the K-factor- depends on the phase of flight. In Figure 1, the en-route value of 10^{-7} has been selected, corresponding to a *K*-factor of 5.33:

$$\text{En route protection level} = 5.33 \cdot \sigma$$

All considerations above concern a one-dimensional position error distribution. Because the horizontal and vertical navigation channels are considered separately, the vertical position is indeed one-dimensional. However, the horizontal position is two-dimensional. To avoid further complication, the following conservative approach is taken [SARPS]. In the horizontal plane, the direction in which the error distribution has the largest variance is used to represent the horizontal position error distribution.

5. SBAS INTEGRITY VERIFICATION

5.1 Gaussian assumption

It has been a long tradition in the design of, research on, and analysis of navigation systems that error sources are assumed to have a zero mean gaussian probability density function (in the absence of failures). This assumption has been instrumental in the development of real-life systems as it allows for a relatively simple metric for trade-off purposes in system designs, and has been the basis for the design of (almost?) all algorithms in GNSS-based systems, including SBAS and GBAS, including the SBAS integrity equation.

Despite the promises of the central limit theorem, not all error sources in GNSS follow a gaussian distribution. Multipath, which is one of the dominant error sources in differential GNSS systems, is a good example of a non-gaussian source. The multipath error has clearly defined minimum and or maximal values that are (analytically) determined by the physics of the system.

Moreover, error sources are not always zero mean, especially not when observed over a relatively short period of time. Again, multipath is a good example. Because requirements are often specified per operation, and averaging over multiple operations is not allowed, this implies that even in the error-free situation the measurement deviations are not free from biases.

5.2 Use of measurement data for verification

Although measurements will always be an important element of any system performance verification, it has been realised that the use of measurements alone will never provide significant amounts of knowledge on the tails of the error distributions up to the 10^{-7} level that would be required. It is demonstrated in [Shively00] that in order to have a 50% chance of even observing a single measurement to exceed the alert limit, about 10^7 independent samples will be required. About $4.5 \cdot 10^7$ samples would be necessary to make this a 99% chance.

As Shively shows for an GBAS system, gathering large numbers of independent samples from the error distributions is limited by a number of factors such as the use of carrier-smoothing, repeatability of multipath errors and the dependence of the ranging performance on satellite elevation. For a typical GBAS, Shively estimates that per receiver a total number of only between 800 and 8000 range measurement error samples will be available per year.

Although an SBAS system might differ substantially from a GBAS system as far as the number of independent samples that can be obtained is concerned, the amount of independent samples will surely fall orders of magnitudes short of the required numbers to use them to fully validate the integrity performance. As Shively puts it referring to the LAAS system: "*it is obviously not practical to directly characterize the tails of the LAAS error distribution in the region of interest for the vertical protection level*".

5.3 Simulations and models

Even when the error distributions would be approximately gaussian, the tails of the distribution can never be verified to be gaussian by the use of measurement data alone. because measurements alone provide insufficient data to back-up a claim on system conformance to its integrity requirements, simulations and models will be needed as additional tools in the system certification procedure.

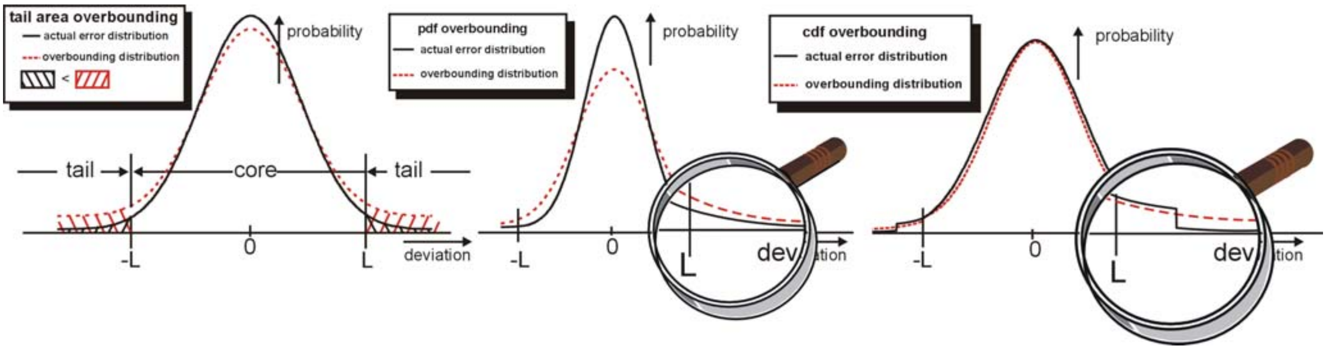

Figure 2. The three different ways to overbound an actual (non gaussian) error distribution by a gaussian modelled ‘overbounding’ distribution

In the remainder of this paper, a particular model based strategy is discussed that has received ample attention over the last couple of years. The work that has been performed on the so-called ‘overbounding problem’ can be seen in the context of the previous section as a sensitivity study of (mainly) the GBAS integrity equation for the gaussian assumption: when the tails of the distribution cannot be proven to obey this assumption, it is desirable to understand what the consequences would be when the range error distributions are not exactly gaussian, and under which conditions the SARPS integrity equations can still be used.

6. OVERBOUNDING

The SBAS integrity equation is build upon the assumption of zero mean gaussian corrected range error distributions. When the actual distribution is not (or can not proven to be) zero mean gaussian, the question is whether the SARPS integrity equation can still be used. The answer is obvious: yes, it can be used as long as this does not cause the operational requirements to be violated. In other words: the integrity equation can be used, as long as its guaranteed not to give an overly optimistic assessment of system integrity. This leads to the following questions:

- for which standard deviation does a gaussian distribution ‘overbound’ the actual position error distribution in the sense that use of this gaussian in the integrity equation is conservative
- which standard deviation should be broadcasted for the range error distributions in order to provide the user with an overbounding gaussian for his position distribution

Three different answers are readily provided to the first of these questions. Although the first answer (tail area overbounding) is a straightforward implementation of the requirement in the position domain, it leads to no solution for the second question. Therefore, other, more restrictive answers have been thought in an attempt to bridge the gap between range and position requirements. However, only the third solution has lead to the providence of some guidance of solving this second problem, when the real range error distributions obey certain constraints.

6.1 Tail area overbounding in the position domain

The implication for the relation between the overbounding gaussian position error distribution that is provided by the SARPS integrity equation with the broadcast standard deviation and the actual position error distribution is depicted in Figure 2 (left). The position error distribution is divided in two regions, which will be called the core and the tails of the distribution. The start of the tail area is indicated with the letter *L*. For tail area overbounding, it is required that the tail area of the overbounding distribution contains more energy (likelihood) than the tail area of the actual distribution, implying that the probability that the alert limit is exceeded in reality will be smaller than the SARPS integrity equation predicts. When the standard deviation of the overbounding distribution is sufficiently large, it will always be possible to overbound the actual distribution in this sense.

Tail area overbounding is practical in the position domain and gives an exact answer to the question whether the SBAS integrity equation is conservative or not. Unfortunately, it turned out not to be a practical condition when the relation between the overbound in the position domain is to be related to an overbounding condition on the signal in space, that is, in the range domain. Hence, two sufficient overbounding alternatives were considered that are more restrictive – and thus might lead to more conservative standard deviations for the overbounding distribution - but also more practical. They will be called *pdf overbounding* and *cdf overbounding* after [DeCleene00].

6.2 Pdf overbounding

For pdf overbounding, one distribution is said to overbound the actual error distribution when its tail values are consistently larger than the tail of the actual distribution, as depicted in Figure 2 (middle).

Because the total area under any distribution curve should necessarily equal one, a distribution that overbounds in the tail necessarily underbounds elsewhere. The overbounding can only be valid in the tails of the distribution. To link an overbounding in the position domain to an overbound in the measurement domain proved to be very hard. [DeCleene00] reports that numerical analysis indicated that the tail probability in the position domain depends heavily on the core of the distribution in the measurement domain. This makes it impossible to find a general relationship between L-values in the measurement domain and the L-value (that equals the alert limit) in the position domain. It therefore proved impossible to characterise a general error distribution by a simple overbounding function in combination with an overbounding parameter L as the value of L is distribution dependent.

6.3 Cdf overbounding

The cumulative distribution function (cdf) represents the energy in the tail of the distribution rather than the probability density. The cdf(x) equals the area under the distribution from $-\infty$ to x. The tail area overbound discussed above is a bound that can be written in terms of the cdf as:

$$cdf_{overbounding}(-L) > cdf_{actual}(-L) \quad \text{(left tail)}$$
$$1-cdf_{overbounding}(-L) > 1-cdf_{actual}(-L) \quad \text{(right tail)}$$

However, for cdf overbounding in the sense DeCleene has used it, this conditions above should be not just be obeyed for a single value L, but for all positive L. Hence, cdf overbounding does not rely on the existence of a special point that separates the tails from the core: it always overbounds the tail whatever starting point is chosen.

An attempt to depict the concept of cdf-overbounding has been made in Figure 2 (right). In the tail area of this figure, it can be seen that the overbounding pdf not everywhere exceeds the actual one. Cdf overbounding is therefore less restrictive than pdf-overbounding, as pdf overbounding is a sufficient, but not a necessary, condition for cdf overbounding. Conceptually, cdf-overbounding exploits the fact that it is always advantageous to make errors smaller than the assumed model, even if that implies that a particular error value becomes more likely.

Recently, it has been proven in [DeCleene00] that when the real error distribution of the range error measurements is zero mean, symmetric and strictly unimodal (having only one single local maximum value), the position distribution is overbounded by the gaussian distribution of the SBAS integrity relation when the overbounding corrected range error distributions are used.

The importance of this proof lies in the fact that a condition has been found that is less restrictive than the gaussian assumption, but still enables the SBAS provider to provide values for the UDRE and UIVE that guarantee integrity at the user to be preserved. Although this condition is sufficient, they might not be necessary, and the search for less restrictive conditions is continued.

Unfortunately, the validation assumption is not really solved by as the following problems still remain:

- conditions of symmetry and strict unimodality will be difficult to assure, especially for the tails of the distributions
- actual distributions might not be zero mean

One of the possible ways to take a non-zero mean into account is by inflating the standard deviation of the overbounding distribution to compensate for this. International research is still ongoing on the exact amount of inflation that would be needed.

7. CONCLUDING REMARKS

Detailed knowledge on the tails of the error distributions cannot be obtained from measurement data. The assumption of gaussian noise on which the SBAS integrity equation is based can therefore not be proven.

The work on the overbounding problem has shown considerable progression with the concept of cdf overbounding. Conditions for the validity of the equation have been found that are less restrictive than the gaussian assumption. Still, the problem with (data based) verification has not been solved completely satisfactory with the current proposals, as detailed knowledge on the tails of the error distribution remains required.

ACKNOWLEDGEMENTS

This work is partly based on a study performed by Reelektronika b.v. for Eurocontrol.

DISCLAIMER

The views expressed in this article are those of the authors and do not reflect the official policy or position of Eurocontrol.

REFERENCES

[DeCleene00] DeCleene, B.; “Defining Pseudorange Integrity-Overbounding”, Proceedings of ION-GPS, Salt Lake City, Utah, U.S.A., September 2000

[SARPS] AN-WP/7556 Addendum No.1 Appendix, Draft Amendment To Annex 10 To The Convention On International Civil Aviation, Volume I – Radio Navigation Aids, International Standards And Recommended Practices, Aeronautical Telecommunications

[Shively00] Shively, C.A. and R. Braff; “An Overbound Concept For Pseudorange Error From The LAAS Ground Facility”, Proceedings of the IAIN World Congress in association with the U.S. ION Annual Meeting, San Diego, CA, U.S.A., 26-28 June 2000

[Walter97] Walter, T., P. Enge and A. Hansen; “A Proposed Integrity Equation for WAAS MOPS”, Proceedings of ION-GPS Meeting 1997