THE JOURNAL OF NAVIGATION (2005), **58**, 83–103. © The Royal Institute of Navigation
DOI: 10.1017/S0373463304003029 Printed in the United Kingdom

A Method of Over Bounding Ground Based Augmentation System (GBAS) Heavy Tail Error Distributions

Ronald Braff and Curtis Shively

(*Center for Advanced Aviation System Development at The MITRE Corporation*)
(Email: rbraff@cox.net)

The purpose of this paper is to describe a statistical method for modelling and accounting for the heavy tail fault-free error distributions that have been encountered in the Local Area Augmentation System (LAAS), the FAA's version of a ground-based augmentation system (GBAS) for GPS. The method uses the Normal Inverse Gaussian (NIG) family of distributions to describe a heaviest tail distribution, and to select a suitable NIG family member as a model distribution based upon a statistical observability criterion applied to the FAA's LAAS prototype error data. Since the independent sample size of the data is limited to several thousand and the tail probability of interest is of the order of 10^{-9}, there is a chance of mismodelling. A position domain monitor (PDM) is shown to provide significant mitigation of mismodelling, even for the heaviest tail that could be encountered, if it can meet certain stringent accuracy and threshold requirements. Aside from its application to GBAS, this paper should be of general interest because it describes a different approach to navigation error modelling and introduces the application of the NIG distribution to navigation error analysis.

KEY WORDS

1. GBAS. 2. LAAS. 3. Tail error distributions. 4. Normal Inverse Gaussian.

1. INTRODUCTION. To meet the stringent accuracy and signal integrity requirements for precision approach and landing, GPS requires a local ground-based augmentation system (GBAS). A GBAS enhances accuracy by providing a single correction to each pseudorange measurement that accounts for common errors, such as incorrect satellite ephemeris and clock bias, and unknown propagation delays due to the ionosphere and troposphere. It ensures the integrity of the satellite signals by employing monitors to detect errors due to satellite failures (e.g., distorted signal waveform) as well as ground facility errors. A top-level description of the FAA's GBAS, the Local Area Augmentation System (LAAS) Ground Facility (LGF), can be found in (Braff, 1997–98). Detailed requirements of the LAAS Category I (CAT I) airborne equipment are given in (RTCA, 2000). The CAT I LGF has already been specified (FAA, 2002). The CAT II/III LAAS requirements are under development. This paper is concerned with an aspect of integrity that involves the LGF's contribution to the fault-free errors in the pseudorange corrections, and

the illustrative application is to CAT I. The characterization of these errors is problematic because their tail probability has to be bounded by order 10^{-9} (CAT I) and 10^{-11} (CAT II/III), and it is obviously impossible to acquire a sufficient independent sample size to make such estimates from data.

1.1. *Over bounding of fault-free errors.* The fault-free errors are due to satellite signal multipath and receiver/antenna noise. From analysis of data, the tail probability of these errors even at the $\pm 3\sigma$ points of their distribution is sometimes found to be significantly heavier than that of the Gaussian distribution. The LGF addresses this heavy-tail problem by multiplying the root-mean-square (RMS) error of each correction by an "inflation factor" greater than 1 to establish the broadcast estimate of the RMS error (called σ_{pr_gnd}) such that the probability of exceeding the lateral (LPL) and vertical (VPL) protection levels is smaller than or equal to an allocated requirement. This process is called "error over bounding." The σ_{pr_gnd} value for each of the corrections is transmitted to the aircraft where they are the inputs to standard equations that calculate VPL and LPL for the fault-free errors (RTCA, 2000). The VPL and LPL are then compared to corresponding alert limits (VAL and LAL). An alert is given if any of these protection levels exceed their corresponding alert limits. The fault-free integrity check in the aircraft has the effect of ruling out poor satellite geometries that significantly amplify the fault-free fault errors in the position domain (dilution-of-precision effect).

A procedure for establishing the inflation factor, I_f, for over bounding is to determine its solution in the following equation

$$\text{Prob}\{R > K_{\sigma_{pr\_gnd}}\} = 2 \int_{-\infty}^{-K I_f \sigma} f_r(r)dr = 2\Phi(-K) \tag{1}$$

f_r: assumed zero mean and symmetric probability density function (PDF) of the correction errors
K: required multiple of sigma, K = 5·8 for CAT I GBAS to achieve probability $6{\cdot}67 \times 10^{-9}$
Φ: Gaussian cumulative distribution function (CDF) with zero mean and unit variance
I_f: Inflation factor that causes the tail probability to equal $2\Phi(-K)$
σ: RMS error of the corrections, $\sigma_{pr\_gnd} = I_f \sigma$

It is seen that the effect of the inflation factor is to increase the tail points to $\pm K I_f \sigma$ so that the required Gaussian tail probability for K is achieved at these points.

The over bounding challenge for heavy tails is to find a suitable PDF (f_r) in (1) that can be used in the derivation of the inflation factor. The FAA has been sponsoring research on two approaches to derive such a PDF. The preferred approach is to base the PDF on physical principles. The other one is to find a method to base the PDF on statistical principles. The latter approach, the subject of this paper, serves as a substitute in the event that the more difficult physical method is not developed in time when LAAS is first implemented.

1.2. *Physically derived fault-free error model.* From the physics of nominal multipath and reference antenna gain pattern, the multipath contaminated signal at the

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

antenna output is derived. The multipath model could represent the topology of an actual site. This signal model is then the input to a model of receiver signal processing including receiver noise. The output is a time series of the navigation sensor error due to multipath and receiver noise. This time series is then sampled and processed to form an error distribution. Errors from diffuse and specular multipath are being considered. So far, it has been found difficult to derive an acceptable error distribution by this method. References (Brenner, 1998) and (Pervan, 2000) contain early discussions of this type of modelling for LAAS.

1.3. *Statistically derived fault-free error model.* From the observation of large error behaviour, a heavy tail error distribution is estimated using a statistical method. The method has to be conservative due to insufficient data. The conservatism of the estimation method leads to an inflation factor that is expected to be larger than one that would be derived from a physical model. If the inflation factor is too large, system availability could be significantly decreased.

There has been interest in estimating heavy-tail distributions for navigation errors. For example, in Campos, 2002 a family of exponential distributions of the form, $F_b(x) = A \exp(-a|x|^b)$, $0 < b \leqslant 2$, is introduced. In this expression $b = 2$ and $b = 1$ correspond to the Gaussian and Laplacian distributions, respectively. It was found in Campos, 2002 that the distribution with $b = 0{\cdot}5$ provides the best fit to an aircraft altitude distribution. Years ago, it was shown that the Laplacian distribution could have a physical interpretation as an approximation of a mixture of Gaussian distributions of different standard deviations due to the pooling of error data from navigation systems having different accuracies (Parker, 1966). If the distribution of standard deviation values is uniformly distributed the error distribution approaches the Laplacian. Gaussian mixtures are a general class of heavy-tail distributions that have been studied in detail (Cornell, 1990).

The approach taken in this paper does not postulate an error distribution that is formed by matching it to the data's distribution. Rather it starts with the recognition that the tail probability is mathematically limited. This limitation is illustrated by the well-known Chebyshev inequality when the mean is zero, $\text{Prob}\{|x| > K\} \leqslant K^{-2}$, where K is the standard deviation multiplier. If certain reasonable constraints are placed on the probability density function (PDF), it could be possible to further decrease this upper bound tail probability. The constraints on the PDF are based on characteristics of the Gaussian distribution. These characteristics are symmetry about the central point, strictly decreasing, greater than zero at all points, and differentiable at all points, including the central point. To date, our organization's research has derived an upper bound probability for only the constraints of symmetry and non-increasing PDF. Its form is a Dirac delta function located at the central point combined with a uniform distribution (Marshall, 2003). The appendix describes this PDF. Although this distribution is a mathematical construct and is not even physically realizable, it does provide an upper bound that is somewhat smaller than the Chebyshev. The resulting inequality is $\text{Prob}\{|x| > K\} \leqslant (4/9)\, K^{-2}$.

As stated previously, it is impossible to attain a sufficient independent sample size to estimate the actual tail probability. Therefore, the $(4/9)\, K^{-2}$ upper bound is used as a benchmark of the conservatism of the tail probability derived from any candidate error distribution that is based on observation of the existing data. For example, considering the family of exponential distributions in Campos, 2002, the PDF has an infinite derivative at the central point when $b < 1$ so it has no resemblance to the

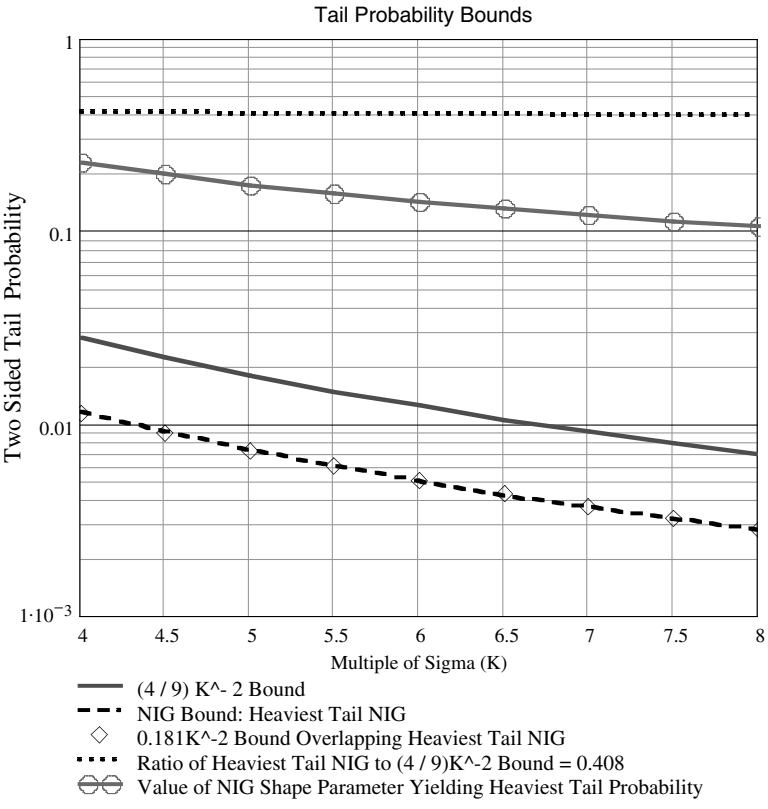


Figure 1. Tail probability upper bounds.

observed LAAS correction error distributions which are not overly steep at the central point. Therefore, by our adopted constraints, that family of distributions has to be restricted to be in the interval 1 (Laplacian) $\leqslant b \leqslant 2$ (Gaussian), and the Laplacian provides the heaviest tail probability. The Laplacian tail probability at the CAT I $\pm 5{\cdot}8\sigma$ points ($2{\cdot}74 \times 10^{-4}$) is about 50 times smaller than the (4/9) K^{-2} tail probability bound (0·0132).

The Normal Inverse Gaussian (NIG) (Hanssen, 2002) is a family of distributions that match the four aforementioned constraints and provide a wide variety of distributions. As will be shown in the next section, a modification of the standard form of the NIG produces a family of distributions containing a heaviest tail probability. This heaviest tail probability is closely approximated by $0{\cdot}181K^{-2}$ for all K, which is about 0·4 of the (4/9) K^{-2} upper bound. Therefore, a conjecture is made that if the theoretical upper-bound distribution were further constrained to be strictly decreasing, non-truncated and differentiable at all points then its tail probability would be even closer to the NIG heaviest tail distribution. Therefore, the NIG heaviest tail distribution is assumed to be a reasonable estimate of the least upper bound tail probability.

Figure 1 contains plots of the heaviest tail NIG and $(4/9)K^{-2}$ bounds that illustrate the foregoing discussion. Note that the ratio of the two bounds is approximately constant for all values of K. Included in the figure is the value of the shape parameter

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

of the NIG that yields the maximum tail probability for each value of K. Note, that this parameter does not vary much with change in the K value.

The inflation factor for the heaviest tail NIG for CAT I (K = 5·8) = 12·9 is much too large to be practical. Therefore, the scope of this paper is to use a statistical method, based on the observability of large errors, to derive a practical inflation factor. A member of the family of NIG distributions provides the model for the estimation, and the heaviest tail probability member is used to analyze a worst-case mismodelling of the distribution.

1.4. *Contents of paper*. The contents of the paper and their order are as follows:

- Description of NIG family of distributions, and its extension to the distribution of the average error from M reference receivers to gauge tail thinning from the M – 1 convolutions
- Definition of a large error observability criterion that is used as the rationale for the selection of the NIG family member for inflation factor estimation
- Comparison of selected NIG family member to error data collected at the FAA's LAAS Test Facility (LTF)
- Transformation of N range domain error distributions from N satellites to the position domain to gauge tail thinning from an additional N – 1 convolutions
- Derivation of inflation factor
- Description and analytical evaluation of a Position Domain Monitor (PDM) that could be used to provide a real-time check on any mismodelling
- Summary and discussion

2. NORMAL INVERSE GAUSSIAN DISTRIBUTION.

The general form of the NIG contains four parameters, one to account for symmetry, one for a bias and two for shape (Hanssen, 2002). Setting the former two parameters to zero yields a symmetric, strictly decreasing, differentiable and zero mean PDF given by

$$f(r,\alpha,\delta)=\frac{\alpha\delta}{\pi}\,\frac{exp(\alpha\delta)}{\sqrt{r^2+\delta^2}}\,K_1\left(\alpha\sqrt{r^2+\delta^2}\right)$$
$$Variance:\ \sigma^2=\frac{\delta}{\alpha} \tag{2}$$

r: fault-free errors in the pseudorange corrections
α, δ: distribution shape parameters with dimensions length^{-1} and length, respectively
K_1: modified Bessel function of the second kind, degree 1

The Fourier transform of the PDF in (2) (Hanssen, 2002) is

$$\Phi(\omega,\alpha,\delta)=\exp(\alpha\delta)\exp(-\delta\sqrt{\alpha^2+\omega^2}) \tag{3}$$

To derive the PDF of the errors of the average of the pseudorange corrections from M reference receivers, M – 1 convolutions of (2) are needed. Such convolutions can be accomplished by multiplying Fourier transforms and then taking the inverse transform. The multiplication is facilitated by the exponential forms in (3). The resulting

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

Fourier transform is given by

$$\Phi(\omega, \alpha, \delta)_M = \exp((M\alpha)\delta) \exp\left(-\delta\sqrt{(M\alpha)^2 + \omega^2}\right) \quad (4)$$

The PDF in (2) is modified as follows to ensure that the variance remains constant as the shape parameters are varied (Braff, 2003). Let σ_0 be the RMS of the PRC errors from one reference receiver. A relationship for any σ_0 is established as $\alpha = \delta_0/\sigma_0$, $\delta = \sigma_0\delta_0$, where δ_0 is a dimensionless parameter that establishes the shape of the distribution and the heaviness of its tails. The inverse Fourier transform is derived by substituting these values into (4) and noting the relationship between (2) and (3). The resulting PDF for the errors of the average of M corrections is then given by

$$f_{avg}(r, \delta_0, M) = \frac{M\delta_0^2}{\pi} \frac{\exp(M\delta_0^2)}{\sqrt{r^2 + \sigma_0^2\delta_0^2}} K_1\left(\frac{M\delta_0}{\sigma_0}\sqrt{r^2 + \sigma_0^2\delta_0^2}\right) \quad (5)$$

Equation (5) is the working equation that will be used in the analyses.

The two-sided tail probability is defined as

$$\text{Prob}\{R > K\sigma\} = 2\int_{K\sigma}^{\infty} f_{avg}(r, \delta_0, M)dr \quad (6)$$

The FAA's GBAS, Local Area Augmentation System (LAAS), nominally broadcasts the average correction from $M = 4$ reference receivers. The tail probabilities for $M = 1$ (single reference receiver correction) and $M = 4$ are plotted in Figure 2 for $K = 5{\cdot}8$, the value used in the CAT I protection level equations.

In Figure 2, the NIG tail probability for $M = 1$ increases with decrease in shape parameter (δ_0), until it reaches a maximum at $\delta_0 = 0{\cdot}148$. It has the same tail probability as the Laplacian distribution when δ_0 is ≈ 1. As δ_0 becomes large, the tail probability approaches that of the Gaussian. The $M = 4$ plot illustrates how averaging the corrections thins the tails of the distribution with respect to multiples of $\sigma_0/\sqrt{M}$ except when $\delta_0 < 0{\cdot}1$. Note that for all δ_0, the $M = 4$ tails are thinner than the $M = 1$ tails when the tail points are defined by equal magnitudes rather than as multiples of sigma.

3. RANGE DOMAIN OBSERVABILITY OF LAAS CORRECTION ERRORS.

In the collection and analysis of LAAS correction error data it has been found that, due to independent sample size limitation, the tail probability can only be observed out to the $\pm 3 - \pm 4\ \sigma_{sample}$ points of the distribution. This limitation is due to the sorting of data into satellite-elevation-angle bins, temporal correlation, and the diurnal repeatability of multipath error over at least several days. Even if the core of the distribution up to $\pm 4\ \sigma_{sample}$ tested to be Gaussian, it would be risky to assume that further out the distribution remained Gaussian or followed some other distribution unless there were a physical basis for this assumption. For CAT I, a model of the tail probability up to the $\pm 5{\cdot}8\ \sigma_{sample}$ points and a means for accounting for any mismodelling are needed.

The proposed method addresses this problem by selecting a mathematical distribution (e.g., an NIG family member) such that if its tails were not heavy enough to

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

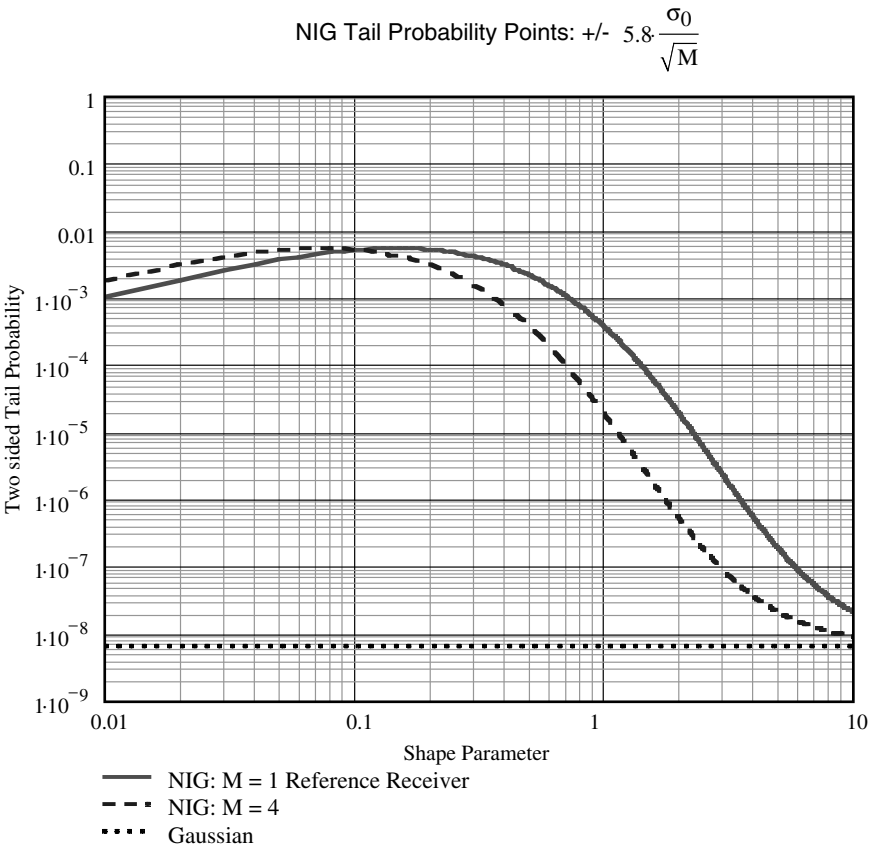


Figure 2. NIG tail probabilities as a function of the shape parameter.

bound those of the true distribution, then observation of the data would reveal, with high probability, at least one error value that is greater than a selected value. This technique is expected to produce an inflation factor that is larger than that produced by a physical-based model, but still contains some uncertainty due to relatively small sample size.

To select an appropriate distribution, observability as applied here needs to be defined. It is defined as the probability of observing at least one error beyond the $\pm K\sigma_{sample}$ points when the independent sample size is n. It is quantified as

$$P_{obs}(K) = 1 - (P_0(K))^n \tag{7}$$

$P_{obs}(K)$: probability of observing at least one error beyond $\pm K\sigma_{sample}$ points
$P_0(K)$: probability not observing an error beyond $\pm K\sigma_{sample}$ points per single observation for the selected distribution

$$P_0(K) = 2 \int_0^{K_{\sigma_{sample}}} f_{avg}(r, \delta_0, M)dr \tag{8}$$

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

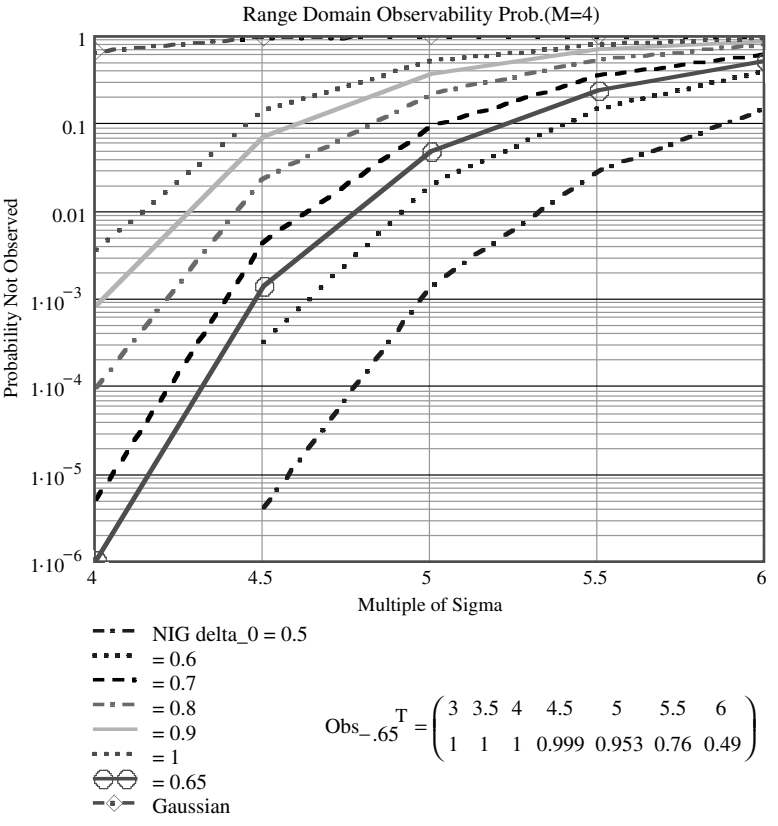


Figure 3. Unobservability of large errors (sample size = 6699, average of four reference receivers).

No errors were observed beyond $\pm 4\ \sigma_{sample}$ in the FAA data collection. For this analysis, the selected criterion for observability is $P_{obs}(4{\cdot}5) = 0{\cdot}999$. This means that it is near certain that the selected distribution at least bounds the tail at the $\pm 4{\cdot}5\sigma_{sample}$ points. The $\pm 4{\cdot}5\sigma_{sample}$ points rather than the $4{\cdot}0\sigma_{sample}$ points are selected to provide a more conservative distribution as will be discussed below. Dealing with covering lack of observability up to the $\pm 5{\cdot}8\sigma$ points is also discussed below.

3.1. *Selection of NIG family member.* The next step is to determine which NIG family member satisfies the $P_{obs}(4{\cdot}5) = 0{\cdot}999$ observability criterion. Figure 3 contains plots of the probability of not observing at least one value (1 − observability) beyond a given multiple (K) of σ_{sample} for an independent sample size of 6699. This sample size corresponds to the two heaviest-tail bins in the FAA's LAAS Test Prototype (LTP) data collection campaign (Warburton, 2002). The plots are for the average of 4 reference receivers which generated the LTP data, and is the nominal LAAS configuration.

In Figure 3, the top-most curve represents the Gaussian distribution. Its probability of not being observed beyond $4{\cdot}5\ \sigma_{sample}$ is practically 1. This implies that any distribution that is close to Gaussian would most likely be unobservable in the tail. Therefore, a Gaussian or near-Gaussian assumption would need to be based on physical reasoning.

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

Table 1. Observability of Selected NIG ($\delta_0 = 0{\cdot}65$).

Tail Points (±)	3·0	3·5	4·0	4·5	5·0	5·5	6·0
Observability	1	1	1	0·999	0·953	0·76	0·49

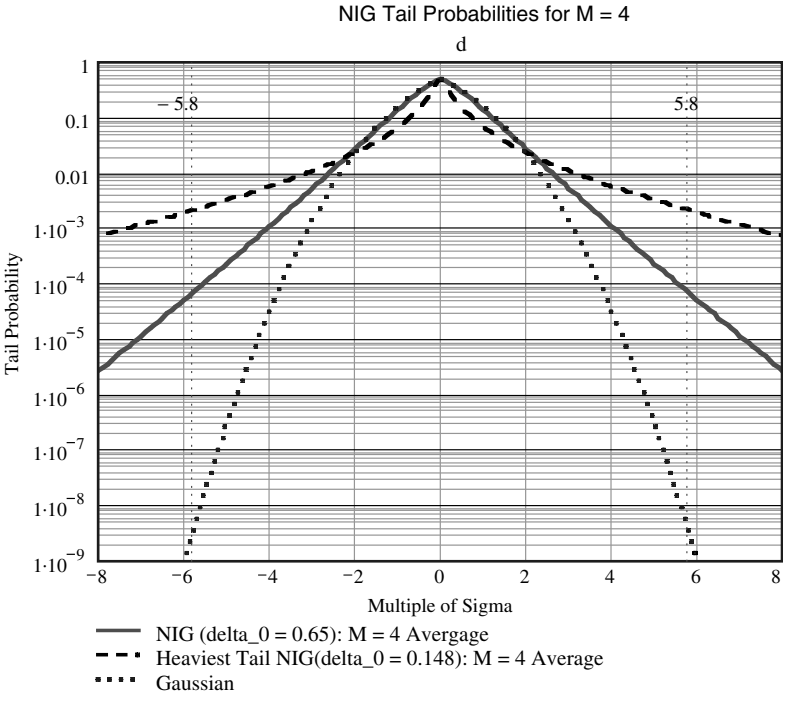


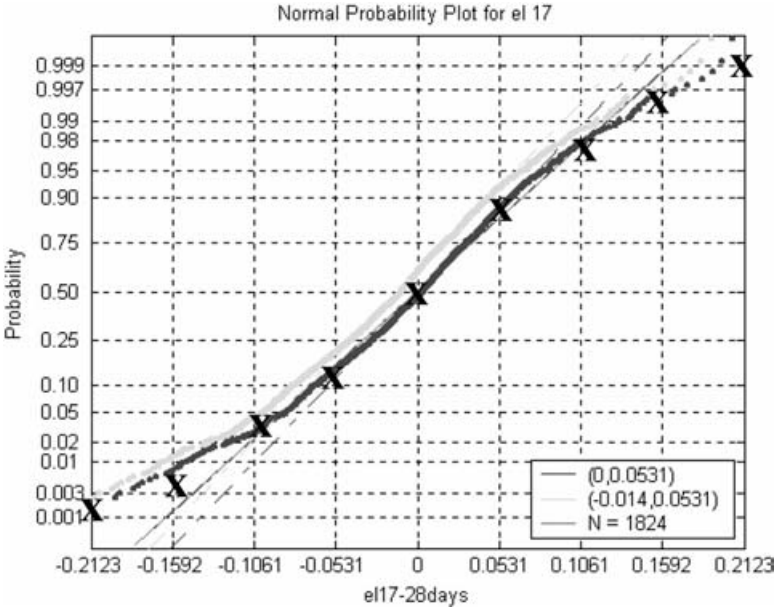
Figure 4. Comparison of tail probabilities.

The remaining curves are for a range of NIG family members. The NIG for $\delta_0 = 0{\cdot}65$ meets the observability criterion. Its observability for various tail points is given in Table 1.

Referring to Figure 3, if the $\pm 4{\cdot}0\sigma_{sample}$ points had been selected then the NIG for $\delta_0 \approx 0{\cdot}9$ would have been selected. That distribution is less conservative in that it has thinner tails than the $\delta_0 = 0{\cdot}65$ distribution.

To provide a perspective of how conservative the NIG ($\delta_0 = 0{\cdot}65$) tail probability is in general beyond the $\pm 4{\cdot}5\ \sigma_{sample}$ points, it is compared in Figure 4 to the heaviest tail NIG and the Gaussian tail. Recall in the introduction that an argument is made for considering the heaviest tail NIG to be close to the mathematical upper bound when reasonable constraints on the PDF were applied. Referring to Figure 4, the tail probability (sum of both sides) at the $\pm 5{\cdot}8\ \sigma_{sample}$ points is $1{\cdot}4 \times 10^{-4}$ ($\delta_0 = 0{\cdot}65$) and $4{\cdot}2 \times 10^{-3}$ ($\delta_0 = 0{\cdot}148$), and their ratio is 0·033. The ratio indicates that, at worst, there is a mathematical potential for two orders-of-magnitude of mismodelling in the tail probability. This should be tempered by the common sense observation that the model tail probability at the $\pm 5{\cdot}8\ \sigma$ points is of the order of 10^{-4}, which is quite heavy, being about five orders-of-magnitude heavier than the Gaussian tail. However, if the tail probability uncertainty is still not acceptable for CAT I application

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press



NIG CDF ($\delta_0 = 0.65$, M = 4):

K = −4	−3	−2	−1	0	1	2	3	4
0.0011	0.0052	0.026	0.134	0.5	0.866	0.974	0.995	0.999

Figure 5. Comparison of FAA data CDF from elevation bin 17 to NIG (Courtesy FAA William J. Hughes Technical Center (Warburton, 2002)).

after the inflation factor is applied, an independent monitor could be employed as will be discussed below.

3.2. *Correspondence to data.* It is noted that the analysis up to this point has not referred to any error distribution derived from data except for the observability of a large error with respect to sample size. Figure 5 contains the CDF derived from one of the elevation bins having the heaviest $\pm 3 - \pm 4$ σ_{sample} tail probabilities (Warburton, 2002). Each elevation bin was sampled every two weeks for nine months. The comparisons are with the CDF where the mean has been removed (dark solid line, "0, 0·0531"). The NIG values are displayed as "Xs" superimposed on the figure. Note that the NIG ($\delta_0 = 0{\cdot}65$) is a good representation of the data. Although it should be a necessary condition that the selected distribution should resemble the data, since there is no known physical reason for the resemblance, no explicit credit is taken for it.

4. DETERMINING INFLATION FACTOR TO ACHIEVE GAUSSIAN OVER BOUND OF TAIL PROBABILITY.

In this section, the inflation factor, I_f, implicitly defined in (1), that causes the two-sided tail probability at the $\pm 5{\cdot}8\sigma_{pr\_gnd}$ points to be equal $2\Phi(-5{\cdot}8) = 6{\cdot}63 \times 10^{-9}$ is derived. It is first derived in the range domain for the distribution representing the average of

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

$M=4$ corrections. In DeCleene, 2000 it is proven that if the range domain distribution is over bounded then the resulting position domain distribution is also over bounded. The average correction distribution is over bounded rather than the individual reference receiver corrections because the average correction has thinner tails due to the $M-1$ convolutions involved in the transformation from a single reference receiver error distribution to the distribution of the average of M receivers. Further reduction of the error distribution tails occurs during the transformation to the position domain. Therefore, a general method for position domain error over bounding is also discussed.

4.1. *Over bounding of average correction.* The average correction inflation factor is calculated implicitly by substituting from (5), $f_{avg}(r, 0{\cdot}65, 4)$, into (1). The result is $I_f=2{\cdot}2$.

4.2. *Over bounding in the position domain.* For efficiency of exposition, only the vertical position error is considered. The same analysis applies to the lateral error. The vertical position RMS error can be expressed in a simplified form as

$$\sigma_v=\sqrt{\sum_{n=1}^{N}\sigma_{v,n}^2} \tag{9}$$

$\sigma_{v,n}$: vertical component of RMS correction error due to satellite n
N: number of satellites in solution ($N\geqslant 4$)

The vertical position error distribution depends upon the satellite geometry, error distributions of the N range measurements, and the weights used in the least-squares-position solution. Therefore, there would need to be a different inflation factor for each satellite configuration. To circumvent this problem, a worst-case tail probability in the position domain is needed. This worst case occurs when the RMS error from a single satellite ($\sigma_{v,max}$) dominates the total RMS error. The worst case distribution can be formulated based on the constraint that the projection elements in the row vectors of the position solution matrix sum to zero in order to remove the receiver clock error. Using this constraint, a condition for the least upper bound was predicted and then verified by simulation (sample size $=10^5$). The condition is that the ratio of the largest RMS error to the total RMS error is limited to a specific value.

$$\frac{\sigma_{v,max}}{\sqrt{\sum_{n=1}^{N}\sigma_{v,n}^2}}\leqslant\frac{1}{\sqrt{1+\frac{1}{N-1}}} \tag{10}$$

This condition can be satisfied by a projection vector containing N elements of the form $\left[\sigma_{v,max}-\frac{\sigma_{v,max}}{N-1}-\frac{\sigma_{v,max}}{N-1}\ldots..-\frac{\sigma_{v,max}}{N-1}\right]$. For comparison the best case occurs when all of the $\sigma_{v,n}$ are equal.

The PDF for the worst-case vertical position solution is obtained by the convolution of the N PDFs, where individual values of $\sigma_{v,n}$ are set equal to the elements of the above vector and substituted as the values of σ_0 in (5) with $\delta_0=0{\cdot}65$ and $M=4$. The convolution is done most easily and much faster by the Fast Fourier Transform (FFT) method. The resulting vertical position PDF was constructed from a cubic spline interpolation of the inverse FFT. The PDF is then substituted into (1) for the over bound calculation.

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

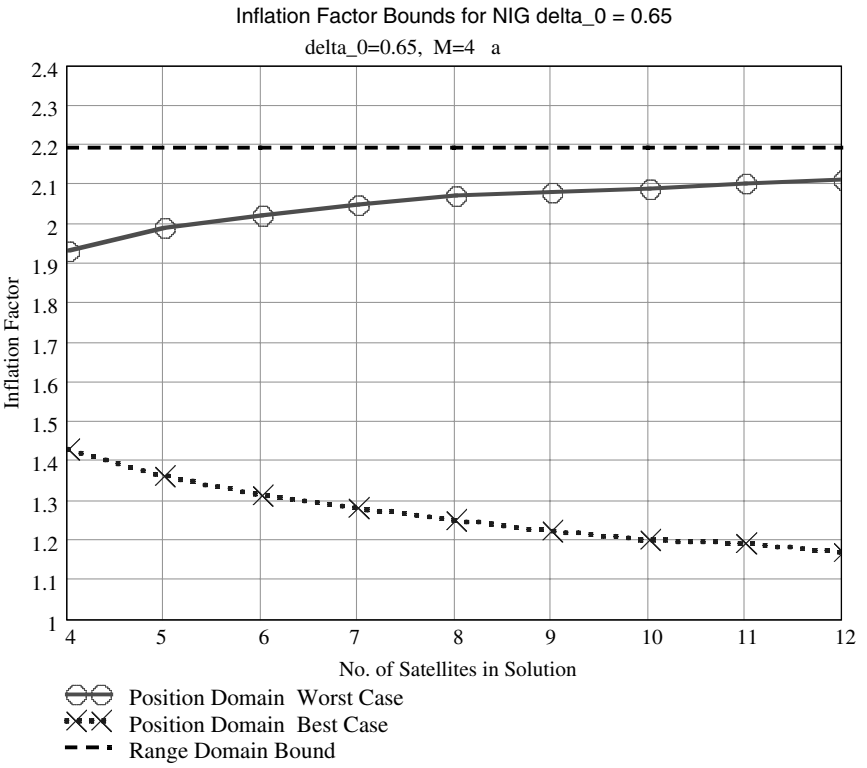


Figure 6. CAT I inflation bounds as a function of the number of satellites in the position solution (M = 4 reference receivers).

4.3. *Over bound results.* Figure 6 contains plots of inflation factor as a function of the number of satellites in a position solution.

The plot with the O symbols is for the worst case tail probability in a position domain solution. The resulting range of inflation factors is narrow, varying from about 1·9 to 2·1 as N varies from 4 to 12. As can be seen, a slight reduction of the inflation factor is achievable if over bounding is done in the position domain using this general method. The best case shown at the bottom of the figure is too optimistic in that actual and simulation data indicate that satellite geometries produce results that are more toward the worst case.

The increase of worst-case inflation factor with increase in N is surprising. An explanation is that as N increases the tails of the resulting position distributions narrow due to the multiple convolutions; however, the ratio in (10) also increases. Apparently the tail thinning cannot quite offset the increase in the ratio.

4.4. *Inflation factor for smaller sample sizes.* The results so far are for the NIG model where $\delta_0 = 0{\cdot}65$ and sample size equal to 6699. The NIG model needed for smaller sample sizes is determined by applying the $P_0(4{\cdot}5) = 0{\cdot}999$ observability criterion to smaller sample sizes. Table 2 contains the independent sample sizes needed to satisfy the $P_0(4{\cdot}5) = 0{\cdot}999$ as a function of various values of the shape parameter δ_0. The inflation factor is then calculated for each of the δ_0 values.

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

Table 2. Sample Sizes Needed to Satisfy $P_0(4{\cdot}5) = 0{\cdot}999$ Observability Criterion

Shape Parameter (δ_0)	0·25	0·35	0·45	0·55	0·65
Sample Size	1260	1890	3000	4650	6700

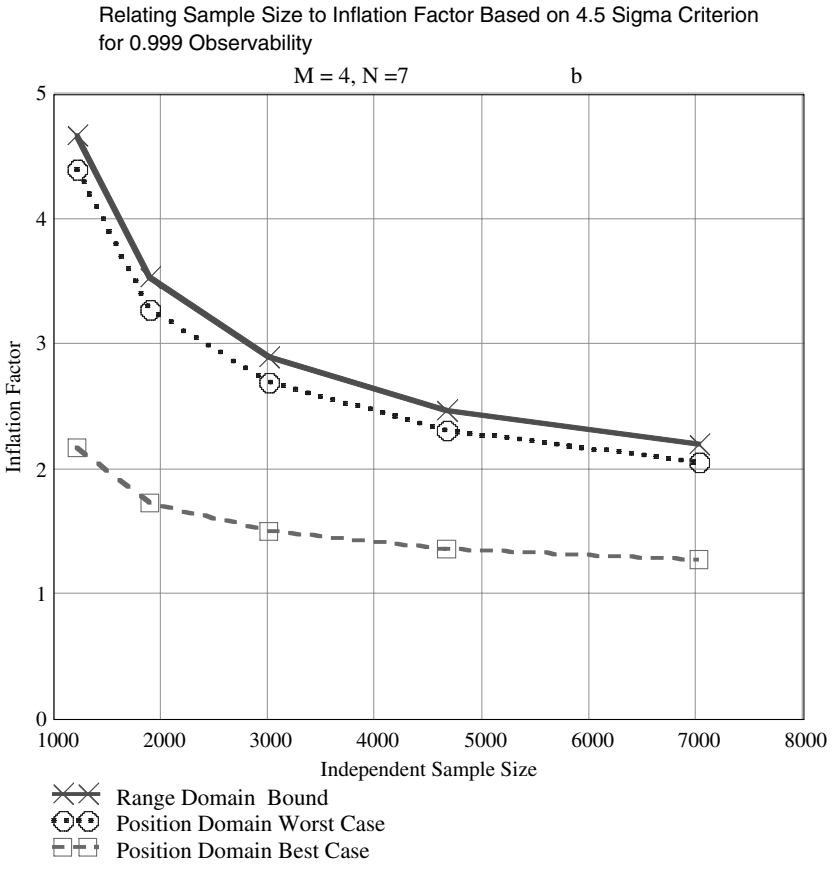


Figure 7. Relating independent sample size to inflation factor.

Note how the values of δ_0 decrease with decrease in sample size. For example, for a sample size of about 2000, the NIG model would be $\delta_0 = 0{\cdot}35$. Figure 7 is a plot of inflation factor as a function of sample size for the $N = 7$ satellite position solution. Recall from Figure 2 that the tail probability increases as δ_0 decreases. Thus as the sample size decreases, the inflation factor increases as expected. From a system availability consideration, it is desirable for the inflation factor to be no more than about 2. In Figure 7, referring to the position domain worst case, it is indicated that the independent sample size would have to be greater than 6000–7000 to meet this goal.

5. POSITION DOMAIN MONITOR FOR MITIGATING MISMODELLING.

A position domain monitor (PDM) is an independent GBAS receiver that is located away from the reference receivers. It generates

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

position solutions using the transmitted pseudorange corrections. The test statistic is the differences between the position solution and the known antenna location. Thus the PDM has the potential to detect unexpected large non-common fault free errors, such as those due to site environmental changes as well as mismodelling of the initially assumed error distribution. That is, the PDM can detect errors due to the LGF or change in its environment, but not common errors due to the satellites. There can be hundreds of different combinations of satellites available for a position solution. For example, with 9 satellites in view there are 382 possible combinations of 4 or more satellites. There seems to be no need to process all satellite combinations. The PDM only processes those satellite combinations that produce protection levels that are just below the alert limit and enough combinations are chosen to include all of the satellites. There is no need to process those combinations that produce protection levels greater than the alert limit because they would be excluded by the airborne protection level monitors. It is assumed that those satellite combinations whose protection levels are well below the alert limit need not be processed because large correction errors would be more likely to be detected through their amplification by the poorer satellite geometries that cause the protection levels to approach the alert limit.

As will be shown below by analysis, for the PDM to be effective its accuracy must be comparable to the accuracy obtained from the averaging of the $M=4$ reference receivers. It is proposed to attain this accuracy by increasing the PDM reference receivers' smoothing time from 100 to 400 s and to remove the repeatable satellite multipath that occurs every day at the same sidereal time. These concepts are currently being validated through the PDM prototype that is being developed at the FAA's William J. Hughes Technical Center.

5.1. *PDM effectiveness in detecting mismodelling of the distribution tails.* The PDM effectiveness is determined by comparing the integrity risk of mismodelling without the PDM and with the PDM. This risk is defined as the probability that a ground facility generated fault-free error in the position solution exceeds $K\sigma_{pr\_gnd}$ ($K=5{\cdot}8$ for CAT I) and the PDM threshold is not exceeded. Assuming the correction and PDM errors are each NIG distributed with $M=4$ and $M=1$ reference receivers, respectively, the vertical risk equation is given by

$$H_{0\_risk} = 2 \int_{K I_f \sigma_{v,avg}}^{\infty} \int_{-K_T \sigma_{v,pdm} - v}^{K_T \sigma_{v,pdm} - v} f_{v\_corr}(v,\ \delta_0\, 4,\ N)\, f_{v\_pdm}(e,\ \delta_0\, 1,\ N)\, de\, dv \qquad (11)$$

v: represents the vertical position error due to the fault-free LGF errors
e: represents the vertical error due to the PDM fault-free errors
f_{v_corr}: PDF of ground facility fault-free component of vertical errors
f_{v_pdm}: PDF of PDM fault-free vertical errors
$\sigma_{v,pdm}$: vertical position RMS error of PDM test statistic
$\sigma_{v,avg}$: vertical position component RMS error of average corrections from 4 reference receivers
I_f: inflation factor
K_T: PDM threshold multiplier
N: number of satellites in position solution

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

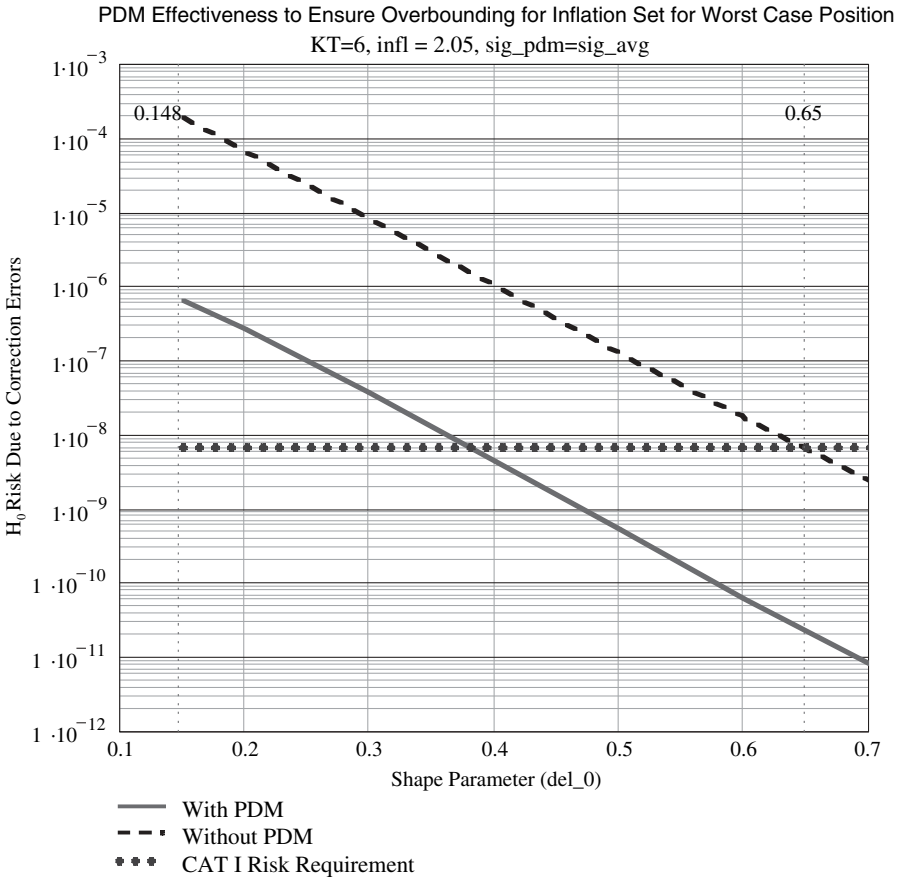


Figure 8. PDM mitigation of tail probability mismodelling.

$\sigma_{v\_pdm} = \sqrt{(1 + C_{pdm}^2)\sigma_{v,avg}^2}$

C_{pdm}: ratio of PDM RMS error to $\sigma_{v,avg}$ (a parameter of the analysis)

From (10), $\sigma_{v,avg} = \sigma_{v,max}\sqrt{1 + \frac{1}{N-1}}$, and without loss of generality $\sigma_{v,max}$ can be set equal to any value.

The worst case position domain heavy tail for an $N = 7$ satellite position solution is the selected example. $I_f = 2{\cdot}05$ corresponds to NIG with $\delta_0 = 0{\cdot}65$. The PDFs in (11) are derived from (5) ($\sigma_0 = 1$) using multiple convolutions of the form

$$f_{v,corr}(v, \delta_0, M, N) = \frac{f_{avg}(r/\sigma_{v1}, \delta_0, 4)}{\sigma_{v1}} \oplus \frac{f_{avg}(r/\sigma_{v2}, \delta_0, 4)}{\sigma_{v2}} \ldots \oplus \frac{f_{avg}(r/\sigma_{vN}, \delta_0, 4)}{\sigma_{vN}} \tag{12}$$

$\sigma_{v1} = \sigma_{v,max}$, $\sigma_{vn} = \frac{\sigma_{v,max}}{N-1}$ $(n = 2, 3 \ldots N)$

$\oplus$: symbol for convolution

Figure 8 contains plots of the integrity risk, with and without the PDM, as a function of δ_0. $0{\cdot}148 \leqslant \delta_0 < 0{\cdot}65$ represents mismodelling in the sense that the actual tail probability is heavier than that assumed in the model. Recall that the heaviest

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

tail occurs when $\delta_0 = 0{\cdot}148$, and this is assumed to be the heaviest tail that can be encountered based on PDF constraints discussed in the Introduction. The PDM threshold multiplier is set to 6 and the RMS PDM error is assumed equal to that of the four reference receivers. The rationale and issues with these PDM parameter assumptions are discussed below.

The dashed curve indicates the integrity risk without the PDM. The maximum possible risk is of the order of 10^{-4} which is somewhat more than four orders-of-magnitude greater than the risk requirement. The solid curve representing the integrity risk with the PDM indicates that the maximum possible risk would be two orders-of-magnitude greater than the risk requirement, and allows the risk requirement to be met for mismodelling that is represented by $0{\cdot}4 < \delta_0 < 0{\cdot}65$. Referring to Figure 2, it is seen that for $\delta_0 = 0{\cdot}4$, the tail probability at the $\pm 5{\cdot}8\sigma$ points for $M = 4$ reference receivers is about 10^{-3}. From the data collected so far there is no indication of such a significant heavy tail. Therefore, if the PDM receive function can be designed to be as accurate as the average of the reference receivers and its threshold multiplier can be limited to about 6, the PDM should be effective in mitigating mismodeling.

5.2. *Sensitivity of PDM risk mitigation to accuracy*. The integrity risk is analyzed with respect to the accuracy of the PDM relative to the average of the reference receivers. The variation in accuracy is accomplished by varying the PDM relative accuracy parameter, C_{pdm}, defined under (11). Figure 9 contains plots of integrity risk for C_{pdm} values of 1, 1·5 and 2. $C_{pdm} = 2$ would represent the expected PDM accuracy relative to that of four reference receivers if there were no accuracy enhancements. As can be seen in the figure, the PDM would provide essentially no risk mitigation for mismodelling if its accuracy were not enhanced.

For significant risk mitigation, as assumed for PDM performance, the PDM has to be at least as accurate as the average of the reference receivers ($C_{pdm} = 1$). For the PDM to completely mitigate mismodelling down to the assumed heaviest possible tail probability ($\delta_0 = 0{\cdot}148$), calculation indicates that the PDM relative accuracy would have to be about 0·1, which is deemed impossible to attain.

5.3. *Sensitivity of PDM risk mitigation to threshold setting*. The integrity risk is analyzed with respect to the threshold setting multiplier K_T with the PDM accuracy assumed to be equal to that of the average of the reference receivers. Figure 10 contains plots of the integrity risk for $K_T = 5, 6 \ldots 8$. Examination of the figure indicates that to achieve two orders-of-magnitude mismodelling mitigation for the heaviest tail, K_T should be no greater than 7.

The threshold cannot be set to just satisfy an integrity risk requirement. It must also satisfy a continuity risk requirement. To obtain some insight into the impact of continuity risk, the following assumptions are made:

- Allowable PDM false alert rate about once per year → false alert rate $\sim 10^{-4}$/h
- In estimating the false detection probability per independent PDM measurement, assume 100 s reference receiver smoothing time yields ~ 18 independent measurements/h
- Assume 10 satellite combinations are processed for each independent measurement
- False detection probability per independent measurement needs to be $\leqslant 10^{-4}/(10 \times 18) = 5{\cdot}6 \times 10^{-7}$ per independent measurement to meet the assumed allowable false alert rate.

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

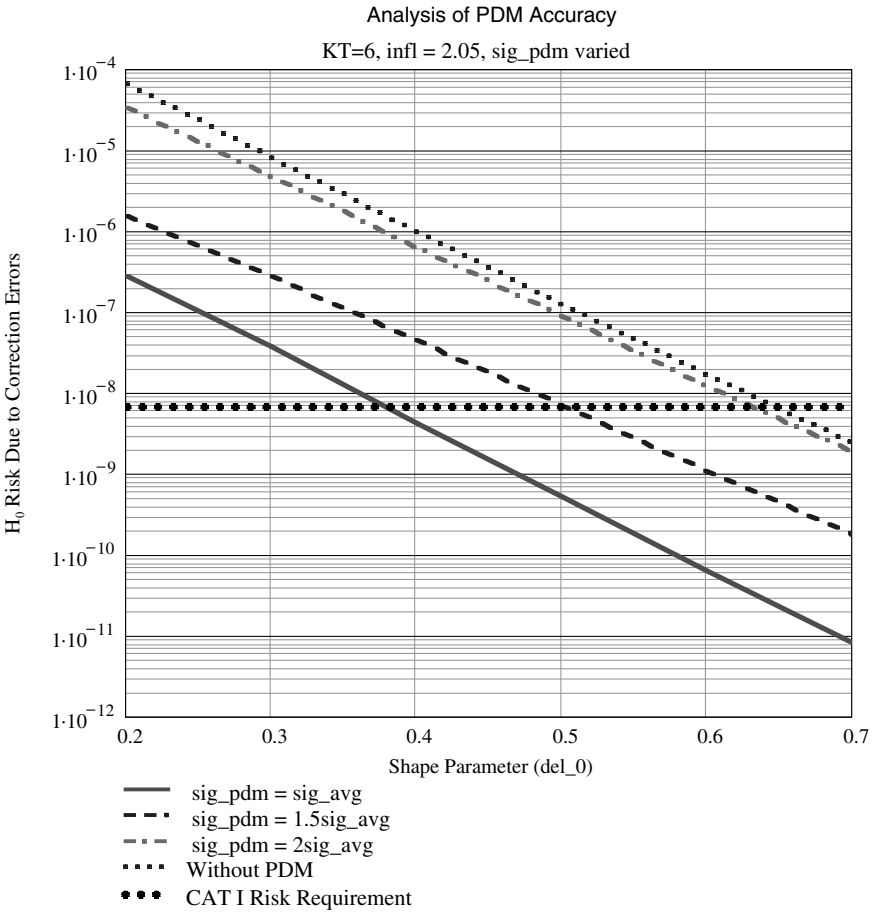


Figure 9. PDM accuracy impact on mismodelling mitigation.

If the test statistic distribution were Gaussian, the false detection probability per independent event would be 2×10^{-9} and $2{\cdot}5 \times 10^{-12}$ for $K_T = 6$ and 7, respectively, which would meet both the integrity and continuity risk requirements. If the test statistic were the NIG with $\delta_0 = 0{\cdot}65$ then $K_T = 11{\cdot}4$, which would only meet the continuity requirement. Since the NIG was chosen as an over bounding distribution for integrity, it may be too conservative for continuity risk estimation. For example, the ''B-value monitor'' in the LGF, which processes pseudorange corrections, has a specified K multiplier in the range 5–6 (FAA, 2002). The feasibility of K_T less than 7 will be determined from PDM prototype data when the accuracy enhancements are completed.

6. SUMMARY. An analysis is performed to determine an upper bound on the tail probability of the fault-free errors of the GBAS corrections that are due to errors attributed to the ground facility, such as multipath and noise. An argument is made that the upper bound tail probability is approximately $0{\cdot}181\ K^{-2}$ for a single

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

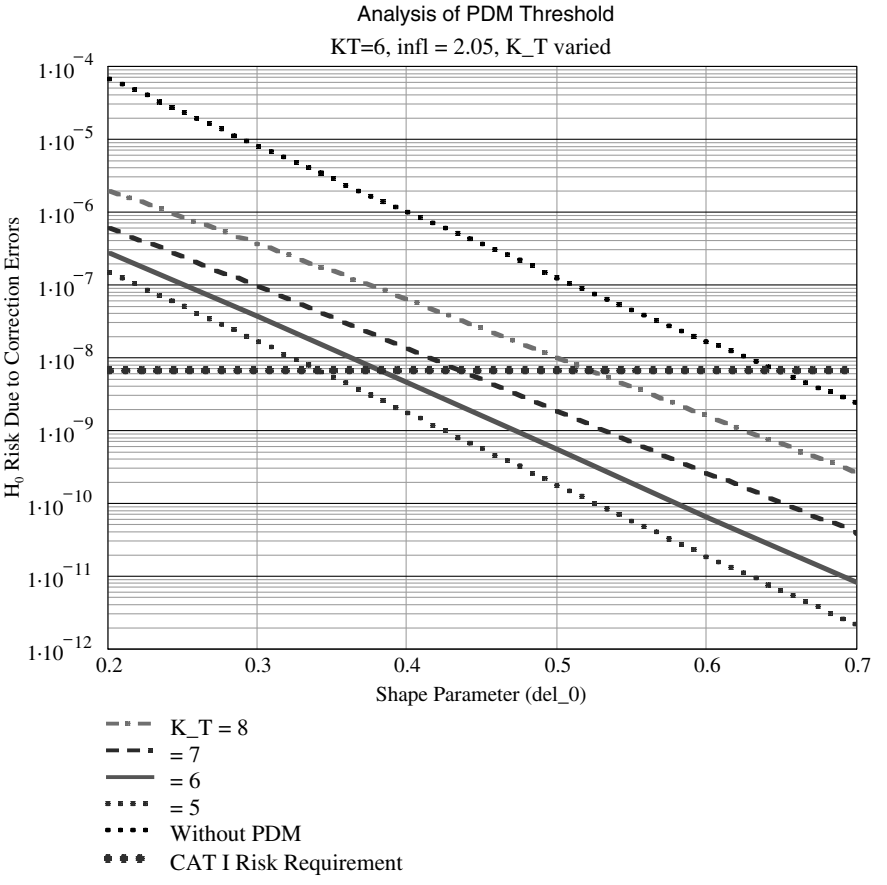


Figure 10. PDM threshold setting impact on mismodelling mitigation.

reference receiver, where K is the standard deviation multiplier. This value is the maximum tail probability of the family of Normal Inverse Gaussian (NIG) distributions. The actual value of the upper bound is in the interval $0{\cdot}181\,K^{-2} \leqslant \mathrm{Prob}_{\mathrm{tail}} \leqslant 0{\cdot}444\,K^{-2}$ ($4/9 \approx 0{\cdot}444$). The $0{\cdot}444\,K^{-2}$ value is an upper tail probability for a non-increasing PDF. In terms of the CAT I value, $K = 5{\cdot}8$, $0{\cdot}0054 < \mathrm{Prob}_{\mathrm{tail}} \leqslant 0{\cdot}0132$). The NIG distribution with tail probability 0·0054 would have been sufficient to derive the inflation factor. Unfortunately, this would result in an inflation factor of ≈13, which is much too large to be practical.

To circumvent this difficulty, a large fault free error observability criterion is introduced that is a function of the independent sample size and assumes that the underlying distribution is a member of the NIG family of distributions. The observability criterion is based on the results of a LAAS prototype data collection campaign conducted by the FAA, where no fault-free errors greater than 4-sigma were observed. The NIG family provides a wide representation of tail probabilities and has the same desirable attributes as those of the Gaussian distribution, including ease of computation. Assuming the differential corrections are the average of four independent ground facility reference receivers and using the observability

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

analysis, the NIG family member with shape parameter $\delta_0 = 0{\cdot}65$ is selected. The CDF of this distribution is compared to the worst-case CDF from the FAA data collection, and it is observed that it provides a good fit. However, although satisfying, no credit is taken for this fit because there is no identifiable physical reason for the fit.

The selected NIG tail probability for a four receiver average is $1{\cdot}4 \times 10^{-4}$. The corresponding Gaussian tail probability at the $\pm 5{\cdot}8$-sigma points is $6{\cdot}63 \times 10^{-9}$. To "over bound" the NIG to yield the Gaussian tail probability an inflation factor of 2·2 is needed. In GBAS the inflation factor multiplies the estimated standard deviation of the corrections to account for the heavy tails of the error distribution.

Since the standard transformation of the pseudorange errors to the position domain is a linear operation, there is the potential for further tail thinning through the multiple convolutions. Note, for the average pseudorange correction from four receivers there are three convolutions. If there were N satellites in the position solution there would be an additional $N-1$ convolutions. However, there are an enormous number of different solutions due to the vast varieties of satellite geometries. Therefore, an upper bound on the tail probability in the position domain is formulated based on an inherent characteristic of the position solution row vectors. Using this upper bound, an inflation factor varying from 2 to 2·1 was derived, depending on the number of satellites in view.

Although the selected NIG distribution is considered conservative ($1{\cdot}4 \times 10^{-4}$ for CAT I), it is based on observability of a limited independent sample size (6699). Therefore, there is a chance that the selected NIG is a mismodelled distribution. Since the tail probability of the NIG increases towards its maximum as the shape parameter δ_0 decreases, the mismodelling is defined as the real distribution with NIG $0{\cdot}148 \leqslant \delta_0 < 0{\cdot}65$. For detecting mismodelling, the concept of a position domain monitor (PDM) is introduced. The PDM is an independent ground receiver that is located away from the reference receivers. It includes the broadcast pseudorange corrections in its position solutions. The test statistic is the differences between the position solutions and the known location of the PDM antenna. For the PDM to be effective, mathematical analysis indicates that the accuracy of the PDM receiver needs to be at least as good as that of the average corrections, and its threshold needs to be no greater than 6–7 σ_{test}. It is the authors' opinion that the heaviness of the selected NIG tails is sufficient for CAT I application and that the PDM should not be needed for CAT I.

7. DISCUSSION. The derived inflation factor of 2 to account for the heaviness of the tails is large due to conservative assumptions. If this were the only inflation factor to be considered, there would be more than enough margin to accommodate it because the measured RMS of the corrections is usually somewhat less than one-half of the requirement. However, there are other inflation factors that also multiply the estimated pseudorange correction RMS error, such as those that have to account for confidence and bias. It is the authors' opinion that the only way to potentially obtain a smaller inflation factor to account for tail probability heaviness is to base it on a physical model that statistically accounts for multipath, antenna and receiver processing as discussed in the paper's introduction. The development

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press

of such a model is more difficult than the statistical model discussed herein. The method described herein is developed as a substitute until a suitable physical model is developed.

ACKNOWLEDGEMENTS

John Warburton (FAA) provided the data and is providing the engineering and physical resources for the PDM prototype at the FAA Technical Center. Dr. Taehwan Kim brought the Normal Inverse Gaussian distribution to the authors' attention. Dr. Chris Varner provided some of the concepts of the PDM and is designing and building the PDM prototype. J.P. Fernow provided a detailed review of an earlier version of this paper. Dan O'Laughlin provided a detailed review of this paper.

REFERENCES

Braff, R. (1997–98), Description of the FAA's local area augmentation system (LAAS). *NAVIGATION, The Journal of The Institute of Navigation*, **44**, No. 1.

Braff, R. (2003), A method of fault-free error over bounding using a position domain monitor. Proceedings of *The Institute of Navigation Technical Meeting*, Anaheim, California.

Brenner, M., Reuter, R. and Schipper, B. (1998), GPS landing system multipath evaluation techniques and results. Proceedings of *The Institute of Navigation GPS 1998*, Nashville, Tennessee.

Campos, L. and Marques, J. (2002), On safety metrics related to aircraft separation. *The Journal of Navigation*, **55**, 59.

Cornell, J. (1990), *Experiments With Mixtures, Designs, Models and Analysis of Mixture Data*. John Wiley and Sons.

DeCleene, B. (2000), Defining pseudorange integrity-over bounding. Proceedings of *The Institute of Navigation GPS 2000*, Salt Lake City, Utah.

FAA (2002), Category I local area augmentation system ground facility. *Specification*, *FAA-E-2937*, April 17, 2002.

Hanssen, A. and Oigard, T. A. (2002), The normal inverse gaussian distribution model for heavy tailed stochastic processes. Proceedings of *IEEE International Conference on Acoustics, Speech and Signal Processing*.

Marshall, J. (2003), Worst case probability density functions for over-bounding LAAS fault-free errors. The MITRE Corporation, (Unpublished).

Parker, J. B. (1966), The Exponential Integral Frequency Distribution. *The Journal of Navigation*, **19**, 526.

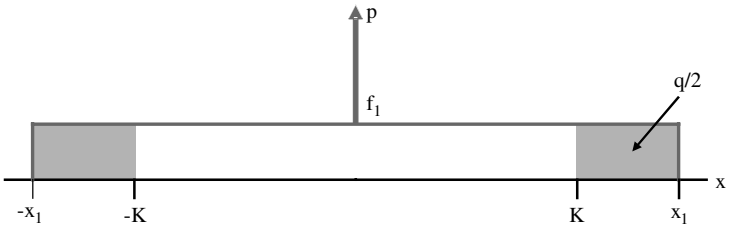
Pervan, B., Pullen, S., and Sayim, I. (2000), Sigma estimation, inflation and monitoring in the LAAS ground system. Proceedings of *The Institute of Navigation GPS 2000*, Salt Lake City, Utah.

RTCA SC-159 (2000), Minimum operational performance standards for GPS local area augmentation system airborne equipment. *RTCA DO-253*.

Warburton, J. (2002), Sigma pseudorange ground establishment, over bounding and monitoring: support data. *FAA Presentation to RTCA SC-159*, December 2002.

APPENDIX: HEAVY TAIL PROBABILITY UPPER BOUND

An attempt was made to derive a heavy tail probability upper bound that is smaller than the Chebyshev inequality, $\text{Prob}\{|x|>K\} \leqslant K^{-2}$; where K is the standard deviation multiplier. So far a PDF has been derived that produces an upper bound for the constraints of non-increasing and unit area and any variance. The form of the PDF is a Dirac delta function at the origin combined with a rectangular distribution. This form was derived from an application of the calculus of variations, and its derivation is given in an unpublished report (Marshall, 2003). Figure 11 contains a description of this PDF. The upper bound tail probability is $(4/9)K^{-2}$. Since no GBAS distributions even remotely resemble this distribution, it is only considered as a means for a mathematical establishment of an upper bound.



Constraints: Non-increasing PDF, area = 1, variance = 1

Parameter values: $X_1 = (3/2)K$, $f_1 = (4/9)K^{-3}$, $p = 1 - (4/3)K^{-2}$, $q = (4/9)K^{-2}$

Figure 11. Definition of non-increasing PDF used to establish upper bound on tail probability.

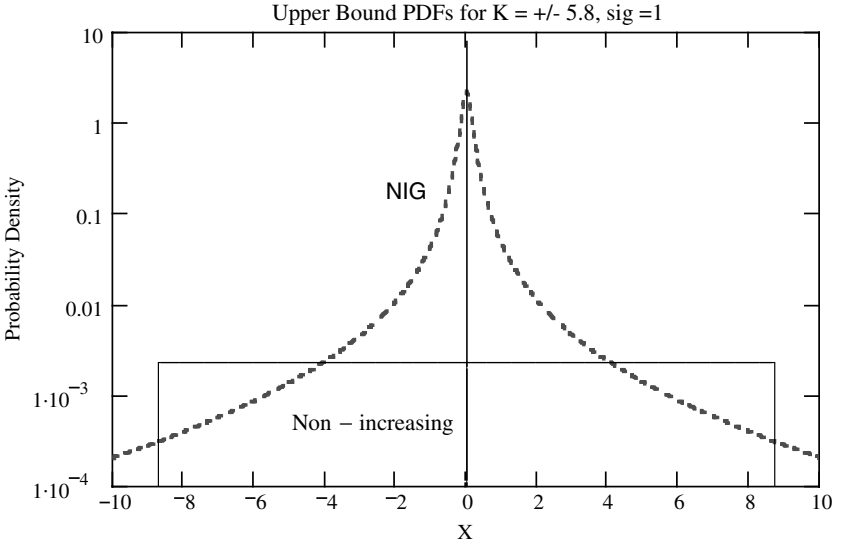


Figure 12. Comparison of upper bound probability density functions.

The maximum tail NIG ($\delta_0 = 0{\cdot}148$) PDF is plotted on a logarithmic scale in Figure 12 along with the non-increasing PDF model in Figure 11. Note, the NIG PDF resembles the non-increasing PDF in that it has a sharp peak and near-level tails. The maximum tail NIG should also be considered only as a means to derive mathematically an upper bound heavy tail probability for a strictly decreasing PDF.

https://doi.org/10.1017/S0373463304003029 Published online by Cambridge University Press