

Monitoring Measurement Noise Variance for High Integrity Applications

Samer Khanafseh, Steven Langel, Fang-Cheng Chan, Mathieu Joerger
and Boris Pervan, *Illinois Institute of Technology*

ABSTRACT

In this paper, a methodology is developed to ensure the integrity of navigation systems given that the measurement noise characteristics might change under different environments. For demanding aviation applications such as autonomous shipboard landing or autonomous airborne refueling, accuracy and integrity of carrier phase positioning and cycle resolution are essential. Integrity risk is computed from the estimate error covariance matrix, which is ultimately derived from the measurement noise covariance matrix. Therefore, any changes in the noise characteristics must be properly accounted for in the integrity risk computation. In this work, a monitor that detects sudden increases in the noise variance is proposed along with an explicit quantification of its effect on integrity risk.

INTRODUCTION

Least squares estimators and Kalman filters guarantee a bounded estimate error variance provided that the measurement errors can be accurately modeled as white noise processes with a variance greater than or equal to the actual variance. The assumed value of the measurement noise standard deviation (or sigma) is used by the navigation algorithm to estimate position and compute protection levels. Nominal values of sigma can be obtained from a prior analysis of data collected over several hours, days, or even months. The data length is primarily dependent on how much noise variation must be encompassed. For example, in the Local Area Augmentation System (LAAS), data is collected over months to capture the impact of non-Gaussian errors, sample size effects, seasonal variations, group delays, and phase center variations on the measurement noise distribution [1]. However, such an analysis can never properly account for all possible antenna environments that may be encountered. The potential danger is that sudden changes in the surroundings (for example weather debris) could lead to situations where the true measurement noise standard deviation exceeds the

nominal value used in the estimator. If unaccounted for, these situations can ultimately lead to optimistic protection levels which jeopardize the system integrity risk.

In response, we propose a monitor that always checks if the actual sigma exceeds the nominal value. However, instantaneous changes in the measurement noise characteristics will not be immediately detectable by the monitor. This results from the fact that it takes a finite amount of time to acquire a suitable data set that will yield a meaningful standard deviation. Therefore, the navigation algorithm must use an inflated sigma (relative to the nominal value) which allows sufficient time for the monitor to detect these sudden changes while simultaneously guaranteeing meaningful protection levels. Qualitatively, the inflated sigma is chosen to provide acceptable levels of availability and probability of missed detection for the monitor.

In this work, we provide a derivation that relates integrity risk or probability of hazardous misleading information (HMI) to the inflation factor, the minimum detectable variance increase, mean time to detect (MTTD), and the mean time between failures (MTBF), where a failure is the event where the actual measurement noise standard deviation exceeds the nominal value. This derivation is used to impose requirements on the MTTD and the minimum detectable sigma of the monitor for a given integrity risk allocation, MTBF and inflation factor; the allowable inflation factor being driven by availability. These requirements are then used in designing a variance shift monitor based on the Cumulative Sum (CUSUM) monitoring technique. The performance of this monitor is compared to the requirements and we demonstrate that such a monitor has the potential to detect variance shifts while meeting the integrity requirements. This paper therefore consists of two parts: the first where the monitor requirements are derived from the integrity risk allocations; and the other is where the monitor design is discussed. Using example numbers for requirements, we also show performance results for such a monitor given the requirements.

HMI PROBABILITY

Before discussing the monitor design in detail, it is instructive to introduce some fundamental concepts that will be used repeatedly throughout this paper. The notion of a protection level occurs frequently in aviation applications, and it is defined as a statistical bound on the estimate error. Under fault-free conditions, it can be computed by multiplying the standard deviation of a coordinate estimate error σ_x (coming from the estimator) by an integrity factor k_{ff} , that is derived from the required fault-free integrity risk, I_{H0} (1).

$$PL = k_{ff} \sigma_x \quad (1)$$

where k_{ff} is determined from I_{H0} through the relation

$$k_{ff} = \Phi^{-1}\left(\frac{I_{H0}}{2}\right) \quad (2)$$

where Φ is the cumulative distribution function for a standard normal random variable.

Although this paper is concerned with anomalous conditions (and hence not fault-free), we still use Equation (1) to construct the PL. Because σ_x is computed using an inflated measurement noise standard deviation, σ_b^2 , Equation (1) will remain valid for a certain amount of time even if a sudden increase in the true measurement noise standard deviation, σ_i^2 , has occurred. It is important to note that σ_b^2 is only used to give the monitor time to detect if σ_i^2 is larger than the nominal value, σ_n^2 . Therefore, the monitor test statistic and threshold are solely based on σ_i^2 and σ_n^2 .

In the case of variance monitoring, a hazardously misleading information (HMI) situation occurs if σ_i^2 exceeds σ_n^2 , the monitor's test statistic is less than the threshold and therefore did not detect the anomaly, and the true estimate error exceeded the protection level. Mathematically, the probability of HMI can be expressed as:

$$P(HMI, S) = P\left\{|\delta x(\sigma_i)| > PL(\sigma_b), q(\sigma_i) < T(\sigma_n), S\right\} \quad (3)$$

Using this equation, the goal is to derive requirements for the monitor, such as how fast it should detect various events S that result in $\sigma_i > \sigma_n$. First, we expand Equation (3) using the law of total probability

$$P(HMI, S) = P\left\{|\delta x(\sigma_i)| > PL(\sigma_b), q(\sigma_i) < T(\sigma_n) | S\right\} P\{S\} \quad (4)$$

In order to simplify the first term on the right hand side of (4), we assume that the event where the estimate error exceeds the protection level and the event where the test statistic is less than the threshold are statistically independent. This assumption is valid if the test statistic is computed using a measurement that is one independent epoch delayed from the measurement that is used in the estimation of x . This delay will be accounted for in the mean time to detect (by increasing it by the period of one independent epoch) for the monitor as we will see in the second part of this paper. With that assumption, (4) becomes,

$$P(HMI, S) = P\left\{|\delta x(\sigma_i)| > PL(\sigma_b) | S\right\} P\left\{q(\sigma_i) < T(\sigma_n) | S\right\} P\{S\} \quad (5)$$

The second and third term in the previous equation represent the probability of missed detection (PMD). In this work we treat the PMD as a state probability such that the system is either in the state of $q < T$ and $\sigma_i > \sigma_n$ or not. This probability has a binomial distribution that can be represented as a function of the failure rate (or MTBS in our case) and exposure time (MTTD). This probability is evaluated using the exponential function as:

$$P\left\{q(\sigma_i) < T(\sigma_n), S(\sigma_i > \sigma_n)\right\} = 1 - e^{-\frac{MTTD}{MTBS}} \quad (6)$$

For the first term on the right hand side of (5), let us assume that the measurement error distribution is Gaussian as shown in Figure 1. This assumption is reasonable and is often made in aviation applications given that the tail probabilities of the actual distribution are still bounded by a Gaussian. In this case, the probability that the actual coordinate estimate error exceeds the corresponding protection level can be represented by the shaded area in Figure 1. Mathematically, this can be computed using the tail probability function (Q) as,

$$P\left\{|\delta x(\sigma_i)| > PL(\sigma_b) | S\right\} = 2Q\left(\frac{PL(\sigma_b)}{\sigma_x(\sigma_i)}\right) \quad (7)$$

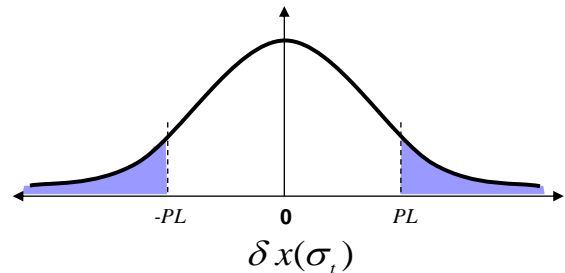


Figure 1: Gaussian distribution of the estimate error δx

Substituting the expression for PL from (1) into (7), yields:

$$P\left\{|\delta x(\sigma_i)| > PL(\sigma_b) \mid S\right\} = 2Q\left(\frac{k_{ff} \sigma_x(\sigma_b)}{\sigma_x(\sigma_i)}\right) \quad (8)$$

For a least squares estimator with an observation matrix \mathbf{H} and measurement noise covariance matrix \mathbf{R} , the estimate vector \mathbf{x} is:

$$\mathbf{x} = (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{z} \quad (9)$$

where \mathbf{z} is the measurement vector, $(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1}$, is the pseudo-inverse of \mathbf{H} which will be denoted as \mathbf{S} , and \mathbf{R} is a diagonal matrix with different variances on the diagonal for different satellites as shown in Equation 10.

$$\mathbf{R} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_j^2 \end{bmatrix} \quad (10)$$

Remember that the estimator will populate \mathbf{R} using the values $\sigma_{b,m}^2$, where $m = 1, 2, \dots, j$. In this case, the estimate error covariance matrix \mathbf{P} is given by

$$\mathbf{P} = \mathbf{H}^+ \mathbf{R} \mathbf{H}^{+T} \quad (11)$$

In the event that \mathbf{R} is not diagonal due to a double differencing operation, the transformation matrix can be extracted and absorbed in \mathbf{H}^+ . If the variance of one state is of interest (for example, the i^{th} state for which PHMI is computed), σ_x^2 can be written as:

$$\sigma_x^2 = \mathbf{H}_i^+ \mathbf{R} \mathbf{H}_i^{+T} \quad (12)$$

where \mathbf{H}_i^+ is the i^{th} row of \mathbf{H}^+ . Exploiting the diagonal structure of \mathbf{R} and expanding (12) in a series form we get:

$$\sigma_x(\sigma_b) = \sqrt{\sum_{m=1}^n H_{i,m}^{+2} \sigma_{b,m}^2} \quad (13)$$

where $H_{i,j}^+$ is the (i, m) element of \mathbf{H}^+ . The same steps can be followed to show that the variance of the actual estimate error (using the true variance $\sigma_{t,m}^2$ instead of $\sigma_{b,m}^2$) is:

$$\sigma_x(\sigma_t) = \sqrt{\sum_{m=1}^n H_{i,m}^{+2} \sigma_{t,m}^2} \quad (14)$$

Substituting (13) and (14) in (8), we get

$$P\left\{|\delta x(\sigma_i)| > PL(\sigma_b) \mid S\right\} = 2Q\left(\frac{k_{ff} \sqrt{\sum_{m=1}^n H_{i,m}^{+2} \sigma_{b,m}^2}}{\sqrt{\sum_{m=1}^n H_{i,m}^{+2} \sigma_{t,m}^2}}\right) \quad (15)$$

Notice that the probability in (15) can be evaluated as long as we have access to the satellite geometry. However, it is preferred to have an expression for this probability that is independent of geometry to derive the monitor requirements. Multiplying and dividing both terms inside the square roots in the numerator and denominator of (15) by $\sigma_{n,j}^2$ results in:

$$P\left\{|\delta x(\sigma_i)| > PL(\sigma_b) \mid S\right\} = 2Q\left(\frac{k_{ff} \sqrt{\sum_{m=1}^n H_{i,m}^{+2} \sigma_{n,m}^2 \frac{\sigma_{b,m}^2}{\sigma_{n,m}^2}}}{\sqrt{\sum_{m=1}^n H_{i,m}^{+2} \sigma_{n,m}^2 \frac{\sigma_{t,m}^2}{\sigma_{n,m}^2}}}\right) \quad (16)$$

In order to eliminate the geometry, we must make some conservative assumptions that provide an overbound for the probability in (16). This overbound is achieved if the numerator is underbounded and the denominator is overbounded. So, if we assume that the minimum value of the buffer ratio $\sigma_{b,j}/\sigma_{n,j}$ is used for all satellites ($f_{b,\min}$) and that the maximum fault ratio $\sigma_{t,j}/\sigma_{n,j}$ is also used for all satellites ($f_{t,\max}$), these ratios can be extracted out of the summation and the square roots, yielding the new expression:

$$P\left\{|\delta x(\sigma_i)| > PL(\sigma_b) \mid S\right\} \leq 2Q\left(\frac{k_{ff} \sqrt{\sum_{m=1}^n H_{i,m}^{+2} \sigma_{n,m}^2 f_{b,\min}}}{\sqrt{\sum_{m=1}^n H_{i,m}^{+2} \sigma_{n,m}^2 f_{t,\max}}}\right) \quad (17)$$

$$\text{where } f_{b,\min} \equiv \min_m \left(\frac{\sigma_{b,m}^2}{\sigma_{n,m}^2}\right) \text{ and } f_{t,\max} \equiv \max_m \left(\frac{\sigma_{t,m}^2}{\sigma_{n,m}^2}\right).$$

This can then be simplified after eliminating the common square root terms as:

$$P\left\{|\delta x(\sigma_i)| > PL(\sigma_b) | S\right\} \leq 2Q\left(\frac{k_{ff} f_{b, \min}}{f_{i, \max}}\right) \quad (18)$$

Substituting (6) and (18) into (5), an upper bound on the probability of HMI, $\overline{P\{HMI, S\}}$ can be expressed as,

$$\overline{P\{HMI, S\}} = 2\left(1 - e^{-\frac{MTTD}{MTBS}}\right) Q\left(k_{ff} \frac{f_{b, \min}}{f_{i, \max}}\right) \quad (19)$$

MONITOR REQUIREMENTS DERIVATION

Recall that the goal of the first part of this paper is to derive requirements on the variance monitor, given the integrity risk requirements. These requirements are MTTD for various fault sizes ($f_{i, \max}$) for different buffer ratios ($f_{b, \min}$). In order to illustrate how to derive such requirements, we will use an example integrity risk tree as shown in Figure 2. This tree is simply an example to demonstrate the methodology in deriving monitor requirements and does not represent the actual requirements for any particular application. However, the values were chosen to be reasonably close to those used in aviation. From the fault-free integrity risk budget, the integrity multiplier k_{ff} is found to be 4.53 for a fault free integrity risk of 6×10^{-6} . A budget of 1×10^{-6} is allocated to the sigma monitor.

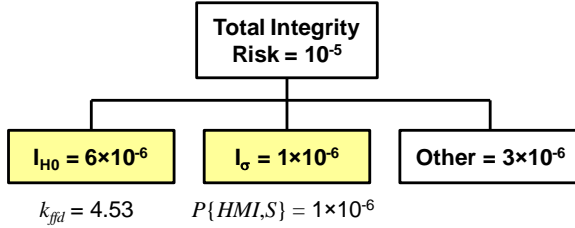


Figure 2: Example integrity risk tree budget.

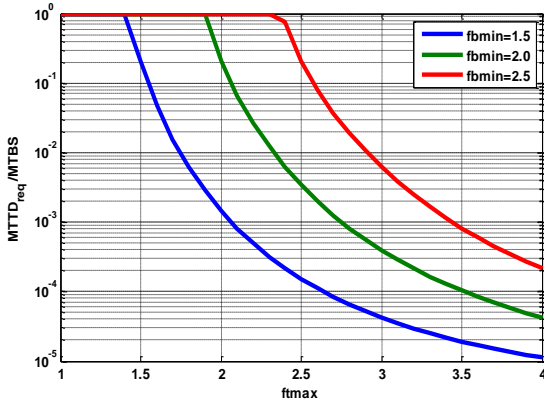


Figure 3: Variance monitor requirements

Using this budget in the left hand side of equation 19 with k_{ff} , the relation between the ratio of required MTTD ($MTTD_{req}$) to the $MTBS$, $f_{i, \max}$ and $f_{b, \min}$ can be expressed as shown in Figure 3. $MTTD_{req}/MTBS$ ratios that exceed 1 are excluded because they imply that the required time to detect is larger than the mean time between failures, and hence a monitor is not necessary. This figure shows three monitor requirement curves corresponding to three buffer ratios (1.5, 2 and 3). For example, if the buffer ratio $f_{b, \min}$ is 2, no monitor is necessary for fault ratios $f_{i, \max}$ that are less than 1.8. If the fault ratio increases to 2, the monitor has 200 hours to detect it, while such a monitor has to detect a fault ratio of 3 in only 18 minutes. We will revisit these curves during the monitor design phase discussed in the next section.

VARIANCE MONITOR

In order to monitor the variance it is necessary to extract the measurement noise, v , which can be accomplished several ways. For example, in dual frequency applications, the ionosphere-free code-minus-carrier operation can be used to extract the noise in the code measurements. Having the measurement noise extracted from the measurements, several methods have been developed in the literature to detect variance shifts [2-6].

One intuitive method is to estimate the variance (or standard deviation) directly from v . Although estimating the variance is straightforward and its distribution is known to be chi-squared with one degree of freedom, the number of samples required to obtain a reliable variance estimate is considerably large. For example, [2] claims that at least 18 independent samples are required to estimate the variance within an acceptable reliability. Therefore, if we assume that the multipath autocorrelation time constant is on the order of 100 seconds, one hour would be required to have a reliable estimate of the variance (18 independent measurements at 200 sec apart). Also, it has been shown in [2, 4] that the performance of the Cumulative Sum (CUSUM) monitor is generally better than that of the direct variance estimator. Therefore, the CUSUM monitor is considered in this work.

CUSUM monitoring was introduced in the 1940's by Wald [5], and uses a log-likelihood ratio test to detect mean and variance shifts. It is widely used in quality control, operational research, and manufacturing engineering. Lee and Pullen [2-4] proposed using the CUSUM monitor to detect mean and variance shifts in GPS reference receiver measurements (B-values) for LAAS. For those interested in the theoretical background of the CUSUM monitor and its design, it is recommended to consult [5, 6].

The variance monitor test statistic of the CUSUM for any observable Y_m is [5]:

$$C_m^+(j) = \max(0, C_m^+(j-1) + Y_m(j) - k_\sigma^+) \quad (20)$$

The initial value for C^+ can either be 0 (standard CUSUM) or any other value as a fraction of the threshold (Fast Initial Response FIR CUSUM). For variance shift detection, the observable for satellite m at epoch j is the weighted square of the measurement noise $v_m(j)$ (21).

$$Y_m(j) = \left(\frac{v_m(j) - \mu_v}{\sigma_{n,m}} \right)^2 \quad (21)$$

where μ_v is the mean of v_m and $\sigma_{n,m}$ is the nominal standard deviation of v_m . GPS measurement noise is usually assumed to have zero mean, and therefore, (21) can be simplified to:

$$Y_m(j) = \left(\frac{v_m(j)}{\sigma_{n,m}} \right)^2 \quad (22)$$

The constant k_σ^+ in (20) is derived based on the designed value for the detection variance as:

$$k_\sigma^+ = 2f_{t,\max}^2 \frac{\ln(f_{t,\max})}{f_{t,\max}^2 - 1} \quad (23)$$

The derivation of this formula can be found in [5]. In this work, we designed the CUSUM monitor for a fault ratio of 2 ($f_{t,\max} = 2$ in Equation 23), which results in $k_\sigma^+ = 1.848$.

The CUSUM monitor alarms when the test statistic C_m^+ exceeds a threshold T . The threshold for the CUSUM is computed based on the assumption that the observables Y_m (and hence the measurement noise v_m) are independent. However, due to multipath, measurement noise is correlated. As a rule of thumb, measurements taken at twice the multipath autocorrelation time constant can be assumed to be independent. In this work, we assume a time constant of 100 sec for multipath, implying independence for measurements taken 200 sec apart. When the input observables to the CUSUM are independent, C_m^+ becomes a Markov process. In Markov processes, the current state is only dependent on the previous state. Therefore, the MTTD for the CUSUM monitor is computed using the transition matrix for the Markov process in conjunction with properties of Markov chains. That said, the MTTD computation is still an iterative process. The details behind the computation of MTTD can be found in [5, 6]. To define the threshold,

the required probability of false alarm (1×10^{-6} in this example) that meets the continuity requirements is used to compute T as 30. With k_σ^+ and T being determined, the CUSUM monitor is complete and its performance can be quantified.

In order to quantify the performance of this CUSUM monitor, different MTTD are computed for different fault ratios ($f_{t,\max}$) ranging from 1.3 to 7. In order to plot the resultant MTTD from the CUSUM on the same chart as the requirements (shown in Figure 3), a value for MTBS needs to be assumed. Since events where the measurement noise variance exceeded the nominal value are rare, MTBS is assumed to be 3600 hours (5 months). The MTTD performance curves are plot against $MTTD_{req}$ in Figure 4.

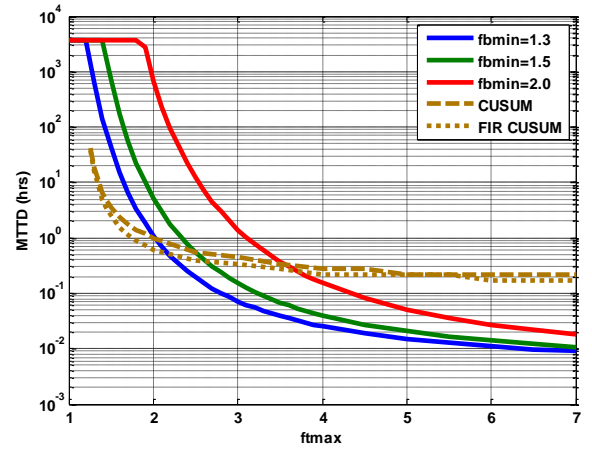


Figure 4: Performance of the CUSUM monitor compared to the requirements derived in Figure 3.

The figure shows that the performance of the CUSUM monitor does not meet the requirements for fault ratios above 3.5 even when the nominal standard deviation is inflated by a factor of 2 ($f_{b,\min} = 2$). When investigated, it was found that even the fastest monitor that would detect instantly (with one independent epoch delay to make the test statistics independent from the measurement error) would result in an MTTD of 200 sec (0.056 hour). This number is considerably higher than the requirement curves. In fact, in order to know the required MTTD for a long range of fault ratios, which is what the monitor should detect, the limit of (19) is taken when $f_{t,\max}$ goes to infinity (24).

$$\lim_{f_{t,\max} \rightarrow \infty} \left\{ 2 \left(1 - e^{-\frac{MTTD}{MTBS}} \right) Q \left(k_{\beta l} \frac{f_{b,\min}}{f_{t,\max}} \right) \right\} = \left(1 - e^{-\frac{MTTD}{MTBS}} \right) \quad (24)$$

Therefore, the relationship between MTBS and the required integrity budget for the sigma monitor (PHMI) is:

$$MTBS = -\frac{MTTD_{\min}}{\ln(1 - P\{HMI, S\}_{req})} \cong \frac{MTTD_{\min}}{P\{HMI, S\}_{req}} \quad (25)$$

With the minimum allowable MTTD being 0.056 hours (as shown before) and the required PHMI of 10^{-6} , the minimum MTBS should be around 6.5 years, which is unrealistic. Alternatively, if PHMI is to be relaxed based on an MTBS of 3600 hours, it should be relaxed to 1.56×10^{-5} . In either case, this analysis shows the relationship between the required PHMI and MTBS for any reasonable monitor to work. If relaxing the required PHMI is not an option due to tight total integrity budget, the MTBS assumption must be revisited.

So far, it has been assumed that the MTBS is constant regardless of what the fault ratio is. This is not realistic because larger faults have a smaller likelihood of occurrence than smaller faults. To test the influence of changing MTBS from a constant value to a variable with different fault magnitudes, the following exponential expression is used as a model

$$MTBS = 1000 e^{1.2(f_{t,max} - 1)} \quad (26)$$

Notice that this MTBS is smaller than the constant one assumed earlier (3600 hours) for $f_{t,max} < 2.1$. Using this MTBS, the requirement curves are plotted versus the monitor performance in Figure 5. It can be seen that the monitor was fast enough to detect all faults and meet the requirements given that the buffer ratio is at least 1.5. Therefore, if the nominal standard deviation is inflated by 1.5 times and the designed sigma monitor is implemented with the assumed parameters, the integrity risk requirements will be met. Further analysis to investigate the impact of any measurement noise inflation due to the monitor on the availability will be conducted in the future for an example application like shipboard landing. Also, the basis of MTBS relation with fault magnitude will be derived more thoroughly.

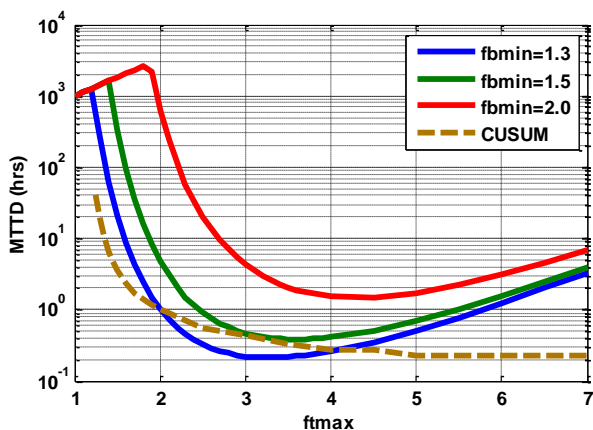


Figure 5: Performance of the CUSUM based on varying MTBS as a function of $f_{t,max}$ compared to the requirements derived in Figure 3.

CONCLUSIONS

In this paper, we developed a method to quantify the integrity risk of measurement variance monitoring. This method is capable of determining the amount of inflation that is required on the nominal values of the measurement noise used in the computation of integrity risk and protection levels. This method also provides requirements in terms of mean time to detect for different fault sizes for the monitor to meet. In this work, we demonstrated how a CUSUM monitor can be used to detect measurement noise variance shifts given the integrity and continuity risk requirements for the navigation system. Although the initial performance was not satisfactory, it was shown this was primarily due to the fact that a constant failure rate was used for all fault magnitudes. In response, we demonstrated the influence of changing the failure rate as a function of the fault size and showed that a feasible monitor can be designed that meets stringent aviation requirements.

ACKNOWLEDGMENT

The authors gratefully acknowledge the Naval Air Systems Command (NAVAIR) of the US Navy for supporting this research. The authors would also like to specifically thank Sam Pullen from Stanford University for his help and guidance in this research. The opinions presents in this paper are those of the authors and do not necessarily represent those of NAVAIR or any other affiliated agencies.

REFERENCES

- [1] Sayim, I., Pervan, B., "LAAS Ranging Error Overbound for Non-Zero Mean and Non-Gaussian Multipath Error Distributions," *Proceedings of the 59th Annual Meeting of The Institute of Navigation and CIGTF 22nd Guidance Test Symposium*, Albuquerque, NM, June 2003, pp. 490-499.
- [2] Lee, Jiyun, Pullen, Sam, Xie, Gang, Enge, Per, "LAAS Sigma-Mean Monitor Analysis and Failure-Test Verification," *Proceedings of the 57th Annual Meeting of The Institute of Navigation*, Albuquerque, NM, June 2001, pp. 694-704.
- [3] Pullen, S., Lee, J., Xie, G., Enge, P., "CUSUM-Based Real-Time Risk Metrics for Augmented GPS and GNSS," *Proceedings of the 16th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS/GNSS 2003)*, Portland, OR, September 2003, pp. 2275-2287.

- [4] Lee, Jiyun, Pullen, Sam, Enge, Per, "Sigma-Mean Monitoring for the Local Area Augmentation of GPS," *IEEE Transaction on Aerospace and Electronic Systems* Vol. 42, No. 2, April 2006, pp. 625-635.
- [5] D. Hawkins and D. Olwell, *Cumulative Sum Charts and Charting for Quality Improvement*, New York: Springer-Verlag, 1998.
- [6] M. Basseville and I. Nikiforov, *Detection of Abrupt Change: Theory and Application*. Englewood Cliffs, N.J. Prentice-Hall, 1993.