LAAS Sigma-Mean Monitor Analysis and Failure-Test Verification

Jiyun Lee, Sam Pullen, Gang Xie, and Per Enge

Stanford University

ABSTRACT

The Local Area Augmentation System (LAAS) is a ground-based differential GPS system being developed to support aircraft precision approach and landing navigation with guaranteed integrity. Stanford University has designed, implemented and tested a LAAS ground Facility (LGF) prototype, known as the Integrity Monitor Testbed (IMT), which is used to insure that the LGF meets its requirements for navigation integrity. One significant integrity risk is that the mean of the pseudorange correction error distribution becomes non-zero or that its standard deviation (sigma) grows to exceed the broadcast correction error sigma (σ_{pr_gnd}) during LAAS operation. Real-time mean and sigma monitoring is necessary to help insure that the true error distribution is bounded by a zero-mean Gaussian distribution with the broadcast sigma value.

In addition to ensuring that the error distribution based on the broadcast sigmas overbounds the true error distribution under nominal conditions, mean and sigma monitoring is needed to detect violations due to unexpected anomalies with acceptable residual integrity risk. Both mean/sigma estimation and Cumulative Sum (CUSUM) methods are useful in this respect. For sigma monitoring, estimation more rapidly detects small violations of σ_{pr_gnd}, but the "fast-impulse-response" (FIR) CUSUM variant more promptly detects significant violations that would pose a larger threat to user integrity.

Based on these analytical results, mean and sigma estimation and CUSUM methods have been implemented in the IMT and have been tested under both nominal and failure conditions. Under nominal conditions, both sigma estimates and CUSUMs stay below the relevant detection thresholds for all visible satellites in the IMT datasets we have tested. In failure testing, both sigma estimation and CUSUM methods reliably detect injected sigma violations, although both methods are limited by the 200-second interval between independent B-values. Similar results were obtained in testing of the mean monitors. Overall, both methods work smoothly and predictably for sigma and mean monitoring to maintain user integrity under both nominal and failure conditions.

1.0 INTRODUCTION

The Local Area Augmentation System (LAAS) will support precise and safe navigation that meets the requirements of Category I (and later Category II and III) aircraft precision approach [7,8,17]. The LAAS Ground Facility (LGF) integrity monitoring is responsible for detection and removal of anomalies such as ground-based or space-based system failures. User navigation integrity is quantitatively appraised by the position bounds that can be ensured with an acceptable level of integrity risk. In this regard, aircraft using LAAS corrections compute the vertical protection level (VPL) and the lateral protection level (LPL) at the aircraft as position error limits assuming a zero-mean, normally distributed fault-free error model for the broadcast pseudorange corrections. User integrity thus relies on the standard deviations of pseudorange-correction errors that are broadcast to users along with the corrections, since these "sigmas" are used in the calculation of VPL and LPL. The bounding standard deviation of correction error, or σ_{pr_gnd}, is broadcast for each satellite approved by the LGF [8,13,17].

To validate the safety of the broadcast σ_{pr_gnd}, the LGF must verify in real time that a Gaussian distribution with zero mean and the broadcast sigma overbounds the true (unknown) error distribution. In this process, special care must be taken regarding the uncertainty of the true error distribution. The main sources of uncertainty are mean and sigma estimation error during site installation and non-stationary error distributions caused by environmental changes that affect multipath, as well as the fact that the tails of the true error distribution may not be exactly Gaussian. This uncertainty is accounted by broadcasting an inflated sigma such that the broadcast distribution overbounds all possible error distributions out to the probabilities assumed in the computation of VPL and LPL [1,2].

Because of the direct connection between the broadcast σ_{pr_gnd} and user integrity, real-time sigma monitoring is necessary to detect anomalies, or failure events where, during operation, the true sigma exceeds the broadcast σ_{pr_gnd}. This paper presents two different sigma monitoring algorithms (sigma estimation and CUSUM) that have been implemented in the Integrity Monitor Testbed (IMT) and also reports the nominal and failure test results of these methods. In failure testing, sigma violations are induced by modifying stored IMT receiver packets collected under nominal conditions to represent sigma anomalies, and the altered measurements are injected back into the IMT in post-processing mode [10].

For the same reason, monitoring of the true mean in real-time is needed to detect unforeseen conditions that cause the true mean to become substantially non-zero. This paper also describes the mean estimation and CUSUM algorithms implemented in the IMT and reports their test results. The results of both nominal and failure tests demonstrate that the sigma-mean monitoring algorithms can detect failures large enough to threaten user integrity and are integrated smoothly with the Executive Monitoring (EXM) Phase II logic in the IMT that excludes such anomalies [6,10].

2.0 SIGMA MONITORING

Sigma monitoring plays an important role in ensuring that the possibility of the true sigma exceeding the broadcast sigma poses no significant integrity risk to LAAS users. This possibility exists not only because of nominal sigma uncertainty but also because real-time sigma monitoring is limited by the number of independent samples that can be collected in one satellite pass. Thus, sigma monitoring cannot detect all cases in which the broadcast σ_{pr_gnd} no longer bounds the true sigma. In order to account for this, additional sigma inflation beyond that needed to bound nominal sigma uncertainty may be necessary to provide margin so that sigma monitoring can meet the LAAS Ground Facility integrity requirements [8,17].

Among many proposed algorithms for monitoring sigma in real-time, the following two different algorithms have been implemented in the Stanford IMT. These algorithms, combined with inflation of the nominal σ_{pr_gnd}, are sufficient to accomplish detection of anomalies that cause the true sigma to exceed the broadcast sigma during LAAS operations.

2.1 Sigma Estimation Method

The real-time sigma estimation method estimates sample standard deviations of the pseudorange correction error from LGF B-values computed in the Multiple Reference Consistency Check (MRCC) for each visible satellite. Since the B-values represent pseudorange correction differences across reference receivers (ideally, the pseudorange corrections from all reference receivers should be the same for a given satellite), the B-values represent pseudorange correction errors that would exist if a given reference receiver has failed [6,8,9]. The normalized values of 'B_{pr_normal}', or B-values divided by their theoretical sigmas (σ_{Bpr}), are the inputs to sigma estimation:

$$B_{pr\_normal} = \frac{B_{pr}}{\sigma_{B_{pr}}} \qquad \sigma_{B_{pr}} = \frac{\sigma_{\text{pr\_gnd}}}{\sqrt{M(n)-1}} \tag{1}$$

Under normal conditions, the estimated sigma ($\hat{\sigma}_B$) has a chi-square distribution with $A-1$ degrees of freedom, where A is the number of independent samples used to derive the estimate. .

$$(A-1)\frac{\hat{\sigma}^2_{B_{PR}}(n)}{\sigma^2_{B_{PR}}(n)} \sim \chi^2(A-1) \tag{2}$$

$$\hat{\sigma}_{B_{PR}}(m,n) = \sqrt{\frac{1}{A-1}\sum_{k=1}^{A}\left\{\left[B_{PR,k}(m,n)-\mu_{B_{PR}}\right]^2\right\}} \tag{3}$$

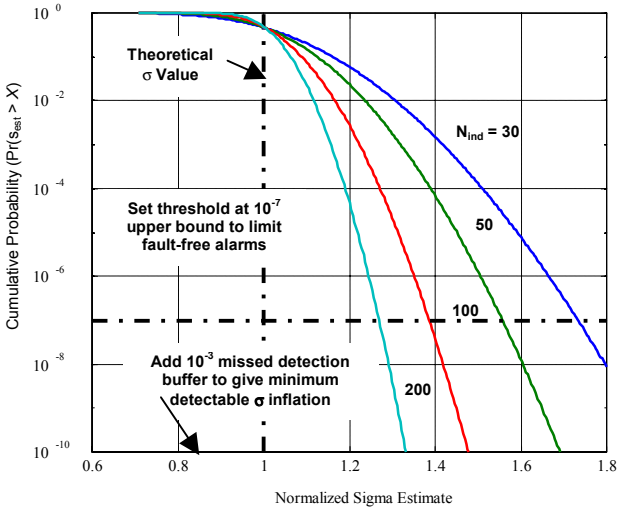

Figure 1: Chi-Square Distribution of Sigma Estimate

Figure 1 shows the resulting cumulative distributions for varying numbers of independent samples. As expected, more samples provide tighter distributions on $\hat{\sigma}_B$. The detection threshold is set based on this chi-square distribution to provide an acceptably low fault-free alarm rate (10^{-7}, based on a sub-allocation of the specified Category I continuity risk allowed per 15-second interval) [2,8]. The estimated sigmas are compared to this time-dependent threshold, which is lowered as more

independent samples are collected. Any alerts are passed on to the Executive Monitoring Phase 2 function (EXM-II) for resolution.

2.2 Cumulative Sum Method

Cumulative Sum (CUSUM) monitoring is very simple and is relatively easy to analyze, and it can be shown to be "optimal" in terms of minimizing time-to-alert under specified failure conditions [16]. It is thus commonly used in manufacturing, where the goal is to detect poor-quality products with reasonably low missed-detection and false-alarm rates (but nowhere near as low as required by the LGF). The idea is to maintain running sums of statistically-independent quality metrics with approximately known distributions under nominal conditions. A 'windowing factor' k is subtracted from the running sum at each update. This factor is chosen to minimize the time-to-alert for a particular failure case with a specific out-of-control distribution. If the targeted fault case is a large deviation from nominal, k will be large as well to reduce the sensitivity of the CUSUM to smaller anomalies. If the targeted fault is closer to nominal performance, k gets smaller, but the price is more fault-free alerts unless the CUSUM threshold h is increased to compensate.

As implemented in the IMT, the CUSUM method collects cumulative sums (C_n^+) of squared and normalized B-values (Y_n) for each receiver channel tracking a GPS satellite and is updated every 200 seconds. Note that updates must be statistically independent in time, as incrementing a CUSUM with highly-correlated inputs greater than the k factor will cause it to quickly exceed the threshold as similar values are added one after another. In this case, each independent epoch corresponds to two carrier-smoothing time intervals. The CUSUM starts at zero or a head-start value of $H^+ > 0$ and then increments each epoch by the size of the monitored input Y minus the desired 'failure slope' k that is based on a target out of control sigma (σ_1) that represents 'failed' performance. Since the CUSUM is sensitive to only one direction, separate upward "+" and downward "–" CUSUMs are generally used, although this is not true if sigma is the input, as decreasing sigma is not a concern.

$$Y_n = \left(\frac{B_{PR} - \mu_{0,B}}{\sigma_{0,B}} \right)^2 \tag{4}$$

$$\begin{aligned} C_0^+ &= 0 \quad (\text{or use 'head start': } H^+) \\ C_n^+ &= \max\left(0,\ C_{n-1}^+ + Y_n - k\right) \end{aligned} \tag{5}$$

$$k_{\text{sigma}}^+ = -\frac{\ln(\sigma_0) - \ln(\sigma_1)}{\left(2\sigma_1^2\right)^{-1} - \left(2\sigma_0^2\right)^{-1}} \tag{6}$$

If the CUSUM falls below zero on a given epoch, it is reset to zero. If the sum is above zero at any update epoch, the CUSUM is compared to a fixed threshold (h) that does not vary with time. If it accumulates to above the threshold, an alert is issued.

While CUSUM behavior is more complicated than sigma estimation, the fact that CUSUMs follow the Markov property (since a running sum is incremented every epoch, the distribution of the CUSUM state at epoch i depends only on its state at the previous epoch i–1 and the distribution of the incrementing value at epoch i) makes them straightforward to analyze via Markov Chains (MCs) [16]. Under either nominal or specified failure conditions, a "one-step" MC transition matrix **P** can be derived to give the probability of going from each discretized CUSUM state between zero and the threshold h on epoch i to each possible state on epoch i+1. From **P**, one can compute the steady-state distribution of the MC and thus determine how long, on average, it takes for the MC to exceed a given value (for the threshold h, this gives the average run length (ARL)). In addition, multiplying **P** by itself d times (e.g., computing $\mathbf{P}^d$) gives the transition probabilities between epoch i and epoch i+d, which allows one to determine, from the failure-state MC, the number of epochs required to exceed the threshold with a given missed-detection probability.

The normal solution process is to guess a threshold value h, form the nominal MC, and then solve for the nominal ARL, iterating on h until the nominal ARL is the inverse of the desired fault-free alert rate. Then, changing the MC to represent the failed or "out-of-control" state, compute the out-of-control ARL (or mean time to detect), and successively multiply the **P** matrix by itself until it shows a probability of exceeding the threshold that is one minus the desired missed-detection probability [2,16].

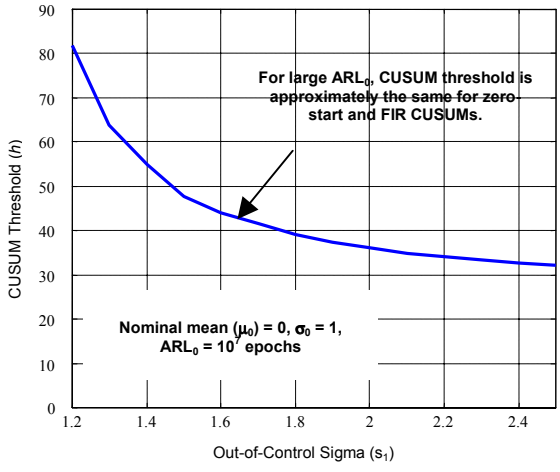

Figure 2: Thresholds for CUSUM Sigma Monitor

2.3 *Comparison of Methods*

As noted above, modeling the CUSUM as a Markov chain allows us to determine the probability of exceeding the threshold h under failure conditions at any future epoch and thus to determine how soon the failure is detected with a probability of 0.999, or, conversely, a missed-detection probability (P_{MD}) ≤ 0.001 [1,2]. For sigma estimation, to find the sigma increase that can be detected with a desired P_{MD}, an additional buffer is added to the thresholds in Figure 1. This buffer is determined by moving the mean of the distribution to a trial value above the detection threshold and moving leftward to determine the probability of falling below the threshold until the value is found that makes this probability equal to 0.001 [10].

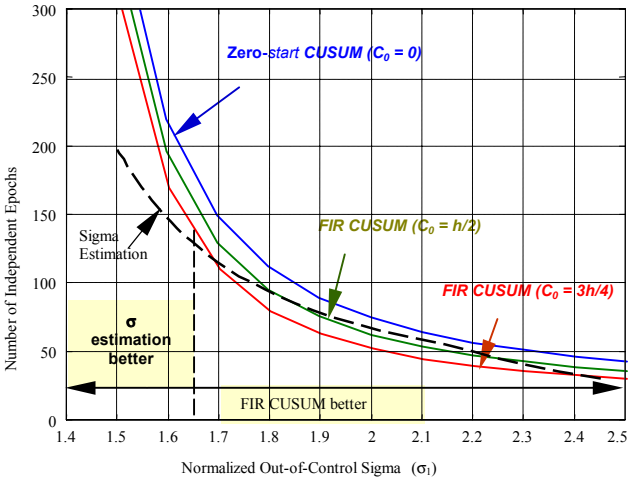

Figure 3: Time-to-Alert for CUSUM and Sigma Estimation Monitors

Figure 3 compares the times-to-detect with P_{MD} ≤ 0.001 for these three CUSUMs and the sigma estimation monitor based on potential sigma violations as a function of the normalized out-of-control sigma, which is the ratio between the actual sigma and the theoretical sigma. The results show that sigma estimation is best for $\sigma_1 < 1.68$, but the FIR CUSUM $H^+ = 3h/4$ is superior for higher sigma values, which is important because larger sigma violations lead to larger integrity threats. The FIR CUSUM achieves faster detection by initializing the CUSUM to a non-zero value closer to the threshold every time the CUSUM resets (goes below zero). The CUSUM has an additional advantage not represented in this plot: CUSUM monitoring begins immediately, whereas sigma estimation in the IMT requires that 18 independent epochs be observed before threshold checks can begin (before then, the sigma estimate is too unreliable to be compared to a chi-square-based threshold).

Note that the existing MRCC screen of IMT B-values on an epoch-by-epoch basis (comparing B-values at each epoch to fixed thresholds) has some utility as a sigma monitor, but its times-to-alert are much longer than those for the CUSUM or for sigma estimation. For example, if $\sigma_1 = 2$, Figure 3 shows that the FIR CUSUM time-to-alert with 99.9% probability is about 60 epochs. For MRCC, this anomaly translates into an actual normalized B-value threshold of 2.8, or half of the specified normalized threshold of 5.6 [8,17]. The probability of a normalized B-value magnitude exceeding 2.8 is about 0.0051 per independent epoch; thus the mean time-to-alert is approximately 1/0.0051 ≅ 196 epochs, and the time-to-alert with P_{MD} = 0.001 is the solution for x to $(1 - 0.0051)^x = 0.001$, which is $x \cong 1351$ epochs.

3.0 NOMINAL TESTING

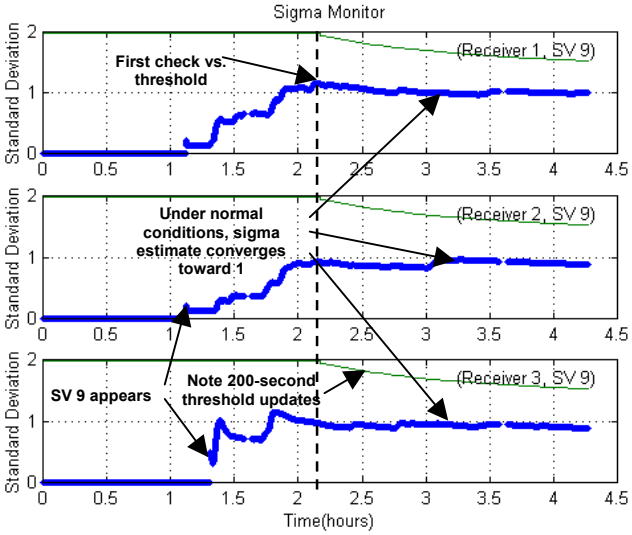

Figure 4: Sigma Estimation Results from IMT Nominal Data

In order to test these sigma monitor algorithms, sigma estimation and CUSUMs are both implemented in the Stanford IMT. Figure 4 shows the results of applying the sigma estimation algorithm to the IMT under nominal conditions. The dark curves show the normalized sigma estimate of a satellite (SV 9) on three reference receivers, and the light curves show the detection thresholds, which get smaller over time as the number of independent samples increases. As mentioned above, monitoring starts after 18 independent B-values have been collected, which corresponds to one hour with a 200-second interval between independent updates. The normalized sigma estimates stay well below the detection thresholds and converge toward 1 over time. Thus, the theoretical sigmas, which depend on the satellite elevation angle, appear to be good estimates. Very similar results have been obtained from the other satellites in this and other IMT datasets.

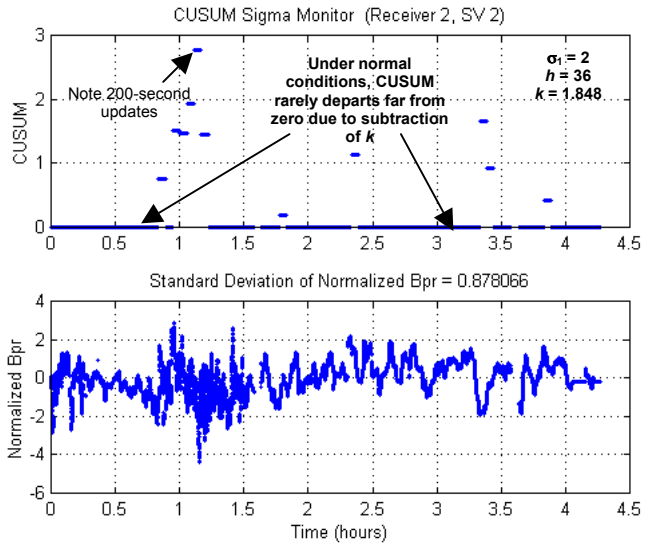

Figure 5: Zero-Start CUSUM Results from IMT Nominal Data

The 'zero-start' CUSUM and FIR CUSUM variants have also been tested with the same IMT data under nominal conditions. The top plot in Figure 5 displays zero-start CUSUM results for satellite 2 and IMT reference receiver (RR) 2, and the lower plot shows the normalized B-values from (1) that fed the CUSUM. The CUSUM in this case is targeted at an out-of-control sigma twice that of the theoretical sigma ($\sigma_1 = 2$), which gives a high windowing factor ($k = 1.848$). The CUSUM rarely departs far from zero due to subtraction of k at each independent B^2 update and stays well below the threshold (h) of 36.

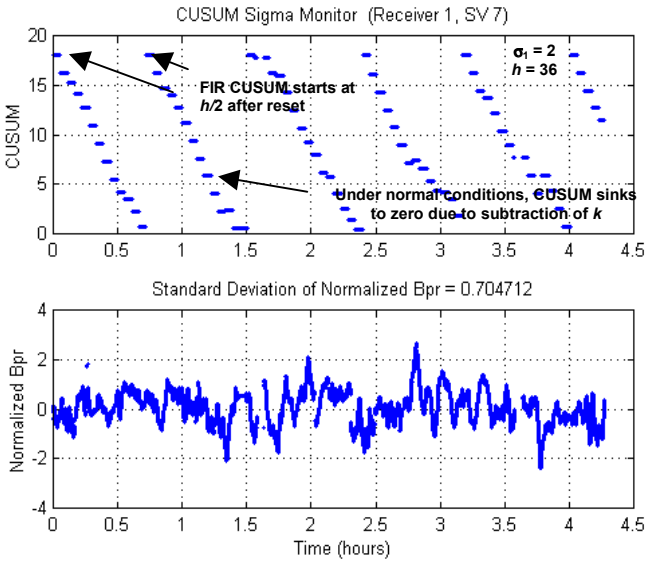

Figure 6: FIR CUSUM Results from IMT Nominal Data

The FIR CUSUM result of satellite 7 and IMT receiver 1 is shown on the top plot of Figure 6. In this case, the CUSUM is initialized at $h/2 = 18$ and is reset there every time the CUSUM falls below zero. Recall that the CUSUM is updated every 200 seconds so that successive updates are statistically independent. Under nominal conditions, the CUSUM slowly falls toward zero and is then reset, since the normalized B^2 is usually below k and k is subtracted off at each epoch. The other satellites tracked by this IMT dataset show very similar patterns for both zero-start CUSUM and head-start CUSUMs. The threshold of 36 is never threatened, and no flags are observed at all.

4.0 FAILURE TESTING

In failure testing, controlled errors are injected into IMT to test the detection of anomalies with the current sigma monitoring algorithms. Inserting errors into stored nominal receiver packets previously collected by the IMT antennas induces sigma violations. Modified nominal IMT data represent an increased-pseudorange-sigma anomaly using code-minus-carrier method.

$$PR_{raw}(\text{failed}) = (PR_{raw}\text{-minus-ADR}) \times (L - 1) + PR_{raw}(\text{nominal}) \quad (7)$$

This equation uses pseudorange-minus-carrier measurements (with a polynomial fit to ionosphere divergence removed) to roughly estimate the nominal error in PR_{raw} [18]. By adding $(L-1)$ times this error estimate to the nominal measured PR_{raw}, the nominal PR error sigma is approximately increased to L times the previous value (stored PR_{raw} values are modified and input back into the IMT in post-processing mode) [10].

Figure 7 shows the results of applying the sigma estimation algorithm under failure conditions. The PR error on a single satellite (SV 2) and a single receiver (RR 2) is increased to $L = 3$ times the nominal error. The dark line of the second plot in Figure 7 shows the normalized sigma estimate, which exceeds the detection threshold through the whole IMT run. For the purpose of this test, Executive Monitoring (EXM) logic is not active for sigma monitoring flags, such that sigma values are estimated over time without a reset. After integrating sigma monitor flags into EXM, the flagged measurement will be excluded by EXM, and its sigma value will be reset upon recovery of the measurement. EXM fault-isolation logic has been tested and clearly demonstrated by prior and current work [3,6]. Since the first check starts when 18 independent B-values have been collected, sigma estimates for RR 1 and RR 3 do not exceed the threshold over time (the initial transient is ignored – threshold checks must wait until enough independent samples have been collected for the sigma estimates to be reliable). However, in addition to the sigma estimates of the error-injected satellite and receiver, the sigma estimates of

nominal RR 1 and RR 3 also converge to values over 1 due to the fact that the B-values are correlated across the three receivers (only two of the 3 B-values for a given satellite are independent).

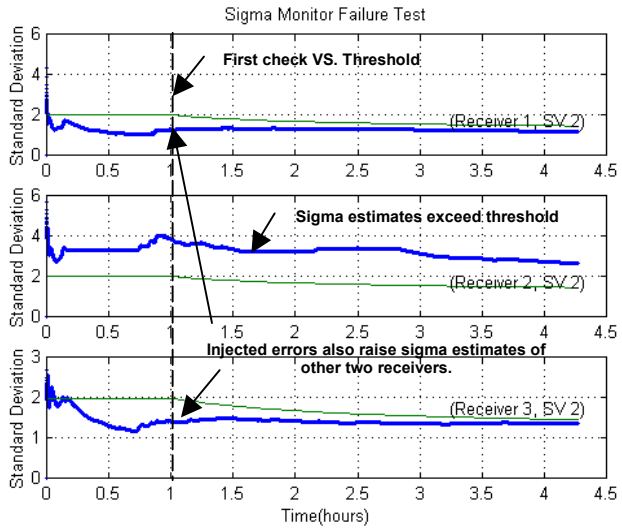

Figure 7: Sigma Estimation Results from Failure Test

IMT sigma failure test has been done with various inflated sigma values. Setting the increased error factor L to 1.7 on the same satellite (SV 2) and IMT RR 2, the sigma estimates remain just under thresholds, meaning that no flag is issued before the end of the run. Sigma estimates on this satellite for the other two receivers appear essentially nominal.

Figure 8 shows the result of applying the FIR CUSUM variant to failure-injected IMT data for the same satellite shown in the previous plot (SV 2) and IMT reference receiver 2. The CUSUM is targeted at an out-of-control sigma twice that of the theoretical sigma (σ_I = 2), which gives the same windowing factor (k = 1.848) as the previous nominal case. Based on this windowing factor and ARL, the detection threshold (h) was determined to be 36. The FIR CUSUM, initialized at $h/2$ = 18, adds up the increased normalized B-values due to severe error factor L = 3. The CUSUM crosses the threshold on the third epoch, which corresponds to 600 seconds (10 minutes) after fault injection. This is much faster than sigma estimation for a newly-rising satellite because of the 1-hour delay before sigma-estimate threshold checks can be made. Since EXM is not active in order to demonstrate how CUSUM responds with respect to injected failures, SV 2 and RR 2 are not excluded, and the CUSUM continually grows regardless of subtraction of the windowing factor at each independent B^2 update. The flat line on the lower plot indicates that the normalized B-values are isolated by the Multiple Reference Consistency Check (MRCC) at this point (EXM is not active on this check). In failure tests like this one where severe errors are injected, very similar results have been obtained from the other satellites tracked by the IMT in this dataset.

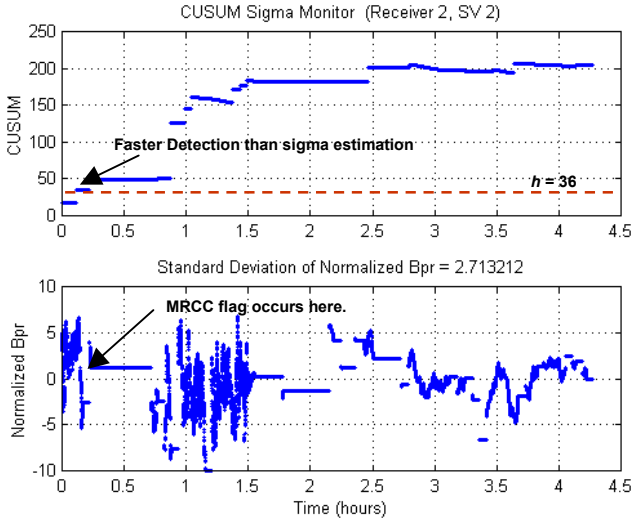

Figure 8: FIR CUSUM Results from IMT Failure Test

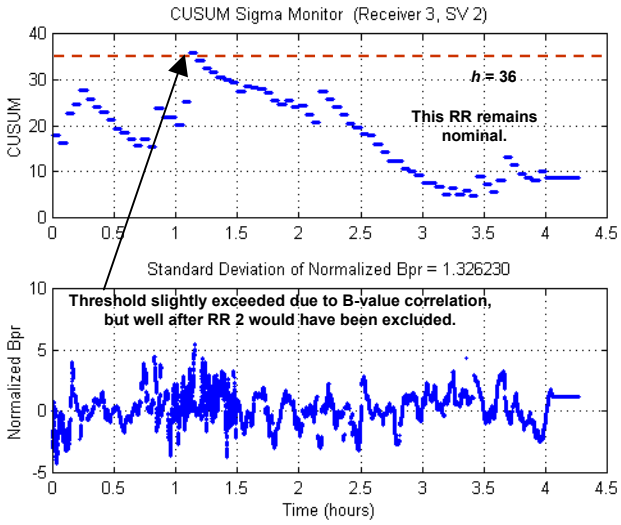

Figure 9: FIR CUSUM Results of Nominal RR from IMT Failure Test

The top plot in Figure 9 shows the FIR CUSUM of the nominal reference receiver (RR 3) and the same satellite (SV 2) affected by injected PR errors on RR 2 and SV 2. In this case, the FIR CUSUM slightly exceeds the threshold (h = 36) at epoch 21 due to B-value correlation among three reference receivers and then declines toward a value near zero. This does not occur when EXM is implemented because the source of the failure (SV 2 on RR 2) is excluded after 10 minutes, and following this exclusion, the B-values on RR 1 and RR 3 return to normal.

IMT CUSUM failure testing also has been done with various inflated sigma values. With a moderate error factor of L = 1.7 on the same satellite (SV 2) and IMT

reference receiver 2, the FIR CUSUM exceeds the threshold on the 21st epoch (1.1 hours after the CUSUM starts). Neither CUSUM nor MRCC flags appear on the non-failed receivers (RR 1 and RR 3). In the $L = 1.4$ case, neither sigma estimation nor CUSUM detects any violation. This fault is too small to be reliably detected during the 4-hour IMT run used here, as is predicted by the theoretical result in Figure 3.

Overall, the CUSUM times-to-detect are much shorter (typically well under one hour) for large anomalies than those of the sigma estimation method, which requires waiting one hour before 18 independent samples are collected. Moreover, a FIR CUSUM with a "head start" at $3h/4$ would detect violations quicker than a FIR CUSUM with initialization at $h/2$ under the same failure conditions but is slightly less robust under fault-free conditions. We have also tested multiple CUSUM monitors tuned to target different values of normalized out-of-control sigma ($\sigma_1 = 1.7$ and 2.3), but these do not improve the time-to-detect measurably over a single CUSUM with $\sigma_1 = 2$. A subset of these failure tests have been rerun after integration with IMT EXM-II logic, and these tests confirmed that the IMT properly excludes measurements that triggered CUSUM and/or sigma estimation alerts.

5.0 MEAN MONITORING

Similar to sigma monitoring, real-time mean monitoring is required to detect possible protection-level violations due to non-zero means of the true pseudorange-correction errors. As with sigma monitoring, a common approach is to estimate real-time sample means. In this respect, the mean estimation method has been analyzed and compared with a mean CUSUM algorithm, which proved to be a better way to detect severe sigma violations.

5.1 Mean Estimation Method

The mean estimation method derives sample means from LGF B-values for each visible satellite. As with sigma estimation, normalized B-values from (1) are the inputs as for sigma estimation. The results are compared to time-dependent thresholds that are set based on the normal distribution of the sample mean ($\hat{\mu}_{B_{PR}}$) as a function of the number of independent measurements (A).

$$\hat{\mu}_{B_{PR}}(n,i) \sim Normal\left(\mu_{B_{PR}}, \frac{\sigma_{B_{PR}}}{\sqrt{A}}\right) \quad (8)$$

Note that using the B-values as inputs to mean monitoring (both for estimation and CUSUM) limits the observability of non-zero means to cases where mean violations occur on only one reference receiver. A common-mode failure that causes the same non-zero mean to occur on all three receivers is not observable from B-values and must be made extremely improbable to meet the LGF integrity allocation to multiple-receiver failures [17].

5.2 Mean Cumulative Sum Method

The input (Y_n) for the mean CUSUM is the normalized B_{PR}, whereas the normalized B_{PR}^2 is the input for the sigma CUSUM.

$$Y_n = \frac{B_{PR} - \mu_{0,B}}{\sigma_{0,B}} \quad (9)$$

A windowing parameter (k) is subtracted from the normalized B-values based on the target out of control mean (μ_1) that represents 'failed' performance, and the result is reset if the resulting sum falls below zero.

$$k^+_{\text{mean}} = \frac{\mu_0 + \mu_1}{2} \quad (10)$$

Note that, unlike sigma violations, threatening mean violations can be either positive or negative; thus two parallel CUSUMs (C_n^+ and C_n^-) are needed for each measurement so that violations in either direction will be detected.

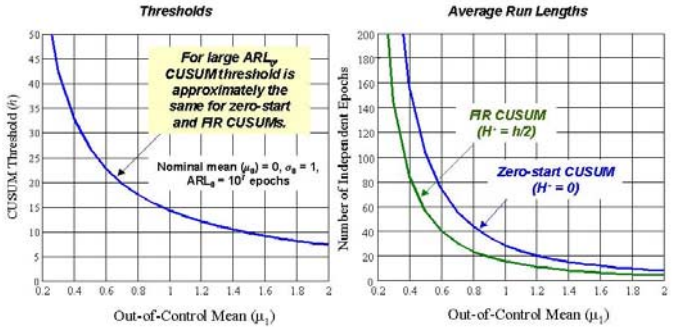

Figure 10: Thresholds and Failure-State ARLs for CUSUM Mean Monitor

The CUSUM threshold (h) is found by numerical search to match the desired average run length (ARL) and k value, which is derived in equation (10) given the target out-of-control mean (μ_1). As a function of μ_1, Figure 10 shows the resulting thresholds to achieve ARL = 10^7 independent epochs for the CUSUM mean monitor on the right and the resulting failure-state ARLs for the zero-start ($H^+ = 0$) and FIR ($H^+ = h/2$) CUSUMs on the left. The out-of-control ARL for the FIR CUSUM is significantly better than the zero-start CUSUM; and as with the sigma CUSUM, it is possible to increase H^+ beyond $h/2$ to decrease detection time further with little nominal ARL penalty. Again, the nominal thresholds for these two CUSUMs are essentially the same due to the fact that very

large thresholds needed to achieve ARL = 10^7 under nominal conditions.

Figure 11 compares the performance of mean estimation to that of several varieties of CUSUM monitoring based on the time to detect potential non-zero means with a missed-detection probability of 0.001 or below as a function of the out-of-control Mean. The results show that FIR CUSUM methods are superior to the mean estimation method in detecting any mean violations. Note that the zero-start CUSUM is slightly worse than mean estimation, but the *h*/2 FIR CUSUM is significantly better for all μ_1 and can be improved further (with slight loss of robustness under fault-free conditions) by increasing H^+ above *h*/2. Both CUSUM and mean estimation methods detect larger mean violations almost simultaneously, though the FIR CUSUM with a "head start" *3h/4* achieves faster detection.

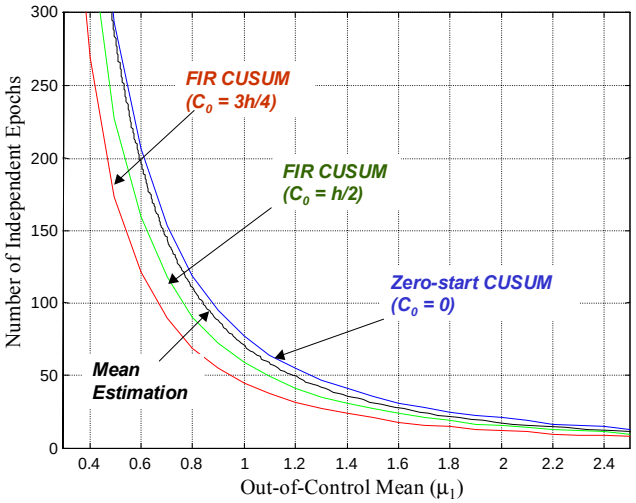

Figure 11: Mean Estimation and CUSUM Monitor Performance

5.3 Mean Monitoring Test Results

5.3.1 Nominal Test Results

Although the analytical results of mean monitoring algorithms show that FIR CUSUMs are always superior to mean estimation, both approaches are implemented in the Stanford IMT for the purpose of testing the theory. Figure 12 shows the results of applying the mean estimation algorithm to the IMT under nominal conditions. A single satellite (SV 2) is shown here on all three IMT reference receivers. The dark curves show the normalized mean estimate, and the light curves show the detection thresholds (monitoring starts after 6 independent B-values have been collected, or after 20 minutes). Note that the normalized mean estimates stay well below the detection thresholds (which get smaller over time as the number of independent samples increases) and that the mean estimates converge over time toward zero. Mean estimation restarts when there is no B-value due to the GPS satellite 'setting' and lock being lost due to the elevation angle dropping too low.

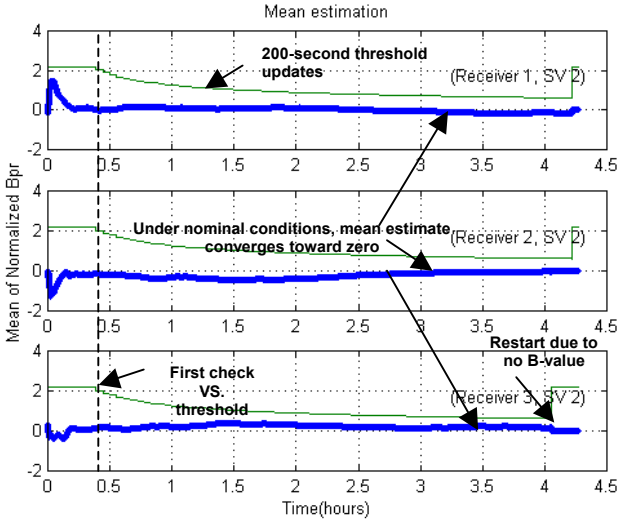

Figure 12: Mean Estimation Results from IMT Nominal Data

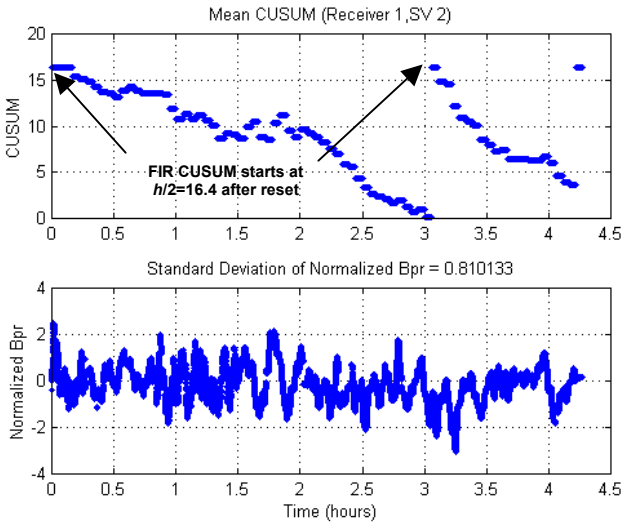

Figure 13: Mean FIR CUSUM Results from IMT Nominal Data

The top plot in Figure 13 shows the result of applying the Mean FIR CUSUM to the same IMT data shown in the previous plot, and the lower plot shows the normalized B-values that fed the CUSUM. The Mean CUSUM in this case is tuned to an out-of-control mean μ_1 = 0.4, which gives a windowing factor k = 0.2. Based on the desired average run length (ARL) and k value, the Mean CUSUM threshold (h = 32.85) is computed by the Markov Chain method using numerical search. As before, the CUSUM is updated every 200 seconds, which makes each update statistically independent. In this case, the CUSUM is initialized at $h/2$ = 16.4 and is reset there every time the CUSUM falls below zero. As shown in the sigma CUSUM cases, under nominal conditions, B^2 is usually

below k; thus the CUSUM slowly falls toward zero and is then reset. The same thing occurs to the negative CUSUM, which is not shown. Again, the other satellites included in this IMT dataset show very similar results, and no flags are generated by either mean estimation or mean CUSUM methods.

5.3.2 Failure Test Results

Both mean estimation and mean CUSUMs have been tested under failure conditions to verify the capability of mean monitoring to detect threatening anomalies. Controlled bias errors are inserted into stored nominal receiver packets previously collected by the IMT antennas, which induces mean violations. The bias added to PR_{raw} at every epoch is pre-selected to be L times the nominal sigma of the error in PR_{raw}, and stored PR_{raw} values with this bias added are input back into the IMT in post-processing mode.

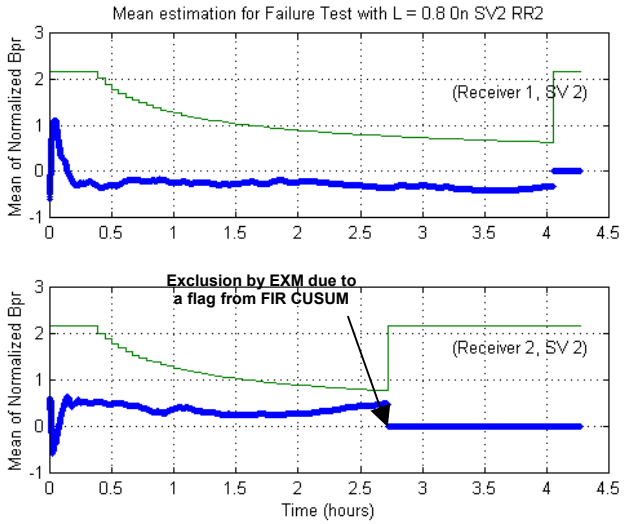

Figure 14: Mean Estimation Results from IMT Failure Test

IMT mean monitoring has been tested with three different non-zero mean values (L = 0.4, 0.8, and 1.2) with EXM-II included (so that flagged measurements are removed). Figure 14 shows the results of applying the mean estimation algorithm under failure conditions. The PR error mean on a single satellite (SV 2) and a single receiver (RR 2) is increased to 0.8 times the nominal error. Note that even though mean estimate of normalized B-values (the dark line of the lower plot in Figure 14) does not exceed the detection threshold, it is set to zero and starts again at 2.72 hours. This reset time exactly matches the time that the CUSUM in Figure 15 (C = 33.17) exceeds the threshold (h = 32.85). Since EXM logic is active, the mean estimate value is reset to zero at the same time as the CUSUM is reset to its initialization value after being excluded by EXM. The top plot in Figure 15 shows that the FIR mean CUSUM of the nominal reference receiver (RR 1) and the same satellite (SV 2) is not affected much by the injected PR error on RR 2 and SV 2. Since the bias is in the positive direction, no flags are seen in the negative CUSUM.

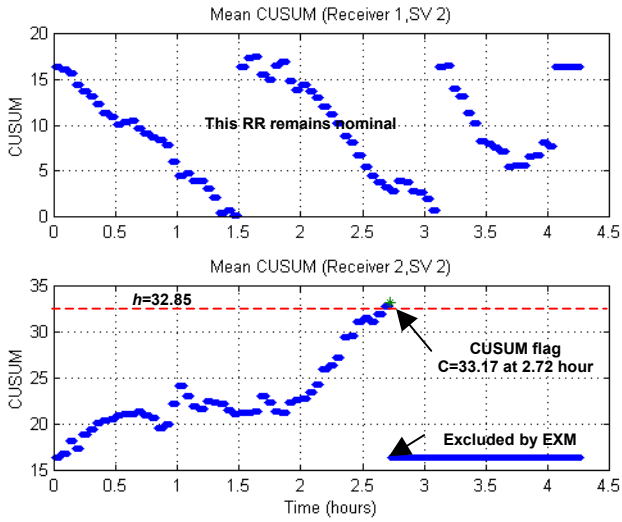

Figure 15: Mean FIR CUSUM Results from IMT Failure Test

When L is lowered to 0.4 on the same satellite (SV 2) and IMT RR 2, no flag is generated by either mean estimation or CUSUM methods, which matches the predicted performance in Figure 11. With a severe mean error with L = 1.2 on the same SV 2 and RR 2, the positive FIR CUSUM exceeds its threshold 1.65 hours after the fault was injected (and when the CUSUM was started). Neither CUSUM nor mean flags appear on the non-failed receivers (RR 1 and RR 3).

6.0 SUMMARY

This paper demonstrates that the Sigma-Mean monitor has been successfully implemented in the Stanford LAAS Integrity Monitor Testbed prototype and can detect threatening mean or sigma violations such that it provides navigation integrity to users. It also summarizes the well-known algorithms for mean and sigma estimation, and analyzes the CUSUM algorithms for mean and sigma monitoring. The test results of both methods under nominal and failure conditions generally agree with analytical predictions. The Sigma-Mean monitor has been smoothly integrated with EXM logic within the Stanford IMT and accomplishes removal of single-channel anomalies, allowing other nominal measurements to continue to be used.

We found that, in most cases, FIR CUSUMs are superior to mean and sigma estimation, although sigma estimation should still be used to detect relatively small sigma violations. Further improvement of FIR CUSUM performance is possible with higher 'head-start' (H^+).

However this causes the fault-free alarm rate to increase. Given this trade-off, the optimum head start is yet to be determined, but the $h/2$ head start implemented in the IMT is a reasonable compromise. While CUSUM *mean* times-to-detect are well under one hour for large violations. However, the time-to-detection with $P_{MD} \leq 0.001$ is somewhat longer.

Since small mean and sigma violations are not detectable, some inflation of the nominal sigma (roughly a factor of 1.5 – 1.7) is needed to provide margin for sigma-mean monitoring so that anomalies too small to be detected are not hazardous to users. With this amount of inflation, the performance achieved by the CUSUMs appears good enough to meet the LGF Specification requirement that "non-minimal-risk" anomalies (those that cause the computed protection levels to be well below reality) be detected within one to three hours with 99.9% reliability, while the mean detection times will be typically under one hour. If faster times-to-detect are desired, additional sigma inflation could be implemented, but "diminishing returns" applies above an inflation factor of 2.0 because the CUSUM time-to-detect does not improve much further.

ACKNOWLEDGEMENTS

The constructive comments and advice regarding this work provided by many other people in the Stanford GPS research group are greatly appreciated. The authors gratefully acknowledge the Federal Aviation Administration Satellite Navigation LAAS Program Office (AND-710) for supporting this research. The opinions discussed here are those of the authors and do not necessarily represent those of the FAA or other affiliated agencies.

REFERENCES

[1] B. Pervan, *et.al*., "Sigma Estimation, Inflation, and Monitoring in the LAAS Ground System," *Proceedings of ION GPS 2000*. Salt Lake City, UT., Sept. 19-22, 2000.

[2] S. Pullen, *et.al*., "The Use of CUSUMs to Validate Protection Level Overbounds for Ground Based and Space Based Augmentations Systems," *Proceedings of ISPA 2000*. Munich, Germany, 18-20 July 2000.

[3] M. Luo, S. Pullen, et.al., "Development and Testing of the Stanford LAAS Ground Facility Prototype," *Proceedings of ION NTM 2000*. Anaheim, CA., 26-28 January 2000, pp.210-219.

[4] B. Pervan and I. Sayim, "Issues and Results Concerning The LAAS σ_{pr_gnd} Overbound," IEEE PLANS 2000, San Diego, CA, 13-16 March 2000.

[5] R. Braff, "Description of the FAA's Local Area Augmentation System (LAAS)," *Navigation*. Vol. 44, No. 4, Winter 1997-1998, pp. 411-424.

[6] G. Xie, S. Pullen, *et.al*., "Integrity Design and Updated Test Results for the Stanford LAAS Integrity Monitor Testbed,", ION NTM 2001, Albuquerque, N.M., June 11-13, 2001.

[6] B. DeCleene, "Proof for ICAO Overbounding Requirement," Paper distributed to ICAO Overbounding Subgroup, 2 December 1999.

[7] RTCA (SC-159/WG-4), *Minimum Aviation System Performance Standards for the Local Area Augmentation System (LAAS)*. Washington D.C., RTCA/DO-245, 28 September 1998.

[8] *Specification: Performance Type One Local Area Augmentation System Ground Facility*. U.S. Federal Aviation Administration, Washington, D.C., FAA-E-2937, 21 September 1999.

[9] "FAA LAAS Ground Facility (LGF) Functions," Version 2.4. LAAS KTA Group, Unpublished Manuscript, September 9, 1998.

[10] S. Pullen, M. Luo, *et.al.*, "GBAS Validation Methodology and Test Results from the Stanford LAAS Integrity Monitor Testbed,", *ION GPS 2000*, 19-22 September 2000, Salt Lake City, UT, pp. 1191-1201.

[11] R. Braff, C. Shively, "An Overbound Concept for Pseudorange Error from the LAAS Ground Facility," *Proceedings of IAIN World Congress/ION 56th Annual Meeting*. San Diego, CA., June 26-28, 2000.

[12] B. Pervan, S. Pullen, M. Luo, "Navigation Integrity Monitoring for DGPS-Based Precision Landing of Aircraft," *Proceedings of the 1999 American Control Conference*. San Diego, CA., June 2-4, 1999.

[13] *Minimum Operational Performance Standards for GPS/Local Area Augmentation System Airborne Equipment*. Washington, D.C., RTCA/DO-253, SC-159, WG-4A, Jan. 11, 2000.

[14] P. Enge, "Local Area Augmentation of GPS for the Precision Approach of Aircraft,", Proceedings of the IEEE, Vol. 87, No. 1, January 1999

[15] M.R. Spiegel, *Theory and Problems in Probability and Statistics*. New York: The McGraw-Hill Book Co. (Schaum's Outline Series), 1975.

[16] D.M. Hawkins, D.H. Olwell, *Cumulative Sum Charts and Charting for Quality Improvement*. New York: Springer-Verlag, Inc., 1998.

[17] *Specification: Category I Local Area Augmentation System Non-Federal Ground Facility*. U.S. Federal Aviation Administration, Washington, D.C., FAA/AND710-2937, May 31, 2001.

[18] M. Luo, S. Pullen, "LAAS Reference Receiver Correlation Analysis," Stanford University, Unpublished Manuscript, Feb. 17, 1999.