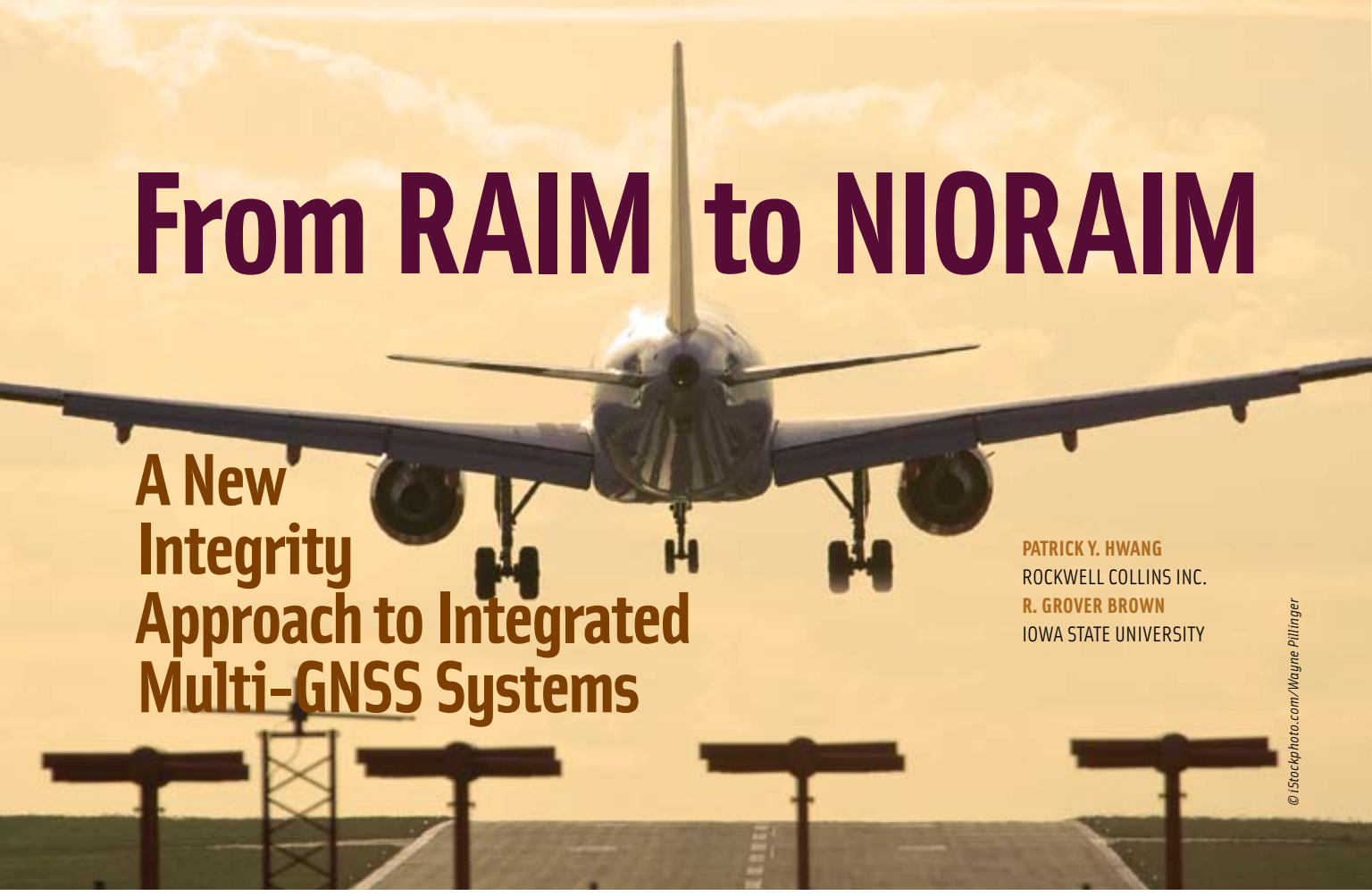
From RAIM to NIORAIM

A New Integrity Approach to Integrated Multi-GNSS Systems

PATRICK Y. HWANG
ROCKWELL COLLINS INC.
R. GROVER BROWN
IOWA STATE UNIVERSITY



As the GNSS world becomes one in which multiple systems are used simultaneously, the matter of detecting erroneous signals from satellites brings both new risks and opportunities. In safety-critical applications such as civil aviation, improved availability of integrity — typically achieved using receiver autonomous integrity monitoring (RAIM) — would help optimize approach and landing operations. However, creating a methodology that yields a useful level of protection, while ensuring the timely detection of a rare unhealthy signal occurring in one of the GNSS systems, is not a trivial problem. But a new approach to integrity may provide an efficient solution to optimizing that protection level.

Exciting times lie ahead in the world of satellite navigation systems.

A vastly improved Global Positioning System will soon be available to the civil community, GLONASS is becoming more reliable with each passing year, and Galileo and Compass are on their way not far behind. Thus, in the near future navigation systems engineers will have more pseudorange measurements to work with than any of us would have dreamed of a decade ago.

Given this situation, the obvious question then arises: "What shall we do with this wealth of range data?" Irrespective of accuracy considerations, this bounty of measurement redundancy can make receiver autonomous integrity monitoring — or RAIM — a more robust means of assuring system integrity — that is, of ensuring that only "healthy" GNSS signals are used in navigation and positioning solutions.

This article will address how we might most effectively accomplish this end.

Background

How could anybody write a learned technical article on something as simple as cross-comparing two independent GNSS solutions? In safety critical situations, prudent navigators have been comparing the readings of two similar instruments against each other since antiquity.

This is, of course, a qualitative concept, and a good one!

Only when we attempt to quantify the concept in terms of strict statistical requirements and instrument accuracies does the problem becomes interesting mathematically. RAIM does exactly that. It quantifies the concept of using a self-consistency check of redundant measurements to assure navigational integrity.

In this article, we describe a new variant of RAIM called Novel Integrity Optimized RAIM or NIORAIM (pronounced "nee-oh-raim") that can be applied to enhance the availability of the consistency check. We will discuss the NIORAIM concept in detail a little later, but first let's consider the practical considerations of working with multiple GNSS signals in the integrity domain.

The availability of alternative GNSS constellations, both in the present and future, has stirred a great deal of interest towards combining measurement information for enhanced performance. In general, the more measurements there are, the better the positioning, navigation, or timing solution. For integrity monitoring, however, other subtle considerations arise over how to combine these diverse signals and pseudoranges.

From an optimal performance standpoint, one would be inclined to simply expand the standard form of RAIM (or NIORAIM) to include all satellite measurements from multiple constellations in a centralized manner. This probably affords the best results in terms of lowest integrity limits, but this approach is also limiting in that it assumes only one satellite failure at a time; multiple failures will require special accommodation that greatly increases the complexity of the solution and degrades the performance accordingly.

In an ION GNSS 2005 paper titled "GPS and Galileo with RAIM or WAAS for Vertically Guided Approaches," (see Additional Resources section at the end of this article) Y. C. Lee of Mitre and coauthors suggested that a more conservative — albeit less optimal — approach would be to do the self-consistency check at the solution level.

In describing a GPS/Galileo joint constellation, Lee proposed what is essentially a cross-compare of two independent solutions, one from GPS and the other from Galileo. In contrast to the standard, more centralized RAIM solution, Lee's proposal is a decentralized (or federated) solution. His approach allows for multiple satellite failures to occur within one system, but only one system can have a failure or failures at any one time.

Tight accuracy/integrity applications have raised concerns about the multitude of failure types that might occur, however rarely, under seemingly "normal" conditions. The decentralized approach provides a clearer and stronger partitioning between the solutions independently derived from GPS and from Galileo.

This philosophy is not without precedent. Most aircraft systems achieve high integrity by means of producing multiple independent solutions from clearly partitioned sensors. Often, this is done with some compromise in accepting the

decentralized solution as compared to the centralized one.

The purpose of this article is to pursue this decentralized approach to RAIM and analyze its performance, especially from the viewpoint of NIORAIM. In the remaining discussion, we will be using vertical position during precision approach as an example application. The numerical values used for the false alarm and missed detection rate specifications are taken from a more recent (2007) paper by Lee and M. P. McLoughlin, which is also listed in the Additional Resources section.

In our discussion, we do not intend to promote numerical values for any particular application. Rather, we consider the methodology presented here to be the important contribution of the article.

It is, however, worth mentioning that this methodology can also be applied to any two-system combination where the position parameter of interest is a scalar. For example, it would apply equally well to the along-track or cross-track integrity assurance problem (with different numerical values, of course).

RAIM and NIORAIM

RAIM came into being with the publication of RTCA/DO-208 in 1991. That document spelled out the recommendations of RTCA Special Committee 159 for using GPS as a supplemental navigation aid in U.S. civil aviation.

These recommendations were translated into practice through the Federal Aviation Administration (FAA) Technical Standard Order TSO C129, which added some extensions to the RTCA proposal. In both documents, RAIM was specified as the means for assuring system integrity. RAIM continues to this day as one of the primary means of assuring GPS integrity in civil aviation.

The mechanics of RAIM can be divided into three distinct steps:

(a) First and foremost comes the *failure detection algorithm*. This consists of forming a test statistic from the redundant measurement suite, and then testing it against some threshold criterion to determine "failure" or "no failure." Here, the algorithm must satisfy rigid statistical requirements on false alarm rate and missed detection probability in order to declare that RAIM is "available."

(b) In addition to requirement (a), RAIM must also provide to the flight crew a measure of the *integrity limit*, that is, a worst-case bound at 10^{-7} probability that is being protected by the RAIM.

For example, in the early days of RAIM when the most demanding applicable phase of flight was nonprecision approach, the horizontal position error being protected had to be calculated online, and this had to be less than the specified alert limit of 0.3 nautical miles in order for the RAIM measurement consistency check to be valid. Otherwise, RAIM had to be declared "not available". (An interesting note: the calculated protection level requirement was not included in the original RTCA/DO-208 document. This came about later.)

(c) Finally, if sufficient measurement redundancy is available, the faulty measurement must be identified and eliminated from the measurement suite within the time-to-alert specification.

We will be concentrating on the two-system, one-state RAIM problem in this article; so, we will only be concerned with steps (a) and (b). However, with a modest effort, the methods presented here could also be extended to higher-order and more complex cases.

In the simplest form of GPS RAIM, we usually assume that we have a snapshot set of linearized redundant measurement equations with three components of position and a clock bias as the unknowns. We also assume that the allowable false-alarm and missed detection rates are specified.

A suitable test statistic for failure detection can be formed in a number of ways, and the equivalences among the various methods have been discussed at some length in the literature. (See, for example, the articles by R. G. Brown and by R. S. Y. Young and G. A. McGraw, cited in Additional Resources.) We will not pursue this further here. Our approach throughout will be the parity-space method in the referenced article by R. G. Brown and G. Y. Chin.

In the usual RAIM, we choose the position solution that minimizes the mean-square error, given the measurement geometry and the assumed measurement error variances. Then, once we have decided to use the parity vector as the test statistic, the remaining problem is to compute the achieved protection level (say, in meters), and, hopefully, the calculated protection level will be less than the specified alert limit for the application at hand.

The point of all this is that once we choose to optimize on position error, we simply have to live with whatever comes out of the protection-level calculation — there is no "wiggle room" left, and the RAIM availability may not be the best that we can do under the circumstances.

In a paper published in 2006 and cited in Additional Resources, the present authors proposed an alternative approach to RAIM in which they showed that rms position accuracy can, in selected applications, be traded for reduced protection level. That is, by a judicious choice of measurement weights in the position solution, one can achieve a reduced protection level with only a modest loss in rms position accuracy.

The benefit of this is, of course, improved RAIM availability. The authors dubbed their new RAIM method, NIORAIM to distinguish it from the usual conventional RAIM.

Now, let us turn to the application of NIORAIM to the two-system scenario. This may seem to be a ridiculously simple problem at first glance. However, we will see presently that the optimization to achieve the smallest possible protection level is not at all trivial.

Parity Space Statistical Test

We begin with the rudimentary linearized measurement equation for the two-system case:

$$\underbrace{\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\mathbf{H}} x + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}}_{\boldsymbol{\varepsilon}} + \underbrace{\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}}_{\mathbf{b}} \quad (1)$$

The two measurements z_1 and z_2 are the vertical position estimates coming from two completely independent satellite navigation systems (e.g., GPS and Galileo), x is the true vertical position (scalar in this case), ε_1 and ε_2 are the respective random errors, which are assumed to be zero-mean Gaussian with known σ_1 and σ_2 (perhaps from real-time VDOP calculations), and b_1 and b_2 are bias components due to system "failures" of any sort.

We will make the usual assumption that either b_1 or b_2 can be non-zero, but not both simultaneously. (We should note that either b_1 or b_2 could be due to multiple satellite failures within their individual systems. We are only interested here in the end result as it reflects into each individual system's vertical position computation.)

The first step in the parity-space approach is to condition the measurement equations such that the random noise components have equal standard deviations, or *sigmas*. The resulting measurement equation is now

$$\underbrace{\begin{bmatrix} \lambda z_1 \\ z_2 \end{bmatrix}}_{\mathbf{z}'} = \underbrace{\begin{bmatrix} \lambda \\ 1 \end{bmatrix}}_{\mathbf{H}'} x + \underbrace{\begin{bmatrix} \lambda \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}}_{\boldsymbol{\varepsilon}'} + \underbrace{\begin{bmatrix} \lambda b_1 \\ b_2 \end{bmatrix}}_{\mathbf{b}'} \quad (2)$$

where the sigma ratio is denoted as λ, i.e.,

$$\lambda = \frac{\sigma_2}{\sigma_1} \quad (3)$$

Note that the "multiply through" operation does not change the physical measurement situation in any way.

The next step is to normalize on σ_2, i.e., let

$$\sigma_2 = 1 \quad (4)$$

Then, all distance calculations are now in σ_2 units. (This becomes a convenience later in composing a lookup table for protection level as a function of λ.)

The equations for forming the estimate $\hat{x}_{WLS}$ and the test statistic p are

$$\hat{x}_{WLS} = \left[(\mathbf{w}\mathbf{H}')^{\mathrm{T}}\,\mathbf{w}\mathbf{H}'\right]^{-1}(\mathbf{w}\mathbf{H}')^{\mathrm{T}}\,\mathbf{w}\mathbf{z}' \quad (5)$$

(Note: This is equivalent to another form $\hat{x}_{WLS} = \left[\mathbf{H}'^{\mathrm{T}}\mathbf{V}\mathbf{H}'\right]^{-1}\mathbf{H}'^{\mathrm{T}}\mathbf{V}\mathbf{z}'$ *if we choose a weighting matrix of* $\mathbf{V} = \mathbf{w}\mathbf{w}^{\mathrm{T}}$*, where* $\mathbf{w}$ *is assumed to be diagonal.)*

$$p = \mathbf{P}\mathbf{z}' \quad (6)$$

Note especially that $\hat{x}_{WLS}$ is formed as a general weighted least-squares estimate rather than the usual least-squares estimate

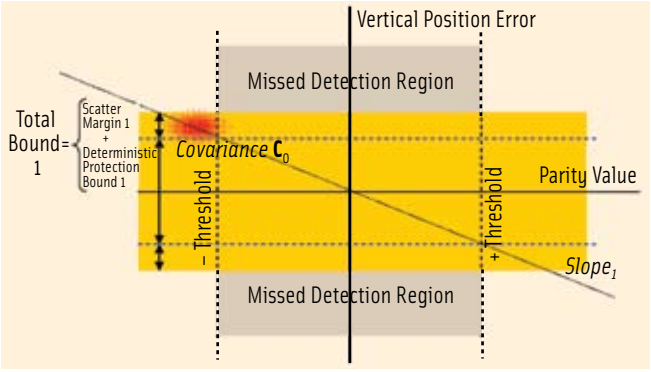


FIGURE 1 In plotting parity value (p) versus vertical position error (e), the protection bound is represented within the shaded region and combines a deterministic component as well as a smaller statistical component. (Total Bound 1 shown for only one half of a two-sided bound.)

where $\mathbf{w} = \mathbf{I}$. The weight matrix $\mathbf{w}$ that yields the best protection level will be considered later.

The $\mathbf{P}$ matrix is called the parity transformation matrix, and it takes the conditioned measurement vector $\mathbf{z}'$ into the test statistic p. It is easily verified that

$$\mathbf{P} = \begin{bmatrix} \frac{-1}{\sqrt{1+\lambda^2}} & \frac{\lambda}{\sqrt{1+\lambda^2}} \end{bmatrix} \quad (7)$$

is a satisfactory $\mathbf{P}$ matrix in this application. (The only requirements are that $\mathbf{P}$ be orthogonal to $\mathbf{H}'$ and that $\mathbf{P}$ is normalized.) The test statistic is now formed as

$$\begin{aligned} p &= \mathbf{P}\mathbf{z}' \\ &= \begin{bmatrix} \frac{-1}{\sqrt{1+\lambda^2}} & \frac{\lambda}{\sqrt{1+\lambda^2}} \end{bmatrix} \begin{bmatrix} \lambda z_1 \\ z_2 \end{bmatrix} \quad (8) \\ &= \frac{\lambda}{\sqrt{1+\lambda^2}}(z_2 - z_1) \end{aligned}$$

or, in terms of the random noises and biases

$$p = \frac{\lambda}{\sqrt{1+\lambda^2}}(\varepsilon_2 + b_2 - \varepsilon_1 - b_1) \quad (9)$$

It is comforting to find that the formal mathematical parity-space approach says that the test statistic should simply be the *difference between the two measurements*, appropriately scaled. This is what we expected from the beginning!

However, now the scale factor on the difference tells us precisely where to set the threshold to achieve the specified false alarm rate. This is routine once we know σ_1 and σ_2, and assume Gaussian statistics throughout. (Note that p is the usual scalar normal random variable *with sign*.)

Calculation of the Estimate, Estimation Error, Slopes and Vertical Integrity Limit. In the weighted least squares position estimate equation (Equation 5), we let the weighting matrix $\mathbf{w}$ be:

$$\mathbf{w} = \begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix} \quad (10)$$

Recall that the weights in weighted least squares are relative. Therefore, the 1,1-term can be arbitrarily set to unity. The 2,2-term is written as the product of positive terms α and λ, and the reason for introducing α will be apparent in a moment.

Using the definition of $\mathbf{w}$ and noting that $\mathbf{z}' = [\lambda z_1 \; z_2]^T$, $\hat{x}_{WLS}$ can now be rewritten in terms of the raw measurements z_1 and z_2 as

$$\hat{x}_{WLS} = \left(\frac{1}{1+\alpha^2}\right)z_1 + \left(\frac{\alpha^2}{1+\alpha^2}\right)z_2 \quad (11)$$

Note that the weighting coefficients of z_1 and z_2 involve only α, and not λ. This is convenient because then we can think of holding λ fixed in our parametric study and vary the weighting by adjusting the α parameter.

We are, of course, interested in the estimation error, and it can be written (recalling Equation 5) as

$$\begin{aligned} e &= \hat{x}_{WLS} - x \\ &= \underbrace{\left[(\mathbf{w}\mathbf{H}')^T \mathbf{w}\mathbf{H}'\right]^{-1}(\mathbf{w}\mathbf{H}')^T}_{\mathbf{A}_w} \mathbf{w}(\mathbf{H}'x + \boldsymbol{\varepsilon}' + \mathbf{b}') - x \\ &= \mathbf{A}_w \mathbf{w}(\boldsymbol{\varepsilon}' + \mathbf{b}') \quad (12) \end{aligned}$$

Note that e does not depend on the true value of x, and this is how it must be in the failure detection scenario.

Both p (Equation 9) and e (Equation 12) are Gaussian random variables, and their means are obtained from the slope equations and the assumed bias in the measurement domain. The slope equations are:

$$Slope_1 = \left.\frac{e}{p}\right|_{bias\ in\ 1} = \frac{-\sqrt{1+\lambda^2}}{\lambda(1+\alpha^2)} \quad (13)$$

$$Slope_2 = \left.\frac{e}{p}\right|_{bias\ in\ 2} = \frac{\alpha^2\sqrt{1+\lambda^2}}{\lambda(1+\alpha^2)} \quad (14)$$

If we put the bias b in measurement z_1,

$$\left.\begin{aligned} p_{mean} &= \begin{bmatrix} \frac{-1}{\sqrt{1+\lambda^2}} & \frac{\lambda}{\sqrt{1+\lambda^2}} \end{bmatrix} \begin{bmatrix} \lambda b \\ 0 \end{bmatrix} = -\frac{\lambda b}{\sqrt{1+\lambda^2}} \\ e_{mean} &= slope_1 \cdot p_{mean} \end{aligned}\right\} \text{Bias in } z_1 \quad (15)$$

Similar equations apply when we put the bias in z_2.

While the slopes form the deterministic component of the relationship, the statistical portion is made up of a noisy scatter with a bivariate Gaussian density function. This distribution has a covariance matrix $\mathbf{C}$ that is given by:

$$\mathbf{C} = \begin{bmatrix} 1 & \frac{\alpha^2\lambda^2 - 1}{(1+\alpha^2)\lambda\sqrt{1+\lambda^2}} \\ \frac{\alpha^2\lambda^2 - 1}{(1+\alpha^2)\lambda\sqrt{1+\lambda^2}} & \frac{1+\alpha^4\lambda^2}{\lambda^2(1+\alpha^2)^2} \end{bmatrix} \quad (16)$$

Note that the off-diagonal correlation terms in Equation 16 become zero when $\alpha\lambda = 1$, i.e., the usual RAIM scheme in which the $\mathbf{w}$ weights are uniform. Let us designate this special case of $\mathbf{C}$ with $\alpha\lambda = 1$ to be called $\mathbf{C}_0$.

Figure 1 shows a conceptual example of $\lambda = ½$ for $\mathbf{w} = \mathbf{I}$ (uniform weighting). It graphically depicts the $slope_1$ relationship between p and e, and shows a noisy scatter of data with a bivariate Gaussian distribution of covariance $\mathbf{C}_0$ and a bias in Measurement 1.

The effect of a failure of this measurement on the vertical position may be bounded by the additive combination of two components, one deterministic from the intersection of $slope_1$ with the threshold and the other an incremental statistical component that we call a *scatter margin* due to the noisy scatter described previously. (Note that the threshold is predefined by false alarm rate requirements).

Calculating the total protection bound (which includes the contribution of both aforementioned components) with reasonable precision is no cakewalk. We will refer to this total protection bound as *vertical integrity limit* or VIL (sometimes also called the *vertical protection limit*). That is, we are assured that any vertical error larger than VIL will be detected within the given specifications.

In Figure 1, the noisy scatter is shown as an elliptical smear of samples representing the Monte Carlo scatter at one particular location corresponding to a given bias. This scatter retains its shape regardless of the bias.

Conceptually, the scatter can be thought of as moving up along the slope line as we increase the bias. In doing so, it is clear that a small bias will produce very little encroachment of the scatter into the missed detection region.

An increasingly larger bias would push the scatter well into the detection region and out of the missed detection region. Thus, it is the mid-range biases that are troublesome, and we need to be able to compute the missed detection rate associated with the portion of the scatter that intrudes into the missed detection region. This can be done with numerical integration.

Fortunately, both p and e are Gaussian random variables (with sign); so, their joint probability density function can be written out explicitly, given the assumed biases and covariance matrix. Thus, numerical integration is quite feasible in this case. The necessary parameters are the biases and the covariance matrix of p and e, $\mathbf{C}$ as given in Equation 16.

Thought about another way, we can use numerical integration of the Gaussian density function to find the horizontal edge such that the probability within the missed detection region is exactly equal to the specified allowable rate.

But that is by no means the end of it!

That edge, or protection level, is valid only for that assumed bias. Because we follow a worst-case protection philosophy for integrity monitoring, this same process must be repeated over and over for different biases. The total bound is the largest of these protection levels (horizontal edges) thus computed.

Although we almost never identify what that worst-case bias actually turns out to be, we do assume that the computed protection bound has considered the worst-case bias that gave rise to the largest protection level.

Next, if we put the bias b in measurement z_2,

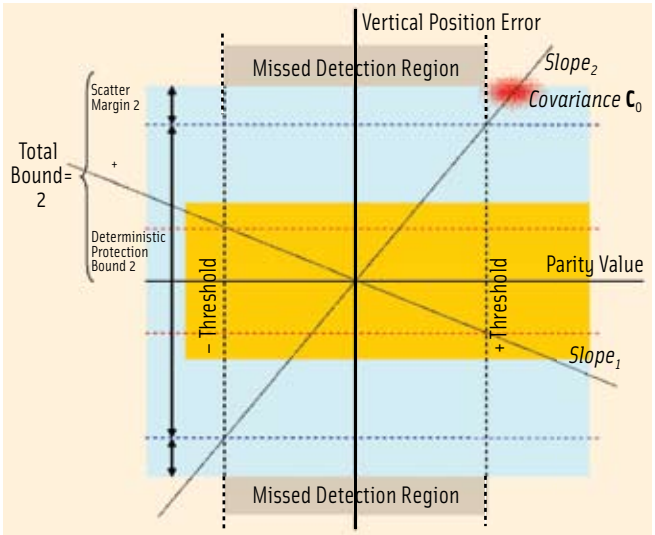


FIGURE 2 If we superimpose $Slope_2$ on the plot of $Slope_1$ from Figure 1, the total bound from Measurement 2 dominates as affording p oorer (larger) protection bound and must thus define the VIL to account for the worst case between the two measurements. (Total Bound 2 shown for only one half of two-sided bound.)

$$p_{mean} = \left[\frac{-1}{\sqrt{1+\lambda^2}} \quad \frac{\lambda}{\sqrt{1+\lambda^2}}\right]\begin{bmatrix}0\\ b\end{bmatrix} = \frac{\lambda b}{\sqrt{1+\lambda^2}} \Big\} \text{Bias in } z_2 \qquad (17)$$

$$e_{mean} = slope_2 \bullet p_{mean}$$

in **Figure 2**, we superimpose the contribution of a failure on Measurement 2 onto that of Measurement 1, while still dealing with the $\lambda = ½$ and $\mathbf{w} = \mathbf{I}$ case. Note that the protection bound with the bias on Measurement 2 is much larger than for the Measurement 1 case, primarily because $slope_2$ is much larger between the two.

Now, it may appear at first glance that all we have to do to achieve a better balance between the two protection bounds is to decrease the α from 2 to 1. From the slope equations, it can be seen that this will equalize the slopes; as we decrease α, one goes "up" and the other goes "down" until they are equal when $\alpha = 1$.

Again, at first glance, it would appear that we now have perfect symmetry in both the slopes and scatter, and that we have achieved the optimum (that is, minimum) VIL. However, the fallacy in this thinking is that as we decrease α from the $\alpha\lambda = 1$ condition, the cross-correlation term in the $\mathbf{C}$ matrix (Equation 16) becomes non-zero, and the respective scatters become skewed as shown in **Figure 3**.

This destroys the symmetry by the way that the respective scatters encroach into the missed detection region. Thus, the $\alpha = 1$ (equal slopes) condition does not, in fact, yield equal VILs! The net result is that we are not "there" yet; we must search further for the truly best α that will minimize the VIL.

Next, consider a conceptual search experiment. Suppose we pick a particular sigma ratio λ, say $\lambda = ½$, and hold it fixed throughout the experiment. Then, consider a test set of discrete

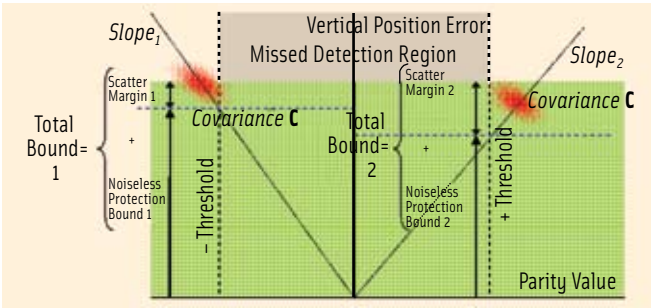


FIGURE 3 If the noisy scatter is skewed, the scatter margin depends on the sign of the statistical correlation as compared to the sign of the slope (In general, the scatter margin is smaller if the two quantities are of like sign.) Due to this asymmetry, equalizing the protection bounds does not arise from equalizing slopes and thus must be computed numerically and optimized accordingly.

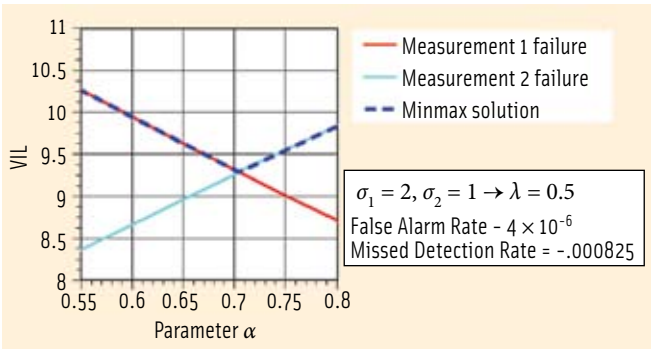


FIGURE 4 Minmax plot shows the VIL constitutes the larger of the two individual protection bounds computed for failure on each measurement. This adheres to accommodating the worst case situation.

α values of a range of interest, say $\alpha = 0.55, 0.60, 0.65, 0.70, 0.75, 0.80$. Finally, choose one of these, say $\alpha = 0.55$, and with λ and α held fixed, put the bias on measurement 1, and choose approximate test values for the bias and VIL such that the scatter is pushed into the missed detection region far enough to give a miss detection rate somewhere near the specified allowable rate.

Now do a numerical search by successive perturbations of bias and VIL until the calculated missed detection rate (by integration) matches the desired allowable value. The resulting VIL represents a worst-case VIL for the given λ and α. (As a final check, one could do a small perturbation of bias in either direction and the calculated miss detection rate should decrease for both perturbations.)

If we repeat this trial-and-error perturbation experiment for each of the other chosen values of α, it would provide sample points for the monotonically decreasing portion of the plot shown in **Figure 4**.

Next, repeat the collection of experiments over again with the bias put on Measurement 2 rather than Measurement 1. This would provide data points for the curve shown with a monotonically increasing line in Figure 4. Bear in mind that each of the points on both plots represents worst-case situations of bias for particular values of α.

The intersection of these two plots represents the item of special interest in NIORAIM. This is truly a minmax optimization problem. We look for the minimum of the maximums (worst-case bias situations), and this occurs at the intersection of the two plots. The value of α there then provides the appropriate weighting of the measurements z_1 and z_1 (via Equation 11) to yield the best possible (i.e., minimum) VIL.

With the false alarm rate set at 4×10^{-6} per sample and allowable missed detection rate set at 0.000825, the optimum values for α and VIL worked out to be

$$\alpha = 0.704$$
$$\text{VIL}_{\text{minmax}} = 9.28$$

All of this conceptual effort has provided just one entry into our lookup table. The effort must be continued for other values of λ to complete the table and a typical completed table is shown as **Table 1** for false alarm rate = 4×10^{-6} and missed detection rate = 0.000825.

The search procedure described here was admittedly crude and clumsy. This was intentional in order to emphasize the minmax nature of the optimization problem. Suffice it to say, more efficient ways exist to accomplish the search numerically (offline, of course).

The Tradeoff: Comparing NIORAIM & RAIM

We mentioned earlier that NIORAIM optimizes on the protection level, whereas conventional RAIM optimizes on rms error under normal no-failure conditions and then simply accepts whatever VIL comes out of the analysis. Of course, the price one pays for improved protection (i.e., lower VIL) is increased rms position error under normal no-failure conditions.

We will now consider the tradeoff for a nominal set of allowable missed detection rate and false-alarm rate specifications. Let's begin by defining RAIM as being one where the position estimate is obtained using the usual least-squares formula

$$\hat{x}_{LS} = \left[(\mathbf{H}')^{\mathrm{T}}\mathbf{H}'\right]^{-1}(\mathbf{H}')^{\mathrm{T}}\mathbf{z}' \qquad (18)$$

where $\mathbf{z}'$ is the conditioned measurement 2-tuple after "multiplying through" to equate the measurement sigmas. This, of course, minimizes the mean-square position error, and the linear weighting of the two measurements is set accordingly.

The estimate equation is different for NIORAIM (see Equation 5), because $\mathbf{w}$ is not equal to $\mathbf{I}$. The test statistic is another matter, though. We use the same equation for NIORAIM and RAIM, and repeat it here for convenience:

$$p = \mathbf{P}\mathbf{z}' \quad \text{(for both RAIM and NIORAIM)} \qquad (6)$$

In order to make a fair comparison between RAIM and NIORAIM, we will use the whole p and $\hat{x}$ (i.e., including sign) rather than just magnitudes as is usually done when horizontal position is the variable of interest.

So, in brief, the difference between NIORAIM and RAIM appears only in how we weight the measurements in the estimate equation, and this, in turn, affects the integrity limit calculations.

For convenience in comparison, the two estimate equations can be rewritten in terms of the raw measurements z_1 and z_2. From Equations 11 and 18:

NIORAIM:

$$\hat{x}_{WLS} = \left(\frac{1}{1+\alpha^2}\right)z_1 + \left(\frac{\alpha^2}{1+\alpha^2}\right)z_2 \qquad (11)$$

RAIM:

$$\hat{x}_{LS} = \left(\frac{\lambda^2}{1+\lambda^2}\right)z_1 + \left(\frac{1}{1+\lambda^2}\right)z_2 \qquad (19)$$

It might appear at first glance that the NIORAIM $\hat{x}_{WLS}$ does not depend on λ. It does, however, because the α in NIORAIM is a very special value that comes from the minmax solution for the best VIL, and this, in turn, involves λ.

Consider for a moment the special case where the poorer measurement of the pair (i.e., z_1) has a sigma that is twice that of z_2. Thus, in this case, λ = ½. From the table lookup of Section 4, we find that for this λ, the best α is: α = 0.704. Then, for these numerical values, the two estimate equations become

$$\hat{x}_{WLS} = 0.668\,z_1 + 0.332\,z_2 \qquad (20)$$

λ	α	VIL
1.000	1.000	5.9492
0.909	0.950	6.2526
0.833	0.907	6.5654
0.769	0.869	6.8862
0.714	0.836	7.2144
0.667	0.807	7.5485
0.625	0.782	7.8879
0.588	0.759	8.2320
0.556	0.739	8.5804
0.526	0.721	8.9314
0.500	0.704	9.2856
0.476	0.690	9.6430
0.455	0.676	10.0014
0.435	0.664	10.3630
0.417	0.653	10.7257
0.400	0.642	11.0905
0.385	0.633	11.4564
0.370	0.624	11.8238
0.357	0.616	12.1924
0.345	0.609	12.5625
0.333	0.602	12.9326
0.323	0.596	13.3038
0.313	0.588	13.6770
0.303	0.584	14.0503
0.294	0.578	14.4236
0.286	0.574	14.7990
0.278	0.569	15.1723
0.270	0.565	15.5477
0.263	0.560	15.9253
0.256	0.556	16.3007
0.250	0.553	16.6782
0.244	0.549	17.0558
0.238	0.546	17.4333
0.233	0.543	17.8108
0.227	0.540	18.1894
0.222	0.537	18.5680
0.217	0.534	18.9477
0.213	0.532	19.3273
0.208	0.529	19.7064
0.204	0.527	20.0866
0.200	0.524	20.4663

* Assumes a threshold of 4.611 meters and allowable missed detection probability of 0.000825

TABLE 1. λ, α, and VIL values for false alarm rate = 4×10^{-6} and missed detection rate = 0.000825*

$$\hat{x}_{LS} = 0.200\, z_1 + 0.800\, z_2 \qquad (21)$$

Note that in RAIM, the better measurement z_2 (in terms of rms error) is given the most weight in forming $\hat{x}_{LS}$. This is the natural thing to do.

On the other hand, with NIORAIM, the relative weighting is reversed. The poorer measurement is given more weight. This seems counterintuitive at first glance; however, on second thought, it makes sense. Imagine a gradually increasing failure occurring in z_2 (the better measurement), and remember that the test statistic reacts to this failure the same way in both RAIM and NIORAIM. Now, according to Equations 20 and 21, the failure reflects into the RAIM estimate much faster than into the NIORAIM estimate. So, at detection, the RAIM estimation error will have grown to be much larger than the NIORAIM one. Thus, the protection level is certainly poorer (that is, larger) in RAIM than it is in NIORAIM.

Of course, a small price must be paid for the improvement in failure protection with NIORAIM over RAIM. **Figure**

Appendix: Derivation of Equations

$$\mathbf{H}' = \begin{bmatrix} \lambda \\ 1 \end{bmatrix}$$

$$\mathbf{w} = \begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix}$$

$$\hat{x}_{WLS} = \left[(\mathbf{wH}')^{\mathrm{T}}\,\mathbf{wH}'\right]^{-1}(\mathbf{wH}')^{\mathrm{T}}\,\mathbf{wz}'$$

$$\hat{x}_{WLS} - x = \underbrace{\left[(\mathbf{wH}')^{\mathrm{T}}\,\mathbf{wH}'\right]^{-1}(\mathbf{wH}')^{\mathrm{T}}\,\mathbf{w}}_{\mathbf{A}_{\mathrm{w}}}\boldsymbol{\varepsilon}'$$

$$\mathbf{A}_{\mathrm{w}} = \left\{\left(\begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix}\begin{bmatrix} \lambda \\ 1 \end{bmatrix}\right)^{\mathrm{T}}\begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix}\begin{bmatrix} \lambda \\ 1 \end{bmatrix}\right\}^{-1}\left(\begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix}\begin{bmatrix} \lambda \\ 1 \end{bmatrix}\right)^{\mathrm{T}} = \frac{1}{\lambda^2 + \alpha^2\lambda^2}\begin{bmatrix} \lambda & \alpha\lambda \end{bmatrix}$$

$$Slope_1 = \frac{\mathbf{A}_{\mathrm{w}}\mathbf{w}\begin{bmatrix} b \\ 0 \end{bmatrix}}{\mathbf{P}\begin{bmatrix} b \\ 0 \end{bmatrix}} = \frac{\frac{1}{\lambda^2 + \alpha^2\lambda^2}\begin{bmatrix} \lambda & \alpha\lambda \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix}\begin{bmatrix} b \\ 0 \end{bmatrix}}{\begin{bmatrix} \frac{-1}{\sqrt{1+\lambda^2}} & \frac{\lambda}{\sqrt{1+\lambda^2}} \end{bmatrix}\begin{bmatrix} b \\ 0 \end{bmatrix}} = \frac{\frac{b\lambda}{\lambda^2(1+\alpha^2)}}{\frac{-b}{\sqrt{1+\lambda^2}}} = \frac{-\sqrt{1+\lambda^2}}{\lambda(1+\alpha^2)}$$

$$Slope_2 = \frac{\mathbf{A}_{\mathrm{w}}\mathbf{w}\begin{bmatrix} 0 \\ b \end{bmatrix}}{\mathbf{P}\begin{bmatrix} 0 \\ b \end{bmatrix}} = \frac{\frac{1}{\lambda^2 + \alpha^2\lambda^2}\begin{bmatrix} \lambda & \alpha\lambda \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix}\begin{bmatrix} 0 \\ b \end{bmatrix}}{\begin{bmatrix} \frac{-1}{\sqrt{1+\lambda^2}} & \frac{\lambda}{\sqrt{1+\lambda^2}} \end{bmatrix}\begin{bmatrix} 0 \\ b \end{bmatrix}} = \frac{\frac{b\alpha^2\lambda^2}{\lambda^2(1+\alpha^2)}}{\frac{b\lambda}{\sqrt{1+\lambda^2}}} = \frac{\alpha^2\sqrt{1+\lambda^2}}{\lambda(1+\alpha^2)}$$

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{bmatrix}$$

$$C_{11} = E\left\{\mathbf{P}\boldsymbol{\varepsilon}'(\mathbf{P}\boldsymbol{\varepsilon}')^{\mathrm{T}}\right\} = \mathbf{P}E\left\{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}'^{\mathrm{T}}\right\}\mathbf{P}^{\mathrm{T}} = \mathbf{PP}^{\mathrm{T}} = 1$$

$$C_{22} = E\left\{\mathbf{A}_{\mathrm{w}}\mathbf{w}\boldsymbol{\varepsilon}'(\mathbf{A}_{\mathrm{w}}\mathbf{w}\boldsymbol{\varepsilon}')^{\mathrm{T}}\right\} = \mathbf{A}_{\mathrm{w}}\mathbf{w}E\left\{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}'^{\mathrm{T}}\right\}\mathbf{w}^{\mathrm{T}}\mathbf{A}_{\mathrm{w}}^{\mathrm{T}}$$

$$= \frac{\begin{bmatrix} \lambda & \alpha\lambda \end{bmatrix}}{\lambda^2(1+\alpha^2)}\begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix}\begin{bmatrix} \lambda \\ \alpha\lambda \end{bmatrix}\frac{1}{\lambda^2(1+\alpha^2)} = \frac{1+\alpha^4\lambda^2}{\lambda^2(1+\alpha^2)^2}$$

$$C_{12} = E\left\{\mathbf{A}_{\mathrm{w}}\mathbf{w}\boldsymbol{\varepsilon}'(\mathbf{P}\boldsymbol{\varepsilon}')^{\mathrm{T}}\right\} = \mathbf{A}_{\mathrm{w}}\mathbf{w}E\left\{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}'^{\mathrm{T}}\right\}\mathbf{P}^{\mathrm{T}} = \frac{1}{\lambda^2(1+\alpha^2)}\begin{bmatrix} \lambda & \alpha\lambda \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \alpha\lambda \end{bmatrix}\begin{bmatrix} \frac{-1}{\sqrt{1+\lambda^2}} \\ \frac{\lambda}{\sqrt{1+\lambda^2}} \end{bmatrix} = \frac{\alpha^2\lambda^2 - 1}{\lambda(1+\alpha^2)\sqrt{1+\lambda^2}}$$

Note: $E\{\cdot\}$ is a statistical expectation operator and $E\left\{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}'^{\mathrm{T}}\right\} = \mathbf{I}$ signifies unity noise covariance.

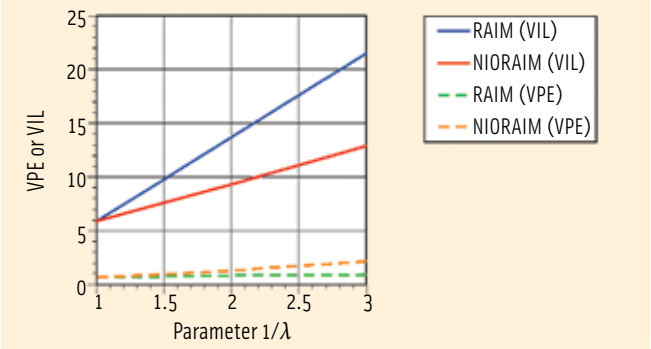


FIGURE 5 A comparison between NIORAIM and RAIM solutions of the vertical integrity limit (VIL) and vertical position error (VPE) are shown as a function of the parameter $1/\lambda$, which corresponds to the ratio of σ_1/σ_2. The differences between NIORAIM and RAIM become more evident as the disparity between σ_1 and σ_2 rises. The lowering of the NIORAIM VIL as compared to the RAIM VIL is far more significant than the increase of the NIORAIM VPE as compared to the RAIM VPE.

5 compares the VILs and vertical position errors (VPE) of RAIM and NIORAIM using typical values of missed detection and false alarm rates set at:

Allowable missed detection rate = 0.000825

Maximum false alarm rate = 4×10^{-6} (Threshold = ±4.611)

For example, consider the situation when the sigma of the poorer measurement is twice that of the better one. As Figure 5 shows, the VIL is reduced from 13.7 to 9.3 in going from RAIM to NIORAIM. This is a reduction of about 32 percent — not dramatic but certainly significant.

Figure 5 also shows the corresponding loss in rms accuracy for the same 2:1 ratio of measurement sigmas. The increase in rms error goes from 0.89 for RAIM to 1.38 for NIORAIM. This is about 55 percent, which seems a bit large in percentage terms. However, if σ_2 is just a meter or so, the loss of another half meter or so might be quite acceptable in order to achieve better availability.

Conclusions

A rigorous methodology has been presented for analyzing the effectiveness of NIORAIM when applied to the two-system problem, along with results comparing NIORAIM with conventional RAIM. These show that a significant reduction in VIL is obtainable for typical civil avionics integrity specifications. The improvement in VIL comes with a modest loss in rms accuracy during normal no-failure conditions, which may or may not be acceptable, depending on the situation at hand.

The authors do not mean to recommend NIORAIM for any specific application. The methodology presented here is simply offered for fair consideration along with other alternatives in any particular application. No attempt has been made here to evaluate the availability rates. This is obviously an important facet in any integrity assurance application, but we will leave it to others to address this problem.

In this article, we have explained multi-GNSS NIORAIM in terms of a two-system combination. However, the same methodology can be readily extended to situations involving even more systems, and this may prove to be of greater importance as additional systems become operational.

In such cases, the computational effort turns out to be more intensive; so, the use of a lookup table certainly becomes even more valuable. A three-system lookup table, for example, would become two-dimensional (in λ) instead of the one-dimensional lookup table shown in this paper for the two-system case.

Even though the preparation of this lookup table will be considerably more involved, the real-time operation of deriving the VIL given the associated sigmas of the three systems is expected to be quick and efficient.

Additional Resources

[1] Brown, R. G., "A Baseline RAIM Scheme And A Note On The Equivalence Of Three RAIM Methods," *The Institute of Navigation Redbook Series, Vol. 5*, pp. 101–116, 1998

[2] Brown, R. G., and G.Y. Chin, "GPS RAIM: Calculation of Thresholds and Protection Radius Using Chi-Square Methods – A Geometric Approach," *The Institute of Navigation Redbook Series, Vol. 5*, pp. 155–178, 1998

[3] Hwang, P. Y., and R. Grover Brown, "RAIM-FDE Revisited: A New Breakthrough In Availability Performance With NIORAIM (Novel Integrity-Optimized RAIM)," *NAVIGATION, Journal of the Institute of Navigation*, Vol. 53, No. 1, pp. 41–51, Spring 2006

[4] Lee, Y. C., and M. P. McLaughlin, "Feasibility Analysis of RAIM to Provide LPV-200 Approaches with Future GPS," ION GNSS 20th International Technical Meeting of the Satellite Division, Fort Worth, Texas, USA, September 2007

[5] Lee, Y. C., and R. Braff, J. P. Fernow, D. Hashemi, M. P. McLaughlin, D. O'Laughlin, "GPS and Galileo with RAIM or WAAS for Vertically Guided Approaches," ION GNSS International Technical Meeting of the ION Satellite Division, Long Beach, California, USA, September 2005

[6] Young, R. S. Y., and G. A. McGraw, "Fault Detection and Exclusion Using Normalized Solution Separation and Residual Monitoring Methods," *NAVIGATION, Journal of the Institute of Navigation*, Vol. 50, No. 3, pp. 151–169, Fall 2003

Authors

Patrick Y. Hwang is principal systems engineer with the Advanced Technology Center of Rockwell Collins in Cedar Rapids, Iowa. He has more than 25 years experience in advanced navigation systems technology and design. Hwang has collaborated with this paper's coauthor in several technical papers on GPS and Kalman filtering applications and also in a reference textbook on applied Kalman filtering.

R. Grover Brown is a Distinguished Professor Emeritus in Engineering, Iowa State University. His many publications in the field of navigation span a period of more than 40 years. In the early 1990s, he chaired the Working Group of RTCA Special Committee 159 during the period when the committee issued the recommendation that receiver autonomous integrity monitoring (RAIM) be the primary means of assuring integrity of GPS signals for use in U.S. civil aviation. IG