The Integrity of GPS

Richard B. Langley

University of New Brunswick

How truthful is GPS? Can you believe the position that your GPS receiver computes? The GPS Standard Positioning Service is designed to provide a horizontal position accuracy of at least 100 meters, but such accuracy cannot be guaranteed 100 percent of the time. Satellite or ground system failures could cause a receiver to use erroneous data and compute positions that exceed its normal accuracy level. This month's column explores the different approaches to ensuring GPS signal integrity, including satellite self-checks, receiver autonomous integrity monitoring, and augmented systems.

"Innovation" is a regular column featuring discussions about recent advances in GPS technology and its applications as well as the fundamentals of GPS positioning. The column is coordinated by Richard Langley of the Department of Geodesy and Geomatics Engineering at the University of New Brunswick, who appreciates receiving your comments as well as topic suggestions for future columns. To contact him, see the "Columnists" section on page 4 of this issue.

The performance of any navigation system is characterized by its accuracy, availability, continuity, and integrity. From a safety point of view, integrity is arguably the most important factor. Without some assurance of a system's integrity, we have no way of knowing whether the information we receive is correct: How are we to know whether a navigation system is actually achieving its advertised accuracy and not misleading us with faulty information?

Because of possible disastrous consequences when relying on GPS for navigating marine vessels or aircraft, researchers have directed a considerable amount of attention to GPS integrity during the past 15 years, as the system has increasingly become central to navigation operations. Efforts have focused on developing both GPS control system enhancements and designs for user-receiver independent assessment of GPS integrity through a concept known as receiver autonomous integrity monitoring (RAIM). This article provides an overview of those efforts and highlights the factors one should be aware of when considering the integrity of GPS.

PERFORMANCE PARAMETERS

Before investigating the various approaches used to ensure some level of GPS system integrity, let's define and quantify navigation system integrity in general as well as the associated concepts of accuracy, availability, and continuity.

Accuracy. Perhaps the most obvious navigation system requirement, accuracy describes how well a measured value agrees with a reference value. Ideally, the reference value should be the true value, if known, or some agreed-upon standard value. The accuracy of a clock, for example, is determined by how well it keeps time compared with a standard clock, such as an atomic clock maintained by a national timing laboratory. In terms of GPS positioning, a reference value might be the published coordinates of a geodetic reference mark.

A measurement's error is simply the difference between the obtained value and the reference value. If we make a series of repeated measurements and calculate the mean value, the difference between the mean and the reference value is called the *bias* or systematic error. Therefore, we usually take accuracy to mean the absence of bias, and we measure inaccuracy by the size of the bias.

Note that accuracy is not necessarily the same thing as precision. Precision denotes a measurement quality that describes how well repeated measurements agree with themselves rather than with a reference value. In other words, it is determined by the scatter or dispersion of the measurements. We have various ways of quantifying precision, including standard deviation, variance, range, and confidence and probability intervals. Accuracy, however, cannot be calculated solely from measurement values. If we can calibrate a system's bias, though, or if the bias is negligible, then we can, with caution, interpret precision estimates as accuracy estimates.

Availability. A navigation system's availability refers to its ability to provide the required function and performance within the specified coverage area at the start of an intended operation. In many cases, system availability implies signal availability, which is expressed as the percentage of time that the system's transmitted signals are accessible for use. In addition to transmitter capability, environmental factors such as anomalous atmospheric conditions or interfering signals might affect availability. According to the 1996 Federal Radionavigation Plan, the GPS Standard Positioning Service will be available at least 99.85 percent of the time, based on a global average.

Continuity. Ideally, any navigation system should be continuously available to users. But, because of scheduled maintenance or unpredictable outages, a particular system may be unavailable at a certain time. Continuity, accordingly, is the ability of a total navigation system to function without interruption during an intended period of operation. More specifically, it indicates the probability that the system will maintain its specified performance level for the duration of an operation, presuming system availability at the beginning of that process.

Integrity. The integrity of a navigation system refers — just as it does to a person — to its honesty, veracity, and trustworthiness. A system might be available at the start of an operation, and we might predict its continuity at an advertised accuracy during the operation. But what if something unexpectedly goes wrong? If some system anomaly results in unacceptable navigation accuracy, the system should detect this and warn the user. Integrity characterizes a navigation system's ability to provide this timely warning when it fails to meet its stated accuracy.

Traditionally, some component of the navigation system or an independent monitoring unit assures integrity by examining the transmitted signals and providing a timely warning if they are out of specification. For example, Loran-C provides system integrity by monitoring timing accuracy. Stations that exceed the system tolerance of (nominally) 100 nanoseconds transmit special blinking signals, which commence within 60 seconds of detecting an anomaly. VHF omnidirectional range aviation systems use an independent monitor to supply system integrity and remove a signal from use within 10 seconds of an out-of-tolerance condition. Integral monitors in instrument landing system and microwave landing system facilities similarly exclude anomalous signals from use within one second.

GPS INTEGRITY

GPS achieves integrity and protects users against system anomalies and failures by relying on satellite self-checks and monitoring by the U.S. Department of Defense's Operation Control Segment Master Control Station (MCS), as well as signal assessment by users. Thus, GPS has both integral and independent mechanisms to assure integrity.

Satellite Self-Checks. GPS satellites internally monitor themselves for some, but not all, anomalies. These include navigation data errors, selective availability (SA) and antispoof (AS) failures, and certain types of satellite clock failures. If such internal failures are detected, satellites notify users within six seconds.

Navigation data and SA failures, for example, can occur because of data corruption. Satellite navigation electronics are susceptible to damage from the space environment. Heavy ion cosmic rays and energetic particles from the sun can ionize silicon material when they pass through it, causing a bit flip or bit hit (also known as a single event upset) in memory devices and thereby corrupting stored data. To prevent bit hits from affecting navigation or SA data in Block II and IIA satellites, these data are stored in specially hardened, electrically-alterable, read-only memory (EAROM) that is almost impervious to bit hits.

A satellite's navigation processor refreshes scratch-pad, random-access memory every six seconds with data pulled from the EAROM. In the unlikely event that EAROM data are corrupted and the appropriate navigation data cannot be found, the satellite will output (for the next six seconds) default navigation data, which consists of alternating ones and zeroes in words three through 10, with invalid parity of the affected words. (Each word in the satellite navigation message contains parity bits that are used to verify correct transmission of the navigation message from the satellite to the receiver.) If a satellite cannot find SA data, nonstandard C/A- and P-codes are modulated onto the carrier in place of the standard codes for the following six seconds. Therefore, the longest that a loss of navigation data could persist because of a bit hit would be six seconds.

For further protection, a satellite resets its processor every 24 seconds so that, if the address pointer is accidentally moved to an erroneous position in memory, it will recover to a known location within 24 seconds.

The Block IIR satellites use a different equipment architecture to protect against bit hits. They have a watch dog monitor (WDM) that regulates the functioning of the processor and decides when a processor reset must occur. If the WDM performs a processor reset, the satellite transmits nonstandard codes until it either automatically reverts to standard codes or MCS commands it to do so after resolving the cause of the reset.

Users are also informed of some anomalies by way of the navigation message handover word (HOW). The HOW contains an alert flag in bit 18 that informs civilian ("unauthorized") users that the satellite's user range error may be worse than indicated in the message subframe 1.

Master Control Station. The GPS constellation is monitored by the MCS at Schriever Air Force Base (formerly known as Falcon Air Force Base). Using data collected by five monitor stations distributed around the globe, MCS assesses GPS performance every 15 minutes by conducting tolerance and validation checks of the measured pseudoranges using a Kalman-filter, error-management process.

In some circumstances, ranging errors could go undetected by this process for as long as 29 minutes. To mitigate any potential problems from such long delays, MCS installed new software in February 1995 to check incoming range measurements every six seconds. When the software detects an anomaly, it raises an alarm. Reaction to the alarm typically occurs within one minute. After confirming an anomaly, MCS staff renders the offending satellite's L-band signals untrackable. MCS accomplishes this using a technique known as SATZAP, which sends a single command telling the satellite to change its pseudorandom noise (PRN) number to PRN 37 — a nonoperational number. The SATZAP procedure takes five minutes or fewer to accomplish and helps resolve integrity anomalies within 10 minutes of initial detection — assuming good monitor station and ground antenna visibility. Because this procedure also affects control segment tracking, MCS uploads new navigation message information to the satellite with a "bad" space vehicle health tag before reassigning the satellite its original PRN number.

The use of nonstandard ranging codes and the meaning of message alert flags and satellite health codes are described in the Interface Control Document (ICD)-GPS-200. Interestingly, MCS has determined that some GPS receivers are not fully ICD-GPS-200 compliant. That is, some receivers may use signals or navigation message data from a satellite marked unhealthy or transmitting nonstandard codes.

In addition to altering satellite navigation messages to inform users of health problems, MCS issues Notice Advisories to Navstar Users (NANUs) that report satellite outages as well as planned service losses caused by maintenance.

RAIM

Although the GPS command segment and the satellites themselves provide a reasonable level of integrity for most GPS purposes, it is insufficient for aviation, as anomalies could go undetected for too long a period. It typically takes MCS 5–15 minutes to remove a satellite with a detected anomaly from service. Furthermore, if a satellite is not in view of one of the ground stations (the ground stations provide only 92 percent tracking coverage), an anomaly could go undetected for more than 10 minutes before MCS can analyze the situation and take remedial action.

Aircraft traveling at high speeds can quickly deviate from a flight path if provided with faulty navigation information. The need for a higher level of integrity for GPS use in aviation led to the concept of having a GPS receiver independently or autonomously establish system integrity. This concept is known as receiver autonomous integrity monitoring, or RAIM — the term introduced by Rudy Kalafus in 1987.

RAIM attempts to answer two questions: Is there a GPS satellite failure? If so, which is the errant satellite? If GPS is used solely for supplemental navigation, then answering the first question is sufficient because an alternative navigation system is available. However, if GPS is used for primary-means navigation, both questions must be answered to identify and remove the failed satellite from the solution, allowing the aircraft to safely proceed. Answering either question necessitates measurement redundancy; that is, more than the minimum four measurements required for a position solution. We need measurements from at least five satellites to detect a satellite anomaly and a minimum of six to remove the faulty satellite from the navigation solution.

Similar to any system used to assess GPS integrity, a RAIM technique must interpret information about pseudorange errors in terms of induced horizontal position error and make a decision as to whether the level of error is acceptable by comparing it to the allowable radial error for a particular phase of flight. This allowable error is termed the *alarm limit*. If the alarm limit is exceeded, a RAIM-equipped receiver must issue a timely warning.

Researchers have proposed a variety of RAIM techniques for assessing integrity based on some kind of self-consistency check on receiver measurements. These techniques fall into two categories: those that use past as well as present measurements along with assumptions about receiver motion (averaging or filtering schemes) and those that use only current measurements (snapshot schemes).

SNAPSHOT APPROACHES

We will review here only the snapshot schemes as these have gained the most acceptance by the aviation community. These approaches have the advantage of not relying on possibly questionable assumptions about how the navigation filter attained its current state. The most commonly-used snapshot schemes are the range comparison, least-squares-residuals, and parity methods.

Range Comparison. This RAIM technique is relatively straightforward. Assuming a receiver has six satellites in view, it can determine a position solution using any four (to solve for the three coordinates of the receiver and its clock offset). From this position solution, the receiver can then estimate the other two satellite ranges and compare these estimates to the actual measurements. Based on the differences, the receiver can compute a test statistic and compare it with a predetermined or user-defined decision boundary (such as an ellipse, in the case of two redundant observations). If the statistic lies inside the decision boundary, the receiver declares no failure; if it lies outside, it activates an alarm.

Least-Squares Residuals. Employing this method, a receiver uses all satellites in view to compute a navigation solution. It then calculates the residuals, or the differences between the measurements and the predicted measurements, based on the solution. The residuals are then squared and summed. The receiver directly compares the resulting sum of squared errors (SSE) with a predefined threshold value. If the SSE is less than the threshold, there is no failure; if greater than the threshold, a failure exists.

Parity. A more formal RAIM technique mathematically transforms GPS measurements to a parity vector. A receiver can use this vector's magnitude (or its square) as a test statistic. In fact, the magnitude squared is identical to the SSE. Consequently, in RAIM failure detection, the least-squares-residuals and parity methods yield indistinguishable statistics, and, with similar threshold settings, they provide the same results.

It can be shown that the range comparison method is also equivalent to the parity method. In essence, then, all three of these RAIM techniques produce the same failure decisions, and their differences are mostly conceptual. The preference of one over another is usually for computational convenience.

RAIM AVAILABILITY

Satellite geometry constitutes an important aspect of RAIM. For RAIM to be available, a GPS receiver requires sufficient observation redundancy with geometries typical of good navigation solutions. In other words, it is necessary to screen out poor (inadmissible) detection geometries that might make it difficult (because of a noisy solution) to detect a failed satellite.

One possible way to determine satisfactory satellite geometry involves adopting a maximum permissible horizontal dilution of precision ($HDOP_{max}$) value and declaring all geometries with HDOP greater than $HDOP_{max}$ inadmissible. Such a determination could be made at the start of a flight and recalculated en route. If at anytime insufficient redundancy exists, RAIM, having excluded inadmissible geometries, would be unavailable. Aviators would then have to employ an alternate navigation system.

Although conceptually simple, this approach has some drawbacks (including uncertainty about how a particular satellite range error projects into a receiver's horizontal coordinates) and GPS researchers have proposed better methods of screening. One such approach computes the HDOPs associated with each of the available satellite subset solutions as well as the HDOP for the available satellites' full least-squares solution. A receiver computes the square roots of the differences between the squares of the subset HDOPs and the least-squares HDOP, selecting the largest of these as the (inverse) measure of satellite-geometry quality for failure detection purposes. If the result is greater than a predetermined threshold value, the geometry is registered inadmissible.

An aircraft GPS receiver must have at least five satellites in view above an elevation mask angle of 7.5 degrees to provide RAIM. This condition is not always satisfied with the existing GPS constellation, and in these situations, RAIM would be unavailable. Redundancy can be increased by adding measurements from other systems, however, such as GLONASS and onboard inertial units or barometric altitude measurements, which effectively provide an additional range value. With such auxiliary measurements, RAIM may be available where, with GPS measurements alone, it would not be.

EXCLUSION AND ISOLATION

Thus far we have only considered failure detection. That is, we have only been con-

FURTHER READING

For navigation integrity requirements and advertised integrity capabilities of various systems, see

- *1996 Federal Radionavigation Plan (FRP),* DOT-VNTSC-RSPA-97-2 and DOD-4650.5, U.S. Department of Transportation and U.S. Department of Defense, Washington, D.C., available from the National Technical Information Service, Springfield, Virginia, 1997.

For online versions of the 1996 FRP in HTML and PDF formats, see

- <http://www.navcen.uscg.mil/policy/frp1996/>.

For a discussion of the techniques used by the GPS satellites and the Master Control Station to maintain GPS integrity, see

- "Inside GPS: The Master Control Station," by C. Shank and J.W. Lavrakas, in *GPS World*, Vol. 5, No. 9, September 1994, pp. 46–54.

For in-depth review articles about RAIM, see

- "Receiver Autonomous Integrity Monitoring," by R.G. Brown, in *Global Positioning System: Theory and Applications,* Vol. 2, edited by B.W. Parkinson and J.J. Spilker Jr., Progress in Astronautics and Aeronautics, Vol. 164. American Institute of Aeronautics and Astronautics, Washington, D.C., 1996, pp.143–165.
- "Performance of Standalone GPS," by J.L. Leva, M.U. de Haag, and K. Van Dyke, in *Understanding GPS: Principles and Applications,* edited by E.D. Kaplan, Artech House Publishers, Norwood, Massachusetts, 1996, pp. 237–320.

For a compendium of papers about RAIM published between 1986 and 1996, see

- *Global Positioning System: Papers Published in Navigation,* Vol. 5, reprinted by The Institute of Navigation, Alexandria, Virginia, 1998.

For a discussion of aviation RAIM requirements, see

- *Minimum Operational Performance Standards for Airborne Supplemental Navigation Equipment Using Global Positioning System (GPS),* RTCA/DO-208, prepared by Special Committee 159, RTCA, Inc., Washington, D.C., 1991.

cerned with the first RAIM question: Is there a GPS satellite failure? For GPS to be used as a primary means of navigation, however, RAIM must also be able to remove the faulty satellite from the navigation solution. The two means of accomplishing this are termed *fault detection and isolation (FDI)*, if the faulty satellite is actually identified, and *fault detection and exclusion (FDE)*, if the faulty satellite is simply excluded from a navigation solution.

The FDE approach avails itself of higher measurement redundancies than required for basic RAIM fault detection. To illustrate how FDE operates, let's assume that eight satellites are in view, but the receiver determines solutions using only a six-satellite subset. If the receiver detects a failure, it examines the position solutions from all six-satellite subsets and selects the one that satisfies its self-consistency test. This solution excludes the faulty satellite along with a likely healthy satellite, but this is inconsequential because the remaining six satellites provide adequate navigation accuracy. With eight satellites in view, the FDE approach could handle two simultaneous failures, or if nine satellites were in view, three simultaneous failures.

Methodologies for solving the isolation problem are still evolving. One FDI approach builds on the RAIM parity method, taking advantage of the fact that the parity vector's direction helps identify the failed satellite.

AVIATION REQUIREMENTS

RAIM is a necessary component for GPS aviation applications. The U.S. Federal Aviation Administration (FAA) has mandated that any GPS receiver used for supplemental aircraft navigation incorporate integrity monitoring. Many other jurisdictions have adopted a similar requirement. Table 1 lists the FAA integrity performance requirements for the various phases of flight.

Currently, with enhancements to RAIM, aviators can even use GPS as a primary-means navigation system for the en route oceanic phase of flight as well as in remote areas (see Table 2 for navigation requirements of different flight phases). Enhancements include an FDE provision, use of a 0-degree elevation mask angle, and, in some areas, barometric aiding. Employing GPS for sole-means navigation during other phases of flight will require augmentation systems to reduce the probability of attaining misleading information within a one-hour period to 10^{-7}.

Table 1. FAA position integrity performance requirements for GPS use as a supplemental navigation system

Phase of flight	Alarm limit[1] (nautical miles)	Maximum allowable alarm rate[2] (per hour)	Time to alarm[3] (seconds)	Minimum detection probability[4]
En route	2.0	0.002	30	0.999
Terminal	1.0	0.002	10	0.999
Nonprecision approach	0.3	0.002	10	0.999

[1]Positional error magnitude that cannot be exceeded and for which the integrity monitoring system provides protection by warning the pilot.
[2]Maximum alarm rate is set so that alarms do not become a nuisance during operations.
[3]Maximum allowable time from the onset of failure until its annunciation.
[4]One minus the miss probability (1–x) that a failure occurs and is not detected and displayed to the pilot.

AUGMENTED GPS SYSTEMS

Augmented GPS systems have additional built-in integrity monitoring that can benefit GPS users.

DGPS. Marine differential GPS (DGPS) systems, such as the network operated by the U.S. Coast Guard (USCG), use an onsite integrity monitor to check satellite-signal validity and provide an independent assessment of satellite health. At times, a USCG DGPS station may be able to extend the period of a faulty satellite's use by transmitting accurate corrections. If the system determines that a satellite is unsuitable for navigation purposes, it notifies users of an out-of-tolerance condition within five seconds of detection. This broadcast alarm is actuated by transmitting pseudorange and pseudorange-rate correction values of –10485.76 meters and –4.096 meters per second, respectively (the largest negative values allowed by the message format). User receivers assess DGPS signal-integrity through parity checks using the same scheme as that used for GPS satellite navigation messages.

WAAS. The Wide Area Augmentation System (WAAS), in addition to providing higher positioning accuracies to aviation users, will afford integrity enhancements necessary for all phases of flight, up to and including Category I approaches. Upon completion, the WAAS ground system will continuously assess the integrity of GPS satellite signals as well as its own corrections, warning WAAS users when a failure is encountered. For Category I operation, for example, an alarm will be provided within 5.2 seconds of failure. This WAAS feature increases availability by obviating the requirement for redundant position calculations.

Table 2. Controlled airspace navigation requirements

Phase / Subphase	Position accuracy (95%)[1] Horizontal	Vertical
En route/terminal		
Oceanic	12.4 NM	—
Domestic	2.8 NM	—
Terminal	1.7 NM	—
Approach/landing		
Nonprecision	0.3 NM	—
Precision		
Category I	18.2 m	7.7–4.4 m
Category II	6.5 m	1.7 m
Category III	4.1 m	0.6 m

[1]Expressed in nautical miles (NM) or meters (m)

LAAS. The Local Area Augmentation System (LAAS) will provide an even higher level of integrity. For a Category III landing, the probability of an undetected system failure cannot exceed 5×10^{-9}. One approach to providing the required integrity employs pseudolites — low-power, ground-based transmitters functioning as pseudo-GPS satellites — in conjunction with a conventional-DGPS airport reference station. Situated in pairs on either side of a runway approach path, pseudolites can provide an aircraft with enough ranging sources to initialize DGPS to centimeter-level accuracy with the required degree of integrity.

CONCLUSION

All too often we get hung up on the stated or claimed accuracy of a positioning or navigation system, forgetting that occassionally the actual real-time accuracy may be worse than the stated accuracy because of system failures. Integrity protects us against these rare occurrences. Stand-alone and augmented GPS offer various levels of integrity protection depending on system use, and continuing research will enhance GPS integrity even further. With such protection, we may safely use GPS to dock a vessel in a crowded harbor or land a jumbo jet in the fog. ■

ACKNOWLEDGMENTS

Thanks to Nav Canada's Ross Bowie and Rob Butler for helpful comments on a draft of this article.