GNSS Solutions:

Quantifying the performance of navigation systems and standards for assisted-GNSS

"GNSS Solutions" is a regular column featuring questions and answers about technical aspects of GNSS. Readers are invited to send their questions to the columnist, **Dr. Mark Petovello**, Department of Geomatics Engineering, University of Calgary, who will find experts to answer them.

Mark Petovello is an Assistant Professor in the Department of Geomatics Engineering at the University of Calgary. He has been actively involved in many aspects of positioning and navigation since 1997 including GNSS algorithm development, inertial navigation, sensor integration, and software development.

Email: mark.petovello@ucalgary.ca

What are the differences between *accuracy, integrity, continuity,* and *availability,* and how are they computed?

These four words describe parameters that quantify the performance of navigation systems. The terms are not unique to satellite navigation, as they have been used for many years with respect to other navigation systems and, with broader definitions, throughout the practice of engineering. This column will describe the use of these terms for navigation and focus on their applicability to GNSS.

Accuracy is the navigation performance parameter that is most commonly used and is the easiest to understand. It is a measure of the error, or the deviation of the estimated position from the unknown true position, of a given navigation tool or system. More precisely, accuracy is a statistical quantity associated with the probabilistic distribution of navigation error.

Depending on the system and its intended application, this quantity can be expressed in somewhat different ways. For example, many military systems express accuracy in terms of "circular error probable" (CEP) in two dimensions or "spherical error probable" (SEP) in three dimensions. This represents the median position error — it exceeds 50 percent of all position errors and falls below the other 50 percent.

More commonly used in civil applications are "1-sigma" and "2-sigma" error limits. In the case of position errors that follow Gaussian distributions, these limits express the 63rd and 95th percentiles of navigation errors, respectively, for a one-dimensional parameter (e.g., altitude). For higher-dimension parameters, such as 2-D horizontal position, the one- and two-sigma limits represent lower percentiles than in the one-dimensional case and can be computed from a chi-squared distribution if the underlying range-domain errors are Gaussian.

Because the Gaussian distribution describes most navigation system error distributions fairly well out to the 95th percentile, many accuracy descriptions use "95%" and "2-sigma" numbers interchangeably. However, this convention should not be taken to mean that the underlying error distribution is actually Gaussian, particularly in the "tails" beyond 2-sigma, where the norm is for variations from the Gaussian distribution to exist.

Accuracy is obviously a value of paramount importance when selecting among candidate navigation systems and deciding what use can be made of their measurements. Because accuracy defines errors under typical conditions, it expresses what users will experience in normal, everyday use.

Beyond Accuracy

To varying extents, the other three parameters described next express relatively rare phenomena that may not be noticed by typical users (unless system performance is far short of what is required) and thus are mostly evaluated by offline analysis and simulation.

Integrity relates to the level of trust that can be placed in a navigation system. Here, "trust" refers to reliance that gross errors (errors much larger than the accuracy of the system) can be avoided.

In practice, this concept is expressed quantitatively using three sub-parameters. The first is *integrity risk,* which denotes the probability (per operation or per unit of time) that the system generates an unacceptable error without also providing a timely warning that the system's outputs cannot be trusted. Such an event is called "loss of integrity" (LOI) in some contexts and "misleading information" (MI) in others. Depending on the potential consequence of "misleading information," a word expressing the severity of the consequence may be added, such as "hazardously misleading information" or HMI.

The second sub-parameter is the *alert limit,* which defines the magnitude of error that, if exceeded, is unacceptable from a safety standpoint. Alert limits generally come from requirements for particular applications and are expressed in terms of position error bounds. Multiple bounds, such as a horizontal alert limit (HAL) and a vertical alert limit (VAL), often exist depending on the intended application.

The third sub-parameter is *time to alert* (TTA), which is defined as the time between the occurrence of potential misleading information (i.e., one or more alert limits are violated) and the time at which an alert or correction (i.e., exclusion of the failed measurements) is issued to the system user to protect him or her from the underlying problem.

Note that the TTA "clock" begins ticking not when a failure takes place but when this failure causes navigation errors to grow to the extent that one or more alert limits might be exceeded.

Losing Integrity

Loss of integrity generally occurs in two ways. The first of these is for a position error to exceed its alert limit under "nominal" conditions without any particular system fault or anomaly taking place. The probability of this is typically extremely low given the gap between typical system accuracy and errors large enough to be unsafe, but it cannot be ruled out.

Because no failure or anomaly is involved in this form of integrity loss, there is nothing to detect. Thus, in such cases, integrity monitoring in general and the time-to-alert metric in particular do not apply, at least in theory. In practice, because many degrees of "off-nominal" circumstances exist between "nominal" and "faulted" conditions, detection is possible, but this possibility is normally not considered in analysis.

The second and more likely possibility for loss of integrity is a fault

Copyright iStockphoto.com/Mark Evans

or anomaly occurring that leads to loss of integrity if detection and exclusion do not occur within the time to alert. In this case, the *prior probability* of the fault occurring is an important contributor to integrity risk, as is the *probability of missed detection,* or the probability that the fault will not be detected before the time to alert expires.

Although detailed analysis of these scenarios must be done offline, the results of these analyses can be encapsulated into real-time user calculations of *protection levels,* which express the position-error bounds that can be protected to the required probability level of integrity risk.

By ensuring that these protection levels are no greater than the corresponding alert limits (i.e., confirming that, in the vertical position axis, the vertical protection level or VPL ≤ VAL), integrity can be verified in real-time for the set of GNSS satellites available to the user.

Continuity concerns the reliability of the position outputs of a navigation system. *Continuity risk* is the probability that the system will stop providing navigation outputs of the specified quality during a given operation or time interval, assuming that the outputs were present and of specified quality at the beginning.

Loss of continuity can occur when a navigation system (or an individual GNSS satellite) simply stops working or broadcasting signals. In these cases, it is obvious that something went wrong.

In other cases, continuity is lost because of the actions of one or more integrity monitors in detecting a real or imaginary fault, and then either failing to exclude the affected measurements (leading to complete loss of navigation) or implementing an exclusion that leaves the remaining measurements incapable of meeting the required performance.

For applications that cannot be interrupted without some level of danger, such as aircraft precision approach and landing under Category III weather minima, loss of continuity poses its own safety hazards.

Two additional parameters are useful in analyzing continuity performance. The *probability of false alert* (or *false detection* or *fault-free alert*) is the probability that one or more integrity monitors issue an alert (leading to measurement exclusion and possibly continuity loss) when no underlying fault is present.

A *critical satellite* for GNSS applications refers to one whose loss due to sudden failure or exclusion by the integrity monitor (whether needed to protect integrity or not) would cause loss of continuity. Because, in most cases, accuracy predictions and integrity protection levels can be computed in advance of a given operation, we can usually determine which satellites (and how many) of the set in view are critical and if the resulting implied continuity risk is acceptable.

Finally, *availability* expresses the likelihood that the other three performance parameters previously defined meet the requirements of a particular application. The most common definition of availability is the long-term average probability (subject to certain conditions) that the accuracy, integrity, and continuity requirements are simultaneously met.

This does not always apply in practice, however. For example, some applications do not require that continuity be assured at the time the operation is conducted. Also, many common applications only have accuracy requirements. Therefore, when availability is mentioned, the requirements that must be met for a given operation (at a given moment in time) to be declared "available" should be clear or commonly understood.

All Together Now

To compute a long-term average probability of a GNSS system meeting all of its requirements simultaneously, simulations of satellite orbits and visibility at one or more user locations are conducted. These simulations normally include the possibility that one or more satellites are unhealthy for one reason or another, including planned maintenance, unplanned failures that are still being corrected, and "end-of-life" outages of a satellite that has not yet been replaced with a new one.

Because GNSS satellite ground tracks typically repeat (e.g., every 24 hours for today's GPS satellites), simulations need only cover the period before the ground tracks (and thus user satellite geometry) begin to repeat.

For a given constellation state and time epoch, the relevant performance metrics are checked, and either a "1" (available) or "0" (unavailable) is stored. The result of each epoch is then multiplied by its probability of occurrence (based on published or estimated satellite-outage probabilities) to calculate a long-term average probability.

Other definitions of availability, different from a long-term average probability, are sometimes used. One example is sometimes called "operational availability" and refers to the maximum duration of unavailability given some worst-case system state. This is a useful additional quantity to know for systems with very high availability requirements.

For example, for GNSS systems designed to aid in landing airplanes in bad weather, knowing that the long-term availability of the system is, for example, approximately 0.999 would be important — suggesting an average period of unavailability of about 86 seconds per day, 44 minutes per month, or 8.76 hours per year. But users would also derive an important benefit from knowing that, for all scenarios where no more than three satellites in the nominal 24 primary orbit slots in the GPS constellation are unhealthy, the maximum consecutive duration of unavailability is 75 minutes.

When assessing availability, one should always remember that the resulting numbers are only as good as the assumptions (error models, satellite constellations, worst-case outages and probabilities, etc.) that go into them. For this reason, all results from availability analyses must make clear all relevant assumptions and, ideally, how the results would differ if key assumptions were changed.

SAM PULLEN

Dr. Sam Pullen is a senior research engineer at Stanford University, where he is the manager of the local area augmentation system (LAAS) research effort. He has supported the FAA in developing LAAS and wide area augmentation system (WAAS) concepts, technical requirements, integrity algorithms, and performance models since he received his Ph.D. in Aeronautics and Astronautics from Stanford in 1996. His research involves optimal aerospace system design and verification under uncertainty.

What is assisted GNSS? What changes are needed in assistance service standards to support the new near-future GNSSs such as GLONASS and QZSS?

Assisted GNSS (AGNSS) is a technology that enables faster position determination in an AGNSS-enabled handset than could be achieved using the broadcast GNSS satellite data only.

When a handset positions itself, it requests assistance data from the network. This assistance data includes, among other things: navigation models for ephemerides and clock corrections, reference location, ionosphere models, reference time, and optionally differential corrections for high-accuracy positioning and data bit assistance for high sensitivity.

When the assistance data is delivered over a telecom system architecture's *application layer* connection (usually a TCP/IP connection), the typical position fix times are in the order of 10–20 seconds compared to 40–60 seconds using autonomous methods or even longer fix times in weak signal conditions. In the best case, the handset already has the assistance data in its memory — for example, in the form of extended ephemeris — and, as a result, the position fix time can be as short as few seconds.

Currently the multi-GNSS cellular standards in 3GPP GERAN (GSM), 3GPP RAN (UMTS) and OMA SUPL 2.0 (Application Layer) support only GPS L1 C/A-code–based Standard Positioning Service and all Galileo Open Service signals. The work to include other GPS signals such as L5, L2C, and L1C signals and those from other GNSSs, such as GLONASS, satellite-based augmentation systems (SBASs) and Japan's regional Quasi Zenith Satellite System (QZSS) is starting in 3GPP and OMA forums this autumn. Adding new systems is straightforward due to the multi-GNSS framework already in place in the assistance data specifications from earlier releases.

One of the most promising activities is the work towards OMA SUPL 2.1 that promises to bring convergence to the currently very messy situation in application layer–based AGNSS solutions. The current OMA SUPL 2.0 specification is largely a collection of cellular network–specific protocols applied to the application layer that offer very different levels of performance from each other, none of which is optimally suited for the application layer.

The OMA SUPL2.1 solution also offers a unique opportunity to enable hybrid use of GNSS-based methods with non-cellular and non-GNSS positioning solutions, such as Wi-Fi positioning because the key principle in the design is its independence of cellular network specific protocols and their limitations.

Other Performance Improvements

The new GNSSs obviously address the issue of low signal availability in urban canyons. Accuracy improvements can be achieved, first of all, by the introduction of carrier-phase positioning–based methods to the assistance service standards. For example, RTK-based carrier phase services are already

GNSS Solutions *continued on page 50*

Acronyms

AGNSS	Assisted GNSS
EDGE	Enhanced Data rates for Global Evolution
GERAN	GSM EDGE Radio Access Network
GNSS	Global Navigation Satellite System
GSM	Global System for Mobile Communications
OMA	Open Mobile Alliance
PPP	Precise Point Positioning
QZSS	Quasi-Zenith Satellite System
RAN	Radio Access Network
SBAS	Satellite Based Augmentation System
SUPL	Secure User Plane Location
UMTS	Universal Mobile Telecommunications System

in place in a plethora of countries and, hence, the mass market applications could now also start exploiting the full advantage of this technology and existing services. The OMA SUPL 2.1 solution is intended to support streaming of assistance data between the handset and the assistance data server enabling, amongst other things, continuous transfer of carrier-phase assistance from the server to the handset.

The new assistance standards also improve positioning availability in terms of novel types of navigation models. For example, the multi-GNSS assistance standards support long-term navigation models that may be valid several days in the future. Moreover, the multi-GNSS assistance standards support non-native formats, resulting in the performance harmonization across different GNSSs. This is achieved by presenting the orbit models for all the GNSSs in the same format.

Over the longer term, AGNSS standards are open for the addition of even more advanced assistance data. To name a few, these could include real-time or predicted ionosphere and troposphere maps enabling precise point positioning (PPP) in handsets.

JARI SYRJÄRINNE AND LAURI WIROLA

Jari Syrjärinne received his M.Sc. and Ph.D. from Tampere University of Technology, Finland, majoring in digital signal processing and applied mathematics. Since 1999 he has been at Nokia Inc. working with various positioning technologies from wireless network–based positioning methods to GNSS positioning. He is currently involved in work on assisted GNSS and algorithms for hybrid positioning, running multiple R&D projects in these areas.

Lauri Wirola, received his Master of Science degree from Tampere University of Technology, Finland, majoring in electrophysics. In 2005 he joined the positioning group at Nokia. His present research interests include multi-GNSS positioning, RTK, and PPP as well as AGNSS. He is also involved in the location service standardization in the Open Mobile Alliance. Moreover, he is currently undertaking postgraduate studies in modern electromagnetism and mathematics. IG