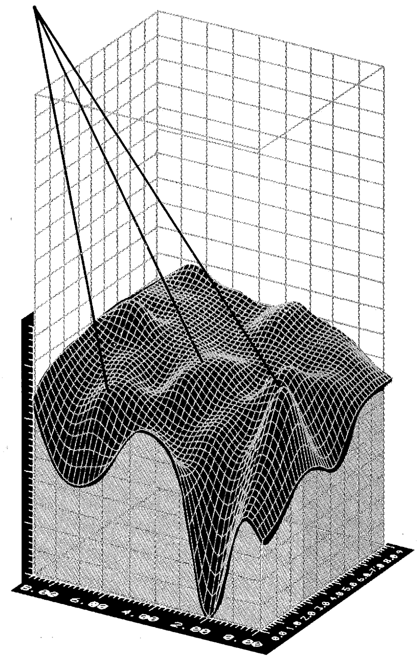
THE LIMITATIONS OF GPS

Alfred Kleusberg &
Richard B. Langley

Reprinted from GPS WORLD March/April 1990



THE LIMITATIONS OF GPS

Alfred Kleusberg & Richard B. Langley
University of New Brunswick

"Innovation" is a regular column in GPS World *commenting on GPS technology, product development, and other issues and needs of the GPS community. The first column, "GPS: A Multipurpose System" in the January/February 1990 issue, focused on the many capabilities of GPS. This second column looks at the down side – what are the limitations of GPS? Future contributions will explore topics such as selective availability, GPS and solar activity, GPS signal structure, GPS and electronic charts, GPS and geographical information systems, and the combined use of GPS and GLONASS. We welcome readers' comments and topic suggestions for future columns.*

Unfortunately, GPS is not a panacea; its performance in certain environments and for particular applications can be quite limited. In these cases GPS may give us no answer, the wrong answer, or an answer with insufficient accuracy when we ask, "Where are we?" We will illustrate these limitations using the examples introduced in the first "Innovation" column. In three plausible scenarios, that article described the use of GPS for positioning, navigation, and surveying. But we glossed over some potential problems with GPS that might have resulted in unhappy endings for our little stories.

The positioning capabilities were illustrated by describing how a leaking underground gas main in Gotham City was located through the combined use of GPS and the city engineering department's geographical information system (GIS). A major assumption in this scenario was that *GPS signal reception* would not be a problem in the concrete canyons of a major urban area.

The navigation capability of GPS was described using a fictional supertanker approaching the port of Rotterdam. The ship was safely guided by a GPS-based navigation system around all navigation hazards from the open sea right to the harbor dock. This happily uneventful voyage was made possible because the reliability of the GPS navigation system was guaranteed. Appropriate measures had been taken to detect possible system malfunctioning. Without properly monitoring *GPS signal integrity,* it would be dangerous to rely on GPS navigation in a hazardous environment.

In surveying, GPS is used to measure distances between survey markers. As was illustrated in the third example, GPS is a prime candidate for the most demanding surveys, such as those carried out to determine the deformations of the earth's crust during and after earthquakes. In the future it may even be possible to use GPS to measure the small precursor deformations that occur just before the earthquake strikes. In this case, GPS could be part of an earthquake prediction and warning system. However, without sufficient *GPS signal accuracy,* these small deformations may remain below the detectability threshold of GPS surveys.

THREE LIMITATIONS

This column will look first at the three particular problem areas that have been identified here: poor GPS signal reception, loss of GPS signal integrity, and limited positioning accuracy. We will also look at ways to overcome some of these limitations.

GPS signal reception. Proper functioning of a GPS receiver requires the undisturbed reception of signals from at least four GPS satellites. These signals propagate from the satellites to the receiver antenna along the line of sight and cannot penetrate water, soil, walls, or other obstacles very well. Therefore, GPS cannot be used for subsurface marine navigation nor for underground positioning and surveying – for example in mines or tunnels. In surface navigation and positioning applications, the signal can be obstructed by trees, buildings, and bridges. In many cases, this signal shading will be transitory, and therefore will not severely hamper the positioning. However, in the inner city streets of urban areas lined with skyscrapers, the "visibility" of the GPS satellites is very limited. In such areas, the signals can be obstructed for extended periods of time or even continuously unavailable.

Transitory signal shading by topography and bridges can also occur in coastal and inland water navigation. Depending on the location of the GPS antenna on board and the motion of the vessel, even the vessel's own superstructure can block the signal temporarily. In airborne applications, signal shading through the aircraft fuselage and wings can happen at high banking angles.

Essentially, users of a GPS-based navigation system have two options to handle such situations. They may forecast that these problems will not last for long, in which case a user takes no action and hopes that nothing serious happens during the time of signal loss. Alternatively, users may invest in hardware to add another independent navigation system, which will bridge those periods when GPS is unable to provide a navigation solution.

GPS signal integrity. A GPS receiver/processor computes position and time from range measurements to the GPS satellites, using satellite positions derived from information encoded in the transmitted signal – the "satellite message." With one measurement to each of four satellites there will usually be a unique receiver position solution. However, wrong satellite positions or wrong range measurements will result in an incorrect calculation of receiver position. If the faulty signals are not detected, the user will not know that the displayed position is wrong. Obviously, this situation has the potential for disaster. Therefore, the issue of GPS signal integrity monitoring has become a major concern, particularly within the civil aviation community.

At least two schemes have been proposed to detect faulty GPS signals and to warn the system's users. The *GPS Integrity Channel (GIC)* concept employs stationary GPS receivers at known locations. Knowledge of the precise receiver locations allows the direct monitoring of the quality of the received GPS signals. If any signal irregularities or gross errors are detected, a warning message is immediately broadcast to the navigational users of GPS through, for example, a separate satellite communications link.

The second scheme, *Receiver Autonomous Integrity Monitoring (RAIM),* does not require any special monitoring stations. RAIM is based on a user's receiver measuring ranges to more than four satellites. Since only four ranges are required for a position fix, the additional measurements can be used to identify signal abnormalities and to alert users. This type of integrity monitoring improves with the number of satellites used for ranging, and would be well suited for a combined GPS/GLONASS navigation system with a total of 48 satellites in orbit.

GPS signal accuracy. In any measuring system there is a limit to the accuracy with which measurements can be made. The accuracy may be limited by how the measuring system is constructed, the laws of physics, or how the system is used. The global positioning system is no exception.

A GPS receiver essentially measures the time required for a signal to travel from the satellite to the receiver. This travel time is converted to a range measurement by multiplying it by the speed of light. However, this measurement is corrupted by a number of different errors, each of which can be expressed as a *user equivalent range error (UERE).*

We attempt to remove these errors from the measurements in one of two ways. The first way is by directly subtracting the error from the measurements. These corrections are computed using mathematical equations or models describing how we think the errors affect our measurements. A number of such models are included in the programs of a GPS receiver's microprocessor. Even more sophisticated models are used by scientists and engineers who are interested in high accuracies and who process GPS data in their own computers.

The other way we can reduce UEREs is in the way we make our measurements. We will illustrate this approach with some examples later in this column. Any error that is not removed from the range measurements will lead to errors of similar or even larger size in the computed receiver positions.

TYPES OF ERROR

GPS measurements can be influenced by errors introduced at the satellite where the signal is generated and transmitted, errors caused as the signal travels from the satellite to the user's receiver, and errors introduced when the measurements are made in the receiver.

Satellite errors. Two different effects associated with the satellite contribute to range errors. As mentioned earlier, GPS satellites transmit, via the satellite message, information on the position of the satellite — known as an *ephemeris* — which is used in the computation of the receiver position. This satellite ephemeris is predicted from previous observations of the satellites by the GPS control stations. Because ephemeris predictions cannot be made with absolute accuracy, the satellite positions computed from the message are contaminated by the *satellite ephemeris error.*

As we have seen, the range between satellite and receiver is obtained from the signal travel-time measurement. In simplified terms, the travel-time measurement is obtained by comparing a clock in the satellite and a clock in the receiver. Although the clocks used in the GPS satellites are of the highest quality, they are not perfect. Therefore, the travel-time measurement and the corresponding range between satellite and receiver will be contaminated by a *satellite clock error.* The combined effect of satellite orbit and clock errors on range measurements is on the order of a few meters, on average.

The atmosphere claims its toll on the GPS signal twice.

To make matters worse, the U.S. Department of Defense plans to deliberately degrade the accuracy of both the broadcast ephemerides and the clocks of the GPS satellites. This action is termed selective availability (SA) and is an effort to restrict the accuracy capability of most civilian users of GPS. The degradations will be carried out using a classified code and only those with access to the code will be able to remove the effects of SA from their range measurements. SA will increase the satellite-dependent range errors to a few tens of meters, on average.

Signal propagation errors. The second category of errors is introduced when the signals from the GPS satellites pass through the earth's atmosphere on their way to the receiver on or near the earth's surface. The atmosphere affects the signals by changing the speed at which they travel, a phenomenon known as refraction. This leads to an error in the range derived from the measured signal travel time.

The atmosphere claims its toll on the GPS signal twice. The signal is first affected by the ionosphere, the uppermost part of the atmosphere, which contains large numbers of electrically charged particles. The number of particles is not constant. Therefore, the effect on the signal travel time, the *ionospheric refraction error*, varies from day to day and from place to place. In some circumstances, it can amount to as much as 30 meters in the vertical direction. For signal reception at the horizon, this number typically increases by a factor of three.

The state of the ionosphere cannot be easily predicted. But a basic property of the ionospheric error is that it is different for signals of different frequency. This feature can be used for error correction and will be discussed further on in this column.

GPS signals are also affected by the lowest part of the atmosphere — the troposphere — giving rise to the *tropospheric refraction error.* This error also varies temporally and spatially, and is about 2.3 meters in the vertical direction. For signals received near the horizon it may increase by a factor of 10. The tropospheric error is not frequency-dependent, but it can be calculated with an accuracy of a few centimeters from observations of atmospheric conditions at the receiver site.

Receiver errors. The third group of range errors originates in the GPS receiver or its antenna. They include a *receiver clock error* and measurement *noise.* Compared to the satellite clock error discussed earlier, the receiver clock error is usually very large and has to be treated as an unknown quantity, like the receiver position. It is for this reason that we need to measure the signals from four satellites to determine a three-dimensional position. The measurement noise depends on the type of "ruler" used for the range measurements. GPS signals contain three rulers with different resolutions: coarse acquisition (C/A) code, P-code, and the carrier phase. These rulers permit a range measurement precision of a few meters, a few decimeters, and a few millimeters, respectively.

Two additional receiver-dependent effects are multipath errors caused by a signal reflected from buildings or other objects interfering with a direct signal from the satellite, and errors caused by movement of the electrical reference point or phase center of the antenna as it receives signals from different directions. Both of these effects are very difficult to handle and must be avoided by careful siting of the antenna during observations and careful design and construction of the antenna by the manufacturer.

Once GPS is fully operational, it will provide two classes of GPS service: the Precise Positioning Service (PPS) for military and other authorized users and the Standard Positioning Service (SPS) for everyone else. The main difference between these two types of service is the measurement accuracy. PPS users will measure ranges using the P-code ruler on both frequencies transmitted by the satellites. The use of two frequencies allows removal of the ionospheric refraction error from the range measurements. PPS users also have access to the SA codes and can remove the deliberately introduced satellite ephemeris and clock errors. The total remaining range error for PPS users is about five meters.

SPS users will measure ranges using the C/A-code ruler on one signal frequency only. They can remove some part of the ionospheric refraction error only by using crude prediction models. Because they do not have access to the code for SA, they will not be able to remove the SA errors. These errors will limit the SPS range measurement accuracy to 30–50 meters.

GPS Geometry. So far we have talked about errors in the ranges measured by a GPS receiver. Another factor also affects the accuracy of GPS positions: satellite configuration geometry. If the satellites we are receiving signals from are all bunched together in the sky, our position accuracy will tend to be quite poor. On the other hand, if they are more or less spread out in the sky, our accuracy will be much better.

The effect of satellite configuration geometry is expressed by a quantity called the dilution of precision (DOP) factor and is largely a function of the number of available satellites. Once the complete GPS constellation of satellites is in place, we can generally anticipate DOP values of 3 to 5. If we multiply the range measurement error by the DOP factor, we get an estimate of the positioning accuracy. Multiplying the typical range measurement error by the typical full-constellation DOP means we should be able to obtain average point-position accuracy at the 20-meter level for PPS users and at the 100-meter level for SPS users.

In general, an increase in position accuracy does not come for free.

IMPROVING GPS ACCURACY

These accuracy figures can be considerably improved in relative positioning: the determination of the position of one or more remote GPS receivers with respect to a fixed base station at a known location. By continuously monitoring the difference between the ranges measured from the satellite signals and the ranges computed using the known base station position, the range measurement error can be determined. Because most of these errors will be similar for the other receivers at the remote stations, they can be subtracted from the range measurements at the remote sites. This procedure handles even the errors deliberately introduced by SA. Resulting relative position errors are about 10 meters for SPS users.

Another approach is employed in very precise surveying applications of GPS. In surveying, we determine the relative position between several static survey markers. For this task, usually the finest of the three GPS rulers is used: the carrier phase. To exploit the high precision of this ruler, we must attempt to model all errors down to the level of a few centimeters. For some of the errors this can be a very difficult proposition. But we have a trick up our sleeve that can eliminate or reduce errors not easily corrected by our models: differencing.

Differencing is the simple subtraction of a range measurement at one station from the simultaneous range measurement at another station. This eliminates some of the range errors because they are common to the two measurements being differenced (satellite clock errors) and reduces other errors drastically because they are almost the same in the two measurements (atmospheric refraction and orbital errors). The error removal/reduction is more likely to be successful if the simultaneously ranging receivers are close together than if they are far apart.

This technique is the standard approach followed by geodesists and others interested in measuring high-precision relative positions between two or more receivers. Resulting accuracies for relative positions depend on the distance between the stations and the sophistication of the measurement analysis. The accuracies vary between several millimeters over a distance of a kilometer and several centimeters over distances of hundreds of kilometers.

CONCLUSION

In this article, we have looked at potential signal reception and signal integrity problems when using GPS for positioning and navigation, and we have shown ways to overcome these difficulties. We have also looked at the position accuracy achievable with GPS, which ranges from 100 meters with SPS and selective availability implemented to 20 meters with PPS. However, because PPS is not expected to be available to the civilian user community, those users must follow other approaches if they need better accuracy. Relative positioning techniques can improve accuracy to 10 meters with SPS and to a few centimeters using the carrier-phase measurements.

In general, an increase in position accuracy does not come for free. It usually goes hand in hand with an increase in equipment cost, logistics problems, and data processing complexity. Users must consider those factors when deciding which mode of GPS positioning to employ. Despite its limitations, however, GPS is still the best all-around positioning system available today. ■