Elliott D. Kaplan
The MITRE Corporation
Bedford, Massachusetts

THE GLOBAL POSITIONING SYSTEM (GPS)

An overview

Functional description

GPS is a space-based radio navigation system that provides worldwide three dimensional position, velocity, and Universal Coordinated Time (UTC) information to users equipped with GPS receivers. Developed by the Department of Defense (DOD), the system was primarily designed for military use; however, the number of civilian users is expected to exceed those in the military. The system provides two services: the Standard Positioning Service (SPS) for the civilian community and the Precise Positioning Service (PPS) for the use by the military and select government agencies (e.g., the Central Intelligence Agency, Drug Enforcement Agency, and Federal Bureau of Investigation). PPS access is controlled through cryptography. GPS implementation and operational policies are delineated in the Federal Radionavigation Plan (FRP), prepared jointly by both the DOD and the Department of Transportation (DOT).

Principle of operation

Let's begin our discussion of GPS with a look at the principle of operation. In **Figure 1**, we wish to determine vector $\underline{U}$, which represents a user receiver's position with respect to the center of the Earth. Vector $\underline{R}_S$ represents the distance and angular orientation from the center of the Earth to the satellite. Vector $\underline{R}_S$ is calculated using ephemeris (position) data broadcast by the satellite. The user-to-satellite distance vector, $\underline{\rho}$, is:

$$\underline{\rho} = \underline{R}_S - \underline{U} \tag{1}$$

Thus, the magnitude of vector $\underline{\rho}$, is:

$$\|\underline{\rho}\| = \|\underline{R}_S - \underline{U}\| \tag{2}$$

Let, ρ represent the magnitude of vector $\underline{\rho}$:

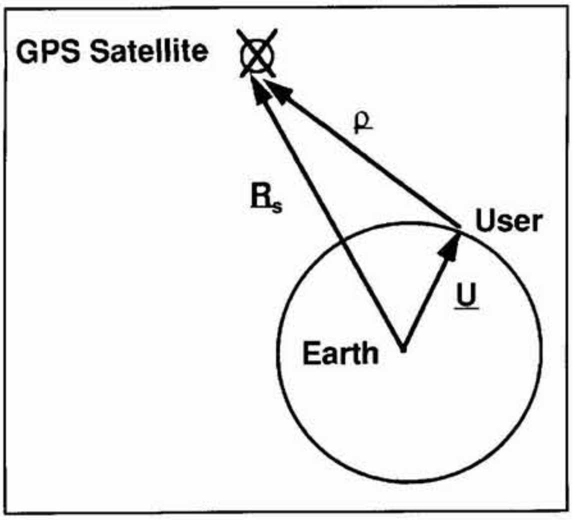

Figure 1. User position vector representation.

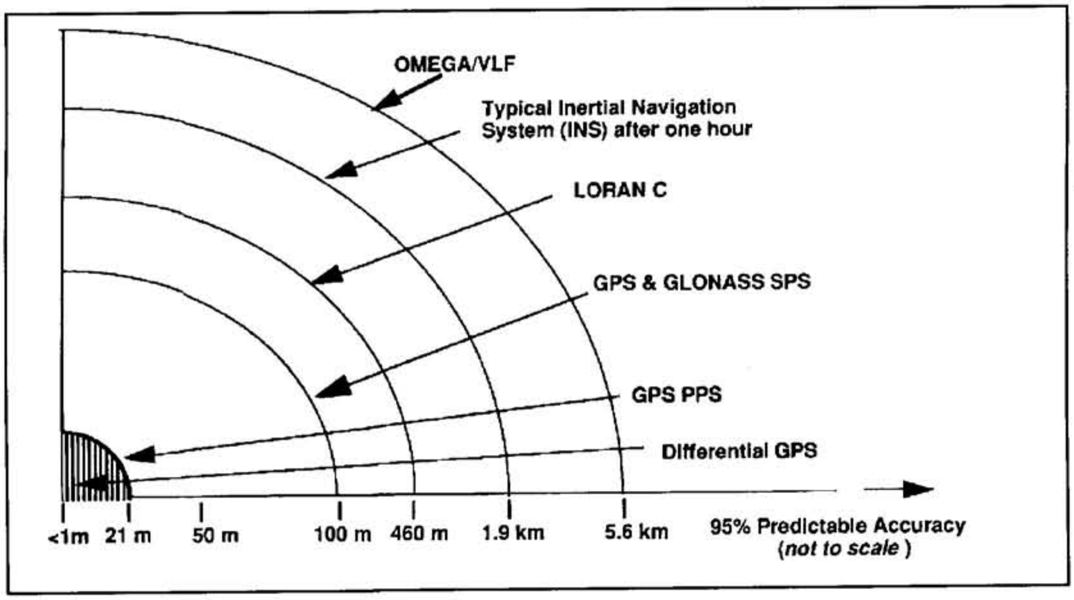

Figure 2. Accuracy comparison of GPS with other navigation systems.

$$\underline{\rho} = \|\underline{R}_s - \underline{U}\| \tag{3}$$

Distance ρ is computed by measuring the propagation time required for a satellite-transmitted ranging code to reach the receiver. Within the receiver, an identical coded ranging signal is generated and shifted in time until it achieves correlation with the satellite-generated ranging code. The magnitude of the time shift is the satellite-to-user range. If the satellite clock and the receiver clock were perfectly synchronized, the correlation process would yield the true satellite-to-user range; however, this is not the case. Each satellite contains a highly accurate cesium or rubidium atomic clock assumed to be synchronized to GPS system time.* (GPS system time is an internal reference that is transparent to the user.) The receiver contains a less accurate crystal clock set approximately to GPS system time. Thus, the range determined by the correlation process is denoted as the pseudorange, R, because it contains the true satellite-to-user range, ρ, and an offset attributed to the difference between GPS system time and user clock. Therefore, **Equation 3** can be rewritten as:

$$R - c\Delta t = \|\underline{R}_s - \underline{U}\| \tag{4}$$

where c is the speed of light and Δt represents the advance of the receiver clock with respect to GPS system time.

To determine user position in three dimensions (U_x, U_y, and U_z) and the offset Δt, pseudorange measurements are made to four satellites resulting in:

$$R_i - c\Delta t = \|\underline{R}_{si} - \underline{U}\| \tag{5}$$

where i denotes satellites 1, 2, 3, and 4.

This equation can be expanded into the following set of equations, with four equations and four unknowns (U_x, U_y, U_z, and Δt):

$$R_1 - c\Delta t = \sqrt{(X_1 - U_x)^2 + (Y_1 - U_y)^2 + (Z_1 - U_z)^2} \tag{6}$$

$$R_2 - c\Delta t = \sqrt{(X_2 - U_x)^2 + (Y_2 - U_y)^2 + (Z_2 - U_z)^2} \tag{7}$$

$$R_3 - c\Delta t = \sqrt{(X_3 - U_x)^2 + (Y_3 - U_y)^2 + (Z_3 - U_z)^2} \tag{8}$$

$$R_4 - c\Delta t = \sqrt{(X_4 - U_x)^2 + (Y_4 - U_y)^2 + (Z_4 - U_z)^2} \tag{9}$$

where X_i, Y_i, and Z_i denote the ith satellite's position in three dimensions.

Several closed form solutions published in the open literature or various iterative techniques such as Newton-Raphson are used to solve these nonlinear equations for the unknown quantities.[1]

Comparison of GPS to existing navigation systems

The benefits of GPS are readily observed

---

*Satellite clocks are free running and each has a unique time bias from GPS system time. However, these time biases are computed by the GPS ground monitoring network (described later) and are uplinked to the satellites for rebroadcast to the users. Because the corrections are applied within the user receiver to synchronize each ranging signal to GPS time, the assumption is considered valid.

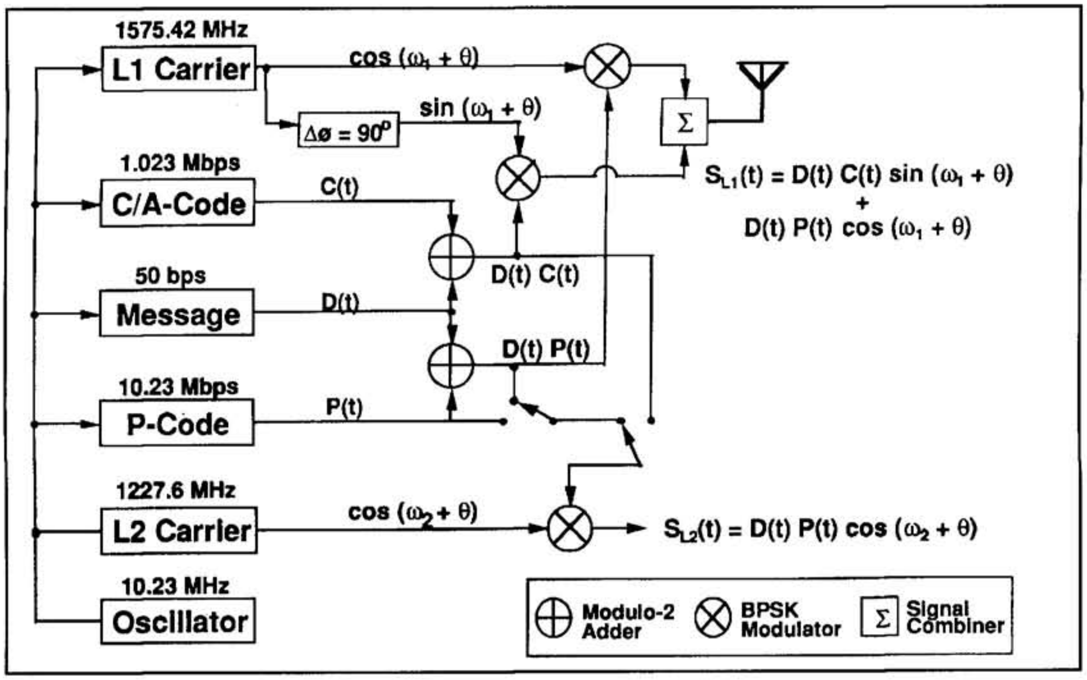

Figure 3. Satellite signal generation.

when compared to some of today's widely used navigation systems. Key features of these systems—such as accuracy, coverage, and frequency—are provided below. An accuracy comparison of these systems is provided in **Figure 2**.

OMEGA. Aside from GPS, OMEGA is the only radio navigation aid providing worldwide coverage. (The TRANSIT satellite system provides two-dimensional worldwide coverage, but DOD plans to terminate TRANSIT operations in 1996.) OMEGA is jointly operated by the DOD and six other nations. Broadcasting very low frequency (VLF) CW signals in the frequency region of approximately 10 kHz to 14 kHz, OMEGA stations provide horizontal guidance only. Predictable horizontal position accuracy varies from 3.7 km to 7.4 km, (95 percent). According to the FRP, the DOD is expected to discontinue OMEGA usage in 1996. OMEGA civilian operations are expected to continue until 2005.[2]

LORAN-C. LORAN-C is a radio navigation system widely used for civilian/military aviation, land, and marine applications. Unlike OMEGA, LORAN-C coverage is regional (e.g., contiguous United States, lower sections of Canada, portions of Japan, etc.). LORAN-C transmissions are at 100 kHz and provide horizontal guidance with a predictable accuracy of 460 m. DOD is discontinuing its LORAN-C requirement starting in 1995. Civilian use of LORAN-C within CONUS is expected to continue until 2015. The U.S. Coast Guard will terminate overseas LORAN-C operations at the end of 1994.[2]

Inertial Navigation. Unlike LORAN-C and OMEGA, an inertial navigation system (INS) is not a radionavigation system, but rather an electromechanical mechanism that provides navigation guidance based on Newton's Laws of Motion. An INS uses gyroscopes, accelerometers, and associated electronics (including a computer) to sense turning rates and accelerations associated with the rotation of the platform. The accelerometers measure individual horizontal and vertical acceleration components, while the gyroscopes stabilize the accelerometers in a desired orientation. These accelerations are integrated by the computer to obtain platform position and velocity. The computer also calculates orientation corrections caused by the motion over the Earth, the rotation of the Earth, and other factors. INS position and velocity accuracies degrade over time. A nominal position drift error rate is 1.9 km/hour (95 percent). If the INS is installed in an integrated navigation system with one or more external navigation aids, such as GPS and/or LORAN, the external navigation aid can be used to update the INS and correct the drift error.[3]

Global Navigation Satellite System (GLONASS). GLONASS is the Russian counterpart to GPS. GLONASS provides worldwide coverage, however, its accuracy performance is optimized for the northern latitudes (Western Europe and Russia, for example). GLONASS accuracy performance is specified as identical to that of GPS SPS.

GPS SPS. As stated earlier, GPS SPS is the primary service provided to the civilian com-

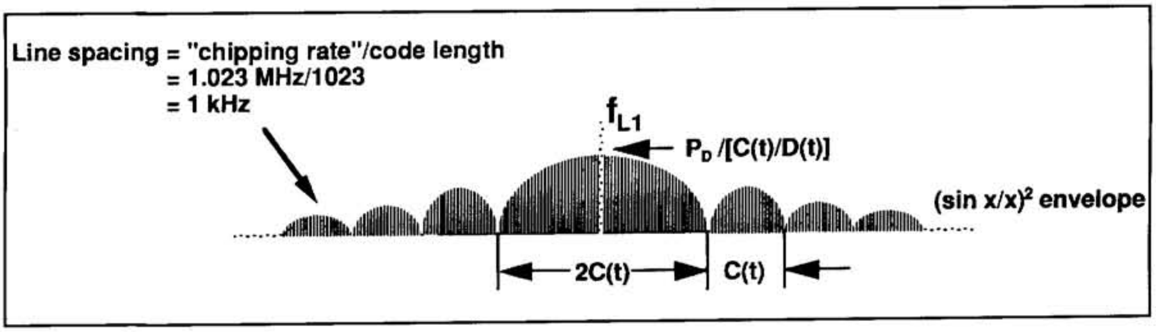

Figure 4. Effects of spreading data.

munity worldwide. The service provides accuracies on a daily basis within 100 m (2 drms,* 95 percent probability) in the horizontal plane and 140 m (95 percent probability) in the vertical plane. Further, the UTC output is within 340 nsec (95 percent probability).[2]

GPS PPS. The GPS PPS is designated for authorized users (for example, select government agencies and the military) equipped with cryptographic capabilities. On a daily basis, PPS horizontal accuracy is 21 m (2 drms, 95 percent probability) in the horizontal plane and 29 m (95 percent probability) in the vertical plane. The UTC output is within 200 nsec (95 percent probability).[2]

Differential GPS (DGPS). DGPS is an accuracy enhancement technique that provides horizontal and vertical positioning accuracies better thanthose of the PPS. Further elaboration on DGPS is provided later in the text, but it is important to note that some DGPS techniques provide measurement accuracy on the order of millimeters.

GPS system components

GPS is comprised of three segments: space, operational control, and user. A brief description of each segment is provided below:

Space Segment. The space segment consists of 24 satellites positioned in six orbital planes with each plane containing four satellites. Satellite orbit altitude is about 20,186 km (10,900 nmi) above the Earth's surface. At present, the constellation consists of 24 Block (BLK) II/IIA (production) satellites and 1 BLK I (prototype) satellite. The system reached initial operation capability (IOC) in mid-December 1993. IOC marked the availability of the SPS. The full operational capability (FOC) constellation is in place now and consists of the 24 BLK II/IIA satellites. The last BLK II/IIA production satellite was launched in March 1994. Declaration of FOC will follow an extensive system "check-out" and is expected sometime in early 1995. Once FOC is attained, four to nine satellites will be in view at any time. (The number of satellites in view is a function of user location, time of day, and satellite outages.) BLK IIR replenishment satellites are in production. The first launch is anticipated for 1996. This launch date is based on BLK II/IIA reliability predictions. A BLK IIF follow-on program for the replacement of the BLK IIR series is being planned.

Operational Control Segment (OCS). The OCS consists of a master control station (MCS) at Falcon AFB, Colorado and monitor stations in the following locations: Ascension, Diego Garcia, Falcon AFB, Hawaii, and Kwajalein. A backup MCS is located at Gaithersburg, Maryland. The OCS monitors the health and status of the constellation and uploads navigation data to the satellites for retransmission to users. The monitoring stations passively track and collect data from all GPS satellites in view. This data is relayed every 6 seconds to the MCS via the Defense Satellite Communications System (DSCS). MCS data processing examines monitoring station data for anomalies and computes new navigation parameters (for example, almanac [i.e., coarse satellite position data], ephemeris [i.e., fine satellite position data], health, satellite clock corrections, etc.). The new navigation parameters are disseminated to all monitoring stations via DSCS and uploaded to the constellation via an S-band link. Only three monitoring station sites (Ascension, Diego Garcia, and Kwajalein) have an upload capability.

User segment. The user segment consists of GPS receiving equipment. Today GPS receivers come in a variety of forms ranging from a single board of "credit card" size to shock-mounted Navy shipboard units with a "footprint" of approximately 1500 sq. cm that weigh 31.8 kg (70 lbs.) each. Single-board receivers

---

*2 drms is the radius of a circle centered about the receiver equal to twice the root-sum-square of the variances around the x and y axes:

$$2\ \text{drms} = \sqrt{\sigma_x^2 + \sigma_y^2} \qquad \textbf{(10)}$$

The percentage of fixes contained within the 2 drms circle varies between approximately 95.5 percent and 98.2 percent, depending on the ellipticity of the error distribution. For GPS, the 95 percent value is used.[2]

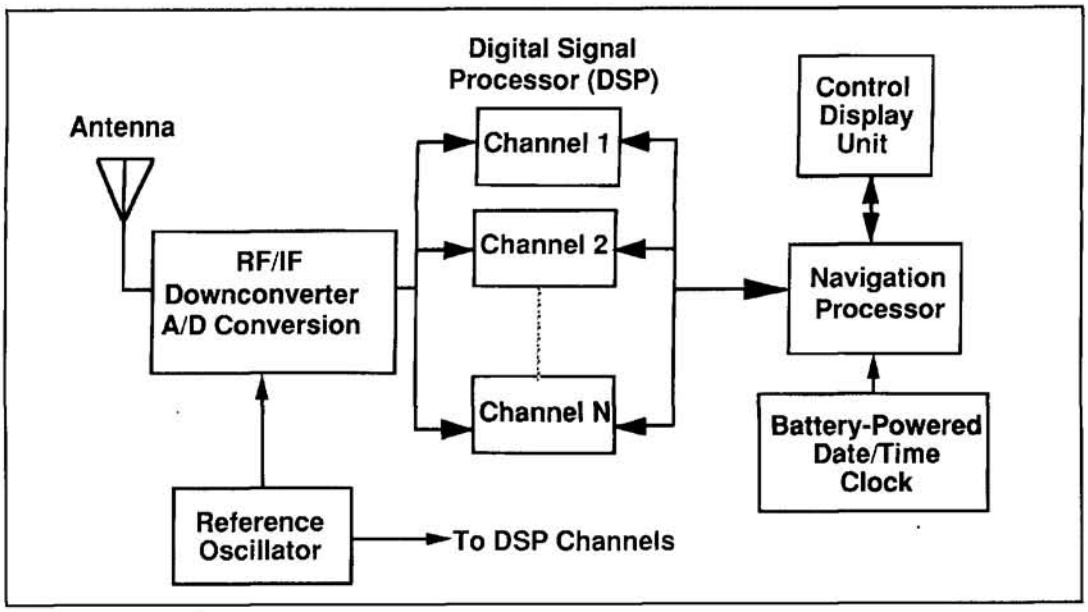

Figure 5. Generic SPS GPS receiver.

are available from several manufacturers. These devices are primarily for SPS OEM applications and require an antenna, display, I/O electronics, and power. There are also numerous self-contained SPS handheld receivers on the market designed for aviation, land, and marine applications. Aviation receivers are designed for a high dynamic environment. Rockwell-Collins is currently producing the PPS Precision Lightweight GPS Receiver (PLGR) for the DOD.Aside from handhelds panel mount receivers for aviation and marine applications are also popular. Unlike handhelds, panel mount receivers rely on platform power rather than batteries. In addition, panel mount units offer hands free operation and provide security in terms of a fixed mounting.

The majority of military airborne GPS receiver configurations are composed of several components: antenna, RF/IF downconverter assembly, receiver, and control/display unit. In some tactical platform configurations, an adaptive array antenna and associated downconverter provide an anti-jam capability. This array is electronically steerable and creates nulls in the antenna pattern in the direction of the jammer.

GPS satellite signal characteristics

Satellite signal generation is illustrated in **Figure 3**. Each satellite contains a 10.23-MHz atomic clock (oscillator) that provides reference timing for all onboard signals. (Depending on the spacecraft type, there can be up to four reference oscillators on board; however, only one

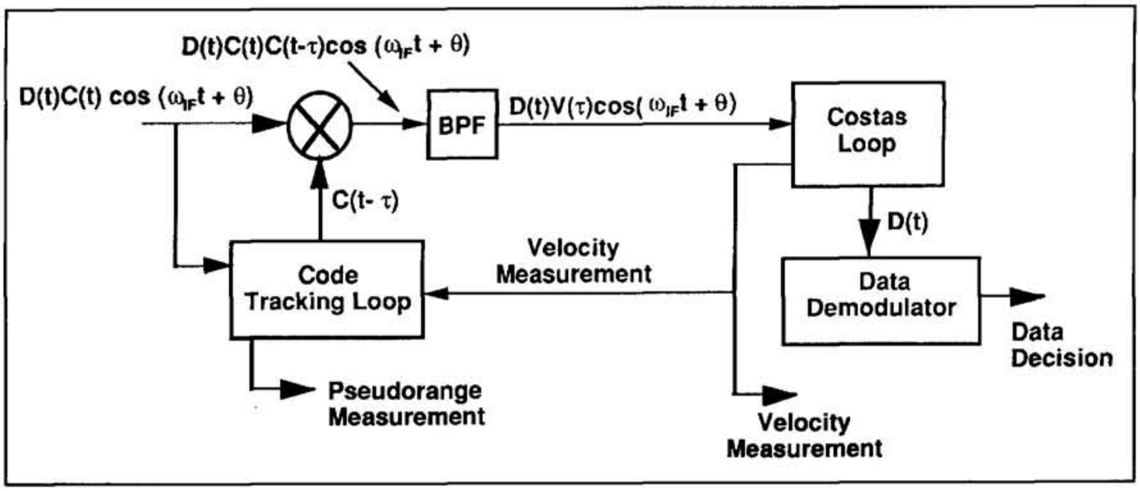

Figure 6. Direct sequence "despreading" technique.

Amateur Radio and GPS

A compact Global Positioning System (GPS) unit can be a handy piece of equipment for the amateur radio operator who likes to use the grid locator system popular with VHF+ aficionados. Introduced in Germany over 40 years ago to encourage activity on the VHF+ ham bands, the grid locator scheme evolved into a worldwide awards and location system. Today's program is called the Maidenhead Grid Locator System, and is used daily by hams all over the world to describe their location.

Grid locator designators are determined by longitude and latitude. The designator itself contains up to six places—two letters followed by two numbers, followed by two letters. The grid system divides the Earth's surface into 324 "fields" of 20 degrees of longitude and 10 degrees of latitude. The fields are identified by two letters—AA through RR. Each field is then divided into 100 "squares" of 2 degrees of longitude x 1 degree of latitude; these are identified by two numbers from 00 to 99. Each square is further divided into 576 subsquares of 5 minutes of longitude x 2.5 minutes of latitude. Subsquares are identified by two letters (AA-XX). The trick for the amateur radio operator is to figure out, as accurately as possible, his location within the grid system.

While it's certainly possible to determine your particular grid locator designator using tools like the *ARRL Grid Square Map for North America* and the *ARRL World Grid Square Atlas*; shareware software like *BD, HAMGRID, LOCCALC,* and *SQUARES*; or U.S. Geological Survey maps; these inexpensive methods are not always the easiest. Enter the handheld GPS receiver. This little gadget, though considerably more expensive, lets you obtain the information you need to find your grid locator designator quickly and with a fair degree of accuracy. GPS handheld units rely on the same technology and 24-satellite system as their larger, more complex counterparts. A unit's accuracy is based on the number of satellite signals it receives. The GPS receiver must receive a minimum of three satellites to obtain the signal necessary to produce a two dimensional location, which will allow the circuitry in the unit to calculate your location. However, it takes a minimum of four satellitesto determine accurate elevation .

Several manufacturers offer handheld GPS units. Prices range from $500 and up. One such unit is shown in **Photo A**. For more information on handheld global positioning systems check out the catalogs of amateur radio equipment suppliers and manufacturers of navigational aids.

Terry Littlefield, KA1STC
Editor

BIBLIOGRAPHY

1. Joseph Lynch, N6CL, *The VHF "How To" Book*, Chapter 15, CQ Communications, Inc., 1994.
2. Arthur Lange, W6RXG, David Sprague, and Arthur Woo, "Scout GPS ™ Application Note 6," Trimble Navigation, Austin Texas. To obtain a copy, call 1-800-959-9567

Photo A. The Trimble Scout GPS ™ handheld unit has the Maidenhead Grid Locator System built into its software. The unit's RS-232I/O port lets you upload and download location data to and from your computer.

SEGMENT SOURCE	ERROR SOURCE	P-CODE PSEUDO RANGE ERROR 1σ [m]
SPACE	Clock and nav, subsystem stability	3.3
	Other	0.5
CONTROL	Ephemeris prediction model implementation	4.2
	Predictability of satellite perturbations	1.0
	Other	0.9
USER	Ionospheric delay	2.3
	Tropospheric delay compensation	2.0
	Receiver noise and resolution	1.5
	Multipath	1.2
	Other	0.5
SYSTEM UERE	Total (Root-sum-square)	6.6

Figure 7. P(Y) code UERE budget.

SEGMENT SOURCE	ERROR SOURCE	C/A-CODE PSEUDO RANGE ERROR 1σ [m]
SPACE	Selective Availability	30.3
	Clock and nav, subsystem stability	3.3
	Other	0.5
CONTROL	Ephemeris prediction model implementation	4.2
	Predictability of satellite perturbations	1.0
	Other	0.9
USER	Ionospheric delay	5.0
	Tropospheric delay compensation	0.8
	Receiver noise and resolution	1.0
	Multipath	1.2
	Other	0.5
SYSTEM UERE	Total (Root-sum-square)	31.3

Figure 8. C/A code UERE budget (estimate).

is used as the primary timing source. The others provide redundancy.) The reference timing signal is upconverted (multiplied by 154) to create the L1 carrier at 1575.42 MHz. The L1 carrier is designated as the primary navigation frequency. The 10.23-MHz source is also multiplied by 120 to form the L2 carrier at 1227.6 MHz. The L2 carrier is designated as the secondary navigation frequency. Each satellite broadcasts on L1 and L2. Carrier L1 is separated into quadrature components. The in-phase and 90-degree phase-shifted components are represented by $\cos(\omega_1 t+\theta)$ and $\sin(\omega_1 t+\theta)$, respectively. Here ω_1 denotes the L1 carrier frequency and θ represents oscillator phase noise and drift. In addition to carrier frequency generation, the 10.23-MHz source is also used to form the coarse/acquisition (C/A) and precision (P) ranging code signals, and the navigation data message (NAV data message). Both ranging codes appear to consist of a random pattern of binary digits and are denoted as pseudorandom codes. The pseudorandom code ones and zeros are equal in magnitude, but reversed in sign. A one is usually represented by +1 volt, while a zero is generally denoted by a -1 volt.

The C/A (ranging) code is a 1023 bit Gold code clocked at 1.023 MHz. Thus, the code has a 1 msec epoch and repeats constantly. Each satellite is assigned a unique Gold code. Gold codes were selected for GPS because a subset of codes within the family have low cross correlation properties with respect to one another. This is an important property because the satellite broadcast format is code division multiple access (CDMA) with each satellite transmitting its C/A-code on L1. Thus, there are multiple C/A-codes on L1 simultaneously. In order to make a pseudorange measurement to a particular satellite, the receiver must be able to distinguish one code from another. C/A-codes are primarily used by the civilian community for SPS measurements.

The P (ranging) code is a 267-day sequence with a bit rate of 10.23 MHz. Each satellite is assigned a unique one-week segment of the P-code that repeats every midnight Saturday/Sunday. Each P-code segment has low cross-correlation properties with respect to other P-code segments and Gold codes. Once FOC is declared, P-code transmissions will be encrypted to form the Y-code. The Y-code will be accessible to authorized users through cryptography. Y-code cross-correlation properties are assumed to be identical to those of the P-code. (Y-code properties are classified and I have no knowledge of them.)

The NAV data message contains information regarding the health and status of the satellites, as well as data to aid in the satellite-to-user measurement process. NAV data message parameters include almanac, ephemeris, satellite clock corrections, GPS system time, UTC offset from GPS system time, and C/A-code to P(Y)-code handover information. Authorized receivers usually acquire the satellite transmission using the C/A-code and then switch to P(Y)-code tracking via the use of handover information. The NAV data message is broadcast at 50 Hz. Messages are 12.5 minutes in length and cyclic.

The NAV data message is modulo-2 added to both the C/A-code and the P-code forming products D(t) C(t) and D(t) P(t), respectively (see **Figure 3**). The D(t) P(t) product binary phase shift key (BPSK) modulates the in-phase component of the L1 carrier while product D(t) C(t) BPSK modulates the phase shifted component. This modulation technique is called direct sequence spread spectrum. The modulo-2 addition effectively "spreads" the 50-Hz data, D(t), across approximately the same amount of spectrum required by the multiplying ranging code. The ranging codes are also called "chipping" codes. The bit rate of the chipping code is referred to as the "chipping" rate. In addition, signal power density is decreased by the ratio of the chipping rate to the data rate. **Figure 4** depicts the spectrum of the C/A-code chipping the NAV message data on L1. The carrier has been suppressed due to the double balanced mixer and balanced spreading code. That is, the C/A-code has an approximately equal number of ones and zeros. This property also resides in the P-code. The outputs of both BPSK modulators are summed and fed to a right-handed circularly polarized (RHCP) antenna for broadcast-to-user receivers. This signal transmitted by the ith satellite on L1 is:

$$S_{L1,i}(t) = P_i(t)D(t)\cos(\omega_1 t + \theta) + C_i(t)D(t)\sin(\omega_1 t + \theta) \qquad \textbf{(11)}$$

where:

$P_i(t)$ = P code on ith satellite
$C_i(t)$ = C/A code on ith satellite
$D_i(t)$ = NAV data message on ith satellite

L2 carrier modulation is similar; however, the modulation format is selectable by the OCS. Either the D(t) C(t) product or P(t) with or without the modulo-2 addition of D(t) is chosen to modulate the L2 carrier. For the most part, the D(t) P(t) product is employed. The modulated carrier is broadcast to user receivers through a different RHCP antenna than the one used for the L1 signal broadcast. The L2 signal transmitted from the ith satellite is:

$$S_{L2,i}(t) = P_i(t)D(t)\cos(\omega_2 t + \theta) \qquad \textbf{(12)}$$

Receiver measurement process

Figure 5 is a block diagram of a generic SPS GPS receiver. Satellite signals are received via the antenna, which is RHCP and provides near hemispherical coverage. Typical coverage is 160 degrees with near unity gain. In general, SPS receivers process only C/A-code on L1 because the P-code on both L1 and L2 will be encrypted once the system reaches FOC. The CDMA signals are then downconverted to an intermediate frequency (IF). Following a second downconversion to baseband, the signals are sampled by an analog-to-digital converter at a frequency greater than or equal to twice the C/A-code bandwidth (2.046 MHz). In some receivers, sampling occurs at the reference oscillator rate, 10.23 MHz. Because the samples aren't synchronized to the received C/A codes, oversampling ensures that the signal will be sampled throughout and not mistakenly near a code chip transition.[4] The samples are forwarded to the digital signal processor (DSP). The DSP contains N parallel channels to track the carriers and codes from up to N satellites simultaneously. (N generally ranges from 5 to 12 in modern SPS receivers.) Each channel contains carrier and code tracking loops to perform pseudorange and velocity measurements, as well as NAV demodulation. Prior to satellite signal tracking, the navigation processor predicts which satellites should be available. This prediction is based on: a) stored almanac data, b) last known values of user position, velocity, and time (PVT), and c) receiver clock time. The receiver attempts to achieve code correlation with those satellites predicted by the almanac data. Since the almanac data predicts that a subset of the satellite constellation is available for navigation, only a subset of the internal ranging code generator's codes is used for satellite signal correlation.[4] If almanac data is not available and/or the initial PVT data is inaccurate, the receiver "searches the sky" by adjusting its code generator to acquire at least one satellite. Once this satellite is acquired, up-dated almanac parameters are obtained from the NAV message. The receiver uses this new data to predict which satellites should be in view.[5]

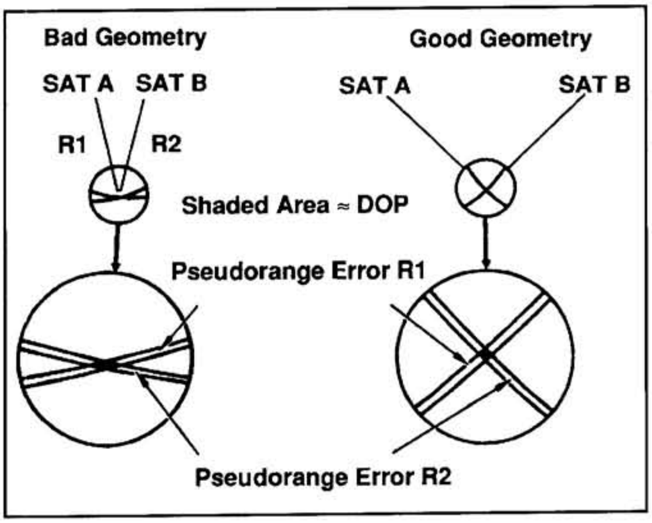

Figure 9. Dilution of precision.

A block diagram of the carrier and code tracking loops for a single receiver channel is shown in **Figure 6**. The downconverted SPS signal, $D(t)\ C(t) \cos(\omega_{IF} t + \theta)$, is input to the mixer and code tracking loop. The code tracking loop generates a pseudorandom code, $C(t-\tau)$ identical to the one being received from the satellite. Time delay τ represents the timing offset between the satellite-generated code and receiver-generated replica code. Correlation is achieved when $\tau=0$. The time shift required to align the code sequences is equal to the time required for the satellite signal to propagate to the receiver. This propagation time, multiplied by the speed of light, is the pseudorange. The pseudorange also contains measurement errors contributed by the clock bias error, initial code synchronization error, ionospheric and tropospheric delays, multipath, and other errors. (Elaboration on pseudorange errors appears in the following section.) After mixer multiplication, the product, $D(t)C(t)C(t-\tau) \cos(\omega_{IF} t + \theta)$, is bandpass filtered. The bandpass filter is equivalent to an integrator and passes only the data modulation and residual doppler frequency uncertainty. The output of the bandpass filter is represented by $D(t)\ V(\tau) \cos(\omega_{IF} t + \theta)$. Where $V(\tau)$ is the autocorrelation function of the code:

$$V(\tau) = \mathrm{Avg}[C(t)C(t-\tau)] = 1 - \frac{\tau}{T} \qquad (13)$$

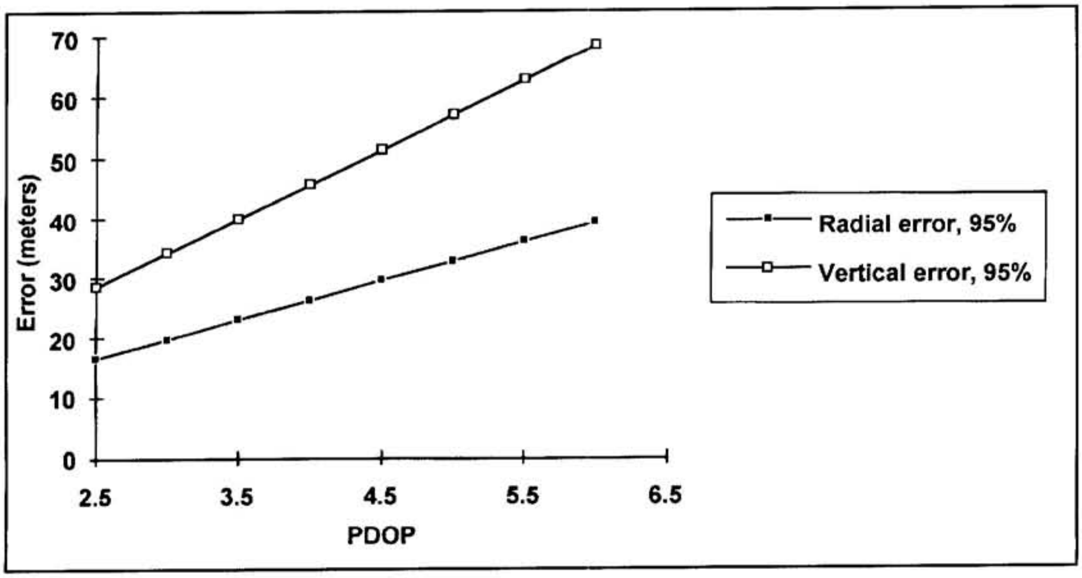

Figure 10. PPS horizontal and vertical error.

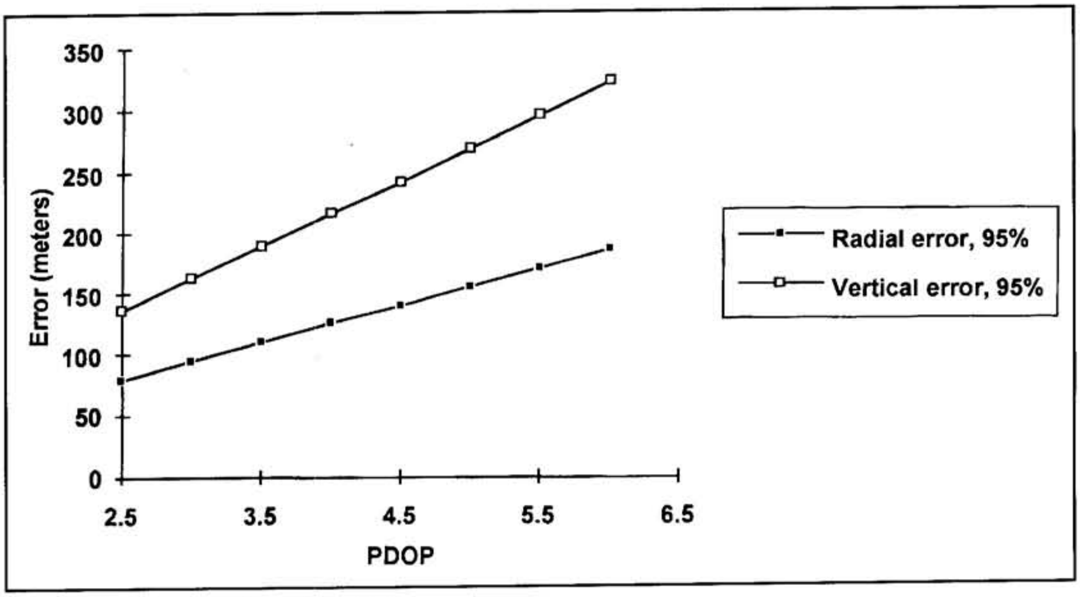

Figure 11. SPS horizontal and vertical error.

The output of the bandpass filter is applied to the Costas loop.

The Costas loop generates an intermediate frequency (IF) close to the expected IF. The difference between the downconverted carrier (expected IF) and the internally generated IF is the Doppler frequency associated with the relative velocity between the satellite and the receiver. It's important to note that this velocity includes the host vehicle dynamics. IF signal tracking occurs after code correlation because the despreading process recovers the carrier.

The Costas loop then adjusts the internally generated IF to match the downconverted carrier frequency.[5] This process yields the relative velocity between the satellite and the receiver. The receiver differences the adjusted IF with the downconverted carrier to obtain the magnitude of the Doppler frequency. This velocity data is used to aid the code tracking loop. NAV data is integrated over the 20 msec bit period within the Costas loop. The integrator output is forwarded to a data decision circuit where it's determined if the data is a one or a zero.

SEGMENT SOURCE	ERROR SOURCE	P-CODE PSEUDO RANGE ERROR 1σ [m]
SPACE	Clock and nav, subsystem stability	-
	Predictability of satellite perturbations	-
	Other	-
CONTROL	Ephemeris prediction model implementation	-
	Other	-
USER	Ionospheric delay	0.2
	Tropospheric delay compensation	0.2
	Alignment error	0.4
	Receiver noise and resolution	0.8
	Multipath	1.2
	Other	-
SYSTEM UERE	Total (Root-sum-square)	1.5

Figure 12. DGPS UERE budget (estimate).

Pseudorange and velocity measurements, as well as the NAV data from each channel, are input to the navigation processor for computation of the navigation solution (e.g., user position, velocity, and time).

The navigation processor contains a Kalman filter that computes optimal estimates of user position, velocity, and time. The filter is generally updated at a 1 Hz rate with new pseudorange and velocity measurements. Receivers designated for aircraft precision approach and other high dynamic applications have measurement Kalman filter update rates of at least 5 Hz. Filter navigation solutions are forwarded to a control display unit (CDU). The CDU permits operator data entry, display status and navigation solution parameters, and usually accesses numerous navigation functions such as waypoint entry, time-to-go, etc.

Performance Accuracy/Primary Error Sources

GPS measurement accuracy is a function of pseudorange measurement error and satellite geometry. Pseudorange measurement error is denoted as the user equivalent range error (UERE). UERE is modeled as a zero mean Gaussian random variable and is comprised of space, control, and user (receiver) segment error contributions. Descriptions of constituent error sources are provided below:

Selective Availability (SA). Under current DOD policy, SA is implemented to deny full system accuracy to unauthorized users (i.e., the civilian community). SA "dithers" each satellite's 10.23 MHz reference oscillator output, corrupting pseudorange measurement accuracy. SA also encrypts the NAV data message so key elements of the navigation data solution (e.g., satellite clock corrections and ephemeris) are reported inaccurately. Authorized users remove SA effects through cryptography.[5,6]

Satellite Clock Error. As mentioned earlier, the 10.23-MHz satellite reference oscillators are "free-running" cesium or rubidium atomic standards. Although these clocks are highly stable, they may deviate up to about 1 msec from GPS system time.[7] An offset of 1 msec translates to a 300 km pseudorange error. This effect is mitigated by the generation of satellite clock correction estimates by the OCS. Since these corrections are estimates of the actual satellite clock errors some residual error remains.

Relativistic Effects. Both Einstein's general and special theories of relativity are factors in the pseudorange measurement process. Relativistic effects result from differences in the gravitational potential between the satellite and user (general theory), as well as differences in satellite-user relative velocity (special theory). The satellite is subjected to only six percent of the gravitation force felt at the Earth's surface; whereas, in most applications the user receiver experiences 100 percent of the Earth's gravitational force. In addition, the satellite velocity is approximately 3.7 km/sec (12,000 ft/sec) while the user receiver velocity is generally on the order of 0.3 km/sec (1,000 ft/sec), or less. Thus, the satellite clock "ticks" at a different rate than the receiver clock.[8] As a result, the satellite signal speeds up as it approaches the Earth causing the observed frequency to increase. To compensate for this time dilation effect, the satellite clock frequency is adjusted to 10.22999999545 MHz (10.23 MHz nominal) before launch.[9]

Ephemeris Prediction Model Implementation. Satellite ephemeris is determined by four OCS monitoring stations simultaneously making pseudorange measurements to a particular satellite. Each station is located at a surveyed position and contains an atomic clock referenced to GPS system time. This process can be considered an inverse GPS measurement system where the monitoring stations are the "satellites" and the satellite is the "user." Therefore, the same set of four simultaneous equations developed in the section on Principle of Operation can be used to solve for the satellite position.[7] At the MCS, optimal estimates of ephemeris data for all satellites are computed and uplinked to the satellites with other NAV data message parameters for rebroadcast to the user. As in the case of the satellite clock corrections, these corrections are estimates and contain a residual error.

Unmodeled Space Perturbations. Unmodeled space perturbations related to uncertainties in the solar radiation pressure and lunar-solar gravitational forces act on the spacecraft, contributing to ephemeris error.

Ionospheric delay. The time delay associated with the satellite signal wavefronts traversing the ionosphere is a function of the ionospheric total electron count (TEC), as well as ray bending. Both effects are due to refraction and are dependent on the character of the ionosphere and the satellite elevation angle. Almost three times as much delay is incurred when viewing satellites at low elevation than at the zenith. For a signal arriving at vertical incidence, the delay ranges from about 10 nsec at night to as much as 50 nsec during the day. At low satellite viewing angles (0 through 10 degrees), the delay can range from 30 nsec at night and up to 150 nsec during the day.[9] Severe delays can be incurred at the geomagnetic equator and the poles, or during magnetic storms caused by solar flares.

Ionospheric delays are inversely proportion-

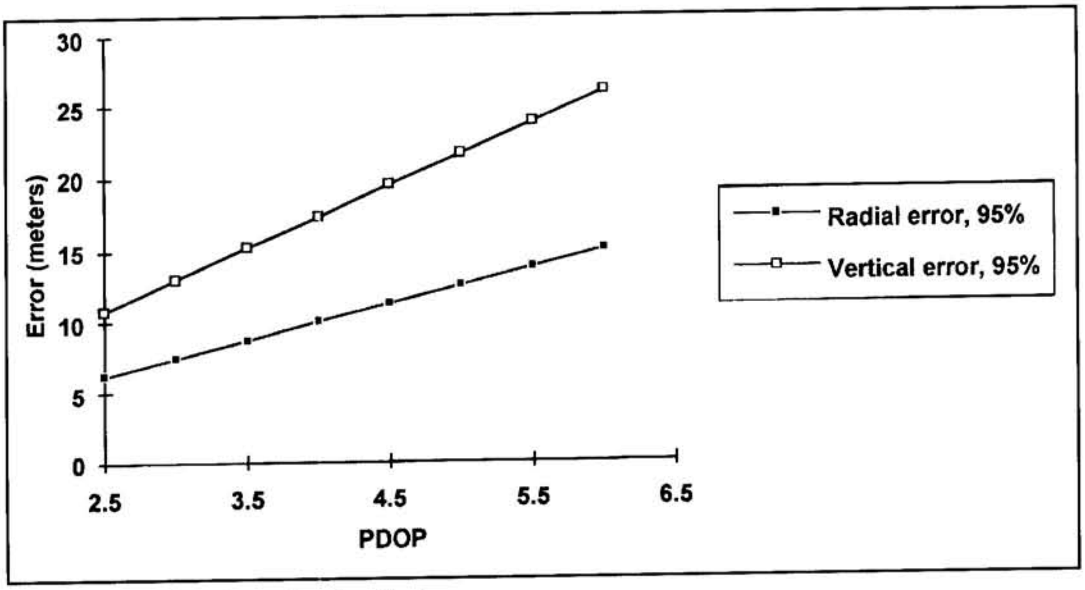

Figure 13. DGPS horizontal and vertical error.

al to the square of the transmitted frequency. As a result, the ionospheric delay can be estimated by making pseudorange measurements to the same satellite on both L1 and L2. The estimated difference in the corresponding time-of-arrival measurements, $\Delta\tau$, is used to calculate the ionospheric delay incurred by each frequency, τ_{dL1}, or τ_{dL2}.

$$\tau_{dL1} = \Delta\tau \, \frac{f_{L2}^2}{f_{L1}^2 - f_{L2}^2} \tag{14}$$

$$\tau_{dL2} = \Delta\tau \, \frac{f_{L1}^2}{f_{L1}^2 - f_{L2}^2} \tag{15}$$

These estimated delay corrections are subtracted from pseudorange measurements made by each frequency; however, because these corrections are only estimates, residual errors are usually present.

Most SPS receivers process C/A-code on L1 only. These receivers compensate for ionospheric delay with an onboard model that uses ionospheric delay correction broadcast in the NAV data message. **Reference 8** reports that the delay can be reduced by approximately 50 percent by employing the onboard compensation model.

Tropospheric Delay. Refraction from the satellite signal transiting the troposphere also induces a time delay into pseudorange measurements. The wavefront is refracted from water vapor and dry atmosphere. These effects are prevalent at low satellite elevation angles because the troposphere is somewhat denser closer to the Earth. Also, the delay is increased when the user receiver is near the Earth's surface rather than at altitude. Compensation for tropospheric delay can be achieved with a mathematical model such as Essen and Froome[9] or Modified Hopfield.[10] Residual tropospheric delay errors contribute to pseudorange error.

Receiver Noise and Resolution. Receiver noise induces errors into pseudorange measurements. Most contemporary high-end receivers are designed to minimize measurement noise. Resolution refers to range error incurred within the ranging code correlation process.

Multipath. The GPS signal is reflected from various surfaces and structures surrounding the antenna. Depending on the angle of incidence and the surface of the reflecting material, reflected ray amplitudes can approximate that of the direct ray (i.e., desired signal). These reflected rays may be incident on the antenna and can combine with the direct ray. Within the receiver, reflected rays tend to distort direct ray ranging signals and induce measurement errors.

UERE Budget. The UERE budget for the P(Y)-code is shown in **Figure 7**. This budget was prepared by the GPS Joint Program Office (JPO) and serves as a guideline for all DOD GPS users.[5] (This budget was developed in the 1980s and has not been updated even though all aspects of today's space, control, and user segments have been improved.) Each budget component is modeled as a zero mean Gaussian random variable over a 24-hour period. Hence, the root-sum-squared (RSS) total of the Gaussian distributions remains Gaussian distributed with a 1-σ value of 6.6 m.

Figure 8 is an estimate of a contemporary SPS receiver C/A-code UERE budget. The space and control segment contributions are identical to those in the P(Y)-code budget except for the inclusion of SA. SA is the dominant error source in SPS C/A-code pseudorange measurements. The value of SA used in this budget was derived from a UERE budget provided in **Reference 11**. With respect to the P(Y)-code budget, user segment error contributions differ for receiver noise and resolution, as well as atmospheric delay components. The value of ionospheric delay is an estimate that reflects delay compensation by an internal receiver model. I believe that today's SPS receiver manufacturers can obtain a better level of tropospheric delay compensation than reported in the 1980-era GPS JPO UERE budget. My estimate of contemporary receiver combined noise and resolution error is also included.

Satellite Geometry. As stated above, navigation solution accuracy is a function of UERE and satellite geometry. For example, assume a receiver's altitude and clock offset from GPS system time are known. The receiver's position can then be obtained by making ranging measurements to two satellites. **Figure 9** shows that the area of uncertainty in user position is minimized when the satellites form a right angle to one another. **Figure 9** indicates that the area of uncertainty increases as the angle formed by the satellites becomes greater than 90 degrees. The error contributed by satellite geometry to a user navigation solution is represented in terms of geometric dilution of precision (GDOP). GDOP is a function of the pointing vectors formed by the user-to-satellite geometric relationship. GDOP is a dimensionless factor that varies with satellite orbits, number of operational satellites, user position, and observation time. GDOP represents the square root of the trace (i.e., diagonal) of a covariance matrix containing estimates of pseudorange measurement errors in the receiver Kalman filter. The trace contains the variances of position errors in three dimensions and receiver clock offset error. In the representation below, it is assumed that all pseudorange errors are independent, zero mean, and have unity variance.

$$GDOP = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_t^2} \qquad \textbf{(16)}$$

In this form, GDOP doesn't have simple physical meaning and can't be used to compute user position or time accuracy. Decomposing GDOP results in:

Position Dilution of Precision (PDOP)

$$PDOP = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \qquad \textbf{(17)}$$

Horizontal Dilution of Precision (HDOP)

$$HDOP = \sqrt{\sigma_x^2 + \sigma_y^2} \qquad \textbf{(18)}$$

Vertical Dilution of Precision (VDOP)

$$VDOP = \sigma_y \qquad \textbf{(19)}$$

Time Dilution of Precision (TDOP)

$$TDOP = \sigma_t \qquad \textbf{(20)}$$

PDOP represents the contribution of satellite geometry to position error in three dimensions. The lateral and vertical contributions are represented by HDOP and VDOP, respectively. TDOP is a measure of error contributed by the satellite geometry incurred when determining the receiver clock bias error.

Computation of Measurement Accuracy. The product of the UERE and DOP factor yield measurement accuracy in a particular dimension. For example, the 2 drms value of horizontal measurement accuracy for a PPS receiver with an HDOP of 1.5 is computed as follows:

$$\begin{aligned} \textit{Horizontal error, 2drms}(95\%) &= 2\ UERE\ HDOP \\ &= 2 * 6.6 * 1.5 \text{ m} \\ &= 19.8 \text{ m} \qquad \textbf{(21)} \end{aligned}$$

Therefore, approximately 95 percent of all measurements will fall within a 19.8 m radius circle centered about the user. VDOP is generally larger than HDOP, and on average may be approximated as:

$$VDOP \approx 1.73\ HDOP \qquad \textbf{(22)}$$

Therefore,

$$PDOP = \sqrt{VDOP^2 + HDOP^2} \qquad \textbf{(23)}$$

$$PDOP \approx \sqrt{HDOP^2 + HDOP^2} \qquad \textbf{(24)}$$

$$\approx 2HDOP$$

Using this approximation, PPS horizontal and vertical errors were calculated for PDOP values ranging from 2.5 to 6. These curves are illustrated in **Figure 10**. **Figure 11** denotes SPS accuracy for the same range of PDOP and the C/A-code UERE delineated in **Figure 7**.

Availability

"Availability is an indication of the ability of a system to provide usable service within the specified coverage area."[2] For most GPS applications, availability is limited by accuracy (DOP criteria). The user needs to be aware of the maximum DOP anticipated for the duration of system usage (i.e., aircraft precision approach, vessel harbor navigation, land or marine surveying, etc.). For example, if the maximum horizontal error for a particular land survey using DGPS is 10 m, 95 percent, the maximum tolerable HDOP is:

Horizontal error, 2drms(95%) = 2 UERE HDOP

$$10 = 2 * 1.5 * \mathrm{HDOP}_{\max} \qquad \textbf{(25)}$$

Therefore,

$$\mathrm{HDOP}_{\max} = 3.33$$

DOP can be predicted by off-the-shelf computer programs. Using data obtained from any GPS bulletin board service, these programs compute all DOP types for a given location and time. The user must also take into account any failed satellites and/or scheduled maintenance actions. The U.S. Coast Guard provides GPS constellation data and status via a dial-up computer bulletin board at (703)313-5910.

Integrity

"Integrity is the ability of a system to provide timely warnings to users when the system should not be used for navigation."[2] A single GPS satellite illuminates approximately one third of the Earth; thus, any satellite failure will affect a large number of users. The OCS checks received monitoring station data for anomalies and issues warnings to users in the event a failure is detected. Integrity is enhanced by the use of differential GPS, DGPS, (described in the next section) and receiver autonomous integrity monitoring, RAIM. RAIM is a receiver-embedded algorithm that uses five or more satellites with good PDOP for fault detection. Other algorithms exist that incorporate measurements from at least six satellites to provide a fault detection and isolation capability.

Differential GPS (DGPS)

DGPS is a technique that offers improvements in GPS performance. DGPS enhances position accuracy by removing correlated errors between two or more receivers viewing the same satellites. Removal of correlated errors enables precise determination of the distance between receivers. Absolute or relative distance measurements can be obtained. DGPS requires processing for correlated error extraction. In real-time applications, a data link capability is required. Error removal can use position, pseudorange, and/or carrier phase information.

Two popular techniques of implementing DGPS are local area DGPS (LADGPS) and wide area DGPS (WADGPS). LADGPS employs receivers located within line-of-sight (LOS) of each other. Correlated errors between two receivers can be removed by operating one GPS receiver (denoted as the reference receiver) at a surveyed location, comparing the electronically derived GPS position to the surveyed position, and broadcasting the difference in positions (i.e., error) to suitably equipped users within the area to correct their navigation solutions. This technique is generally not used because it restricts both receivers to view the same satellites or the reference receiver must compute corrections for all possible combinations of satellites in view. If eight satellites are in view and measurements to four or more satellites are required for a navigation solution, there are 163 possible combinations of four satellites or more to compute. This technique is generally prohibitive for real-time applications due to data latency, unless the user set is commanded to track the same set of satellites as the reference receiver.

Modern reference receivers compute the difference between the "true" reference receiver-to-satellite range (i.e., distance from the surveyed location to the ephemeris-reported position) and the pseudorange. Range differences (i.e., pseudorange measurement errors) for all satellites in view are transmitted via a communications link to users for incorporation in their pseudorange measurements, thereby improving user position accuracy. A ground-based integrity monitoring station assesses the quality of the correction broadcast before and during their transmissions and provides an alert if the correction broadcasts do not meet performance specifications. In general, existing PPS and SPS DGPS equipment performance is similar and can be represented by the estimated UERE budget delineated in **Figure 12**. LADGPS horizontal and vertical position errors are shown in **Figure 13**.

WADGPS is a means for improving user navigation sensor accuracy over a large area (e.g., CONUS and Canada, France, etc.). WADGPS is presently in development by the FAA and is comprised of local monitoring stations distributed across the large area, a central processing facility (CPF), a communications ground Earth station, and geostationary satellites. The system provides ephemeris and satellite clock bias error corrections for each GPS satellite, plus ionospheric time delay parameters. Ephemeris and satellite clock errors, as well as ionospheric delay data, are generated by the ground monitoring stations and transmitted to the CPF via a terrestrial communications link. At the CPF, ephemeris and satellite clock corrections are computed and then forwarded to the ground Earth station for uplinking to geostationary satellites for rebroadcast to suitably equipped users within the satellite coverage area. The corrections are applied to the user receiver pseudoranges and collected ephemeris data in order to improve user navigation accuracy. It's envisioned that WADGPS accuracy performance will be an improvement to standalone GPS, but somewhat less than that provided by LADGPS.

The WADGPS network will also include an integrity failure and user warning capability. Integrity and WADGPS correction data will be generated by the ground monitoring network. One important current design goal is to provide an integrity failure warning to users via geostationary satellite within six seconds.

Applications

A multibillion dollar GPS market is projected. Applications range from leisure hiking to space shuttle landing guidance. Land uses include surveying, vehicle tracking, stolen vehicle recovery, public transportation, and fleet management. Combining GPS data with digitized maps permits a user to visualize her/his real-time position. There are also, several private entities broadcasting DGPS corrections over FM subcarriers. This service requires the purchase of an FM data receiver and is available for a monthly fee. The fee structure is based on the degree of accuracy the user requires.

GPS receivers are in wide use by the commercial/leisure marine community. Some manufacturers offer receivers that float. At present, the U.S. Coast Guard is establishing a coastal DGPS network that includes Alaska, most of CONUS, Hawaii, the Great Lakes, and the U.S. Caribbean Islands. Corrections are broadcast using the existing marine beacon system on approximately 300 kHz.

Current aviation usage encompasses enroute, terminal, and nonprecision approach applications. These phases of flight use standalone GPS. GPS applicability to precision approach is under investigation. The use of GPS with a communications link for automatic aircraft position reporting is expected to revolutionize air traffic control (ATC). This technique is known as automatic dependent surveillance (ADS). At present, aircraft use INU and OMEGA for oceanic en route navigation. These positions are transmitted on an hourly basis via high frequency voice to ATC. ADS would automatically transmit an aircraft's GPS position to ATC approximately every five minutes over a digital data link, providing more accurate and timely position information.

GPS is also used in a multitude of military missions. Several military applications are listed below:

Rendezvous
All-weather air drop
Search and rescue
Close air support
Range Instrumentation
Weapons delivery
Target acquisition aid
Target reconnaissance
Missile guidance
Bomb scoring

Closing comments

GPS is here. The system is operational and has revolutionized the science of navigation by offering better positioning technology than ever. A person can now locate herself/himself nearly anywhere in the world with a relatively low-cost receiver. The system continues to prove its military utility as evidenced in Operation Desert Storm. New civilian and military applications are arising almost daily, and the number of civilian users is expected to outnumber those in the military. Within the next few years, GPS will influence some part of our daily lives. ■

REFERENCES

1. R. Langley, "The Mathematics of GPS," *GPS World Magazine*, July/August 1991.
2. DOD/DOT, *1992 Federal Radionavigation Plan*, January 1993.
3. E. Maloney, *Dutton's Navigation and Piloting*, Naval Institute Press, 1978.
4. R. Keegan, *The New Magnavox GPS Receiver Technology*. Paper presented at Institute of Navigation Technical Meeting, September 1990.
5. NAVSTAR-GPS Joint Program Office, *NAVSTAR GPS USER EQUIPMENT—INTRODUCTION*, Department of Defense, February 1991.
6. K. Doucet and Y. Georgiadow, "The Issue of Selective Availability," *GPS World Magazine*, September/October 1990.
7. R. Milliken and C. Zoller, "Principle of Operation of NAVSTAR and System Characteristics," *GPS Paper in Navigation, Volume 1*, Institute of Navigation, 1980.
8. T. Logsdon, *The NAVSTAR Global Positioning System*, Van Nostrand Reinhold, 1992.
9. J. Spilker, Jr., "GPS Signal Structure and Performance Characteristics," *GPS Papers in Navigation, Volume 1*, Institute of Navigation, 1980.
10. J. Collins, et al., *GPS Theory and Practice*, Second Edition, Springer-Verlag, 1993.
11. Radio Technical Commission for Aeronautics (RTCA) Special Committee 159, *Minimum Operational Standards for Airborne Supplemental Navigation Equipment Using GPS*, RTCA, 1991.