JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 20, No. 1, January–February 1997

von Kármán Lecture

Origins, Evolution, and Future of Satellite Navigation

Bradford W. Parkinson
Stanford University, Stanford, California 94305-4085

The Global Positioning System (or NAVSTAR) has been called the most significant civil spinoff of the cold war. It evolved from technology efforts in the Navy and in the Air Force but was almost canceled before gaining approval by the Department of Defense. Although it offers phenomenal accuracy (centimeters for landing aircraft), it still requires some expansions to satisfy integrity, continuity, and availability requirements. Most of all, to be formally accepted (and certified) internationally, the world will need greater confidence in commitments by the U.S. and greater participation by the world community.

I. Background and Predecessors of Global Positioning System

THE satellite-based Global Positioning System (GPS) has been called the ninth utility. Knowing where you are to the width of a street has spawned a host of new technologies, new uses, and new industries. With a current commercial production rate of over 60,000 sets per month, this system is perhaps the most significant civil spinoff of the cold war.

The operational GPS is the product of the imaginations of a small group of U.S. Air Force officers and supporting civilians, meeting in the Pentagon, over the long Labor Day weekend of 1973. But those imaginations were inspired by a rich history of prior proposals and technology, which had been developed by many talented people.

This paper[1,2] will recount how that historic Pentagon meeting came to happen. How GPS was then developed is also discussed, what has been accomplished to date is summarized, and forecast is made of the next major wave of development, which has been made possible by GPS. To start this historical journey, let us go back to 1957.

A. Dawn of the Space Age

The American public was both fascinated and shocked in October of 1957 to find that the Soviet Union had launched the world's first artificial space satellite, known as Sputnik. On Nov. 3, in the same year, a second jolt occurred with the launch of another Soviet satellite. Sputnik 2 weighed 508 kg and included a dog named Laika. At the same time the U.S. suffered through a series of humiliating launch failures. Finally, on Jan. 31, 1958, Wernher von Braun and the Army Ballistic Missile Agency launched the tiny (13.6 kg) Explorer 1 satellite. This began the U.S.'s venture into space.

In the infancy of that era we were ignorant of many of the essential technologies and processes needed for a successful space program. Component reliability, space radiation, and orbit determination were all opportunities to gain new knowledge. In this quest were found the keys to the development of the first space-based navigation system.

B. One Thing Leads to Another: Creation of Transit

Robert Cannon, the former Air Force Chief Scientist has shared his law of consequence, which explains why everything happens. Simply stated (it always is), the law is: "One thing leads to another." This principle is aptly illustrated by the next events in the development of space-age knowledge that led to GPS.

1. Sputnik Orbital Determination

The advent of Sputnik triggered a flurry of activities to discover the details of the satellite, especially an interest in the exact nature of its orbit. Two researchers (William H. Guier and George C. Wieffenbach) at the Johns Hopkins Applied Physics Laboratory (APL) had carefully studied the radio signal being radiated by Sputnik and were struck by the regularity of its features. Perhaps the most interesting were the pronounced changes in Doppler shift produced by an overflight. Of course, this was caused by accelerations along the line of sight, which were enhanced by the satellite's high speed and low orbital altitude.

Guier and Wieffenbach developed a computer program, which determined the entire Sputnik 1 orbit from Doppler shift data recorded at one site during a single pass of the satellite. This achievement was possible because the satellite's orbit obeyed Kepler's laws of astrodynamics, and the rotation of the Earth removed any ambiguity in the solution. The results were consistent and much simpler in principle than those obtained by either the British or American conventional tracking networks.

Bradford W. Parkinson of Stanford University, the original Department of Defense (DOD) Global Positioning System (GPS) Program Director, has a broad background in management, modern control, astrodynamics, simulation, avionics, and navigation. He manages the NASA/Stanford Relativity Mission, Gravity Probe B and also directs Stanford research on innovative uses of GPS. Degrees are from the U.S. Naval Academy (B.S., 1957), the Massachusetts Institute of Technology (M.S., 1961), and Stanford University (Ph.D., 1966). He is a distinguished graduate of the U.S. Naval War College and was head of the Department of Astronautics and Computer Science at the U.S. Air Force Academy. From 1966–1968 he was an academic instructor for the U.S. Air Force Test Pilot School. From 1972 to 1978 he led concept development and directed the NAVSTAR Joint Program Office for which he received the DOD Superior Performance Award for Best Program Director (1977). He retired as Colonel in 1978. From 1979 to 1980 he served as a group Vice President for Rockwell International. He was vice president and general manager of Intermetrics Inc. He is Chair of the NASA Advisory Council and a member of the Presidential Commission on Air Safety and Security. Parkinson is a member of the American Astronomical Society, Institute of Electrical and Electronics Engineers, Institute of Navigation, and Royal Institute of Navigation (RION). He has received many distinguished awards and authored more than 80 papers on guidance, navigation, and control. He is a Fellow of the AIAA and the RION and a member of the National Academy of Engineering.

Presented as the von Kármán Lecture at the AIAA 34th Aerospace Sciences Meeting, Reno, NV, Jan. 15–19, 1996; received March 20, 1996; revision received Sept. 14, 1996; accepted for publication Sept. 16, 1996.

2. Inversion: Invention of Transit (Where Are We in Two Dimensions?)

One thing led to another when Frank T. McClure recognized that the technique of Guier and Wieffenbach could be inverted: if the satellite's orbit were already known, a radio receiver's unknown position could be determined accurately from the same type of Doppler measurements. Satellites could be developed specifically for the purpose of providing location. Thus the Transit navigation concept was born. A key advantage of this concept is that it provides worldwide coverage, since periodic updates can be obtained with a single, polar-orbiting satellite.

McClure convinced Richard B. Kerschner (then head of APL's Polaris support division) of the concept's viability, and the newly named Transit program was rapidly developed by the Johns Hopkins APL under the sponsorship of the U.S. Navy. Kerschner directed a program that, by mid-1960, had launched two prototype navigation satellites. The system was further refined, and the first position fix was computed by a Polaris submarine (the major application of Transit) in January 1964, using the satellite illustrated in Fig. 1.

As illustrated by Fig. 2, the Transit satellites are in circular, polar orbits about 1075 km high, circling the Earth every 107 min. This constellation of orbits forms a birdcage within which the Earth rotates, carrying us past each orbit in turn. Whenever a satellite passes above the horizon, the user has the opportunity to obtain a single horizontal position fix. The average time interval between fixes varies

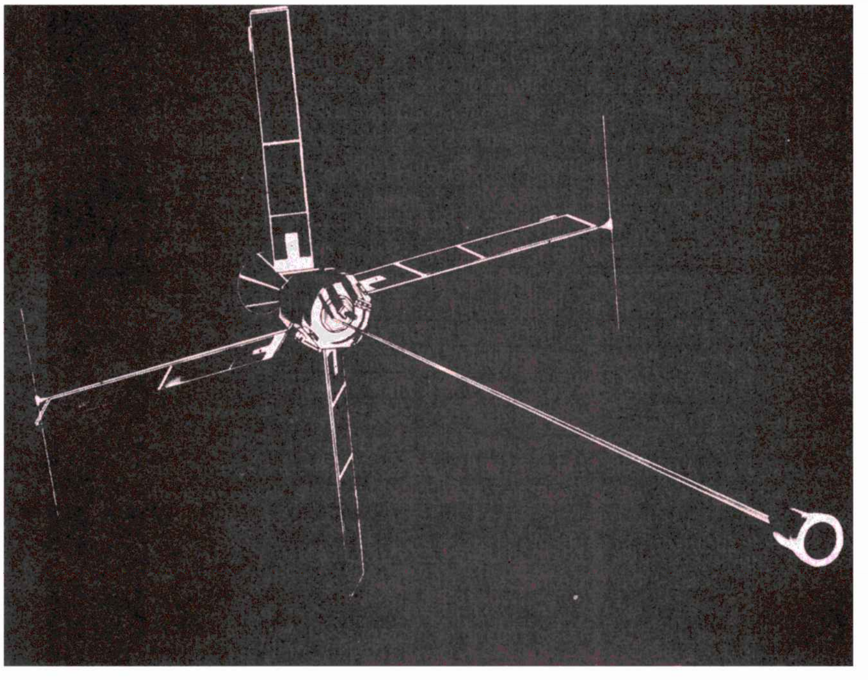


Fig. 1 Physical configuration of Transit satellites. They featured gravity-gradient stabilization and two-frequency broadcast for ionospheric corrections.

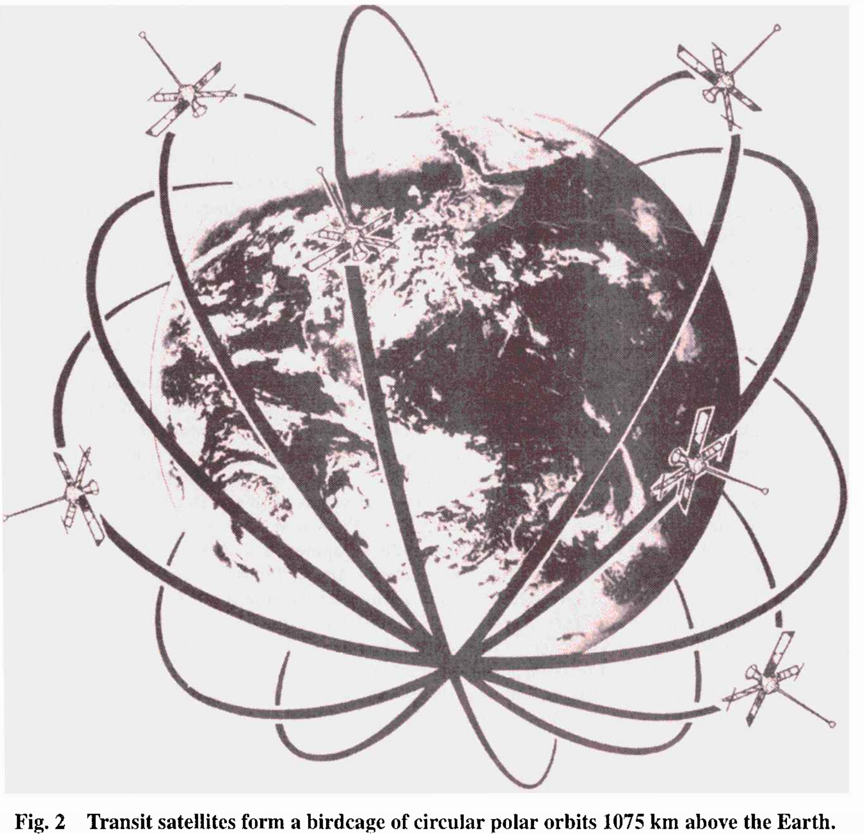


Fig. 2 Transit satellites form a birdcage of circular polar orbits 1075 km above the Earth.

between 35 and 100 min, depending on latitude and the number of operational satellites.

In the operational system, each Transit satellite provides four to six position updates per day for every user. Additional satellites can increase the update frequency and improve the daily distribution of these updates. Note that the number of satellites is limited by the mutual radio interference that they generate.

3. Contributing Technology: Effects of Free Electrons and Prediction of Satellite Location

To build the Transit, the Navy had to develop a substantial number of new technologies and concepts. These included 1) providing the data needed to analyze the Earth's gravity field, 2) testing ionospheric refraction correction techniques, 3) developing a reliable and stable frequency source, and 4) developing reliable mechanical and electronic satellite construction techniques. All of these were important but, as will be explained, the first two laid an essential foundation for the development of GPS. An accurate model of the Earth's gravity field (item 1) is needed for satellite-based navigation because accurate predictions of satellite location are required. The Transit system both depended on this knowledge and helped to improve it. Under APL's sponsorship, Richard Anderle of the Naval Surface Weapon Center rapidly developed the foundations of modern satellite geodesy, which is the science of accurate orbital prediction.

Free electrons in the ionosphere (about 190 n mile above the Earth) produce significant delays in electromagnetic signals. Because these delays are inversely proportional to frequency squared, a signal, which is simultaneously broadcast at two separate frequencies, can be used to calibrate and remove the effect in real time. This technique was pioneered by Transit (item 2), which broadcast on both 150 and 400 MHz. Later, GPS was to use the same technique.

4. Limitations of Transit

There are two principal components of error in a Transit position fix. First is the inherent system error, and second is error introduced by unknown user's (Navy ship's) motion during the satellite pass. The inherent system error can be measured by operating a Transit set at a fixed location and observing the scatter of navigation results.

Dual frequency, static (not moving) Transit receivers typically produce results whose horizontal accuracies are 15–25 m (the two-dimensional rms error). Less expensive single-frequency receivers, which do not measure and remove ionospheric refraction errors, typically achieve horizontal results of 80–100 m (two-dimensional rms).

The second source of position fix error is introduced by unknown motion during the satellite pass. The exact error is a complex function of satellite pass geometry and direction of the velocity error, but a reasonable rule of thumb is that 0.2 n mile of position error will result from each knot of unknown ship velocity. Transit can also be used in aircraft and on land, but user accuracies are degraded by uncertainties in altitude and in three-dimensional velocity.

Thus, the Transit system is not three dimensional, provides only periodic updates, and has degraded accuracy when used by a moving (particularly an airborne) vehicle. These limitations led to the development of GPS. One thing again leading to another.

5. Status of Transit

To complete the Transit story before returning to the main narrative, let us fast forward to the 1980s. Even as the U.S. Government was planning in this period to replace Transit with GPS, additional Transit satellites were still being launched, two at a time. The last of these dual launches occurred in August 1988, at which time there were more active Transit satellites in orbit than at any time during the first two decades of system operation. Because Transit satellites are exceptionally reliable, the system will be near the peak of its capability when it is retired at the end of 1996.

C. TIMATION: Calibrating the Fourth Dimension

By 1972, another Navy satellite system was extending the state of the art by orbiting very precise clocks. Known as TIMATION, these satellites were developed beginning in 1964 under the direction of Roger Easton at the Naval Research Laboratory (NRL). They were principally used to provide very precise time and time transfer between various points on the Earth. In addition, the concept could provide navigation information. The broadcast signals used a technique called side-tone ranging (STR), which broadcasts a variety of synchronized tones to resolve phase ambiguities and directly measure range.

Initially, these spacecraft used very stable quartz-crystal oscillators; later models developed under the GPS program were to orbit the first atomic frequency standards (rubidium and cesium). The atomic clocks were capable of frequency stabilities of several parts in 10^{12} (per day) or better. This level of frequency stability greatly improves the prediction of satellite orbits (ephemerides). Such capability also is used to extend the time between required control segment updates for the operational GPS satellites.

1. TIMATION Concept

The TIMATION project was begun to explore the idea of providing both precise time and accurate position to passive terrestrial observers using range rather than Doppler measurements. In this context, passive means that the user only listens to the satellite broadcasts and does not emit any signals.

Passive ranging required very stable spaceborne clocks that could be regularly updated by a master clock on the ground to synchronize the multiple satellites needed for worldwide coverage. The satellite clocks would be linked to the user's receiving equipment through the one-way ranging signals. Overdetermined data would allow measurement of the difference between user and satellite clocks as part of the navigation process. These concepts were important contributions to GPS and paralleled developments by the U.S. Air Force under a program known as 621B, which will be described subsequently.

2. TIMATION I

The TIMATION I satellite[3] was a small, gravity-gradient stabilized, power-limited satellite that was launched in May 1967 into a 500-n mile polar orbit from the Western Test Range as a secondary payload. An artist's concept of the satellite in orbit is shown in Fig. 3. Its size was approximately 8 × 16 × 32 in., and it weighed 85 lb. In operation it required 12 W of power, and the standby power was 2.5 W.

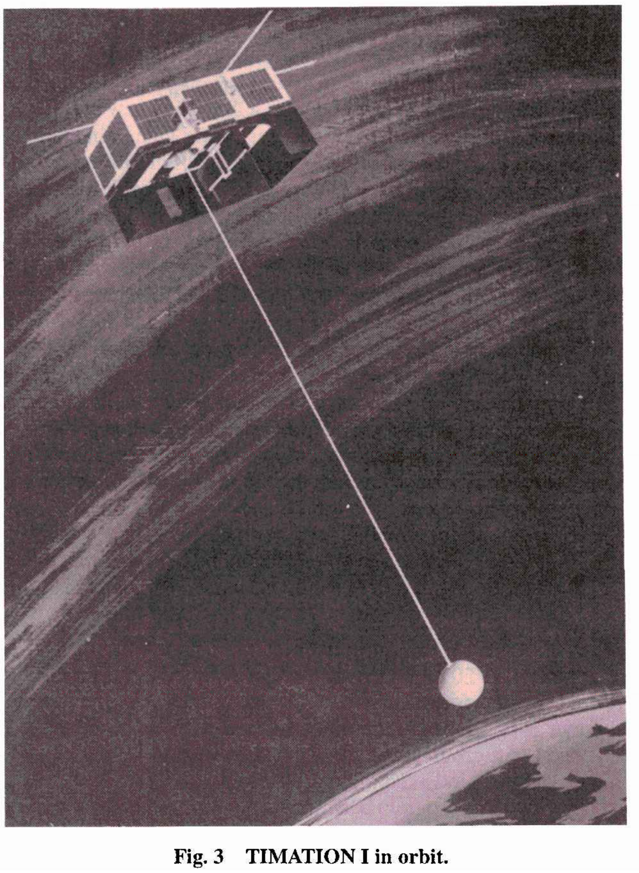


Fig. 3 TIMATION I in orbit.

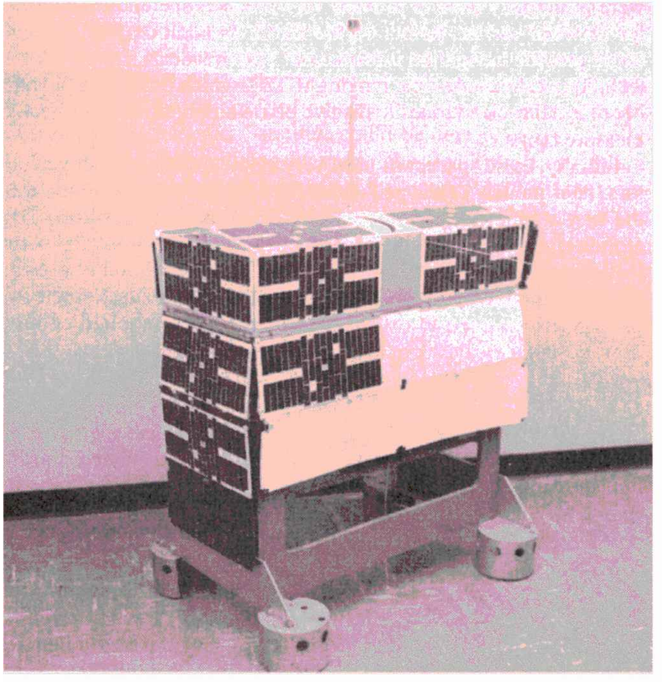


Fig. 4 TIMATION II satellite extended the technology of space-borne precision clocks.

Using a stable space-borne crystal oscillator, a variety of experiments were conducted with TIMATION I to investigate navigation and positioning with different platforms in January–May 1968.

3. TIMATION II

A second satellite, TIMATION II (shown in Fig. 4), was larger, was designed for continuous operation, and was also gravity-gradient stabilized. It was launched into a 500-n mile polar orbit in September 1969 as a secondary payload similar to TIMATION I. It incorporated lessons learned from that first satellite. This included STR signals that were transmitted at both 150 and 400 MHz so that corrections for ionospheric refraction could be made.

D. Defense Navigation Satellite System Definition

In 1968, the Joint Chiefs of Staff (JCS) issued new requirements for precisely locating military forces worldwide. These requirements and the capabilities beginning to be offered by space technology led to the establishment of a Department of Defense (DOD) Navigation Satellite Executive Steering Group to examine the best way to address these requirements with a purpose of establishing a DOD-wide capability.[4,5] Because the most stringent of the JCS navigation requirements were for aircraft, these became the driving parameters.

Design and fabrication of the third in the series of TIMATION satellites was begun in April 1971 to support this DOD program. This satellite was intended to be an experimental version to demonstrate Defense Navigation Satellite System (DNSS) technology and included two experimental rubidium-vapor frequency standards. As the GPS program later developed, this satellite was redesignated as navigation technology satellite (NTS) I and was included in the GPS program. This satellite was launched in July 1974. The on-orbit tests of the experimental rubidium units showed that they had good potential to be used in the operational GPS system.[6]

E. U.S. Air Force Project 621B

Transit and TIMATION were essential foundations for GPS. The third essential forerunner for GPS was a U.S. Air Force program known as 621B. This program was directed by an office in the Advanced Plans group at the Air Force's Space and Missile Organization (SAMSO) in El Segundo, California.

By 1972, this program had already demonstrated the operation of a new type of satellite ranging signal based on pseudorandom noise (PRN). Successful aircraft tests had been run at Holloman Air Force Base to demonstrate the PRN technique. The tests used simulated satellite transmitters located on the floor of the New Mexican desert.

The PRN signal modulation was essentially a repeated digital sequence of fairly random bits, ones or zeros, that possessed certain useful properties. The sequence could be easily generated by using a shift register or, for shorter codes, by simply storing the entire sequence of bits. A navigation user could detect the start (phase) of the repeated sequence and use this for determining the range to a satellite. These signals could be detected even when their power density was less than 1/100th that of ambient noise. Furthermore, all satellites could broadcast on the same nominal frequency because properly selected PRN coding sequences were nearly orthogonal (uncorrelated).

The ability to reject noise also implied a powerful ability to reject most forms of jamming or deliberate interference. In addition, a communication channel could be added by inverting the whole sequence at a slow rate and using these inversions to indicate the ones or zeros of digital data. This slow communication link (50 bps) allowed the user to receive ephemeris (satellite location) and clock information.

The original Air Force concept visualized several constellations of highly eccentric satellite orbits with 24-h periods. Alternative constellations were nicknamed the egg-beater, the rotating X, and the rotating Y configurations because of their resulting ground traces. Whereas these designs allowed the system to be deployed gradually (for example, to cover North and South America first), they had high (and undesirable) line-of-sight accelerations. Initially, the concept relied on continuous measurements from the ground to keep the signals time synchronized. Later the TIMATION clock concept was added because the synchronizing link would have been quite vulnerable. As we will see, GPS did adopt the TIMATION clocks to remove any reliance on continuous ground contact.

II. Failure and Synthesis: GPS is Conceived

Program 621B was the immediate predecessor to the effort that eventually developed the GPS. However, the transition to a true DOD-wide program did not occur until it came perilously close to cancellation, as will be described. A 10-year perspective is described in Ref. 7.

A. Joint Program Office Formed, 1973

In the early 1970s, a number of changes in the systems acquisition process had begun to be adopted for the DOD. These changes, recommended by David Packard, were to have a profound effect on NAVSTAR and other major DOD programs. To increase efficiency and reduce interservice bickering, joint programs were formed, which forced the various services to work together. GPS was one of the earliest examples of this new philosophy. It was decreed to be a joint program, with a Joint Program Office (JPO) located at the Air Force's SAMSO and to have multiservice participation (with the Air Force as the lead service).

The first Program Director was Bradford W. Parkinson, supported by Deputy Program directors—eventually from the Army, Navy, Marine Corps, Defense Mapping Agency, Coast Guard, Air Logistics Command, and NATO. Also continuing their support of 621B were a small cadre of engineers from the Aerospace Corporation under Walter Melton. Parkinson was directed to develop the initial concept as a joint program and to gain approval of the DOD to proceed with full-scale demonstration and development. The joint-service character of the program is illustrated in Fig. 5.

B. Lucky, Chance Encounter

Malcomb Currie was appointed head of Defense Development, Research, and Engineering (DDR&E), third in the DOD chain of command, in early 1973 as part of the incoming administration. He had been living in Los Angeles prior to his appointment and to complete his move he made numerous trips to Los Angeles in the initial months.

One legitimate, official purpose of these trips was to review programs at SAMSO. After a few trips, he had completed all of the high-level reviews that were available, and so the head of SAMSO, Gen. Kenneth Schultz, suggested that he receive an in-depth review

Fig. 5 Joint program included deputy directors from all services. Parkinson is in discussions with his deputy, Bill Huston of the U.S. Navy. Models of the NTS-II and phase one GPS satellites are on the table. The civilian is Frank Butterfield of the Aerospace Corporation. (Photograph courtesy of U.S. Air Force.)

by Parkinson on the space-based navigation concept, then known as 621B. This resulted in a remarkable meeting with the number-three man in all of the U.S. DOD spending about three hours in a small office with a lowly Colonel, talking about engineering, technology, and the wide applications of the proposed system.

With a doctorate in physics, Currie was a keen and quick study, and he had acquired a great deal of space experience from his years at Hughes. The outcome was that the GPS program enjoyed his steadfast support. Without intervention by this key decision maker, the Air Force would have killed the program in favor of additional airplanes. The pivotal (and coincidental) meeting with Parkinson was destined to be an essential factor in gaining system approval. During the early, vulnerable stages of the program, a number of other key supporters emerged who were willing to fight for program approval. These included the Air Force Chief Scientist, Michael Yarymovich, and the President of the Aerospace Corporation, Ivan Getting.

C. Failed Defense System Acquisition and Review Council

Fortunately, the first attempt to gain system approval failed in August 1973. The program that was brought before the Defense System Acquisition and Review Council (DSARC) at that time was not representative of a joint program. Instead it was packaged as the 621B system. As head of DDR&E, Currie made the decision to fail the DSARC.

Immediately after the meeting, Parkinson was called into Currie's office for a private meeting. Currie reiterated his strong support for the idea of a new satellite-based navigation system, but requested that the concept be broadened to embrace the views and requirements of all services. He also told Parkinson that, by making it a truly joint program, he felt confident that it would pass the decision process.

D. New Concept: Synthesis of GPS

Applying a broad, joint-services philosophy, Parkinson and the JPO immediately went to work. In keeping with this new philosophy, Parkinson decided to formulate the new concept at a neutral location. Over the Labor Day weekend of 1973, he assembled about a dozen members of the JPO on the fifth floor of a very quiet Pentagon. He directed the development of a new design that employed the best of all available satellite navigation system concepts and technologies. The result was a system proposal that was not exclusively the concept of any prior system but was instead a synthesis of all of them. The details of the proposed GPS constellation are outlined subsequently.

The multiservice heritage of the new concept precluded any factual basis for further bickering, because all contending parties had contributed to the design. From that point forward, the JPO acted as a true multiservice enterprise, with officers from all military branches attending reviews and meetings that had previously been Air Force only. The TIMATION program was included as an active and important part of the technology development effort.

E. GPS System Description and Technical Design

The operational GPS system of today is virtually identical to the one proposed in 1973. The satellites have expanded their functionality to support additional military capabilities; the orbits are slightly modified, but the equipment designed to work with the original four satellites would still perform that function today. The following section will provide an overview of the system design. The system configuration is illustrated in Fig. 6.

1. Principles of System Operation

The fundamental navigation technique for GPS is to use one-way ranging from the GPS satellites, which are also broadcasting their estimated positions.[8] Ranges are measured to four satellites simultaneously in view by matching (correlating) the incoming PRN signal with a user-generated replica signal and measuring the received phase against the user's (relatively inaccurate) crystal clock. This was the technique pioneered by 621B. With four satellites and appropriate geometry, four unknowns can be determined. They are typically: latitude, longitude, altitude, and a correction to the user's clock. If either altitude or time are already known, a lesser number of satellites can be used.

Each satellite's future position is estimated from ranging measurements taken at a set of worldwide monitoring stations. (The operational control system uses five monitor stations, which are located at Colorado Springs, Ascension Island, Diego Garcia, Kwajalein, and Hawaii.) These ranging measurements use the same signals that are employed by a typical user receiver. Using sophisticated prediction algorithms, the master control station forms estimates of future satellite locations and future satellite clock corrections. For the uploads, which occur daily or (optionally) more frequently, the combined predictions for satellite clock and position have been measured to have an average rms error of 2–3 m. These satellite position estimates have demonstrated reasonable errors even after three days (this limit case gives 24.3 m of expected ranging error). This system concept is shown in Fig. 6.

2. GPS Ranging Signal

The GPS ranging signal was derived from the Air Force 621B program. It is broadcast at two frequencies: a primary signal at 1575.42 MHz (called L1) and a secondary broadcast at 1227.6 MHz (called L2). These signals are generated synchronously, so that a user who receives both signals can directly calibrate the ionospheric group delay and apply appropriate corrections. However, most civilian users will only use the primary (or L1) frequency for reasons that will be explained.

Potentially, both the signal at the L1 frequency and the signal at L2 can each have two modulations at the same time (called phase quadrature). Current implementation has two modulations on the higher frequency (L1) but only a single (protected) modulation on L2. The two modulations are as follows.

1) Clear acquisition (C/A) code, broadcast at 1.023 MHz, is the principal civilian ranging signal, and it is always broadcast in the clear (unencrypted). It is also used to acquire the much longer P code. The use of this signal is called the standard positioning service (SPS). It is always available although, as described subsequently, it may be somewhat degraded.

2) Precise code (P, sometimes called the protected code), a very long code, is broadcast at 10 times the rate of C/A, 10.23 MHz. Because of its higher modulation bandwidth, the code ranging signal is somewhat more precise. This signal provides the precise positioning service (PPS). The military has decided to encrypt this signal in such a way that it will not be available to the unauthorized user. This feature is known as antispoofing. When encrypted, the P code

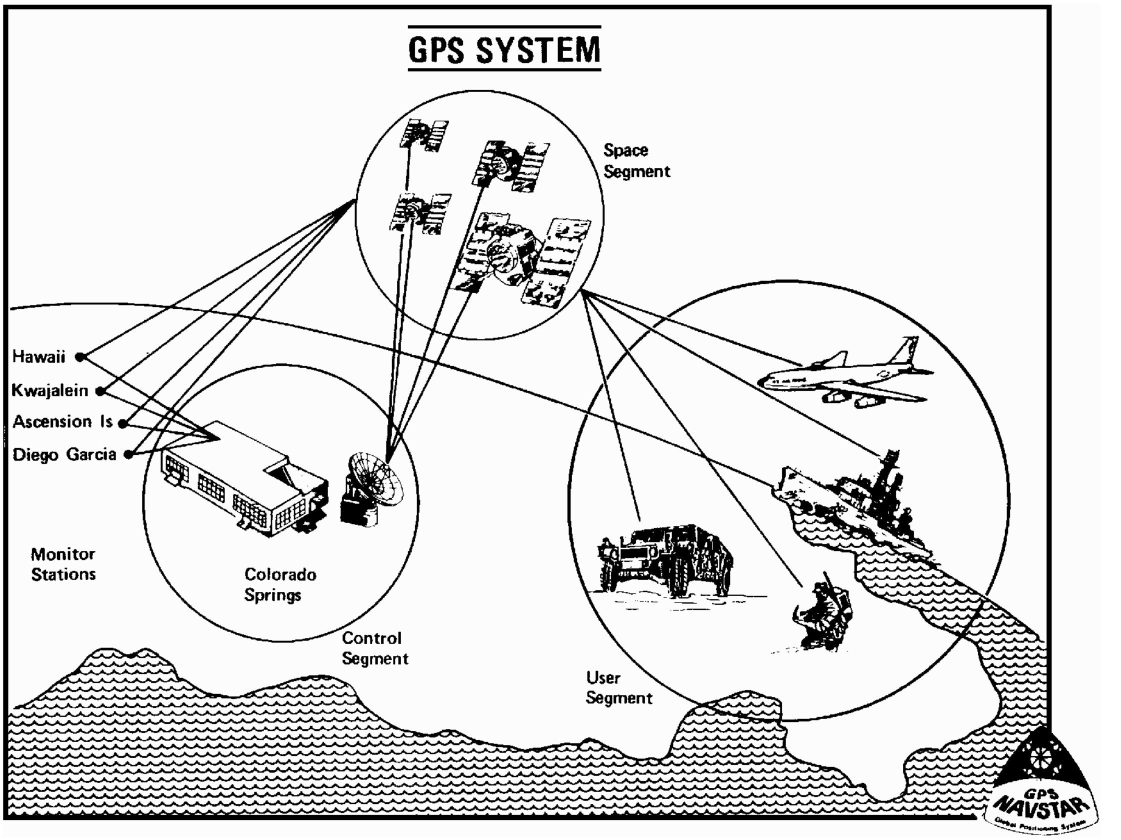


Fig. 6 System configuration of GPS showing the three fundamental segments. (Drawing courtesy of U.S. Air Force.)

becomes the Y (or P/Y) code. (There are provisions in the Federal Radio-Navigation Plan for users with critical national needs to gain access to the P code.)

a) Selective availability. The military operators of the system have the capability to intentionally degrade the accuracy of the C/A signal by desynchronizing the satellite clock or by incorporating small errors into the broadcast ephemeris. This degradation is called selective availability (S/A). The magnitude of these ranging errors is typically 20 m, and results in rms horizontal position errors of about 50 m (one-sigma). The official DOD position is that SPS errors will be limited to 100 m (two-dimensional rms), which is about the 97th percentile. A technique known as differential GPS (DGPS) can overcome this limitation and potentially provide civilian receivers with accuracies sufficient for precision landing of aircraft.

b) Data modulation. One additional feature of the ranging signal is a 50-bps modulation, which is used as a communications link. Through this link, each satellite transmits its location and the correction that should be applied to the spaceborne clock, as well as other information. (Although the atomic clocks are extremely stable, they are running in an uncorrected mode. The clock correction is an adjustment that synchronizes all clocks to GPS time.)

3. Satellite Segment

a) Satellite orbital configuration. The orbital configuration approved at DSARC in 1973 was a total of 24 satellites, 8 in each of three circular rings with inclinations of 63 deg (see Fig. 7). The rings were equally spaced around the equator, and the orbital altitudes were 10,980 n mile. This altitude gave two orbital periods per sidereal day (known as semisynchronous) and produced repeating ground traces with satellites positioned 4 min earlier each day. The altitude was a compromise among user visibility, the need to periodically pass over the Continental U.S. ground/upload stations, and the cost of the U.S. Air Force spacecraft launch boosters.

Three rings of satellites were initially selected because it would be easier to have orbital spares; having only three such spares (one in each ring) would allow easy replacement of any single failure in the whole constellation. This configuration provided a minimum of 6 satellites in view at any given time, with a maximum of 11. As a result of this redundancy, the system was robust in the sense that it could tolerate occasional satellite outages.

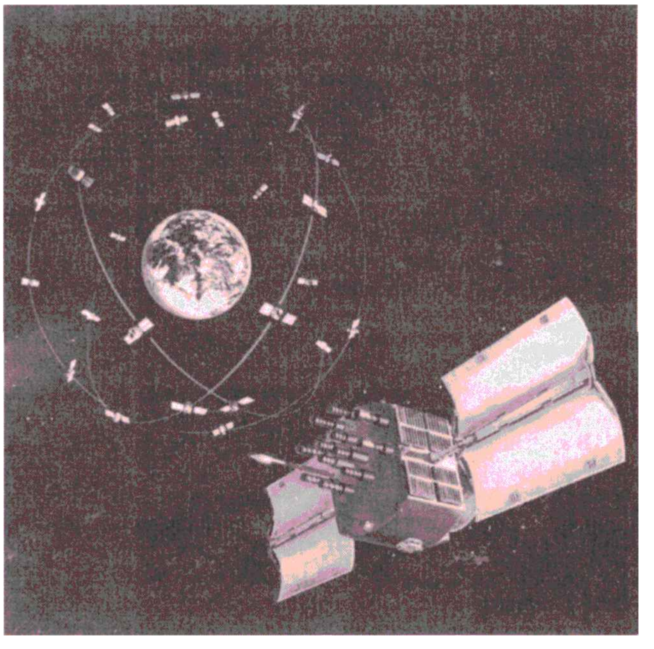


Fig. 7 Original GPS orbital configuration of three rings of eight satellites each. The final operational configuration has the same number of satellites arranged in six rings of four satellites. (Drawing courtesy of U.S. Air Force.)

Two changes have since been made to the original constellation proposal. The inclinations have been reduced to 55 deg, and the number of orbital planes have been increased to six, with four satellites in each. The number of satellites, including spares, remains 24. This constellation still gives six or more satellites in view all the time, virtually everywhere in the world. This assumes that there are no outages, and all satellites more than 5 deg above the horizon can be seen.

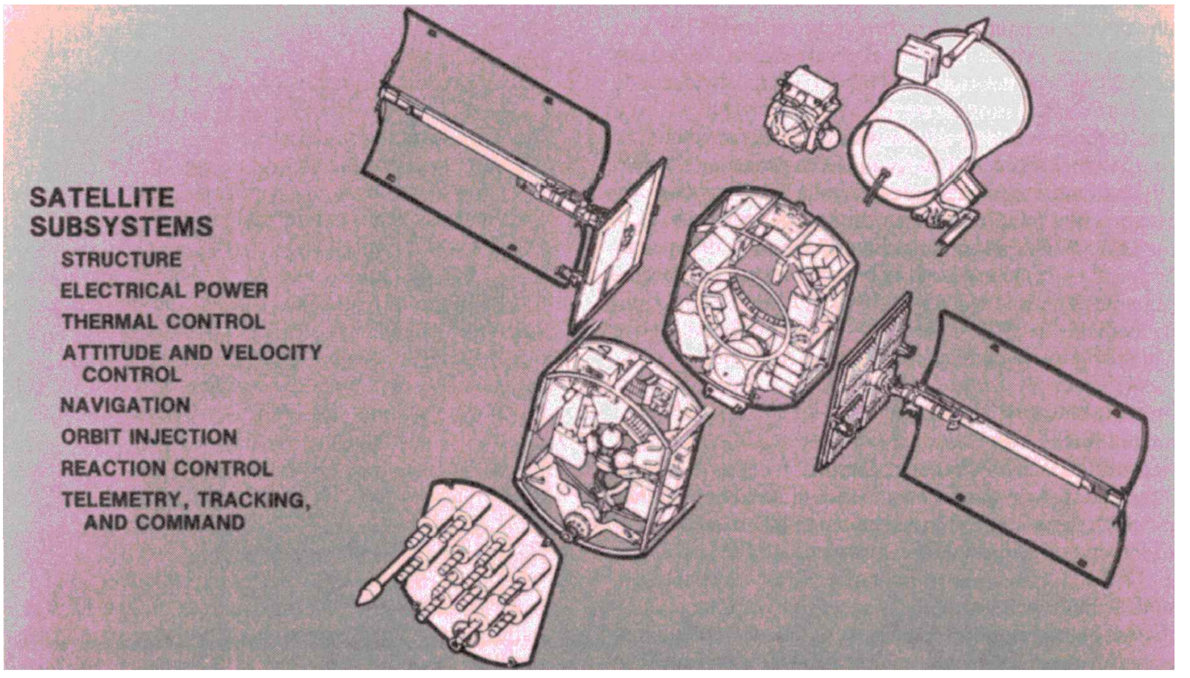


Fig. 8 Breakaway view of the GPS phase one satellite design. Satellite characteristics: weight at booster-satellite separation = 1636 lb, weight at insertion into final orbit = 982 lb, antenna span = 17.5 ft, design life = 5 years, and life of consumables = 7 years. (Drawing courtesy of U.S. Air Force.)

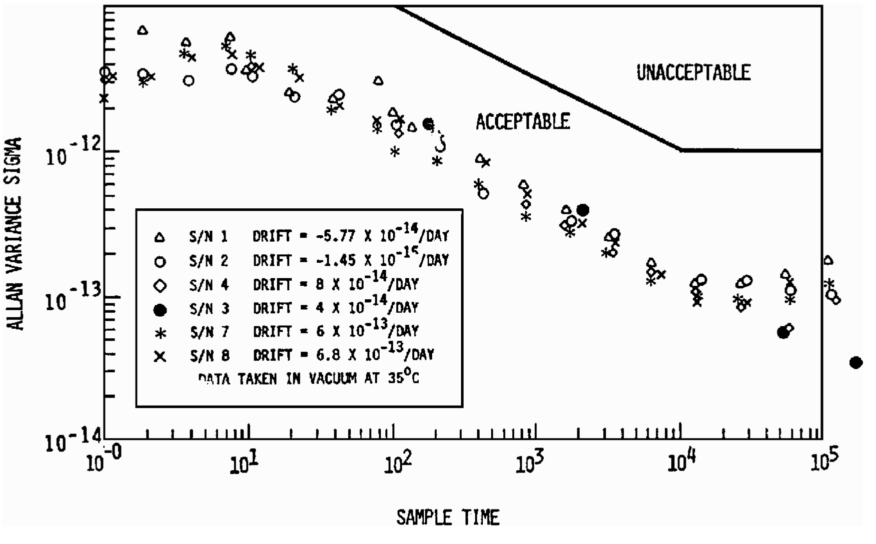


Fig. 9 Space-qualified rubidium-cell frequency-standard performance. These units were developed by Rockwell as a derivative of a clock designed by Efratom Inc. under contract to the NRL. (Data courtesy of U.S. Air Force.)

In the U.S. typical coverage is 6–8 or more in view. The minimum number of 4 satellites (to navigate) is therefore reasonably assured.

Many users would like to have 6 or more to perform cross checks on the integrity of the positioning solution. To attain this at all times is obviously more difficult. This so-called availability problem will be addressed later.

b) Satellite design. All GPS satellites are attitude stabilized on all three axes and use solar panels as the primary power source. The ranging signal is radiated through a shaped beam antenna; by enhancing the received power at the limbs of the Earth, compensation is made for space losses. The user therefore receives fairly constant power for all local elevation angles. (The requirement for received power on L1 is −163 dBW into an isotropic, circularly polarized antenna on the primary frequency.) The satellite design is generally doubly or triply redundant, and the phase one satellites have demonstrated average lifetimes in excess of 5 years (and in some cases over 12). This satellite design is shown in Fig. 8.

c) Satellite autonomy: atomic clocks. A key feature of the GPS design is that the satellites need not be continuously monitored and controlled. To achieve this autonomy, the satellites must be predictable in four dimensions: three of position and one of time. Predictability in orbital position is improved because the high-altitude orbits are virtually unaffected by atmospheric drag. Many other factors that affect orbital position must also be considered. For example, lunar–solar perturbations, solar pressure, and outgassing can all have significant effects.

When GPS was conceived, it was recognized that the most difficult technology problem facing the developers was probably the need to fly accurate long-lived timing standards to ensure that all of the satellites' clocks remained synchronized. As mentioned, the TIMATION program had been developing frequency standards for space, and so this effort was continued and extended.

GPS has traditionally used two types of atomic clocks: rubidium and cesium. Phase one test results for the rubidium cell standard are shown later. A key to outstanding satellite performance has been the stability of the space-qualified atomic clocks, which have exceeded their specifications. They have measured stabilities of one part in 10^{13} over periods of 1–10 days.[9] Test results are shown in Fig. 9.

d) Ionospheric errors and corrections. Free electrons in the ionosphere create a delay in the modulation signal (PRN code). It is not unusual to find delays of over 30 m at lower satellite elevation angles. There are two techniques for correcting this error. The first is to use an ionosphere model. The model parameters are broadcast as part of the GPS 50-bps message. This model is typically accurate to a few meters of vertical ranging error.

The second technique uses both broadcast frequencies and the inverse square law behavior to directly measure the delay. By differencing the code measurements on each frequency; the delay on L1 is approximately 1.546* (difference in delays on L1 and L2). (The delay at 1575 MHz is found as the difference in delay multiplied by $[f_2^2/(f_1^2 - f_2^2)]$ because it is proportional to the inverse frequency squared. The frequency of L2 is 1227 MHz.) This technique is only available to a P/Y code receiver (since only the P code has L2 modulation) or to a codeless (or cross-correlating) receiver.

III. Incubation and Birth: GPS from 1973 to 1978

A. Approval to Proceed with GPS

To gain approval for the new concept, Parkinson began to contact all those with some stake in the decision. After interminable rounds of briefings[16] on the new approach were given to offices in the Pentagon and to the operating armed forces, a successful DSARC was held on Dec. 17, 1973, only three months after GPS was conceived. [Gen. Kenneth Schultz was particularly incensed with the endless presentations that had to be made in the Washington arena. The situation with any bureaucracy is that many can say no and few (if any) can say yes. To bring the nay sayers to neutral, extended trips from Los Angeles to Washington were necessary for Parkinson.] Approval to proceed was granted in a memorandum dated Dec. 22, 1973.

B. Fast Start to Development

The first phase of the program originally included four satellites (one was the refurbished qualification model), the launch vehicles, three varieties of user equipment, a satellite control facility, and an extensive test program.

By June of 1974, the satellite contractor, Rockwell International, had been selected, and the program was well underway. Magnavox, which had been a key participant in the user equipment for 621B (along with Hazeltine), was selected to develop the user equipment under subcontract to General Dynamics, which was also responsible for developing the satellite control segment and the pseudosatellites for the Yuma range.

The initial types of user equipment included sequential (the Y set) and parallel (the X set) satellite-tracking military receivers, as well as a civilian-type set for utility use by the military (the Z set). The development test and evaluation was extensive, with a laser tracking range set up at the Army's Yuma Proving Ground. An independent evaluation was then performed by the Air Force's Test and Evaluation Command.

To maintain the focus of the program, the GPS JPO adopted the following simple and direct motto.

The mission of this program is to:

1) drop 5 bombs in the same hole and

2) build a cheap set that navigates (<$10,000),

and don't you forget it!

The $10,000 price goal was considered very ambitious!

The program developed rapidly; the first prototype operational satellite was launched in February 1978 (44 months after contract start). The design is shown in Fig. 10. By this time, the initial control segment was deployed and working, and five types of user equipment were undergoing preliminary testing at the Yuma Proving Ground. The initial user equipment types had been expanded to include a five-channel set developed by Texas Instruments and a highly jam-resistant set developed by Rockwell Collins.

C. Further Technology Development

The NAVSTAR GPS Development Concept Paper 133, approved in 1973, called for NRL to continue the technology efforts begun in the TIMATION project under JPO direction. A cesium-standard satellite clock was selected as the most promising candidate for meeting the ultimate requirements of the program. The NRL program was to aid in demonstrating the feasibility of the system concept by constructing NTS and to advance the state of the art in navigation satellite technology.

The second technology satellite built by NRL, NTS-II, was the first satellite launched in the NAVSTAR GPS program that was specifically built to the GPS concept. Shown being placed atop the Atlas F launch system at Vandenberg in Fig. 11, it was launched in June 1977. NTS-II contained the first two prototype cesium-beam frequency standards to be flown in space. An essential component was the navigation payload, which was provided by the JPO.

While NTS-II was short lived, the performance of the NTS-II units was a frequency stability of 2 parts in 10^{-13} per day, giving a time error of about 20 ns a day.

D. Needed: A Few More Good Satellites

Because only four satellites were initially approved by the DOD, including a refurbished qualification model, there was a need for spare satellites. (Recall that the minimum number for

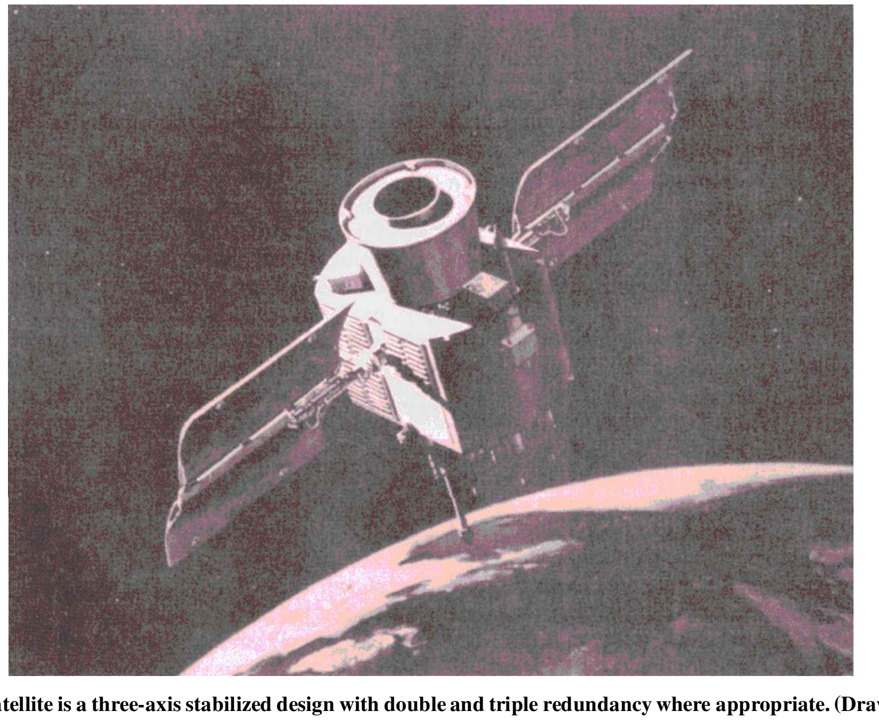


Fig. 10 Phase one GPS satellite is a three-axis stabilized design with double and triple redundancy where appropriate. (Drawing courtesy of U.S. Air Force.)

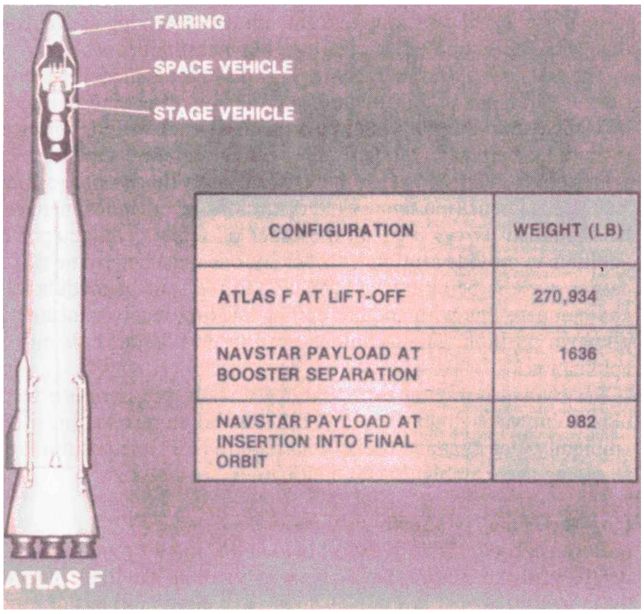


Fig. 11 NTS-II being placed on an Atlas-F booster at Vandenberg Air Force Base.

three-dimensional navigation is four.) Any launch or operational failure would have gravely impacted the phase one demonstration program. Authorization for additional GPS satellites was urgently needed!

The Navy's Transit program inadvertently solved this problem. This chain of events unfolded when Transit requested funds for upgrading certain Transit satellites to a PRN code similar to that used by GPS. The purpose was to provide accurate tracking of the Trident (submarine-launched ballistic missile) booster during test firings into the broad ocean areas. Robert Cooper of DDR&E requested a series of reviews to address whether GPS could fulfill this mission.

The GPS solution was to use a signal translator on the Trident missile bus, which would relay the GPS modulations to the ground on another frequency. The central issues were whether the ionosphere could be adequately calibrated (because it was a single-frequency system, the ionosphere could not be directly measured) and whether the translated signal could be recorded with sufficient fidelity (it required digitizing at 60 MHz).

During the third (and capstone) review for Cooper, Parkinson (supported by James Spilker and John Klobuchar) was able to present convincing arguments that a GPS solution could solve the Trident problem provided two additional satellites were authorized. Cooper immediately made the decision to use GPS. He directed the transfer of $60 million from the Navy to the Air Force. This approved two additional satellites and thereby greatly expanded the phase one test time as well as significantly reduced program risk. This little-known event also eliminated the possibility of an upgraded Transit program competing with the fledgling GPS.

E. Why Is It Called NAVSTAR or GPS?

There has been much speculation regarding the origin of the names GPS and NAVSTAR. The GPS title originated with Gen. "Hank" Stehling, who was the Director of Space for the Air Force Deputy Chief of Staff research and development in the early 1970s. He pointed out to Parkinson that navigation was an inadequate descriptor for the proposed concept. He suggested that Global Positioning System or GPS would be a better name. The JPO enjoyed his sponsorship, and this insightful description was immediately adopted.

The title NAVSTAR came into being in a somewhat similar manner. John Walsh (an Associate Director of DDR&E) was a key decision maker when it came to the budget for strategic programs in general, including the proposed satellite navigation program. In the contention for funding, his support was not as fervent as the JPO would have liked. During a break in informal discussions between John Walsh and Col. Brent Brentnall (the program's representative at DOD), Walsh suggested that NAVSTAR would be a nice sounding name. Brentnall passed this along as a good idea to Parkinson, noting that if Walsh were to name it, he would undoubtedly feel more protective toward it. Parkinson seized this opportunity, and since that time the program has been known as NAVSTAR, the Global Positioning System. While some have assumed that NAVSTAR was an acronym, in fact it was simply a pleasant name that enjoyed the support of a key DOD decision maker. [It is noted that TRW apparently had advocated a navigational system for which NAVSTAR was an acronym (Navigation system timing and ranging). This may have been in Walsh's subliminal memory, but was not part of the process. It was never used as an acronym.]

IV. Coming of Age: GPS from 1978 to the Present

In the early years, the GPS program enjoyed only lukewarm support from the U.S. Air Force because it was viewed as a DOD program. As a result, many attempts were made to cancel its development. This led to key actions to broaden its usefulness and, hence, its support. Initially, it was felt that the full system could be made operational by 1984. Principally due to funding restrictions and redirection, the system was not able to become operational until 1994. Had the phase one satellites simply gone into production, this could have occurred almost 10 years earlier.

A. Add Ons and Delays, the Perils of Innovation

To expand political support, additional operational payloads were incorporated into the baseline design. From its vantage point in space, the NAVSTAR satellites potentially could time nuclear explosions on the ground from many directions and pinpoint their locations. This capability was added along with various others.

These innovations were essential for continuing support but also contributed to delays in the program because they required extensive modifications to the satellites. Steady progress has been made in GPS satellite development, but the real revolution has been the diminishing cost and increased capabilities of GPS user equipment.

B. S/A

To protect the U.S. against hostile use of GPS by enemies, the civilian signal is deliberately degraded in accuracy as described earlier. The name given this degradation is S/A. This action has been a source of international tension over the potential adaptation of GPS as an official navigation standard. Because these effects can be calibrated out by a receiver at a known location on the ground (see later discussions of DGPS), the effectiveness of this technique to deny accuracy has come under severe questioning.

V. Status of GPS: Performance and Test Results

GPS offers two standard services in terms of accuracy. As mentioned previously, the service for military users is known as the PPS and is not affected by the deliberate degradation of S/A. The civilian service is known as the SPS. Both of these standard applications rely on the tracking of the modulation (or code) with no ground-based corrections. More sophisticated techniques apply range corrections computed from ground receivers located at known positions. These so-called differential techniques can attain relative accuracies of better than a centimeter. The following sections briefly summarize the current situation.

A. Standard GPS Navigation Performance

The performance capabilities of standard GPS are primarily affected by two things: satellite geometry (which causes geometric dilution) and ranging errors. Under the assumption of uniform, uncorrelated, zero-mean ranging error statistics, this can be expressed as

$$\text{position error} = (\text{geometric dilution}) * (\text{ranging error})$$

1. Satellite Geometry: Geometric Dilution Effect

Geometric dilution can be calculated for any instantaneous satellite configuration as seen from a user at a particular location. For a 24-satellite constellation and a three-dimensional fix, the world median value of the geometric dilution factor (for the nominal constellation) is about 2.5. This quantity is usually called position dilution of precision (PDOP). Typical dilution factors can range from 1.5 to 8. The variations in this dilution factor are typically much greater than the variations in ranging errors.

2. Ranging Errors

Ranging errors are generally grouped into six major causes: 1) satellite ephemeris, 2) satellite clock, 3) ionospheric group delay, 4) tropospheric delay, 5) multipath, and 6) receiver measurements. With S/A turned off, all errors for single-frequency SPS are nearly identical in magnitude to those for single-frequency PPS except for receiver measurement errors (which decrease with increasing bandwidth). Dual frequency, which is only available on PPS, can reduce the third error (due to the ionosphere) to about 1 m. This is summarized in Table 1. (Multipath errors are generally negligible for path delays that exceed one and one-half modulation chips, expressed as a range. Thus, P code receivers reject reflected signals whose path delay exceeds 150 ft. For the C/A signal, the number is 1500 ft, giving a slight advantage to the P code. However, it is usually reflections from very close objects that are the main sources of difficulty.)

3. Expected Positioning Accuracies for PPS and SPS

The product of the rms PDOP and the ranging error for a single satellite gives the three-dimensional rms position error. Because the value of rms PDOP averages 2.5–3.0 (depending on user location and the assumptions on minimum visible satellite elevation angle, or mask angle), one can multiply 2.5 times the values in Table 1 to find typical positioning errors. Note that PDOP is the same for all users, civilian and military.

A second way to specify accuracy is by spherical error probable (SEP). The SEP is defined to be the radius of the sphere that contains 50% of the errors. For horizontal errors only, similar concepts of circular error probable (CEP) and horizontal rms error can be used to specify accuracy. Typically, the horizontal rms error is about 1.2 times the CEP.

Without the degradation of S/A, SPS would provide solutions with about 50% greater error than single-frequency PPS due to uncompensated ionospheric effects and somewhat greater receiver noise (due to the narrower band C/A signal). It is reasonable to expect that rms horizontal errors for SPS with S/A off would be less than 15 m.

4. Positioning Accuracy Summary for Code Tracking Receivers

Table 2 summarizes the expected positioning accuracies for GPS. It includes a conservative allowance for ionospheric errors.

B. High Accuracy/Carrier Tracking Receivers

A special feature of GPS, which initially was not generally understood, is the ability to obtain an extremely precise ranging signal by reproducing and tracking the rf carrier wave (1575.42 MHz). Because this signal has a wavelength of 19 cm (7.5 in.), tracking it to 1/100th of a wavelength provides a precision of about 2 mm. Generally, carrier tracking techniques can be used in two ways. For normal use, carrier tracking can smooth code tracking and greatly reduce the noise content of code ranging measurements.

The other use of carrier tracking is in a differential mode. There are several variations of this, including surveying, direct measurement of vehicle attitude (with multiple antennas), and various forms of dynamic differential. Modern receivers can attain 2-mm tracking precisions for this second use, but unfortunately this is not accuracy. Reflected signals (multipath) and distortions of the ionosphere can be significant errors (i.e., on the order of a few centimeters). In addition, to provide centimeter accuracy, one must determine which carrier cycle is being tracked (relative to the start of modulation) and compare this with another carrier tracking receiver located at a known position. The technique used to do this depends on the application.

Surveyors, averaging over time for a static position, use techniques known as double or triple differencing to resolve this cycle ambiguity. For dynamic users (who require real-time positioning), resolving cycle ambiguities is a bit harder.

Table 1 Nominal ranging errors for various classes of service

Class of service	Ranging accuracy, m: Single freq.	Dual freq.
PPS	5	3
SPS no S/A	6	N/A
SPS with S/A	20	N/A

Table 2 Expected positioning accuracies for various GPS operating conditions

	PPS		SPS[a] Est. capability	
	Spec. value, m	Measured (static), m	No S/A, m	With S/A, m
Ranging accuracy	6	2.3	6	20
CEP (Horiz.)	——	4.6	12	40
SEP three-dimensional	——	8.3[b]	22	72

[a]SPS results are believed to be conservative.
[b]For dynamic PPS users reported to be less than 10 m.

C. Worldwide Test Results for the PPS

Because each of the five worldwide GPS monitoring stations is continuously measuring the ranging errors to all satellites in view, these measurements are a convenient statistic of the basic, static accuracy of GPS. Table 3 summarizes over 11,000 measurements taken from Jan. 15 to March 3, 1991 during the Desert Storm operation of the Gulf War.[10] The S/A feature was not activated during this period. Note that the PPS results presumably would not be affected by S/A at all.

During this period, one satellite (PRN 9) was ailing but is included in the solution, making the results somewhat worse than they would otherwise be. By dividing the overall SEP by the rms PDOP, an estimate of the effective ranging error can be formed. The average of these results is 2.65 m. [This number is probably somewhat better than an average receiver would measure for several reasons. Monitor station receivers are carefully sited to avoid multipath. The receivers are of excellent quality and are not moving. Also, since the monitor station measurements are used to update the ephemeris, there may be some tuning to make the predictions match any peculiarities (e.g., survey errors) at the monitor station locations. Nonetheless, an average ranging error of 2.3 m is an impressive result.] This should be compared to the specification of 6 m.

D. DGPS Techniques

The ultimate level of accuracy for GPS is attained by calibrating ranging errors with direct, real-time measurements from a calibrating receiver at a known location. These corrections are then applied in the user's receiver to eliminate correlated errors. Errors correlated between user and reference are the dominant sources, particularly when S/A is activated. These techniques are called DGPS since the resulting position is effectively a measurement that is relative to the assumed position of the calibrating receiver. For a detailed explanation of DGPS see Ref. 11.

Not addressed here are the various techniques for transmitting corrections to the user. This is an important consideration in the design of a system and, due to delays or bit errors, can be the largest source of DGPS error. The next sections will summarize some of the various DGPS techniques and provide estimates of their accuracy.

1. Local Area DGPS

This is the simplest and most widely used technique. A receiver is placed in a known location and solves for the corrections (as ranges) to each of the satellites in view. These are then transmitted in some convenient way to the user, who applies them to more accurately solve for his location. This is illustrated in Fig. 12.

A well-designed DGPS system using code measurements can correct ranges down to 0.5–2 m of residual error. When coupled with geometric dilution (note that it still applies), the resulting navigation or positioning errors are on the order of 1–5 m, provided that the user is within 50 km of the calibrating receiver. The efficacy of corrections at longer ranges is discussed in the next section.

The U.S. Coast Guard has deployed a local area DGPS system which uses marine radiobeacons [nondirectional beacons (NDBs)] as data links. These links have added a modulation to the beacon

Table 3 PPS accuracies and implied ranging errors, measured by the GPS monitor stations during desert storm; S/A is off

Criteria	All	Colorado Springs	Ascension Island	Diego Garcia	Kwajalein	Hawaii
SEP three-dimensional, m	8.3	7.8	6.8	9.0	9.1	9.0
CEP[a] two-dimensional, m	4.5	4.5	3.8	5.1	4.6	5.0
RMS PDOP	3.6	3.9	3.4	3.9	3.4	3.3
Est. range error, m[b]	2.6[c]	2.4	2.3	2.7	2.8	3.1

[a]CEP equals the radius of a circle that would contain 50% of the errors. It is the two-dimensional analog of SEP.
[b]This row is formed as follows. Horizontal dilution of precision (HDOP) is approximately PDOP divided by 1.7, and CEP is rms horizontal error divided by 1.2. Because ranging error is horizontal rms error divided by HDOP, this is approximated as 2.04 times CEP divided by PDOP.
[c]Note that this is slightly larger than the result in Table 2, perhaps because of the ailing satellite 9.

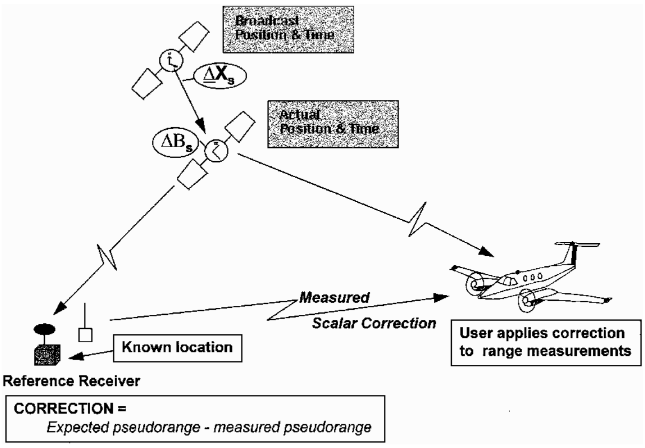


Fig. 12 Local area DGPS, the simplest form of DGPS.

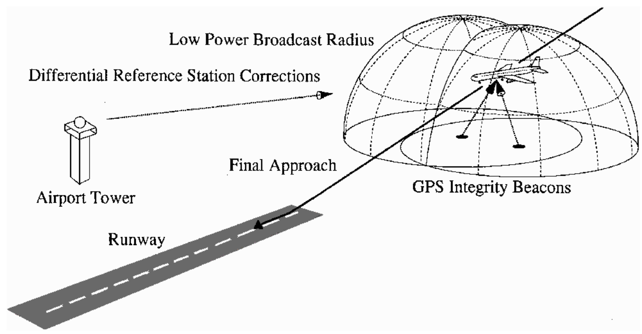


Fig. 13 Carrier DGPS using the Stanford IBLS; accuracies of 2–4 cm are routinely attained.

signal, which provides the corrections in a standardized format. These can be used by marine and other users. Accuracies of 5 m or better have been attained.

Even greater accuracy can be attained by user carrier-tracking receivers provided that the cycle ambiguity can be resolved. There are several ways to do this. A technique that has been demonstrated at Stanford University uses simple, low-power ground transmitters (called integrity beacons) to resolve cycles with very high success rates. If the cycles have been resolved, the ranging errors are reduced to a few centimeters, and dynamic position fixes of better than 10 cm have been demonstrated. This technique [called carrier-tracking DGPS (CDGPS)] is shown as part of the Stanford integrity beacon landing system (IBLS) in Fig. 13.

2. Wide Area DGPS

To increase the operating area of DGPS, a technique called wide-area DGPS (WADGPS) is used. This concept uses multiple calibrating (or reference) receivers to develop vector corrections for the various error sources. Existing systems cover the Gulf of Mexico and the South China Sea. Accuracies of 3–5 m have been reported using WADGPS.

The U.S. Federal Aviation Administration (FAA) is currently embarked on a program to field a WADGPS system called wide area augmentation system (WAAS). It will cover the U.S. and provide corrections through a satellite data link to aircraft and other users.

3. Time and Frequency Transfer

The pioneering use of GPS was probably time and frequency transfer, because a single satellite within the common view of two receivers can easily resolve time to the microsecond or better. This is not surprising since TIMATION was initially focused on providing precise time transfer. The current time transfer capability is reported by the U.S. Naval Observatory to be 20 ns or better.

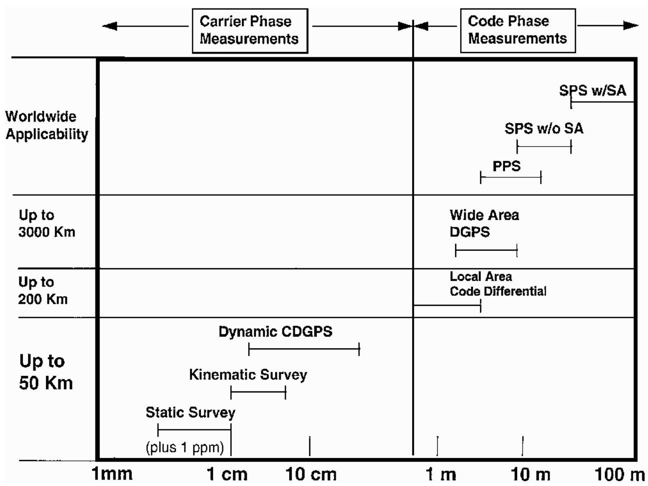


Fig. 14 Range of accuracies attainable with various forms of GPS: horizontal accuracy (1σ).

4. Static Survey

Surveying was the first commercially significant market for GPS because periodic daily coverage provided important economic benefits. By 1986, commercial application in this field was in full swing, and surveying pioneered many of the techniques later employed by dynamic DGPS users. Surveyors rely on carrier tracking and resolve ambiguities by observing satellite motion over times of 30 min to an hour.

Commercially advertised accuracies are 1–2 mm plus one part in 10^6 of range, and this performance is routinely attained. A newer technique called kinematic survey promises much more rapid resolution of integers using dual-frequency receivers.

5. Summary of GPS Accuracies

A summary of the full range of GPS capabilities is shown in Fig. 14.

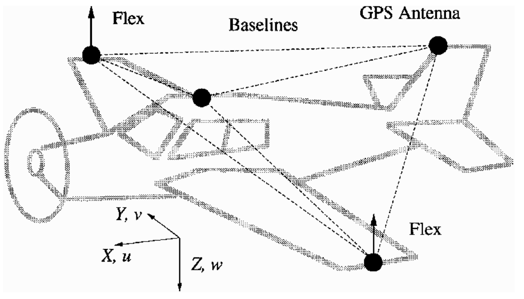
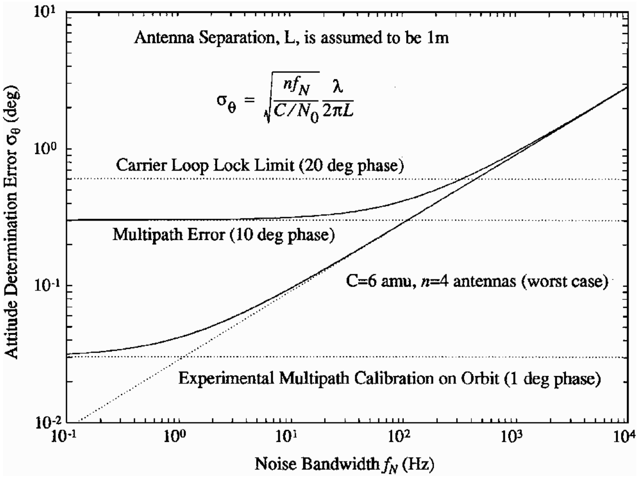


Fig. 15 Antenna configuration for measuring aircraft attitude using CDGPS.



$$\sigma_\theta = \sqrt{\frac{nf_N}{C/N_0}}\frac{\lambda}{2\pi L}$$

Fig. 16 Accuracies of vehicle attitude measurements using CDGPS.

E. Vehicle Attitude

A variation of DGPS uses carrier tracking and multiple receiver antennas attached to a single receiver to dynamically measure vehicle attitude. A typical configuration is shown in Fig. 15. This application was pioneered and demonstrated by Clark Cohen of Stanford University.

The capability to measure attitude is particularly important for vehicle control applications. The accuracy of the measurement is a function of several parameters, including baseline length and sampling rate. Figure 16 shows these tradeoffs. Typical aircraft accuracies have been demonstrated to be 0.1 deg (one sigma) for all three axes. Attitude rate accuracies of 0.5 deg/s (one sigma) are also attainable.

VI. Cold War Responses: Crickets and GLONASS

Because Transit and GPS both were principally developed for U.S. military purposes, a response from the (then) U.S.S.R. was not surprising. The Soviet systems will now be briefly described. Since the breakup of the U.S.S.R., these navigation systems have continued to be supported by Russia.

A. Russia's Cicada (Tsikada)

Tsikada (Cicada) is a passive Doppler satellite navigation system similar to the U.S. Transit system. The motivation for its development was similar as well: its principal use is for warship navigation. The orbiting Tsikada system usually consists of 4 active satellites of the Cosmos-1000 type. These satellites have an orbital inclination of 83 deg with an orbital period of 105 min. Like Transit, they broadcast signals on two frequencies: 150 and 400 MHz.

The Tsikada operational concept is very similar to Transit, with typical horizontal position accuracies of 0.2 miles (two-dimensional rms). Tsikada provides all-weather global coverage for an unlimited number of users. The system continues in operational status: a rocket launched from Plesetsk on Jan. 24, 1995, carried a new Tsikada satellite as one of its payloads.

B. Russia's Global Navigation Satellite System

Development of the Russian (Soviet) satellite radionavigation system (SRNS), the global navigation satellite system (GLONASS), started in the 1970s on the basis of the development and successful operation of the Soviet low-orbital SRNS Tsikada described above. Results of fundamental research on high-precision orbit prediction, general relativity effects, increasing long-term stability of space-qualified atomic clocks, refraction of radio waves in the troposphere and ionosphere, and digital signal processing techniques were incorporated into the design of the new system.

After a large-scale study, the research, development, and experimental work on elements of the system were completed and deployment began. The first space vehicles of the GLONASS series (Cosmos-1413, Cosmos-1414, and Cosmos-1415) were launched into orbit on Oct. 12, 1982. The size of the space segment of GLONASS has continually increased, with one to two launches each year at first and even more launches in the last few years.

By the end of 1988, there were 6 completely functioning satellites in orbit, which were enough to begin full-scale testing of the system. By 1991, there were 12 functioning satellites, enough to give continuous global two-dimensional position fixes.

The GLONASS is intended to provide location, velocity, and precise time for naval, air, land, and other types of users. Like GPS, it was designed for unlimited use by military and civilian users.

Like GPS, the GLONASS is open for use by foreign users as well. It has been pledged that the system will keep its basic characteristics unchanged and will be free to the world for at least the next 15 years. These proposals were given by Soviet (Russian) representatives to International Civil Aviation Organization (ICAO) on May 9, 1988, along with simultaneous disclosure of all technical characteristics necessary for development, manufacturing, and operation of user equipment. GLONASS has now been declared operational. It is planned that by the year 2000 GLONASS will be the basic navigation and precise time aid for all vehicles in Russia.

VII. Selected GPS Applications

Acceptance of GPS has been accelerating. For example, commercial GPS sets are currently being manufactured at a rate of over 60,000 per month. Almost half of these are going into Japanese automobiles. Commercial DGPS is also increasingly available. The U.S. Coast Guard is completing its coverage of all major waterways, including the Great Lakes and the Mississippi River, with the NDB (radiobeacon) version of DGPS. In the air, the FAA has begun the certification of GPS avionics. As the number of GPS receivers accelerates, so have the applications. Rather than repeat the usual applications, let us consider some of the more unusual uses that have been reported.

A. Unique and Unusual Applications

There have been many published lists of GPS applications. Most obvious are the usual navigation uses. Aircraft, ships, trucks, and backpackers are included in virtually all phases of motion or location. As fertile imaginations grasp the potential of the ninth utility, additional innovative applications appear. The following are some examples.

A major commercial application is the use of FM stations to inexpensively broadcast DGPS corrections as an additional modulation that is invisible to the normal listener. With coverage extending across virtually the whole U.S., the additional utility of 2–5 m accuracy is enormous. At least two companies are pursuing this.

The magazine GPS World has an annual contest for unique and unusual applications of GPS from which these examples are extracted. Here is a sample of some of the recent winners.

1) Tracking sheep with GPS is used to correlate their eating habits with the radioactive fallout from the Chernobyl accident. One fix per minute is taken that is accurate to within 5 m over an extended period.

2) DGPS is used to map malaria outbreaks in Kenya. Research should indicate areas and circumstances to be avoided, hopefully reducing the incidence of the disease.

3) Tracking and coordinating the movements of large, parallel overhead cranes, which are used to move lumber, are used to prevent adjacent cranes from crashing into each other.

4) Oil spills are tracked using buoys that are equipped with GPS and a radio system to notify an oil spill response team of the location. Of course, the buoy will tend to drift with the spill.

5) The location and evaluation of the health of electrical power poles are tabulated. This replaces an error-prone manual

data-capture process and is an example of the extensive use of GPS for geographic information systems.

B. Worldwide Humanitarian Use

Throughout the world, conflicts of today and yesterday have inflicted a terrible legacy on the landscape. That legacy is over 20 million buried land mines. The current situation in Bosnia is only the latest; Cambodia, Kuwait, and Somalia are still in our recent memory. According to the Public Broadcasting System, one person in 280 in Cambodia has been injured by a land mine. A potential application of GPS is the clearing and identification of safe corridors through these minefields. The submeter positioning capability of DGPS can ensure identification of cleared areas. Robotic devices for clearing mines could be operated under closed-loop control with DGPS. It may be ironic that a system conceived for war could be used for such an important peacekeeping application.

VIII. Challenges for GPS

Although GPS was designed to be robust, expanded expectations have illuminated aspects that probably call for increased capabilities or resilience. The following sections outline these challenges.

A. Air: Integrity Challenge

Integrity is the technical term used by the FAA to describe the confidence in a measurement of aircraft position. It is usually measured as integrity risk, which is the probability that the error of the indicated position exceeds some threshold error value. There is great sensitivity to this aspect of navigation, because a significant position error can lead to substantial loss of life and erosion of confidence in airline travel.

1. Integrity Problem

While autolanding an aircraft (FAA category III), the positioning errors cannot exceed a maximum safe error limit more than once in a billion landings. Whereas this is the most stringent stated requirement, virtually all civil uses have some implied or stated integrity specification. Therefore, the challenge is to provide a positioning signal that meets these difficult requirements.

GPS satellites will eventually fail and create holes in the constellation. These outages could be a major cause of reduced system integrity until they can be replaced. A rash of generic satellite failures would be difficult to immediately replace.

2. Potential Integrity Solutions

Potential solutions to this challenge are being explored by the FAA and others. They include the following.

1) Ground monitor: all DGPS systems are naturally integrity monitors.

2) Cross check in the user receiver [called receiver autonomous integrity monitoring (RAIM)] using redundant ranging measurements from more than the minimal set of satellites. Usually, at least six measurements are needed for a high confidence in integrity.

3) Additional or supplemental navigation satellites can be of enormous benefit, especially when possible outages are considered. Strong arguments can be made for civil supplements.

B. All Users: Availability Challenge

For most users, four satellites must be available for a navigation solution. If the user is to determine integrity using cross checks with redundant satellites (RAIM), generally 6 satellites or more must be available. Availability is defined to mean that the necessary ranging signals are available to commence an operation requiring positioning at a specified performance level. Availability requirements are determined by the particular application. For some aircraft applications, better than 99% availability is required.

1. Availability Problem

Less than the minimum required number of satellites (4–6) may be available for a variety of reasons.

1) Satellite outages produce a hole in the usual coverage. As the satellites continue their orbits, this hole will move, so that the local outage will not be permanent.

2) Local geography may mask the satellites. This includes mountains, buildings, and vegetation.

3) Vehicle attitude may cause the user antenna pattern to have insufficient gain in certain directions.

4) Local radio interference may prevent the user from receiving the GPS signal.

2. Potential Availability Solutions

A user can reduce the availability problem by increasing his antenna coverage (e.g., dual antennas), or by using other measurements (such as precise time or altitude) to reduce the requirement for four satellites in view.

A more universal solution to this problem is to supplement GPS with additional ranging sources. Supplementary satellites could be placed in GPS type orbits or at geosynchronous altitude. By making the satellite coverage more dense, both outages and local shading impacts would be lessened.

Another solution, particularly useful for the landing of aircraft, is the use of GPS type transmitters from the ground. These are generally called pseudosatellites or pseudolites. One particular type of pseudolite is the Stanford integrity beacon, which also provides a means of resolving integers for the highly accurate CDGPS aircraft landing system shown in Fig. 13.

C. Ground: Continuity Challenge

Continuity for positioning systems means that an operation is not interrupted because of a lapse in measurements after the operation begins. This is somewhat different from availability, which requires that positioning measurements are available at the beginning of an operation. Continuity and integrity may be, to different degrees, safety issues, whereas availability tends to address the economy or efficiency of a system.

1. Continuity Problem

Continuity is particularly a problem for ground users due to intermittent shading of the satellite signals. Travel through cities exposes the ground user to urban canyons, which may limit the number of satellites to one or two. Tilting of the ground vehicle can aggravate the continuity problem because of modified antenna coverage. Complicating the reduced coverage problem is the need to reacquire GPS after an outage.

2. Potential Continuity Solutions

Ground users can supplement GPS with wheel counters and magnetic compasses to automatically flywheel through low-visibility periods. Additional GPS satellites (or supplementary payloads on other satellites as already described) would provide substantial help. Ground transmitters will only be useful in a local area because the GPS signal is strictly line of sight.

IX. Next Wave: Coupling Precise GPS Positioning to Vehicle Guidance and Control

A. Rationale

Most initial uses for GPS were as replacement technology for applications that were already established. DGPS has rapidly improved accuracy, while improved receiver technology has expanded the set of measurable quantities. A single GPS receiver can now measure 13 dimensions of position (more properly states) for an airplane or other vehicle. This is summarized in Table 4.

These 13 simultaneous GPS measurements of vehicle state represent opportunities for expanded use. This next wave will probably include closed-loop control of a wide variety of vehicles using the power of the 13 GPS dimensions. Accurate control using navigation satellites (ACUNS) is the name given by Parkinson to this set of applications. The value of this class of uses includes greater utility, safety, and productivity. This section will briefly explore these opportunities.

B. Vehicle Control Status

A number of examples of automatic control using GPS are currently being developed. Their status and prospects are summarized in the following.

Table 4 Capability of a single GPS aviation receiver: the 13 dimensions

State	Accuracy	Comments
Three dimensions of position		
Non-DGPS	20–50 m	Nonprecision approach
Local DGPS	1–5 m	Precision approach
Local CDGPS	5–10 cm	Automatic landing
Three dimensions of velocity		
Non-DGPS	0.3 m/s	Improved guidance
Local DGPS	0.05 m/s	Precision approach
Local CDGPS	0.02 m/s	Automatic landing
Three dimensions of attitude	0.1 deg	Improved and automatic guidance
Three dimensions of attitude rate	0.5 deg/s	Improved and automatic guidance
Precise time	$<1 \mu s$	Time coordinated operations

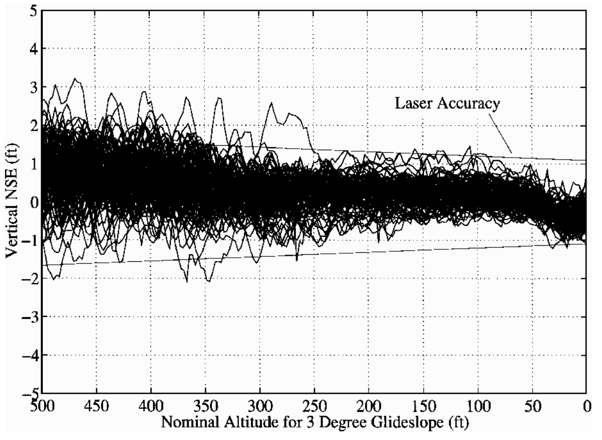


Fig. 17 Vertical GPS sensor error for 110 autolandings of a United Boeing 737.

1. Autolanding of Aircraft

Enormous FAA attention has focused on this application since the announcement that the U.S. would no longer support the development of the microwave landing system. Major FAA-sponsored contractors are working to establish GPS-based landing feasibility and to select the best concept. In October of 1994, Stanford University teamed up with United Airlines, under FAA sponsorship, to demonstrate full category III autolanding with a commercial transport, a Boeing 737-300. The resulting vertical accuracy is shown in Fig. 17.

The estimated 2–4 cm position accuracy was obtained using the variety of pseudolites called integrity beacons. The GPS system certainly was more accurate than the laser tracker, which calibrated the 110 resulting landings. Current analysis shows that the landing system based on integrity beacons is the only configuration that provides the integrity required of the FAA category III landing systems.

Surprisingly, the issue for autolanding is not accuracy but is instead the high demand for integrity. As mentioned, there must be a negligible ($<10^{-9}$) probability of the system indicating a position outside the specified protection limit. A number of organizations are working on solutions. There is high confidence that this aspect of the next wave of GPS development will be successful. It should lead to a reliable autolanding system that will be available at hundreds of airports that do not yet have this capability.

2. Autonomous Model Aircraft and Helicopters

Research has also explored GPS-based guidance for unmanned air vehicles. Two examples at Stanford have been a large model airplane and a model helicopter.

Figure 18 shows a flight in which a model aircraft took off, flew a square pattern, and returned to a landing (within less than half a meter of the designated path) without any human intervention. The winds were large relative to aircraft velocity, yet the model plane held to its flight path within 0.5 m except in the turns.

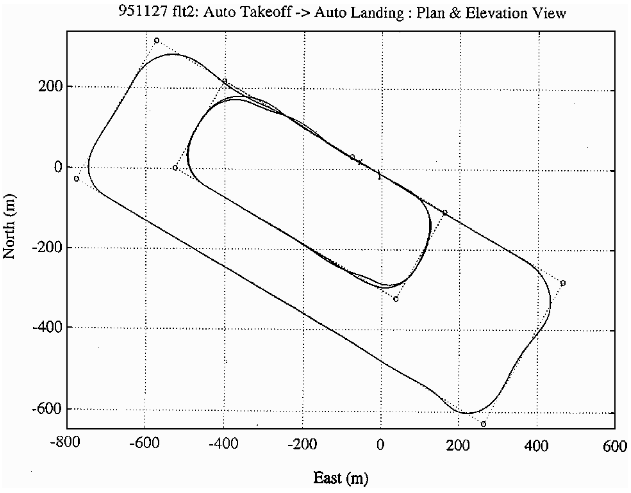


Fig. 18 Fully autonomous flight of a model aircraft with GPS-based navigation and control, including takeoff and landing.

3. Autonomous Farm Tractor

The operation of farm tractors is very manpower intensive. It is exacting yet very repetitive: some fields in California require over 2 h to make a single pass around them. A current program at Stanford is aimed at providing closed-loop guidance of such tractors, initially supervised by an onboard operator. It is hoped that full robotic operation will be also feasible with proper attention to safety and integrity. With CDGPS, potential positioning accuracies of a few centimeters should be sufficient for the most stringent farm tractor operations. The Stanford program has already demonstrated closed-loop control of large John Deere tractor (on a rough field) to an accuracy of 3–5 cm.

C. Some Further Guidance Opportunities

1. Aircraft Guidance With or Without a Pilot in the Control Loop

a) Parallel runways. Key airports such as San Francisco cannot use both of their main runways for blind landings during periods of low visibility. The main reason is the navigation inaccuracy of the instrument landing system (ILS) during the approach at long distances (5–10 mile) from the airfield. DGPS accuracy does not significantly degrade with distance. In addition, GPS can support gradually curved approaches whose separation distances are large at the longer distances and only gradually converge down to touchdown, thereby minimizing risk and user hazard. Beam-riding systems such as ILS cannot do this.

b) Collision avoidance. With the precise three-dimensional position and velocity that GPS provides to an airplane, a generalized broadcast of these quantities to other aircraft in the vicinity allows all users to calculate the probability of collision. If necessary, evasive maneuvers can be undertaken and closely watched using the same information.

c) Autonomous cargo aircraft. Whereas the potential of larger unmanned aircraft may seem farfetched, there are economic incentives for use, particularly for cargo aircraft. One could foresee an evolution in which large aircraft fly with only one pilot as an emergency backup operator for cargo or overnight mail. Of course, issues of safety, reliability, and integrity would have to be resolved. The military has flirted with remotely piloted vehicles (RPVs), most recently in Bosnia. As our society increasingly expects minimal risk to humans, the military will be correspondingly motivated to avoid risks to human life through use of GPS-guided RPVs.

An analogous civilian application is autonomous crop spraying. CDGPS not only has the accuracy for this mission but, when coupled to atmospheric data, the wind vector (and chemical dispersion patterns) can be calculated in real time. This would reduce the hazard of misapplication and would probably increase the efficiency of dispersion.

2. Automotive Guidance Opportunities

The intelligent transportation systems program is a major new Congressional effort that will have a substantial impact over the next 20 years. One of its key technologies is mobile positioning, which has GPS as a key ingredient. Uses will span everything from vehicle surveillance to emergency notification.

3. Equipment Guidance: Construction and Agriculture

The use of GPS for farm tractors has already been mentioned, but this can be expanded into the precise control of virtually all forms of heavy construction equipment using CDGPS. Examples include mining, road building, and pile driving. Some pioneers are now starting to experiment with these applications.

X. Navigation Satellites, an International Resource: Need for Cooperation

A. Major Issue

Like it or not, GPS is truly an international resource. The U.S. DOD views this with concern, since GPS navigation can potentially be used by an enemy. This legitimate fear has led to the continuation of S/A long after it has been shown to be ineffective against a user who employs the crudest form of DGPS.

The major issue is this: GPS has shown a capability vastly superior to any competing technology, yet it will not be established as a standard until the international community is comfortable with the U.S. commitment to uninterrupted service and to some form of shared control. The U.S. will not be comfortable with shared control unless they have some means of ensuring that the system will not be used against them by an enemy and will not relinquish control without others sharing the economic burden.

All non-U.S. countries do not feel so uncomfortable with the current situation, but enough members of ICAO are concerned (particularly in Europe) that GPS alone probably cannot yet be established as an international standard. The details are many, but the essence of the problem is the perception that they cannot depend on GPS.

Lurking in the background is a world community of GPS and DGPS users that is expanding at 1 million per year and accelerating in rate of growth. Any major changes or additions to GPS that are not compatible with the pre-existing user equipment will cause howls of discontent. Any attempt to legislate or otherwise rule that their equipment is unusable will be met with great resistance. In essence, they have voted with their pocketbooks.

B. Path to a Solution

To date, this issue has mostly been considered by ICAO, although it is broader than aviation. Perhaps it is time for the U.S. to propose a solution that addresses these international concerns. Such an agreement would be the basis for a truly international GPS [or, as it is called, the Global Navigation Satellite System (GNSS)]. It might include the following elements.

1) Regions and individual countries would be encouraged to supplement GPS with their own (probably geosynchronous) satellites. Such satellites would broadcast, as a minimum, a GPS SPS signal using an on-board atomic clock. (The stable clock would ensure that the resulting ranging accuracy would be consistent with the existing GPS system; hence, it would not degrade accuracy.) They would also include DGPS corrections as part of an integrity message. Prototyping of this has already begun in the U.S., and Japan has proceeded with a prototype satellite system called MTSAT.

2) All countries that have contributed in proportion to their gross national product would be deemed authorized users and would be entitled to a proportional (to their contribution) vote in setting international standards for GNSS.

3) An international oversight board for GNSS would be set up with powers as determined by common agreement. Operation of individual elements would be delegated to contributing states but would be monitored to ensure that minimum standards of quality are met.

GPS is truly an instrument to unify locations and peoples throughout the world. When global was selected as its first name, this was the intent. Hopefully this unification will be fostered with an international spirit of trust and cooperation.

Acknowledgments

In closing, I would like to acknowledge and thank the people who worked for and with me as part of the original JPO. They overcame adversity, ignorance, and uncertainty. They steadfastly adhered to the vision. The destiny of engineers and builders is to be rapidly forgotten by the public. I, for one, will not forget either their sacrifices or their achievements. The historical discussion draws heavily from the first volume of *Theory and Applications of GPS*. It also makes use of "A History of Satellite Navigation," for the non-GPS material. I thank my coauthors of that article for their contributions.

References

[1]Parkinson, B. W., "Introduction and Heritage of NAVSTAR, the Global Positioning System," *Global Positioning System: Theory and Applications*, Vol. 1, Progress in Aeronautics and Astronautics, edited by B. Parkinson, J. Spilker, P. Axelrad, and P. Enge, AIAA, Washington, DC, 1996, Chap. 1.

[2]Parkinson, B. W., Stansell, T., Beard, R., and Gromov, K., "A History of Satellite Navigation," *Navigation, Journal of the Institute of Navigation*, Vol. 42, No. 1, 1995, pp. 109–164.

[3]Anon., "The TIMATION I Satellite," Naval Research Lab., NRL Rept. 6781, Space Applications Branch, Washington, DC, Nov. 1968.

[4]Anon., "Navigation Satellite Constellation Study, Final Report," RCA Astro Electronics Div., RCA Rept. AED R-3632F, Princeton, NJ, Jan. 1971.

[5]Anon., "Medium-Altitude Navigation Satellite System Definition Study," RCA Astro Electronics Div., RCA Rept. AED R-3665-F, Princeton, NJ, April 1971.

[6]McCaskill, T., and Buisson, J., "NTS-1 (TIMATION III) Quartz-and-Rubidium Oscillator Frequency Stability Results," Naval Research Lab., NRL Rept. 7932, Washington, DC, Dec. 1975.

[7]Parkinson, B. W., and Gilbert, S. W., "NAVSTAR: Global Positioning System—Ten Years Later," *Proceedings of the IEEE*, Vol. 71, No. 10, 1983, pp. 1177–1186.

[8]Parkinson, B. W., "The Global Positioning System," *Bulletin Geodesique*, Paris, No. 53, 1979, pp. 89–108.

[9]Bowen, R., Swanson, P. L., Winn, F. B., Rhodus, N. W., and Fees, W. A., "GPS Control System Accuracies," *Global Positioning System, Vol. III*, Inst. of Navigation, Washington, DC, 1986, p. 250.

[10]Sharrett, Wysocki, Freeland, Brown, and Netherland, "GPS Performance: An Initial Assessment," *Proceedings, ION GPS-91*, Institute of Navigation, Washington, DC, 1991.

[11]Parkinson, B. W., and Enge, P., "Differential GPS," *Global Positioning System: Theory and Applications*, Vol. 2, Progress in Aeronautics and Astronautics, edited by B. Parkinson, J. Spilker, P. Axelrad, and P. Enge, AIAA, Washington, DC, 1996, Chap. 1.