

12-1-2006

# What Does *Height* Really Mean? Part IV: GPS Orthometric Heighting

Thomas H. Meyer

*University of Connecticut Department of Natural Resources Management and Engineering*, [thomas.meyer@uconn.edu](mailto:thomas.meyer@uconn.edu)

Daniel R. Roman

*National Geodetic Survey*

David B. Zilkoski

*National Geodetic Survey*

Follow this and additional works at: [https://opencommons.uconn.edu/nrme\\_articles](https://opencommons.uconn.edu/nrme_articles)



Part of the [Natural Resources Management and Policy Commons](#)

---

## Recommended Citation

Meyer, Thomas H.; Roman, Daniel R.; and Zilkoski, David B., "What Does *Height* Really Mean? Part IV: GPS Orthometric Heighting" (2006). *Department of Natural Resources and the Environment Articles*. 5.  
[https://opencommons.uconn.edu/nrme\\_articles/5](https://opencommons.uconn.edu/nrme_articles/5)

# What Does *Height* Really Mean?

## Part IV: GPS Heighting

**Thomas H. Meyer, Daniel R. Roman, David B. Zilkoski**

**ABSTRACT:** This is the final paper in a four-part series examining the fundamental question, “What does the word *height* really mean?” The creation of this series was motivated by the National Geodetic Survey’s (NGS) embarking on a height modernization program as a result of which NGS will publish measured ellipsoid heights and computed Helmert orthometric heights for vertical bench marks. Practicing surveyors will therefore encounter Helmert orthometric heights computed from Global Positioning System (GPS) ellipsoid heights and geoid heights determined from geoid models as their published vertical control coordinate, rather than adjusted orthometric heights determined by spirit leveling. It is our goal to explain the meanings of these terms in hopes of eliminating confusion and preventing mistakes that may arise over this change. The first paper in the series reviewed reference ellipsoids and mean sea level datums. The second paper reviewed the physics of heights culminating in a simple development of the geoid in order to explain why mean sea level stations are not all at the same orthometric height. The third paper introduced orthometric heights, geopotential numbers, dynamic heights, normal heights, and height systems. This fourth paper is composed of two sections. The first considers the stability of the geoid as a datum. The second is a review of current best practices for heights measured with the Global Positioning System (GPS), essentially taking the form of a commentary on NGS’ guidelines for high-accuracy ellipsoid and orthometric height determination using GPS.

JULIET : And not impute this yielding to light love,  
Which the dark night hath so discovered.

ROMEO: ...Lady, by yonder blessed moon I vow,  
That tips with silver all these fruit-tree tops--

JULIET: O, swear not by the moon, th’ inconstant moon,  
That monthly changes in her circle orb,  
Lest that thy love prove likewise variable.

[William Shakespeare: Romeo and Juliet—The Balcony Scene (Act 2, Scene 2)]

### Vertical Datum Stability

Stability is a desirable quality for a datum, meaning that a datum ought not to change with time—this is a concept well understood by surveyors. The purpose of this series of papers is to explore issues pertaining to determining orthometric heights with GPS technology at the accuracy on the order of centimeters; so if the datums to which the height systems are referred vary by this amount or more, then these effects must be taken into account and removed. Therefore, let us consider the geoid in this light: is the geoid stable or does it change with time and, if so, how quickly and by how much?

An investigation into the variability of the geoid is equivalent to an investigation into the variability of the Earth’s gravity potential field; it is a subject in the field of geodynamics. Changes in Earth’s gravity are caused by changes in (1)

---

**Thomas H. Meyer**, Department of Natural Resources Management and Engineering, University of Connecticut, Storrs, CT 06269-4087. Tel: (860) 486-2840; Fax: (860) 486-5408. E-mail: <thomas.meyer@uconn.edu>. **Daniel R. Roman**, National Geodetic Survey, 1315 East-West Highway, Silver Springs, MD 20910. E-mail: <Dan.Roman@noaa.gov>. **David B. Zilkoski**, National Geodetic Survey, 1315 East-West Highway, Silver Springs, MD 20910. E-mail: <Dave.Zilkoski@noaa.gov>.

---

the Earth’s diurnal rotation that produces the centrifugal force component of gravity; (2) the Earth’s mass and its distribution; or (3) the spatial arrangement of objects massive enough and near enough that their gravitational fields have a discernable effect on the geoid.

### Changes in the Earth’s Rotation

The Earth’s diurnal rotation is not constant in velocity or direction. It is known that the length

of the day is decreasing by about two milliseconds per century and that there are seasonal variations (with periods on the order of a month) on the same order (Vanicek and Krakiwsky 1996, p. 68). Consequently, the Earth's centrifugal force is likewise diminishing and variable. However, these variations are far too small (on the order of  $10^{-12}$  radians  $s^{-1}$ ) to change the Earth's centrifugal force at a discernable level in faster than a geologic time frame.

The rotational axis of the Earth slowly traces a circle on the celestial sphere, the same motion that can be observed in a spinning top. This motion is called **precession**. The Earth's precession is caused by the equatorial bulges not aligning in the plane of the **ecliptic** (the plane in which the Earth orbits the sun), thereby giving rise to a torque from the gravitational attraction of the sun (Vanicek and Krakiwsky 1996, p. 59). The Earth's precession is slow, with its axis returning to a previous orientation once in approximately 25765 years, a period known as a **Platonic year**. Likewise, the equatorial bulges are not aligned with the Moon's orbital plane, which is inclined  $5^{\circ}11'$  to the ecliptic. The intersection of the Moon's orbital plane with the ecliptic is known as the **nodal line**, and the nodal line rotates once in 18.6 years, the **Metonic cycle**. This constant realignment of the Moon with the Earth also affects the orientation of the Earth's rotational axis, causing a motion called **nutation** (Vanicek and Krakiwsky 1996; Volgyesi 2006, p. 61).

There are additional, smaller perturbations as well. The motion of the Earth's rotational axis in the celestial reference frame affects astrometric and satellite observations but not gravity because, although the direction of the centrifugal force vector is changing, this change was brought about by a motion of the Earth itself, so the relative change is zero. However, actual movement of the rotational axis relative to the Earth's crust itself (known as "Polar Motion" or "Polar Wobble") does affect gravity, because the direction of the centrifugal force vector in this case is changing relative to the Earth's crust. These small changes are only on the order of a few nanoGals, well below the noise level of most gravity measurements.

## Changes in the Earth's Mass

The Earth's mass can increase or decrease, and it can be redistributed. Concerning the former, the Earth does gain mass almost continuously due to a stream of space debris entering the atmo-

sphere and, occasionally, striking the Earth's surface. Similarly, the Earth is constantly losing mass as gaseous molecules too light to be bound by gravity drift off into space (e.g., helium gas). Neither the addition nor the removal of mass changes the Earth's gravity field enough to be of concern in this paper.

The Earth's mass is redistributed in various ways including postglacial rebound, melting ice caps and glaciers, the Earth's fluid outer core, the oceans (Cazenave and Nerem 2002), and earthquakes. For example, earthquakes can be caused by the motion of tectonic plates along their margins, and this motion causes a change in the Earth's shape. Earthquakes can cause a measurable change in the Earth's rotation velocity, and thus its gravity, by changing one of its moments of inertia (Chao and Gross 1987; Smylie and Manshina 1971; Soldati and Spada 1999). The Sumatra, Indonesia, earthquake of December 26, 2004 was such an event. It decreased the length of day by 2.68 microseconds, shifted the "mean North pole" about 2.5 cm in the direction of 145 degrees East Longitude, and decreased the Earth's flattening by about one part in 10 billion (Buis 2005). The uplift of plates due to tectonic or postglacial activities affects ellipsoidal heights, as well as having a smaller gravity-based effect which changes the geoid. The National Geodetic Survey is planning to engage in research which tracks the time-dependent changes of the geoid due to these effects.

## Tides

People who have been at an ocean shore for half a day or more have had the opportunity to watch the ocean advance inland and then retreat back out to sea. This motion is caused primarily by the gravitational attraction of the Moon and, to a lesser degree, the Sun. Therefore, the definition of **tide** found in NGS' Geodetic Glossary may be somewhat surprising.

**Tide** (1) Periodic changes in the shape of the Earth, other planets or their moons that relate to the positions of the Sun, Moon, and other members of the solar system.

Note that this definition is not about the oceans, *per se*. Instead, it speaks of, among other things, a change of the shape of the *Earth itself*, the **Earth tide** or **body tide**. It is commonplace knowledge that the Moon moves the oceans; it deforms them to set them in motion. But, what is probably not so well known is that the Earth's core, mantle, and crust have their shape deformed

in a manner similar to the deformation of the oceans. The NGS definition continues:

In particular, (2) those changes in the size and shape of a body that are caused by movement through the gravitational field of another body. The word is most frequently used to refer to changes in size and shape of the Earth in response to the gravitational attractions of the other members of the solar system, in particular, the Moon and sun. In such cases, three different tides are usually distinguished: the atmospheric tide, which acts on the gaseous envelope of the Earth; the earth tide, which acts on the solid Earth; and the ocean tide (usually simply called “the tide”), which acts on the hydrosphere.

The effects of the tides are numerous and complicated, so perhaps the first question to consider is whether the tides cause enough of an effect to be of concern. Is the earth tide large enough to affect the geoid in any practical way? It happens that there are two high and low earth tides each day, with the highest being on the order of a 50 cm displacement from its undeformed shape (Moritz 1980, p. 477)! So, the answer is “yes;” we must take tides into consideration.

Tides on the Earth arise due to the influences from all celestial bodies. The Sun and the Moon produce the largest effects by far, but the other planets have a discernable affect, albeit too small to impact GPS positioning (Wilhelm and Wenzel 1997, p.11). All celestial bodies create tides in the same way, the only difference being the details of how these manifest themselves. Therefore, we will consider the effects created by the Moon, with the understanding that they apply to any celestial body with the appropriate change of mass and distance variables.

## Tidal Gravitational Attraction and Potential

According to Newton, force gives rise to motion by accelerating mass. The gravitational force of the Moon on the Earth itself is found using Equation (II.2):

$$\mathbf{F}_E = -\frac{GMm\hat{\mathbf{r}}}{|\mathbf{r}|^2} \quad (\text{IV.1})$$

where:

$\mathbf{r}$  = a vector from the Moon’s center to the Earth’s center (note the negative sign in Equation (IV.1) reversing the direction of the vector

so that the force is directed from the Earth’s center towards the moon’s center);

$M, m$  = the mass of the Earth and the Moon, respectively; and

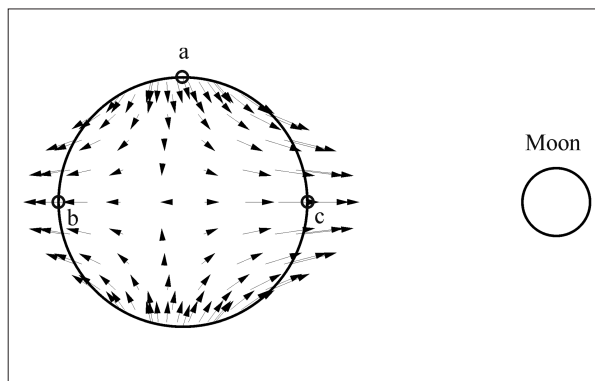
$\mathbf{F}_E$  = the gravitational force vector produced by the Moon on the Earth.

The gravitational force of the Earth exerted on the Moon can be found simply by defining  $\mathbf{r}$  to have the opposite direction, so the magnitudes of the two forces are equal. The gravitational attraction of the Earth on the Moon causes the Moon to orbit the Earth rather than to move off into space. However, Equation (IV.1) also means that *the Earth is orbiting the Moon*, but this motion is much less obvious due to the difference in masses of the two bodies. If we take  $5.9742 \times 10^{27}$  g to be the Earth’s mass,  $7.38 \times 10^{25}$  g to be the Moon’s mass, and  $3.84 \times 10^8$  m to be their mean separation, then the barycenter of the Earth–Moon system can be found to be at a point on a line connecting their two centers approximately  $4.69 \times 10^6$  m from the Earth’s center. This point is inside the Earth, being about 73.5 percent of the length of the GRS 80 semimajor axis.

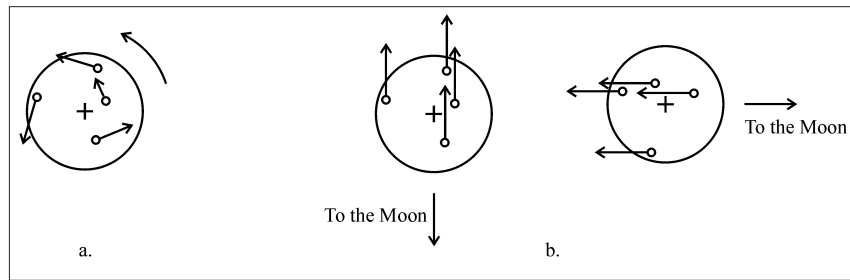
It is critical to understand the nature of the motion of the Earth’s orbiting the Moon. The diurnal rotation of the Earth, the source of days and nights, is a rotation around its axis, which is nominally the North Pole. Points on a rigid rotating body that are on different radii move in different directions and at different instantaneous linear velocities (see Figure IV.1a). However, a rigid body can rotate around only one axis at any moment in time. Therefore, the Earth does *not* rotate about the Earth–Moon barycenter. To understand this orbital motion, envision someone waxing a tabletop with a cloth by rubbing it in a circular motion, such that their fingers remain parallel to some wall in the room. If the circular motion of the cloth has a fairly small radius, then the point around which the cloth is moving is always beneath the cloth, just as the motion of the Earth around the barycenter has its center at a point within the Earth. Now, it is apparent that every point on the cloth is actually moving with the same velocity (same direction and speed). Similarly, the orbital motion of the Earth around the Moon gives rise to a *constant* acceleration that is always directed opposite to the line connecting the Earth’s center to the Moon’s. In particular, *everywhere and everything on and in the Earth is accelerating away from the Moon as if the Earth were moving in a straight line along the instantaneous axis between them*; see Figure IV.1b. This acceleration gives rise to a component of

observable gravity that is at most 3.4 percent of the total acceleration (Vanicek and Krakiwsky 1996, p. 125).

The moon's gravitational attraction gives rise to a force at any particular place on the Earth that is directed (approximately<sup>1</sup>) along the line from the point of interest to the Moon's center. In contrast, the orbital acceleration experienced at that place is always parallel to the line connecting the Earth–Moon centers, so these forces are not generally parallel to each other. Furthermore, places on the side of the Earth opposite the Moon experience a smaller attraction than places on the same side as the Moon due to being closer to the Moon, giving rise to the asymmetry evident in Figure IV.2. Each of the vectors in Figure IV.2 indicates the force vector of the place located at the tail of the vector resulting from the combination of the orbital acceleration and the Moon's attraction at that place.



**Figure IV.2.** Arrows indicate force vectors that are the combination of the moon's attraction and the Earth's orbital acceleration around the Earth–Moon barycenter. This force is identically zero at the Earth center of gravity. The two forces generally act in opposite directions. Points closer to the Moon experience more of the Moon's attraction whereas points furthest from the Moon primarily experience less of the moon's attraction; c.f. (Bearman 1999, pp. 54-56; Vanicek and Krakiwsky 1996, p. 124).

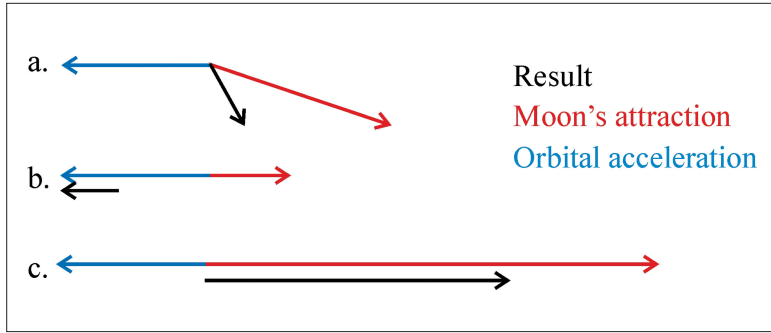


**Figure IV.1.** Velocities of points moving on a rigid rotating body. Panel (a) presents the instantaneous velocity vectors of four places on the Earth; the acceleration vectors (not shown) would be perpendicular to the velocity vectors directed radially towards the rotation axis. The magnitude and direction of these velocities are functions of the distances and directions to the rotation axis, shown as a plus sign. Panel (b) presents the acceleration vectors of the same places at two different times of the month, showing how the acceleration magnitude is constant and its direction is always away from the Moon.

Figure IV.3 shows the details of the vector addition of three points of interest from Figure IV.2. Orange vectors are the Moon's attraction; their non-parallelism with the orbital acceleration vectors, shown in blue, is greatly exaggerated. The vector result of the addition of these two vectors is shown in black. Figure IV.3a represents the situation at point *a*, which is located at the top of the circle in Figure IV.2. The Moon's attraction is the most non-parallel with the orbital acceleration at this place and its antipodal counterpart. Given the roughly equal magnitude of the orbital acceleration and Moon attraction forces, their component in the direction of the Moon largely cancels at *a*, leaving a small result oriented sharply towards the Earth's middle. Figure IV.3b represents the situation at *b* which is located at the point furthest from the Moon. The Moon's attraction is parallel but opposite in direction with the orbital acceleration at this place. The orbital acceleration is moderately stronger than the Moon's attraction here, creating the force primarily responsible for the lower high tide of the day. Figure IV.3c represents the situation at *c* which is located at the point closest to the Moon. The Moon's attraction is considerably stronger here than the orbital acceleration, creating the force that is primarily responsible for the higher tide of the day (see (Vanicek and Krakiwsky 1996, p. 124; Bearman 1999, pp. 52-61).

The magnitude and direction of the Moon's attraction is periodic due to the nature of its orbit around the Earth. The situation is com-

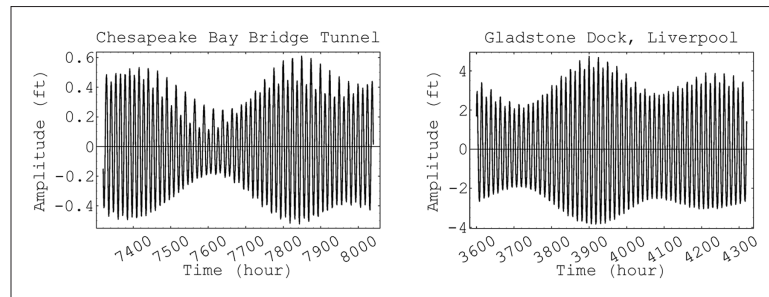
<sup>1</sup> The moon is too close to the Earth for this to be exact. The actual direction of the vector would be determined by triple integrating over the Moon's mass, and approximately end up pointing at the Moon's center of mass, approximate because the Moon is not a perfectly homogeneous sphere.



**Figure IV.3.** Details of force combinations at three places of interest; c.f. Figure IV.2.

plicated but made tractable by accounting for individual **tidal constituents**. It is possible to decompose the Moon's attraction into individual **constituents**, a constituent being a sinusoid with a particular amplitude, frequency and phase that arises due to a particular phenomenon. As discussed by Boon (2004), some of the prominent tidal constituents are caused by:

- The inclination of the Moon's orbital plane with respect to the ecliptic giving rise to the lunar declination (topic–equatorial) cycle;
- The Sun's attraction giving rise to the spring–neap cycle;
- The eccentricity of the Moon's orbit giving rise to the perigean–apogean cycle; and
- The precession of the lunar nodes giving rise to the metonic cycle.



**Figure IV.4.** Two simulations of tide cycles illustrating the variety of possible affects.

Simple ocean tide models include as few as six constituents; complicated models can incorporate more than 100 (Wilhelm and Wenzel 1997). These models produce tidal predictions such as those shown in Figure IV.4. The predictions in Figure IV.4 use constituents from Boon (2004, pp. 97-102) and clearly show higher high water, lower high water, higher lower water, and lower low water, as well as many longer-period variations.

Up to this point we have been concerned with gravity *force*. We now consider how tides affect gravity *potential* because, after all, the geoid (an equipotential surface) is a principle datum of interest, hence we must examine how these temporal changes come into play. The gravitational potential field created by the Moon at some point of interest can be expressed as an infinite series of which only the second term is important for tides. This

second term  $W_2$  takes the form of the expression (Vanicek 1980, p. 5, Equation (12)):

$$W_2 \approx D \left[ \overbrace{\cos^2 \phi \cos^2 \delta \cos 2t}^{\text{sectorial}} + \overbrace{\sin 2\phi \sin 2\delta \cos t}^{\text{tesseral}} + \overbrace{3(\sin^2 \phi - 1/3)(\sin^2 \delta - 1/3)}^{\text{zonal}} \right] \quad (\text{IV.2})$$

where:

$D$  = Doodson's constant (Doodson 1922);

$\phi$  = geocentric latitude;

$\delta$  = the declination of the Moon; and

$t$  = the Moon's hour angle (see any standard work on celestial mechanics for exact definitions of  $\delta$  and  $t$ ).

Doodson's constant is given by Vanicek (1980, p. 4, Equation (7)) as:

$$D = \frac{3}{4} Gm \frac{R^2}{r_m^3} \quad (\text{IV.3})$$

where:

$G$  = the universal gravitation constant;

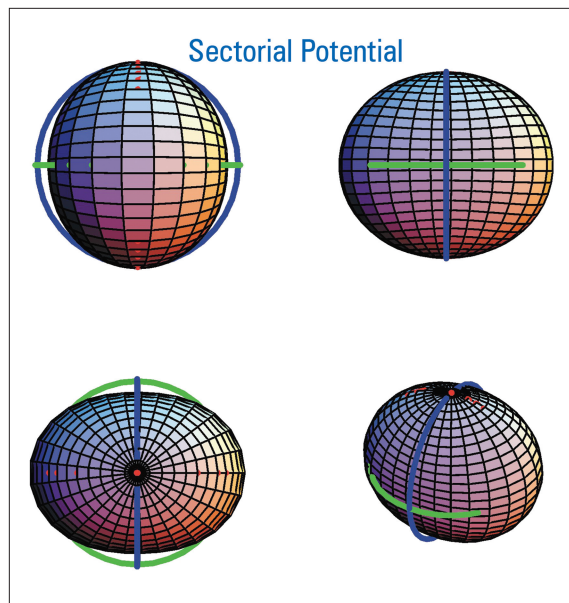
$R$  = the mean (equivoluminous) radius of the Earth; and

$r_m$  = the mean distance to the Moon.

$D$  has a value of approximately  $2.6277 \times 10^7$  cm mgal. Equation (IV.2) consists of three terms within the brackets. The first term contains **sectorial** constituents; the

second term contains **tesseral** constituents, and the third term contains **zonal** constituents. These three components are shown in Figures IV.5-7 and their combination in Figure IV.8. Sectorial constituents vary in longitude (time), much like the sectors of an orange, and give rise to the two daily tides. Tesseral constituents possess both latitude and longitude components and give rise to patterns resembling the tessellation of a checker





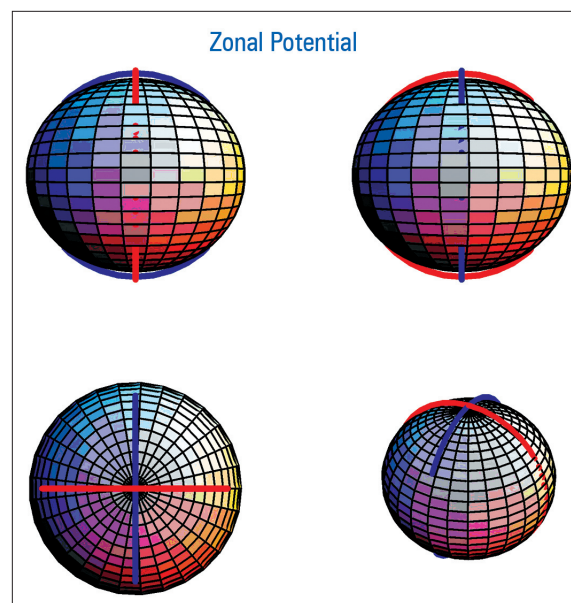
**Figure IV.5.** The sectorial constituent of tidal potential. The green line indicates the Equator. The red and blue lines indicate the Prime Meridian/International Date line and the 90/270 degree meridians at some arbitrary moment in time. In particular, these circles give the viewer a sense of where the potential is outside or inside the geoid. The oceans will try to conform to the shape of this potential field and, thus, the sectorial constituent gives rise to the two high/low tides each day.

board. The zonal constituents do *not* vary in time and give rise to so-called “permanent” tides.

### Body Tides

The first clear evidence of body tides came from the measurement of ocean tides, which showed that they were consistently about two-thirds as high as Newton’s physics predicted. It was eventually shown that the missing one-third was due to deformation of the Earth itself, moving with the oceans (Melchior 1974). The tides of the solid Earth behave in the same manner as the ocean tides, but in a simpler manner because the Earth deforms like an elastic solid at the frequencies of tides, rather than with all the freedom of a liquid, like the oceans.

It is remarkable that the effect of the Moon’s potential field upon the Earth can be described with such high accuracy by such a simple equation as Equation (IV.2); compare this with the effort necessary to determine the geoid! The simplicity of Equation (IV.2) is because (1) the Moon is far enough away to be treated as a point mass, and (2) the motion of the Moon is very accurately described by celestial mechanics.

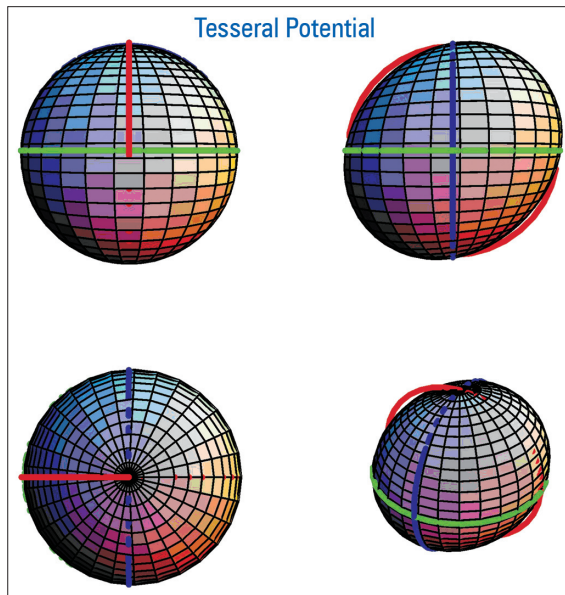


**Figure IV.6.** The zonal constituent of tidal potential. The red and blue lines are as in Figure IV.5; the equatorial green line is entirely inside the potential surface. The zonal constituent to tidal potential gives rise to latitudinal tides because it is a function of latitude.

Therefore, no gravity observations are needed to determine the potential from the Moon; it all falls out of the mathematics.

The parameters that describe the response of the Earth’s shape and gravitational potential field to tidal forces are called **Love** and **Shida** numbers, which are empirically derived. They are used in equations similar to Equation (IV.2) and sufficiently capture the deformation of the Earth so that tidal affects may be removed from geoid models, gravity observations, GPS observations, and other geodetic quantities (Vanicek 1980). However, it should be noted that the permanent tides (those portions of the tidal equations which describe the non-time-varying, or “permanent” deformations) are not *completely* determinable empirically. There are two components of this permanent tide: first, the permanent deformation of Earth’s geopotential field due to the existence of the permanent (non-zero time-averaged) Sun and Moon and second, the permanent deformation of Earth’s geopotential field due to the existence of the permanent deformation of Earth’s crust (which, in turn, is due to the existence of the permanent Sun and Moon).

The first part (called the “direct” component of the permanent Earth tide) is computable empirically, as it deals solely with the Sun’s and Moon’s mass affecting the Earth’s geopotential

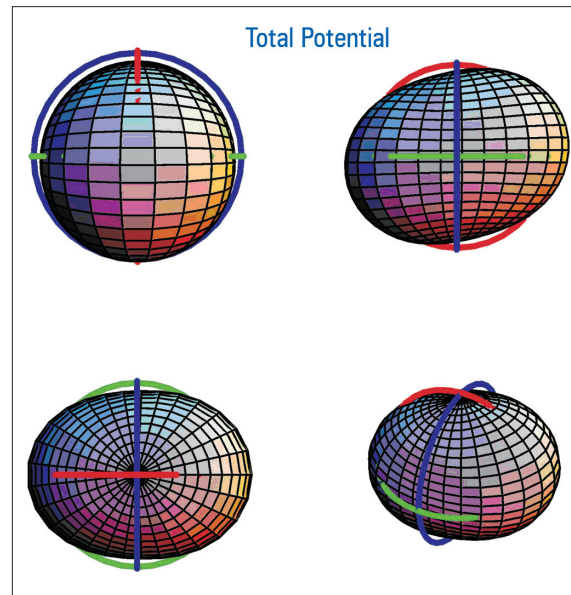


**Figure IV.7.** The tesseral constituent of tidal potential. The green, red and blue lines are as in Figure IV.5. The tesseral constituent to tidal potential gives rise to both longitudinal and latitudinal tides, producing a somewhat distorted looking result, which is highly exaggerated in the figure for clarity. The tesseral constituent accounts for the Moon's orbital plane being inclined by about five degrees from the plane of the ecliptic.

field. The second part is not computable empirically. This is because the permanent deformation of the Earth's crust can not be directly observed. The Earth's crust perpetually ("permanently") exhibits a deformation due to the permanent existence of the Sun and Moon. Because we can not observe how the crust would react without a permanent Sun and Moon, we can not determine empirically how much permanent deformation actually exists (that is, we can not determine a "zero degree Love number" for the Earth), and thus can not compute what the effect of this permanent crustal deformation is on the Earth's geopotential.

### Ocean Tides

Ocean tides affect the geoid by redistributing the mass of the oceans, which has the following effects. First, the redistribution of the water in the oceans creates a discernible change in the geoid. Second, the weight of the water deforms the Earth below it, in addition to the tidal potential also deforming the Earth (Vanicek 1980, pp. 9-12). The deformation of the Earth due to tidal loading can also be modeled by certain Love numbers that parameterize Equation (IV.2). The



**Figure IV.8.** The total tidal potential is the combination of the sectorial, zonal, and tesseral constituents. The green, red and blue lines are as in Figure IV.5. The complicated result provides some insight into why tides have such a wide variety of behaviors.

liquid nature of the oceans allows dramatically more complexity in their response to gravitational attraction and, consequently, its modeling is likewise more complex.

## Global Navigation Satellite System (GNSS) Heighting

Global navigation satellite systems, such as the European Union's Galileo system, the Russian *Global'naya Navigatsionnaya Sputnikovaya Sistema* (GLONASS), and the U.S. Global Positioning System (GPS) offer, in conjunction with a highly accurate model of the gravimetric geoid, the potential of determining orthometric heights with centimeter accuracy without conventional leveling. The prospect of establishing vertical control in remote locations without running levels to established distant bench marks holds great promises of time savings, and therefore, cost savings. These savings are the reward for surveyors who practice GPS heighting and were a primary motivation for this series. According to Zilkoski et al. (2000), "GPS-derived orthometric heights can now provide a viable alternative to classical geodetic leveling techniques for many applications."

Deriving orthometric heights from ellipsoid heights is mathematically very simple. As explained in the previous papers, a geoid height is the geo-



metrical separation (distance) from some reference ellipsoid to the geoid, an ellipsoid height is the geometrical separation from some reference ellipsoid to a point of interest, and an orthometric height is the length of the plumb line from the geoid to a point of interest. Were plumb lines straight lines and if they were normal to the reference ellipsoid, these three definitions would immediately lead to an exact relationship:

$$H = h - N \quad (\text{IV.4})$$

where:

$H$  = orthometric height;

$N$  = geoid height; and

$h$  = ellipsoid height.

However, plumb lines are curved and not normal to reference ellipsoids, in general. Therefore, we cannot be correct in using an equality relationship and must instead write:

$$H \approx h - N \quad (\text{IV.5})$$

Although Equation (IV.5) is not exact, it is close enough for most practical purposes (Hein 1985; Henning et al. 1998; Vanicek et al. 1999; Zilkoski 1990; Zilkoski and Hothem 1989). For example, an extreme case of a two-arc-minute deflection of the vertical would introduce less than two millimeters of error in the orthometric height (Tenzer et al. 2005, p.89), based on Equation (IV.5).

Much of the information from this series is contained within Equation (IV.5) (Hwang and Hsiao 2003; Kao et al. 2000; Sun 2002). For example, the choice of the reference ellipsoid is important. Local geodetic reference ellipsoids are generally not geocentric, so their normal directions could differ significantly from those of ellipsoids that are geocentric insofar as was possible at the time of their creation. It is important not to mix heighting systems. The GPS surveyor must therefore use a reference ellipsoid of a datum that matches the reference ellipsoid of the gravimetric geoid model. In the U.S., NGS recommends using GEOID03 which is modeled relative to the NAD 83 datum (which uses the GRS 80 ellipsoid). Therefore, for example, GPS heighting should not be done with GEOID03 and the WGS 84 datum. Also, because Equation (IV.5) is an approximation rather than an equality (due to the non-parallelism of the equipotential surfaces), dynamic/orthometric corrections will have to be applied to (the purely geometric) spirit leveling measurements (Strang van Hees 1992).

In theory, GPS heighting is simple: determine an ellipsoid height with a GPS receiver and sub-

tract the geoid height, which is provided by a gravimetric geoid model, to obtain the approximate orthometric height. In practice, things are more complicated. This fourth paper now presents a survey of GPS heighting error sources and best practice guidelines put forth by NGS and other authors in the peer-reviewed literature. Although this paper depends in large part on previous work by Zilkoski and others at the NGS (Zilkoski et al. 1997), it is not our intention to restate that material verbatim (Zilkoski et al. 2000). Instead, this final paper will provide commentary on the guidelines and explanations why some of the recommendations were made. We will emphasize the key issues necessary for achieving the accuracies in those guidelines and provide examples from the literature that illustrate them, when possible. More detailed and comprehensive treatments include (Leick 1995; Hofmann-Wellenhof et al. 1997; Seeber 2003; Hofmann-Wellenhof and Moritz 2005).

## Error Sources

Effective GPS heighting depends upon having an understanding of the measurement error budget and acting in such a manner as to eliminate or mitigate those errors. Error sources have been grouped in three main categories: satellite position and clock errors, signal propagation errors, and receiver errors (Seeber 2003). We will discuss these error sources and explain what, if anything, can or should be done about them according to best practices reported in the current literature. Although it is beyond the scope of this paper to review GNSS as a whole, the reader is referred to the large existing literature on the topic, such as (Hofmann-Wellenhof et al. 1997; Leick 1995; Seeber 2003; van Sickle 1996) and collections of articles published by the U.S. Institute of Navigation (ION). However, before discussing these error sources, we present issues that arise due to the Earth itself.

## Geophysics

There are several issues pertaining to the Earth itself that factor into GNSS heighting. Most of these pertain to the dynamic shape of the Earth but one arises simply because the Earth is opaque at the radio frequencies broadcast by GNSS satellites.

### *No Satellites Below*

We begin by explaining why it is that GNSS positioning cannot be expected to be as accurate for vertical coordinates as for horizontal ones.

Currently operational GNSS satellites, abbreviated as SV for “space vehicle,” are stationed in orbital planes inclined from the equator by 55 degrees (for GPS) or 64.8 degrees (for GLONASS). Consequently, any place on Earth is always surrounded by SVs, above and below. However, the Earth completely blocks signals from SVs below the horizon from reaching a receiver; the radio signals cannot penetrate solid rock. Therefore, receivers on the ground cannot detect signals from SVs below the horizon. As a result, while it is possible to be surrounded on all azimuth points by SVs, one cannot be surrounded on all zenith angles (essentially none greater than 90 degrees). Consequently, the local vertical is not controlled as well as the local horizontal. As stated by Brunner and Walsh (1993), “We note that even without any tropospheric propagation errors, an inherent geometrical weakness exists in the GPS baseline results that usually makes the determination of height differences worse by a factor of about 3 compared with the horizontal baseline components.” Therefore, we cannot expect the best GNSS heighting to be as accurate as the best GNSS horizontal positioning.

#### *Earth Tides, Ocean and Atmospheric Loading*

GNSS post-processing software often includes tide corrections which remove these effects, creating a “tide-free” system. See the opening discussion for more elaboration.

#### *Crustal Motion*

Plate tectonics constantly move the Earth’s crust both horizontally and vertically. Horizontal motions can be accounted for by modeling the position and velocity of fiducial stations and then interpolating to places of interest. The NGS Horizontal Time-Dependent Positioning (HTDP) software (Snay 2003; Snay 1999) allows U.S. users to reconcile control coordinates published in the past with current position measurements that have moved due to plate motion, including earthquakes. Of particular note to heighting, a vertical equivalent, VTDP, has been created for the lower Mississippi valley and the northern Gulf Coast (Shinkle and Dokka 2004). Vertical crustal motion includes both tectonic crustal motion and anthropogenic factors, such as liquid extraction resulting in ground subsid-

ence (Gabrysch and Coplin 1990), which complicates matters considerably.

### Satellite Position and Clock Errors

We now begin a discussion of the GNSS error budget. Because GNSS positioning is accomplished by a process similar to trilateration there are two key pieces of information upon which GNSS positioning depends: signal propagation time and the location of the SVs. Signal propagation time is used to infer the range from the SVs to a receiver antenna’s phase center, and SV locations are used as the coordinates of the known points in the trilateration scheme. However, the signal propagation time is biased due to an immeasurable time offset between GPS time and a receiver’s internal clock; this results in a “pseudo-range” rather than the actual range. The implications of this will be discussed below. Any errors in locating a SV and any inconsistencies in the clocks on board the SVs that govern its operation result directly in positioning errors.

#### *Orbit Errors and Ephemerides*

Knowing the position of the satellites at any given moment in time is a cornerstone of how GNSS positioning works. The satellites themselves should be perceived as being moving monuments because pseudo-range positioning (positioning using pseudo-ranges) is based on trilateration: given three (or more) known locations and a distance from those locations to the point of interest, determine the coordinates of the point of interest.<sup>2</sup> Therefore, since the satellites are in motion, it does not suffice to publish a single set of coordinates for them. Instead, ephemerides are created for each SV so that the processing software can determine SV positions at the moment of transmission, which form the basis for the trilateration.

In broad strokes, GNSS ephemerides come in two types: broadcast and precise. Broadcast ephemerides, as the name implies, are broadcast by the SVs and read by GNSS receivers as they operate. Broadcast ephemerides are essentially highly educated, physics-based guesses about the future locations of the SVs based on their past locations and velocities. Precise eph-

<sup>2</sup> In fact, three known locations and distances do not uniquely determine a three-dimensional position; the problem is reduced to a selection between two solutions. One of these solutions will either be deep inside the Earth or in outer space and can be discarded by inspection for terrestrial GNSS positioning. See Awange and Grafarend (2005) for novel solutions of this problem based on Groebner bases.

emerides are produced by observing the SVs and deducing their positions after-the-fact. Needless to say, broadcast ephemerides are not as accurate as precise ephemerides. The accuracy of the broadcast ephemerides is currently around one to three meters (Seeber 2003).

The International GNSS Service (IGS) provides three types of precise ephemerides, which differ by how much time elapses before they are available (IGS 2005). The most accurate are the “final ephemerides” which are updated weekly with a latency of about 13 days and have accuracy reported to be better than 5 cm (Seeber 2003). The “rapid ephemerides” are updated daily with a 17-hour latency and an accuracy around 5 cm. Ultra-rapid ephemerides are updated four times daily with a latency of either 3 hours (observed half) or none (predicted half) with an accuracy around 25 cm. It can be shown that the error introduced into computed positions varies by baseline length as a function of ephemeris accuracy: the longer the baseline, the more accurate the ephemeris needs to be (Eckl et al. 2002, Seeber 2003, p. 305). For high-accuracy GPS heighting, final precise ephemerides are required by NGS guidelines (Zilkoski et al. 1997).

#### *Satellite Clock Errors*

Although GNSS satellites have onboard atomic time standards that are highly accurate and precise, they are not perfect. Like all clocks, atomic clocks drift and experience unpredictable jumps, albeit very small ones (Diddams et al. 2004; Flowers 2004). GPS time is a weighted average of the clocks in the controlling station on Earth and the GPS satellite clocks. Each SV clock is monitored for its offset from GPS time, and this time bias estimate is included with the ephemerides, both broadcast and precise, to be accounted for in the positioning software.

### GPS Signal Propagation Delay Errors

GNSS ranges are inferred by measuring a (biased) elapsed time from the satellite to the receiver; it is biased due to an immeasurable time offset between GPS time and a receiver’s internal clock. This elapsed time interval is scaled to be a distance by multiplying by the speed of light. Although the speed of light is constant in a vacuum, electromagnetic waves propagating through media can be delayed and refracted. GNSS signals propagate through the Earth’s atmosphere and are affected by the iono-

sphere and the troposphere. Both of these atmospheric layers delay the signals, thus introducing timing/ranging errors.

#### *Ionosphere Delays*

The ionosphere is a high-altitude (roughly 50 km to 1000 km above the Earth’s surface) part of the atmosphere that is composed of charged particles that have been ionized by solar radiation. The ionosphere refracts radio signals in a manner similar to how water in a glass refracts light, such that a pencil appears to have a sharp bend in it. It happens that the ionosphere refracts radio-frequency electromagnetic waves of different frequencies differently. Consequently, it delays the two GPS broadcast frequencies, L1 and L2, differently. This difference can be detected by dual-frequency receivers and subsequently virtually eliminated by post-processing. For more details consult, for example, Brunner and Walsh (1993), Hofmann-Wellenhof et al. (1997), Leick (1995), and Seeber (2003). Single-frequency receivers cannot detect the ionosphere delays, but differencing processing on short baselines can cancel out most of the error, leaving errors on the order of 1 to 2 ppm of the interstation distance (Seeber 2003). The NGS guidelines require dual-frequency receivers for baselines greater than 10 km, and they are the preferred type of GPS receiver for all observations (Zilkoski et al. 1997).

According to Jakowski et al. (2005, p. 3071), “The space weather is defined as the set of all conditions —on the Sun, and in the solar wind, magnetosphere, ionosphere and the thermosphere—that can influence the performance and reliability of space-borne and ground-based technological systems and can endanger human life.” Space weather can significantly influence the propagation of the SV transmissions through the ionosphere, resulting in a degradation of positioning quality (*ibid*). Dual-frequency receivers are not able to eliminate the problems caused by severe space weather, hence observations should not be performed during severe ionospheric storms. The National Oceanic and Atmospheric Administration (NOAA) includes space weather reporting from its Space Environment Center, which is part of the National Weather Service (<http://www.sec.noaa.gov/>).

#### *Troposphere Delays*

The troposphere is that part of the atmosphere in which weather (in the ordinary sense) occurs. Atmospheric density gradients of the tropo-

sphere, like the ionosphere, refract GNSS radio waves. However, the tropospheric delays do not depend upon the frequency of the electromagnetic waves. No hardware exists today that can directly measure the delay created by the troposphere, so its affect must be accounted for by modeling the troposphere or by treating it as an unknown nuisance variable determined using least squares techniques.

The errors associated with the troposphere are considered the most problematic member of the GNSS heighting error budget. According to Seeber (2003), “[tropospheric delay]... is one of the reasons why the height component is much worse than the horizontal components in precise GPS positioning.” According to Brunner and Walsh (1993), “Tropospheric delay errors mainly affect the accuracy of height differences. Today this must be considered the main limitation of the attainable accuracy using GPS, which seems to be around 2.5 centimeters for height differences of baselines longer than about 50 kilometers.”

Marshall et al. (2001) performed a detailed study of the affect of tropospheric modeling successfulness at addressing the tropospheric delay on baselines from 62 km to 304 km in length. Based on their experiments conducted using the NGS Continuously Operating Reference Stations (CORS), they show significant reductions in height standard deviations by increasing session duration from one to four hours, and that the choice of the tropospheric model has a strong influence on the precision and accuracy of the resulting heights. Some of these models depend upon measured tropospheric parameters such as atmospheric pressure, atmospheric temperature, and relative humidity, i.e., the quantities that determine the static density of the atmosphere and its density gradient. Others rely on standard models of the atmosphere and are parameterized by latitude and day of the year. Another approach is to treat the tropospheric delay as another unknown parameter and estimate it using statistics from the GPS observables. Marshall et al. (2001) concluded that, “Session lengths shorter than two hours contain insufficient GPS data to estimate both heights and nuisance parameters, and hence more accurate weather information is needed to obtain more precise heights for these shorter sessions.” The models showed a large amount of variability among each other and all of them displayed significant individual variability—more than 5 cm. This fact would appear to contradict NGS claims

that following their guidelines should result in 2 cm - 5 cm ellipsoid height accuracy. The difference is the length of the baselines. Marshall’s study had baselines not shorter than 60 km, but NGS requires lines no longer than 10 km. This is an important difference because the unmodeled tropospheric delay error is spatially auto correlated, meaning that the closer two stations are, the more likely they are to “see” the same tropospheric delay. If the delays were exactly the same, they would be canceled by post-processing differencing. To what degree they are not the same, they do not cancel.

According to the current literature, measuring weather parameters is not very helpful. Marshall et al. (2001) found that, “For session lengths greater than two hours, we conclude that sufficiently precise NAD [neutral atmospheric delay] modeling for geodetic activities may be achieved by coupling nuisance parameter estimation with the relatively crude seasonal model.” This means that weather measurements were not needed to achieve sufficiently precise error models. Brunner and Walsh (1993) note that:

“In general, the tropospheric delay models using meteorological ground observations have produced rather poor, and in most cases worse, results compared with the results from the default model values that replaced the actual observed meteorological values. We would like to comment on this surprising finding. Taking accurate meteorological observations is a somewhat difficult task, and frequently large observation errors can occur. In addition, the closeness of the ground and very local micrometeorological conditions severely affect meteorological observations.”

These comments appear to support the conclusions found by Marshall et al. (2001). Recently, Ray et al. (2005) noted succinctly: “To the central question, whether measured surface met data can be used to improve geodetic performance, we find no such utility.” Nevertheless, NGS guidelines require meteorological data to be collected (Zilkoski et al. 1997). It has been shown (Marshall et al. 2001) that “Weather fronts may cause the GPS signal delay to vary by greater than 3 centimeters over a 1-hour period, potentially leading to ellipsoidal height errors exceeding 9 cm.” Surface met data are not collected for modeling purposes. Rather, they are useful for *a posteriori* error detection as they help to spot the passing of a weather front through the surveying network, something that could possibly go unnoticed by the ground crews.



The affect of the troposphere increases with zenith angle. For this reason (amongst others) NGS recommends a 15 degree minimum elevation mask (Zilkoski et al. 1997).

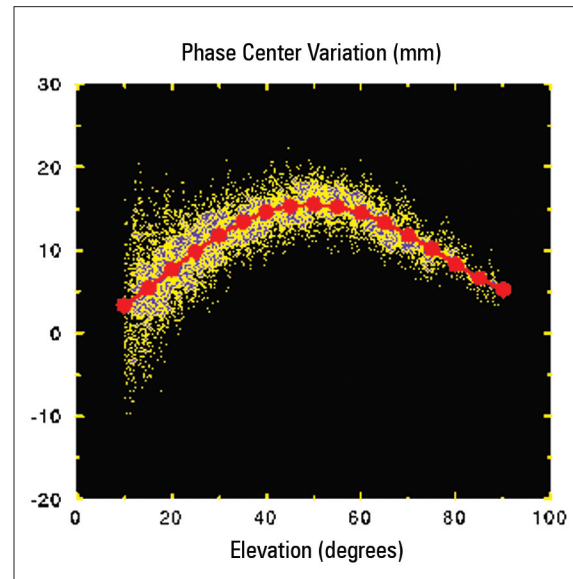
#### *Multipath*

One of the two GNSS observables is **carrier phase**: “carrier” refers to the unmodulated radio signal broadcast by the SVs and “phase” refers to the total number of cycles of the carrier waves from its transmission to its reception, including a partial wavelength at the end. In relative positioning, baselines between phase centers are deduced by differencing phase observations from multiple SVs; see Hofmann-Wellenhof et al. (1997) among many others for more details. Multipath is the situation where GNSS radio signals arrive at the receiver via more than one path. This happens by the signal reflecting from some surface such as a chain link fence, a building, a car, or the ground. According to Seeber (2003), “Multipath influences on carrier phase observations produce a phase shift that introduces a significant periodic bias of several centimeters into the range observation... Their propagation into height errors may reach +/- 15 cm (Georgiadou and Kleusberg 1988)”. Multipath also affects pseudo-range derived positions, introducing errors potentially on the order of meters.

Multipath can be reduced by antenna design, principally choke rings and ground planes, and by elevation masks. Multipath is more likely to occur at low elevation angles so, again, NGS recommends a 15 degree minimum elevation mask (Zilkoski et al. 1997). Ground planes are known to reduce multipath, especially spurious signals arriving at the receiver from below, perhaps being reflected off the ground. Likewise, choke ring antennas mitigate multipath by attenuating reflected signals. Therefore, NGS requires ground planes for GPS antennas and recommends choke rings (Zilkoski et al. 1997). There are also software techniques for multipath reduction (e.g., Seeber et al. 1997) that are available in some processing packages and, sometimes, in the receiver itself (Townsend and Fenton 1994).

### Receiver Errors and Interference

No instrument is perfect, and GNSS receivers are no exception. The receivers themselves cannot determine positions exactly, but we know the error sources associated with the receiver hardware. Also, since the presence of electromagnetic noise in the environment has the potential to



**Figure IV.9.** This image depicts the location of a GPS receiver's phase center as a function of the elevation angle of the incoming GPS radio signal.

interfere with the GNSS radio signals, electromagnetic noise requires some attention, too.

#### *Antenna Phase Center Variation*

The electrical phase center of a GNSS receiver antenna is a point in space where the antenna detects the radio signal broadcast from the satellites; it is the point whose coordinates are being determined. That is to say, unless the position is reduced to the antenna reference point (ARP) or a surveying marker, the latitude, longitude, and ellipsoid height reported by the GNSS post-processing package are those of the phase center. Interestingly, the phase center is not on the physical surface of the antenna; indeed, it is not on or in the hardware at all. It is above the antenna and, furthermore, it is not a single location (see Figure IV.9). Although most modern antennas are azimuthally symmetric electrically, local environmental conditions can produce dependences on azimuth. Therefore, phase centers can change with the zenith and azimuth angle of the incoming signal. Additionally, the phase center for L1 is typically different than that for L2 (Mader 1999). Because the phase center is the position being determined by the receiver, as the satellites move, the phase centers move, which is an effect called phase center variation (PCV). As a phase center moves, its coordinates change. If left uncorrected, phase center variations can introduce as much as a decimeter of error into the vertical coordinate. The NGS antenna calibration program has produced models of phase



center variation that are available for downloading at <http://www.ngs.noaa.gov/ANTCAL/> (NGS 2005). These models can be entered into the post-processing software, which will adjust for the effect.

National Geodetic Survey publishes several coordinates for its CORS base stations. Coordinates are currently given in the ITRF00 (epoch 1997.0) and NAD 83 (CORS96) datums for both the ARP and the L1 phase center. Coordinates for the ARP and the phase center are different by several centimeters, typically. For example, the NAD 83 ellipsoid height for the DE6429 NRME COOP CORS L1 phase center is 163.027 m, whereas the ellipsoid height of the ARP is 162.951 m, a difference of 7.6 cm. Surveyors clearly need to be very careful in choosing their control coordinates and know what their post-processing software does with those coordinates.

Some packages may assume that the vertical coordinate refers to some particular place, typically the ARP or the phase center; others allow the user to specify to which place the control coordinates refer. Surveyors should take care to pick coordinates that match the expectations of their software or they will introduce systematic vertical errors by accounting for the phase center-ARP separation incorrectly. Furthermore, some packages have antenna geometry databases to allow the software to compute the distance from the ARP to the phase center. Surveyors should check the values in such a database to verify they are correct by comparing with designs provided by manufacturers or by information on the aforementioned NGS antenna calibration website.

Also, GNSS observation files often allow for marker offsets. Some CORS base station RINEX observation files have offsets that reflect the phase center-ARP separation, typically a negative number a few centimeters in magnitude. Surveyors will need to zero these offsets if their processing software assumes the control coordinates refer to the ARP and computes the offsets automatically via the antenna geometry database. If they are not zeroed, the software will account for the distance from the phase center to the ARP twice, introducing a several-centimeter blunder into the vertical control coordinate. Such a blunder can be extremely difficult to find if the processing package does not give a complete account (report) of how the vertical coordinate was determined. The NGS processing software, PAGES, does report all the offsets that go into determining the spatial location of the phase center, so the surveyor knows whether all the control coordinates and offsets are consistent.

Additionally, as CORS stations are increasingly being used in local surveys, it is likely that a mixture of antenna types will occur in a single survey. Any azimuthal PCV inconsistencies among the antennas will not cancel in the differencing processing unless the same inconsistency occurs for all antennas. Therefore, it is important to orient all antennas in the survey to the North so that any residual azimuthal effects are canceled. CORS antennas are already oriented to the North, which means that surveyors need only be concerned about their own antennas.

#### *Electromagnetic interference and signal attenuation*

The radio signals currently broadcast by the GPS satellites are relatively low power, around 50 watts. Although GNSS signals occupy a protected frequency band, nearby sources of broadband electro-

Error	Remedy
Orbit errors and clock errors	Use final precise ephemerides; Double differencing of phase observations eliminates orbit and clock errors
Ionospheric delay	Use dual-frequency receivers; Can be reduced on short baselines by differencing phase observables
Tropospheric delay	Modeled or determined in post processing; Longer observation times yield better results; Can be reduced on short baselines by differencing phase observables
Multipath	Avoid multipath-prone locations; Use a ground plane or choke ring antenna
Phase center variation	Use antenna calibration models; Orient antennas to North; Check antenna offsets and antenna geometry databases to ensure consistency with control coordinates
Electromagnetic noise	L5 receivers; Avoid problem sites if possible

**Table IV.1** Summary of error sources and recommended remedies.

magnetic noise can overwhelm them (Johannessen 1997; Butsch 2002), thus causing decreased signal to noise ratios, increased difficulty or prevention of GNSS signal acquisition, and loss of signal tracking (Seeber 2003, p. 320). Power transmission lines, television and radio stations, and radar installations are possible examples of such noise sources. To help address this problem, the GPS modernization program includes a third, higher-power frequency (L5) which is expected to reduce this problem (Hatch et al. 2000). Unfortunately, new receivers will probably have to be purchased when enough satellites have been placed in orbit to make using L5 practical and to take advantage of its potential. In the mean time, surveyors should occupy sites that are not directly below electromagnetic noise sources, if possible. Overhead vegetation that comes between the receiver and the SVs can also attenuate or block the SV transmissions, causing the same problems as with decreased signal to noise ratios (Spilker 1996; Meyer et al. 2002).

### Error Summary

Table IV.1 provides a summary of error sources and recommended remedies.

## NGS Guidelines for GPS Ellipsoid and Orthometric Heighting

NGS has guidelines and suggested practices that, if followed exactly, are intended to achieve ellipsoid / orthometric height network accuracies of 5 cm (95 percent confidence level) and ellipsoid / orthometric height local accuracies of 2 cm and 5 cm (95 percent) (Zilkoski et al.1997; Zilkoski et al.2000). The **local accuracy** of a control point is defined as:

“... a value expressed in cm that represents the uncertainty in the coordinates of the control point relative to the coordinates of the other directly connected, adjacent control points at the 95 percent confidence level. The reported local accuracy is an approximate average of the individual local accuracy values between this control point and other observed control points used to establish the coordinates of the control point” (Zilkoski et al.1997).

The **network accuracy** of a control point is defined as:

“... a value expressed in cm that represents the uncertainty in the coordinates of the

control point with respect to the geodetic datum at the 95 percent confidence level. For NSRS network accuracy classification, the datum is considered to be best supported by NGS. By this definition, the local and network accuracy values at CORS sites are considered to be infinitesimal, i.e., to approach zero.” (*ibid*)

This section presents an overview of these guidelines and of currently available U.S. geoid models and how local geoid modeling is used in practice.

### Three Rules, Four Requirements, Five Procedures

The National Geodetic Survey created a series of rules, requirements and procedures to derive orthometric heights using GPS (Zilkoski et al.1997; Zilkoski et al.2000). We now review this material.

#### Three Rules

Rule 1. Follow NGS' guidelines to establish GPS-derived ellipsoid heights (Zilkoski et al. 1997) when performing a GPS survey;

Rule 2. Use NGS's latest National Geoid Model, i.e., GEOID03 (Roman et al. 2004), when computing GPS-derived orthometric heights; and

Rule 3. Use the latest National Vertical Datum, i.e., NAVD 88 (Zilkoski et al. 1992), height values to control the project's adjusted heights.

We note that GEOID03 is a **hybrid** geoid model for the conterminous U.S. and, as such, has been custom-crafted to fit properly with the NAVD 88 level surface (Milbert 1991; Milbert and Smith 1996; Roman et al. 2004; Smith and Milbert 1999; Smith and Roman 2000; Smith and Roman 2001; Smith 1998). Inferior results would likely result from using a geoid model that had not been so fitted. There are many studies on how to apply local geoid models for surveying purposes; for example see Amod and Merry (2002), Corchete et al. (2005), Featherstone and Olliver (2001), Forsberg et al. (2002), Fotopoulos (2005), Luo et al. (2005), Pellinen (1962), Soykan and Soykan (2003), and Tranes et al. in press).

Some of these are studies were across very limited areas (Soykan and Soykan 2003; Tranes et al. in press) in which the geoid could be adequately modeled with simple polynomial models. The others are local improvements over global models for regions as large as Iberia (Corchete et al. 2005), Hong Kong (Luo and Chen 2002; Luo et al. 2005; Zhan-ji and Yong-qi 2001), the

Caribbean Sea (Smith and Small 1999), Taiwan (Hwang and Hsiao 2003), and the British Isles (Featherstone and Olliver 2001; Forsberg et al. 2002; Iliffe et al. 2000; Iliffe et al. 2003), where a simple polynomial model will not suffice. These approaches depend upon absolute and relative gravity measurements.

Although it can be shown that completely rigorous orthometric heighting also depends on such data (Tenzer et al. 2005), collecting them is impractical for most surveyors. Fortunately, U.S. surveyors need not resort to such efforts because GEOID03 has been shown to be accurate at the 2 cm (95 percent confidence) level for the continental U.S (Roman et al. 2004). Although newer versions are planned to be released in the future, GEOID03 is sufficient for GPS orthometric heighting at the 2 cm and 5 cm accuracy levels as put forth by NGS, thus eliminating the need for U.S. surveyors to create their own gravimetric geoid models.

#### *Four Requirements (Control)*

##### **Requirement 1**

GPS-occupy stations with *valid* NAVD 88 orthometric heights; stations should be evenly distributed throughout (the) project.

##### **Requirement 2**

For project areas less than 20 km on a side, surround project with *valid* NAVD 88 bench marks, i.e., minimum number of stations is four; one in each corner of the project.

##### **Requirement 3**

For project areas greater than 20 km on a side, keep distance between *valid* GPS-occupied NAVD 88 bench marks to less than 20 km.

##### **Requirement 4**

For projects located in mountainous regions, occupy *valid* bench marks at the base and summit of mountains, even if distance is less than 20 km.

NGS guidelines repeatedly stress the need to tie to *valid* NAVD 88 bench marks, although (unfortunately) the criteria for validity are not discussed. Obviously, bench marks without NAVD 88 heights are not valid. This disqualifies NGVD 29 heights or bench marks tied to tide gauges. A valid bench mark is one that has been tied into NAVD 88 and has not been disturbed either by natural and human forces in such a way as to render its published NAVD 88 height inconsistent with the remainder of the network. Caution should be used in areas of ground subsidence or

uplift, such as along the U.S. Gulf Coast or in California, for example.

GNSS heighting can take advantage of four-dimensional markers, where they exist. The National Geodetic Survey has conducted “GPS-on-bench-mark” field surveys as part of its height modernization program, thereby establishing many four-dimensional markers: geodetic latitude, longitude, ellipsoid height, and Helmert orthometric height. For example, according to the data sheet for Y88 (PID LX3030) in Connecticut, Y88 is vertical First-Order, Class II; Horizontal Order A and ellipsoid Fourth Order, Class I. Four-dimensional bench marks are very useful for GNSS adjustment software packages because they eliminate the need to estimate any of the four coordinates (usually either ellipsoid or orthometric height) with a model.

Occupying bench marks at the bases and summits of mountains helps overcome error sources in geoid models typically caused by a lack of gravity measurements at such places (Featherstone and Alexander 1996; Allister and Featherstone 2001; Dennis and Featherstone 2002; Featherstone and Kirby 2000; Goos et al. 2003; Kirby and Featherstone 2001; Zhang and Featherstone 2004).

#### *Five Procedures*

1. Perform a 3-D minimum constraint least squares adjustment of the GPS survey project, i.e., constrain one latitude, one longitude, and one orthometric height value.
2. Using the results from the adjustment in procedure 1, detect and remove all data outliers.

Repeat procedures 1 and 2 until all outliers have been removed.

3. Compute differences between the set of GPS-derived orthometric heights from the minimum constraint adjustment (using the latest national geoid model, i.e., GEOID03) from procedure 2 above and published NAVD 88 bench marks.
4. Using the result from procedure 3 above, determine which bench marks have valid NAVD 88 height values. This is the most important step in the process. Determining which bench marks have valid heights is critical to computing accurate GPS-derived orthometric heights.
5. Using the results from procedure 4 above, perform a constrained adjustment fixing one latitude and one longitude value and all valid NAVD 88 height values.

Correctness is ascertained by repeatability in GPS heighting.

## Discussion and Summary

GNSS surveying is becoming more commonly used for vertical control. GNSS heighting can be attractive from a cost perspective because it offers the possibility of reducing or eliminating the need for leveling runs and trig-heighting, which are very costly. Although GNSS heighting is not a panacea, the prospect of establishing high-quality vertical control in a remote site without running levels to distant bench marks is very attractive.

Unfortunately, traditional training in leveling does not adequately prepare a surveyor to perform GNSS heighting because the two techniques are nearly completely different. For example, different instruments are used for each technique; the concept of a leveling route does not exist in GNSS heighting; they have different error budgets; they reference different vertical datums; and they are even based on different conceptualizations of height itself.

This series presented concepts such as reference ellipsoids, vertical datums, mean sea level, level surfaces and the geoid, gravity and potential, and orthometric vs. geometric vs. ellipsoid heights. From these concepts come applications such as why some reference ellipsoids are suitable as vertical datums while others are not; what is a GNSS receiver really doing when used for heights and how to integrate its measurements with those of a spirit level, and what is an orthometric correction. Finally, this last paper presented practical aspects of GNSS heighting based on suggested practices given by the National Geodetic Survey in light of its height modernization program. This paper considered network design and control, observation strategies, the role and application of geoid models, and the integration of leveled heights with GNSS-determined heights.

Although there are many issues affecting GNSS-determined orthometric heights, we believe the key points are these. GNSS heighting depends on using consistent control, control from a single, modern datum such as NAVD 88. For example, mixing heights in NGVD 29 and those referenced to a mean sea level station with NAVD 88 heights would violate this rule. Since orthometric heights are derived from ellipsoid heights by subtracting the geoid height from

them, the geoid model must be referred to the same heighting system as the control. Currently, in the United States, GEOID03 is the correct model to use, although surveys over very small areas can also benefit from polynomial-based geoid models derived from GPS-on-bench-mark observations.

The primary error factor is the difficulty to measure and model wet zenith delay. In arid regions the wet zenith delay is very small, and short occupations (even as short as 30 minutes) have been used successfully. In humid regions, this is seldom true. It has also been shown by several investigators that collecting meteorological measurements for the purpose of tropospheric delay modeling is ineffectual. However, these measurements should be collected for use as evidence regarding which baselines need to be re-observed. It has been shown by Marshall et al. (2001) that "Weather fronts may cause the GPS signal delay to vary by greater than 3 centimeters over a 1-hour period, potentially leading to ellipsoidal height errors exceeding 9 cm." Therefore, weather observations are useful not so much for tropospheric modeling as they are for detecting that a weather front may have passed through unnoticed. The key for reducing tropospheric delay errors to acceptable levels is to keep baselines very short, less than 10 km in length. By doing so the delay at both ends is nearly the same, and it is subsequently removed by post-processing differencing. The accuracy of GNSS heighting on long baselines is currently limited by wet zenith delay errors.

The importance of antenna modeling cannot be overstated, as well. Ellipsoid height errors as much as 10 cm for certain antennas can be introduced simply by failing to include phase center variation correction models in the processing. It is critical to check the database of the post-processing software to ensure that the antenna geometry is entered correctly and that a PCV model is used. Similarly, when using RINEX observations, make sure that the offsets that may come with those data have correct signs for the conventions of your software and that they ultimately refer to your control coordinates, which can be either ARP or phase center. Any mistakes here will introduce a several-centimeter bias in all baselines with an endpoint at the receiver.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the careful and constructive reviews of this series by Dr.



Dru Smith, Chief Geodesist of the National Geodetic Survey.

## REFERENCE LIST

- Allister, N.A., and W.E. Featherstone. 2001. Estimation of Helmert orthometric heights using digital bar-code leveling, observed gravity and topographic mass-density data over part of Darling Scarp, Western Australia. *Geomatics Research Australia* 75: 25-52.
- Amod, A., and C.L. Merry. 2002. The use of the two-dimensional spherical FFT for quasi-geoid modeling in South Africa. *Survey Review* 36(285): 508-20.
- Awange, J. L., and E.W. Grafarend. 2005. *Solving algebraic computational problems in geodesy and geoinformatics*. New York, New York: Springer. 333p.
- Bearman, G. 1999. *Waves, tides and shallow-water processes*, 2nd ed. Boston, Massachusetts: Butterworth-Heinemann. 224p.
- Boon, J.D. 2004. *Secrets of the tide: Tide and tidal current analysis and applications, storm surges and sea level trends*. Chichester, U.K.: Horwood Publishing. 212p.
- Brunner, F., and W. Walsh. 1993. Effect of the troposphere on GPS measurements. *GPS World* 4(1): 42-51.
- Buis, A. 2005. NASA details earthquake effects on the Earth. Jet Propulsion Laboratory, <http://www.jpl.nasa.gov/news/news.cfm?release=2005-009>.
- Butsch, F. 2002. Radiofrequency interference and GPS: A growing concern. *GPS World* 13(10): 40-9.
- Cazenave, A., and R.S. Nerem. 2002. Redistributing Earth's mass. *Science* 297 (5582): 783-4.
- Chao, B.F., and R. S. Gross. 1987. Changes in the Earth's rotation and low degree gravitational field induced by earthquakes. *Geophysics Journal of the Royal Astronomical Society* 91: 569-96.
- Corchete, V., M. Chourak, and D. Khattach. 2005. The high-resolution gravimetric geoid of Iberia: IGG2005. *Geophysical Journal International* 162 (3): 655-84.
- Dennis, M. L., and W.E. Featherstone. 2002. Evaluation of orthometric and related height systems using a simulated mountain gravity field. In: *3rd Meeting of the International Gravity and Geoid Commission: Gravity and Geoid 2002 - GG2002*, Thessaloniki, Greece. Section VI.
- Diddams, S.A., J.C. Bergquist, S.R. Jefferts, and C.W. Oates. 2004. Standards of time and frequency at the outset of the 21st century. *Science* 306(5700): 1318-24.
- Doodson, A.T. 1922. The harmonic development of the tide-generating potential. In: *Proceedings of the Royal Society of London, Series A* 100, pp. 305-29.
- Eckl, M.C., Snay, R.A., Soler, T., Cline, M.W. and Mader, G.L. 2002. Accuracy of GPS-derived relative positions as a function of interstation distance and observing-session duration. *Journal of Geodesy* 75 (12): 633-640.
- Featherstone, W.E., and K. Alexander. 1996. An analysis of GPS height determination in Western Australia. *The Australian Surveyor* 41(1): 29-34.
- Featherstone, W.E., and J.F. Kirby. 2000. The reduction of aliasing in gravity anomalies and geoid heights using digital terrain data. *Geophysical Journal International* 141(1): 204-12.
- Featherstone, W.E., and J.G. Olliver. 2001. A review of geoid models over the British Isles: Progress and Proposals. *Survey Review* 36(280): 78-100.
- Flowers, J. 2004. The route to atomic and quantum standards. *Science* 306 (5700): 1324-30.
- Forsberg, R., G. Strykowski, J.C. Iliffe, M. Ziebart, C.A. Cross, C.C. Tscherning, P. Cruddace, O. Finch, C. Bray, and K. Stewart. 2002. OSGM02: A new geoid model of the British Isles. In: *Gravity and Geoid 2002, 3rd Meeting of the International Gravity and Geoid Commission*, Thessaloniki, Greece. pp. 132-7.
- Fotopoulos, G. 2005. Calibration of geoid error models via a combined adjustment of ellipsoidal, orthometric and gravimetric geoid height data. *Journal of Geodesy* 79(1-3): 111-23.
- Gabrysch, R.K., and L.S. Coplin. 1990. Land-surface subsidence resulting from ground-water withdrawals in the Houston-Galveston region, Texas, through 1987. Harris-Galveston Coastal Subsidence District Report of Investigations 90-01. 53p.
- Georgiadou, Y., and A. Kleusberg. 1988. On carrier signal multipath effects in relative GPS positioning. *Manuscripta Geodaetica* 13: 172-9.
- Goos, J.M., W.E. Featherstone, J.F. Kirby, and S. A. Holmes. 2003. Experiments with two different approaches to gridding terrestrial gravity anomalies and their effect on regional geoid computation. *Survey Review* 37(288): 92-112.
- Hatch, R., J. Jung, P. Enge, and B. Pervan. 2000. Civilian GPS: The benefits of three frequencies. *GPS Solutions* 3(4): 1-9.
- Hein, G.W. 1985. Orthometric height determination using GPS observations and the integrated geodesy adjustment model. *NOAA Technical Report No 110 NGS 32*. National Oceanic and Atmospheric Administration, Rockville, Maryland. 16p.
- Henning, W.E., E.E. Carlson, and D.B. Zilkoski. 1998. Baltimore County, Maryland, NAVD 88 GPS-derived orthometric height project. *Surveying and Land Information Systems* 58(2): 97-113.
- Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins. 1997. *GPS theory and practice*, 4th ed. New York, New York: Springer. 389p.
- Hofmann-Wellenhof and H. Moritz. 2005. *Physical Geodesy*. New York, New York: SpringerWienNewYork. 403p.
- Hwang, C., and Y.S. Hsiao. 2003. Orthometric corrections from leveling, gravity, density and elevation data: A case study in Taiwan. *Journal of Geodesy* 77(5-6 ): 279-91.
- International GNSS Service 2005. IGS data & products. [<http://igsb.jpl.nasa.gov/components/prods.html>].



- Iliffe, J.C., W.J. Griffiths, and E.L. Message. 2000. Localized geoid determination for engineering control surveys. *Survey Review* 35 (275): 320-8.
- Iliffe, J.C., M. Ziebart, P.A. Cross, R. Forsberg, G. Strykowski, and C. C. Tscherning. 2003. OSGM02: A new model for converting GPS-derived heights to local height datums in Great Britain and Ireland. *Survey Review* 37(290): 276-93.
- Jakowski, N., S.M. Standov, and D. Klaehn. 2005. Operational space weather service for GNSS precise positioning. *Annales Geophysicae* 23: 3071-9.
- Johannessen, R. 1997. Interference: Sources and symptoms. *GPS World*: 8(11): 45-8.
- Kao, S., R. Hsu, and F. Ning. 2000. Results of field test for computing orthometric correction based on measured gravity. *Geomatics Research Australia* 72: 43-60.
- Kirby, J., and W.E. Featherstone. 2001. Anomalous large gradients in the "Geodata 9 Second" Digital Elevation Model of Australia, and their effects on gravimetric terrain corrections. *Cartography* 30(1): 1-10.
- Leick, A. 1995. *GPS satellite surveying*, 2nd ed. New York, New York: John Wiley & Sons. 560p.
- Luo, Z., and Y. Chen. 2002. Evaluation of geo-potential models EGM96, WDM94, GPM98CR in Hong Kong and Shenzhen. *Journal of Geospatial Engineering* 4(1): 21-30.
- Luo, Z., J. Ning, Y. Chen, and Z. Yang. 2005. High precision Geoid models HKGEOID-2000 for Hong Kong and SZGEOID-2000 for Shenzhen, China. *Marine Geodesy* 28(2): 191-200.
- Mader, G.L. 1999. GPS antenna calibration at the National Geodetic Survey. *GPS Solutions* 3(1): 1521-1886.
- Marshall, J., M. Schenewerk, R. Snay, and S. Gutman. 2001. The effect of the MAPS weather model on GPS-determined ellipsoidal heights. *GPS Solutions* 5(1): 1-14.
- Melchior, P. 1974. Earth tides. *Geophysical Surveys* 13: 275-303.
- Meyer, T.H., J.E. Bean, C.R. Ferguson, and J.M. Naismith. 2002. The effect of broadleaf canopies on survey-grade horizontal GPS/GLONASS measurements. *Surveying and Land Information Science* 62(4): 215-24.
- Milbert, D. G. 1991. Computing GPS-derived orthometric heights with the GEOID90 geoid height model. In: *Technical Papers of the 1991 ACSM-ASPRS Fall Convention*, Atlanta, Georgia. pp. A46-55.
- Milbert, D. G., and D. A. Smith. 1996. Converting GPS height into NAVD 88 elevation with the GEOID96 geoid height model. In: *Proceedings of GIS/LIS '96 Annual Conference and Exposition*, Denver, Colorado. pp. 681-92.
- Moritz, H. 1980. *Advanced physical geodesy*. Tunbridge Wells, U.K.: Abacus Press. 500p.
- NGS (National Geodetic Survey). 2005. GPS antenna calibration. [<http://www.ngs.noaa.gov/ANTCAL/>]
- Pellinen, L.P. 1962. Accounting for topography in the calculation of quasigeoidal heights and plumb-line deflections from gravity anomalies. *Bulletin Geodesique* 63: 57-65.
- Ray, J., M. Morrison, S. Hilla, W. Dillinger and G. Mader. 2005. Geodetic sensitivity to surface meteorological data: 24-hr and 6-hr observing sessions. *GPS Solutions* 9(1): 12-20.
- Roman, D.R., Y.M. Wang, W. Henning, and J. Hamilton. 2004. Assessment of the New National Geoid Height Model, GEOID03. *Surveying and Land Information Science* 64(3): 153-62.
- Seeber, G., F. Menge, C. Volksen, G. Wubbena, and M. Schmitz. 1997. Precise GPS positioning improvements by reducing antenna and site dependent effects. In: *IAG Symposium No. 115*, Rio de Janeiro, Brazil. pp. 237-44.
- Seeber, G. 2003. *Satellite Geodesy*. 2nd edn. New York: Walter de Gruyter. 589 pp. pp.
- Smith, D.A. and Milbert, D.G. 1999. The GEOID96 high-resolution geoid height model for the United States. *Journal of Geodesy* 73 (5): 219-236.
- Shinkle, K.D. and Dokka, R.K. 2004. *Rates of vertical displacement at benchmarks in the lower Mississippi valley and the northern Gulf Coast*. NOAA Technical Report NOS/NGS 50, U.S. Department of Commerce. 147 pp.
- Smith, D. A., and D. R. Roman. 2000. NAVD 88 Helmert orthometric heights from NAD 83 GPS heights and the GEOID99 high resolution geoid height model. In: *2000 Conference of the American Congress on Surveying and Mapping*, Little Rock, Arkansas.
- Smith, D.A., and D.R. Roman. 2001. GEOID99 and G99SSS: 1-arc-minute geoid models for the United States. *Journal of Geodesy* 75(9-10): 469-90.
- Smith, D.A. 1998. There is no such thing as "The" EGM96 geoid: Subtle points on the use of a global geopotential model. *IGeS Bulletin* 8: 17-28.
- Smith, D.A., and H.J. Small. 1999. The CARIB97 high-resolution geoid height model for the Caribbean Sea. *Journal of Geodesy* 73(1): 1-9.
- Smylie D.E., and L. Manshina. 1971. The elasticity theory of dislocation in real Earth models and changes in the rotation of the Earth. *Geophysics Journal of the Royal Astronomical Society* 23: 329-54.
- Snay, R. 2003. Horizontal time-dependent positioning. *Professional Surveyor* 23(11): 30-32.
- Snay, R.A. 1999. Using the HTDP software to transform spatial coordinates across time and between reference frames. *Surveying and Land Information Systems* 59(1): 15-25.
- Spilker, J. 1996. Foliage attenuation for land mobile users. In: *Global Positioning System: Theory and Applications*, vol. 1, American Institute of Aeronautics and Astronautics, Inc. (Progress in Astronautics and Aeronautics. Vol. 163), Washington, DC, p. 569-583.

- Soldati, G., and G. Spada. 1999. Large earthquakes and Earth rotation: The role of mantle relaxation. *Geophysical Research Letters* 26: 911-14.
- Soycan, M., and A. Soycan. 2003. Surface modeling for GPS-leveling geoid determination. *Newton's Bulletin* 1: 41-52.
- Strang van Hees, G.L. 1992. Practical formulas for the computation of the orthometric and dynamic correction. *Zeitschrift für Vermessungswesen* 117.
- Sun, W. 2002. A formula for gravimetric terrain corrections using powers of topographic height. *Journal of Geodesy* 76(8): 399-406.
- Tenzer, R., P. Vanicek, M. Santos, W.E. Featherstone, and M. Kuhn. 2005. The rigorous determination of orthometric heights. *Journal of Geodesy* 79(1-3): 82-92.
- Townsend, B. R., and P.C. Fenton. 1994. A practical approach to the reduction of pseudorange multipath errors in a L1 GPS receiver. In: *ION GPS-94*, Salt Lake City, Utah. pp. 1-6.
- Tranes, M.D., T.H. Meyer, and D. Massalski. (in press). Comparisons of GPS-derived orthometric heights using local geometric geoid models. *Journal of Surveying Engineering*.
- van Sickle, J. 1996. *GPS for land surveyors*. Chelsea, Michigan: Ann Arbor Press, Inc.
- Vanicek, P., J. Huang, P. Novak, S. Pagiatakis, M. Veronneau, Z. Martinec, and W.E. Featherstone. 1999. Determination of the boundary values for the Stokes-Helmert problem. *Journal of Geodesy* 73(4): 180-92.
- Vanicek, P. 1980. Tidal corrections to geodetic quantities. *NOAA Technical Report NOS 83 NGS 14*. National Oceanic and Atmospheric Administration, Rockville, Maryland. 30p.
- Vanicek, P., and E. J. Krakiwsky. 1996. *Geodesy: The concepts*, 2nd ed. Amsterdam, The Netherlands: Elsevier Scientific Publishing Company. 697p.
- Volgyesi, L. 2006. Physical backgrounds of Earth's rotation, revision of the terminology. *Acta Geodaetica et Geophysica Hungarica* 41(1): 31-44.
- Wilhelm, H., and H.-G. Wenzel. 1997. *Tidal phenomena*. New York, New York: Springer. 398p.
- Zhan-ji, Y., and C. Yong-qi. 2001. Determination of the Hong Kong gravimetric geoid. *Survey Review* 36(279): 23-34.
- Zhang, K., and W. E. Featherstone. 2004. Investigation of the roughness of the Australian gravity field using statistical, graphical, fractal and Fourier power spectrum techniques. *Survey Review* 37(293): 520-30.
- Zilkoski, D.B. 1990. Establishing vertical control using GPS satellite surveys. In: *Proceedings of the 19th International Federation of Surveying Congress (FIG), Commission 5, Helsinki, Finland*. pp. 282-94.
- Zilkoski, D.B., E.E. Carlson, and C.L. Smith. 2000. A guide for establishing GPS-derived orthometric heights (Standards: 2 cm and 5 cm). DRAFT v1.1. National Geodetic Survey, Silver Spring, Maryland. 30p.
- Zilkoski, D.B., J.D. D'Onofrio, and S. J. Frakes. 1997. Guidelines for establishing GPS-derived ellipsoidal heights (Standards: 2cm and 5cm)—Version 4.3. *NOAA Technical Memorandum NOS NGS-58*. National Geodetic Survey, Silver Spring, Maryland.
- Zilkoski, D.B., J.H. Richards, and G. M. Young. 1992. Results of the general adjustment of the North American Vertical Datum of 1988. *Surveying and Land Information Systems* 52(3): 133-49.
- Zilkoski, D.B., and L.D. Hothem. 1989. GPS satellite surveys and vertical control. *Journal of Surveying Engineering* 115(2): 262-81.



# GLIS

## GIS competition for high schools

The Geographic and Land Information Society is pleased to announce a GIS competition for high schools in the U.S. The contest is sponsored by the Environmental Systems Research Institute (ESRI) and is open to projects aimed at introducing students to GIS applications that lead to better management of land and other natural resources.

GIS projects completed in the preceding year can be entered. Only one project entry per school is allowed.

**Submission deadline is January 8, 2007.  
Winners will be contacted by February 1.**

Awards will be presented at the ACSM-IPLSA-MSPS annual conference and technology exhibition at St. Louis, Missouri, March 9-12, 2007.

**Check out the competition's website  
<http://www.glismo.org/giscompetition/comphome.htm>**