Magneto-ionic polarization and GPS signal propagation through the ionosphere

R. C. Moore[1] and Y. T. Morton[2]

Received 28 February 2010; revised 7 November 2010; accepted 3 December 2010; published 4 February 2011.

[1] Recent progress in high-precision GPS measurements research and applications leads to the study of higher-order ionosphere effects on GPS signal propagation. This paper focuses on second-order ionospheric effects, which are influenced by the presence of the Earth's magnetic field. Due to the presence of Earth's magnetic field, GPS signals may propagate through the ionosphere in two distinct modes: the ordinary mode and the extraordinary mode. These two modes correspond to different refractive indices, and the difference between these refractive indices affects the computation of the second-order error introduced by the ionosphere to GPS code and carrier phase approximations. The first objective of this paper is to clarify a misconception about the different modes of GPS signal propagation through the ionosphere: although the GPS signal is predominantly right-hand circularly polarized, it may propagate through the ionosphere in either the ordinary or the extraordinary mode. The second objective of this paper is to analyze the impact of the different modes of propagation on GPS solutions at different geographical locations.

Citation: Moore, R. C., and Y. T. Morton (2011), Magneto-ionic polarization and GPS signal propagation through the ionosphere, *Radio Sci.*, *46*, RS1008, doi:10.1029/2010RS004380.

1. Introduction

[2] The ionosphere significantly affects observations of the code phase delay and the carrier phase advance of Global Positioning System (GPS) signals at the two GPS frequencies currently in use (1227.60 and 1575.42 MHz) [e.g., *Brunner and Gu*, 1991; *Bassiri and Hajj*, 1993]. The ionospheric errors are dominated by the electron density of the ionospheric plasma along the signal propagation path and result in a ranging error on the order of tens of meters [e.g., *Klobuchar*, 1996]. Additional ionospheric errors are introduced (although to a lesser extent) by the presence of the Earth's magnetic field [e.g., *Brunner and Gu*, 1991; *Bassiri and Hajj*, 1993; *Datta-Barua et al.*, 2006, 2008]. When a predominantly right-hand circularly polarized (RCP) GPS signal propagates through the ionosphere, which is an anisotropic medium, it may propagate as the linear combination of two different modes: the extraordinary (or X) mode and the ordinary (or O) mode, depending on the angle between the GPS wave normal and the Earth's magnetic field. These two modes correspond to two different magneto-ionic polarizations and two different refractive indices, affecting the second-order ionospheric error introduced to GPS ranging calculations. The calculations provided herein demonstrate that accounting for the magneto-ionic polarization of the GPS signal produces a second-order error that is asymmetric about the Earth's geomagnetic equator. The specifics of the asymmetry are location dependent. For instance, we will show that near the geomagnetic equator, signals arriving from the north propagate with O mode polarization, and the second-order ionosphere carrier phase error is positive, implying that the first-order ionosphere error is an underestimation of the total signal delay. On the other hand, GPS signals arriving from the south propagate with X mode polarization, and the second-order ionosphere carrier phase error is negative, implying that the first-order ionosphere error is an overestimation of the total delay.

[3] The difference between the magneto-ionic modes of propagation was overlooked in previous studies. For example, *Bassiri and Hajj* [1993] state that "the (−) and (+) signs in equation (11) correspond to the ordinary and extraordinary waves, respectively. Ignoring the LCP

[1]Department of Electrical and Computer Engineering, University of Florida, Gainesville, Florida, USA.
[2]Electrical and Computer Engineering Department, Miami University, Oxford, Ohio, USA.

0048-6604/11/2010RS004380

component of the GPS signal which has less than 0.35% and 2.5% of the total power for L1 and L2, respectively, only the (+) sign will be of relevance to us in the subsequent analysis." In the paper, equation (11) is the expression that computes the second-order error. The paper mistakenly relates the left-hand circularly polarized (LCP) component of the GPS signal with the ordinary wave propagation mode and failed to recognize that a RCP signal can propagate in both ordinary and extraordinary modes. *Bassiri and Hajj* [1993] were not alone in this misconception: *Brunner and Gu* [1991] provide a very similar description in their Appendix A. This misconception persists today, due at least in part to the propagation of the concept from previous works. As recently as 2008, *Hoque and Jakowski* [2008] state "Since the anisotropic plasma is doubly refracting (indicated by the ±sign in refractive index equation (14)) there are actually two waves. The wave with the upper (+) sign is usually called the ordinary wave, whereas the lower (−) sign is related to the extraordinary wave. The ordinary mode is left-hand circularly polarized, while the extraordinary mode is right-hand circularly polarized [*Hartmann and Leitinger*, 1984]. However, since GPS signals are transmitted in the right-hand polarization [*Parkinson and Gilbert*, 1983], only the results of extraordinary mode are considered here." Additionally, the several articles [e.g., *Hartmann and Leitinger*, 1984; *Datta-Barua et al.*, 2006, 2008; *Hoque and Jakowski*, 2007, 2008] that have critically analyzed second-order ionospheric effects have employed the second-order error terms provided by *Bassiri and Hajj* [1993] and by *Brunner and Gu* [1991], which are inconsistent. We note that several recent articles [e.g., *Kedar et al.*, 2003; *Fritsche et al.*, 2005; *Hoque and Jakowski*, 2008; *Datta-Barua et al.*, 2008; *Petrie et al.*, 2010] employ the same formula as that derived herein to calculate second-order GPS errors, and these analyses are therefore consistent with the work presented here. None of these other works deviate from the concept that GPS signals propagate solely as an X mode polarized signal, however, and none have provided a complete mathematical justification for using this formula.

[4] This paper provides a rigorous Taylor-series expansion of the refractive index and the group refractive index for GPS signals propagating through the ionosphere, which we take to be a cold, collisionless, and magnetized electron plasma. It is demonstrated that the magneto-ionic polarization of the predominantly right-hand circularly polarized GPS signal depends on the direction of the GPS signal k vector with respect to the Earth's magnetic field. To illustrate the effect of the GPS signal magneto-ionic polarization on GPS positioning accuracy, we interpret second-order GPS carrier phase ranging errors for three geographically distinct receiver locations. This dependence is shown to produce an asymmetry in the second-order GPS ranging error about the geomagnetic equator.

2. Magneto-ionic Polarization

[5] Taking the ionosphere to be well represented by a cold, collisionless, and magnetized electron plasma, we first analyze the magneto-ionic polarization of the dominantly right-hand circularly polarized GPS signal. We then apply this analysis to interpret the expressions for the refractive index and the group refractive index, which together determine the GPS observables.

[6] The complete expression for the magneto-ionic polarization is [*Budden*, 1985, p. 70]:

$$\rho = \frac{-\frac{1}{2}jY\sin^2\Theta \pm j\sqrt{\frac{1}{4}Y^2\sin^4\Theta + (1-X)^2\cos^2\Theta}}{(1-X)\cos\Theta} \tag{1}$$

where the radical in the numerator is taken to be positive so that the "+" applies for ordinary mode (O mode) polarized signals and the "−" applies for extraordinary mode (X mode) polarized signals [*Budden*, 1985, p. 88]. In this equation, Θ is the angle between the wave normal (i.e., the k vector) and the Earth's magnetic field. The ordinary and extraordinary modes constitute two characteristic waves which may successfully propagate in the plasma. Any plane wave propagating in the plasma must be resolved into two component waves with the two polarizations. Throughout this derivation, we will take the top sign to apply to O mode and the bottom sign to apply to X mode. If ω is the angular frequency of the propagating GPS signal, X and Y are defined:

$$X = \frac{\omega_{pe}^2}{\omega^2} \quad \text{and} \quad Y = \frac{\omega_{ce}}{\omega}. \tag{2}$$

[7] The quantity ω_{pe} is known as the electron plasma frequency given by:

$$\omega_{pe} = 2\pi f_{pe} = \sqrt{\frac{N_e q_e^2}{\varepsilon_0 m_e}} \tag{3}$$

where N_e is the electron density, q_e is the charge of an electron, ε_0 is the permittivity of free space, and m_e is the mass of an electron.

[8] The quantity ω_{ce} is the electron cyclotron frequency given by:

$$\omega_{ce} = 2\pi f_{ce} = \frac{|q_e|B_0}{m_e} \tag{4}$$

where B_0 is the magnitude of the Earth's magnetic field.

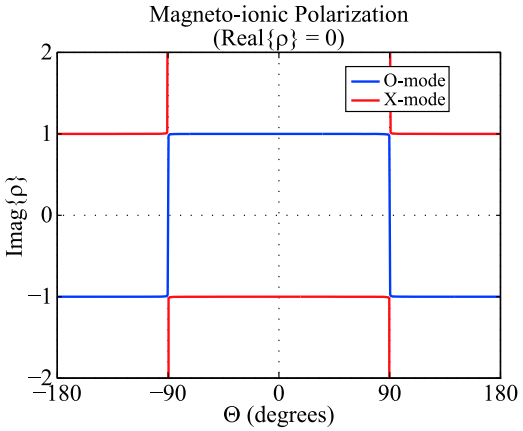

Figure 1. The imaginary part of the magneto-ionic polarization for O and X mode characteristic waves as a function of Θ, the angle between the k vector, and the direction of the Earth's magnetic field. The real part of the polarization term is zero in each case.

[9] Figure 1 shows the dependence of the O and X mode polarizations on Θ, using an electron density of 10^{12} m^{-3}, a B field strength of 2×10^{-5} Tesla, and a signal frequency of 1575.42 MHz. The relationship does not significantly change for the other GPS signal frequency of 1227.60 MHz. It can be seen from Figure 1 that the O mode characteristic wave is right-hand circularly polarized (i.e., $\rho = -j$) when $|\Theta| > \sim 90°$ and left-hand circularly polarized (i.e., $\rho = +j$) when $|\Theta| < \sim 90°$. The opposite is true for X mode characteristic waves. Because the largely right-hand circularly polarized GPS signal must be resolved into a linear combination of these two characteristic waves within the ionosphere, the GPS signal will propagate as an O mode polarized wave when $|\Theta| > \sim 90°$ and as an X mode polarized wave when $|\Theta| < \sim 90°$, as depicted in Figure 2.

[10] Within the ionosphere, the maximum values of ω_{pe} and ω_{ce} are small compared to the frequency of the GPS signal. In fact, at GPS frequencies, we may make the approximations that $X \ll 1$ and $Y \ll 1$, leading to a simplified expression for the magneto-ionic polarization:

$$\rho \approx \pm \frac{j|\cos\Theta|}{\cos\Theta} \tag{5}$$

Near $|\Theta| = 90°$, both X and O mode polarizations are linear (as indicated by $\rho_O \rightarrow 0$ and $\rho_X \rightarrow \pm\infty$ in Figure 1), and the approximate expression in equation (5) is no longer valid. The region $|\Theta| \sim 90°$ constitutes a transition region where the characteristic wave polarizations (both O and X mode) smoothly develop from one sense of circular polarization, through linear polarization, to an oppositely sensed circular polarization. Within the transition region, the right-hand circularly polarized GPS signal will propagate as two independent waves (one X mode and one O mode) through the plasma. Because the two signals propagate with different group and phase velocities, neglecting one of the modes would yield greater modeling inaccuracies. The width (in Θ) of this transition region will vary with altitude together with the magneto-ionic variables, but it can be seen from Figure 1 that the region is expected to be quite thin (in Θ). For the values used in Figure 1, equation (5) is accurate to within 1% of equation (1) for all Θ such that $||\Theta| - 90°| > \sim 1°$. In this paper, we will ignore the transition region to concentrate on the much larger regions dominated by circularly polarized characteristic waves (for which equation (5) is valid).

[11] We now proceed to derive expressions for the refractive index and group refractive index, expanded in powers of $1/f$, following the method presented by *Hartmann and Leitinger* [1984], *Brunner and Gu* [1991], and *Bassiri and Hajj* [1993].

3. The Refractive Index

[12] The refractive index is used to determine the propagation phase of the received GPS signal by integrating over the entire propagation path from satellite to ground, including propagation through the ionosphere. The refractive index in a weakly ionized, collisionless electron magnetoplasma is given by the classical Appleton-Hartree equation [*Appleton*, 1932; *Ratcliffe*, 1959]:

$$n^2 = 1 - \frac{X}{1 - \dfrac{Y^2 \sin^2\Theta}{2(1-X)} \pm \sqrt{\dfrac{Y^4 \sin^4\Theta}{4(1-X)^2} + Y^2\cos^2\Theta}} \tag{6}$$

where X, Y, and Θ are as defined above. Although we may first manipulate this equation to place the radical in

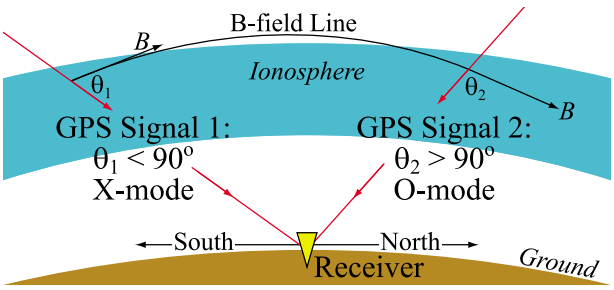

Figure 2. A cartoon depicting two typical scenarios for GPS signal propagation direction (relative to the direction of Earth's magnetic field) and their corresponding modes of propagation.

the numerator, we will instead expand it in its current form, which yields the same result. In the ionosphere, the maximum values of the electron plasma frequency and the electron cyclotron frequency are small compared to the frequencies of the GPS signals. We therefore make the approximations: $X \ll 1$ and $Y \ll 1$. When Θ is not near 90° (i.e., when we are not in the transition region), the following expression is satisfied:

$$\frac{Y^4 \sin^4 \Theta}{4(1-X)^2} \ll Y^2 \cos^2 \Theta. \tag{7}$$

[13] Using Taylor's theorem, the radical in the denominator of equation (6) may be expanded:

$$\sqrt{\frac{Y^4 \sin^4 \Theta}{4(1-X)^2} + Y^2 \cos^2 \Theta} = Y|\cos \Theta|(1+A) \tag{8}$$

where

$$A = \sum_{m=1}^{\infty} \left[\frac{Y^2 \sin^4 \Theta}{4\cos^2 \Theta(1-X)^2}\right]^m \prod_{k=1}^{m} \frac{1.5-k}{k} \tag{9}$$

Here we have forced the radical to be positive, so that the plus and minus signs in equation (6) apply, respectively, to O and X mode polarized signals. The quantity A is an infinite summation over terms with increasingly higher orders of $1/f$. In this case, because the argument of the square root is slightly larger than 1, it may be stated that A is also smaller than 1:

$$A < \frac{Y^2 \sin^4 \Theta}{4\cos^2 \Theta(1-X)^2} < 1 \tag{10}$$

At the end of this derivation, we will decide how many terms of A to retain, making the expression in equation (8) an approximation.

[14] Substituting for the radical and factoring out a $\pm Y|\cos \Theta|$ in the denominator yields the expression:

$$n^2 = 1 - \frac{X}{1 \pm Y|\cos \Theta|\left(1 \mp \frac{Y \sin^2 \Theta}{2|\cos \Theta|(1-X)} + A\right)}. \tag{11}$$

[15] By equations (7) and (10) together with the assumption that $Y \ll 1$, the denominator of the second term on the right is also close to 1, and the expression for the refractive index may be further expanded:

$$n^2 = 1 - X(1+B) \tag{12}$$

where

$$B = \sum_{m=1}^{\infty} \left[\mp Y|\cos \Theta|\left(1 \mp \frac{Y \sin^2 \Theta}{2|\cos \Theta|(1-X)} + A\right)\right]^m \tag{13}$$

The number of terms retained in A and B will be determined by the order of $1/f$ desired in our final result. It is also true in this case that $B \ll 1$ because $Y \ll 1$. We may then approximate the refractive index by expanding using the fact that $X(1 + B) \ll 1$

$$n = 1 + \sum_{m=1}^{\infty} [-X(1+B)]^m \prod_{k=1}^{m} \frac{1.5-k}{k} \tag{14}$$

The expression above may be used to evaluate the refractive index to an arbitrary order of $1/f$. If an infinite number of terms are retained in both A and B, equation (14) is exact for all Θ such that equations (7) is satisfied.

[16] As an example, let us evaluate the refractive index accurate to $1/f^4$. Starting with equation (14) and noting that $X \propto 1/f^2$, we have:

$$n = 1 - \frac{X(1+B)}{2} - \frac{X^2(1+B)^2}{8} \tag{15}$$

Since our goal is to calculate 4 orders of $1/f$, we retain 2 orders of $1/f$ in B to evaluate the second term on the right, but we neglect the B in the third term on the right.

[17] Using equation (13), we expand B to 2 orders of $1/f$ as:

$$B \approx \mp Y|\cos \Theta|\left[1 \mp \frac{Y}{2|\cos \Theta|}\left(1+\cos^2 \Theta\right) + A\right] \tag{16}$$

It is apparent that only 1 order of $1/f$ is needed in A, which (from equation (9)) has a lowest order of $1/f^2$. We therefore take $A = 0$, and B becomes:

$$B \approx \mp Y|\cos \Theta| + \frac{1}{2}Y^2\left(1+\cos^2 \Theta\right). \tag{17}$$

[18] Substituting into equation (15), the refractive index, accurate to $1/f^4$, may therefore be expressed:

$$n = 1 - \frac{1}{2}X \pm \frac{1}{2}XY|\cos \Theta| - \frac{1}{4}X\left[\frac{X}{2} + Y^2\left(1+\cos^2 \Theta\right)\right] \tag{18}$$

which is the same expression provided by *Bassiri and Hajj* [1993] and by *Datta-Barua et al.* [2008].

[19] From the previous discussion related to Figure 1, a right-hand circularly polarized GPS signal, which has a polarization of $-j$, corresponds to X mode polarization

when $|\Theta| < \sim 90°$ and corresponds to O mode polarization when $|\Theta| > \sim 90°$. For $|\Theta| < 90°$, the third term on the right-hand side of equation (18) is $-XY|\cos\Theta|/2$, and for $|\Theta| > 90°$, the same term is $+ XY|\cos\Theta|/2$. This second-order term may therefore be viewed as a piecewise function of Θ. Because $\cos\Theta$ crosses through zero and changes sign at $\Theta = 90°$, however, the refractive index expression for right-hand circularly polarized waves may be simplified: both requirements are satisfied if the third term on the right-hand side of equation (18) is expressed $-XY\cos\Theta/2$. Accounting for this dependence on Θ, the general expression for the refractive index of the right-hand circularly polarized GPS signal as a function of X, Y, and Θ becomes:

$$n^{RH} = 1 - \frac{1}{2}X - \frac{1}{2}XY\cos\Theta - \frac{1}{4}X\left[\frac{X}{2} + Y^2\left(1 + \cos^2\Theta\right)\right] \tag{19}$$

Similarly, the refractive index for the smaller magnitude left-hand circularly polarized component of the GPS signal is:

$$n^{LH} = 1 - \frac{1}{2}X + \frac{1}{2}XY\cos\Theta - \frac{1}{4}X\left[\frac{X}{2} + Y^2\left(1 + \cos^2\Theta\right)\right]. \tag{20}$$

[20] These expressions may be used to calculate the carrier phase error of the GPS signal received on the ground or near the Earth surface, and we will do so after deriving the expressions for the group refractive index.

4. The Group Refractive Index

[21] The group velocity, which is used to determine the GPS signal code phase delay between the satellite and the ground, is determined by the group refractive index. The group refractive index may be expressed as $n_{\text{group}} = n + f(dn/df)$ [*Budden*, 1985, p. 131]. Differentiating equation (18) and using

$$\frac{dX}{df} = \frac{-2f_{pe}^2}{f^3} = \frac{-2X}{f} \tag{21}$$

$$\frac{dY}{df} = \frac{-f_{ce}}{f^2} = \frac{-Y}{f} \tag{22}$$

the group refractive index, accurate to $1/f^4$, is then:

$$n_{\text{group}} = 1 + \frac{1}{2}X \mp XY|\cos\Theta| + \frac{3}{4}X\left[\frac{1}{2}X + Y^2\left(1 + \cos^2\Theta\right)\right] \tag{23}$$

This expression is also the same as that provided by *Bassiri and Hajj* [1993].

[22] The group refractive index, accurate to an arbitrary order of $1/f$, may be calculated using a refractive index approximation of the same order. Although this expression is derived by differentiating the approximation of the refractive index, the same result is found (with tedious algebra) by expanding the complete expression for the group refractive index given by *Budden* [1985, p. 131].

[23] Accounting for the magneto-ionic polarization of the signal as a function of Θ using the same technique applied to the refractive index above, the expressions for the group refractive index for right-hand circularly polarized and left-hand circularly polarized GPS signals are:

$$n_{\text{group}}^{RH} = 1 + \frac{1}{2}X + XY\cos\Theta + \frac{3}{4}X\left[\frac{1}{2}X + Y^2\left(1 + \cos^2\Theta\right)\right] \tag{24}$$

$$n_{\text{group}}^{LH} = 1 + \frac{1}{2}X - XY\cos\Theta + \frac{3}{4}X\left[\frac{1}{2}X + Y^2\left(1 + \cos^2\Theta\right)\right]. \tag{25}$$

5. The GPS Observables

[24] With expressions for the refractive index and the group refractive index in hand, we now derive expressions for the GPS observables, which are dominated by the right-hand circularly polarized portion of the GPS signal transmission. The carrier phase, ϕ, accumulated along the raypath, l, and the total propagation time, τ, from satellite to receiver are given by [*Budden*, 1985, pp. 426, 412]:

$$\phi = \frac{\omega}{c}\int n^{RH}\cos\alpha dl \tag{26}$$

$$\tau = \frac{1}{c}\int n_{group}^{RH}\cos\alpha dl \tag{27}$$

where c is the speed of light in a vacuum, and α is the angle between the ray (i.e., the Poynting vector) and the wave normal as defined by [*Budden*, 1985, p. 77]:

$$\alpha = \tan^{-1}\left[\pm\frac{\frac{1}{2}Y\sin\Theta\cos\Theta(n^2 - 1)}{(1 - X)|\cos\Theta|\sqrt{1 + \frac{Y^2\sin^4\Theta}{4(1-X)^2\cos^2\Theta}}}\right] \tag{28}$$

Converting to a dimension of length, we express the standard GPS observables relative to the line of sight

Table 1. The Three Locations Selected for This Study

Location	Latitude	Longitude
Arecibo, Puerto Rico	18° 20′N	66° 45′W
Jicamarca, Peru	11° 57′N	76° 52′W
Bahia Blanca, Argentina	38° 43′S	62° 16′W

distance, d, between the satellite and the receiver [*Bassiri and Hajj*, 1993]:

$$L = \phi\frac{c}{w} - d = \int \left(n^{RH}\cos\alpha - 1\right)dl \qquad (29)$$

$$P = \tau c - d = \int \left(n^{RH}_{group}\cos\alpha - 1\right)dl. \qquad (30)$$

[25] In equation (28), the argument of the inverse tangent is small, and α may therefore also be expanded in powers of $1/f$. Equation (12) shows that the lowest order of $(n^2 - 1)$ is $1/f^2$, making the lowest order of $Y(n^2 - 1)$ equal to $1/f^3$. Because the lowest order of $\tan^{-1}(x)$ is x, the lowest order of α is also $1/f^3$. The Taylor expansion of $\cos\alpha$ is then $1 + O(1/f^6)$. Thus, for the purpose of calculating values up through $1/f^5$ (and for our example case through $1/f^4$), $\cos\alpha$ may be taken to be equal to 1.

[26] Using equations (19), (24), (29), and (30) together with the approximation $\cos\alpha \approx 1$, the GPS observables may be expressed with accuracy to $1/f^4$:

$$L_i = \mu + m_i\lambda_i - \frac{q}{f_i^2} - \frac{s}{2f_i^3} - \frac{r}{3f_i^4} \qquad (31)$$

$$P_i = \mu + \frac{q}{f_i^2} + \frac{s}{f_i^3} + \frac{r}{f_i^4} \qquad (32)$$

where the subscript i indicates the frequency used, μ includes all other dispersive and nondispersive measurement errors and propagation errors, m represents an unknown integral number of wavelengths, and:

$$q = \frac{1}{2}\int f_{pe}^2 dl = 40.31\int N_e dl = 40.31 \times \mathrm{TEC} \qquad (33)$$

$$s = \int f_{ce} f_{pe}^2 \cos\Theta dl = 7527c\int N_e B_0 \cos\Theta dl \qquad (34)$$

$$r = \int \left[\frac{3}{8}f_{pe}^4 + \frac{3}{4}f_{pe}^2 f_{ce}^2\left(1 + \cos^2\Theta\right)\right]dl = 2437\int N_e^2 dl + 4.738 \times 10^{22}\int N_e B_0^2\left(1 + \cos^2\Theta\right)dl \qquad (35)$$

The final expression provided here for the second-order error term, s, is different from those provided by *Brunner and Gu* [1991], *Bassiri and Hajj* [1993], and *Datta-Barua et al.* [2006], which is:

$$s = \int f_{ce} f_{pe}^2 |\cos\Theta| dl = 7527c\int N_e B_0 |\cos\Theta| dl. \qquad (36)$$

[27] The only difference between equations (34) and (36) is the lack of absolute value signs in the former. This is a subtle, but important, difference. Physically, equation (36) describes the second-order error for a GPS signal that propagates only with Xmode polarization, whereas equation (34) describes the more general case where the GPS signal may propagate with either X or O mode polarization. Because the mode of propagation depends on the signal direction with respect to the Earth's magnetic field, this new expression is expected to be more accurate for receiver sites at and south of the geomagnetic equator, as GPS receivers at these locations will detect a higher number of O mode propagating GPS signals.

[28] We point out that equation (34) has been used in several recent articles. *Kedar et al.* [2003], *Fritsche et al.* [2005], *Hernandez-Pajares et al.* [2007], *Hoque and Jakowski* [2008], *Datta-Barua et al.* [2008], *Morton et al.* [2009a], and *Petrie et al.* [2010] each provide the same expression for the second-order error term. We are not aware of any article providing a mathematical or physical justification for this expression, however. The final expression provided here for the second-order error term, s, is different from those provided by *Brunner and Gu* [1991], *Bassiri and Hajj* [1993], and *Datta-Barua et al.* [2006], surrounding the second-order error term with confusion. This section has attempted to provide a physical and mathematical framework alleviating this confusion.

6. GPS Signal Propagation Mode Impact on Receiver Position Error

[29] GPS carrier phase measurement errors directly translate to receiver position solution error. To access the position error that may be directly attributed to the second-order ionosphere error (which is dependent on GPS signal propagation mode), we first compute the second-order carrier phase error for GPS signals arriving from all possible directions at the three locations listed in Table 1 over a 24 h period. The second-order carrier phase errors are then used to compute the resulting receiver position error in the receiver local ENU (east, north, and vertical) coordinate. We used the same approaches as described by *Morton et al.* [2009b] to

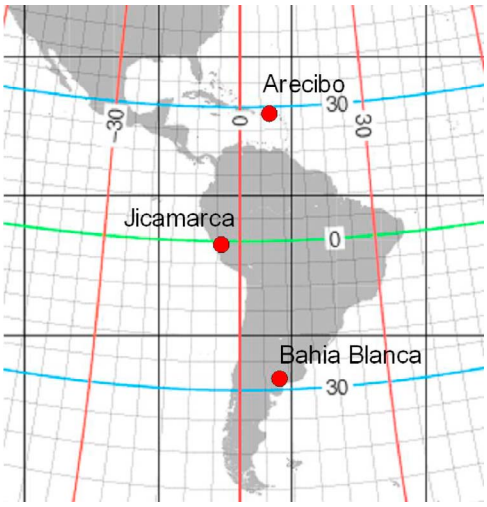

Figure 3. A map showing the locations of the three selected sites for second-order ionosphere error studies. Grid lines depict geomagnetic latitude and longitude.

compute the s values and the receiver position errors. A summary of the computation approaches is provided below for the sake of clarity and completeness.

[30] As shown in equation (34), electron densities and the magnetic field vectors along the GPS signal propagation path are needed to compute s. For each location listed in Table 1, a 24-satellite constellation is used to generate GPS satellite positions in direct view of the receiver over a 24 h period in 10 min increments. The constellation used is as defined by DoD as a baseline constellation to evaluate GPS system performance [*U.S. Department of Defense*, 2008]. The magnetic field vectors were generated along each satellite-receiver signal propagation path (assumed to be line-of-sight paths) using the tenth generation of International Geomagnetic Reference Field (IGRF) model [*Macmillan and Maus*, 2005].

[31] The electron density profiles along GPS signal propagation path were more difficult to obtain. Since the purpose of our computation is to illustrate the impact of the GPS signal propagation mode on receiver position error, we utilized the vertical electron density profiles inferred from the Arecibo incoherent scatter radar (ISR) measurements taken on 13 April 2000, which was a quiet day characterized by a low K_p index (0 to 3.67) throughout the day. The electron density measurement profiles are limited to 56–500 km altitude above Arecibo, Puerto Rico. We extended the vertical measurement profiles to 2000 km altitude by incorporating the IRI model with the ISR measurements [*Bilitza et al.*, 2006; *Morton et al.*, 2009b]. The same vertical electron density profiles were assumed for Jicamarca, Peru, and Bahia Blanca, Argentina. In order to assess the second-order error for signals arriving from arbitrary azimuth and elevation angles, we assume that electron density profiles at all neighboring areas have the same shape and their magnitudes differ from each other by a scale factor. This scale factor can be derived from their total electron content (TEC) values available at the International GNSS Services (IGS) [*Hernandez-Pajares*, 2004]. *Morton et al.* [2009b] describes the detailed process of obtaining the electron density profiles for an arbitrary satellite-receiver path.

[32] We apply the electron density profiles and magnetic field vectors obtained using the above described approach to equation (34) and computed the GPS L1 signal second-order carrier phase error throughout 13 April 2000 for the three locations listed in Table 1. Figure 3 shows the three sites on a geomagnetic latitude and longitude grid to illustrate the rationale in selecting these three locations. Jicamarca, Peru, is located near the magnetic equator so that all GPS signal arriving from the south will be propagating in X mode while those propagating from the north will be in O mode, a nearly 50–50 split. At Arecibo, Puerto Rico, more than 50% of the sky will be filled with signals propagating in X mode, due to its ~30° north magnetic latitude. Bahia Blanca is close to the geomagnetic conjugate point of Arecibo. We expect the majority of the signals arriving at the site to be propagating with O mode polarization. For example, Figure 4 shows the sky plots of the second-order carrier phase error at 1330 h local time at each site. The black lines on each plot indicate the satellite locations for which the second-order error at the receiver is zero. The errors are negative for signals arriving from directions below the black lines, corresponding to X mode propagation, whereas the errors are positive for signals arriving from directions above the black lines, corresponding to O mode propagation. This description is consistent with the cartoon diagram shown in Figure 2.

[33] To isolate the position error due to the second-order carrier phase error, we set all other carrier phase measurement errors, including the first-order ionosphere error (which can be easily obtained with a dual frequency receiver) to zero. If there is no measurement error or propagation error, the true distance between a satellite and a receiver d_k should be equal to the carrier phase measurement r_k (converted to dimension of length):

$$r_k = d_k = \sqrt{(x_k - x)^2 + (y_k - y)^2 + (z_k - z)^2} \qquad (37)$$

where (x_k, y_k, z_k) and (x, y, z) are the kth satellite and the receiver's Earth-Center Earth-Fixed (ECEF) coordinates, respectively. The existence of the second-order ionosphere

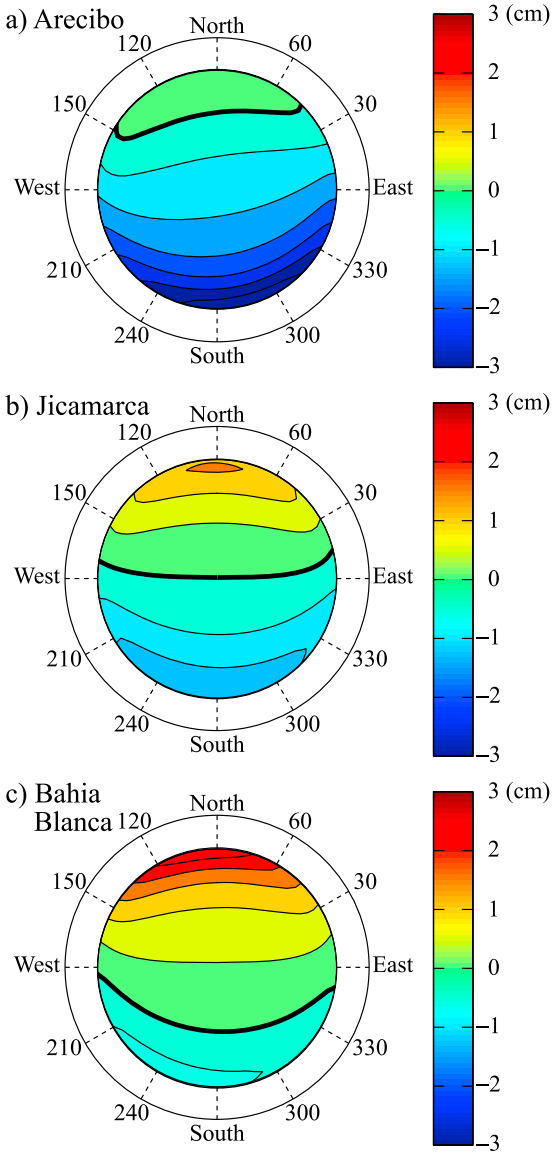

Figure 4. Sky plots of second-order carrier phase errors for (a) Arecibo, Puerto Rico, (b) Jicamarca, Peru, and (c) Bahia Blanca, Argentina, on 13 April 2000, 1330 h local time.

error in the carrier phase-based range measurements leads to the following range equation:

$$\hat{r}_k = d_k - \frac{s_k}{f^3} \tag{38}$$

Applying the conventional linear model approach to solve for the receiver position embedded in equation (38) [*Misra and Enge*, 2005, chapter 6], we obtain the receiver position solution errors:

$$\begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} = \left(G^{\mathrm{T}}G\right)^{-1} G^{\mathrm{T}} \begin{bmatrix} -\frac{s_1}{f^3} \\ -\frac{s_2}{f^3} \\ \vdots \\ -\frac{s_K}{f^3} \end{bmatrix} \tag{39}$$

where $\delta x = x - \hat{x}$, $\delta y = y - \hat{y}$, and $\delta z = z - \hat{z}$, representing the differences between the true receiver position coordinates (x, y, z) and the estimated coordinates $(\hat{x}, \hat{y}, \hat{z})$ if we do not take into consideration the second-order error. To avoid confusion, we call the estimated coordinate $(\hat{x}, \hat{y}, \hat{z})$ the uncorrected position coordinate. K is the total number of satellites in view, and G is the so-called geometry matrix:

$$G = \begin{bmatrix} \delta\vec{l}_1 \\ \delta\vec{l}_2 \\ \vdots \\ \delta\vec{l}_K \end{bmatrix} \tag{40}$$

where $\delta\vec{l}_k$ is a unit vector pointing from the receiver to the kth satellite in view. Using the baseline 24-satellite constellation, we can compute the G matrix for a given instant at each of the three locations listed in Table 1. Combined with the second-order error computed using the method described earlier in this section, equations (39) and (40) can now be used to obtain the receiver ECEF coordinate error. The ECEF coordinate error can be further converted to the receiver local ENU (east, north, and vertically up) coordinate to reveal the receiver east, north, and vertical position errors [*Misra and Enge*, 2005, Appendix 4.A]:

$$\begin{bmatrix} \delta E \\ \delta N \\ \delta U \end{bmatrix} = \begin{bmatrix} -\sin\eta & \cos\eta & 0 \\ -\sin\psi\cos\eta & -\sin\psi\sin\eta & \cos\psi \\ \cos\psi\cos\eta & \cos\psi\sin\eta & \sin\psi \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} \tag{41}$$

where $\delta E = E - \hat{E}$, $\delta N = N - \hat{N}$, $\delta U = U - \hat{U}$, and ψ and η are the latitude and longitude of the receiver location. With δE, δN, δU defined in this manner, equation (41) essentially calculates the position correction terms to be applied to the uncorrected receiver position coordinate to yield the true receiver position (i.e., it is the negative of the position error).

[34] To contrast the impact of the propagation mode on the position correction terms, two sets of second-order correction terms are computed for each location. One is

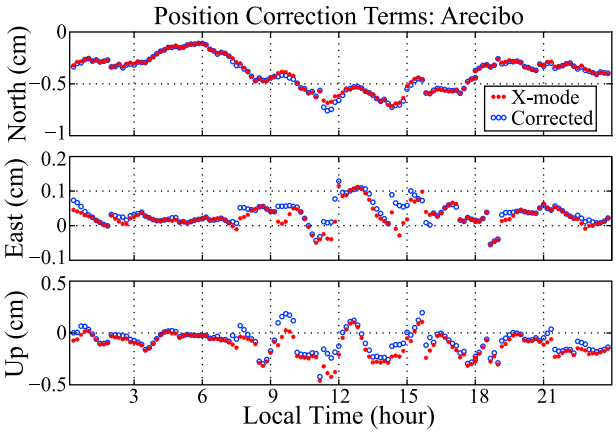

Figure 5. North, east, and vertical position correction terms resulting from the presence of a second-order carrier phase error under two scenarios at Arecibo, Puerto Rico: the "X mode" traces correspond to the second-order carrier phase errors computed assuming only X mode propagation, while the "Corrected" traces result from considering both X and O mode propagation.

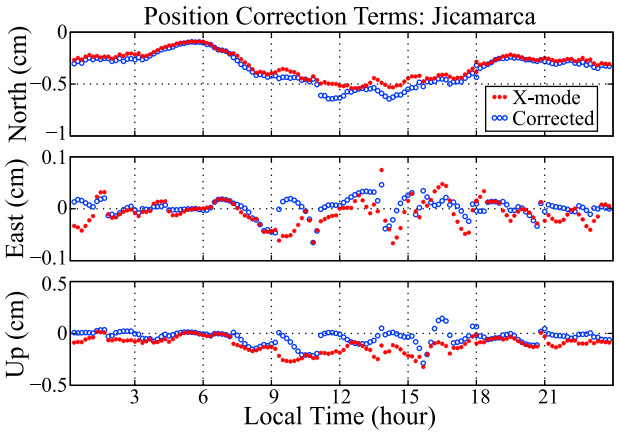

Figure 6. Same as Figure 5 for a receiver located near the magnetic equator at the Jicamarca Radio Observatory in Peru.

based on equation (34) which takes into consideration both X and O mode propagation, while the other is based on equation (36) where only the X mode is allowed. Figures 5–7 show the north, east, and vertical local coordinate correction terms δE, δN, and δU for the three locations, respectively. In each plot, there are two sets of data points: the "Corrected" traces result from applying equation (34), while the "X mode" traces result from applying the incorrect equation (36). For each of these plots, a positive north, east, or vertical correction term indicates that the true position is north, east, or above the uncorrected position, respectively. On the other hand, a negative north, east, or vertical correction term indicates the true position is south, west, or below the uncorrected position.

[35] Figures 5–7 show interesting differences in position correction terms as a result of the signal propagation mode treatment. For example, Figure 5 clearly shows negative north corrections at Arecibo throughout the 24 h, indicating that the true position is south of the uncorrected solution. This is in agreement with previous studies [*Kedar et al.*, 2003; *Fritsche et al.*, 2005]. This phenomenon can be explained by the following reasoning. At Arecibo, Puerto Rico, a larger percentage of the space in the direct GPS satellite viewing area is below the zero-error line. Signals arriving from these directions propagate in X mode with negative second-order carrier phase error, as shown in Figure 4a. As a result, we expect the majority of the signals to have carrier phase advances and therefore, an overestimation of true range. The overestimation occurs to signals arriving from the south of the zero error line shown in Figure 4. If uncorrected, the estimated position will have a northward bias due to the second-order ionosphere error. Also notice that the two sets of solutions are very similar in all three subplots of Figure 5. This is expected because according to Figure 4a, the majority of the signal directions of arrival have less than 90° angles with the B field vector and these signals propagate in the X mode. Only signals from a relatively small area above the zero error line propagate in the O mode.

[36] Near the magnetic equator at Jicamarca, we expect signals propagating from the local southern part of the sky to be X mode polarized (and hence have a negative second-order error) and signals propagating from the

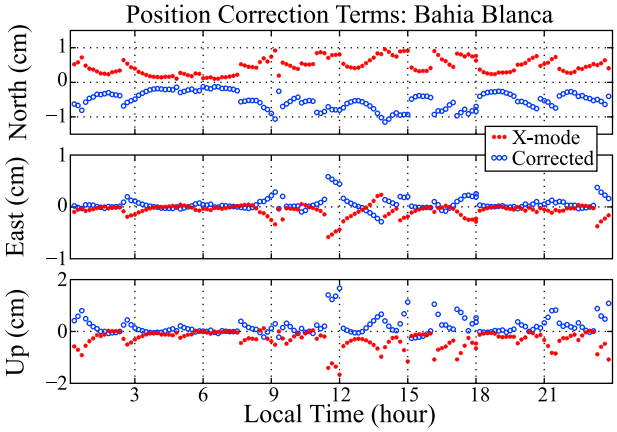

Figure 7. Same as Figures 5 and 6 for a receiver located in Bahia Blanca, Argentina.

local northern part of the sky to be O mode polarized (and have a positive second-order error). Compared to Arecibo, more signals propagate with O mode polarization to Jicamarca. As a result, we expect there to be a somewhat larger discrepancy between the two sets of position correction terms at Jicamarca. A comparison between Figures 5 and 6 demonstrate that this is the case: while the "X mode" and "Corrected" traces in Figure 5 closely follow each other, there are more separations between the traces in Figure 6.

[37] Figure 7 presents the position correction terms at Bahia Blanca, Argentina, which is located in the magnetic southern hemisphere. At this location, a larger percentage of the signals is propagating from the local northern portion of the sky against the magnetic field lines. As can be seen in Figure 4c, the corresponding second-order errors are positive, indicating a carrier phase delay or an underestimation of the range. Therefore, we expect a larger northward position bias if the second-order error is not corrected (i.e., the true position coordinate is southward of the estimated position coordinate). If we treat all signals as propagating only in X mode, the errors will retain the negative sign, however, and instead indicate a southward position bias, as shown by the positive position correction terms in Figure 7. In stark contrast, the "Corrected" trace in Figure 7 properly indicates that the estimated position has a northward bias.

[38] The examples shown in Figures 5–7 illustrate that if the correct signal propagation mode is not taken into consideration, the second-order error correction scheme will cause the most damage to position solutions at locations south of the magnetic equator and the amount of error decreases moving northward. For a quiet day as we selected to show here, the position errors caused by ignoring the mode of propagation can range from a few centimeters in the local afternoon to a few millimeters in the early morning. The largest error components are horizontal errors in the north-south and vertical directions. For dual-frequency receivers where only the first-order error is eliminated, the residual second-order error will result in a northward bias regardless of the location of the receiver. This northward bias is consistent with the overall southward correction trends reasoned by other researchers [*Kedar et al.*, 2003; *Fritsche et al.*, 2005].

[39] If an attempt is made to mitigate the second-order error, however, and only X mode propagation is considered, the consequence will be different depending on the location of the receiver. If a receiver is located in the northern hemisphere, an "X mode only" correction will account for a large portion of the actual second-order error. For a receiver located near the magnetic equator, such a correction will provide very little improvement in the horizontal errors (especially in the north-south direction). For a receiver located in the southern hemisphere, however, the "X mode only" correction is in the wrong direction, and it will further worsen the position error, especially in the north-south and vertical directions.

7. Conclusions

[40] Dual-frequency GPS receivers can only eliminate the first-order error introduced by the ionosphere. For GPS applications requiring centimeter-level accuracy, the second-order ionosphere error needs to be corrected. The second-order error is a function of the electron density distribution, the magnitude of the Earth's magnetic field, and the angle between the Earth's magnetic field vector and the GPS wave normal along the entire signal propagation path. Furthermore, the second-order error is dependent on magneto-ionic mode of the propagating GPS signal. For signals whose wave normal is less than 90° from the local magnetic field vector, the propagation mode is the X mode. Otherwise, the signal propagates in the O mode. There is a common mistaken connection made between the GPS signal polarization and the magneto-ionic mode of propagation. Such a misconception affects the estimation of the second-order impact on receiver position error correction. Several recent papers (discussed above) have used the same formula derived herein to successfully calculate and interpret second-order GPS errors. This paper has provided a physical and mathematical justification for using this formula.

[41] This paper analyzed the relationship between GPS signal polarization and magneto-ionic mode of propagation. Electron density profiles obtained using the Arecibo Incoherent Scatter Radar, TEC maps from the IGS archives, and the IGRF model are used to compute the second-order carrier phase errors near the magnetic equatorial (Jicamarca, Peru), and at two geomagnetically conjugate locations in the northern hemisphere (Arecibo, Puerto Rico) and the southern hemisphere (Bahia Blanca, Argentina). To illustrate the impact of the misinterpretation of propagation mode on position error, we computed position errors based on the assumption that only X mode propagation is allowed for GPS signals. Our computation shows that this incorrect assumption has little effect on the position errors at Arecibo because for a large portion of the sky in direct view of a receiver at this site, signals are propagating in the X mode. For Jicamarca, signals arriving from nearly half of the sky (the south) propagate in the X mode, while signals from the other half of the sky propagate in the O mode. As a result, the second-order carrier phase errors will produce a position error that is not entirely corrected by an "X mode only" interpretation. The wrong assumption on propagation mode will lead to a position estimation that is slightly northward and upward from the true position. The worst outcomes occur in southern hemisphere. At Bahia Blanca, signals arriving from most of the sky are in the O mode. Misinterpretation of these signals for the

X mode will lead to position errors that are nearly reversed in sign in the north-south and vertical directions.

[42] **Acknowledgments.** This work is supported by ONR grant N000141010909 to the University of Florida. This project is also supported by AFOSR grant FA9550-07-1-0354 to Miami University and by partial support from AFRL, Reference Systems Branch, at Wright Patterson Air Force Base. Arecibo Observatory is managed by Cornell University under a cooperative agreement with NSF.

References

Appleton, E. V. (1932), Wireless studies of the ionosphere, *J. Inst. Electr. Eng.*, *71*, 642–650.

Bassiri, S., and G. A. Hajj (1993), Higher-order ionospheric effects on the GPS observables and means of modeling them, *Manuscr. Geod.*, *18*, 280–289.

Bilitza, D., B. W. Reinisch, S. M. Radicella, S. Pulinets, T. Gulyaeva, and L. Triskova (2006), Improvements of the International Reference Ionosphere model for the topside electron density profile, *Radio Sci.*, *41*, RS5S15, doi:10.1029/2005RS003370.

Brunner, F. K., and M. Gu (1991), An improved model for the dual frequency ionospheric correction of GPS observations, *Manuscr. Geod.*, *16*, 205–214.

Budden, K. G. (1985), *The Propagation of Radio Waves: The Theory of Radio Waves of Low Power in the Ionosphere and Magnetosphere*, 669 pp., Cambridge Univ. Press, New York.

Datta-Barua, S., T. Walter, J. Blanch, and P. Enge (2006), Bounding high order ionosphere errors for the dual frequency GPS user, paper presented at ION GNSS 2006, Inst. of Navig., Fort-Worth, Tex., Sept.

Datta-Barua, S., T. Walter, J. Blanch, and P. Enge (2008), Bounding higher-order ionosphere errors for the dual-frequency GPS user, *Radio Sci.*, *43*, RS5010, doi:10.1029/2007RS003772.

Fritsche, M., R. Dietrich, C. Knöfel, A. Rülke, S. Vey, M. Rothacher, and P. Steigenberger (2005), Impact of higher-order ionospheric terms on GPS estimates, *Geophys. Res. Lett.*, *32*, L23311, doi:10.1029/2005GL024342.

Hartmann, G. K., and R. Leitinger (1984), Range errors due to ionospheric and tropospheric effects for signal frequencies above 100 MHz, *Bull. Geod.*, *58*, 109–136.

Hernandez-Pajares, M. (2004), IGS ionosphere WG status report: Performance of IGS ionosphere TEC maps, in *Celebrating a Decade of the International GPS Service, Proceedings of IGS Workshop and Symposium 2004, 1–5 March 2004*, pp. 225–250, Astronom. Inst., Univ. of Bern, Bern, Switzerland.

Hernandez-Pajares, M., J. M. Juan, J. Sanz, and R. Orus (2007), Second-order ionospheric term in GPS: Implementation and impact on geodetic estimates, *J. Geophys. Res.*, *112*, B08417, doi:10.1029/2006JB004707.

Hoque, M. M., and N. Jakowski (2007), Higher order ionospheric effects in precise GNSS positioning, *J. Geod.*, *81*, 259–268.

Hoque, M. M., and N. Jakowski (2008), Estimate of higher order ionospheric errors in GNSS positioning, *Radio Sci.*, *43*, RS5008, doi:10.1029/2007RS003817.

Kedar, S., G. A. Hajj, B. D. Wilson, and M. B. Heflin (2003), The effect of the second order GPS ionospheric correction on receiver positions, *Geophys. Res. Lett.*, *30*(16), 1829, doi:10.1029/2003GL017639.

Klobuchar, J. A. (1996), Ionospheric effects on GPS, in *Global Positioning Systems: Theory and Applications*, vol. 1, edited B. W. Parkinson and J. J. Spilker Jr., pp. 485–515, Am. Inst. of Aeronaut. and Astronaut., Washington, D. C.

Macmillan, S., and S. Maus (2005), International geomagnetic reference field—The tenth generation, *Earth Planets Space*, *57*, 1135–1140.

Misra, P., and P. Enge (2005), *Global Positioning System, Signals, Measurements, and Performance*, 2nd ed., Ganga-Jamuna, Lincoln, Mass.

Morton, Y. T., Q. Zhou, and F. van Graas (2009a), Assessment of second order ionosphere error in GPS range observables using Arecibo incoherent scatter radar measurements, *Radio Sci.*, *44*, RS1002, doi:10.1029/2008RS003888.

Morton, Y. T., F. van Graas, Q. Zhou, and J. Herdtner (2009b), Assessment of the higher order ionosphere error on position solutions, *Navigation*, *56*(3), 185–193.

Parkinson, B. W., and S. W. Gilbert (1983), NAVSTAR: Global Positioning System—Ten years later, *Proc. IEEE*, *71*, 1177.

Petrie, E. J., M. A. King, P. Moore, and D. A. Lavallee (2010), Higher-order ionospheric effects on the GPS reference frame and velocities, *J. Geophys. Res.*, *115*, B03417, doi:10.1029/2009JB006677.

Ratcliffe, J. A. (1959), *The Magneto-Ionic Theory and Its Applications to the Ionosphere*, 206 pp., Cambridge Univ. Press, Cambridge, U. K.

U.S. Department of Defense (2008), *Global Positioning System Standard Positioning Service Performance Standard*, 4th ed., 17 pp., Washington, D. C.

R. C. Moore, Department of Electrical and Computer Engineering, University of Florida, 319 Benton Hall, PO Box 116200, Gainesville, FL 32611, USA. (moore@ece.ufl.edu)

Y. T. Morton, Electrical and Computer Engineering Department, Miami University, Oxford, OH 45056, USA. (mortonyt@muohio.edu)