A review of higher order ionospheric refraction effects on dual frequency GPS

Elizabeth J. Petrie[1], Manuel Hernández-Pajares[2], Paolo Spalla[3], Philip Moore[1], Matt A. King[1]

[1] *School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK*

[2] *Research Group of Astronomy and Geomatics, Technical University of Catalonia (gAGE/UPC), Mod. C3, Campus Nord UPC, Jordi Girona 1, E-08034 Barcelona, Spain.*

[3] *IFAC-CNR, Via Madonna del Piano 10, 50019 Sesto Fiorentino, Italy*

Tel: +44 (0)191 222 7833

Fax: +44 (0)191 222 6502

Email: elizabeth.petrie@ncl.ac.uk

Abstract: Higher order ionospheric effects are increasingly relevant as precision requirements on GPS data and products increase. The refractive index of the ionosphere is affected by its electron content and the magnetic field of the Earth, so the carrier phase of the GPS L1 and L2 signals is advanced and the modulated code delayed. Due to system design the polarisation is unaffected. Most of the effect is removed by expanding the refractive index as a series and eliminating the first term with a linear combination of the two signals. However, the higher order terms remain. Furthermore, transiting gradients in refractive index at a non-perpendicular angle causes signal bending. In addition to the initial geometric bending term, another term allows for the difference that the curvature makes in electron content along each signal. Varying approximations have been made for practical implementation, mainly to avoid the need for a vertical profile of electron density. The magnetic field may be modelled as a tilted co-centric dipole, or using more realistic models such as the International Geomagnetic Reference Field. The largest effect is from the second term in the expansion of the refractive index. Up to several cm on L2, it particularly affects z-translation, and satellite orbits and clocks in a global network of GPS stations. The third term is at the level of the errors in modelling the second order term, while the bending terms appear to be absorbed by tropospheric parameters. Modelling improvements are possible, and three frequency transmissions will allow new possibilities.

Keywords: higher-order effects, ionosphere, GPS, refractive index, signal bending

This is the author version - the final publication is available at www.springerlink.com/content/0132115683q12701/:

Petrie, E., M. Hernández-Pajares, et al. (2011). "A Review of Higher Order Ionospheric Refraction Effects on Dual Frequency GPS." Surveys in Geophysics **32**(3): 197-253. doi 10.1007/s10712-010-9105-z

Contents

1. Introduction

1.1. Significance of higher-order ionospheric corrections

The potential precision of Global Positioning System (GPS) data has developed significantly since the launch of the first GPS satellite in 1978 and since the achievement of initial operational capability with a 24 satellite constellation at the end of 1993. For a quick introduction to GPS see e.g. Kintner and Ledvina (2005) while for further background several textbooks are available such as those by Hofmann-Wellenhof et al. (2001) or Leick (2004). This development in precision has been possible not only due to advances in receiver technology, but also to improved understanding of systematic errors affecting GPS signals and modelling of these errors in scientific GPS processing software. One such systematic error is the effects of the ionosphere on GPS signals. The signals are electromagnetic waves with a pseudo-random code modulated on to a carrier phase, transmitted at two frequencies. Ionosphere related errors are of particular interest; the activity of the ionosphere varies in ~11 year cycles, as well as seasonally and on a daily scale, so the errors have the potential to produce long term systematic variations.

The presence of the free electrons making up the ionosphere alters the refractive index of the atmosphere. The effect is frequency dependent (see Section3 for further details) and results in the phase of the transmitted signal being advanced and the code transmitted along the GPS signal being delayed. In addition to the carrier phase advance and code delay, there are other effects such as scintillation and fading (see e.g. Klobuchar, 1996). Scintillations occur when the GPS signal traverses a region of irregularities in the refractive index, and both phase and amplitude scintillations exist (Kintner and Ledvina, 2005). However, while they are of interest (e.g. Kim and Tinin, 2009), they are outside the scope of this review.

The design of the GPS signals with two transmitted carrier frequencies (L1 and L2) aimed to minimise the effects of the ionosphere by making signal combination possible. However, while combining the two signals in the 'ionosphere-free' linear combination cancels the first term in a series expansion of the refractive index of the ionosphere (see Section 4.2), it leaves errors due to the higher order terms in the series which do not completely cancel. There are also systematic errors due to bending of the signals, caused by the signals passing at an angle through gradients in refractive index. The bending also affects L1 and L2 frequencies differently so they take slightly different paths, meaning that the 'ionosphere-free' linear combination may no longer completely cancel the first refractive index term.

As the relative size of these higher order errors is much smaller – mm to cm at times of high ionospheric content compared to tens of metres for the main first order ionospheric error – until recently the higher order errors were considered negligible. However, geodetic GPS processing is now aiming for rates of movement known to well within 1mm/yr (Plag, 2005). There is a wide range of applications, varying from vertical land motion estimates for calibration of tide gauges and comparison with glacial isostatic adjustment (Bouin and Wöppelmann, 2010), to tectonic strain (e.g. Calais et al., 2005) and long term subsidence monitoring of oil rigs (Plag, 2005). An excellent review of some of the many geodetic applications of GPS can be found in Herring (1999). Due to this developing requirement for precision, higher order ionospheric effects have become an area of active research for the past few years. They are also relevant to the rapidly developing field of radio occultation (Hoque and

Jakowski, 2010), which uses satellite-based receivers to study the atmosphere. A review of higher order ionospheric effects on GPS is thus very timely.

Readers who are primarily interested in the effects of the corrections may wish to omit section 3, dealing with the magneto-ionic theory. Those who wish to understand the derivation of the corrections will hopefully find it provides a useful synthesis of what can be a rather complex literature. Any readers interested purely in the magnitudes and geographic distribution of the effects will find this information in section 7.

1.2. Brief review of historical context

The following is a brief review of most of the significant papers relating to higher order ionospheric corrections on GPS signals. Aspects of the papers will be discussed in much greater depth in later sections, but below is a brief chronological outline of developments in both the refractive index terms and bending terms as applied to GPS.

Hartmann and Leitinger (1984) gave an excellent discussion of the higher order ionospheric effects on signal frequencies above 100 MHz, in terms of residual errors. Brunner and Gu (1991) applied some of the concepts to GPS using ray tracing methods. Bassiri and Hajj (1993) proposed simplifications to the modelling such that the terms could be modelled practically, and limited their consideration to second and third order effects, with no bending effects. Kedar et al. (2003) were the first to consider the effects on GPS coordinates of the second order term. However, they did not estimate the effects on the orbits, so the coordinate results were somewhat misleading. Munekane (2005) performed a semi-analytical simulation looking at coordinates and transformation parameters and noted that the effects of the second order term on coordinates are considerably reduced if orbits are estimated. Fritsche et al. (2005) used a global network of sites to estimate the effects of the second and third order terms on translations over a three year period and coordinate effects for a single day. Steigenberger et al. (2006) used the same data set but showed mean coordinate effects over a three year period. Hernandez-Pajares et al. (2007) implemented a more sophisticated magnetic model and also obtained the Total Electron Content (TEC) along the signal path directly from the GPS results. A study by Petrie et al. (2010b) extended the comparison period to 14 years, which enabled the assessment of effects on GPS rates. This study also considered the effects on co-ordinates and transformation parameters and investigated different modelling options. To look at the behaviour over still longer time periods, simulation is necessary. Palamartchouk (2010) used the International Reference Ionosphere (IRI) model to simulate higher order ionospheric effects on a uniform GPS network over several solar cycles. Pireaux et al. (2010) consider the effects of second and third order errors on time transfer. Morton et al. (2009b) used incoherent scatter radar data from Arecibo to investigate more accurate modelling of the second order error, while Morton et al. (2009a) extended the work to include the third order error and simulation of potential effects on GPS coordinates at Arecibo.

Linked to the issue of the higher order terms of the refractive index series expansion is the geometric bending effect that the changes in the refractive index have on signals transiting the gradient in a non-perpendicular fashion. Hartmann and Leitinger (1984) also include this term as part of their analysis of residual errors. Brunner and Gu (1991) include it in their ray tracing study, but most of the initial practical studies applying ionospheric corrections to GPS mentioned above focused on the initial higher order refractive index expansion terms. Jakowski et al. (1994) studied the effects on the signal and suggested a simplified approach. Hoque and Jakowski (2008b) consider the theory behind the corrections and suggest an additional term due to the TEC

difference between the two signals that arises because they are bent by different amounts and so travel by slightly different paths through the atmosphere. Hoque and Jakowski (2010) extend this work to look at higher-order effects on radio occultation measurements. Petrie et al. (2010a) investigates the potential effects of the bending terms described by Hoque and Jakowski (2008b) on global GPS network processing using the International Reference Ionosphere (IRI) to provide estimates of additional ionospheric parameters.

Studies have also been done on the significance of higher-order ionospheric effects for Very Long Baseline Interferometry (VLBI), (e.g. Herring, 1983; Hawarey et al., 2005) and satellite altimetry (Imel, 1994). At present higher-order ionospheric effects are still considered negligible for these techniques due to the higher frequencies used. For the DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite) system, the situation is more complex, as it has a larger frequency gap and uses Doppler measurements. The draft International Earth Rotations and References System Service (IERS) conventions chapter on models for propagation delays provides a table of transmission frequency and estimated size of higher order ionospheric corrections for "a representative subset of typical frequencies used in radio astronomy (http://tai.bipm.org/iers/convupdt/convupdt_c9.html, accessed 28 Jan 2010: p10, Table 2). It is also a good reference for general information on higher order ionospheric effects.

In this paper both the theory and application of higher order ionospheric corrections to GPS will be discussed, with more emphasis on areas where approaches have differed or there has been confusion or disagreement. To begin, it is necessary to have a basic understanding of the ionosphere, in order to then consider its higher order effects on GPS.

2. The ionosphere

2.1. Characteristics and variability

Definitions of the ionosphere vary, but the following summary by Langley (1998: p134) is suitable for the present purpose:

> 'The ionosphere is that region of the earth's atmosphere in which ionising radiation (principally from solar ultraviolet and x-ray emissions) causes electrons to exist in sufficient quantities to affect the propagation of radio waves. This definition does not impose specific limits on the height of the ionosphere. Nevertheless, it is useful to delineate some sort of boundary to the region. The height at which the ionosphere starts to become sensible is about 50 km and it stretches to heights of 1,000 km or more. Indeed, some would argue for an upper limit of 2,000 km. The upper boundary depends on what particular plasma density one uses in the definition since the ionosphere can be interpreted as thinning into the interplanetary plasma. Although the interplanetary plasma affects the propagation of the signals from space probes and the quasar signals observed in VLBI, it may be considered to lie beyond the orbits of the GPS satellites and therefore will be ignored here.'

The above definition is a fairly simple one. For example, the true situation is more complex than the impression given by "thinning into the interplanetary plasma". However, it provides a good introduction to the basic structures relevant to this paper.

The ionosphere is formed when molecules and atoms in the atmosphere are ionised by radiation and energetic charged particles from the Sun. While the flux of visible energy to the surface of the Earth from the Sun varies little (<0.5%), activity at shorter wavelengths varies by orders of magnitude, depending on the solar activity and number of sunspots on the Sun (Davies, 1990: p28). The number of sunspots and hence the solar and ionospheric activity varies on several timescales, the most prominent being an ~11 year cycle (see Fig. 1).

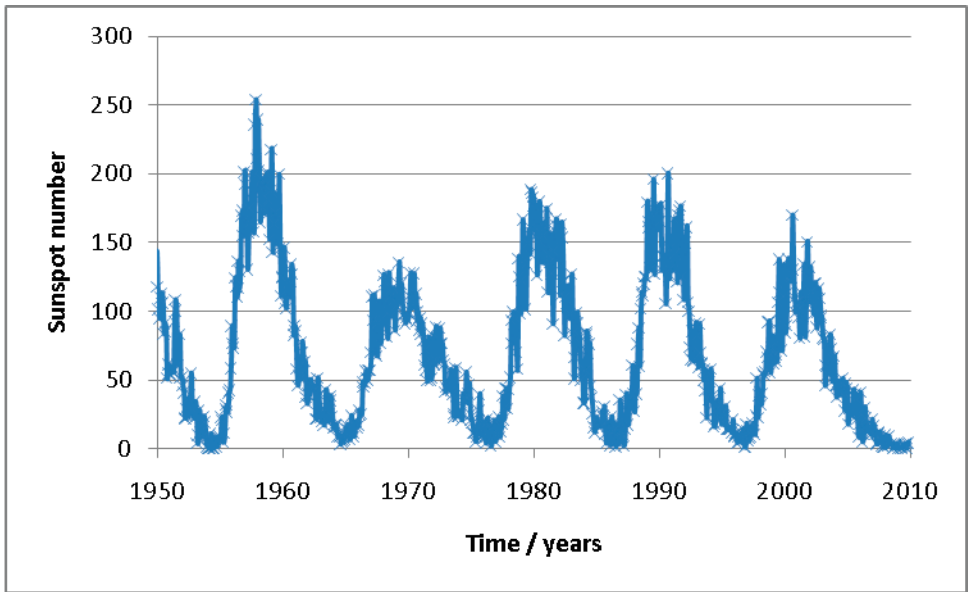


Fig. 1 Variation in solar activity (monthly average sunspot number) since 1950. Data from NASA (http://solarscience.msfc.nasa.gov/greenwch/spot_num.txt, accessed 10 Jan 2010)

The ionosphere has different concentrations of electron density at different heights which are known as the D, E, F1 and F2 layers (see Fig. 2). These layers are associated with ionisation of particular molecules or atoms by certain frequencies of high energy solar radiation.

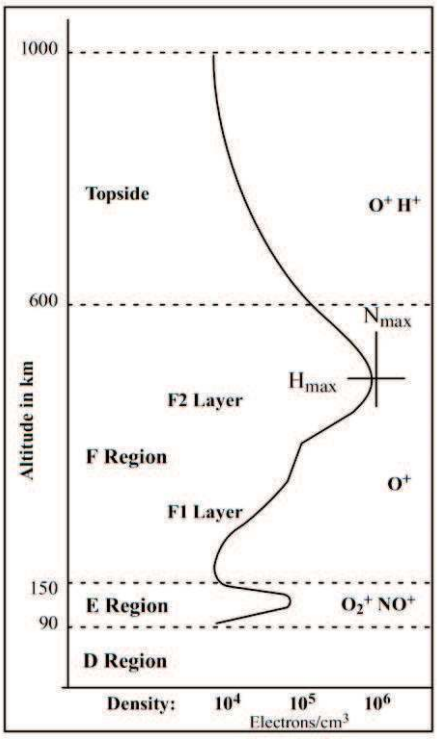


Fig. 2 Layers in the ionosphere. Reproduced from Anderson and Fuller-Rowell (1999). Note that the height boundaries between the ionospheric layers are not generally agreed; the boundaries shown above differ slightly from those suggested by Klobuchar (1996), given in the text

There is also usually a ~27 day periodicity associated with solar rotation (see e.g. Liang et al., 2008; Hocke, 2008; Min et al., 2009), together with seasonal variations in the activity of the ionosphere at different latitudes and heights (e.g. Liu et al., 2009). For more information, see e.g. Rishbeth (2003: Chapter 15) or Davies (1990: Chapter 2).

Klobuchar (1996) provides a summary of the layers existing in the ionosphere and their significance to GPS signal propagation, listing the layers in order of increasing height as follows:

D region, approx 50-90 km. No measureable effect on GPS frequencies.

E region, approx. 90-140km. Minimal effect on GPS frequencies in normal conditions.

F1 region, approx 140-210km

> "The normal F1 region, combined with the E region, can account for up to 10% of the ionospheric time delay encountered by GPS. Diffusion is not important at F1 region heights, and, as with the normal E region, it has a highly predictable density from known solar emissions."

F2 Region, approx. 210-1000km.

> "The F2 region is the most dense and also has the highest variability, causing most of the potential effects on GPS receiving systems. The height of the peak of the electron density of the F2 region generally varies from 250 to 400km, but it can be even higher or somewhat lower under extreme conditions [such as at low latitude, night time or during geomagnetic storms]. … The F2, and to some extent the F1 regions cause most of the problems for radiowave propagation at GPS frequencies."

H+ > 1000km. The protonosphere or plasmasphere (actually a torus rather than a sphere). Low electron density, but extends out to approximately the orbital height of GPS satellites.

> "Estimates of the contribution of the protonosphere vary from 10% of the total ionospheric delay during the daytime hours, when the density of the F2 region is highest, to approximately 50% during the nighttime, when the F2 region density is low."

In summary, the F2 layer is responsible for most of the effect on GPS signals although the contribution of the layer above 1000km, which Klobuchar (1996) terms the protonosphere, may be proportionally higher during the local night. Studies using data from satellites at different elevations can help quantify their respective contributions (e.g. Ciraolo and Spalla, 1997) with further discussion of the effects of the layer above 1000km in e.g. (Mazzella, 2009). For more detailed information on the ionosphere, see e.g. Schunck and Nagy (2009), Kelley (2009), or Kamide and Chian (2007).

2.2. Modelling the ionosphere

In the simplest ionospheric models, the ionosphere is often considered as one or more Chapman layers. The formula for a simple Chapman layer (in a planar, horizontally stratified, isothermal ionosphere, under hydrostatic equilibrium and neglecting the detailed radiative transfer of photons through neutral gas, see for instance Kivelson and Russell, 1995: p183-188) is:

$$\text{Electron density, } N \propto \exp\left(1 - z - \exp\left(-z\right)\right) \quad \text{where} \quad z = \frac{\left(h - h_{\max}\right)}{H_{Ch}} \qquad (1)$$

where:

h is height in the atmosphere, $h_{\max}$ the height of the peak electron density, and H_{Ch} the Chapman scale height (Hartmann and Leitinger, 1984: p124; see also Budden, 1985: p9). As the F2 layer is the most dense, the ionospheric peak electron density is typically written as h_mF_2 and the height of the peak electron density as N_mF_2.

Examples of the form of a Chapman layer with varying Chapman scale heights can be seen in Fig. 3.

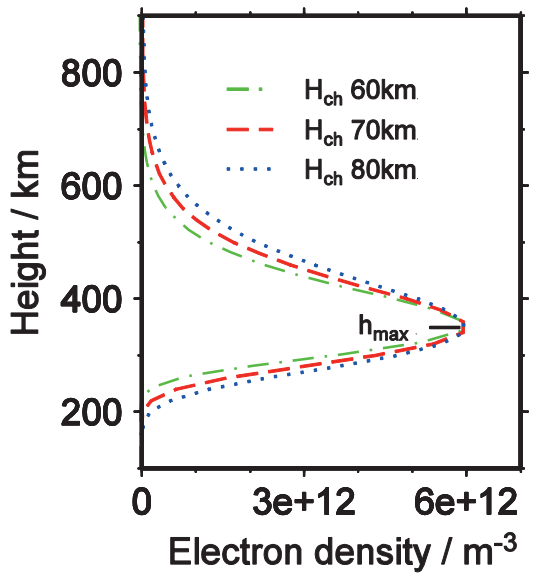


Fig. 3 Examples of Chapman layers with h_{max} = 350 km, peak electron density, N_m = 6 x 10^{12} electrons m^{-3}, and Chapman scale heights H_{Ch} as shown

Chapman layers can also be combined by using different scale heights above and below h_{max}, as was done in the study by Brunner and Gu (1991). In addition to simple representations of the ionosphere such as Chapman layers, more complicated models also exist. Cander (2008) lists some of the many ionospheric models in existence and summarises: "Empirical models, like the global IRI model, the NeQuick model, the European region COSTPRIME, and COSTPROF family of models, attempt to extract systematic ionospheric variations from past data records … Physics-based theoretical and/or numerical models attempt to solve a set of first-principles equations for the ionospheric plasma, starting from the continuity, energy, and momentum equations for electrons and ions. … Parametric models simplify the theoretical models by expressing them in terms of solar-terrestrial parameters and geographical locations, giving a realistic representation of the ionospheric spatial and temporal structure using a limited number of numerical coefficients".

There are also models that combine theory and data, such as the global assimilative ionospheric model (GAIM) (Wang et al., 2004) that combines first principles physics based models with data assimilation. For the future, a discussion of ionospheric imaging using tomography can be found in Bust and Mitchell (2008).

However, while the GPS signal is affected by interaction with the electron content of the ionosphere, it is not just the electron content that affects it, but also the presence of the Earth's geomagnetic field. The next section discusses the theoretical basis for modelling how the signals travel through such a medium.

3. Refractive index for GPS signals transiting the ionosphere

The mathematics describing how electromagnetic waves, such as GPS signals, travel through a partially ionised medium in the presence of a magnetic field has been a subject of interest for decades. While there are more complex versions, the theory that is typically applied for GPS frequencies is termed the magneto-ionic theory. The applications of this theory to GPS will be reviewed here.

The refractive index of a medium defines the phase speed in the material relative to the speed in a vacuum. For a magneto-ionic medium such as the ionosphere, the refractive index depends upon several factors, as this section will demonstrate. The factors include the frequency of the signal, its polarisation, and the angle between the signal and the magnetic field. As will become clear, the decision to transmit GPS signals with circular polarisation reduces the ionospheric effects on the signal.

A helpful in-depth treatment of the magneto-ionic theory can be found in Yeh and Liu (1972), with a simpler overview in Davies (1990). There is an excellent early monograph by Ratcliffe (1959) that describes the previous development of the theory as well as the details of its derivation and some implications. A later work by Budden (1985) also discusses the theory and derivation in some depth, and there are several other works that deal with the subject (e.g. Rawer and Suchy, 1967; Papas, 1965). The different works define coordinate systems and some conventions differently which can be a source of confusion; in the following description, the conventions in Yeh and Liu (1972) are generally followed.

3.1. Defining the ionosphere as a magneto-ionic medium

A magneto-ionic medium is one in which "free electrons and heavy positive ions are situated in a uniform magnetic field and are distributed with statistical uniformity, so that there is no resultant space charge." Ratcliffe (1959: p8). A collection of charged particles in a magnetic field may also be known as a magneto-plasma, and has a characteristic frequency of oscillation, known as the plasma frequency, f_p. Considering electrons only (as all ions are so massive in comparison that their motions are negligible), f_p is related to the electron charge, e, the electron density, N_e, the electron mass, m_e, and the permittivity of free space, ε_0 as follows:

$$f_p^{\,2} = \frac{e^2 N_e}{4\pi^2 \varepsilon_0 m_e}, \qquad (2)$$

or, if the constants are evaluated (Table 1), $f_p^{\,2} = A_p N_e$ where $A_p = 80.62\,\mathrm{m^3 s^{-2}}$.

However, the presence of a magnetic field also influences the magneto-plasma. Charged particles tend to spiral around the lines of magnetic force at a characteristic frequency known as the gyrofrequency, f_g. For electrons, the gyrofrequency depends on m_e, e, and the magnetic field vector, **B**, in the following relationship:

$$\boldsymbol{f}_g = -\frac{e}{2\pi m_e}\mathbf{B} \qquad (3)$$

or

$\boldsymbol{f}_g = A_g\mathbf{B}$, where $A_g = 2.80\text{x}10^{10}$ sA/kg if the constants are combined.

Table 1 Values of physical constants

Symbol	Quantity	Value	Unit
ε_0	permittivity of free space	8.854x10^{-12}	farad/metre $farad : \frac{s^4A^2}{m^2kg}$,
m_e	electron mass	9.107x10^{-31}	kg
e	electron charge	-1.602x10^{-19}	Coulomb *Coulomb: A s*
μ_0	permeability in a vacuum	12.57x10^{-7}	Henry/m *Henry :* $\frac{m^2kg}{s^2A^2}$

When defined as above (Eqn. 2), the plasma frequency is always a positive scalar. However, the gyrofrequency (Eqn. 3) is a vector, representing the direction around which charged particles will orbit in a clockwise manner (and may thus be negative for particles with a positive charge). Klobuchar (1996: p488) states that "the electron gyro frequency … is typically 1.5 MHz; the plasma frequency … rarely exceeds 20 MHz". This translates to a magnetic field strength of ~53500nT (nanoTesla) and electron density of ~ 4.96 x 10^{12} electrons m^{-3}.

To allow consideration of the effects on signals travelling through a magneto-plasma, it is necessary to set out a consistent set of conventions. We can define a coordinate system with orthogonal axes such that the signal propagation direction **k** is along the z-axis, with the magnetic field vector **B** in the z-y plane and where θ is the angle between the **B** and **k** vectors (see Fig. 4).

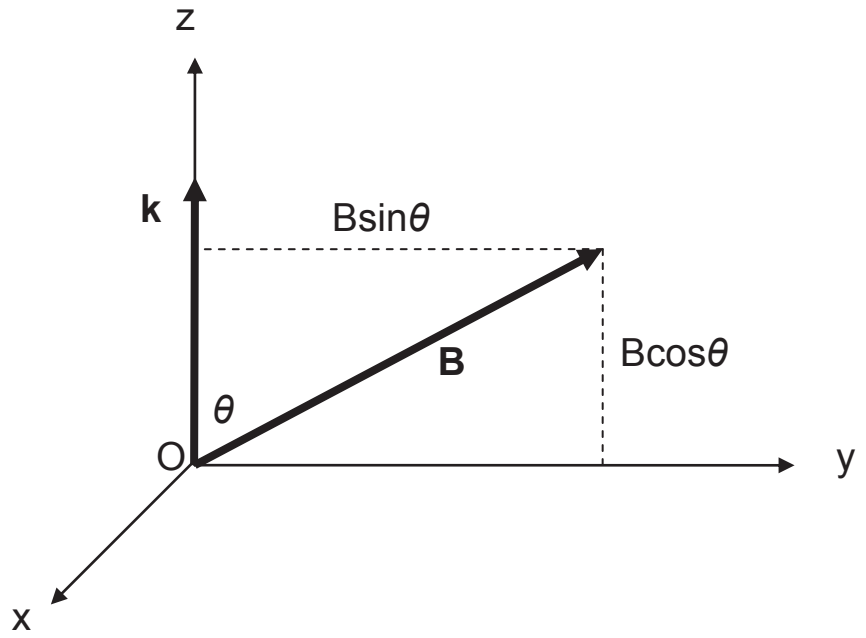


Fig. 4 Coordinate system with signal vector **k** and magnetic vector **B**

It is also worth noting that there is no common convention in the literature (e.g. Ratcliffe, 1959; Yeh and Liu, 1972; Budden, 1985; Davies, 1990) regarding the assignment of the magnetic field vector and signal vector to particular axes orientations, which complicates the process of comparison. Where formulae from different authors vary, it can usually be shown to be due to the different conventions that have been used in the derivation. Having defined the magneto-ionic medium and coordinate system, we now consider the properties of the electromagnetic signal passing through the medium.

3.2. Defining the GPS signal and its polarisation

An electromagnetic signal such as a GPS signal consists of linked time-varying electric and magnetic fields propagating through a medium. It will have a certain polarisation, which can also affect its propagation. A formal definition of the polarisation of an electromagnetic wave is "that property of a wave describing the time-varying direction and relative magnitude of the electric field vector; specifically the figure traced in time by the extremity of the vector at a fixed location in space, and the sense in which it is traced, as observed along the direction of propagation." (Balanis, 2005: p71). The polarisation of signals emitted by an antenna may be linear, circular or ellipsoidal, depending upon the antenna design (see, e.g., the excellent summary by Balanis (2005: p74) for further information).

As set out in the technical document ICD-GPS-200 (2000), GPS signals are effectively right-hand circularly polarised (RHCP), with RHCP defined as the electric vector of the wave being of constant amplitude and rotating clockwise *when viewed from the source of the signal* (see Fig. 5a,d). An electric vector of constant amplitude will appear to trace a circle in a plane perpendicular to the signal as it rotates. While this is the usual definition of RHCP according to the Institute of Electrical and Electronics Engineers (IEEE), it should be noted that in optics the traditional definition is completely the opposite, i.e. RHCP is clockwise motion as viewed by an observer facing the oncoming wave (Born and Wolf, 1999).

So, how can the electric field vector of such a wave be represented mathematically? If we assume that the wave is of a simple harmonic form with RHCP polarization, then using the coordinate system set out above, the two components of the electric vector along the axes perpendicular to the line of travel are E_x and E_y. For the electric vector to be of constant amplitude, a, we need $E_x^{\,2} + E_y^{\,2} = a^2$. To achieve this, the amplitudes of the two components must be the same, and the two components must be out of phase in time by a multiple of $\pi/2$ (see Fig. 5c,d). For the electric vector to rotate clockwise as specified above, the E_y component must lag the E_x component. A simple harmonic wave with clockwise (RCHP) rotation may then be represented mathematically as follows:

$$E_x = a_x \cos(\omega t - kz + \phi)$$
$$E_y = a_y \sin(\omega t - kz + \phi) = a_y \cos\left(\omega t - kz + \phi - \frac{\pi}{2}\right) \qquad (4)$$
$$\text{and } a_x = a_y = a$$

where angular rotation, ω, equals $2\pi f$ (Hz), and angular wave number, k, equals $2\pi/\lambda$ with f the frequency, λ the wavelength of the wave and ϕ is the initial phase. The components of the wave vary in time through the factor ωt and in space along the direction of propagation (z-axis) through kz. Fig. 5a illustrates how the tip of the electric vector rotates clockwise with time

at one point in space, as described by the IEEE definition. In contrast, Fig. 5b shows what the electric vector looks like in space at a fixed instant of time. In space the wave becomes a left hand helix, which can lead to misunderstanding of the RHCP definition. Fig. 5c is a graphical representation of the phase lag between the E_x and E_y components needed to produce a RHCP wave. Fig. 5d has three snapshots of E_x and E_y with time, showing how the electric vector rotates in a clockwise manner with time, that is from the positive x-axis towards the positive y-axis in our coordinate system.

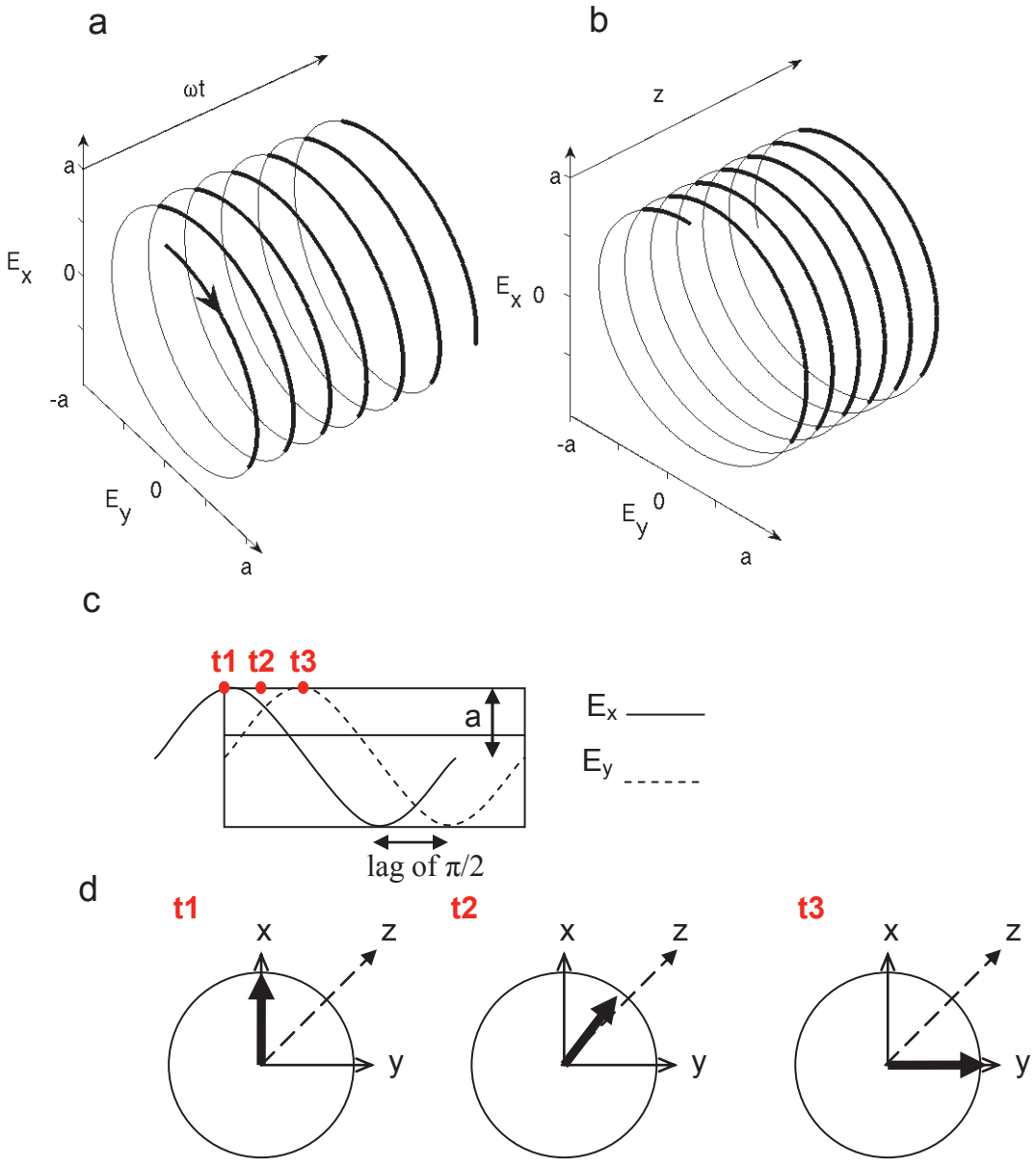


Fig. 5 RHCP wave (IEEE convention). a) RHCP wave demonstrating change with time at a location in space. b) RHCP wave demonstrating the shape in space at one instant of time. c) Graphical representation of an RHCP wave showing how the x and y components of the wave, E_x and E_y, vary with time (see Eqn. 4). d) Rotation of electric vector of an RHCP wave with time. Direction of wave travel is along z, into the page

Polarisation is most often described mathematically by using complex number notation. Using this notation, the components in the x and y directions of electric field of a GPS transmission set out above can be expressed as (with the complex number $i = \sqrt{-1}$):

$$\widetilde{E}_x = ae^{i(\omega t - kz + \phi)}, \text{ and } E_x = \mathrm{Re}\left[\widetilde{E}_x\right]$$
$$\widetilde{E}_y = ae^{i\left(\omega t - kz + \phi - \frac{\pi}{2}\right)}, \text{ and } E_y = \mathrm{Re}\left[\widetilde{E}_y\right] \quad (5)$$

Ratcliffe (1959) shows the magnetic vectors, H_x, H_y are related to the electric vectors and that, when using complex notation, the polarisation of the wave can be represented as follows (see also Yeh and Liu(1972: p163)):

$$R = -\frac{\widetilde{H}_y}{\widetilde{H}_x} = \frac{\widetilde{E}_x}{\widetilde{E}_y} \quad (6)$$

Using Eqn. 6, above, for the polarisation, then for a RHCP wave in the IEEE convention, such as a GPS signal, the polarisation can be described in complex notation as follows:

$$\frac{\widetilde{E}_x}{\widetilde{E}_y} = e^{i\pi/2} = i \quad (7)$$

(while for a LHCP wave, the polarisation would be $\frac{\widetilde{E}_x}{\widetilde{E}_y} = e^{-i\pi/2} = -i$).

For further information see Born and Wolf (1999: p18, 29-31) or for details of complex representation see Hecht (1998: p23-28), though convention differences should be noted. For example, both books use the optics convention rather than the IEEE convention for RHCP and Hecht (1998) defines harmonic waves using (*kz-ωt*) rather than (*ωt-kz*). The equations above are consistent with the conventions in both Ratcliffe (1959) and Yeh and Liu (1972).

3.3. Waves passing through a magneto-ionic medium

An expression for the refractive index of a magneto-plasma can be derived using Maxwell's equations relating electric and magnetic fields to electric charges and currents and the constitutive relations describing how the ionospheric plasma responds to electromagnetic waves. Following the conventions set out above in Fig. 4, the resulting expression for the phase refractive index is:

$$n_{\pm}^2 = 1 - \frac{X}{1 - \frac{Y^2 \sin^2\theta}{2(1-X)} \pm \sqrt{\left(\frac{Y^4 \sin^4\theta}{4(1-X)^2} + Y^2\cos^2\theta\right)}} \quad (8)$$

where X is the normalised plasma frequency, $X = \frac{f_p^2}{f^2} = A_p \frac{N_e}{f^2}$, (see Eqn. 2), and Y is the modulus of the normalised gyrofrequency, i.e. $Y = |\mathbf{Y}| = A_g \frac{|\mathbf{B}|}{f}$ (see Eqn. 3). The expression is generally known as the Appleton-Lassen equation or Appleton-Hartree equation. Though "Appleton-Hartree" is frequently used,

"Appleton-Lassen" would be the preferred nomenclature, as Lassen (1927) independently derived the same formula as Appleton (1932), while Hartree (1931) derived an alternative version of the formula.

The derivation includes a linked formula for the polarisation of waves that can travel through a medium with the given refractive index with their polarisation unchanged (see e.g. Ratcliffe, 1959). Such waves are called characteristic waves, and the linked expression giving their polarisations is:

$$R_{\mp} = \frac{i}{\cos\theta}\left[\frac{Y\sin^2\theta}{2(1-X)} \mp \sqrt{\left(\frac{Y^2\sin^4\theta}{4(1-X)^2} + \cos^2\theta\right)}\right] = \frac{i}{Y\cos\theta}\left(1 - \frac{X}{1-n_{\pm}^2}\right) \qquad (9)$$

R_+, the lower sign in Eqn. 9, is linked to the lower sign of the refractive index formula, n_-, and similarly for the upper signs R_- and n_+.

The equations above for the polarisation and the refractive index have been simplified to exclude the effects of electron collisions (which are generally considered negligible at the frequencies used by GPS). Additional assumptions are necessary for the derivation, including that:

- positive ions are infinitely massive so do not affect the wave;
- the plasma is a 'cold' plasma (i.e. the thermal motions of the electrons are much smaller than the phase velocity of the wave) and
- there is a uniform magnetic field

(see e.g. Davies (1990: p70) for a summary or Yeh and Liu (1972) for further information).

3.4. Evaluating the polarisation formula for GPS signals

The two polarisations of Eqn. 9, R_+ and R_-, thus represent the two possible characteristic waves of the magneto-plasma. Waves that are not one of the characteristic waves are separated into components with the characteristic polarisations. These components may then travel at different speeds through the medium before recombining (which is the origin of Faraday rotation for a linearly polarised wave).

However, when Eqn. 9 is evaluated for the GPS L1 and L2 frequencies, the result is a polarisation of i (or RHCP, see Eqn. 7) using the lower sign (R_+) when $0 \le \theta < \pi/2$, and the upper sign (R_-) when $\pi/2 < \theta \le \pi$. This is true for all values of θ except ~ 89.5-90.5 degrees (see Fig. 6). This is excellent as it implies that for almost all values of θ, GPS signals are characteristic waves.

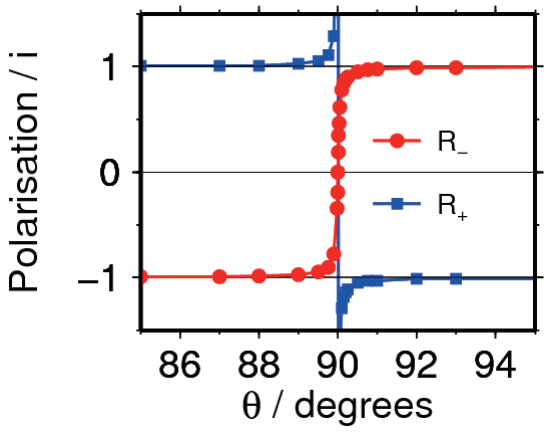


Fig. 1 Polarisations of the characteristic waves at the GPS L1 frequency (multiply y-axis by i) showing changes when θ is close to 90 degrees. Obtained by evaluating Eqn. 9.

For values of θ very close to 90 degrees, the characteristic wave polarisations are no longer i and $-i$ and so no longer RHCP and LHCP (see Fig. 6). This was also recognised by Hartmann and Leitinger (1984) who state:

> 'For frequencies above 100MHz the "characteristic polarisations" are, to a very good approximation, the right-hand and the left-hand circular polarisations except for propagation perpendicular to the vector of geomagnetic field strength (and deviations of some tenths of a degree from this direction). The latter exception has no practical implications'.

This is a very useful result, as it implies that the RHCP GPS signals can travel with unchanged polarisation for almost all values of θ.

3.5. Simplified refractive index for GPS signals

We have seen that for all practical intents and purposes, GPS signals will be one of the two characteristic waves. However, to predict the ionospheric effects on the characteristic wave/GPS signal as it travels, we need the refractive index of the medium. The expression for the refractive index in Eqn. 8 is fairly complex to evaluate, and is usually simplified.

For GPS frequencies, X is generally less than 2.7×10^{-4} and Y is less than 1.3×10^{-3} (using an approximate maximum value for plasma frequency and typical value for gyrofrequency of 20MHz and 1.5 MHz respectively (Klobuchar, 1996), and the L2 frequency, 1227.6MHz).

On rearranging Eqn. 8, then expanding it as a binomial series, discarding terms smaller than 1 part in 10^{-9} as suggested by Brunner and Gu (1991), taking the square root and again expanding and discarding, the following simplified expression for the Appleton-Lassen equation is obtained:

$$n_{\pm} = 1 - \frac{1}{2}X \pm \frac{X|Y\cos\theta|}{2} - \frac{1}{8}X^2 . \qquad (10)$$

Other derivations leading to essentially the same result are summarised in the appendix to the excellent paper by Datta-Barua et al. (2006). The paper also examines the size of the largest discarded terms and confirms that their effects are sub-millimetre.

Now, for the formula to be correct for the GPS RHCP polarisation the lower sign must be used when $0 \le \theta < \pi/2$, and the upper sign when $\pi/2 < \theta \le \pi$. Eqn. 10 is then equivalent to:

$$n = 1 - \frac{1}{2}X - \frac{XY\cos\theta}{2} - \frac{1}{8}X^2 \qquad (11a)$$

or equivalently to:

$$n = 1 - \frac{1}{2}\frac{A_p N_e}{f^2} - \frac{A_p N_e A_g B\cos\theta}{2f^3} - \frac{1}{8}\frac{A_p^2 N_e^2}{f^4} \qquad (11b)$$

on using the notation of Eqns. 2 and 3.

In the above expanded and simplified equation for the refractive index of the ionosphere, the term in f^2 becomes the first order term, I1. On evaluation (from Eqn. 2, Ap = 80.6), it reduces to the more familiar $40.3N_e/f^2$. The second order (I2) and third order (I3) refractive index terms then arise from those in f^3 and f^4 respectively.

3.6. Comparison with previous literature

Linking the derivation above with the previous literature reveals inconsistencies in a few places, but generally only in the justification for the formula used rather than in the end results.

Hartmann and Leitinger (1984: p122) give a formula for the refractive index as follows:

$$n_{l,r} = 1 - \frac{1}{2}\frac{f_p^2}{f^2} \pm \frac{f_p^2 f_g |\cos\theta|}{f^3} - \frac{1}{8}\frac{f_p^4}{f^4} - \frac{f_p^2 f_g^2}{4f^4}\left(\sin^2\theta + 2\cos^2\theta\right) \pm \ldots$$

This formula would agree with the result of the derivation above (Eqn. 11b), with the exception of a missing ½ in the third term which may be a misprint. The equivalent of the fifth term shown here, which Hartmann and Leitinger call ΔS_c, has been discarded in Eqn. 11b above due to its size. The authors also state "It is clearly seen that ΔS_c can be neglected against ΔS_a [I2] and ΔS_b [I3]" Hartmann and Leitinger (1984: p124).

In the subsequent paper by Brunner and Gu (1991) the following statement is made and attributed to Hartmann and Leitinger (1984): 'The ionosphere is a birefringent medium, and thus causes the EM wave to split into two propagation modes, the ordinary and the extra-ordinary mode. The ordinary mode corresponds to the upper (+) sign in Eq. (A.1.) [the Appleton-Hartree formula] and is left-hand circularly polarized, while the extra-ordinary mode corresponds to the lower (-) sign in Eq. (A.1) and is right-hand circularly polarised, *Hartmann and Leitinger [1984]*'. The actual sentence in Hartmann and Leitinger (1984) limits the link between the lower (-) sign and RHCP to the geomagnetic north hemisphere; it reads 'In the geomagnetic north hemisphere the + sign corresponds to the left-hand circular component, the – sign to the right-hand circular component'. This is because, as seen in section 3.5, the appropriate sign depends on the angle between the signal and the magnetic field. While this may seem a somewhat pedantic point to make, it is from this point onwards that some confusion about which sign should be used and why seems to spread in the later literature.

The assertion of Brunner and Gu (1991) that the lower sign in the Appleton-Hartree formula always represents a right-hand circularly polarised wave would seem to apply only to the northern hemisphere. However, if formula A.5 of Brunner and Gu (1991: p214), $n = 1 - \frac{X}{2} - \frac{X^2}{8} \pm \frac{XY}{2}\cos\theta$, is taken and the lower sign is used as they suggest, it is equivalent to that derived above (Eqn. 11a).

Brunner and Gu also suggest that because they derive their formula with no assumptions about the angle θ, it is valid for all values of θ. While the refractive index formula is valid, it should be considered together with the polarisation formula which describes the form of the two characteristic waves. As mentioned in section 3.4, between ~89.5 and ~90.5 degrees neither of the characteristic waves is circularly polarised (see Fig. 6) so the formula does not properly represent (RHCP) GPS signals for a small (and probably insignificant) range of values of θ.

Bassiri and Hajj (1993: p282) give a refractive index expansion up to the fourth inverse power of frequency: $n_{\pm} = 1 - \frac{1}{2}X \pm \frac{1}{2}XY|\cos\theta| - \frac{1}{4}X\left[\frac{1}{2}X + Y^2\left(1-\cos^2\theta\right)\right]$. This simplifies and rearranges to be identical with that derived here (Eqn. 11a). However, when they apply the formula, the assumption is again made that the lower sign represents the extra-ordinary wave and is RHCP. This means that, referring to the derivation above, this formula is only appropriate in certain conditions ($0 \le \theta < \pi/2$, approximately the northern hemisphere). Bassiri and Hajj (1993) consider the issue of the effects on the refractive index formula when θ is near 90 degrees, but again do not mention the linked change in polarisation of the characteristic waves.

Kedar et al. (2003) give a formula for the error in length of the ray path which is corrected from Bassiri and Hajj (1993). It equates to the formula for the refractive index derived above (see Section 4.1). Subsequent papers (e.g. Fritsche et al. (2005), Hernandez-Pajares et al. (2007), Petrie et al. (2010b)) continue to use this formulation, which is consistent with the one derived above (Eqn. 11a) .

4. Correction to GPS signal from ionospheric refractive index expansion terms

4.1. Expressions for phase advance and code delay

Integrating the effect of the refractive index of Eqn. 11b along the curved path leads to the following equation for the ionospheric effects on the carrier phase in units of length:

$$\Phi_1 = \rho + N_1\lambda_1 - \frac{1}{2}\frac{A_p\int N_e dL}{f_1^2} - \frac{1}{2}\frac{A_pA_g\int N_eB\cos\theta dL}{f_1^3} - \frac{1}{8}\frac{A_p^2\int N_e^2 dL}{f_1^4}, \qquad (12)$$

where ρ is the geometric range between the satellite and the receiver, λ the wavelength and dL denotes integrals along the path. In Eqn. 12 and subsequently we denote N_1 to be the integer ambiguity. The subscript here refers to the GPS signal with frequency f_1, often known as L1. An analogous formula holds for the GPS frequency f_2. For the sake of simplicity, Eqn. 12 does not show corrections for non-dispersive errors, such as tropospheric delay, or for the effects of phase wind-up.

Once the values for the constants in A_p and A_g are inserted, the effect on the carrier phase becomes (with units of metres):

$$\Phi_1 = \rho + N_1\lambda_1 - \frac{40.3\int N_e dL}{f_1^2} - \frac{1.1284\times10^{12}\int N_eB\cos\theta dL}{f_1^3} - \frac{812.0\int N_e^2 dL}{f_1^4} \qquad (13)$$

and we can define the first, second and third order refractive index correction terms as:

$$\mathrm{I1} = -\frac{40.3\int N_e dL}{f_1^2},\ \mathrm{I2} = -\frac{1.1284\times10^{12}\int N_eB\cos\theta dL}{f_1^3} \text{ and } \mathrm{I3} = -\frac{812.47\int N_e^2 dL}{f_1^4} \qquad (14)$$

These corrections are for the GPS carrier phase. To implement the corrections in a practical GPS analysis, approximations of the integrals are usually performed to reduce computing time and avoid the need for a profile of N_e with height. Typical approximations are discussed in Section 6.

For the modulated GPS code, the following information is useful:

"A signal, or modulated carrier wave, can be considered to result from the superposition of a group of waves of different frequencies centred on the carrier frequency. If the medium is dispersive, the modulation of the signal will propagate with a different speed from that of the carrier; this is called the group velocity. Corresponding to the phase refractive index, n, we can define a group refractive index, n_g, where: $n_g = n + f\frac{dn}{df}$." (Langley, 1998: p124).

The equivalent formula for the refractive index for the modulated code can then be derived from the formula for the phase:

$$n_g = 1 + \frac{1}{2}\frac{A_p N_e}{f^2} + 2\left(\frac{A_p A_g N_e B\cos\theta}{2f^3}\right) + 3\left(\frac{1}{8}\frac{A_p^2 N_e^2}{f^4}\right) \qquad (15)$$

and the expression for the ionospheric effect on the pseudorange follows,

$$P_1 = \rho + \frac{40.3\int N_e dL}{f_1^2} + 2\frac{1.1284\times10^{12}\int N_e B\cos\theta dL}{f_1^3} + 3\frac{812.47\int N_e^2 dL}{f_1^4} \qquad (16a)$$

which becomes

$$P_1 = \rho + \frac{40.3\int N_e dL}{f_1^2} + \frac{7527.87c\int N_e B\cos\theta dL}{f_1^3} + \frac{2437\int N_e^2 dL}{f_1^4} \qquad (16b)$$

where c is the speed of light in a vacuum.

4.2. 'Ionosphere-free' linear combination

The traditional approach to removing the effect of the ionosphere is the 'ionosphere-free' linear combination, often known as the LC or L3 combination:

$$\Phi_{LC} = \frac{f_1^2}{f_1^2 - f_2^2}\Phi_1 - \frac{f_2^2}{f_1^2 - f_2^2}\Phi_2 \qquad (17)$$

where Φ_i is the expression for the phase from Eqn. 13, and again the units are metres.

While this combination cancels the first order term (assuming that $\int N_e dL$ is identical for both the L1 and L2 frequencies), the second and third order terms do not cancel completely.

Defining K_1 and K_2 equal to $A_p A_g \int N_e B\cos\theta dL$ for the L1 and L2 carrier phases respectively, the effect of I2 on the phase LC combination is:

$$\Delta I2 = \frac{K_2 f_1 - K_1 f_2}{2 f_1 f_2 \left(f_1^2 - f_2^2 \right)}$$, which, if K_1=K_2, becomes $$\Delta I2 = \frac{K}{2 f_1 f_2 \left(f_1 + f_2 \right)}$$. (18)

Similarly, if J1 and J2 are defined as equal to $A_p^2 \int N_e^2 dL$ for the L1 and L2 signals, then

the effect of I3 on the phase LC combination is:

$$\Delta I3 = \frac{J_2 f_1^2 - J_1 f_2^2}{8 f_1^2 f_2^2 \left(f_1^2 - f_2^2 \right)}$$, which, if J_1=J_2, becomes $$\Delta I3 = \frac{J}{8 f_1^2 f_2^2}$$. (19)

Alternative linear combinations of the signals are often made for other purposes (e.g. Hofmann-Wellenhof et al., 2001: p93) and should not be confused with the combination described above.

5. Bending

In addition to the higher order refractive index terms there are also bending terms; when a wave crosses a gradient in refractive index in a direction that is not perpendicular to the surfaces of constant refractive index the wave is refracted or bent. This process will be termed geometric bending in this paper. Due to their differing frequencies, the L1 and L2 signals bend slightly differently, so the slant total electron content (STEC) integrated along each signal path is not quite the same. This difference in STEC means that there is also a small residual effect from the first order term even after forming the LC combination (Hoque and Jakowski, 2008b). In this paper we term this residual correction the "TEC difference bending effect" or dTEC correction.

Neither the geometric bending term nor the dTEC bending term are routinely implemented in practical GPS processing and there is no real consensus as to how the bending terms should be computed. Below we outline where the errors arise. We review approaches for estimating the errors, and compare the results from the various formulae that have been suggested for the geometric bending error for a standard case. While after formation of the 'ionosphere-free' linear combination the dTEC bending correction is the larger of the two bending corrections, for a proper illustration of the effects, both corrections should be modelled. This is particularly so because in the residual range error expression for the phase combination the geometric bending correction is of the opposite sign to the other terms in the expression. Omitting the geometric bending correction may thus make the impact of the bending terms appear larger than it actually is. The first study on applying bending terms to GPS data when determining coordinates is Petrie et al. (2010a, (see Section 7).

Finally, all of the following discussion is based upon a uniform ionosphere, either flat or spherically symmetric. Horizontal gradients do exist and may be fairly abrupt (e.g. Doherty et al., 2004; Stankov et al., 2009) but a method for dealing with such gradients for GPS analysis has not yet been described. Finally, such gradients are not necessarily well characterised on a day to day basis.

5.1. Geometric bending term

The geometric bending term, b_{gm}, is also known as the curvature term or refraction error (Leitinger and Putz, 1988). There are several papers that have considered the concept at the signal level and we review their

approaches below. As each paper uses different ionospheric parameters, we both review the original results and attempt to produce equivalent results for a standard case, for an improved comparison.

As a foundation for the discussion, consider the underlying geometry of the situation. The geometric bending term is the difference between the geometric range (the straight line shown in Fig. 7) and the range when signal bending due to the ionospheric refractive index is taken into account (curved line in Fig. 7).

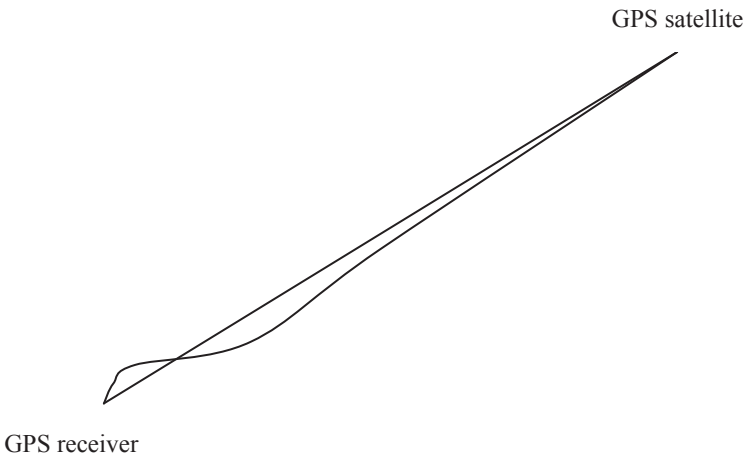


Fig. 7 Difference between curved signal path and geometric straight line range (not to scale)

In their classic paper on range errors, Hartmann and Leitinger (1984) considered bending of the signal path as part of the residual ionospheric errors. They indicated that: "Model calculations have shown that the resulting additional range error ΔS_d [b_{gm}] is approximately proportional to ($1/f^4$)" and stated that by means of series expansions one can show that the following formula is a good approximation:

$$\Delta S_d \cong c^r \left(\frac{A^2}{f^4}\right) I^2 \left(\frac{y}{\tau} - \frac{1}{h_s}\right) \quad \text{with elevation dependence,} \quad c^r = \frac{\tan^2 \chi}{8 \cos \chi} \qquad (20a)$$

or, using the symbols defined in this paper:

$$b_{gm} \cong c^r \left(\frac{A_p^2}{f^4}\right) VTEC^2 \left(\frac{\eta}{VTEC / N_m F_2} - \frac{1}{h_s}\right) \text{with elevation dependence,} \quad c^r = \frac{\tan^2 z'}{8 \cos z'} \qquad (20b)$$

where b_{gm} is the geometric bending term in m; z' (or χ) the zenith angle of the ray at the mean ionospheric height; η (or y) the shape factor; VTEC/N_mF_2 (or τ) represents the ionospheric slab thickness and VTEC is vertical total electron content, h_s is the height of the satellite transmitter and A_p represents physical constants that evaluate to 80.6 (see section 3.1).

The derivation of Eqn. 20a for both planar and spherical layering can be found in Leitinger and Putz (1988). While they indicate in the derivation that using $1/h_s$ (as in Eqn. 20a) may be an approximation most suitable for $h_s \sim 1000$km, their alternative approximations deviate by less than 1mm at an elevation angle of five degrees. As always the sign of the correction should be defined to avoid confusion. For Leitinger and Putz (1988)

'propagation "error" means length of optical path, S_n, minus true (geometric) range, S_O'. One complication in the derivation is that χ is the zenith angle of the ray at the mean ionospheric height rather than the zenith angle of the geometric straight line. However, this difference in angle is very small even in fairly extreme conditions. Calculations using GPS data from the SAC-C satellite at solar maximum, equatorial noon conditions, and occultation scenarios show the typical top value between the transmitter and receiver tangent directions is less than 1 thousandth of a radian, which implies it can be neglected. SAC-C data is available at http://www.gsfc.nasa.gov/gsfc/service/gallery/fact_sheets/spacesci/sac-c.htm.

Hartmann and Leitinger (1984) considered the potential size of this range error, or geometric bending error, for a worst case scenario for several different frequencies. They consider their formula valid for elevation angles greater than 30 degrees and suggest it may need adjusting by using a lower mean ionospheric height at lower angles. We reproduce the part of their Table 6 that is most relevant to GPS as Table 2 below, and also add the equivalent numbers for GPS frequencies.

Table 2 Worst case predictions for b_{gm} after Hartmann and Leitinger (1984), extended to GPS frequencies

Frequency	Elevation angle /deg	b_{gm} ('ΔS_d') /m	Parameters
400MHz	60 45 30	-0.2 -0.5 -1.7	Shape factor = 1 Ionospheric slab 200km thick Mean ionospheric height 400km Transmitter height 1000km
2GHz	60 45 30	-3 x 10-4 -1 x 10-3 -3 x 10-3	
1575.4 MHz GPS L1	**60 45 30**	**-5.5 x 10-4 -1.8 x 10-3 -8.9 x 10-3**	As above, but transmitter height 20200 km If L1 and L2 b_{gm} estimates are combined to form LC:
1227.6 MHz GPS L2	**60 45 30**	**-2.28 x10-3 -7.53 x 10-3 -24.14 x 10-3**	1.38 mm 4.57 mm 14.66 mm

Fig. 8 plots b_{gm} for the GPS frequencies, using the Hartmann and Leitinger (1984) formula, for two scenarios. The first scenario uses the conditions they suggest in their paper, the second is conditions selected by the authors to allow a 'standard' comparison between the various formulae.

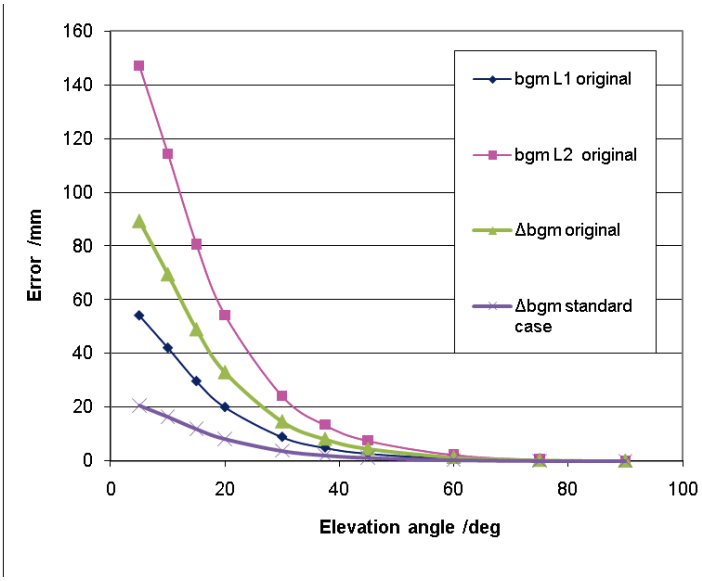


Fig. 8 Geometric bending errors calculated with the formula of Hartmann and Leitinger (1984). Original is using the conditions from the original paper, standard is equivalent to Case 1 of Brunner and Gu (1991), Δbgm shows the effect on the LC combination

Brunner and Gu (1991) also considered bending effects during their study of higher order ionospheric effects on GPS signals using three dimensional ray tracing calculations. They took a

slightly different approach, examining the curvature correction from the straight path for L1, and adding a small perturbation term to account for the difference between the L1 and L2 paths (the derivation for these terms can be found in Gu and Brunner (1990)). They provided results for the LC combination for two different cases. Case 1 was chosen as VTEC = 138 TECU (1TECU=10^{16}electrons/m^2) and N_mF_2 = 6 x 10^{12} electrons m^{-3}. Case 1 was intended to represent 'high Nm[F_2] values which are frequently observed' and we have selected these parameters as the basis for a standard comparison with equations from other papers. Case 2 was chosen to represent a scenario with extremely high ionospheric activity (VTEC = 455 TECU and N_mF_2 = 20 x 10^{12} electrons m^{-3}). They approximated the satellite altitude as 1000km rather than ~20000 km as they assumed that above 1000km the ionosphere would have little effect on signal propagation.

In addition, in their conclusions they note that their method requires information about the structure of the ionosphere and provide "an empirical formula which gives sufficiently accurate results" for residual range error (RRE) due to geometric bending in metres (shown here with slightly altered notation):

$$\Delta b_{gm} = 4.70\times10^{5}\,\frac{A_p^2}{8f_1^4}\,N_m^2\cos^2\beta\cot\beta\,. \qquad (21)$$

Results are shown in Fig. 9. The large increase from Case 1 to Case 2 is as expected, and the empirical formula appears to give reasonably similar results.

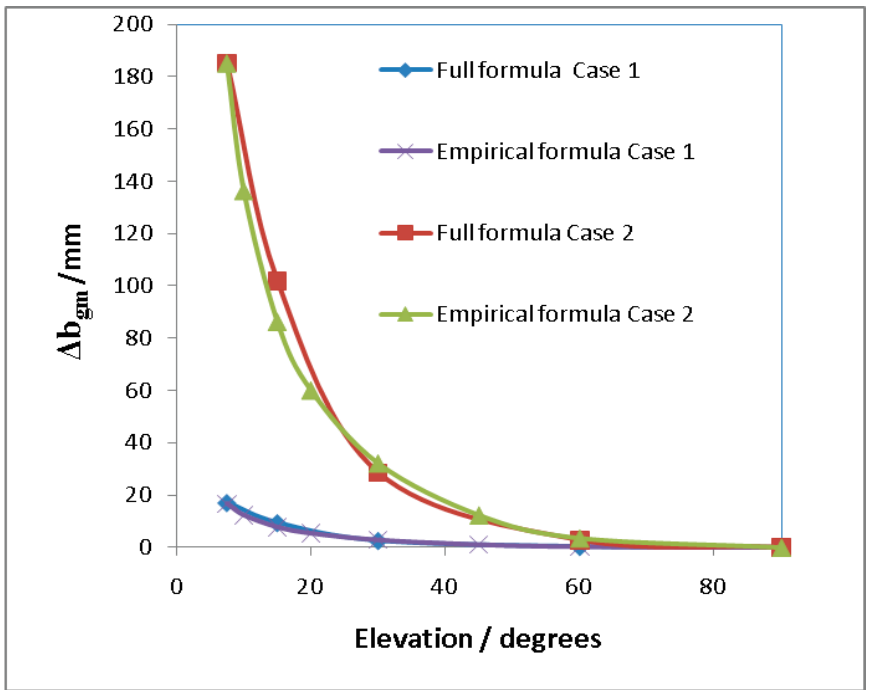


Fig. 9 Bending RRE, Δb_{gm}, after Brunner and Gu (1991) for Case 1, Case 2 and empirical formula (using N_mF_2 as for Case 1, Case 2)

Brunner and Gu (1991) also provide the maximum separation from the straight line path for L1 and L2 signals which can be used to check the separation between the signals (see Table 3).

Table 3 Path separation from straight line path (metres) (Brunner and Gu, 1991)

Elevation angle (degrees)	Case 1 L1	Case 1 L2	Difference	Case2 L1	Case 2 L2	Difference
90	0.0004	0.0007	0.0003	0.0015	0.0024	0.0009
60	4.3190	7.1144	2.7954	14.1979	23.3876	9.1897
30	27.2780	33.4000	6.1220	66.6608	109.8183	43.1575
15	55.5253	91.4643	35.9390	182.5863	300.8688	118.2825
7.5	92.7348	152.7742	60.0394	305.0658	502.9136	197.8478

Jakowski et al. (1994) considered the geometric bending effect and derived an empirical formula for satellites at altitudes of 1000km and 20000km to represent its effect. They investigated several frequencies including those appropriate for GPS. They used a Chapman layer to represent the ionosphere and performed numerical ray tracing on 2000 parameter sets.

Their resulting empirical formula (with coefficients for 20000km) is shown below with slightly altered notation:

$$b_{gm} = \frac{\left[1.319\times10^{8}\cdot STEC\left(foF2\right)^{2} + (-0.105\times10^{5})\left(foF2\right)^{4} hmF2\right]\cdot\left[\left(1-\left(1+\left(-0.678\times10^{-4}\cdot hmF2\right)^{2}\right)\cos^{2}\beta\right)^{-1/2}-1\right]}{f^{4}} \quad \text{(22a)}$$

The units they used are not SI units but are as follows:

Ray path geometric bending error, b_{gm}, in mm, STEC in TECU, critical ionosonde frequency *f*oF2 and signal frequency *f* in MHz, peak height of the F2 layer h_mF_2 in km and elevation angle β in degrees or radians as appropriate.

Substituting into their Eqn. (4) giving the relationship of the peak electron density (N_mF_2) to the critical ionosonde frequency (in SI units), $N_mF_2 = 1.24*10^{-2}(foF2)^2$, the formula becomes:

$$b_{gm} = \frac{\left[1.319\times10^{8}\cdot TEC\frac{NmF2}{1.24\times10^{10}} + (-0.105\times10^{5})\left(\frac{NmF2}{1.24\times10^{10}}\right)^{2} hmF2\right]\left[\left(1-\left(1+\left(-0.678\times10^{-4}\cdot hmF2\right)^{2}\right)\cos^{2}\beta\right)^{-1/2}-1\right]}{f^{4}} \quad \text{(22b)}$$

if peak electron density N_mF_2 is in electrons/m^3 (and converting for foF2 in MHz).

As they recognised that ionospheric parameters such as h_mF_2 and N_mF_2 are often unavailable, they also provided a simplified formula which is only a function of slant TEC (TECU) and elevation angle, again with b_{gm} in mm:

$$b_{gm} = \frac{2.495\times10^{8}\cdot STEC^{2}\left[\left(1-0.8592\cos^{2}\beta\right)^{-1/2}-1\right]}{f^{4}} \quad (23)$$

This formula is in $1/f^4$ and if the 'ionosphere free' linear combination is performed and Eqn. 22a is represented as $b_{gm} = \frac{A}{f^4}$, it reduces to the following residual range error:

$$\begin{aligned}\Delta b_{gm} &= \frac{f_1^2}{f_1^2-f_2^2}\frac{A}{f_1^4} - \frac{f_2^2}{f_1^2-f_2^2}\frac{A}{f_2^4} = \frac{Af_2^2}{f_1^2f_2^2\left(f_1^2-f_2^2\right)} - \frac{Af_1^2}{f_1^2f_2^2\left(f_1^2-f_2^2\right)} \\ &= \frac{A\left(f_2^2-f_1^2\right)}{f_1^2f_2^2\left(f_1^2-f_2^2\right)} = -\frac{A}{f_1^2f_2^2}\end{aligned} \quad (24)$$

Jakowski et al. (1994) found residual range error values for GPS with VTEC of 124 TECU ranging from 9.1mm (10 degrees elevation) through 2.6mm at 25 degrees to 0.4mm at 50 degrees elevation (see their Table 3). They suggest that their full formula would 'calculate the bending error with an accuracy of better than 1mm at elevation angles greater than 10º for GPS users', while their simplified formula performs equivalently 'ignoring unusual ionospheric conditions'.

The recent study by Hoque and Jakowski (2008b) then expands on the previous work by Jakowski et al. (1994) and defines the two terms: 'excess path length' or geometric bending, and 'range error due to TEC difference at two frequencies' (which is called the TEC difference or dTEC bending error in this paper). They derived empirical formulae using raytracing simulations with a spherically symmetric ionosphere and checked the results using comparisons with reconstructed CHAMP profiles.

The formula for the geometric bending or 'excess path length' as it is called by Hoque and Jakowski, is slightly different from that mentioned in the previous work:

$$b_{gm} = \frac{7.5\times10^{-5}\cdot STEC^2\cdot\exp\left(-2.13\beta\right)}{f^4 HF_2\cdot\left(h_mF_2\right)^{1/8}} \qquad (25)$$

in millimetres, where STEC is in TECU, frequency f in GHz, F2 layer scale height HF_2 and peak ionization height h_mF_2 in kilometres and elevation β in radians.

Comparing the three formulae, it can be seen that the dependent parameters differ:

Jakowski et al (1994) full (Eqn. 22b) depends on STEC, N_mF_2, h_mF_2 and 1/cos(elevation angle)

Jakowski et al (1994) approx (Eqn. 23) depends on $STEC^2$, and 1/cos(elevation angle)

Hoque and Jakowski (2008b) full (Eqn. 25) depends on $STEC^2$, a negative exponential of the elevation angle, HF_2 and h_mF_2.

If 70km and 350km are chosen for HF_2 (the F2 layer scale height) and h_mF_2, the F2 layer peak height, as suggested by Hoque and Jakowski (2008b), the '2008 full' expression gives results very similar to the '1994 approx' expression (see Fig. 10).

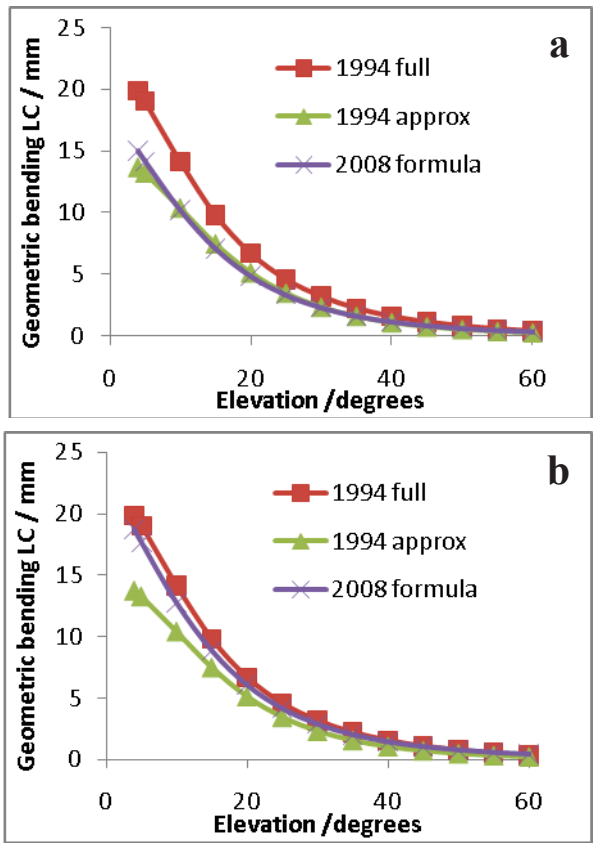


Fig. 10 Δb_{gm} using geometric bending formulae from Jakowski et al. (1994) and Hoque and Jakowski (2008b). a) using h_mF_2 =350km and HF_2 = 70km, b) using h_mF_2=350km and HF_2 = 56km (HF_2 calculated as suggested by Hoque and Jakowski (2008b))

If the six suggested formulae from different authors are compared (Fig. 11), the agreement is reasonable, though the approximate formula from Jakowski et al. (1994) is much lower. All the formulae are related to $1/f^4$. The formula of Hartmann and Leitinger (1984) results in the largest estimate for the error, and is also the only one to specifically include satellite height, probably because the paper was not specific to GPS. Interestingly, if 1000km is substituted for the 20200km used to evaluate the formula in Fig. 11, the result is close to the lowest formula, the approximate formula of Jakowski et al. (1994). Hartmann and Leitinger (1984) suggested their formula might need amendment at angles below 30 degrees. The empirical formula of Brunner and Gu (1991) has a slightly different elevation dependence to the remaining formulae; Δb_{gm} rises more quickly as the elevation decreases below 20 degrees.

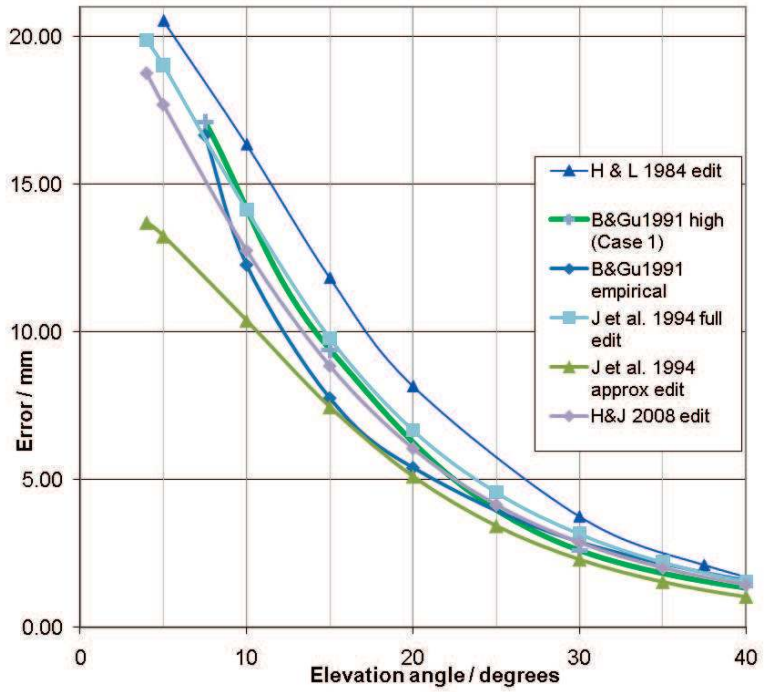


Fig. 11 Comparison of different formulae for Δb_{gm} using standard conditions equivalent to Brunner and Gu (1991) Case 1 ($VTEC$ = 138 TECU, N_mF_2 = 6 x 10^{12} electrons m^{-3}).The other formulae used are as follows: 'H & L 1984 edit', Eqn. 20b; B&Gu1991 empirical, Eqn. 21; J.et al. 1994 full edit, Eqn. 22b; J et al. 1994 approx edit, Eqn. 23; H&J 2008 edit, Eqn. 25

The full formula of Brunner and Gu, the full formula of Jakowski and the formula of Hoque and Jakowski all agree fairly closely (Table 4).

Table 4 Size of Δb_{gm} for formulae from different authors at 10 and 30 degrees elevation, using standard conditions equivalent to Case 1 of Brunner and Gu (1991).

Formula	Δb_{gm} /mm at 10 degrees elevation	Δb_{gm} /mm at 30 degrees elevation
H & L 1984 edit (Eqn. 20b)	16.3	3.8
B & Gu 1991 high (Case1)	14.1	2.6
B & Gu 1991 empirical (Eqn. 21)	12.3	2.9
J et al. 1994 full edit (Eqn. 22b)	14.1	3.2
J et al. 1994 approx edit (Eqn. 23)	10.4	2.3
H & J 2008 edit	12.7	2.9

Given that the Jakowski et al. (1994) paper modelled the ionosphere as a single Chapman layer, and the empirical formulae were derived by ignoring the plasmasphere, while the formula in the Hoque and Jakowski (2008b) paper is based on modelling of three Chapman layers and validated using profiles from radio

occultation measurements from the CHAMP satellite, the expression from the 2008b paper is presumably more correct.

An extension of the formulae of Hoque and Jakowski (2008b) for both the geometric and TEC difference bending terms to encompass the challenges of radio occultation scenarios can be found in Hoque and Jakowski (2010).

5.2. Total Electron Content difference bending term

The TEC difference bending error, b_{TEC} arises because 'due to the dispersive nature of the ionosphere, two radio signals do not follow the same curved path' (Hoque and Jakowski, 2008b). This means that the STEC accumulated along the signal path is not quite the same for the L1 and L2 GPS frequencies. The complete elimination of the first order ionospheric term relies on $STEC_1$ and $STEC_2$ being the same, so the TEC bending term is effectively a residual cancellation error from the first order refractive index term. Hoque and Jakowski (2008b) derive an analytical formula for a Chapman layer, but to represent a more realistic ionosphere have derived an empirical correction formula to determine the excess STEC in addition to the straight path STEC. With slightly altered notation, their formula is:

$$\Delta TEC_i = \frac{0.1108 \cdot STEC^2 \cdot \exp\left(-2.1844\beta\right)}{f_i^2 HF_2 \cdot \left(h_m F_2\right)^{0.3}}, \qquad (26)$$

where β is in radians, HF_2 and h_mF_2 are in km, f_i is in Hz and $STEC$ is in electrons m^{-2}.

To turn this TEC difference into a signal difference, the same formula used for the first order correction is appropriate.

$$b_{TEC} = \frac{40.3 \cdot \Delta TEC_i}{f_i^2} \qquad (27)$$

Elevation angle, slant TEC, and signal frequency are quantities used in estimating I2 and I3, the higher order refractive index expansion terms (see Section 6 for further details). However, the F2 layer scale height, HF_2, and the peak ionisation height, h_mF_2, are not. Hoque and Jakowski (2008b: p11) observe that 'in practical cases, the information about F2 layer scale height and peak density height hmF2 are not easy to estimate.' They go on to suggest that 'using assumptions for HF2 and hmF2, a significant amount of error can be corrected. For HF2 = 70 km and hmF2 = 350 km, about 80% error will be removed on average.' This is calculated when modelling the ionosphere as four Chapman layers. However, they find that when using values from real profiles from CHAMP, with HF_2 and h_mF_2 obtained from radio occultation data, on average 65% of error is corrected by the formula. This suggests that using constant values of HF_2 and h_mF_2 for real data could correct an even lower percentage of the error, and indicates that investigating sources of HF_2 and h_mF_2 estimates could be worthwhile. This is particularly so because the ionosphere has latitudinal (geographical) and seasonal dependence, so systematic biases could be potentially be introduced.

Hoque and Jakowski (2008b) recommend using their Equation (19) to find HF_2, given the other quantities in the equation:

$$STEC \approx \frac{H\left(h_m + R_E\right) N_m \sqrt{2\pi \exp(1)}}{\sqrt{\left(h_m + R_E\right)^2 - \left(R_h + R_E\right)^2 \cos^2 \beta}} \qquad (28a)$$

where H = scale height in m, giving $STEC$ in units of electrons m^{-2} and N_m = maximum ionisation at altitude h_m in electron m^{-3}.

The equation can be rewritten more specifically as:

$$STEC \approx \frac{HF_2\left(h_mF_2 + R_E\right) N_mF_2 \sqrt{2\pi \exp(1)}}{\sqrt{\left(h_mF_2 + R_E\right)^2 - \left(R_h + R_E\right)^2 \cos^2 \beta}}$$

and rearranged to

$$HF_2 \approx \frac{STEC\sqrt{\left(h_mF_2 + R_E\right)^2 - \left(R_h + R_E\right)^2 \cos^2 \beta}}{\left(h_mF_2 + R_E\right) N_mF_2 \sqrt{2\pi \exp(1)}} \qquad (28b)$$

Petrie et al. (2010a) used VTEC to estimate STEC, so chose to use the vertical case to estimate HF_2 from N_mF_2 (the above expression is based on representing the ionosphere in a simplified way as a Chapman layer).

With vertical elevation, the equation simplifies to:

$$HF_2 \approx \frac{VTEC}{N_mF_2\sqrt{2\pi \exp(1)}} \approx \frac{VTEC}{4.13 N_mF_2} \qquad (29)$$

However, Eqns. 26 and 29 still leave the issue of where to obtain estimates of N_mF_2 and h_mF_2. This is discussed in the next section, which discusses implementation of the corrections in practical situations.

6. Implementation

Section 6 reviews and discusses the compromises between analytical rigour and computational efficiency that are often needed when implementing corrections for the higher-order ionospheric terms. To recap, the I1,I2, I3 (and the remaining I4+) terms are the result of expanding the refractive index of the ionosphere as a series. I1 contains the majority of the effect (but is usually cancelled in a two frequency combination, see Eq. 17) and I2 and higher are thus known as higher-order terms. There are also two bending terms: b_{gm}, which is simply the effect of bending due to refraction, and b_{TEC}, which arises when the I1 term is not quite cancelled in the combination due to the frequency dependent curvature of b_{gm}. A summary of the terms, necessary inputs and general implementation choices is given in Fig. 12.

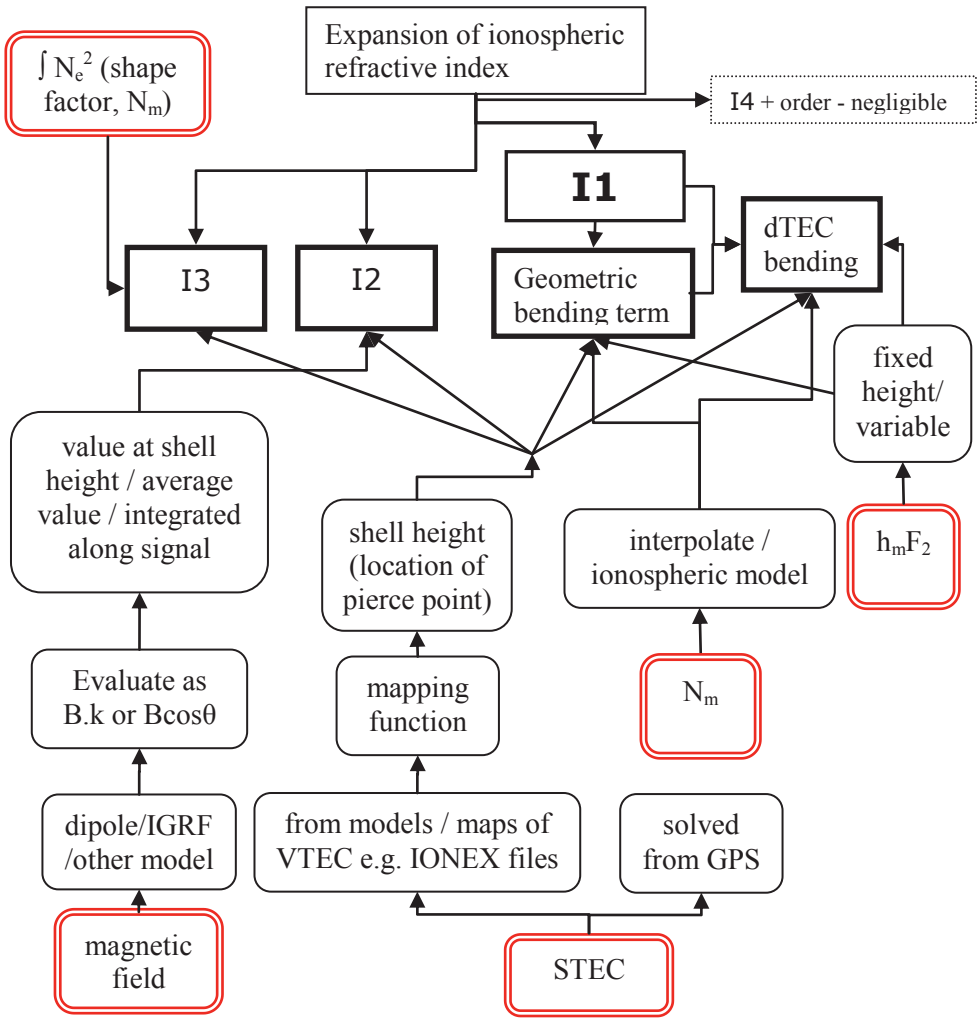


Fig. 12 Summary of corrections, choices made when modelling them and necessary inputs

6.1. Magnetic field

As shown in Fig. 12 and Section 4, the magnetic field contributes to the I2 correction. This section first discusses possible models for the magnetic field, moves on to forming the dot product between the magnetic and signal vectors, before concluding with a discussion of the thin shell model and the effect of taking Bcosθ at a single height.

6.1.1. Magnetic field models

To a first approximation, the Earth is a sphere uniformly magnetised in the direction of the dipole axis (Davies, 1990: p39) so one of the simplest ways of representing the Earth's magnetic field is as a simple co-centric dipole model. Returning briefly to basics, Fowler (1990: p37) derives the equations in spherical geomagnetic coordinates for the components of the magnetic field due to a dipole at the centre of the Earth. The radial component of the field, B_r, is as follows:

$$B_r(r,\theta,\phi) = -\frac{2\mu_0 m \cos\theta}{4\pi r^3} \qquad (30)$$

while 'the component of the field in the θ direction', B_θ, (or south (Parkinson, 1983: p122)) is:

$$B_\theta(r,\theta,\phi) = -\frac{\mu_0 m \sin\theta}{4\pi r^3} \qquad (31)$$

where r is radius, θ is colatitude and φ is longitude. This means that in a local geomagnetic North East Up (NEU) coordinate system, if we define R_E = 6371km as the approximate radius of the Earth, and r_m as the height above R_E, we can use the estimate of $B_{eq} = 3.12 \times 10^{-5}$ nT at the Earth's surface at the geomagnetic equator (Bassiri and Hajj, 1993) to estimate the magnetic dipole vector at a point in the geomagnetic NEU coordinate system:

$$\mathbf{B} = \begin{pmatrix} B_{eq}\left(\frac{R_E}{R_E + r_m}\right)^3 \sin\theta \\ 0 \\ -2B_{eq}\left(\frac{R_E}{R_E + r_m}\right)^3 \cos\theta \end{pmatrix} \qquad (32)$$

One alternative to a dipole model is the more accurate International Geomagnetic Reference Field (IGRF), the latest version of which (IGRF-11) was released in December 2009. The IGRF consists of spherical harmonic coefficients, representing the Earth's main field and its secular variation. Each release incorporates predicted coefficients for five years of secular variation which are then revised to definitive coefficients as measurements are incorporated. The coefficients are available at the following url: http://www.ngdc.noaa.gov/IAGA/vmod/igrf.html (accessed 12 Jan 2010). For further information on the IGRF see e.g. Maus et al (2005), Maus and Macmillan (2005), and the IGRF "health warning" webpage at http://www.ngdc.noaa.gov/IAGA/vmod/igrfhw.html (accessed 12 Jan 2010). Other models such as the World Magnetic Model (WMM) are also available (http://www.ngdc.noaa.gov/geomag/WMM/DoDWMM.shtml, accessed 28 Jan 2010).

It should be noted that the IGRF and WMM model only that part of the field that originates from the Earth's core. This is often known as the 'main field' and represents the vast majority of the magnetic field intensity. Langel (1987) states 'it is now known that the magnetic field at any location near the Earth can be attributed to a combination of three sources located respectively in the Earth's core, in the Earth's crust, and in the Earth's ionosphere and beyond. By far the largest in magnitude is the field from the core, or the 'main' field. Near dipolar in nature, the strength of the main field is approximately 60 000nT (nanoTesla) at the poles and approximately 30 000nT at the equator.' (see e.g. Parkinson (1983: p6) for a table with the typical strengths of the various different components).

One external source of magnetism is from geomagnetic storms. Pireaux et al. (2010: Fig. 6) show an example of the impact of a geomagnetic storm on vertical **B**, with variations of almost 300nT. However, when considered in the context of the background typical quiet night level of over 45000nT at the location the percentage change is less than 1%. In major magnetic storms effects of up to several thousand nT are possible (CM4 website http://denali.gsfc.nasa.gov/cm/, accessed 12 Jan 2010). In most regions, several thousand nT will still be a minor secondary effect in comparison to the strength of the main field, so this is not a major source of error. However, in future it may become possible to use a model such as the CM4 model (Sabaka et al., 2004). This is a "comprehensive model" and includes other sources, such as the ionospheric field, in addition to the main field.

Looking at the magnetic models applied in the studies to date reveals a mixture of approaches. While the initial ray tracing study of Brunner and Gu (1991) did use the IGRF-1985, the subsequent approximation studies (Bassiri and Hajj, 1993; Kedar et al., 2003; Fritsche et al., 2005) used a co-centric tilted dipole model in which the magnetic field can be represented using the equations outlined above. When using a dipole field, the geomagnetic pole must be defined. Bassiri and Hajj (1993) used a geomagnetic north pole at an angle of 78.5N, 291.0E. Kedar et al. (2003) modelled their dipole after that of Bassiri and Hajj. However, Fritsche et al. (2005) used an annually changing geomagnetic pole.

In contrast, Hernandez-Pajares et al. (2007), Palamartchouk (2010), Hawarey et al. (2005), Morton et al. (2009a,b) used the IGRF. Hernandez-Pajares et al. (2007) note that there can be up to 60% difference between the dipole and IGRF models, most prominently in the South Atlantic Anomaly, and plot the I2 correction with the contrasting models for ASC1, a site in the South Atlantic. However, it is challenging to compare results from papers with different processing strategies as it is hard to definitively attribute the differences. As part of a study into comparisons of different aspects of higher order ionospheric modelling, Petrie et al. (2010b) compared otherwise identical processing runs and looked at differences in the results due to using the two models. They found up to 2mm difference in the transformation from the GPS reference frame to ITRF2005 and small but noticeable shifts in the mean coordinates in areas with the largest differences between the two magnetic field models (see Fig. 13).

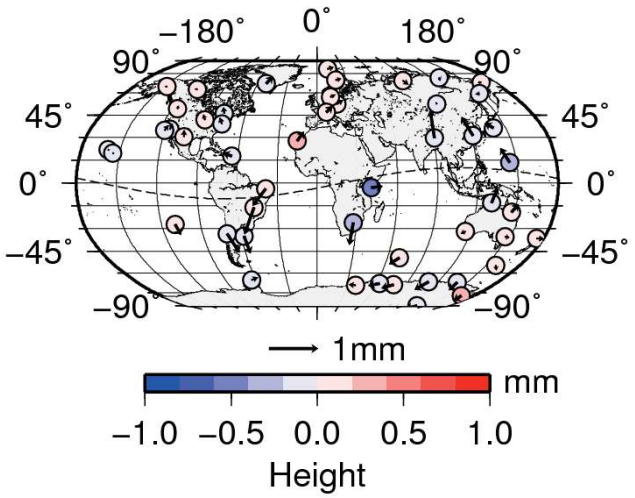


Fig. 13 Mean coordinate differences (2001.0-2004.0) between runs using a co-centric tilted dipole or the IGRF-10 to model the second-order ionospheric correction (dipole – IGRF-10). Sites have > 2.5 years of data. Geomagnetic equator shown as dashed line, arrows represent shifts in plan. Data from Petrie et al. (2010b)

6.1.2. Evaluating $\mathbf{B}\cdot\mathbf{k}$ or B cosθ and coordinate system effects

Estimating $\mathbf{B}\cdot\mathbf{k}$ or B$\cos\theta$ at a point or points along the GPS signal is part of estimating the I2 term (see Section 4.1). Mathematically, the dot product of a vector with a unit vector ($\mathbf{B}\cdot\mathbf{k}$) is equivalent to the size of the first vector multiplied by the angle between the two vectors (B$\cos\theta$). However, some issues can arise in the way the mathematics is implemented. This is particularly true if using a dipole field, as it is naturally in a geomagnetic coordinate system, aligned to the geomagnetic poles rather than the geographic ones.

Evaluation of the magnetic dipole vector at a particular point along the line between the satellite and receiver can be achieved in two ways;

Method 1: Calculation of a point along a line between two sets of coordinates in a particular frame (see section 6.1.3 below) and evaluation of the dipole field at that point. The geographic coordinates for the desired point can be converted to geomagnetic coordinates (see Bassiri and Hajj (1993) Eqs. 22 and 23). These geomagnetic coordinates together with the height of the point are then used to estimate the vector of the magnetic field due to the dipole (Eqn. 32). Once the magnetic vector is known, it can be converted first into a vector in geomagnetic Cartesian coordinates (components along X_{mag},Y_{mag},Z_{mag} axes), then to a geographic Cartesian system (components along X, Y, Z axes). Once in the same reference frame as the satellite-site unit vector, the dot product of the two can be found successfully.

This is probably the best method, but care must be taken to correctly rotate the resulting vector into the same frame as the site-satellite unit vector (or vice-versa). This method is used by Fritsche et al.(2005), Petrie et al. (2010b).

Method 2: Using the azimuth and elevation of the satellite from the site to define the signal vector, together with an adjustment to colatitude. This approach is suggested by Bassiri and Hajj (1993), and is also followed by Kedar et al.(2003), but as outlined below it may be not be as mathematically rigorous as Method 1 at high latitudes.

If Method 2 is used, the satellite site unit vector, **k**, is defined in terms of local (at the site) geomagnetic East, North, Up unit vectors, ($\mathbf{X_m}$, $\mathbf{Y_m}$, $\mathbf{Z_m}$):

$\mathbf{k} = -\left(\cos E_m \sin A_m \mathbf{X_m} + \cos E_m \cos A_m \mathrm{Y_m} + \sin E_m \mathbf{Z_m}\right)$ where E_m is elevation and A_m is azimuth measured from magnetic North (after Eqn. 18, Bassiri and Hajj (1993)).

However, Bassiri and Hajj (1993) seem to take the dot product of **k** with a magnetic dipole vector, obtained in a coordinate system where the colatitude is adjusted to the same value as the desired point (in their case the pierce point), but the longitude is not. In terms of magnitude of **B**, this has no effect, but in terms of $\mathbf{B}\cdot\mathbf{k}$, it introduces an error; the local NEU reference frame at the pierce point is not the same as that at the adjusted point, unless the geomagnetic longitudes are the same (see Fig. 14 below). The error is largest at high geomagnetic latitudes, where a given distance can cross more meridians of longitude.

If using the IGRF, the software available from the National Geophysical Data Center (http://www.ngdc.noaa.gov/IAGA/vmod/igrf.html, accessed 10 Feb 2010) is set up to provide data in a geodetic (WGS84 ellipsoid) or geocentric (spherical) coordinate system as requested.

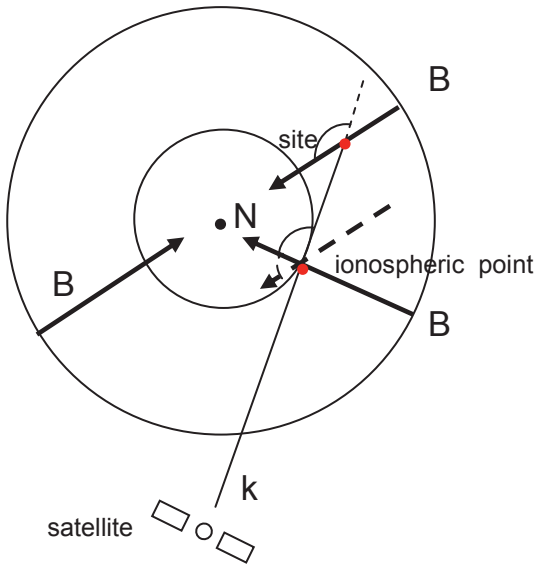


Fig. 14 Effect of longitude difference on $\mathbf{B} \cdot \mathbf{k}$ when assuming East North Up is the same frame at the site and the desired ionospheric point. Viewpoint is from above geomagnetic boreal pole, marked as N. Dotted arrow represents magnetic vector adjusted for colatitude but not for longitude. Circles represent geomagnetic parallels of latitude

6.1.3. Integration at a point – the thin shell model

To be most accurate the terms in the refractive index formula should be integrated over the full path length. However, this is usually not practical, mainly due to the computational burden. The vertical distribution of the electron content in the ionosphere is also not well known on a daily basis, which makes integrating $\int B\cos\theta N_e \mathrm{dL}$ along the length of signal challenging. If $B\cos\theta$ is removed from the integral, $\int N_e \mathrm{dL}$ is simply STEC, which is much easier to obtain. Thus $B\cos\theta$ is usually taken outside the integral; the question is then what to use for the value of $B\cos\theta$. Brunner and Gu (1991) used the average value along the path. In general, it is not easy to average $B\cos\theta$ for practical uses, although Hoque and Jakowski (2008a) calculated a specific set of parameters for approximating the average value specifically over Europe.

The most commonly used solution at present is to calculate the correction as if all the electron content is in a 'thin shell' at a selected height, h_i, and take the value of $B\cos\theta$ at the point where the signal crosses the thin shell. This crossing point is often known as the 'pierce point' (see Fig. 15).

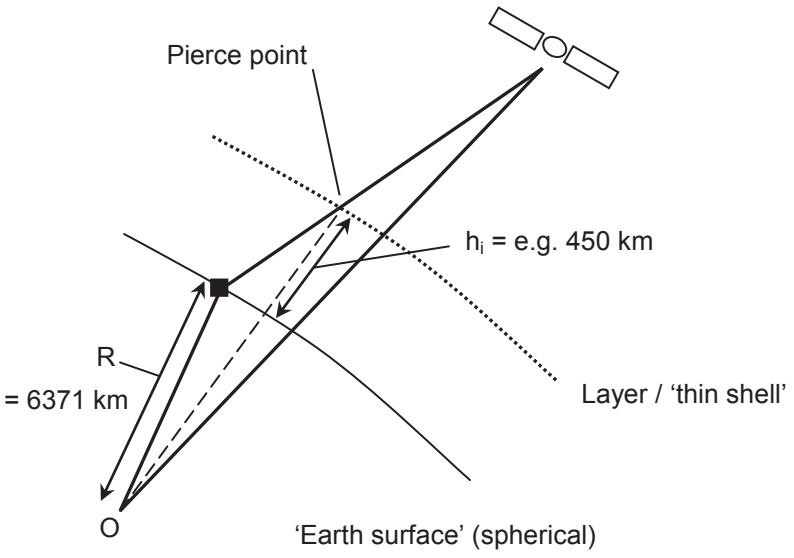


Fig. 15 Pierce point geometry

The coordinates of the point where the straight line path from the satellite to the receiver crosses a shell at h_i can be found as follows:

It can be seen that (Fig. 16):

$$\underline{x} = \underline{x}_{rec} + \lambda\left(\underline{x}_{sat} - \underline{x}_{rec}\right) \tag{33}$$

where $0 \le \lambda \le 1$ (λ represents fraction of distance from receiver to satellite).

It can be shown that:

$$\lambda = \frac{-\underline{x}_{rec} \cdot \left(\underline{x}_{sat} - \underline{x}_{rec}\right) + \sqrt{\left[\underline{x}_{rec} \cdot \left(\underline{x}_{sat} - \underline{x}_{rec}\right)\right]^2 - \left|\underline{x}_{sat} - \underline{x}_{rec}\right|^2 \left[\left|\underline{x}_{rec}\right|^2 - \left(R + D\right)^2\right]}}{\left|\underline{x}_{sat} - \underline{x}_{rec}\right|^2} \tag{34}$$

From the above, it is apparent that if the coordinates of the satellite and the receiver are known, R and D are 6371 and 450 km respectively and the equation for λ is substituted into Eqn. 33, the coordinates of the pierce point can be found.

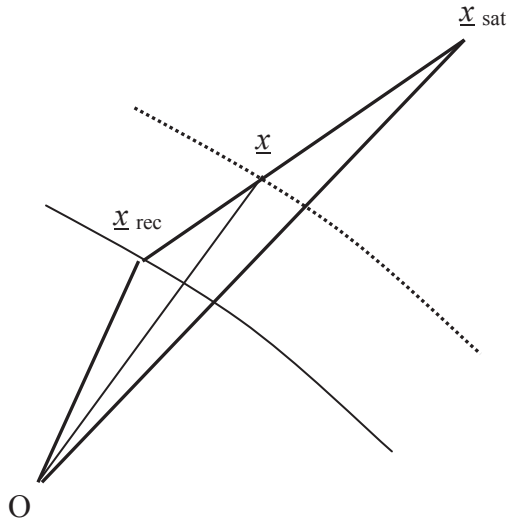


Fig. 16 Vectors in pierce point calculation

The concept of a pierce point is also useful if calculating slant TEC from a global ionospheric map of VTEC, as if the VTEC directly above the receiver is used, it may not correctly represent the ionosphere that the signal has travelled through. To improve the value, typically the VTEC at the coordinates of the pierce point is used (see Section 6.2.1 for further discussion on issues regarding global ionospheric maps).

So, how does the thin shell approximation compare, and is there an optimum shell height? Gherm et al. (2006) and Morton et al. (2009b) show that B$\cos\theta$ can vary substantially with height along potential GPS signal paths. Fig. 17 shows an example of such variation, at Arecibo (Morton et al., 2009b). Strangeways and Ioannides (2002) point out that "since this term is everywhere weighted by the local electron density, the major contribution to the integral will occur where the electron density is greatest, which will be around the F layer peak". This would be expected to be between ~250-550 km, reducing the effect of the variability somewhat. The effects shown in Fig. 17 would also be very geographically dependent. On the other hand, it has been suggested that the height of the first moment of the height distribution of electron density or "center of gravity" of the ionosphere would be appropriate, with a rule of thumb formula of $h_i=h_m+50$km, where h_i is evaluation height (Leitinger and Putz, 1988). The approximation was tested by Bassiri and Hajj (1993) by comparing precise and approximate calculations on an ionospheric model and found to be accurate to within 90%, using a shell height of 300km.

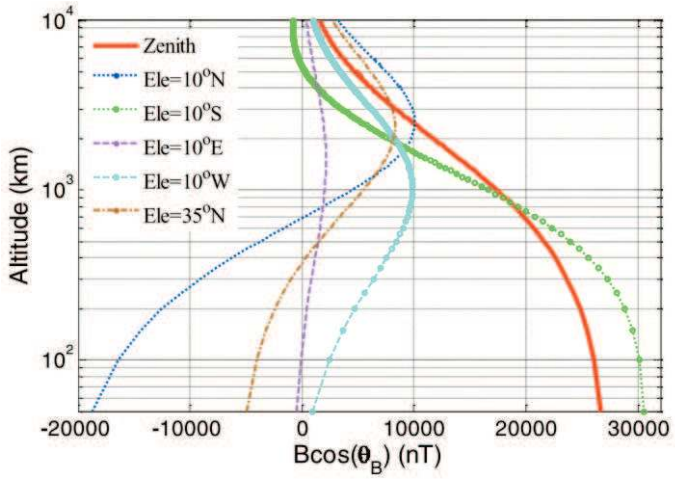


Fig. 17 Bcos(θ_B) versus altitude above Arecibo Observatory (latitude: 18°20′36.6″N, longitude: 66°45′11.1″W) in 2008. Examples are given for GPS signals arriving from zenith, 10 degrees elevation from North, South, East, and West, and 35 degrees from North, respectively. (Reproduced from Morton et al., 2009b)

Hoque and Jakowski (2008b) also assess the effect of using a value of Bcosθ at a particular height, in contrast to calculating Bcosθ along raypaths. For a shell height of 400km, and the ionosphere represented by a Chapman layer with a VTEC of ~143TECU and maximum electron density of 4.96×10^{12} electrons/m^3 at a height of 400km, they find a maximum difference to the second order ionospheric LC correction of ~ ±1.9mm at low elevation, with the maximum correction ~ ± 29mm. This is ~7%, so agrees with the findings of Bassiri and Hajj (1993). They also provide plots showing the geographical distribution of the error, which is generally largest close to the geomagnetic equator.

Strangeways and Ioannides (2002) performed ray-tracing calculations. They investigated the height at which Bcosθ would equal the average value, but found that due to azimuth changes an average value could still underestimate the effect. They finally conclude that "For a variety of Earth station locations, it has been found that by taking the value of $B \cdot |\cos \Theta|$ at 330 km for elevation angles < 25˚ and at 520 km for elevation angles ≥ 25˚ degrees, the field effect can be approximated to an accuracy of 2mm."

However, all three studies mentioned above seem to have assessed the effect using a selected modelled ionosphere, so the question of the effect of selecting a shell height which does not match the modelled peak electron density height, h_{max}, is not completely addressed.

6.1.4. Height of the thin shell

Opinions on the height of the thin layer vary. Hartmann and Leitinger (1984: p124) suggest that "the ionosphere [can be considered] to be a thin layer around a "mean ionospheric height" for which 400km is a good approximation for mid-latitudes and in the absence of severe ionospheric disturbances. ("Thin" means thin compared with the radius of the earth.)".

Bassiri and Hajj (1993) state that "dependence ..[on the height of the assumed thin layer] is weak and a nominal value of H depicting the peak height of the ionosphere, should serve the purpose of obtaining a working formula to approximate the second-order effect". They choose "a representative global average peak height" of 300km. Kedar et al. (2003) also use this approximation, but with a peak height of 400km. Fritsche et al.(2005) quote a value of 400km, but in fact used 450km (Fritsche 2007 pers. comm.). Hernandez-Pajares et al. (2007) also use 450km, as do Petrie et al. (2010b).

If the preferred height for the thin layer is considered to be at the height of the peak electron density, this can be very variable. Klobuchar (1996) reports that "the height of the peak of the electron density of the F2 region generally varies from 250 to 400 km, but it can be even much higher or somewhat lower under extreme conditions". Fig. 18 shows modelled variability in h_mF_2 for a day during ionospheric maximum.

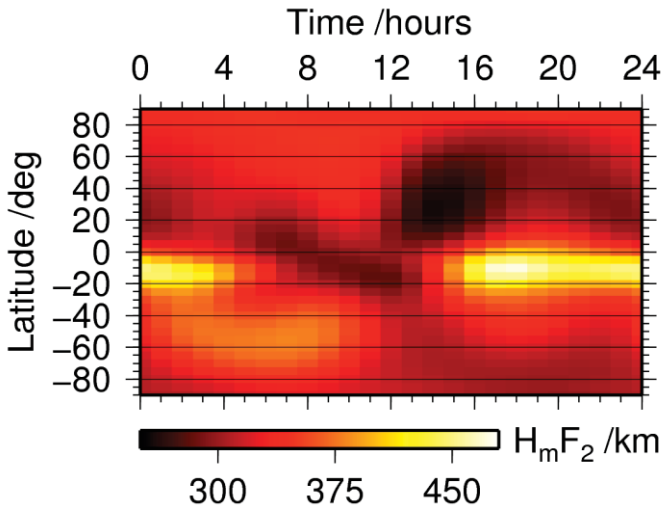


Fig. 18 Variability of estimated h_mF_2 at 270 degrees longitude on DOY 301, 2001 with latitude and time. Data from the IRI2007 model. (Reproduced from Petrie et al., 2010a)

While there have been studies including the effect of shell height when calculating VTEC from GPS data, (Komjathy and Langley, 1996; Horvath and Crozier, 2007; Mushini et al., 2009), there is a lack of published information on how the shell height affects the second order ionospheric correction.

In general the geomagnetic field decreases with increasing height above the Earth surface, so selecting a lower height at which to evaluate $\mathbf{B}\cdot\mathbf{k}$ would tend to increase the size of the correction. Fig. 19 shows an example of the size of the effects that might be expected if a magnetic field (in this case the IGRF) is evaluated for pierce points for a range of heights from 250-500km.

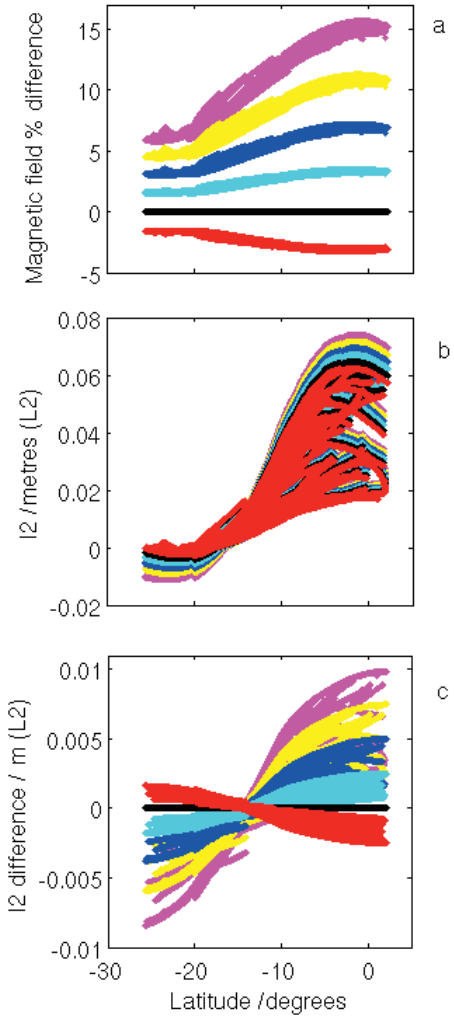


Fig. 19 Effects of changing pierce point height for the magnetic field on the GPS signals received at DARW (12.84S, 131.13E), DOY 301, 2001. a) Magnetic field strength (IGRF-10) percentage difference to value at 450km. b) I2 correction c) I2 difference to value at 450 km. Height values plotted: 250km-magenta, 300km-yellow, 350km-blue, 400km-cyan, 450km-black, 500km-red

The percentage difference in magnetic field strength from the commonly adopted 450km (evaluated as (test height value-450km value)/450km value) is up to ~16%. The difference in I2 is up to ~10mm for L2, but depends on the size of the correction. The percentage difference in I2 is similar to that in the magnetic field (though much larger values are possible when the reference value is close to zero).

Palamartchouk (2010) uses a variable shell height based on the electron density from the IRI2007 model. This is a simulation

study with the IRI model supplying all the ionospheric data, including the TEC. It would be very interesting to see how the simulated values would compare with those using a fixed ionospheric height. However, sufficient ionospheric information may not yet be available to make this possible for practical error reduction.

6.2. Slant Total Electron Content

If we accept the thin layer approximation as a valid simplification, then $B\cos\theta$ can be removed from the integral in the I2 term leaving simply the integral of the electron content along the line of the signal. The bending terms and potentially the I3 term also require the same information, often described as slant total electron content (STEC). STEC can be obtained from available pre-existing ionospheric data, or it can be solved for directly from the GPS signals.

6.2.1. STEC from pre-computed ionospheric data

One method of obtaining the STEC along a given signal path is from a pre-existing dataset. Global ionospheric maps (GIMs) of vertical total electron content (VTEC) are created at different IGS analysis centres and combined to form the IGS product (Feltens, 2003; Hernández-Pajares et al., 2008). The information is stored in the IONEX format (Schaer et al., 1998). In a daily IONEX file there are 13 maps spaced two hours apart, with a resolution of five degrees longitude and two and a half degrees latitude. The combined IGS IONEX files are available from IGS data centres (e.g. CDDIS ftp://cddis.gsfc.nasa.gov/pub/gps/products/ionex/) for day of year 152, 1998 onwards. IONEX files for the period back to 1995 are available from the Centre for Orbit Determination in Europe (CODE), though there are fewer maps per day. While the standard maps are produced purely from GNSS data, Todorova et al. (2008) suggest that combining TEC maps from GNSS and altimetry would lead to improved accuracy over the oceans.

The values of VTEC in the file must be interpolated to obtain a value at the pierce point where the GPS signal crosses the single layer. The coordinates of this point can be found using the calculation outlined in Section 6.1.3. Suggested interpolation strategies are provided by Schaer et al. (1998).

To convert the ionospheric information from VTEC to STEC a mapping function must be used. It is generally appropriate to select the mapping function used to convert the original data to VTEC during the file production. Typically the mapping function used is the single layer or thin shell mapping function:

$$F(z) = \frac{1}{\cos z'} \quad \text{with} \quad \sin z' = \frac{R}{R + h_i}\sin(z) \tag{35}$$

where z is the zenith angle of the signal at the receiver, R is the mean Earth radius (~6371km), h_i is the height of the thin shell above the Earth's surface and z' is then the zenith angle at h_i. For the IGS IONEX files h_i is set as 450km. For CODE IONEX files h_i is 450km after DOY 087 1998, and 400km on and before DOY 086 (Schaer, 1997)). However, after DOY 251 in 2001, the CODE IONEX files use the Modified Single-Layer Model Mapping Function:

$$F(z) = \frac{1}{\cos z'} \quad \text{with} \quad \sin z' = \frac{R}{R + H}\sin(\alpha\, z) \tag{36}$$

with values of R = 6371 km, H = 580.1 km and α = 0.9782 which best approximate the JPL extended slab model mapping function, assuming a maximum zenith distance of 80 degrees (CODE, 2007).

However, the assumptions made when using the simple mapping functions described above can lead to errors. It is usual to assume a 2D distribution of the electron content at a given ionospheric effective height (such as 450 km for the IGS VTEC maps) which is constant in time and space. This assumption is quite reasonable for mid latitude and day time conditions, but it can induce significant errors elsewhere, in particular at low latitudes where the higher order ionospheric effects are more important due to the higher electron content. Komjathy and Langley (1996) and Komjathy (1997) consider errors due to shell height when calculating VTEC, and find errors of up to 1 TECU. However, the studies were considering a period with low solar activity (Oct-Nov 1995) and mid-latitudes, where the height of the ionosphere is less variable.

From Eqn. 35, it is possible to calculate the equivalent VTEC for a STEC of 350 TECU, using a thin layer based at different heights (Table 5) and a signal at ten degrees elevation. A difference of 150 km leads to an error in the resulting VTEC of between 15-20 TECU. The pierce point will also be above a different geographic location, so the coordinates attributed to the VTEC will be different.

Table 5 Variations in estimated *VTEC* with varying thin layer height, h_i, based on an *STEC* of 350 TECU and a signal at 10 degrees elevation.

h_i /km	F	*VTEC* /TECU
300	2.943	118.908
450	2.549	137.305
600	2.295	152.529

Now, for a single signal, resulting in a single value of *VTEC*, the process is reversible, as long as the same value is used for h_i. However, in practice, values are averaged spatially and temporally, so the process is not directly reversible. Also, supposing a perfect map of VTEC was obtained by other means, if an incorrect h_i was used in the mapping process to estimate STEC, the errors described above would apply directly.

An optimal height for a thin layer model, or "effective ionospheric height" can be calculated (see Hernández-Pajares et al. (2005) for further details). Fig. 20 shows a snapshot of effective ionospheric height computed from actual ground GPS data, validated with occultation measurements. Discrepancies in effective height compared to a reference value of 450 km can be seen to be up to ±200 km and larger discrepancies are possible during geomagnetic storms. Fig. 20 also shows that the effective height varies substantially from 450km in equatorial regions, where high TEC and thus large higher order ionospheric effects are found. Mushini et al. (2009) consider this issue for polar regions at ionospheric minimum and comparison to a fixed shell at 350km. The small differences they find (up to 0.3 TECU) would be expected to increase at ionospheric maximum and with a comparison shell at 450km. Finally, there is also a small purely geometric mapping function error from treating the ionosphere as an infinitely thin shell (Smith et al., 2008).

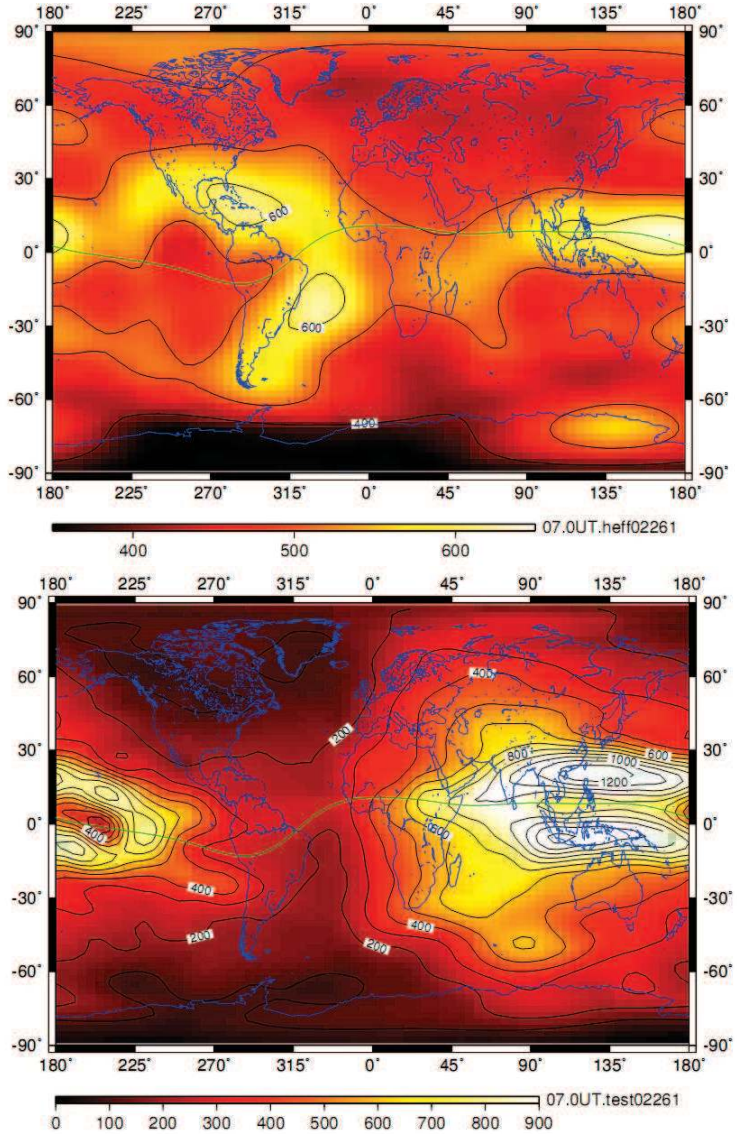


Fig. 20 Example of GPS ionospheric effective height (in km) computed from global IGS data in Solar maximum but geomagnetically quiet conditions, for 0700UT, day 261 of 2002 (top plot). Colour scale from 450km (dark) to 650km (pale).The corresponding VTEC map (in tenths of TECU, from 0 (dark) to 90 TECU (white)) is shown in the second plot (both figures extracted from Hernández-Pajares et al. (2005))

An alternative potential source of pre-computed STEC would be an ionospheric model such as the International Reference Ionosphere (IRI). The IRI is 'based on a wide range of ground and space data. It describes monthly averages of ionospheric densities and temperatures in the altitude range 50 – 1500 km in the non-auroral ionosphere' (Bilitza and Reinisch, 2008) and the latest version at the time of writing is IRI2007.

A comparison between the IRI 2001 model and data from the CHAMP satellite showed general agreement between CHAMP data and the IRI model in the seasonal variation of the electron density and temperature, with some exceptions (Liu et al., 2007). A later study performed a comparison with Slant Total Electron Content (STEC) data and found that IRI2007 is much improved compared to IRI2001 (Coïsson et al., 2008). However, while models such as the IRI or PIM may be useful for simulations (e.g. Palamartchouk, 2010), the TEC data they provide is currently not sufficiently accurate for correcting real data – it represents ionospheric climate, rather than daily (and subdaily) variations in ionospheric weather (see e.g. Leitinger et al.(2000) for a discussion of space weather).

6.2.2. STEC from GPS signals

The alternative method of obtaining STEC directly from the GPS signals received has an attractive simplicity – no outside ionospheric data is necessary. This section will review the method outlined by Hernández-Pajares et al. (2007) for calculating *STEC* for higher order ionospheric effects based on combining phase and code

measurements. A similar method is described in some detail in Komjathy (1997: p89-96). Further discussion of this topic can be found in e.g. Ciraolo et al. (2007) and Brunini and Azpilicueta (2009).

Simple linear combinations can be formed for phase and code by subtraction (using the first order ionospheric approximation, see Eqns. 13, 14 and 16b). We will denote the combinations by LN (to distinguish them from the 'ionosphere-free' linear combination, LC, described above):

$$\Phi_{LN} = \Phi_1 - \Phi_2 = xSTEC + b_{LN}\,, \qquad (37)$$

$$P_{LN} = P_2 - P_1 = xSTEC + D_{LN} + D^{'}_{LN} \qquad (38)$$

where $x = 40.3\left(\frac{1}{f_2^2} - \frac{1}{f_1^2}\right)$, b_{LN} includes the carrier phase ambiguity and interfrequency phase biases, and D_{LN}, $D^{'}_{LN}$ are interfrequency delay code biases (also known as differential code biases (DCBs)) for receiver and satellites respectively. Note that the order of subtraction is different for the phase and code combinations due to the sign of the ionospheric effect.

The phase combination is more precise due to its smaller wavelength, but there is the issue of the unknown integer ambiguities. This can be fixed by aligning the ionospheric carrier phase combinations Φ_{LN} for each continuous satellite-receiver arc to the code combination, P_{LN}, corrected with the interfrequency code delay biases, D_{LN} and $D^{'}_{LN}$, for receiver and satellites respectively. A continuous arc is one where the integer ambiguities remain the same for each successive observation, i.e. no cycle slips occur.

The STEC can thus be computed as follows:

$$STEC \equiv \int N_e dL = \alpha\left(\Phi_{LN} - \left\langle \Phi_{LN} - P_{LN} \right\rangle - D_i - D_i{}'\right) \quad (39)$$

where $\left\langle \Phi_{LN} - P_{LN} \right\rangle$ is the mean value over a continuous satellite-receiver arc, and $\alpha = 1/x = 9.52$ TECU/metre.

Horvath and Crozier (2007) discuss practical issues when implementing this strategy, such as cycle slip detection, with examples of issues from real GPS data. Burrell et al. (2009) discuss issues arising from and methods for performing real time TEC calculation, when averaging over a full satellite receiver arc is not possible. Hoque and Jakowski (2010) discuss higher order ionospheric effects on TEC estimation using radio occultation when the receiver is located on a satellite.

6.2.3. Comparison of methods for obtaining STEC

In order to evaluate the error of directly obtaining the STEC for each given transmitter-receiver dual-frequency observation, compared with the procedures relying on global VTEC maps (see section 6.2.1) we can focus on the following main sources of error (the first order approximation can be used with typically less than 0.1% of error):

Errors that affect the direct approach arise mainly from the code pseudorange measurements, due to multipath and thermal noise. Ciraolo et al. (2007) suggest that the major contributor is the code delay multipath and find

errors up to 5.3TECU. In theory, multipath could be almost cancelled by applying sidereal-day filtering for permanent stations (due to the repeatability of the GPS constellation every ~23h56m, see for instance Larson et al. (2007), paragraphs 13-14 and cited references for details). The code pseudorange thermal noise, σ, (~1 metre for a single ionospheric measurement) can be significantly reduced when the average over the phase-continuous arc of data is taken in the previous equation:

$$\sigma_{average} = \alpha \frac{\sigma}{\sqrt{n}} \qquad (40)$$

where n is the number of measurements in the continuous arc of data used for the carrier phase levelling. In this way the thermal noise error can easily be reduced by an order of magnitude, leaving from one to a few TECUs of error, and it is independent of the magnitude of the STEC.

DCB errors are typically at the level of ~0.1 ns for satellite transmitters and ~1ns for receivers. This implies typical errors in DCBs of less than 5 TECU, although Ciraolo et al. (2007) found intra-day variations of DCBs of up to 8.8TECU. However, DCB errors also affect VTEC maps to a roughly similar extent as the maps are generally estimated simultaneously with DCBs. (see e.g. Hernández-Pajares (2004) for supporting values of DCB and VTEC assessments).

Fig. 21 shows an example comparison of STEC interpolated from an IONEX file with that obtained directly from the GPS signals.

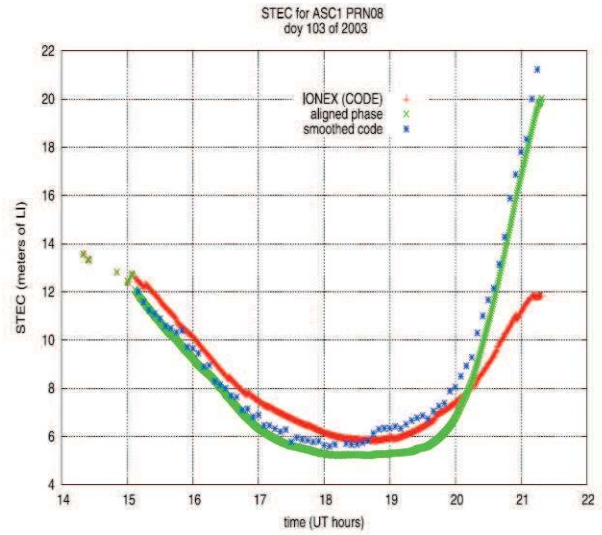


Fig. 21 Comparison of STEC estimated using a global ionospheric map (CODE IONEX file, red line) and using the method of code alignment proposed in Hernández-Pajares et al. (2007) (taking the instrumental delays from the same IONEX file, blue points). Both estimations are compared with the ground truth provided by the geometric-free combination of carrier phases (LI), in green. The data correspond to the ASC1 receiver (14W 08S), DOY 103 of 2003 (reproduced from Hernández-Pajares et al. (2007))

To summarise, it would seem that deriving the STEC directly from the code and phase ionospheric combinations using previously calculated DCBs would usually involve lower errors than using VTEC maps with a mapping function to convert to STEC. This is especially true for low latitude and solar maximum conditions (errors of less than 10 TECU compared to errors of potentially up to about 40 TECU), i.e. when the higher order ionospheric effects are largest. It is also true pre-1998 as available IONEX files have only one map of VTEC per day, so interpolation errors are large. This suggests direct STEC derivation may be preferable for providing the STEC in real-time situations or for runs for computing long time series of geodetic parameters, such as those including higher order ionospheric mitigation. However, no detailed comparison studies have yet been performed.

6.3. Shape factor

For the I3 term (Eqn. 14), obtaining the vertical distribution of N_e in order to integrate N_e^2 is still an issue. To deal with this, Hartmann and Leitinger (1984: p123) introduce the idea of a shape factor.

The shape factor can be defined as $$\eta = \frac{\int_0^{h_s} N_e^2 dh}{N_m \int_0^{h_s} N_e dh} \qquad (41)$$

This definition uses notation after Hartmann and Leitinger (1984) but the form of the definition is as given by Brunner and Gu (1991). However, if rearranged, the definition is the same as the original definition of Hartmann and Leitinger. Using the shape factor or parameter therefore allows $\int_0^{h_s} N_e^2 dh$ to be replaced by $\eta\, N_m \int_0^{h_s} N_e dh$ (effectively VTEC multiplied by N_mF_2 and the shape factor).

Hartmann and Leitinger (1984) suggest for a worst case scenario, estimating higher order ionospheric errors assuming a slab of constant thickness τ and a shape factor 1. They then extend the concept to allow consideration of other shapes – a triangular shape where the electron density increases linearly to a maximum then decreases again has a shape factor of 2/3, while a simple Chapman layer has a shape factor of 0.680.

Brunner and Gu (1991: Table 2 p211) calculated values for the 'shape parameter' for two cases and found that they hardly varied with elevation angle and maximum electron density. They concluded that the shape parameter 'may be assumed with 0.66 as the appropriate value for any profile of the electron density in the ionosphere with suficient accuracy'. Bassiri and Hajj (1993) also use an estimate of 0.66 for the shape parameter, as do Fritsche et al (2005) and Petrie et al. (2010b). Hoque and Jakowski (2008b) derive a similar number analytically for a Chapman function.

Brunner and Gu (1991) and Bassiri and Hajj (1993) used modelled ionospheres, so could extract values of N_mF_2. In order to make practical use of the technique for real data it is necessary to obtain an estimate for the values of N_mF_2 along each GPS signal path. Potential strategies for this are discussed in the next section.

6.4. Peak electron density, N_mF_2 and height of peak electron density, h_mF_2

6.4.1. N_mF_2

N_mF_2 is required for the I3 term, if the shape factor approach described in the previous section is taken. It is also needed for both bending terms, if the expressions of Hoque and Jakowski (2008b) are used.

Fritsche et al. (2005) solved this issue for the I3 term by using the values for the two ionospheric cases from Brunner and Gu (1991) as the basis for a linear interpolation of N_mF_2 (also known as N_m and N_{max}). While the equation in Fritsche et al. (2005) is:

$$N_m\left[m^{-3}\right] = \frac{(20-6)\cdot 10^{12}}{(4.55-1.38)\cdot 10^{18}} \cdot TEC \qquad (42a)$$

where TEC represents STEC, their intended equation (Pireaux et al., 2010) for correct interpolation was:

$$N_{\max} = \frac{(20-6)\times 10^{12}}{(4.55-1.38)\times 10^{18}} \times \left(TEC - 4.55\times 10^{18}\right) + 20\times 10^{12} \qquad (42b)$$

Fritsche et al. (2005) used STEC for the interpolation as a means of allowing for the integral of N_e^2 to be along the signal path. However, as N_mF_2 is the peak electron density and Brunner and Gu (1991) declared that the shape factor hardly varies with elevation, Petrie et al. (2010b) used VTEC, as did Pireaux et al. (2010).

However, a final note is that the interpolation results in negative values of N_mF_2 for values of VTEC of a few TECU and below. While VTEC values of this size result in a completely negligible I3 correction, this method may not be a suitable method of estimating N_mF_2 for other purposes. A serious limitation of this method is that it effectively assumes the profile or 'slab thickness' of the electron density is constant. This assumption would not be true in various cases, particularly during storm conditions (see e.g. Jakowski et al. (1990), Jakowski et al. (1991)).

Alternatively, an estimate of N_mF_2 can be obtained from an ionospheric model such as the IRI. While this method allows for more geographical variability in the shape of the ionosphere rather than a constant ratio between N_mF_2 and VTEC, the IRI is a climatic model with monthly median values, so would not represent actual daily departures from the median due to e.g. storm conditions. Petrie et al. (2010a) use both N_mF_2 and VTEC from the IRI2007 to obtain a scale factor to estimate the bending corrections (see Eqns. 28-29). However, they consider their study to demonstrate the potential size of the bending corrections rather than produce accurate corrections.

The interpolation of N_mF_2 means that the only other quantity that is necessary for the third order term is the STEC, which can be obtained by either of the methods described above in sections 6.2.1 and 6.2.2.

6.4.2. h_mF_2

An estimate of h_mF_2 (the height at which N_mF_2 is located) is needed if applying bending terms using the formulae of Hoque and Jakowski (2008b). Recognising that accurate values for h_mF_2 and HF_2 (or N_mF_2) are not easily obtained, they suggest using constant values of 350 and 70km. However, h_mF_2 can vary significantly (Fig. 18) leading to potentially considerable effects on the bending corrections (Fig. 22). While the values suggested by Hoque and Jakowski (2008b) are a good compromise, the combined LC bending correction, Δb_{comb}, varies from ~9-20 mm at ten degrees elevation for a range of values of h_mF_2 and HF_2 similar to those shown above (Petrie et al., 2010a). Due to this variation, Petrie et al. (ibid) chose to use values of h_mF_2 from the IRI2007 to implement the bending corrections in their study, which applied the suggested equations of Hoque and Jakowski (2008b) to a global GPS network. Another recent study (Gulyaeva, 2009) compared values from IRI 2007 for h_mF_2 to values from topside sounding electron density profiles and found general agreement.

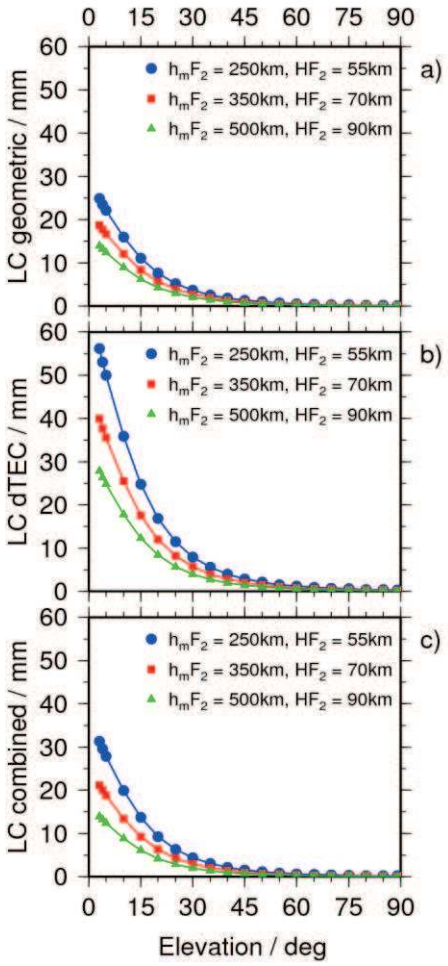


Fig. 22 Sensitivity of the bending terms to changes in HF_2 and h_mF_2. a) Geometric, b) dTEC, c) combined (for the phase LC observable). Calculated for $VTEC$ = 150 TECU. (Reproduced from Petrie et al., 2010a)

6.5. Summary of implementation choices

A summary of implementation choices in the most relevant literature can be found in Table 6.

Three of the major scientific GPS processing software packages are Bernese, GAMIT, and GIPSY-OASIS. Currently the I2 and I3 corrections have been implemented in versions of Bernese (Fritsche et al., 2005) and GAMIT (Petrie et al., 2010b), but are not yet generally released. GIPSY-OASIS can correct for I2 using an altered linear combination (after Brunner and Gu (1991)), which removes the I2 effect but results in variable carrier phase ambiguities.

Table 6 Summary table of implementation choices in the most relevant literature

Paper	Thin shell height	2nd order	Third order	Bending	IGRM/ dipole	Simulated/ real data
Hartmann & Leitinger (1984)	400km centre of slab (p126)	As error	As error	As error	n/a	Estimating errors
(Leitinger and Putz, 1988)	400km as example	no	no	Yes	n/a	theory
Brunner & Gu (1991)	Hi=hm+50km (L&P1988)	yes	yes	yes	IGRF 1985	3D raytracing
Bassiri & Hajj (1993)	300km	yes	yes	no	dipole	Models, chapman layer vs thin shell

Strangeways (2002)	suggest elevation dependent height	full Appleton-Lassen formula			dipole	raytracing
Odijk (2002)	350km	yes	yes	yes	dipole	example calculations
Kedar et al. (2003)	400km	yes	no	no	dipole	real data – equatorial sites, PPP
Fritsche et al. (2005)	400km (pers com. 450km)	yes	yes	no	dipole	real data, global network
Munekane (2005)	use thin shell, no height specified	yes	no	no	IGRF 2000	semi-analytical simulation
Datta-Barua et al. (2006)	350km (magnetic field value) 100km slab	yes	yes	no	IGRF-10	bounding errors using WAAS data
Hoque & Jakowski (2006)	evaluated along signal	yes	no	no	IGRF	One reference location in Germany
Hernández-Pajares et al. (2007)	450km	yes	no	no	IGRM Geopack	real data, simulated global network
Hoque & Jakowski (2008a)	investigate average values	yes	no	no	IGRF 2000	simulation study, Europe
Hoque & Jakowski (2008b)	400km, compare along raypath	yes	yes	yes	IGRF-8	2D raytracing compare CHAMP profiles
Morton et al. (2009b)	along raypath	yes	yes	no	IGRF-10	investigating effects using incoherent scatter radar data at Arecibo
Morton et al. (2009a)	along raypath	yes	yes	no	IGRF-10	simulation at Arecibo of position effects
(Pireaux et al., 2010)	thin shell height not specified	yes	yes	no	IGRM Geopack	example studies for time transfer
Palamartchouk (2010)	variable, based on IRI2007	yes	yes	no	IGRF 2005	simulated even global network
Petrie et al. (2010b)	450km	yes	yes	no	IGRF-10	real data, global network
Hoque & Jakowski (2010)	along raypath	yes	yes	yes	IGRF 2000	effects on radio occultation for 2 CHAMP events
Petrie et al. (2010a)	450km	yes	yes	yes	IGRF-10	real data, global network

7. Effects

Having considered the theory, approximations and implementation methods, this section now reviews the main effects seen in the GPS solutions when some or all of the higher-order ionospheric terms mentioned above are implemented. Different studies have looked at different ionospheric conditions for different purposes. A review of their results below can give a feeling for the general size of the corrections.

7.1. On signals

7.1.1. Maximum size

Datta-Barua et al. (2006; see also Datta-Barua et al., 2008) specifically aim to bound the size of the higher order ionospheric errors – (neglecting bending). With this in mind they choose all parameters to produce larger errors. The value of the magnetic field is taken at 350km, with a shape factor of 1 used for the I3 term. For a day with very active ionospheric conditions (29 October 2003), they find maximum errors for phase of up to **45mm for I2** and **35mm for I3**. Combined, this results in up to 80mm error for the LC phase observable, and up to ~180mm for the LC group observable. While their study is limited to some extent geographically by the data (WAAS 'supertruth data'), these values are unlikely to be exceeded frequently. If we consider where the majority of points lie, rather than their maximum values, the magnitudes become approximately: I2 (phase) < 15mm, I3 (phase) < 5mm, LC (phase) < 20mm, LC (group) < 40mm. The maximum values at 30 degrees elevation are I2 ~ **25mm**, I3 ~ **15mm**, LC (phase) ~ **40mm**.

If we compare these values with others in the literature, Hartmann and Leitinger (1984) were not dealing specifically with GPS signals, and considered higher-order ionospheric terms as residual errors. However, they do attempt to quantify these errors at various frequencies above 100MHz. The closest frequency to GPS in the paper is 2GHz, but if the results are adjusted to the L2 frequency, their equivalent (worst case) estimates for vertical incidence are: I2 (ΔSa) = **±86mm**, I3 (ΔSb) = **-7mm** and discarded term (ΔSc) = -10.6 x 10^{-2}mm. This is based on a VTEC = 200 TECU, f_g = 1.74 MHz (equivalent to ~ 62143nT, which is quite large), and a slab of uniform electron density 200km thick (with shape factor = 1)

When considering non vertical incidence then the height of the slab and transmitter become relevant, and geometric bending is also a factor. The following worst case results are for an elevation angle of 30 degrees, with the electron density centred around 400km and a transmitter height of 1000km: I2 (ΔSa) = **±130mm**, I3 (ΔSb) = **-14mm**, b_{gm} (ΔSd) = **-21mm**. The larger value of I2 compared to the 30 degree value in Datta-Barua et al. (2006) is probably due in part to the large magnetic value but may also be due to calculating for all values of θ or to a higher VTEC. For I3 the value is similar, which suggests a higher VTEC is compensating for the effect of the thicker slab (which means a lower value of N_mF_2).

Brunner and Gu (1991) was the first study to consider GPS specifically. They compared results from their suggested model with results for rigorous raytracing technique. Their results are shown as residual range error (RRE) for LC in relation to their raytracing results, for both the normal 'ionosphere-free' linear combination and their improved model which includes curvature of the ray path. They look at two cases, Case 1 with very high ionospheric activity and Case 2 with extremely high ionospheric activity. Case 1 has a VTEC of 138 TECU and

peak electron density of 6.0 x 10^{12} electrons m^{-3}, while Case 2 has a VTEC of 455 TECU and peak electron density of 20 x 10^{12} electrons m^{-3}. For vertical incidence, the LC (phase) RRE values are: Case1, **3mm**, Case 2, 23.6 mm, while for an elevation angle of 30 degrees they are: **19.6mm**, 99.4 mm. It is worth noting that while the conditions of Case 1 may be reached in equatorial regions during ionospheric solar maximum, those of Case 2 are so high as to be unlikely to be seen even at ionospheric maximum. The Case 1 RRE value at 30 degrees elevation is roughly half that estimated by Datta-Barua et al. (2006), probably due to the VTEC being lower and a smaller contribution from I3, with a shape factor now at 0.66 rather than 1.

Bassiri and Hajj (1993) STEC 100 TECU **I2: ~-16.5mm** (L2 phase), **ΔI2 ~5.5mm** (phase) I3: **~-0.8mm** (L2 phase), ΔI3 ~0.2mm (phase). If we multiply by 1.5 to increase the STEC to 150TECU, this becomes **I2 ~25mm** (L2 phase).

Odijk (2002) examines the potential size of the corrections for the area of the Netherlands, using a maximum value of 100TECU and a shape factor of 0.66. He suggests maximum magnitudes at ten degrees elevation of I2: ~ 43mm (L2 phase) and 48mm (L5 phase); I3: ~2mm (L2 phase) and ~2.3mm (L5 phase); b_{gm}: ~13mm (L1 phase) and ~16mm (L5 phase) in addition to estimating details of likely effects on baselines up to 400km.

Jakowski et al. (1994) estimate a geometric bending effect at 10 degrees elevation and VTEC of 100 TECU of ~10mm (L2 phase). Hoque and Jakowski (2007) estimate ΔI2 geographically as up to ~±10mm again using a VTEC of 100VTEC and 90 degree elevation. Hoque and Jakowski (2008b) estimate (for a global VTEC of ~143TECU and elevation of 5 degrees) ΔI2 to be up to ±25mm, ΔI3 up to ~3mm, Δb_{TEC} up to ~40mm. They note that b_{gm} is "higher for a thin profile with large maximum ionisation than for a thick profile with small maximum ionisation". For varying profile shapes (all with VTEC=100TECU), they find a maximum Δb_{gm} of 17mm. When considering higher order effects on radio occultation measurements, Hoque and Jakowski (2010) find values of up to 50mm for ΔI2 and 12mm for ΔI3. Once the effects of bending are included, they find a combined maximum RRE of 1187mm. In addition, they find that the separation of the signal paths may be over a kilometre.

7.1.2. Geographic distribution

Hoque and Jakowski (Hoque and Jakowski, 2008b) plot geographic patterns of the ΔI2 term with a constant global VTEC for various azimuths, revealing the strong hemispheric asymmetry. The ΔI2 term is predominantly positive in the northern hemisphere and negative in the southern hemisphere. In contrast, Fig. 23 shows how the actual geographical pattern of VTEC for a day during ionospheric maximum is reflected in the patterns for the combined Δb term and for the ΔI2 term.

While the values are not 'maximum values', they are probably close to the highest that could be expected for the last solar maximum (solar cycle 23). It is worth noting that solar cycles do vary in magnitude, however (Fig. 1). When applying the bending terms of Hoque and Jakowski (2010b) to a global GPS network for a sample day with high TEC, Petrie et al. (2010a) note that it is only in equatorial regions that the effect on the combined LC phase reaches over 3 mm.

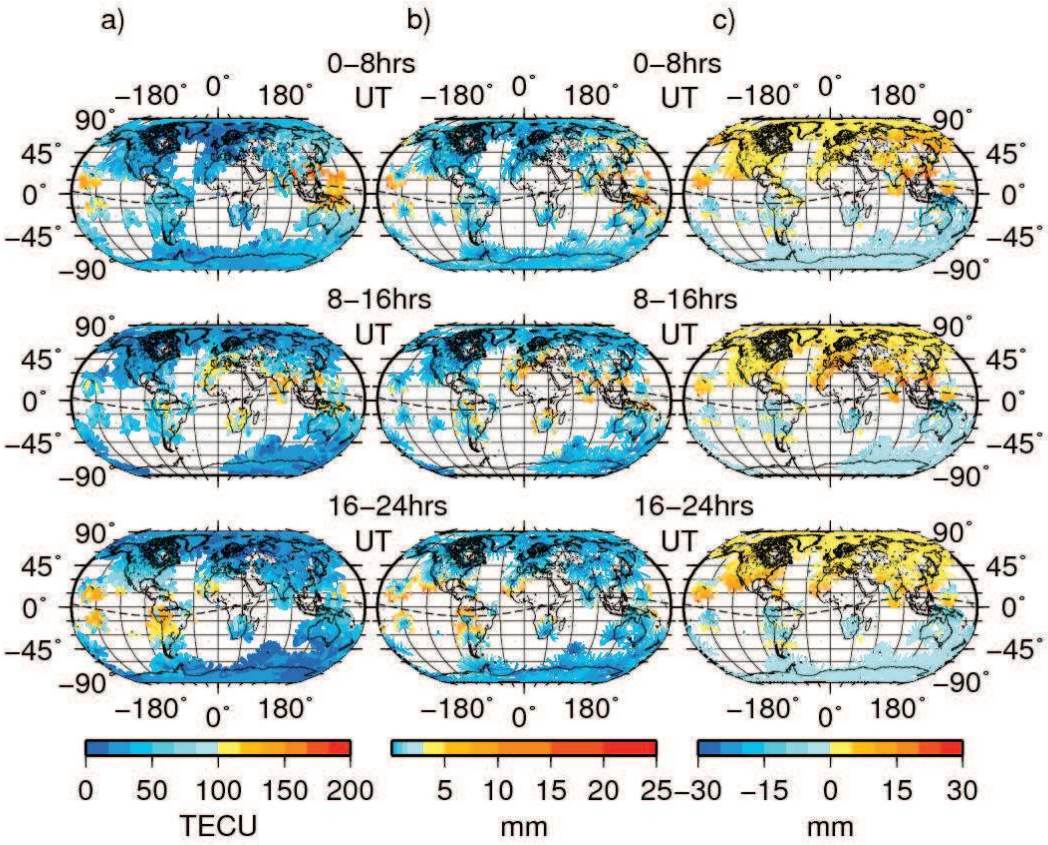


Fig. 23 a) Global distribution of VTEC (TECU), b) combined bending effect on the LC observable (mm) and c) second order ionospheric effect on the LC observable (mm). Data are for one day during ionospheric maximum (DOY 301, 2001) showing values used/modelled in the GPS processing. Values are estimated and plotted geographically at the point where the GPS signal crosses a 450km high 'shell' around the Earth. VTEC data are from IONEX files provided by CODE. (Reproduced from Petrie et al., 2010a)

7.2. On positions and other estimated parameters

7.2.1. Positions and rates

The first paper to consider the potential effects of the I2 term on receiver coordinates was that of Kedar et al. (2003) who found that all stations move southwards when the I2 term is modelled. However, as they noted, their precise point positioning analysis forces the corrections to be absorbed solely by receiver parameters, with satellite orbits and clocks held fixed. Their finding was later revised by the study by Fritsche et al. (2005), which estimated both receiver and satellite parameters. Fritsche et al. (2005) find that the majority of the effect is taken up by the satellite orbits. This equates to a translation of the frame origin in the z-direction (discussed in section 7.3). They find that stations in the southern hemisphere still appear to move southwards, but those in the northern hemisphere remain fixed (Fig. 24).

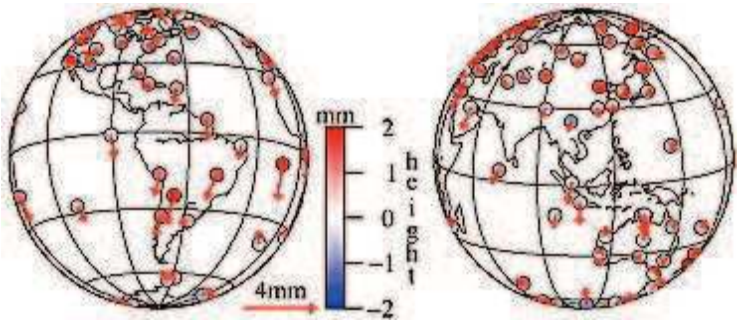


Fig. 24 Residuals after a coordinate transformation (3 translations) on 2 April 2000. A standard solution (1st–order ionospheric term only) is compared with a modified solution in which the 2nd and 3rd–order corrected observations are used to adjust all relevant parameters (including orbits, ERPs, tropospheric delays etc). Reproduced from Fritsche et al., 2005)

A second paper includes a figure of mean coordinate differences over a three year period (2001.0-2004.0) (Steigenberger et al., 2006) reproduced here as Fig. 25.

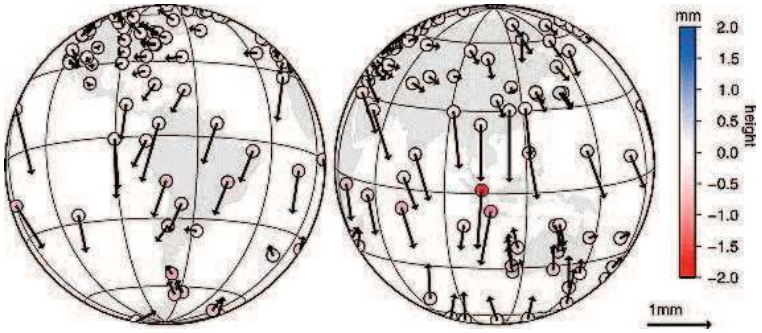


Fig. 25 Mean effect of higher order ionospheric (I2+I3) corrections on global station positions from 3 years GPS observation (2001.0-2004.0). The arrows indicate a predominant southward shift of horizontal station positions. The colour scale indicates the changes in station heights (Reproduced from Steigenberger et al., 2006)

However, a later paper by Hernandez-Pajares et al. (2007) found a slightly altered pattern, with high latitude sites moving north and mid latitude sites moving south (Fig. 26, Fig. 27). They suggested that the difference to the finding of Fritsche et al. (2005) could be attributed either to the fact that they used STEC from GPS signals instead of from GIMs as the earlier paper did, or because their station network was much more globally balanced than that of Fritsche et al (2005), which had many more stations in the northern hemisphere.

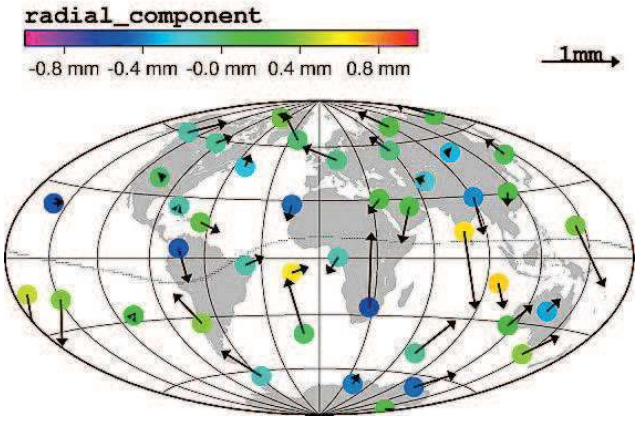


Fig. 26 Mean I2 effect on receiver positions for 21 months of GPS observations during ionospheric maximum (DOY 100 of year 2002 to DOY 365 of year 2003). Only receivers with more than 200 days of data are represented (Reproduced from Hernández-Pajares et al., 2007)

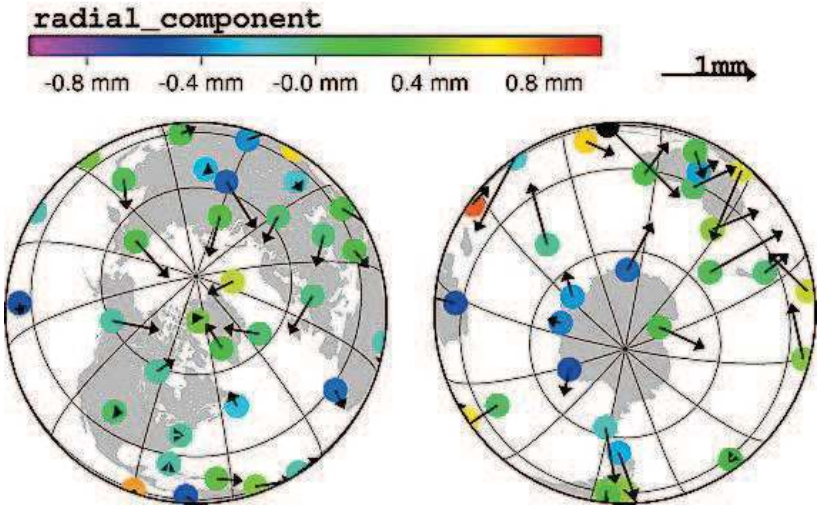


Fig. 27 Polar projections of data in Fig. 26. "In this case, receivers with at least 100 days of data are plotted. The northward shift of the high-latitude receivers is confirmed" (Reproduced from Hernández-Pajares et al., 2007)

The study by Petrie et al. (2010b), over a longer time period, revealed a similar pattern of coordinate shifts (Fig. 28) to that of Hernandez-Pajares et al. (2007). This suggests that it is the balanced site network rather than the method of obtaining STEC that results in the different pattern, as Petrie et al. (2010b) use a similar method to Fritsche et al (2005).

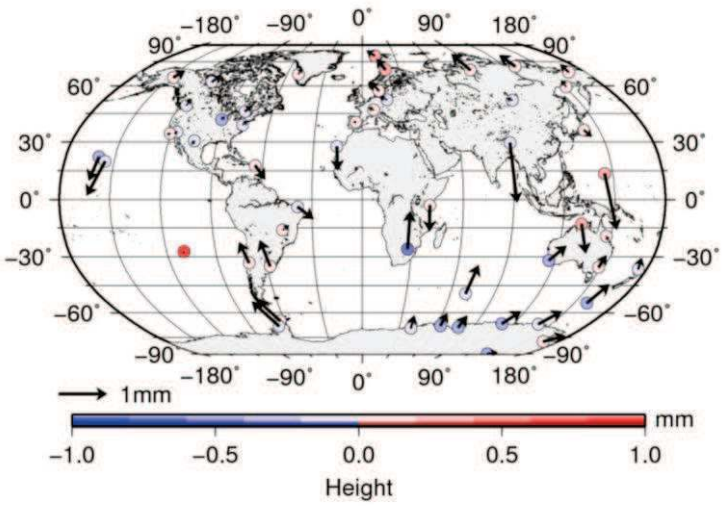
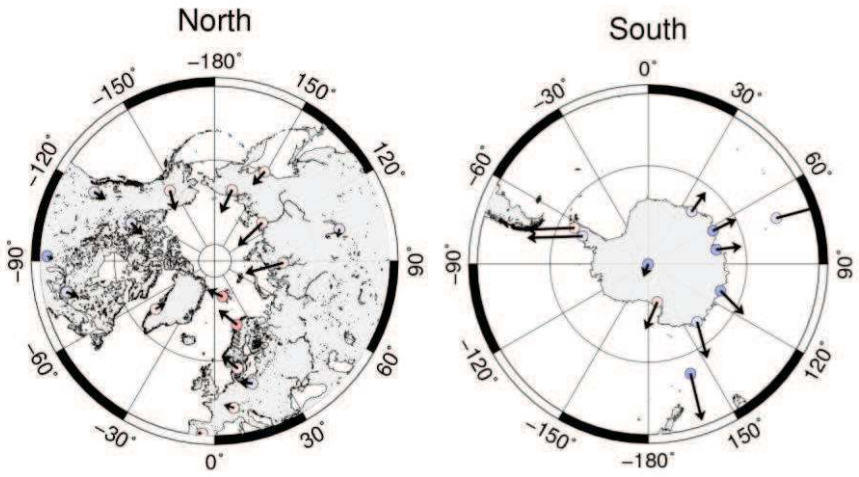


Fig. 28 Mean coordinate differences due to modelling second and third order ionospheric effects (2001.0-2004.0). Sites shown have at least 2.5 years data. Top: Robinson projection, bottom polar projections. Data previously published in Petrie et al. (2010b)



I2 + I3 effects on GPS vertical rates are estimated by Petrie et al. (2010b) at less than 0.4mm/year in magnitude. The sign depends upon whether solar activity is rising or falling (Fig. 29).

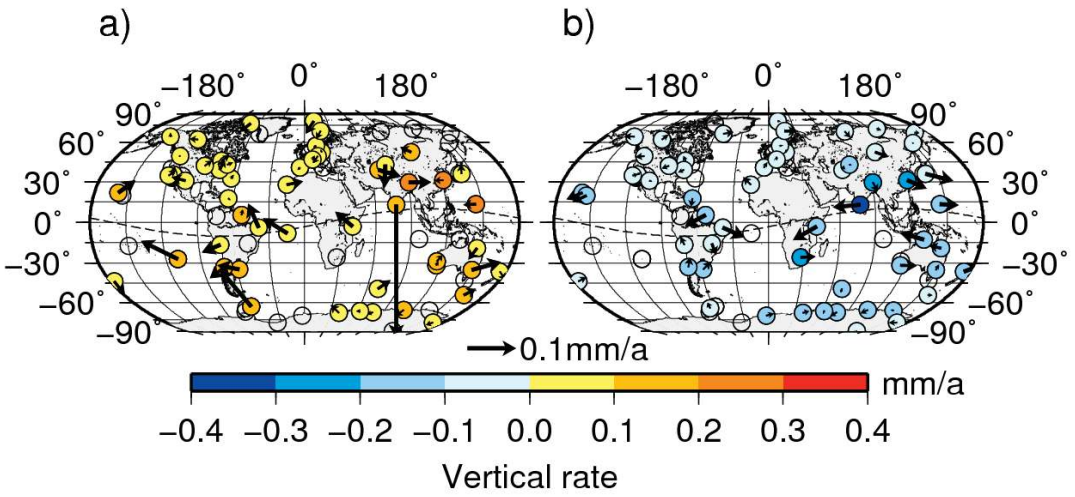


Fig. 29 Velocity differences (run with no higher order corrections minus run with I2 and I3 corrections). a) Velocity differences over the period 1996.0-2001.0. b) Velocity differences over the period 2001.0-2006.0. Geomagnetic equator is shown as dashed line and arrows represent motion in plan. Sites shown have data spanning at least 4.5 years of the 5 year period, with a minimum of 2.5 years of data. Empty circles show sites processed which did not meet these criteria. (Reproduced from Petrie et al., 2010b)

The only study to date of potential bending effects on positions is Petrie et al. (2010a). While the study used the IRI2007 to estimate h_mF_2 and N_mF_2, meaning the corrections were rough approximations, "For the three years spanning ionospheric maximum (2001-2003), the mean difference at any site is less than ±0.3mm", which suggests bending corrections can reasonably be neglected at present. This result would appear to be due to the similar elevation dependence of the LC bending terms and the tropospheric mapping functions.

Sub-daily effects have also been considered by a limited number of authors. Due to the rotation of the Earth with respect to the sun, the peak of ionospheric activity moves around the world and occurs close to 2pm local time (visible in Fig. 23). Hernandez-Pajares et al. (2007) look at subdaily effects during ionospheric maximum on differential positioning of equatorial receivers and find shifts of up to several millimetres in position. Palamartchouk (2010) shows examples of simulated displacements at 04UT and 16UT on the 1 Jan 2000. However, over longer periods coordinate shifts due to modelling higher order ionospheric effects appear limited to less than ~1.5mm. Morton et al. (2009a) also simulate subdaily position changes at Arecibo, in Puerto Rico, and find effects of up to 2cm based on direct error computation, but suggest this would be somewhat reduced when using LC, as the higher order effects are partially cancelled.

7.2.2. Orbits and clocks

The majority of the I2 effect appears to be absorbed by satellite clocks and orbits (Hernández-Pajares et al., 2007). Due to the contrast of day and night time effects, the orbit will "present a northward displacement for dayside observations, when the satellite is at high latitudes, and a southward shift when the satellite is, for instance, at low latitudes or nightside observations" (Hernández-Pajares et al., 2007), see Fig. 30.

They also report that "the I_2 effect on satellite clocks is typically the greatest one: it can be larger than 1 centimeter, i.e., comparable with the reported IGS clock estimations accuracy".

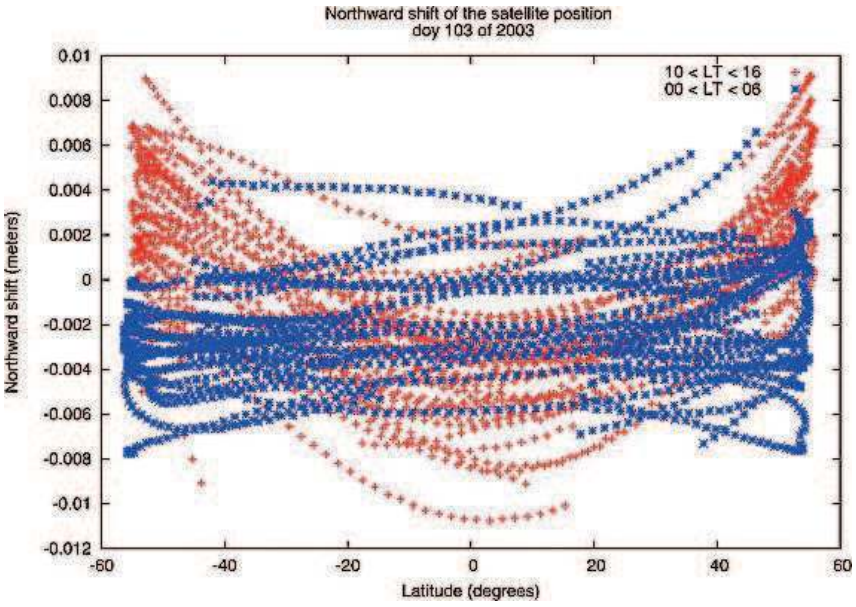


Fig. 30 Representative example of the effect of I2 on the north component of the satellite position. (DOY 103 of year 2003). All satellites basically experience a southward shift, except for those in high latitudes during day time. (Reproduced from Hernández-Pajares et al., 2007)

Pireaux et al. (2010) consider higher order effects on geodetic time and frequency transfer and find effects from I2 of up to ~10ps for receivers in common view, but that I3 can be neglected.

7.3. Reference frame parameters

The transformation to change coordinates in one reference frame into coordinates in another is usually performed using a seven-parameter transformation, also known as a Helmert transformation. The seven parameters are: three translations, one along each of the X, Y and Z axes; three rotations, one around each axis; and a scale factor. The average transformation between two sets of coordinates obtained e.g. by processing with and without the I2 and I3 effects modelled can thus be described in terms of these seven parameters. Below we review the findings of several authors who have investigated these transformations relating to higher order ionospheric effects.

Munekane (2005) performed a semi-analytical analysis and found an annual oscillation of 0.1ppb in scale, with a z-translation offset of 1.8mm and semi-annual oscillation of 1mm. Fritsche et al. (2005) analysed three years of data during ionospheric maximum and showed an effect of up to ~12mm on the z-translation, with similar but smaller effects upon the x and y-translations (Fig. 31). Effects on scale were found to be negligible at ~ -0.1ppb.

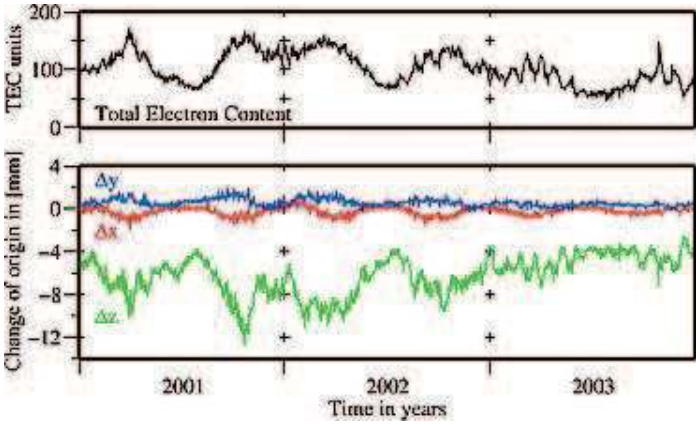


Fig. 31 Top panel: time series of maximum *TEC* values (daily means). Bottom panel: origin component differences (frame without minus frame with higher–order corrections). The differences reveal mean biases of $\Delta x \approx -0.2$ mm, $\Delta y \approx +0.6$ mm and $\Delta z \approx -6.2$ mm for the period 2001.0-2004.0. The temporal representation of the differences is obviously correlated with the *TEC* variations. (Reproduced from Fritsche et al., 2005)

Petrie et al. (2010b) distinguished between the effects of the I2 and I2+I3 corrections on the z-translation and scale and found that the I3 term has very little effect (Fig. 32a, b). Petrie et al. (2010a) extended the comparison to the bending terms and showed that they have a similarly small effect (Fig. 32c, d), leaving the I2 term as the major contributor.

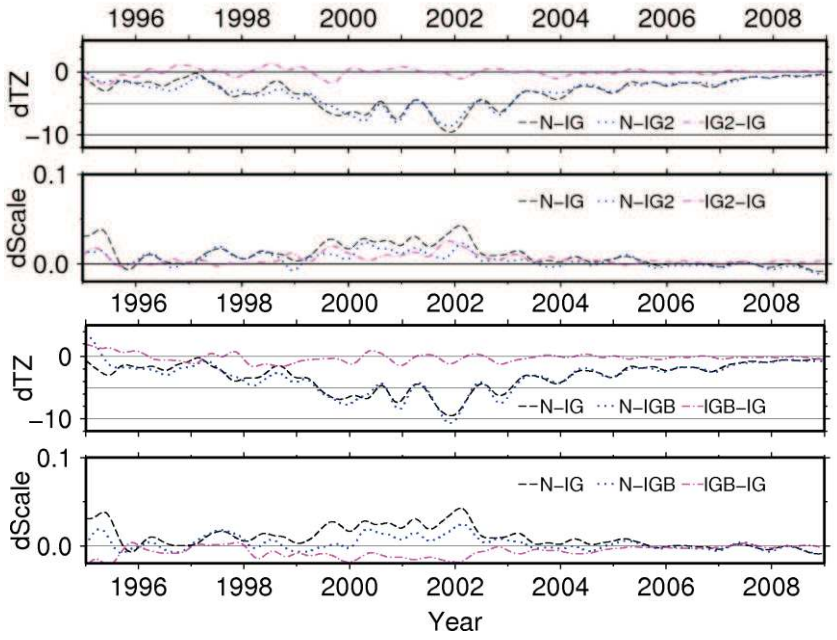


Fig. 32 Z-translation and scale differences when modelling I2 (IG2), I2+I3 (IG) and I2+I3+Bending (IGB).

Top panels: Reproduced from Petrie et al. (2010b). Bottom panels: Reproduced from Petrie et al. (2010a).

Palamartchouk (2010) performed an interesting simulation study with a homogeneously spaced GPS network and the I2 and I3 corrections, modelled using the IGRF v10 and the IRI2007 model. The simulation method used considers the orbits as perfectly known and fixed. As a simulation requires no actual GPS data, the time span could be greatly extended over several solar cycles (1970-2008). The study found that "The z-component of the degree-1 deformation shows long-term changes with the amplitude of about 4 mm and long-term mean of 4.1 mm." and also that "The amplitude of oscillations in z direction reach 2.5 mm for the semi-annual and 3.5 mm for 11-year components when satellite orbits are considered and degree-1 parametrization is used". Oscillations in the z direction found using the Helmert transformation method were even smaller. The effects are considerably smaller than those found by both Fritsche et al. (2005) (Fig. 31) and Petrie et al. (2010b). This may be due to the fixed orbits used in the study or to the homogeneously spaced network, as studies using real data lack sites in ocean regions and tend to be biased at least to some degree towards the Northern Hemisphere. Alternative reasons may be the use of the variable IRI h_mF_2 values for the thin shell height, or the IRI model may represent extremes of electron content poorly. Palamartchouk (2010) also found that "The scale parameter also shows pronounced diurnal oscillations with amplitude of order 0.04 ppb, experimentally undetectable at the moment". Another interesting finding of the study is that using elevation dependent weighting reduces the effects of modelling the second order correction (Fig. 33). However, this weighting was also performed by Petrie et al. (2010b), so cannot explain the difference.

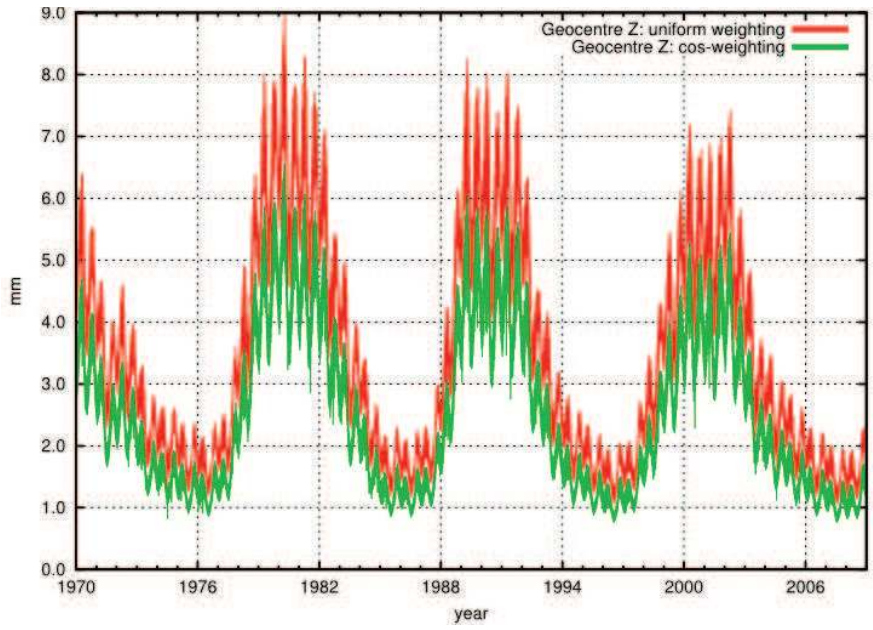


Fig. 33 Simulated degree-1 z-deformation when modelling second order ionospheric effects. (Reproduced from Palamartchouk, 2010)

Rotations were described as negligible by Fritsche et al. (2005) while Palamartchouk (2010) found that "The rotation parameters did not show any dependence on the phase of the solar cycle except that the amplitude of the subdaily oscillations is higher during the solar maximum".

8. Summary

8.1. Current state of the art

The I2 term affects satellite clocks by up to 1cm, and time and frequency transfer for stations in common view by up to ~10ps. Satellite positions are affected by up to a few mm while effects on site coordinates positions are limited to less than ~1.5mm, though subdaily effects may be larger. Rates are affected by up to ~0.3mm/yr. If reference frame parameters are of interest, the z-translation is the most strongly affected, by up to ~12mm. The x and y translations are also affected, but to less than 2mm and the effects on scale are limited to ~0.1ppb. The maximum effects described above are generally based around the top end of the electron densities seen in the last solar cycle (cycle 23). As solar cycles vary in intensity, it is possible that they could be exceeded if there is a very intense solar cycle in future. To summarise, the I2 correction is relevant to precise GPS analysis, though it is not yet performed routinely.

The I2 correction is commonly implemented using the thin shell approximation at a single fixed height, with STEC either estimated from the signals or from maps of VTEC. Using STEC from GPS should be of benefit in real time applications, for dates pre1998 (when IONEX files have less resolution) and in remote areas with limited amounts of GPS data. However, in general using IONEX VTEC files appears to give broadly similar results.

While the dipole model is a simpler representation of the Earth's geomagnetic field than the IGRF, it is not necessarily simpler to implement. While the resulting coordinate differences are not large, the IGRF or similar magnetic model is probably the preferred option.

The I3, b_{gm} and b_{TEC} terms are both smaller than the I2 term and somewhat more complex to implement. While the N_mF_2 interpolation method of Fritsche et al. (2005) was a reasonable approach to obtaining rough estimates of N_mF_2 for long term global GPS processing, it is far from ideal. However, current publically available ionospheric data with a global geographic and sub-daily temporal scale are not sufficient. At present, given the size of the bending terms and the errors in implementing them, further research is required before implementation can be recommended.

8.2. Future developments

In future it will become feasible to determine the free electron density distribution at global scale thanks to large constellations of GNSS receivers on board low Earth orbit satellites (LEOS) augmenting the present FORMOSAT-3/COSMIC constellation). As the ionospheric data available becomes more detailed, it may be possible to look at the impact of variable h_mF_2 height on the appropriate height at which to evaluate the magnetic field effect upon the signal. Eventually, integrating both the magnetic field and the electron content along the signal may become practical. Further investigation into mapping function errors and the differences on a global scale between obtaining STEC from GIMs of VTEC and direct STEC derivation using code alignment would also be of interest.

While the external components of the magnetic field are small compared to the main field, if a model is available for the required date range, they could be included.

To date, studies of ionospheric bending effects upon GPS have used spherically symmetric ionospheres. As the ionosphere has a distinctly unspherical equatorial anomaly, an investigation of its effects, perhaps by a raytracing study would be of interest.

In terms of the effect on transformation parameters and particularly on the z-translation, an investigation into the network dependence of the effects could be of interest. For example, what is the effect on a network consisting of the IGS reference frame sites in comparison to a uniform network?

The desire for precise near-real time processing for applications such as tsunami warning systems (Blewitt et al., 2009) could lead to second-order effects becoming of interest, particularly as solar activity will now increase towards the next ionospheric maximum.

While in this paper we have reviewed the steps taken to deal with higher order corrections to dual frequency GPS, developments to GPS and other positioning systems such as Galileo mean that three or more frequencies will become available in future. An early paper presenting a theoretical analysis for Galileo was written by Groves and Harding (2003), while papers dealing with the first data for Galileo are beginning to appear (Fleury et al., 2009). Technical details of dealing with three frequency combination have also been assessed (e.g. Odijk, 2003; Wang et al., 2005), though noise increase is an issue (IERS, 2009). However, while the increase in frequencies will bring new possibilities, simply dealing with the archive of current data means the second order correction will certainly remain a concern for some time to come.

Acknowledgements

The authors would like to thank the Editor and the two anonymous reviewers for their thoughtful reviews. Elizabeth Petrie was supported by the Natural Environment Research Council (NERC) and would like to thank Cathryn Mitchell for assistance with the literature on magneto-ionic theory and Ian Thomas for comments which improved the clarity of Section 3. Matt King was supported by NERC and RCUK Academic fellowships.

References

Anderson, D, Fuller-Rowell, T. (1999) The Ionosphere, Space Environment Topics, (SE-14), [Online]. Available at: http://www.swpc.noaa.gov/info/Iono.pdf . Accessed: 10 Jan 2010

Appleton EV (1932) Wireless studies of the ionosphere. Wireless Section, Institution of Electrical Engineers - Proceedings of the 7 (21):257-265

Balanis, CA. (2005) Antenna Theory - Analysis and Design (3rd Edition) [Online]. Available at: http://knovel.com/web/portal/browse/display?_EXT_KNOVEL_DISPLAY_bookid=1955&VerticalID=0. Accessed: 7 Sept 2009

Bassiri, S, Hajj, GA (1993) Higher-order ionospheric effects on the global positioning system observables and means of modeling them. Manuscr Geod 18: 280-289

Bilitza, D, Reinisch, BW (2008) International Reference Ionosphere 2007: Improvements and new parameters. Adv Space Res 42 (4): 599-609

Blewitt, G, Hammond, WC, Kreemer, C, Plag, HP, Stein, S, Okal, E (2009) GPS for real-time earthquake source determination and tsunami warning systems. J Geod 83 (3-4): 335-343

Born, M, Wolf, E (1999) Principles of optics - Electromagnetic theory of propagation interference and diffraction of light. Cambridge University Press, Cambridge

Bouin, MN, Wöppelmann, G (2010) Land motion estimates from GPS at tide gauges: a geophysical evaluation. Geophys J Int 180 (1): 193-209

Brunini, C, Azpilicueta, F (2009) Accuracy assessment of the GPS-based slant total electron content. J Geod 83 (8): 773-785

Brunner, FK, Gu, M (1991) An improved model for the dual frequency ionospheric correction of GPS observations. Manuscr Geod (16): 205-214

Budden, KG (1985) The propagation of radiowaves: the theory of radio waves of low power in the ionosphere and magnetosphere. Cambridge University Press, Cambridge

Burrell, A, Bonito, N, Carrano, C (2009) Total electron content processing from GPS observations to facilitate ionospheric modeling. GPS Sol 13 (2): 83-95

Bust, GS, Mitchell, CN (2008) History, current state, and future directions of ionospheric imaging. Rev Geophys 46: RG1003, doi:10.1029/2006RG000212

Calais, E, Mattioli, G, DeMets, C, Nocquet, JM, Stein, S, Newman, A, Rydelek, P (2005) Seismology: Tectonic strain in plate interiors? Nature 438 (7070): E9-E10

Cander, LR (2008) Ionospheric research and space weather services. J Atmos Sol Terr Phys 70 (15): 1870-1878

Ciraolo, L, Azpilicueta, F, Brunini, C, Meza, A, Radicella, S (2007) Calibration errors on experimental slant total electron content (TEC) determined with GPS. J Geod 81 (2): 111-120

Ciraolo, L, Spalla, P (1997) Comparison of ionospheric total electron content from the Navy Navigation Satellite System and the GPS. Radio Sci 32 (7): 1071-1080

CODE (2007) Global Ionospheric Maps from the Centre for Orbit Determination in Europe. http://aiuws.unibe.ch/ionosphere/mslm.pdf. Accessed: 10 Feb 2010

Coïsson, P, Radicella, SM, Ciraolo, L, Leitinger, R, Nava, B (2008) Global validation of IRI TEC for high and medium solar activity conditions. Adv Space Res 42 (4): 770-775

Datta-Barua, S, Walter, T, Blanch, J, Enge, P (2006) Bounding Higher Order Ionosphere Errors for the Dual Frequency GPS User. ION GNSS 19th International Technical Meeting of the Satellite Division. Fort Worth, TX, 26-29 September 2006

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

Davies, K (1990) Ionospheric Radio. Peter Peregrinus Ltd, London

Doherty, P, Coster, AJ, Murtagh, W (2004) Space weather effects of October–November 2003. GPS Sol 8 (4): 267-271

Feltens, J (2003) The activities of the Ionosphere Working Group of the International GPS Service (IGS). 7 (1): 41-46

Fleury, R, Clemente, M, Carvalho, F, Lassudrie-Duchesne, P (2009) Modelling of ionospheric high-order errors for new generation GNSS. Ann Telecomm 64 (9): 615-623

Fowler, CMR (1990) The solid Earth. Cambridge University Press, Cambridge

Fritsche, M, Dietrich, R, Knofel, C, Rülke, A, Vey, S, Rothacher, M, Steigenberger, P (2005) Impact of higher-order ionospheric terms on GPS estimates. Geophys Res Lett 32: , L23311, doi: 10.1029/2005GL02434223

Gherm, VE, Novitsky, R, Zernov, N, Strangeways, HJ, Ioannides, RT (2006) On the limiting accuracy of range measurements for the three-frequency mode of operation of a satellite navigation system. 2nd Workshop on Radio Systems and Ionospheric effects, COST action 296 Mitigation of ionospheric effects on radio systems (MIERIS). Rennes, 3-7 October 2006

Groves, PD, Harding, SJ (2003) Ionosphere propagation error correction for Galileo. J Navig 56: 45-50

Gu, M, Brunner, FK (1990) Theory of the two frequency dispersive range correction. Manuscr Geod 15: 357-361

Gulyaeva, TL (2009) Linkage of the ionospheric peak electron density and height deduced from the topside sounding data. Adv Space Res 43 (11): 1794-1799

Hartmann, GK, Leitinger, R (1984) Range errors due to ionospheric and tropospheric effects for signal frequencies above 100 MHz. Bull Geod 58: 109-136

Hartree DR (1931) The propagation of electromagnetic waves in a refracting medium in a magnetic field. Proc Camb Phil Soc 27:143-162. doi:10.1017/S0305004100009440

Hawarey, M, Hobinger, T, Schuh, H (2005) Effects of the 2nd order ionospheric terms on VLBI measurements. Geophys Res Lett 32: , L11304, doi 10.1029/2005GL022729

Hecht, E (1998) Optics. Addison Wesley Longman, New York

Hernández-Pajares, M (2004) IGS Ionosphere WG status report: performance of IGS Ionosphere, TEC Maps. IGS Technical Meeting, Bern, Switzerland.

Hernández-Pajares, M, Juan, J, Sanz, J, Orus, R, Garcia-Rigo, A, Feltens, J, Komjathy, A, Schaer, S, Krankowski, A (2008) The IGS VTEC maps: a reliable source of ionospheric information since 1998. J Geod 83 (3): 263-275

Hernández-Pajares, M, Juan, JM, Sanz, J (2005) 'Towards a more realistic mapping function', *URSI GA*. New Delhi, October 2005.

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

Herring, T (1983) The precision and accuracy of intercontinental distance determinations using radio interferometry. PhD thesis. Massachusetts Institute of Technology

Herring, TA (1999) Geodetic applications of GPS. Proc. IEEE 87 (1): 92-110

Hocke, K (2008) Oscillations of Global Mean TEC. J Geophys Res 113: A04302, doi: 10.1029/2007JA012798

Hofmann-Wellenhof, B, Lichtenegger, H, Collins, J (2001) Global Positioning System: theory and practice. Springer, Wien

Hoque, M, Jakowski, N (2007) Higher order ionospheric effects in precise GNSS positioning. J Geod 81 (4): 259-268

Hoque, M, Jakowski, N (2008a) Mitigation of higher order ionospheric effects on GNSS users in Europe. 12 (2): 87-97

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

Hoque, MM, Jakowski, N (2010) Higher order ionospheric propagation effects on GPS radio occultation signals. Adv Space Res Adv Space Res 46(2):162–173. doi:10.1016/j.asr.2010.02.013

Horvath, I, Crozier, S (2007) Software developed for obtaining GPS-derived total electron content values. Radio Sci 42: , RS2002, doi:10.1029/2006RS003452

ICD-GPS-200. (2000) Interface Control Document ICD-GPS-200 Navstar GPS Space Segment / Navigation User Interfaces.

IERS. (2009) IERS Conventions update: Chapter 9. Working version updated 16 July 2009. International Earth Rotation Service

Imel, DA (1994) Evaluation of the TOPEX/POSIEDON dual frequency ionosphere correction. J Geophys Res 99 (C12): 24895-24906

Jakowski N, Jungstand A, Lois L, Lazo B (1991) Night-time enhancements of the F2-layer ionization over Havana, Cuba. Journal of Atmospheric and Terrestrial Physics 53 (11-12):1131-1138

Jakowski N, Putz E, Spalla P (1990) Ionospheric storm characteristics deduced from satellite radio beacon observations at 3 European stations. Ann Geophys-Atmos Hydrospheres Space Sci 8 (5):343-351

Jakowski, N, Porsch, F, Mayer, G (1994) Ionosphere - Induced -Ray-Path Bending Effects in Precision Satellite Positioning Systems. Z Satell Position Navig Kommun SPN1/94: 6-13

Kamide Y, Chian AC-L (eds) (2007) Handbook of the solar-terrestrial environment. Springer, Berlin

Kedar, S, Hajj, GA, Wilson, BD, Heflin, MB (2003) The effect of the second order GPS ionospheric correction on receiver positions. Geophys Res Lett 30 (16): 1829, doi:10.1029/2003GL017639

Kelley MC (2009) The earth's ionosphere: Plasma physics and electrodynamics. International geophysics series, vol 96, 2nd edn. Academic Press, San DiegoKim, BC, Tinin, MV (2009) The association of the residual error of

dual-frequency Global Navigation Satellite Systems with ionospheric turbulence parameters. J Atmos Sol Terr Phys 71 (17-18): 1967-1973

Kintner, PM, Ledvina, BM (2005) The ionosphere, radio navigation, and global navigation satellite systems. Adv Space Res 35 (5): 788-811

Kivelson, MG, Russell, CT (eds.) (1995) Introduction to Space Physics. Cambridge University Press, New York

Klobuchar, JA (1996) Ionospheric effects on GPS. In: Parkinson, BW, Spilker, JJ (eds) Global Positioning System: Theory and Applications. Vol. 1 American Institute of Aeronautics and Astronautics Inc., Washington DC, pp 485-515

Komjathy, A (1997) Global Ionospheric Total Electron Content Mapping Using the Global Positioning System, University of New Brunswick Technical Report No. 188. PhD thesis. University of New Brunswick

Komjathy, A, Langley, RB (1996) The Effect of Shell Height on High Precision Ionospheric Modelling Using GPS. Proceedings of the 1996 IGS Workshop International GPS Service for Geodynamics. Silver Springs, MD, 19-21 March 1996.

Langel, RA (1987) The Main Field. In: Jacobs, JA(ed), Geomagnetism. Vol. 1 Academic Press Ltd, London, pp 249-267

Langley, RB (1998) Propagation of the GPS signals. In: Teunissen, PJG, Kleusberg, A(eds) GPS for Geodesy.Springer-Verlag, Berlin, pp 115-149

Larson, KM, Bilich, A, Axelrad, P (2007) Improving the precision of high-rate GPS. 112: B05422, doi:10.1029/2006JB004367

Lassen H (1927) Über den einfluss des erdmagnetfeldes auf die fortpflanzung der elektrischen wellen der drahtlosen telegraphie in der atmosphare. ENT 4:324-334

Leick, A (2004) GPS satellite surveying. Wiley, New Jersey

Leitinger, R, Jakowski, N, Davies, K, Hartmann, GK, Feichter, E (2000) Ionospheric Electron Content and Space Weather: Some Examples. Phys Chem Earth (A) 25 (8): 629-634

Leitinger, R, Putz, E (1988) Ionospheric refraction errors and observables. In: Brunner, FK(ed), Atmospheric Effects on Geodetic Space Measurements. Monograph 12, School of Surveying, University of New South Wales, pp 81-102

Liang, M-C, Li, K-F, Shia, R-L, Yung, YL (2008) Short-period solar cycle signals in the ionosphere observed by FORMOSAT-3/COSMIC. Geophys Res Lett 35: L15818, doi: 10.1029/2008GL034433

Liu, H, Stolle, C, Watanabe, S, Abe, T, Rother, M, Cooke, DL (2007) Evaluation of the IRI model using CHAMP observations in polar and equatorial regions. Adv Space Res 39: 904–909

Liu, L, Zhao, B, Wan, W, Ning, B, Zhang, M-L, He, M (2009) Seasonal variations of the ionospheric electron densities retrieved from Constellation Observing System for Meteorology, Ionosphere, and Climate mission radio occultation measurements. J Geophys Res 114: A02302, doi: 10.1029/2008JA013819

Maus, S, Macmillan, S (2005) 10th Generation International Reference Field. EOS Trans. 86 (16): 159

Maus, S, Macmillan, S, Chernova, T, Choi, S, Dater, D, Golovkov, V, Lesur, V, Lowes, F, Lühr, H, Mai, W, McLean, S, Olsen, N, Rother, M, Sabaka, T, Thomson, A, Zvereva, T (2005) The 10th-Generation International Geomagnetic Reference Field. Geophys J Int 161 (3): 561-565

Mazzella, AJ, Jr. (2009) Plasmasphere effects for GPS TEC measurements in North America. Radio Sci 44: RS5014, doi:10.1029/2009RS004186

Min, K, Park, J, Kim, H, Kim, V, Kil, H, Lee, J, Rentz, S, Luhr, H, Paxton, L (2009) The 27-day modulation of the low-latitude ionosphere during a solar maximum. J Geophys Res 114: A04317, doi: 10.1029/2008JA013881

Morton, YT, van Graas, F, Zhou, Q, Herdtner, J (2009a) Assessment of the Higher Order Ionosphere Error on Position Solutions. Navig 56 (3): 185-193
Morton, YT, Zhou, Q, van Graas, F (2009b) Assessment of second-order ionosphere error in GPS range observables using Arecibo incoherent scatter radar measurements. Radio Sci 44: RS1002, doi:10.1029/2008RS003888
Munekane, H (2005) A semi-analytical estimation of the effect of second-order ionospheric correction on the GPS positioning. Geophys J Int 163 (1): 10-17
Mushini, SC, Jayachandran, PT, Langley, RB, MacDougall, JW (2009) Use of varying shell heights derived from ionosonde data in calculating vertical total electron content (TEC) using GPS - New method. Adv Space Res 44 (11): 1309-1313
Odijk, D (2002) Fast Precise GPS positioning in the presence of ionospheric delays. PhD thesis. Technische Universiteit Delft
Odijk, D (2003) Ionosphere-Free Phase Combinations for Modernized GPS. J Surv Eng 129 (4): 165-173
Palamartchouk K (2010) Apparent geocenter oscillations in global navigation satellite systems solutions caused by the ionospheric effect of second order. J Geophys Res 115 (B3):B03415. doi:10.1029/2008jb006099
Papas, CH (1965) Theory of electromagnetic wave propagation. McGraw-Hill, New York
Parkinson, WD (1983) Introduction to Geomagnetism. Scottish Academic Press, Edinburgh
Petrie EJ, King MA, Moore P, Lavallée DA (2010a) A first look at the effects of ionospheric signal bending on a globally processed GPS network. Journal of Geodesy 84 (8):491-499
Petrie EJ, King MA, Moore P, Lavallée DA (2010b) Higher-order ionospheric effects on the GPS reference frame and velocities. J Geophys Res 115 (B3):B03417. doi:10.1029/2009jb006677
Pireaux S, Defraigne P, Wauters L, Bergeot N, Baire Q, Bruyninx C (2010) Higher-order ionospheric effects in GPS time and frequency transfer. GPS Solutions 14 (3):267-277. doi:10.1007/s10291-009-0152-1Plag, HP (2005) The GGOS as the backbone for global observing and local monitoring: A user driven perspective. J Geodyn 40 (4-5): 479-486
Ratcliffe, JA (1959) The magneto-ionic theory and its applications to the ionosphere - a monograph. Cambridge University Press, Cambridge
Rawer, K, Suchy, K (1967) Radio-Observations of the Ionosphere. In: Flugge, S(ed), Handbuch der physik. Vol. 49/2 Springer-Verlag, Berlin, pp 1-537
Rishbeth, H (2003) Basic physics of the ionosphere. In: Barclay, L(ed), Propagation of Radiowaves.The Institution of Electrical Engineers, London, pp 460
Sabaka, TJ, Olsen, N, Purucker, ME (2004) Extending comprehensive models of the Earth's magnetic field with Ørsted and CHAMP data. Geophys J Int 159 (2): 521-547
Schaer, S. (1997) How to use CODE's Global Ionosphere Maps. Astronomical Institute, University of Berne
Schaer, S, Gurtner, W, Feltens, J (1998) IONEX: The IONosphere map EXchange format version 1. IGS AC workshop, Darmstadt, Germany, February 9-11
Schunk RW, Nagy AF (2009) Ionospheres: Physics, plasma physics, and chemistry. Cambridge atmospheric and space science series, 2nd edn. Cambridge University Press, Cambridge
Smith, DA, Araujo-Pradere, EA, Minter, C, Fuller-Rowell, T (2008) A comprehensive evaluation of the errors inherent in the use of a two-dimensional shell for modeling the ionosphere. Radio Sci 43: RS6008, doi: 10.1029/2007RS003769

Stankov, SM, Warnant, R, Stegen, K (2009) Trans-ionospheric GPS signal delay gradients observed over mid-latitude Europe during the geomagnetic storms of October-November 2003. Adv Space Res 43 (9): 1314-1324
Steigenberger, P, Rothacher, M, Dietrich, R, Fritsche, M, Rülke, A, Vey, S (2006) Reprocessing of a global GPS network. J Geophys Res 111: B05402, doi:10.1029/2005JB003747
Strangeways, HJ, Ioannides, RT (2002) Rigorous calculation of ionospheric effects on GPS Earth-satellite paths using a precise path determination method. 37 (2-3): 281-292
Todorova, S, Hobiger, T, Schuh, H (2008) Using the Global Navigation Satellite System and satellite altimetry for combined Global Ionosphere Maps. 42 (4): 727-736
Wang, C, Hajj, G, Pi, X, Rosen, IG, Wilson, B (2004) Development of the Global Assimilative Ionospheric Model. Radio Sci 39: RS1S06, doi:10.1029/2002RS002854
Wang, Z, Wu, Y, Zhang, K, Meng, Y (2005) Triple-Frequency Method for High-Order Ionospheric Refractive Error Modelling in GPS Modernization. 4 (1-2): 291-295
Yeh, KC, Liu, CH (1972) Theory of Ionospheric Waves. Academic Press, London