

COSMIC GPS Ionospheric Sensing and Space Weather

G. A. Hajj^{1,2}, L. C. Lee³, X. Pi^{1,2}, L. J. Romans^{1,2}, W. S. Schreiner⁴, P. R. Straus⁵, C. Wang²

- 1- Jet Propulsion Laboratory, California Institute of Technology
- 2- Department of Applied Mathematics, University of Southern California
- 3- National Space Program Office
- 4- University Corporation for Atmospheric Research
- 5- The Aerospace Corporation

Abstract

As our civilization becomes more dependent on space based technologies, we become more vulnerable to conditions in space weather. Accurate space weather specification and forecasting require proper modeling which account for the coupling between the sun, the magnetosphere, the thermosphere, the ionosphere and the mesosphere. In spite of the tremendous advances that have been made in understanding the physics behind different space weather phenomena, the ability to specify or predict space weather is limited due to the lack of continuous and extensive observations in these regions. Placing a constellation of GPS receivers in low-Earth orbit, such as the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC), provides an extremely powerful system for continuously and extensively measuring one of these regions, the ionosphere. COSMIC, by use of GPS occultations, will make it possible to obtain continuous and global 3-dimensional images of electron density, irregularities and TIDs in the ionosphere and plasmasphere. COSMIC would provide nearly 5600 globally distributed occultations per day suitable for ionospheric sensing. Occultations can be processed individually to obtain vertical profiles of electron density, with vertical resolution of ~1km, or collectively by means of tomography or data assimilation to obtain 3-D images of electron density or irregularity structure. In this paper we describe the GPS observables for ionospheric sensing and the occultation geometry. Our presentation evolves from discussing simple to more complicated inversion techniques starting with the Abel inversion, gradient-constrained Abel inversion, tomography, and finally data assimilation. In each of these techniques, the accuracy is assessed either via examination of real data from GPS/MET or via simulation. Brief discussions of measuring ionospheric irregularities and TIDs are given.

1. Introduction

As we enter the twenty-first century, we are witnessing revolutions in several areas of technology. Remote sensing of the Earth environment is among those areas, where continuous and global sensing of the Earth's atmosphere is becoming a technological necessity for our daily life. For instance, remote sensing of the neutral atmosphere has tremendously influenced the way Numerical Weather Prediction (NWP) is done today and our reliance on it. And as our dependence on technology in space grows, a similar revolution

will be witnessed in the manner we sense and model space weather. The awareness of the potentially hazardous effects of space weather on technological systems motivated the creation of the National Space Weather Program (NSWP) in the United States and similar programs elsewhere. NSWP places particular emphasis on the need to provide timely, accurate, and reliable space environment observations, specifications, and forecasts. These requirements are similar to what is already accomplished with a great measure of success in the NWP community.

Accurate specification and forecasting of space weather phenomena is very difficult because it requires (1) accurate modeling of the coupling between the sun, the magnetosphere, the thermosphere, the ionosphere, and the mesosphere, (2) continuous, reliable and accurate observations of all of these regions, (3) the ability to assimilate the data into the models in an optimal and self-consistent manner. These are great challenges for the space weather community and require tremendous efforts and advances in several areas of research and technology.

In the next few years we will witness great advances in the way the space environment is sensed as several satellites or constellations of satellites are launched with space sensing instruments of various types. One such mission is the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC), which will consist of 8 microsattellites in 8 evenly spaced orbital planes with circular orbits at ~800km altitude and 72 deg. inclination. Each COSMIC satellite is equipped with a receiver that tracks all Global Positioning System (GPS) satellites in view including those rising and setting behind the Earth's atmosphere and ionosphere. Limb sounding of the Earth's ionosphere with GPS-LEO links was first demonstrated with the GPS/MET experiment and more recently with Oersted and SUNSAT. The COSMIC mission with its 8 microsattellites will provide nearly 5600 ionospheric occultations per day when counting rising, setting and "grazing" occultations. This relatively new approach of sensing the Earth's ionosphere with a GPS receiver in LEO will potentially change the manner by which we observe the ionosphere. GPS observables can be used to measure integrated electron density (known as total electron content—TEC) along the transmitter-receiver line-of-sight and scintillation. With a dense set of measurements, as would be obtained from COSMIC, it will possible to obtain instantaneous global 3-dimensional (3-D) maps of electron densities, ionospheric irregularities and Traveling Ionospheric Disturbances (TIDs) at all times.

This paper has two purposes. First, to provide an overview of ionospheric sensing from GPS space measurements. Second, to discuss means of making use of the extremely rich set of measurements that could become available from COSMIC. We will describe the different methods of processing ionospheric data collected during an occultation in a progressive manner starting from the simplest (Abel inversion) to the most sophisticated (data assimilation). We first review the GPS observables and the various ways in which TEC, with different qualities of noise, can be derived. This is done in section 2. In section 3 we discuss the coverage of COSMIC occultations in the ionosphere from which we will start to get a glimpse that a constellation like COSMIC would provide nothing less than a revolution in ionospheric sensing. In section 4 we review the Abel transform as a means of obtaining electron density profiles from individual occultations and we compare results obtained from

GPS/MET with incoherent scatter radar and ionosonde measurements. In section 5, we discuss a possible improvement to the Abel inversion scheme by use of the “constrained-gradient” inversion and examine its performance through simulation experiments. In section 6, a tomographic scheme is discussed as a means of combining nearby occultations to provide a purely data driven specification of electron density. In section 7, we describe a full-fledged assimilation scheme using a recursive Kalman filter. An Observation System Simulation Experiment (OSSE) is described in that section to assess the value of our assimilation scheme and the impact that a system such as COSMIC might have for space weather. In section 8 a brief discussion is given of the impact of GPS ionospheric measurements, including scintillation and detection of TID, on space weather. We conclude in section 9.

2. Basic Observables and Occultation Geometry

2.1 Occultation Geometry

A LEO-GPS occultation geometry is depicted in Fig. 1. An occultation takes place when the GPS satellite sets or rises behind the Earth’s ionosphere or lower neutral atmosphere as seen by a receiver in LEO. Each occultation therefore consists of a set of links with tangent points ranging from the LEO satellite height to the surface. One occultation takes nearly 4-10 minutes depending on the relative geometry of the LEO and GPS satellites. When the signal passes through the Earth media, it experiences delay, bending and amplitude change (e.g., defocusing or scintillation). These changes can be used to extract information such as profiles of electron density in the ionosphere, or temperature, pressure and water vapor in the lower neutral atmosphere (0-50 km), in the vicinity of the tangent point.

2.2 GPS Observables

The GPS constellation consists of 24 satellites at ~ 26500 km radius, ~ 12 hr period, orbiting in 6 different planes inclined at $\sim 55^\circ$. Each GPS satellite broadcasts two signals at L-band ($f_1 = 1575.42$ MHz and $f_2 = 1227.60$ MHz). A properly equipped receiver will detect amplitude, pseudorange¹ and phase measurements for both frequencies. Range and phase measurements can be modeled as:

$$P_k^{ij} \equiv \rho^{ij} + \delta_k^{ij} + d \frac{TEC_k^{ij}}{f_k^2} + C^i + C^j + v_k \quad (1)$$

$$L_k^{ij} \equiv -\frac{c}{f_k} \Phi_k^{ij} = n_k^{ij} \lambda_k + \rho^{ij} + \delta_k^{ij} - d \frac{TEC_k^{ij}}{f_k^2} + C^i + C^j + \varepsilon_k \quad (2)$$

with

- P_k^{ij} the pseudorange between transmitter i to receiver j ;
- $k = 1$ or 2 for L1 or L2 frequencies, respectively;
- ρ^{ij} the range corresponding to the light travel time (in vacuum) between the transmitter and the receiver;

¹ Pseudorange is an absolute measurement of group delay between the time a signal is transmitted and received. It is the sum of the actual range between the transmitter and the receiver, atmospheric and ionospheric delays and transmitter and receiver clocks offsets.

δ_k^{ij} the extra delay due to neutral atmosphere;
 d is a constant corresponding to the square of the plasma frequency divided by the electron density ($= e^2/(4\pi^2\epsilon_0 m) = 40.3$ in International Standard (IS) units);
 TEC_k^{ij} the TEC along the ray path;
 C^i, C^j transmitter and receiver clock errors, respectively, including constant biases and time dependent terms;
 v_k measurement errors which include the receiver's thermal noise and local multipath;
 L_k^{ij} the recorded phase (in units of distance) between transmitter i and receiver j ;
 Φ_k^{ij} the recorded phase in cycles;
 c the speed of light in vacuum;
 n_k^{ij} an integer which is constant over a connected arc;
 λ_k the operating wavelength;
 ϵ_k phase measurement noise.

Eqs. (1 and 2) ignore (1) transmitting and receiving antenna phase center variations and “wind up” effects [Wu *et al.*, 1993] which can be calibrated or modeled if needed, and (2) higher order ionospheric terms (order $1/f^3$ or higher), which result from the expansion of the Appleton-Hartree formula [see, e.g., Bassiri and Hajj, 1993]. These terms can be ignored for ionospheric retrievals. A subscript on any term in Eqs. (1 and 2) implies that it is frequency dependent. The neutral atmosphere is non-dispersive at radio frequencies; however, because the L1 and L2 “split” in the ionosphere, at a given time they sense slightly different parts of the neutral atmosphere (as depicted in Fig. 1). This split is the reason for having subscripts on δ_k^{ij} and TEC_k^{ij} in Eqs. (1 and 2).

2.3 Ionospheric Observables

There are several methods for extracting the ionospheric delay along the LEO-GPS line-of-sight. The selection of a specific method depends on many factors which include the data noise, desired retrieval accuracy and resolution, and the inversion technique used in retrieving geophysical quantities such as electron density profiles. For completeness, we summarize the different methods below.

2.3.1 Relative TEC from L_1 and L_2 ($TEC_{r,p}$)

By ignoring the split of the L1 and L2 signals, we can calculate the relative $TEC_{r,p}$ (subscript r refers to relative, and p refers to precise) along the LEO-GPS line-of-sight by taking the linear combination

$$TEC_{r,p} = \frac{1}{d} \frac{f_1^2 f_2^2}{(f_1^2 - f_2^2)} (L_1 - L_2) + B + v \quad (3)$$

where B is an undetermined bias which is constant for a connected arc, and v is the TEC noise. The constant multiplying $L_1 - L_2$ is equal to 9.52×10^{16} in IS units. Therefore, each 1 meter of differential delay between L_1 and L_2 corresponds roughly to 10 TECU (1 TECU = 10^{16} e/m³). The TEC noise is related to the L_1 and L_2 noise by

$$\sigma_v = \frac{1}{d} \frac{f_1^2 f_2^2}{(f_1^2 - f_2^2)} \sqrt{\sigma_{v_1}^2 + \sigma_{v_2}^2} \quad (4)$$

The precision of $TEC_{r,p}$ when Anti-Spoofing (AS) is off or with enhanced codeless GPS receivers such as the ones to be carried on COSMIC is better than 0.01 TECU.

2.3.2 Absolute TEC from P_1 and P_2

Absolute $TEC_{a,n}$ (a for absolute, n for noisy) can be derived from P_1 and P_2 by taking the linear combination

$$TEC_{a,n} = \frac{1}{d} \frac{f_1^2 f_2^2}{(f_1^2 - f_2^2)} (P_2 - P_1) + C + \zeta \quad (5)$$

where C is the sum of the transmitter and receiver differential code biases (DCB's), which experience little variation and are usually assumed constant. Much work has been done on solving for these constants [Mannucci *et al.*, 1998], and they are usually known to within 1-2 TECU. The TEC 1- σ noise (ζ) is determined by the P_1 and P_2 noise in a relation similar to Eq. (4). Since range measurements tend to be about 100 times noisier than those of the phase, a standard procedure is to solve for B in Eq. (3) by a least-square minimization of $\Sigma(TEC_{a,n} - TEC_{r,p})^2$, where the sum is over all measurements in a connected arc. This way, we obtain an absolute and precise $TEC_{a,p}$ with a bias which is of order of 1-3 TECU.

2.3.3 Relative TEC from L1 and P_1 ($TEC_{r,n}$)

Thus far, it has been assumed that the GPS receiver tracks both L1 and L2 signals. In many applications, the GPS receiver can be a single frequency receiver tracking L1 only. In this case the $TEC_{r,n}$ can be derived by taking

$$TEC_{r,n} = \frac{f_1^2}{2d} (P_1 - L_1) + B' + \xi \quad (6)$$

where B' is an undetermined bias which is constant over a connected arc. We note that $TEC_{r,n}$ is a relative TEC measurement with noise (ξ) which is dominated by the P1 noise.

2.3.4 Relative TEC from L1 Only ($TEC_{r,p,L1}$)

If the range and clock terms (δ^j , C^i , C^j) in Eq. (2) are known, then the sum of the atmospheric and ionospheric delays contributing to a phase measurement ($\delta^j - d \times TEC^{ij}/f^2$) can be determined up to a constant bias. The range term can be determined from Precise Orbit Determination (POD) of the LEO and GPS satellites. GPS POD is routinely available by use of a ground network of receivers. LEO satellite POD is possible by use of GPS satellites in view. In addition, the clock terms can be solved relative to an accurate clock on the ground by a process similar to "double differencing" where four links (the occulting link and the dashed links in Fig. 1) are differenced such that all clocks cancel out. The details of this technique as it applies to neutral atmospheric sensing are described in [Kursinski *et al.*, 1997]. When the tangent point of the occulted signal is above 80 km, the neutral atmospheric delay is less than 2 mm and can be ignored, therefore a relative and precise TEC can be obtained.

This method has several advantages over the one described in 2.3.1: (1) it requires a single frequency receiver², (2) it is more precise since the $\text{TEC}_{r,p,L1}$ noise ($\sigma_v = (f_1^2/d) \sigma_{v_1} = 6.16 \times 10^{16} \sigma_{v_1}$ in IS units) is significantly smaller than that of Eq. (4), (3) it does not ignore the effect of bending. (Based on GPS/MET data, Hajj and Romans [1998] showed that bending in the ionosphere during solar minimum is < 0.01 deg. If this value reaches up to 0.1 deg. during solar maximum, ignoring bending will result in systematically over-estimating the HmF2—since bending above the F-peak is upward—by up to 5 km). These advantages have to be weighed against the extra complications and time required in ensuring that precise orbits as well as double differencing geometries are available.

L1 and L2 amplitude measurements

In addition to phase and range measurements, the receiver measures the voltage signal-to-noise ratio (VSNR) for both L1 and L2 from which amplitude scintillation can be detected. Scintillation measurements, from which ionospheric irregularity maps can be obtained, are of vital importance to space weather because of their potentially detrimental effects on satellite and ground communications. Amplitude measurements can also be combined with the phase measurements to improve occultation retrievals in various ways including detecting and correcting for atmospheric multipath effects, accounting for diffraction for the purpose of improving the horizontal resolution [see e.g., *Karayel and Hinson, 1997*], and possibly localizing electron density irregularities along the GPS-LEO ray path [*Vorob'ev et al., 1999; Sokolovskiy, 2000*].

3. COSMIC Coverage and Resolution

During one LEO satellite revolution, each GPS satellite will in general set or rise once behind the Earth's atmosphere. Each COSMIC satellite, with nearly 14.6 revolutions per day, will see $14.6 \times 24 \times 2 = 700$ daily occultations. The total number of daily occultations seen by the 8 COSMIC satellites is then over 5600. An example of the distribution of such a coverage based on actual GPS and representative COSMIC orbits is shown in Fig. 2 which has a total of 5616 occultations. These occultations are divided nearly equally between rising, setting, and grazing geometries, which are defined based on the angle between the occultation plane and the LEO orbital plane being between 0-45 deg., 45-135 deg., and 135-180 deg, respectively.

Representing an occultation by one point, as in Fig. 2, does not reveal the complex structure of occultation links in the ionosphere. Fig. 3 shows all the occultation links (at the rate of one per minute) obtained during 98 minutes (the LEO satellite period) as a function of local time. Links obtained from one COSMIC satellite are highlighted in order to see their distribution relative to the LEO plane. Lines that are nearly parallel correspond to links from the same occultation. Setting and rising occultations appear as a dense set of lines (in some cases nearly on the top of each other) parallel to the orbital plane. Grazing occultations

² Sufficiently accurate POD and clock solutions can still be obtained with a single frequency receiver [see e.g., Bertiger et al., 1995].

appear as parallel lines that move considerably (in some cases half way across the globe) during an occultation span. This smearing of the occultation link makes the spherical symmetry approximation (used in the Abel inversion process) invalid for grazing occultations—which is the reason why these occultations are considered less useful for sensing the neutral atmosphere (0-50km)³. However, because the scale height in the ionosphere is nearly 10 times larger than that of the lower neutral atmosphere, the horizontal scale over which the occultation links are sensitive in the ionosphere is nearly 10 times that of the neutral atmosphere (i.e., ~2000 km in the ionosphere vs. 200 km in the lower neutral atmosphere). This fact allows us to combine links from nearby occultations in a tomographic manner when imaging the ionosphere. By examination of Fig. 3, it is possible to see that grazing occultations would be of extreme value to ionospheric sensing since they provide nearly orthogonal cuts to those of the setting and rising occultations.

Even though the numbers of rising, setting and grazing occultations in Fig. 2 are nearly the same (1867, 1878 and 1868, respectively), their distribution as a function of latitude is considerably different as indicated in Fig. 4 which shows the average number of daily occultations per (1000km)² at different latitudes. Based on Fig. 4, we estimate that the distance between occultations for COSMIC (in one day) is on the order of 250-500 km depending on latitude, and on whether grazing occultations are used.

Vertical Resolution

The theoretical vertical resolution of ionospheric profiles obtained from GPS-LEO occultations is set by the diffraction limit, which corresponds to the first Fresnel zone diameter given by

$$D_F = 2\sqrt{\frac{\lambda D_{GPS} D_{LEO}}{D_{GPS} + D_{LEO}}} \quad \lim_{\frac{D_{LEO}}{D_{GPS}} \rightarrow 0} = 2\sqrt{\lambda D_{LEO}} \quad (7)$$

Where D_{GPS} and D_{LEO} are the distance from the tangent point (Fig. 1) to the GPS and LEO, respectively. In the case of COSMIC, $D_{GPS} \approx 30,000$ km and $D_{LEO} \approx 3,000$ resulting in $D_F \approx 1.5$ km. However, a more practical limit on the vertical resolution for the ionosphere is introduced by horizontal structures which, if ignored, can alias into vertical structures. Therefore, the horizontal resolution depends on the retrieval scheme and the ability to capture the large horizontal scale. This issue will be addressed further later.

Topside Ionosphere and Plasmasphere coverage

In addition to the occultation links which probe the ionosphere below COSMIC height, COSMIC will collect positive elevation GPS data at 0.1 Hz rate which measure line-of-sight TEC in the topside ionosphere and plasmasphere. Each COSMIC satellite will observe nearly 10 GPS satellites at positive elevations simultaneously in different directions of the sky providing a very dense coverage and mapping of electron density above the COSMIC altitude.

³ Grazing occultations could still be useful in the neutral atmosphere if assimilated at the bending angle level.

Complementary Ground Coverage

The continuously operating networks of GPS ground receivers provide a very rich set of GPS-TEC ground data complementary to the COSMIC space data. Data from over 200 GPS globally distributed ground receivers, each tracking 8-12 GPS satellites simultaneously (Fig. 5), currently constitute a major network for sensing the ionosphere. These TEC data, for instance, were used to provide the first global-scale ionospheric maps of vertical TEC [Mannucci *et al.*, 1998; Lu *et al.*, 1998].

4. Ionospheric Profiling

4.1. Abel Inversion

When the ionospheric advance (for phase) or delay (for range) ($I_k = \mu \left(\frac{d}{f_k^2} \right) TEC$, see Eqs. 1 and 2) is determined, it is differentiated (after proper smoothing) to obtain the extra Doppler shift induced by the medium. This extra Doppler shift can be used to derive the bending of the signal, α , as a function of the asymptotic miss distance, a (Fig. 1), by assuming a spherically symmetric atmosphere in the locality of the occultation; the relationship between the direction of the signal's propagation and the extra Doppler shift, Δf , is then given by

$$\Delta f = \frac{f}{c} [\mathbf{v}_t \cdot \mathbf{k}_t - \mathbf{v}_r \cdot \mathbf{k}_r - (\mathbf{v}_t - \mathbf{v}_r) \cdot \mathbf{k}], \quad (8)$$

where \mathbf{v}_t and \mathbf{v}_r are the transmitter and receiver velocity, respectively; \mathbf{k}_t and \mathbf{k}_r are the unit vectors in the direction of the transmitted and received signal, respectively; and \mathbf{k} is the unit vector in the direction of the straight line connecting the transmitter to the receiver. Assuming spherical symmetry introduces the extra constraint

$$a = n(\mathbf{r}_t) \|\mathbf{r}_t \times \mathbf{k}_t\| = n(\mathbf{r}_r) \|\mathbf{r}_r \times \mathbf{k}_r\|, \quad (9)$$

where \mathbf{r}_t and \mathbf{r}_r are the coordinates of the transmitter and the receiver, respectively; and n is the index of refraction. Eqs. (8) and (9) can be solved simultaneously in order to estimate $\alpha(a)$. Solving Eqs. (8 and 9) requires knowledge of n at the satellites locations; however, the error acquired by setting $n=1$ is negligible [Hajj and Romans, 1998; Schreiner *et al.*, 1999].

The spherical symmetry assumption can also be used to relate the signal's bending to the medium's index of refraction, n , via the relation [Born and Wolf, 1980, p. 123]

$$\alpha(a) = -2a \int_a^\infty \frac{1}{\sqrt{a'^2 - a^2}} \frac{d \ln(n)}{da'} da', \quad (10)$$

where $a(r) = r \times n(r)$ and r is the radius of the tangent point. This integral equation can then be inverted by using an Abel integral transform given by [see, e.g., Tricomi, 1985, p. 39]

$$\ln(n(a)) = -\frac{1}{\pi} \int_a^\infty \frac{\alpha(d)}{\sqrt{a'^2 - a^2}} da'. \quad (11)$$

Electron density can then be derived by use of the relation (– for phase advance, + for group delay)

$$n = 1 \mu 40.3 \times \frac{n_e}{f^2} \quad (12)$$

Eqs. (8-12) can be approximated and reduced to a single equation by ignoring bending. From Eq. (8) it is easy to establish the following relation between ionospheric bending and advance/delay (I_k)

$$\alpha_k = \frac{dI_k}{dt} \frac{D_{LEO} + D_{GPS}}{D_{LEO} V_{GPS\perp} + D_{GPS} V_{LEO\perp}} = \frac{dI_k}{dt} \frac{dr_s}{dt} = \frac{dI_k}{dr_s}. \quad (13)$$

where d/dt is the time derivative, D_{GPS} and D_{LEO} are defined in Eq. (6), $V_{GPS\perp}$ and $V_{LEO\perp}$ are the GPS and LEO velocity components perpendicular to the GPS-LEO link, r_s is the radius of the tangent point of the straight line connecting the two satellites. By use of the approximation $a = rn \approx r_s$, Eq. (13), and expanding $\ln(n)$ in Eq. (11) we obtain

$$n_e(r_s) = -\frac{1}{\pi} \int_{r_s}^{\infty} \left(\frac{dTEC(r'_s)}{dr'_s} \right) \frac{1}{\sqrt{r_s'^2 - r_s^2}} dr'_s. \quad (14)$$

Note that in both of these formulations, only the derivative of the ionospheric delay or TEC is used and therefore a constant bias is irrelevant.

The upper limit of the integrals in Eq. (11 or 14) requires knowledge of $\alpha(a)$ or the derivative of $TEC(r_s)$ up to the top of the ionosphere. While the GPS satellites are above the ionosphere, this is not true of the COSMIC instrument at ~ 7170 km radius. Different methods have been used to overcome this for the GPS/MET experiment at 7110 km radius. Hajj and Romans [1998] have used an exponential extrapolation of $\alpha(a)$ based on information from $a < 7110$. Schreiner *et al.* [1999] have eliminated the contribution of the ionosphere above the LEO satellite by assuming the occultation to be planar, the GPS satellite to be at infinity, and the ionosphere to be static over the duration of the occultation. In this way the contribution from the topside can be removed by subtracting a TEC measurement taken at point A (Fig. 6) from that at point B.

4.2 Results from the GPS/MET Experiments

Many features and examples of GPS/MET retrievals can be found in [Hajj and Romans., 1998; Schreiner *et al.*, 1999; Straus *et al.*, 1999a and 1999b]. In this section we summarize some key features and present some new comparisons.

4.2.1 Examples of ionospheric retrievals

Fig. 7 compares a GPS/MET profile obtained on May 5, 1995, 2200 local time, with tangent points at about 42N and 282E, with an incoherent scatter radar (ISR) profile from Millstone Hill (42.6N & 288.5E) taken within 20 minutes of the time of the occultation. The general agreement is very good. Discrepancies between the ISR and the occultation profiles can be due to several factors including mis-collocation between measurements, spherical symmetry assumption of the Abel inversion and ISR measurement noise.

Other examples of GPS/MET electron density are shown in Figs. 8.a, 8.b and 8.c which show one, several, and two sharp layers, respectively, at the bottom of the ionosphere⁴. The

⁴ One should be careful, however, not draw too much information from the E layers in Figs 8a-d as some of the vertical structure could be aliased horizontal structure.

sensitivity to ionospheric structures of very short vertical scales, such as sporadic *E*, is a consequence of the high vertical resolution characteristic of the GPS occultation limb viewing geometry. (Data rate for retrievals in Figs. 7 and 8 is 0.1 and 50 Hz for points above and below 120 km, respectively, which explains the much finer resolution below 120 km.) It should be noted, however, that the gross structure in Fig. 8.a below 200 km (i.e., the increase in density towards lower altitudes) is not realistic for the near-midnight local time of that occultation. This feature is most likely due to horizontal gradients which have a stronger effect on the retrieval at lower altitudes both in absolute (because the ray travels a longer distance in the ionosphere) and fractional (because the E-region densities are lower) terms.

4.2.2 NmF2 Comparisons with Ionosonde

Ionospheric profile retrievals were derived at the University Corporation for Atmospheric Research (UCAR) using single frequency TEC data (as described in section 2.3.3) and the Abel transform for 40,000 GPS/MET occultations collected between Apr. 1995-Feb. 1997. Out of those, 4411 occultations had tangent points within 1200 km and 1 hour from an ionosonde measurement and are used for statistical comparisons.

Fig. 9.a shows a scatter plot of NmF2 values obtained from GPS/MET and nearby ionosonde. Fig. 9.b shows a histogram of the difference (NmF2(GPS/MET) – NmF2(ionosonde)) with a standard deviation, σ , of $0.15 \times 10^{12} \text{e/m}^3$. NmF2 values with differences larger than 4σ (a total of 190 values or 4.3% of the entire data set) were excluded from the statistics.

We further examine these NmF2 differences as a function of local time (LT). Fig. 10.a shows the NmF2 values derived from GPS/MET data as a function of LT. Hourly averages are shown by the solid line and their 1σ scatters are shown as vertical bars. The diurnal signature is evident with larger values during the day-time, an average maximum near local noon and a minimum near 400 LT. Hourly standard deviations of the NmF2 GPS/MET and ionosonde differences are shown in Fig. 10.b, and their ratios to the mean NmF2 values are shown in Fig. 10.c. Even though the differences are larger during the day in an absolute sense, the fractional difference is largest ($\sim 40\%$) during the pre-dawn sector and smallest ($\sim 20\%$) around local noon. The larger fractional error of NmF2 during the nightside reflects the fact that ionospheric horizontal gradients are generally larger on the nightside than on the dayside, which causes the spherical symmetry assumption used in the Abel inversion to be more in error. These results are consistent with those found by *Staus* [1999a and 1999b]. Differences of 20% or 40% in NmF2 implies differences of 10% or 20 % in foF2, respectively. It should be noted that these are average statistics for all seasons near solar minimum conditions (1995-1997).

5. “Horizontally-constrained” Inversion

The dominant error in the Abel inversion is due to the spherical symmetry assumption imposed on the ionosphere. However, when information on horizontal structure is available, more general inversion schemes can be constructed. *Hajj et al.*, [1994] presents one possible scheme where the ionosphere is divided into pixels on the basis of radius, latitude and longitude. Pixels at the same layer (i.e., same radius) are then constrained to have a given

functional form with an undetermined scale factor. This functional form can be the same for all layers, or different for different layers. The problem is then reduced to solving for the single unknown scale factor for each layer, reducing the 3-D under-determined problem to a one-dimensional problem with a unique solution.

Following the analysis by *Straus* [1999b], in order to study the effects of horizontal gradients we replace the one dimensional electron density function, $n_e(a)$, with a two-dimensional density $n_e(a,\phi)$, where ϕ is the angle of a point along the ray path from the tangent point as measured from the Earth's center. We further assume that $n_e(a,\phi)$ can be separated into the product of two terms: $n_e(a,0)$, the electron density profile at the tangent point, and $f(\phi)$, a function which provides the ratio of electron densities along the line of sight to the density at the tangent point (note that $f(0) = 1$). It can be readily seen that only the portion of $f(\phi)$ which has even symmetry about the tangent point (we denote by $f_e(\phi)$) contributes to gradient-induced retrieval errors, since errors introduced by the odd terms will cancel upon integrating along the ray path.

To assess the utility of various types of ancillary data for constraining the horizontal structure and improving occultation retrievals, simulations of both occultation limb soundings and three different types of ancillary data have been performed using the RIBG ionospheric model [*Reilly*, 1993] to provide a "known" ionosphere based on climatology. The simulated ancillary data are in-situ plasma density, nadir-viewing extreme ultraviolet airglow, and total electron content maps derived from ground-based GPS receivers. The second type of measurements will be available from an ionospheric photometer sensor on COSMIC. The third type of data is routinely available from Global Ionospheric Maps (GIM) derived at JPL. To focus on correction of errors strictly due to horizontal gradient effects along the line-of-sight for limb-viewing measurements, an idealized occultation geometry is imposed which is only susceptible to this particular error source. That is, we simulate occultations that include no motion of the tangent point throughout the measurement. Although the case of zero tangent point motion never actually occurs in practice due to the motion of the GPS satellite during an occultation, it is most closely approximated by an "in-track" occultation for which the direction to the GPS satellite is in the LEO orbit plane. Use of this idealized occultation geometry also allows us to perform the analysis on a regular spatial grid, avoiding the irregularly spaced spatial distribution of occultations which actually occurs under the realistic conditions generated by the orbital motions of various spacecraft involved.

The simulation proceeds as follows. A latitude/longitude grid, with 5° and 10° spacing, respectively, is constructed and a look direction for all occultations is selected. The driving parameters for the RIBG model (F10.7, Kp, day of year, universal time) are chosen and fixed. For each point on the grid, an in-track occultation (i.e., in the selected look direction) with zero tangent point motion is simulated. Simulated values of the various ancillary measurements are calculated along the plane containing the occultation lines-of-sight. For the cases in which these measurements are made by secondary sensors on the LEO satellite, this corresponds to the "in-track" assumption in which the satellite passes directly over the tangent point for the occultation. Electron density profiles are then determined using the Abel inversion and the gradient-corrected inversions for the three different types of ancillary measurements. The accuracies of the retrievals are judged in terms of NmF2 and HmF2, as

compared to the “true” model values at each of the grid points. The process is repeated for different look directions and ionospheric model drivers.

For this study, an additional simplifying assumption has been made in the case that the gradient information is obtained using the EUV technique. The specific emission simulated is the 911 Å emission associated with radiative recombination of atomic oxygen ions [Meier, 1991; Dymond, *et. al.*, 1997]. The volume emission rate of this feature is proportional to the product of the oxygen ion and electron densities. For the purposes of this analysis, we have assumed that these densities are equal, so that the observed airglow is proportional to the integral of the square of the electron density below the satellite. An estimate of the horizontal gradients is then obtained from the variations in the square root of the simulated observations. To the extent that light or molecular ions are present, particularly at altitudes near the location of the maximum emission rate, this assumption will break down. However, since oxygen is known to be the dominant ion near the F2 peak under nearly all conditions, and since only the relative horizontal changes in the airglow magnitude are used to determine the gradients, our assumptions should introduce an insignificant error most of the time. Although the 911 Å feature can be simulated globally, on the dayside it is difficult to isolate in practice from other dayglow emissions, limiting actual use of this correction technique to the nightside.

Error maps for NmF2 for one particular set of simulation results are shown in Figs. 11. This run corresponds to solar maximum, equinox, geomagnetically quiet conditions at 0 UT, with a northward occultation look direction (i.e., the direction in which the strongest average gradients are expected). The “true” ionospheric model values of NmF2 and HmF2, shown in Fig 11.a and 11.b, provide an indication of the horizontal structures in this simulation. Fractional errors in NmF2 and actual errors in HmF2 obtained when using the simulated EUV measurements to constrain the horizontal structure are shown in Figs. 11.c and 11.d, respectively. A statistical summary of other simulations is given in Table 1. As might be anticipated, the largest NmF2 errors occur in the auroral zone and near the equatorial arcs, where gradients are strongest. Errors in HmF2 follow a similar phenomenology. On average, techniques employing a gradient correction provide significant improvement over the spherically symmetric Abel transform, particularly for NmF2. In general, the TEC and EUV techniques for gradient correction outperform the in-situ based method. This is reasonable, since these data types provide information on gradients more localized to the vicinity of the peak of the layer, whereas the in-situ method relies on extrapolation of gradients measured at the satellite altitude to lower regions.

However, improvement using any of the three gradient correction techniques is not guaranteed in individual cases. For example, consider the point near the peak of the equatorial anomaly near sunset at -10° latitude/ 270° longitude. Fig. 12.a shows ionospheric density contours in the plane of the occultation together with four representative lines of sight corresponding to tangent altitudes of 300, 400, 500, and 600 km. The corresponding true profile at the tangent point is shown in Fig. 12.b together with Abel, in-situ, and TEC-based inversions. Despite the strong gradients in this region, the Abel inversion fortuitously performs best in reproducing the true NmF2. None of the techniques reproduce HmF2 or the bottomside particularly well. The situation can be understood by considering Fig. 12.c,

which shows the actual and estimated horizontal gradient function in the ionosphere, $f_c(\phi)$. Here the solid curves represent the even asymmetries encountered by rays with the same four tangent altitudes as in Fig. 12.a (the largest asymmetry corresponds to the lowest tangent altitude). At the highest tangent altitude, the asymmetry is modeled well by the in-situ based estimate (dashed curve), but the shape of the asymmetry function for lower altitudes is different. For tangent altitudes near HmF2 (between 400 and 500 km), neither of the two asymmetry estimates provides a good match to the actual asymmetries. This occurs primarily because the height of the layer is significantly varying within the region sensed by the occultation. If the height of the layer were not varying, then the gradients in the vicinity of the peak would be accurately reflected in at least the TEC-based measure, and possibly the in-situ-based measure as well.

Table 1. Standard deviations of NmF2/HmF2 errors for different retrieval techniques. The EUV technique can only be applied in practice at night at non-auroral latitudes.

	Abel	In-Situ	EUV	TEC
North-Looking				
Equinox, All	12.2%/11.0 km	7.9%/12.2 km	---	4.2%/7.7 km
Low & Mid Lat./ Night	14.1%/11.7 km	8.7%/12.4 km	7.1%/5.1 km	5.0%/6.5 km
Winter Solstice, All	13.0%/10.4 km	7.8%/12.4 km	---	4.7%/9.1 km
Low & Mid Lat./ Night	11.8%/ 6.6 km	8.4%/ 7.3 km	7.3%/4.5 km	6.2%/5.0 km
Summer Solstice, All	17.8%/10.6 km	7.6%/ 9.1 km	---	5.6%/9.0 km
Low & Mid Lat./ Night	15.2%/ 6.1 km	8.7%/ 5.6 km	6.2%/3.2 km	5.1%/3.7 km
East-Looking				
Equinox, All	6.0%/5.2 km	3.2%/ 5.0 km	---	2.3%/3.9 km
Low & Mid Lat./ Night	5.4%/3.7 km	2.7%/ 2.8 km	2.4%/0.9 km	1.7%/1.2 km
Winter Solstice, All	6.5%/6.4 km	3.5%/ 7.4 km	---	2.8%/6.3 km
Low & Mid Lat./ Night	4.7%/2.0 km	2.1%/ 2.1 km	2.5%/1.2 km	2.2%/1.4 km
Summer Solstice, All	8.4%/5.6 km	3.2%/ 4.1 km	---	2.8%/4.8 km
Low & Mid Lat./ Night	6.4%/2.0 km	2.3%/ 1.5 km	2.2%/0.8 km	1.5%/0.9 km

In addition to the equinox, north-looking case presented here, summer and winter solstice cases have been simulated. The overall statistics for these cases are shown in Table 1 and are similar to the equinox case. However, east-looking statistics, also shown in Table 1, are significantly better for all methods. This is because ionospheric gradients are typically less strong for this direction of viewing. The most significant retrieval errors for the east-looking geometry are near the sunrise terminator, in the auroral zone, and in the equatorial zone near the point where the magnetic equator bends away from the geographic equator, resulting in more significant gradients in the geographic east-west direction.

The simulation results presented in this section do not include ancillary data noise effects and are therefore optimistic. When realistic errors of the ancillary data are considered, the improvement seen in the above simulations may not be as dramatic. For instance, *Schreiner et al.*, [1999] have examined the use of vertical TEC ancillary data to constrain GPS/MET retrievals, but their derived NmF2 did not show any significant improvement over the Abel inversion. The reasons for the discrepancies between the simulations presented here and

statistical results presented by *Schreiner et al.* are not well understood, and further simulations and data analysis are needed to reach more definitive conclusions.

6. Tomographic Inversion

With the straight line propagation assumption, we can combine data from several occultations and obtain a linear system of equations given by

$$A_{m,i}x_i = y_m + \sigma_m \quad (15)$$

where y_m is the m 'th differential TEC measurement (TEC differences between consecutive measurements are used in this process to get rid of all TEC bias terms); x_i is the electron density in a cell i (here we divide the ionosphere into pixels and assume the density to be constant inside each pixel), $A_{m,i}$ is an element in the observation matrix relating the observation m to the density in pixel i ; σ_m is the standard deviation of the differential TEC measurement error.

The tomography problem as stated is in general underdetermined. One approach to regularize this singularity is to augment the measurement Eqs. (15) with *a priori* constraints on the pixel values and the differences between nearest-neighbor pixels

$$x_i = x_i^{ap} + s_i \quad (16)$$

$$x_i - x_j = x_i^{ap} - x_j^{ap} + s_{ij} \quad (17)$$

where i, j are nearest neighbors. These constraints can be handled in parallel with the *bona fide* measurements. The derivative constraints of Eqs. (17) automatically implement spatial interpolation of the solution to regions not probed by the data links. The values of the various sigmas determine the relative strengths of interpolation *vs. a priori* values *vs.* the data itself.

We note that, aside from the effects of ignoring the bending, the spherical symmetry approximation used in section 4.1 can be thought of as the solution to Eqs. (15) and (17) with $s_{i,j} \rightarrow 0$ for pixels i and j in the same layer (no horizontal variation allowed) and $s_{i,j} \rightarrow \infty$ for pixels i and j in different layers (complete flexibility in the vertical direction). Therefore, the approach used here serves as a natural extension to the Abel inversion solution for combining nearby occultations.

For the purpose of illustrating this approach we choose to combine data from occultations that are taking place during one revolution of GPS/MET. Figs. 13.a and 13.b show side and top views of the occultation links during one orbital revolution of the LEO orbit and the grid used in doing the tomography. Examples of 2-D images obtained during consecutive revolutions are shown in Fig. 14. The plane shown corresponds to that of the GPS/MET orbital plane (at 70° inclination). Zero degrees corresponds to local noon at the equator, and the reconstructed peak density is at roughly 11 N. latitude and 12:20 local time, May 4, 1995. No ground data were used in this inversion. The *a priori* electron density was set to zero for the first reconstruction and to the prior solution for the 2nd, 3rd, and 4th reconstructions. Even with the sparse coverage of GPS/MET, some of the main features expected in the ionosphere, such as the position of the peak and the scale heights, are reflected in the reconstruction.

(Other examples of GPS/MET tomographic reconstruction from GPS/MET data are also given by *Rius et al.* 1997 and *Rius et al.*, 1998.) With the continuous COSMIC coverage, 4-dimensional (3 spatial and 1 temporal) global maps of ionospheric density with high vertical (1-2 km) and horizontal (200-500 km) resolution will become available for the first time.

7. Data assimilation

While computerized ionospheric tomography (CIT) can be very powerful in capturing certain ionospheric structures that can be missed by physical modeling, it has a major limitation that is fundamental to space weather: the inability to forecast the ionospheric state. A more powerful approach with the potential to overcome this limitation is data assimilation, similar to what is done in NWP. In the presence of global and continuous data, a physical assimilation model will provide two main functions: (1) the necessary framework for proper interpolation in time and space, and (2) the ability to propagate in time based on the model's dynamics. Furthermore, because of the dynamical coupling of the thermosphere-ionosphere-magnetosphere system, data assimilation can provide adjustment to other physical parameters such as neutral wind, plasma drift, density and temperature of neutral constituents, which are of paramount importance as inputs to ionospheric models for successful forecasting.

7.1 General Approach

The success of ionospheric modeling depends mainly on accurate knowledge of the forces (drivers) which enter into the hydrodynamic plasma equations for the ionosphere. These include solar EUV and UV radiation, magnetospheric electric fields, particle precipitation, dynamo electric fields, thermospheric neutral densities, temperatures and winds. Therefore, ideally, one needs a coupled magnetospheric-thermospheric-ionospheric model that assimilates global and continuous data in all these regions. This is a goal of the space weather program which is likely to be reached and become operational (just like operational weather models) in a few decades. However, a first step toward reaching that goal is to learn to do data assimilation optimally in each region independently, before moving on to the more complicated coupled models. Furthermore, it is important in data assimilation to carefully study the impact of each new data type and its noise on the stability and accuracy of the model.

In order to assess the use of GPS-TEC data obtained from ground and space on ionospheric specification and forecasting, we start with a simplified low- and mid-latitude ionospheric model and use the Kalman filter in the context of an Observation System Simulation Experiment (OSSE). An OSSE is a controlled simulated experiment where data are synthesized based on the “true” state of the medium being studied, and then assimilated after adding the proper kind of noise. OSSEs are commonly used in NWP to assess the impact of new data types and/or new assimilation schemes [*Atlas*, 1997].

In our OSSE scenario (Fig. 15) a “true” ionospheric state is generated based on model 1 (to be described later). This “truth run” is meant to represent the true state of the ionospheric electron density. Data are then generated by assuming a distribution of ground and space receivers tracking GPS. Measurement and representation error covariances are assigned to the synthesized data. A perturbation run is generated by a different model (model 2). The

assimilation run uses the apriori state produced by the perturbation run and model 2 for forward propagation in time, in conjunction with the data. The impact of the assimilated data is then evaluated by comparing the difference between the assimilation run from the truth (error 2) to the difference between the perturbation and the true runs (error 1). More detailed descriptions of each of these steps are given below along with some results. But first we introduce some basic terminology and summarize the Kalman filter.

7.2 Basic Terminology and the Kalman Filter

We introduce the following definitions (commonly used in NWP)

x_k^t	The <i>true state</i> : a discrete representation of the true underlying continuum of the ionospheric state (electron density) at time k
$x_k^a = \langle x_k^t / m_k^o, x_k^f \rangle$	The <i>analysis</i> : an estimate of x_k^t given measurements at time k , and a forecast x_k^f
$x_k^f = \langle x_k^t / m_{k-1}^o \rangle$	The <i>forecast</i> : an estimate of x_k^t given measurements up to time $k - 1$
$m_k^o = H_k x_k^t + \varepsilon_k^o$	The observations are related linearly to the true state through an <i>observation operator</i> H_k . ε^o is an <i>observational error</i>
$\varepsilon_k^o = \varepsilon_k^m + \varepsilon_k^r$	The observational error is composed of the <i>measurement error</i> , ε^m , and a <i>representativeness error</i> , ε^r
$x_{k+1}^t = \Psi_k x_k^t + \varepsilon_k^d$	The true state at time $k+1$ is related to the state at time k via the <i>forward model</i> Ψ_k and a <i>model error</i>

Let M_k , R_k and Q_k be the measurement, representativeness and model error covariances, respectively. Then the Kalman filter can be summarized by the following set of Equations:

$$x_k^a = x_k^f + K_k (m_k^o - H_k x_k^f) \quad (18.a)$$

$$K_k = P_k^f H_k^T (H_k P_k^f H_k^T + R_k + M_k)^{-1} \quad (18.b)$$

$$P_k^a = P_k^f - K_k H_k P_k^f \quad (18.c)$$

$$x_{k+1}^f = \Psi_k x_k^a \quad (18.d)$$

$$P_{k+1}^f = \Psi_k P_k^a \Psi_k^T + Q_k \quad (18.e)$$

K is known as the Kalman gain, P^a and P^f are the analysis and the forecast covariances.

Some of the terms defined here will be explained in detail below (see also *Howe et al.*, 1998; *Ruffini et al.*, 1999). During the data assimilation process, time is discretized (indexed by k in Eqs. 18) where the state during a time step (taken to be 15 minutes in our analysis) is assumed constant. At time $k = 0$, starting with a forecast (apriori) state, x_0^f , a forecast state covariance, P_0^f , and a set of observations, m_0^o , we can update the state by use of Eqs. (18.a-c) to obtain the analysis and its covariance at time $k = 0$. Using the forward model Ψ_k we can propagate the analysis and its covariance from time 0 to time 1 via Eqs. (18.d and 18.e) to obtain the forecast state and covariance to be used as apriori for the next time step, and so on.

7.3 Dynamical Modeling

Our physical modeling is based on a simplified version of the University of Sheffield Plasmasphere/Ionosphere Model (USPIM) [Bailey and Shellek, 1990; Bailey et al., 1993] for middle and low-latitude regions. In our model we solve the conservation of mass and momentum equations for a plasma accounting for production, loss, transport of the major ion species in the F-region (O^+), neutral wind, and electric forcing. These equations can be written as

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{V}) = (P - L) \quad (19)$$

$$-\nabla(nk_B T) + nM\mathbf{g} + cn(\mathbf{E} + \mathbf{V} \times \mathbf{B}) - nM\nu(\mathbf{V} - \mathbf{U}) = 0 \quad (20)$$

n is the ion number density; \mathbf{V} is the velocity; P and L are production and loss rates, respectively; k_B is Boltzmann's constant; T is temperature, M is the ion mass, \mathbf{g} is gravitational acceleration, \mathbf{E} and \mathbf{B} are electric and magnetic fields, respectively; ν is the collision frequency for moment transfer between the atomic oxygen ion and the neutral particles; \mathbf{U} is neutral wind. An equation similar to Eq. (20) can be obtained for electrons, and after ignoring terms that are multiplied by the electron's mass, we obtain

$$-\nabla(n_e k_B T_e) - en_e(\mathbf{E} + \mathbf{V}_e \times \mathbf{B}) = 0 \quad (21)$$

In addition we also have

$$n_e = n, \quad n_e V_e = nV \quad (22)$$

We solve for the ion and electron densities by solving the above equations and making use of the empirical or parameterized models of the thermosphere (MSIS), thermospheric winds from the Horizontal Wind Model (HWM), solar EUV (SERF2, [Tobiska, 1991]), and electric fields [e.g., Heppner and Maynard, 1987; Fejer and Scherliess, 1997].

Traditionally, the equations of many ionospheric models are rewritten in a moving Lagrangian coordinate frame. The motion of this coordinate frame is dictated by the plasma drift perpendicular to the geomagnetic field lines. That approach introduces significant computational efficiency by transforming a time-dependent partial differential equations in 3-D into a family of time-dependent differential equations in 1-D. However, this 1-D approach introduces a significant difficulty to data assimilation, since the measurements are taken in 3-D space from ground-to-satellite and satellite-to-satellite links, making the mapping between data and the model parameter space very difficult to construct. In addition, the moving frame creates technical difficulties in applying some well-established data assimilation algorithms such as the adjoint approach.

Because of these considerations, we solve Eqs. (19-22) in a fixed Eulerian frame. Figure 16 shows the region being modeled and its division into pixels. These pixels are bounded by constant geomagnetic field lines, constant geomagnetic potential lines and local time, where the \mathbf{B} field is that defined by an eccentric tilted dipole.

7.4 "Truth" Run

A "true" ionospheric state was created as a function of latitude, height and local time. The 2-D region modeled and the p-q coordinates are depicted in Fig. 16. The region modeled is

bounded by [-30N,+30N] geomagnetic latitude, 100-1000 km altitude, with 2 deg. horizontal and 40 km vertical resolution (a total of ~1100 pixels). The run was done for 16 hours between 700 and 2300 LT (after an initialization run of 72 hours) based on EUV, $\mathbf{E} \times \mathbf{B}$ zonal drift, electron temperature, and neutral densities, temperatures and winds, derived from climatological models for solar maximum summer solstice conditions. The $\mathbf{E} \times \mathbf{B}$ zonal drift used is shown in Fig. 17 (red curve).

7.5 Perturbation Run

A perturbation run (to be used as an apriori) was also generated for the same region and time period by use of the same physical model described in Sec. 7.3 except that some driving forces were modified. In particular, we use a different vertical $\mathbf{E} \times \mathbf{B}$ drift (shown in blue in Fig. 17), and reduce the neutral density of atomic oxygen by a factor of 2. The true run, perturbed run and the difference are shown at 700 LT in Fig. 18. The electron densities for the perturbed run are nearly half of those of the “truth” run and they remain so at all local times, mainly due to reducing the neutral atomic oxygen density by a factor of 2.

7.6 Generation of Data

For the purpose of our study we assumed 4 GPS transmitters and 6 GPS receivers: 5 ground receivers evenly separated by 20 deg. tracking GPS at the rate of one data point every 3 minutes, and one LEO receiver at 700 km altitude with fore and aft antennas tracking the GPS satellites at the rate of 0.1 Hz. All stations and satellites are assumed to stay in the same 2-D magnetic longitudinal plane at different local times of the runs, which implies that the satellites are precessing at the same rate as the Earth. This is clearly not the case in real life, but with a constellation of LEO satellites such as COSMIC, it is not far fetched since one can combine data from different LEO and GPS satellites at various local times. In practice, the problem is inherently 3-D, but our aim here is to simplify the problem sufficiently in order to demonstrate the use of ionospheric data assimilation. A further simplification is made by using absolute TEC (as suppose to differential TEC) as our measurements and ignoring any biases (which are normally ~1-3 TECU as described in Sec. 2.3.2). Fig. 19 shows the data links for the first two hours of our OSSE. The color indicates the time when the data were collected. The dense samples correspond to LEO data acquisition starting from 30 deg. elevation relative to the LEO local horizon plane down to the surface. Some of the links do not pass through the region being modeled and were therefore discarded.

Given the receiver-transmitter geometry, synthetic TEC values are generated based on the “underlying continuum of the true” electron density (which will be explained more later). Line-of-sight TEC values are shown in Fig. 20.a for the entire 17 hours of the OSSE. The large TEC values (up to 600 TECU) are associated with occultation links. Each of the narrow U shapes in Fig. 20.a correspond to two occultations, one rising and one setting as seen in Fig. 20.b. Since our OSSE did not include a plasmasphere, TEC is nearly 0 at the beginning or the end of an occultation. On average, there are nearly 200 ground data links and 2 occultations every 1 hour.

7.7 Observational Errors

A major part of data assimilation is to assign appropriate covariances to the measurements and representativeness errors. Measurement errors are due to instrumental noise. Representativeness error is due to the finite representation of the model to the underlying continuum both in time and space. To understand the latter type of error better we consider the difference between TEC data generated based on a very fine space and time gridding (which is the actual data being assimilated) and TEC data for the same links and same driving forces but for the coarser spatial (shown in Fig. 16) and temporal grid (15 minute). The difference in TEC for each link is a measure of the representativeness error and is shown in Fig. 21.a and 21.b.

Based on the histogram of the representativeness error, and by excluding differences larger than 20 TECU, we estimate a 1- σ representativeness error of 2 TECU. Therefore, in our assimilation we take R in Eq. (18.b) to be $\mathbb{1} \times \sigma_P^2$ where $\mathbb{1}$ is the unit matrix and $\sigma_P = 2$ TECU. We also take M in the same equation to be $\mathbb{1} \times \sigma_D^2$ where $\sigma_D = 0.1$ TECU.

7.8 Assimilation Runs

In order to carry out the assimilation process according to Eqs. (18), we still need to define the covariance for the apriori state, P_0 , and the model error covariance, Q_k . We choose P_0 and Q_k to be diagonal matrices with elements given by

$$P_0 = \text{Diag}(n_{e_1}^2(0), n_{e_2}^2(0), n_{e_3}^2(0), \Lambda) \quad (23.a)$$

$$Q_k = 0.1 \text{Diag}(n_{e_1}^2(k), n_{e_2}^2(k), n_{e_3}^2(k), \Lambda) \quad (23.b)$$

where $n_{e_i}(k)$ is the electron density in pixel i at time k . Fig. 22.a shows the “true” electron density at 1900 LT. Fig. 22.b shows the electron density “analysis” obtained by using the recursive formulas of Eqs. (18). Fig. 22.c shows the difference between the truth and analysis. The “true” electron density state, including the equatorial anomaly, is very well captured by the analysis, with fractional error of $\sim 5\%$ near the maximum densities. This is encouraging because it implies that when there is a sufficiently dense set of TEC links, as would be the case with COSMIC, it is still possible to get accurate electron densities even if the driving forces are poorly known (e.g., substantially reduced neutral density and no prereversal enhancement in vertical drift). However, as demonstrated below, the physics still has a significant impact on the inversion.

For operational space weather purposes, some practical considerations such as computer memory and time limitations ought to be considered. Even though we will not discuss these issues in detail here, we do point out that no existing operational weather model (for the neutral atmosphere) uses the full Kalman filter (i.e., keeps and updates the full covariance of the state, which is usually of order 10^6 parameters). For the recursive Kalman filter 16-hour run presented above, it takes nearly 8 hours of CPU time on a dual processor DEC Alpha to assimilate ~ 4000 links. As a first attempt to reduce this processing time, we tested an approximate solution to the Kalman filter where Eq. 18.c is replaced by

$$P_k^a = P_k^f \quad (24)$$

while the other equations in (18) are kept the same. This saves us the expensive operation of computing $K_k H_k P_k^f$; the state covariance is still updated through Eq. (18.e). This reduces the computer run time from 8 hours to 1 hour without any significant impact on the solution as shown in Fig. 22.d and f. The reason for the similarity in the two solutions is that the rank of H_k (the observation matrix for data taken within 15 minutes) is small. Therefore, the impact $K_k H_k P_k^f$ has on updating P_k^f through Eq. (18.c) is much smaller than the impact of the forward propagation of P_k^a and the addition of the model error covariance through Eq. (18.e). This is the first evidence that the physical model plays an important role in the assimilation.

To further assess the impact of the physical model we test a further approximation where the recursive inversion is done using Eqs. (18.a, 18.b, 24) and

$$x_{k+1}^f = x_k^a \quad (25.a)$$

$$P_{k+1}^f = P_k^a \quad (25.b)$$

where no forward model is used. We refer to this approximation as tomography, although this is not strictly accurate, since in tomography Eq. (18.c) can still be used. The analysis of this approximation and the associated error are shown in Figs. 22.f and g. The degradation of the reconstruction relative to the previous solution is apparent. The benefit of data assimilation can be seen in Fig. 23, which is a measure of the performance of each technique: no data assimilation (perturbed run), tomography, approximate, and full Kalman.

8. Benefits to Space Weather

There are two important aspects to space weather, one scientific and one operational. On the scientific side the focus is on trying to improve our understanding and modeling of space weather, while on the operational side the focus is on the users and systems that might be affected by the space weather environment. COSMIC will make a tremendous contribution to both these aspects in three main areas: (1) Electron density specification and forecasting—which have been the main focus of our discussion thus far, (2) ionospheric scintillation and irregularities, and (3) TIDs. These three areas are not entirely decoupled, but for the purpose of clarity we divide our discussion below into the contribution of COSMIC to each of these areas.

8.1 Electron Density Specification and Forecasting

Accurate electron density specification allows us to better understand and model several very important space weather related phenomena. Most importantly, it will allow us to understand the response of the ionosphere to geomagnetic storms and substorms during their growth, main and recovery phases. Measurements of electron density will improve our understanding of the time scale over which the ionosphere responds to certain changes in the driving forces. Continuous and global measurements of the ionosphere and plasmasphere are needed for better characterization of diurnal, monthly, seasonal, and inter-annual cycles and for improving climatological models. In addition, electron density measurements will improve modeling of several space weather phenomena [see e.g., *Schunk and Sojka, 1995*] such as the production and evolution of boundary and auroral blobs, the generation of spread-F, sporadic E and intermediate layers. In addition to direct measurement of electron density,

several inferences concerning the thermosphere and the magnetosphere can be made from knowledge of the time evolution of electron density. For instance, NmF2 and HmF2 and their variation in time can be used to infer dynamo electric field and thermospheric wind.

On the operational side, specification of electron density can be used to estimate ionospheric delay along any line-of-sight which can be used for single frequency GPS users, deep space observations in the microwave band (such as from the Deep Space Network), single frequency altimeters (such as Geosat Follow On—GFO) and military surveillance applications.

8.2 Ionospheric Scintillations and Irregularities

Most of our discussions thus far has focused on electron density specification and forecasting from space TEC data. Equally important to space weather are GPS phase and amplitude scintillation measurements. Similarly to TEC measurements, scintillation measurements integrated over different lines-of-sight can be combined in a tomographic manner to produce 3-D maps of irregularities [Kunitsyn and Tereshchenko, 1992]. These measurements can also be used to localize electron density irregularities along the GPS-LEO ray path if the data quantity is insufficient for a tomographic localization [Vorob'ev *et al.*, 1999; Sokolovskiy, 2000].

Furthermore, by assimilating scintillation measurements into an ionospheric scintillation model, much can be learned about the mechanism of the generation and dissipation of these irregularities. For instance, ground-based GPS TEC measurements, in conjunction with NCAR's TIE-GCM, revealed that an irregularity event occurred across the northern U.S. continent during a major geomagnetic storm in October 1995 [Pi *et al.*, 2000]. The observed irregularities are believed to be due to the penetration of plasma convection and strongly sheared flow down to the middle latitudes, which not only deepened the middle latitude ionospheric trough but also may have produced plasma undulations and Kelvin-Helmholtz instabilities. An example of ionospheric irregularities observed during a geomagnetic storm through the global GPS network using the ROTI index [Pi *et al.*, 1997] is shown in Fig. 24.

8.3 Traveling Ionospheric Disturbances

Observations of the ionospheric E- and F-regions have shown the presence of oscillations resembling atmospheric planetary waves, typically observed in the mesosphere and lower thermosphere. These have indicated a 2-day oscillation in foF2, in addition to longer period waves of 5, 10 and 16 days, suggesting a causal connection between events in the neutral atmosphere and the ionosphere [Forbes and Zhang, 1996; Harris and Vincent, 1993]. In their analysis, Forbes and Zhang use a database of foF2 and foE from ionosonde measurements extending over a long period of time, while they approximate HmF2 by an empirical relation. Their database is limited to 20 geographical locations corresponding to 20 ionosonde stations.

Continuous, global and accurate recovery of electron density at all heights will facilitate a detailed study of the quasi two-day (QTD) wind oscillation as well as longer period

oscillations at different altitude ranges. Because the GPS occultation technique provides accurate measurements of neutral atmospheric density below 60 km and electron density above that height, these measurements are particularly suitable to establishing a causal connection, if any, between waves generated at the stratosphere and lower mesosphere and those in the ionosphere.

As an example, Fig. 25.a illustrates the sensitivity of radio occultation to atmospheric waves. Close agreement in both amplitude and phase with a radiosonde sounding taken some 300 km away (the profile shown in blue is in almost exact agreement with the GPS/MET profile) implies that the wave has a horizontal wavelength that is much larger than 300 km. The horizontal and vertical scales of this wave are consistent with a 5-day westward propagating Rossby-gravity wave [Kursinski *et al.*, 1996]. On the other hand, electron density derived in the region between 60-120 (25.b) shows an oscillation that might be indicative of a possible causal connection between the stratospheric wave and that in the D- and E-region. It should be noted that even though the electron density estimated in Fig. 25.b could be biased (due to the spherical symmetry assumption used in the retrieval), the point-to-point structure of the profile should be accurate.

COSMIC will provide the possibility to detect and analyze individual wave events in order to quantify and catalog the wave parameters, and to study and identify the relation between the source event and the resulting gravity wave and TID.

9. Conclusion

COSMIC GPS occultations in the ionosphere will yield accurate 3-D specification of electron density, ionospheric irregularities and global maps of TID's. The ~5600 global daily occultations from COSMIC can be processed individually to produce high-resolution vertical profiles of electron density, or collectively by properly assimilating them into a physically based ionospheric model. This kind of continuous global 3-D specification of the ionosphere will have tremendous payoffs for improving our understanding and the modeling of various space weather phenomena, as well as serving operational needs such as communication and navigation. Just as the number of GPS ground receivers has grown substantially over the past decade, the number of space-based GPS receivers collecting occultations will also witness substantial growth in the next decade, to the point where GPS ionospheric occultation measurements will truly revolutionize the manner in which we sense the ionosphere and our understanding of it.

Acknowledgement

We thank Peter Lyster of the Data Assimilation Office for several helpful conversations on data assimilation. We thank Graham Bailey of the University of Sheffield for providing the USPIM code and for assistance in working with it.

References

- Atlas R., Atmospheric observations and experiments to assess their usefulness in data assimilation, *J. Meteor. Soc. of Japan*, 75 (1B) 111-130, 1997.
- Bailey, G. *et al.*, A modeling study of the equatorial topside ionosphere, *Ann. Geophys.*, 11, 263, 1993.
- Bailey, G., and R. Sellek, A mathematical model of the Earth's plasmasphere and its application in a study of He⁺ at L = 3, *Ann. Geophys.*, 8, 171, 1990.
- Bassiri S. and G. A. Hajj, Higher-order ionospheric effects on the Global Positioning System observables and means of modeling them, *manuscripta geodaetica*, 18:280-289, 1993.
- Bertiger W. and S. Wu, Single frequency GPS orbit determination for low earth orbiters, *Proc. of ION Tech. Meeting*, Institute of Navigation, Santa Monica, Jan. 1996.
- Born, M., and E. Wolf, *Principles of Optics*, 6th ed., Pergamon, Tarrytown, N. Y., 1980.
- Dymond, K. F., S. E. Thonnard, R. P. McCoy, and R. J. Thomas, An optical remote sensing technique for determining nighttime F region electron density, *Radio Sci.*, 32:1985-1996, 1997.
- Fejer, B. G., and L. Scherliess, Empirical models of storm time equatorial zonal electric fields, *J. Geophys. Res.*, 102, 24047, 1997.
- Forbes J. M. and X. Zhang, Quasi 2-day oscillation of the ionosphere: A statistical study, *J. Atmo. and Solar-Terr. Phys.*, 59:1025-1034, 1997.
- Hajj G. A., R. Ibañez-Meier, E. R. Kursinski, L. J. Romans, Imaging the Ionosphere with the Global Positioning System, *Internat. J. of Imag. Sys. and Tech.*, 5, 174-184, 1994.
- Hajj G.A. and L.J. Romans, Ionospheric Electron Density Profiles Obtained With the Global Positioning System: Results From the GPS/MET Experiment, *Radio Sci.*, 33:175-190, 1998.
- Harris T. J. and R. A. Vincent, The quasi-two-day wave observed in the equatorial middle atmosphere, *J. Geophys. Res.*, 98(10), 1993.
- Heppner, J. P., and N. C. Maynard, Empirical high-latitude electric-field models, *J. Geophys. Res.*, 92 (A5), 4467-4489, 1987.
- Howe, B. M., K. Runciman, J. A. Secan, Tomography of the ionosphere: Four-dimensional simulations, *Radio Sci.*, 33(1):109-128, 1998.
- Karayel E. T., D. P. Hinson, Sub-Fresnel-scale vertical resolution in atmospheric profiles from radio occultation, *Radio Sci.*, 32(2), pp 411-423, 1997.
- Kunitsyn V. E., E. E. Tereshchenko, Determination of the turbulent spectrum in the ionosphere by a tomographic method, *J. Atmo. and Solar-Terr. Phys.*, 54: (10) 1275-1282, 1992
- Kursinski, E. R., et al., Initial results of radio occultation observations of Earth's atmosphere using the Global Positioning System, *Science*, 271, 1107-1110, 1996.
- Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. Hardy, Observing earth's atmosphere with radio occultation measurements using the global positioning system, *JGR*, 102(D19), 23429-24465, 1997.
- Lu, G., X. Pi, A. D. Richmond, and R. C. Roble, Variations of total electron content during geomagnetic disturbances: A model/observation comparison, *Geophys. Res. Lett.*, 25, 253, 1998.
- Mannucci A.J., B.D. Wilson, D.N. Yuan, C.M. Ho, U.J Lindqwister, and T.F. Runge, A global mapping technique for GPS-derived ionospheric total electron content measurements, *Radio Sci.* 33(3), 565-583, 1998.

- Meier, R. R., Ultraviolet spectroscopy of and remote sensing of the upper atmosphere, *Space Sci. Rev.*, 58, 1-185, 1991
- Pi, X., A. J. Mannucci, U. J. Lindqwister, and C. M. Ho, Monitoring of global ionospheric irregularities using the worldwide GPS network, *Geophys. Res. Lett.*, 24, 2283, 1997.
- Pi, X., G. Lu, B. A. Iijima, A. J. Mannucci, L. Sparks, B. D. Wilson, A. D. Richmond, and R. G. Roble, Midlatitude ionospheric disturbances and irregularities during a geomagnetic storm: A joint study using GPS and TIE-GCM, to be submitted to *J. Geophys. Res.*, 2000.
- Reilly, M. H., A transionospheric radio propagation model, *Proc. of the Ionospheric Effects Symposium*, Alexandria, VA, May 4-6, 1993.
- Rius, A., G. Ruffini and A. Romeo, Analysis of ionospheric electron density distribution from GPS/MET occultations, *IEEE Trans. On Geosc. & Remote Sensing*, 36: (2) 383-394, 1998.
- Rius, A., G. Ruffini and L. Cucurull, Improving the vertical resolution of ionospheric tomography with GPS occultations, *GRL*, 24: (18) 2291-2294, 1997
- Ruffini, G., L. Cucurull, A. Flores, A. Rius, A PIM-Aided Kalman filter for GPS tomography of the ionospheric electron content, *Phys. and Chem. of Earth, C-Solar-Terr. and Planet. Sci.* 24: (4) 365-369 1999
- Schreiner, W. S., S. V. Sokolovskiy, C. Rocken, D. C. Hunt, Analysis and validation of GPS/MET radio occultation data, *Radio Sci.* 34: 949-966 1999.
- Schunk R. W. and J. J. Sojka, Ionosphere-thermosphere space weather issues, *J. Atmo. and Solar-Terr. Phys.*, 58:1527-1574, 1996.
- Sokolovskiy, S. V., "Inversions of radio occultation amplitude data", *Radio Sci.*, 35(1), pp 97-105, 2000.
- Straus, P. R., Correcting GPS occultation measurements for ionospheric horizontal gradients, *Proc. of Iono. Effect Symp.*, Alexandria, VA, June 1999b.
- Straus, P. R., Ionospheric remote sensing using GPS, *AIAA Conference Proc.*, (AIAA-99-0628), Reno, NV, Jan 13, 1999a.
- Tobiska, W. K., Revised solar extreme ultraviolet flux model, *J. Atmos. Terr. Phys.*, 53, 1005, 1991.
- Tricomi, F. G., *Integral Equations*, Dover, Mineola, N. Y., 1985.
- Vorob'ev V.V.; A. S. Gurvich, V. Kan, S. V. Sokolovskii, O.V. Fedorova, A. V. Shmakov, Structure of the ionosphere based on radio occultation data from GPS Microlab-1 satellites: Preliminary results, *Earth Observation and Remote Sensing*, 15(4), pp 609-622, 1999.
- Wu J. T., S. C. Wu, G. A. Hajj, W. I. Bertiger and S. M. Lichten, Effects of antenna orientation on GPS carrier phase, *manuscripta geodaetica*, 18:91-98, 1993.

Figure Captions:

Fig. 1: Occultation geometry defining a , r , α and the tangent point and showing the split of the L1 and L2 signals due to the dispersive ionosphere.

Fig. 2: Representative daily COSMIC occultation coverage.

Fig. 3: All COSMIC occultation links in 98 minutes (at the rate of 1 link per minute). Each line corresponds to the surface projection of a LEO-GPS link with the LEO satellite on one end and a point intersecting a 700 km ionospheric shell on the other end. Dark lines are from one COSMIC satellite, gray lines are from the other 7 COSMIC satellites.

Fig. 4: Average number of rising, setting, grazing and total daily occultations from COSMIC per $(1000\text{km})^2$

Fig. 5: A map showing the distribution of the global GPS network. Circles around the stations enclose projection areas at 450 km altitude, in which GPS data are taken at elevation angles higher than 10 degrees.

Fig. 6: Absolute TEC between A and B determined by subtracting TEC measurements taken at point A from that taken at point B,

Fig. 7: Electron density profiles from GPS/MET and Millstone Hill ISR. (after *Hajj and Romans*, 1998)

Fig. 8: GPS/MET profiles of electron density indicating the ability of GPS occultations to resolve sharp layers in the ionosphere. (after *Hajj and Romans*, 1998)

Fig. 9: Scatter plot of NmF2 obtained from GPS/MET profiles and nearby ionosonde (a), and a histogram of the difference (b). The $1-\sigma$ for the NmF2 difference is $0.15 \times 10^{12} \text{e/m}^3$.

Fig. 10: (a-top) GPS/MET derived NmF2 for nearly 4200 occultations used in the statistics of Fig. 9 as a function of LT. Hourly averages and scatters are shown by the solid line and the vertical bars. (b-middle) Hourly standard deviation of GPS/MET and ionosonde NmF2 differences. (c-bottom) Hourly fractional differences of GPS/MET and ionosonde NmF2 obtained by dividing values in 10.b over the averages in 10.a.

Fig. 11: (a-upper-left) Global NmF2 distribution in units of particles/cm^3 for the solar maximum, equinox case discussed in the text. (b-upper-right) HmF2 distribution under the same conditions. (c-lower-left) Fractional errors in simulated NmF2 retrievals of idealized occultations under northward viewing conditions using horizontal gradients from EUV data (only possible at night in actual practice); and (d-lower-right) same as (c) but for HmF2 error. (after *Straus*, 1999b)

Fig. 12: (a-top) Contours of constant ionospheric density from the RIBG model together with occultation sensor lines-of-sight at -10° latitude, 270° longitude tangent point under solar

maximum, equinox conditions. Lines of sight are for tangent altitudes of 300, 400, 500, and 600 km. (b-middle) True vertical electron density profile at this location (solid), together with Abel (dotted), in-situ (dashed), and TEC (dash-dot) based inversions. (c-bottom) True values of $f_e(\phi)$ for the four lines of sight in (a) together with in-situ (dashed) and TEC (dash-dot) based estimates. (after *Straus* 1999b.)

Fig. 13: Sideview (top panel) and topview (bottom panel) of 10 sec. occultation links obtained from one GPS/MET orbital period, and the grid used in generating the images in Fig. 14.

Fig. 14: A sequence of 2-D tomographic reconstructions of electron density obtained in the GPS/MET orbital plane from nearly coplanar occultations taking place at consecutive revolutions on May 4-5, 1995. Angle 0 corresponds to local noon.

Fig. 15: A flowchart summarizing the OSSE scenario.

Fig. 16: Modeled region of the ionosphere and its parameterization with p (lines along magnetic field lines) and q (lines perpendicular to magnetic field lines).

Fig. 17: Two patterns of vertical $E \times B$ drift used to generate “true” and “perturbed” ionospheric runs. The red curve showing a prereversal enhancement was used for the “true” run.

Fig. 18: Electron density based on “true” run (top), perturbed run (middle) and the difference (bottom) at 700 LT.

Fig. 19: Data links for the first two hours of our OSSE. The color indicates the time when the data was collected, where 0 correspond to 700 local time.

Fig. 20: (a) Line-of-sight TEC generated based on the “true” electron density for 5 ground stations and 1 LEO satellite tracking 6 GPS satellites. (b) An enlargement of (a) for the period 1400-1420 LT showing two occultation tracks in detail. The peaks correspond to the time when the tangent point of the occultation link is near the F-peak.

Fig. 21: (a) Representativeness errors for the links of Fig. 20.a, and (b) their histogram.

Fig. 22: (a-g from top to bottom) (a) “True” electron density, and “analysis” electron density and corresponding error using the full Kalman (b-c), the approximate Kalman (d-e) and tomography (f-g). All densities are shown for 1900 LT.

Fig. 23: L1 norm of the difference between the perturbed or analysis state from the truth.

Fig. 24: Global ionospheric irregularities measured by the GPS global network during a 1995 magnetic storm. The rate of TEC index (ROTI) [*Pi et al.*, 1997] measures dual-frequency GPS phase fluctuations. One ROTI corresponds to 1 TEC Unit change per minute.

Figure 25: Atmospheric temperature (a) and electron density profiles (b) obtained from a GPS/MET occultation at 12:40 UT on 4 May 1995 in the south Pacific (8S, 168E). Also shown is a radiosonde profile obtained from a ship at 350 km from the occultation confirming the presence of what might be a Rossby-gravity wave. (Fig. (a) is after *Kursinski et al.*, 1996)