GPS and Ionosonde Data Fusion for Ionospheric Tomography

Karen Q.Z. Chiang, *Cornell University*
Mark L. Psiaki, *Cornell University*

BIOGRAPHIES

Karen Q.Z. Chiang is a GNC Engineer at MDA Corporation's Advanced Space Missions and Robotics Laboratory. She received a B.S. in applied physics from Columbia University in 2009, and a Ph.D. in Aerospace Engineering from Cornell University in 2014, specializing in dynamics, systems, and controls. Her areas of interest include estimation applications for remote sensing, as well as guidance and navigation systems for autonomous vehicles and spacecraft.

Mark L. Psiaki is a Professor in the Sibley School of Mechanical and Aerospace Engineering. He received a B.A. in Physics and M.A. and Ph.D. degrees in Mechanical and Aerospace Engineering from Princeton University. His research interests are in the areas of GNSS technology, remote sensing applications, and integrity, spacecraft attitude and orbit determination, and general estimation, filtering, and detection.

ABSTRACT

GPS measurements are combined with ionosonde measurements in an estimation problem for the state of the quiescent local ionosphere. This estimator has been developed to remotely sense the ionosphere above the High Frequency Active Auroral Research Program (HAARP) heater facility in Gakona, AK. The measurement model used by the estimator consists of a refractive ray-tracing model of ionosonde measurements for the electron density profile and of dual-frequency GPS measurements. A set of parameter sensitivity calculations augment the ray-tracing solutions in order to facilitate standard estimation-based model inversion calculations. The resulting algorithm determines an optimal parameterization of the ionosphere's electron density profile. Ionosonde data provide information about bottomside ionospheric layers, and trans-ionospheric GPS signal data enable observability of topside structure and a simple characterization of its variations with respect to latitude and longitude through implicit correlation of slant total electron content (TEC) integrals with the bottom side characterization. Experimental ionosonde and GPS data have been recorded and used to fit the ionosphere models. The result are ionosonde virtual height measurement fits within 8km, and frequency-differenced GPS pseudorange fits within 0.2 TEC units.

INTRODUCTION

Improved methods are sought to map the full altitude, latitude, and longitude electron density profile of the ionosphere above a certain locale, especially when the profile has been disturbed by, for example, the High Frequency Active Auroral Research Program (HAARP) heater in Gakona, AK. This capability would benefit areas of atmospheric science where an accurate ionospheric profile is needed. Ionosonde data may be used to generate sufficient representations of the profile below the peak F2 density. Until recently, a measurement-based characterization of the profile above this point could be obtained only from a space-based topside sounder [1] or from an incoherent scatter radar. Dual-frequency GPS data, however, makes topside reconstruction possible, albeit with higher levels of uncertainty.

Previous work has been performed that approximates topside behavior with simple model extrapolation, such as α-Chapman profile matching at the F2 peak [3][4]. Other authors have fused ionosonde data with GPS TEC measurements, but for the purpose of separating topside and bottomside TEC [5][6] by arithmetic differencing and employing an empirically determined upper transition height. The topside TEC determines the topside profile, but in somewhat of an ad hoc manner rather than by using optimal estimation techniques directly on both the ionosonde and TEC data.

The present paper's main contribution is the development and testing of a technique that optimally fuses GPS TEC data and ionosonde data to estimate a single parameterized local electron density profile. The optimal states mostly characterize the vertical profile above the ionosonde/GPS-receiver location, but also include simple latitude and longitude dependencies.

The ionosphere refracts radio waves propagating through its magnetoplasma, which affects group delay, carrier phase advance, and the geometric propagation path. The extent of each effect is dependent on wave frequency. Given the design of the dual-frequency GPS signal, the ray-path ge-

ometry dependency is negligible, and the other two dependencies can be modeled to first order in terms of the slant total electron content (TEC) along the line-of-sight (LOS) from a ground receiver to a GPS satellite.

One might like to use Ref. [2], however, to show that even the simplest of vertical electron density parameterizations, the Chapman parameterization, is not observable through from GPS slant TEC data alone. By including ionosonde group delay measurements, an optimal approach can be taken to simultaneously reconstruct both topside and bottomside.

The remainder of this paper is divided into seven sections. The first reviews Hamilton's equations for refractive ray-tracing of radio waves through a magnetoplasma. The next describes the parameterization chosen to model the ionosphere and its layers. The third and fourth sections discuss the algorithms that are used, in conjunction with the Hamiltonian ray-tracing equations, in order to model the ionosonde ray-path and observables and the observables' GPS integrals given the ionospheric electron density parameterization. This section also explains how to compute the partial derivative sensitivities of these observables with respect to electron density profile parameters. The fifth section presents the nonlinear optimal estimation methods used to fuse the ionosonde and GPS data in order to estimate the ionospheric parameterization. The sixth gives the ionospheric parameter estimation results that are obtained by this algorithm when using real data from a quiescent ionosphere. The last section contains the paper's conclusions.

HAMILTON'S EQUATIONS FOR REFRACTIVE RAY-TRACING IN A MAGNETOPLASMA

Refractive ray-tracing of a radio wave with frequency ω entails solving a coupled set of differential equations for the ray path, $\mathbf{r}$, and the wave-front unit direction vector $\boldsymbol{k}$. $k_0 = \omega/c$ is the free-space wave number. The ray-tracing equations are expressed with Hamiltonian formulations that are dependent on $\boldsymbol{r}$, $\boldsymbol{k}$, and k_0 [7]:

$$\frac{d\boldsymbol{r}}{dP'} = \frac{-[dH/d\boldsymbol{k}]^T}{dH/dk_0} \tag{1}$$

$$\frac{d\boldsymbol{k}}{dP'} = \frac{[dH/d\boldsymbol{r}]^T}{dH/dk_0} \tag{2}$$

where $P' = ct$ is the range-equivalent group delay, chosen as the independent variable for seamless integration through any possible spitze's at the top of an ionosonde path.

The Hamiltonian H is a function that must remain constant along a valid ray path and that can be derived from rewriting the dispersion relation. Two alternate valid forms of the Hamiltonian are used in the present study:

$$H_1(\boldsymbol{r},\boldsymbol{k};k_0,\boldsymbol{p}) = \frac{1}{2}\left[\frac{c^2}{\omega^2}\|\boldsymbol{k}\|^2 - n^2(\boldsymbol{r},\boldsymbol{k};k_0,\boldsymbol{p})\right] \tag{3}$$

$$\begin{aligned} H_2(\boldsymbol{r},\boldsymbol{k};k_0,\boldsymbol{p}) = real\Bigg\{&\Big[(U-X)U^2 - Y^2U\Big]\frac{k^4}{k_0^4} + X(\boldsymbol{k}\cdot\boldsymbol{Y})^2\frac{k^2}{k_0^4} \\ &+ \Big[-2U(U-X)^2 + Y^2(2U-X)\Big] \\ &\frac{k^2}{k_0^2} - X(\boldsymbol{k}\cdot\boldsymbol{Y})^2\frac{1}{k_0^2} + \Big[(U-X)^2 - Y^2\Big](U-X)\Bigg\} \end{aligned} \tag{4}$$

Where $X(\boldsymbol{r};k_0,\boldsymbol{p}) = \omega_{pe}^2/\omega^2$ is the squared ratio of the plasma frequency to the radio wave frequency, and $\boldsymbol{Y}(\boldsymbol{r};k_0) = Y\hat{\boldsymbol{b}} = (\omega_{ce}/\omega)\hat{\boldsymbol{b}}$ is a vector with magnitude equal to the ratio of the electron gyrofrequency to the radio signal's frequency and with direction parallel to the local magnetic field. $U = 1 - iZ = 1 - i\nu/\omega$, where ν is the electron collision frequency, is taken to be 1 in the assumed lossless magnetoplasma. $\boldsymbol{p}$ is a vector of parameters that characterize the electron density profile $N_0(\boldsymbol{r},\boldsymbol{p})$. It is needed to determine

$$\omega_{pe}^2 = \frac{q_e^2}{m_e\epsilon_0}N_0(\boldsymbol{r},\boldsymbol{p}) \qquad = 3.183\times10^3 N_0(\boldsymbol{r},\boldsymbol{p}) \tag{5}$$

The specific definition of $\boldsymbol{p}$ will vary depending on the model used. The parameter vector associated with the model of the present algorithm will be discussed in a later section. In Eq. (3), n is the phase index of refraction, given by the Appleton-Hartree formula for $Z = 0$ [8]:

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

where θ is the angle between wave vector $\boldsymbol{k}$ and magnetic field direction vector $\hat{\boldsymbol{b}}$.

The first Hamiltonian, the one in Eq. (3), is appropriate for free space, which may defined as regions where $X \le X_{min}(Y)$, with the index of refraction significantly different from zero. For ray-tracing through regions with higher electron densities and near-zero index of refraction, that are characterized by $X \ge X_{max}(Y)$, the second Hamiltonian formulation is used, possibly with a sign change. H_2 is especially needed in the neighbourhood of a spitze (singular point of reflection), illustrated in the ray path of Fig. 1, because this latter Hamiltonian and the corresponding ray-tracing differential equation in Eqs. (1)-(2) do not become a singularity in this case.

The bounds for X differ depending on whether the radio wave is ordinary (O-mode) or extraordinary (X-mode). For some chosen Hamiltonian transition tuning constants a_X and b_X, these limits take the form:

$$X_{min}(Y) = \begin{cases} a_X & \text{if O-Mode or } Y > 1 \\ a_X(1-Y) & \text{Otherwise} \end{cases} \tag{7}$$

$$X_{max}(Y) = \begin{cases} b_X & \text{if O-Mode or } Y > 1 \\ b_X(1-Y) & \text{Otherwise} \end{cases} \tag{8}$$

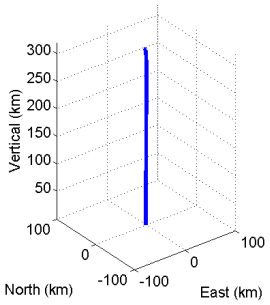

Figure 1: Example ray-path solution typical of an ionosonde signal, with a spitze that occurs at the upper reflection point, shown in both normal and magnified views.

where reasonable values for a_X and b_X are 0.1 and 0.25.

In between these two thresholds, i.e., $X_{min}(Y) < X < X_{max}(Y)$, the Hamiltonian used is a linear combination of Eqs. (3) and (4). Specifically, the transition Hamiltonian is

$$H_{12} = H_1\,[1-\lambda(\alpha)] \pm \lambda(\alpha) H_2 \tag{9}$$

where the H_2 term is positive if the quadratic polynomial in $\left(\frac{\|\boldsymbol{k}\|}{k_0}\right)^2$ that is equivalent to the 2^{nd} Hamiltonian in Eq. (4) has a non-negative slope at the root where $\left(\frac{\|\boldsymbol{k}\|}{k_0}\right)^2 = n^2$. Otherwise, it is negative. This creates a smooth transition, and avoids the partial derivatives of the two Hamiltonians being canceled out by each other. The new quantities introduced in Eq. (9) are defined by

$$\lambda(\alpha) = 10\alpha^3 - 15\alpha^4 + 6\alpha^5 \tag{10}$$

$$\alpha = \frac{X - X_{min}(Y)}{X_{max}(Y) - X_{min}(Y)} \tag{11}$$

These quantities combine in Eq. (9) in a way that causes a smooth transition from the free-space H_1 Hamiltonian to the high-electron-density H_2 Hamiltonian. The transition is a quintic spline in X that transitions from H_1 to H_2 with continuous 0^{th}, 1^{st}, and 2^{nd} derivatives in X.

The ray-tracing equations (1)-(11) model electromagnetic wave propagation in the zero-wavelength approximation limit of a lossless magneto-plasma, using the lossless Appleton-Hartree refractive index equation [8]. The Appleton-Hartree formula is used directly in the free-space case of Eq. (3), and effectively in the case of Eq. (4). Both H_1 and H_2 are non-dimensional, and are zero-valued along a valid ray path, as is their linear combination, H_{12}. H_1, H_2, or H_{12} can be substituted into Eqs. (1) and (2) for H.

The magneto-ionic parameters X and Y are required to complete this system of equations. Y is dependent on the background magnetic field, which can be obtained from a model such as the International Geomagnetic Reference Field (IGRF). X varies with electron density, the model for which will be covered in the following section.

IONOSPHERIC ELECTRON DENSITY PROFILE PARAMETERIZATION

There are several potential models that could be used to describe the ionospheric profile. Any candidate $N_0(\boldsymbol{r},\boldsymbol{p}$ profile must be continuous in $\boldsymbol{r}$ with continuous first and second derivatives in order to be useable in the ray-tracing equations (1)-(11). The need for first derivatives is obvious from the form of Eqs. (1) and (2). The need for second derivatives arises for two reasons. The first is the desire to use a Newton/shooting method to solve the underlying nonlinear two-point boundary value problem for ionosonde ray-tracing. The second reason for using second derivatives arises from the need to solve for Jacobian first-partial-derivative sensitivities of ionosonde ray-tracing solutions with respect to elements of $\boldsymbol{p}$.

One of the simplest and most common ionospheric electron density models that meets the continuity requirement is a superposition of Chapman layers. Other composite electron density models, such as the International Reference Ionosphere (IRI), may also be chosen. While these models have various advantages, Booker's "skeleton" model [9] has been chosen for its balance of simplicity, flexibility, and physical meaning. These qualities make it very adaptable to an estimation problem. The Booker model has a spline-like dependency on altitude when considered in the $ln[N_0(\boldsymbol{r},\boldsymbol{p}]$ form, it is based on nodes that may be placed at a series of critical altitudes. These nodes, along with transition scale-heights information give rise to a sufficiently smooth vertical profile. An analysis of this model can be found in Ref. [9]. Only one convenient adaptation will be presented in this paper.

Booker's vertical electron density profile can be fully described by $3m+2$ parameters, where m is the number of nodes. By adding two more parameters, this model can be augmented to include a first-order dependence of vertical total electron content (VTEC) on latitude and longitude. The total $3m+4$ variables comprise the ionospheric Booker

profile parameter vector:

$$\boldsymbol{p} = \begin{bmatrix} z_1 \\ \ln(z_2 - z_1) \\ \vdots \\ \ln(z_m - z_{m-1}) \\ \ln\left(\frac{1}{Z_{0,1}}\right) \\ \ln\left(\frac{-1}{Z_{m,m+1}}\right) \\ \ln(N_{01}) \\ \vdots \\ \ln(N_{0m}) \\ \ln(\tilde{z}_1) \\ \vdots \\ \ln(\tilde{z}_m) \\ \dfrac{\partial \ln(VTEC)}{d\phi} \\ \dfrac{\partial \ln(VTEC)}{d\lambda} \end{bmatrix} \tag{12}$$

where z_n is the altitude, N_{0n} is the electron density, and $\tilde{z}_n$ is the transition height for node $n = 1, \ldots, m$. The transition heights control the intervals over which the Booker skeleton slope of the $\ln(N_0)$ versus z function is smoothed to create an analytic function $N_0(z, \boldsymbol{p})$ that yields electron density, given an altitude and a parameterization. $Z_{0,1}$ and $Z_{m,m+1}$ are the first and last Booker skeleton signed scale heights, and ϕ and λ denote, respectively, latitude and longitude.

Without smoothing, the skeleton profile between the successive nodes n and $n+1$ is simply characterized by

$$\frac{1}{N}\frac{dN}{dz} = Z_{n,n+1}^{-1}; z_n < z < z_{n+1} \tag{13}$$

with respect to height z that is calculable from $\boldsymbol{r}$, along with ϕ and λ, according to the WGS-84 Earth model. Incorporating the full set of parameters yields the relations

$$ln\left[\frac{N_0(z)}{N_{00}}\right] = \frac{z - z_0}{Z_{0,1}} + \sum_{n=1}^{m}\left(\frac{\tilde{z}_n}{Z_{n,n+1}} - \frac{\tilde{z}_n}{Z_{n-1,n}}\right)\ln\frac{1 + \exp\left\{\dfrac{z - z_n}{\tilde{z}_n}\right\}}{1 + \exp\left\{\dfrac{z_0 - z_n}{\tilde{z}_n}\right\}} \tag{14}$$

$$\ln(N_{00}) = \ln(N_{00,nom}) + \frac{\partial \ln(VTEC)}{\partial \phi}\Delta\phi + \frac{\partial \ln(VTEC)}{\partial \lambda}\Delta\lambda \tag{15}$$

where $N_{00,nom}$ is the nominal electron density at reference height z_0, which is a weighted sum of the nodes:

$$z_0 = \sum_{n=1}^{m}\left(\frac{N_{0n}}{\sum_{k=1}^{m} N_{0k}}\right)z_n \tag{16}$$

The heavy use of natural logarithms in this version of Booker's model has the effect of ensuring the sign of various quantities or quantity differences that must be positive for a sensible profile. The ionospheric parameter vector $\boldsymbol{p}$ consists of the states that the present algorithm estimates. For each set of parameters, the coupled system of differential equations, Eqs. (1) and (2), can be solved to give the ray path, the algorithm for which is given in the subsequent section.

IONODSONDE RAY-TRACING SOLUTION FOR A GIVEN PROFILE PARAMETERIZATION

Although the ray-tracing method may be adapted for GPS signals, it does not produce significant discrepancies from a straight-line path due to the high frequency of the GPS L-band signals. The ionosonde ray-tracing case will be presented because the corresponding ray paths can be complicated and require care solution of Eqs. (1)-(2).

The ray-tracing differential equations in (1) and (2) can be assembled into a general state-space model that represents a nonlinear two-point boundary value problem:

$$\frac{d\boldsymbol{x}_{ray}}{dP'} = \boldsymbol{f}(\boldsymbol{x}_{ray}, \boldsymbol{p}) \tag{17}$$

$$\text{subject to} \quad \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\boldsymbol{x}_{ray}(0) = \begin{bmatrix} \boldsymbol{r}_i \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & I & 0 \end{bmatrix}\boldsymbol{x}_{ray}(P'_f) = \boldsymbol{0},$$

$$\text{and} \quad H(\boldsymbol{r}_i, \boldsymbol{k}_i; k_0, \boldsymbol{p}) = 0$$

where $\boldsymbol{x}_{ray} = [\boldsymbol{r} \quad \boldsymbol{k} \quad P]^T$ is the state vector, $\boldsymbol{r}_i$ is the known initial transmission point of the ray-path, P'_f is the unknown total range-equivalent group delay from the transmitter to the reflection point, and $\boldsymbol{k}_i = [0 \quad I \quad 0]\boldsymbol{x}_{ray}(0)$ is the initial wave vector. P'_f may be directly compared to virtual height data from the ionosonde. The second constraint in Eq. (17) is the boundary condition stipulates that at reflection, the final wave vector is zero.

This nonlinear problem can be solved by first-guessing the two unknown components of $\boldsymbol{k}_i$, which dictate its off-zenith and azimuth pointing angles, using the initial condition on the Hamiltonian in Eq. (17) in order to compute magnitude of $\boldsymbol{k}_i$, and numerically integrating using Runge-Kutta techniques up to P'_f. A wise guess would be an exactly vertical vector. At the end of the integration, the terminal reflection boundary condition is checked. If it has not been satisfied, then the initial direction of $\boldsymbol{k}_i$ and total group path P'_f are adjusted iteratively using Newton's method within this shooting scheme for solving the boundary value problem.

This shooting method requires calculating the sensitivities of the solution to the initial $\boldsymbol{k}_i$ direction variables and to P'_f. The sensitivities to each of the two independent initial directions of $\boldsymbol{k}_i$ can be determined by solving the following

initial value problem:

$$\frac{d}{dP'}\left(\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta}\right) = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}_{ray}}\right|_{\boldsymbol{x}_{ray}(P'),\boldsymbol{p}} \frac{\partial \boldsymbol{x}_{ray}}{\partial \eta} \tag{18}$$

where the initial condition $\left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta}\right|_0$ must obey

$$\begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta}\right|_0 = 0 \tag{19}$$
$$\left.\frac{\partial H}{\partial \boldsymbol{r}}\right|_0 [I \quad 0 \quad 0] \left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta}\right|_0 + \left.\frac{\partial H}{\partial \boldsymbol{k}}\right|_0 [0 \quad I \quad 0] \left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta}\right|_0 = 0$$

Given that Eq. (19) is 5 equations in 7 unknowns, there are only two possible linearly-independent values of $\left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta}\right|_0$: The two that correspond to initial directional variations of $\boldsymbol{k}_i$. Suppose that these are called $\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_1}$ and $\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_2}$. The variable η is a placeholder for of these two directional elements of $\boldsymbol{k}_i$.

The sensitivity differential equations represented by Eqs. (18)-(19) are integrated twice, once for $\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_1}$ and once for $\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_2}$, i.e., once for each independent directional sensitivity of $\boldsymbol{k}_i$. These integrations occur simultaneously with the numerical integration of Eq. (17). One can use a $4^{th}/5^{th}$-order Runge-Kutta approximations of the solutions for $\boldsymbol{x}_{ray}$ and $\partial \boldsymbol{x}_{ray}/\partial \eta$. This ODE integrator has been augmented to take advantage of the fact that the Hamiltonian must remain zero along a valid ray-path. After each $4^{th}/5^{th}$-order Runge-Kutta step, a numerical solver step is taken back towards the $H = 0$ manifold along a direction perpendicular to the local $H = constant$ manifold in the event that the $H \neq 0$ at the end of the step due to the build-up of numerical error in the integrator.

The Newton's method of solving the shooting problem works with the final two sensitivities $\left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_1}\right|_{P'_f}$ and $\left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_2}\right|_{P'_f}$. It also works with $\left.\frac{\partial \boldsymbol{x}_{ray}}{\partial P'_f}\right|_{P'_f} = f\left[\boldsymbol{x}_{ray}(P'_f), \boldsymbol{p}\right]$. These three sensitivities can be used in the following linearized version of the terminal boundary condition

$$[0 \quad I \quad 0]\left[\boldsymbol{x}_{ray}(P'_f) + \left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_1}\right|_{P'_f} \Delta\eta_1 + \left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_2}\right|_{P'_f} \Delta\eta_2 + \left.\frac{\partial \boldsymbol{x}_{ray}}{\partial P'_f}\right|_{P'_f} \Delta P'_f\right] = 0. \tag{20}$$

Solution of these three linearized terminal boundary conditions for the three unknowns $\Delta\eta_1$, $\Delta\eta_2$, and $\Delta P'_f$ gives the Newton increments for these quantities. The new initial state is $\boldsymbol{x}_{ray}(0) + \left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_1}\right|_0 \Delta\eta_1 + \left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \eta_2}\right|_0 \Delta\eta_2$ with a second-order adjustment to the length of $\boldsymbol{k}_i$ in order to ensure that the initial value of H is zero, and the new terminal group delay is $P'_f + \Delta P'_f$. The iteration is then repeated until the Newton increments $\Delta\eta_1$, $\Delta\eta_2$, and $\Delta P'_f$ all approach zero.

An outer step-size control algorithm has been included in the Newton shooting procedure. It uses the values $\alpha_{step}\Delta\eta_1$, $\alpha_{step}\Delta\eta_2$, and $\alpha_{step}\Delta P'_f$ to update $\boldsymbol{x}_{ray}$ and P'_f for each Newton increment, with the step length α_{step} chosen in the range 0 to 1 in a way that ensures decrease of the sum of the squares of the errors in the terminal boundary condition $[0 \quad I \quad 0]\boldsymbol{x}_{ray}(P'_f) = 0$.

Note that it can be helpful to use various non-dimensionalizations within the numerical integration of Eq. (17) and within Newton's method. The version implemented here replaces the ray-path position by its non-dimensionalized position $\boldsymbol{r}/P'_f$, and it replaces the wave vector by its non-dimensionalized form $\boldsymbol{k}/k_0$. The corresponding modifications to the other problem equations are straightforward.

The chosen $4^{th}/5^{th}$-order Runge-Kutta method has the capability to perform automatic step-size adjustment, but this feature is used only in an outer loop that executes outside the basic Newton/shooting-method solution of the nonlinear two-point boundary value problem. Otherwise, the Runge-Kutta step size adjustment could produce spurious results during the Newton step-size adjustment associated with α_{step}. The Runge-Kutta step size adjustment is designed to take small steps where there is a significant numerical integration error indicated by the $4^{th}/5^{th}$-order comparison or by the Hamiltonian deviations from 0 caused by each step prior to re-enforcement of the $H = 0$ constraint. Other two-point boundary value solvers are also possible substitutes for this portion of the algorithm . The algorithm presented has been found to work very well, usually converging in a few iterations.

Sensitivities with respect to aiming parameters and P'_f are needed first to solve shooting problem for increments to aiming parameters and P'_f. Once the nonlinear shooting problem is solved, then the linearized two-point boundary value solution is used to derive sensitivities of observables with respect to ionosphere model parameters. To this end, the elements of $\boldsymbol{p}$ may also be substituted in for η in Eq. (17). The sensitivities to $\boldsymbol{p}$ are needed in order to find its optimal values. The constraint for this version of Eq. (17) is

$$[0 \quad I]\left\{\left.\frac{\partial \boldsymbol{x}_{ray}}{\partial \boldsymbol{p}}\right|_{P'_f} + \boldsymbol{f}[\boldsymbol{x}_{ray}(P'_f), \boldsymbol{p}]\frac{\partial P'_f}{\partial \boldsymbol{p}}\right\} = \boldsymbol{0} \tag{21}$$

The same principles that solve the nonlinear shooting problem are applied to solve for $\frac{\partial \boldsymbol{x}_{ray}}{\partial \boldsymbol{p}}$ for P' ranging from 0 to

P'_f and for $\dfrac{\partial P'_f}{\partial \boldsymbol{p}}$. The latter is used in the estimation problem for $\boldsymbol{p}$.

GPS INTEGRAL SOLUTION FOR A GIVEN PROFILE PARAMETERIZATION

Although the same ray-tracing principles may be applied to GPS signals traversing the ionosphere, due to the relatively small deviation of the ray path from the direct (LOS), only perturbations to group delay integrated along the straight-line path between transmitter and receiver need be considered. These perturbations occur when the group-delay index of refraction differs from 1, and can be integrated over the path s to give the total effect [10]:

$$\begin{aligned}\delta P'_{fGPS}(\omega, \boldsymbol{p}) &= \int n'(\omega, \boldsymbol{p}) - 1 \; ds \\ &= \int n(\omega, \boldsymbol{p}) + \omega \frac{\partial n(\omega, \boldsymbol{p})}{\partial \omega} - 1 \; ds\end{aligned} \tag{22}$$

The total perturbation will depend on the GPS broadcast frequency. The start of the path can be found by determining the satellite's position in its orbit with broadcast ephemerides, and the end of the path is at the receiver. As before, the sensitivities of these variables can also be calculated by numerically integrating them concurrently with Eq. (22). $\partial(\delta P_{fGPS})/\partial \boldsymbol{p}$ and $\partial(\delta P'_{fGPS})/\partial \boldsymbol{p}$ are required to optimally estimate the ionospheric electron density distribution parameter vector $\boldsymbol{p}$ using the group delay measurements differenced between the L1 and L2 frequencies.

KALMAN FILTER DATA FUSION FOR OPTIMAL ESTIMATION OF IONOSPHERIC PARAMETERS

In order to find the optimal estimates of the profile parameters in $\boldsymbol{p}$, data from the ionosonde and dual-frequency GPS are fused using a Square Root Extended Kalman Filter (SREKF). The SREKF is known to be a numerically stable form of the extended Kalman filter (EKF), which is itself an approximation of the optimal Bayesian filter for nonlinear estimation problems. For details of the EKF and Square Root Information Filter, see Refs. [11], [12], and [13]. Only the models and initialization procedures for the filter will be given.

Measurement and Dynamic Models

The ray-traced ionosonde virtual heights and the integrated GPS L1/L2 frequency-differenced group delays, which are all functions of the state vector $\boldsymbol{p}$, are stacked as the vector of measurements that define the nonlinear measurement function $\boldsymbol{h}$:

$$\boldsymbol{h}(\boldsymbol{p}, DCB_{RX}) = \begin{bmatrix} P'^{1}_{f}(\omega_1, \boldsymbol{p}) \\ \vdots \\ P'^{n_I}_{f}(\omega_{n_I}, \boldsymbol{p}) \\ \Delta \boldsymbol{P}'^{1}_{fGPS}(\boldsymbol{p}) \\ \vdots \\ \Delta \boldsymbol{P}'^{n_{GPS}}_{fGPS}(\boldsymbol{p}) \end{bmatrix} \tag{23}$$

where DCB_{RX} is the unknown inter-frequency GPS receiver bias. The quantities $P'^{i}_{f}(\omega_i, \boldsymbol{p})$; $i = 1, \ldots, n_I$ are the virtual heights from the ionosonde, modeled by numerical integration of Eq. (17) for each sounding frequency ω_i. The frequency-differenced GPS group delay observables are:

$$\begin{aligned}\Delta \boldsymbol{P}'^{j}_{fGPS}(\boldsymbol{p}) &= \boldsymbol{P}'^{j}_{fL2} - \boldsymbol{P}'^{j}_{fL1} \qquad ; \quad j = 1, \ldots, n_{GPS} \\ &= \Delta \boldsymbol{\Psi}^{j}(\boldsymbol{p}) - c(DCB^{j} + DCB_{RX}) + \Delta \boldsymbol{w}^{j}_{\Psi}\end{aligned} \tag{24}$$

where the vectors indicate the possibility of stacking of multiple samples in time. Each satellite may have its own time series of measurements; the bookkeeping of time is only necessary for determining satellite location – the starting point of path integration. The vector $\Delta \boldsymbol{\Psi}^{j}(\boldsymbol{p}) = \delta \boldsymbol{P}'^{j}_{fGPS}(\omega_{L2}, \boldsymbol{p}) - \delta \boldsymbol{P}'^{j}_{fGPS}(\omega_{L1}, \boldsymbol{p})$ contains the differences in group delay between the L1 and L2 GPS frequencies for satellite j as modeled using Eq. (22). These differential quantities are used for the GPS data because differencing removes the unknown effects of satellite to receiver geometry, receiver clock error, and satellite clock error. These differences only contain total electron content (TEC) information. Moreover, it is assumed that the differential pseudoranges have undergone carrier-smoothing, via one of the many existing methods [14],[15], such that the measurements entering into Eq. (24) are both precise and ambiguity-free.

Equation (24) includes real-world effects on its right-hand side that model the ways in which actual receiver-generated frequency-differenced group delay varies from their theoretically modeled values in $\Delta \Psi^{j}(\boldsymbol{p})$. These differences arise from various instrument biases and noise terms. DCB^{j} is satellite j's (inter-frequency) differential code bias, which can be obtained from the Center for Orbit Determination in Europe (CODE) [16]. DCB_{RX}, as before, is the differential code bias of the receiver. $\Delta \boldsymbol{w}^{j}_{\Psi}$ and Δw_{Φ} is a general white measurement noise term for the j^{th} satellite.

DCB_{RX} is unknown and needs to be estimated along with $\boldsymbol{p}$. Therefore, an augmented state vector $\boldsymbol{x}_k = [\boldsymbol{p} \quad DCB_{RX}]^{T}_{k}$ is constructed for the SREKF measurement model, which takes the final form:

$$\boldsymbol{y}_k = \boldsymbol{h}(\boldsymbol{x}_k) + \boldsymbol{w}_k \tag{25}$$

where $\boldsymbol{y}_k$ is a stacked vector consisting of the raw ionosonde virtual height data and frequency-differenced pseudorange and carrier phase data from GPS, at filter step

k. The zero-mean Gaussian white-noise vector $\boldsymbol{w}_k$ is comprised of the ionosonde virtual height measurement noise components for the sounding frequencies $\omega_1,...,\omega_{n_I}$, and the frequency-differenced GPS pseudorange measurement noise components $\Delta\boldsymbol{w}_\Psi^1,...,\Delta\boldsymbol{w}_\Psi^{n_G PS}$. This noise vector has an associated covariance matrix

$$R = \text{diag}\left(\sigma_1^2,\ldots,\sigma_{n_I}^2,\boldsymbol{\sigma}_{\Psi 1}^2 I,\ldots,\boldsymbol{\sigma}_{\Psi n_{GPS}}^2 I\right) \quad (26)$$

which has diagonal entries comprised of the variances of the measurement noise terms. Ionosonde precision is typically within a few kilometres, and therefore a standard deviation around $\sigma = 1km$ would be a good choice. Similarly, each $\sigma_\Psi = 0.01m$ are reasonable values to use for the measurement uncertainty in frequency-differenced, carrier-smoothed GPS pseudorange.

This effort seeks to estimate snapshots of the ionospheric background electron density profile that span the order of a few minutes. Therefore, the dynamic model for the parameter vector $\boldsymbol{p}$ is assumed to be a random walk model:

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k + \boldsymbol{v}_k \quad ; \; v \sim \mathcal{N}(0,\epsilon I) \quad (27)$$

where ϵ is on the order of 10^{-6}. This small value of white process noise covariance has the effect of modeling the parameters in p as being nearly constant over the data spans of interest.If a longer span is needed, a Markov-process dynamic model for the $\boldsymbol{p}$ vector like that in Ref. [2] might prove useful.

The inter-frequency GPS receiver bias is modeled as being constant. Thus, its dynamic model takes the form:

$$DCB_{RX,k+1} = DCB_{RX,k} \quad (28)$$

Filter Initialization

The initial guess for the ionospheric parameter vector $\boldsymbol{p}$ is generated by utilizing the auto-scaled numerical real height function from the ionosonde's software, ARTIST-5. Figure 2 is an example ionogram processed by ARTIST-5. The real heights are used in a Levenberg-Marquardt algorithm that minimizes the cost

$$J(\boldsymbol{p}) = \frac{1}{2}\sum_{i=1}^{M}\left[N_{Ri} - N_0(z_i,\boldsymbol{p})\right]^2 + \frac{1}{2}w\sum_{n=1}^{m}\left[z_{ng} - z_n(\boldsymbol{p})\right]^2 \quad (29)$$

where N_{Ri} for $i = 1,\ldots,M$ is the electron density at the real height z_i determined by the ionosonde's software. This is compared to the aforementioned Booker density model $N_0(z_i,\boldsymbol{p})$. z_{ng} is an input guess of the n^{th} node altitude z_n. Examples of node placement can be found in Ref. [9] and in Fig. 3, where the blue curve denotes the logarithmic skeleton profile that simply joins the nodes without smoothing, i.e., only uses the z_n and N_{0n} terms of $\boldsymbol{p}$. To avoid an ill-conditioned problem, the nodes are typically the approximate inflection points in the real density function. The red dashed curve is the full logarithmic profile

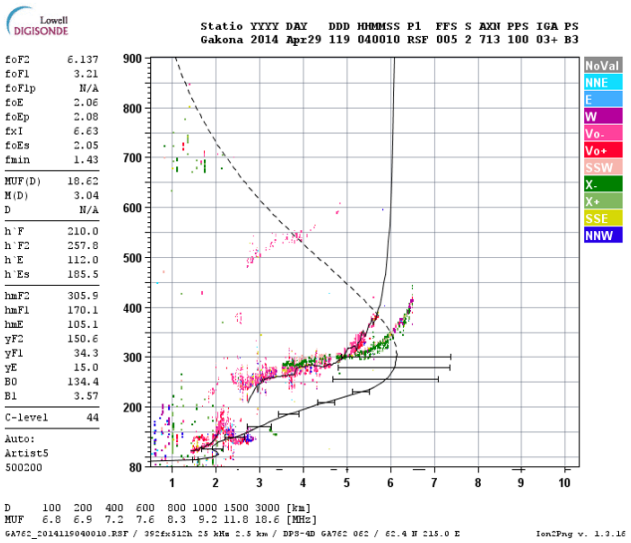

Figure 2: Example ionogram with virtual and real height profiles auto-scaled by ARTIST-5. Virtual heights are automatically fitted to the lowest echoes.

that incorporates the $Z_{0,1}$, $Z_{m,m+1}$, and $\tilde{z}_n$ terms as well. w is a positive scalar that weights the cost contributions of the differences between each final optimized altitude node value and z_{ng}. First guesses of other ionospheric parameters in the p vector used to initialize this Levenberg-Marquardt algorithm are produced by polynomial fitting and interpolation.

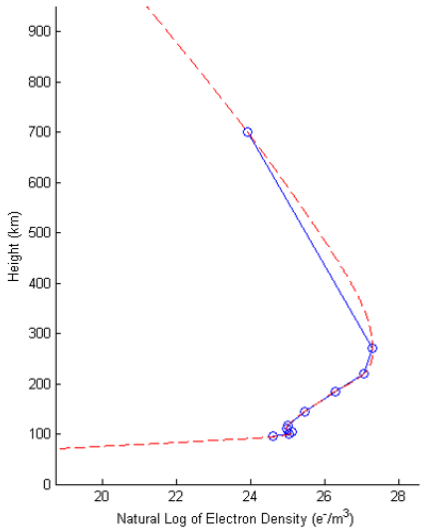

Figure 3: Example log of density vs. height function with nodes.

The resulting guess $\boldsymbol{x}_g = [\boldsymbol{p} \quad \boldsymbol{0}]^T$, optimized to minimize the cost function in Eqn. (29) for $\boldsymbol{p}$, is typically close to the optimal estimate of $\boldsymbol{x}$ produced by the SREKF of the

previous section for the bottom-side ionosphere, i.e. below the peak density altitude. However, the top-side carries higher uncertainty because only ionosonde information is incorporated into $\boldsymbol{x}_g$, and the topside distribution from the ionosonde data is based on a number of assumptions that are only rough approximations of reality. This discrepancy can be accounted for by a higher variance attributed to the initial electron density for the node in the top-side and attributed to the N_{Ri} vs. z_i ionosonde values for i values from the topside model. The initial state covariance matrix would then be

$$P_{xx} = \text{diag}\left(\sigma^2_{z_1}, \sigma^2_{z_2-z_1}, \ldots, \sigma^2_{z_m-z_{m-1}}, \sigma^2_{Z_{0,1}}, \sigma^2_{Z_{m,m+1}}, \sigma^2_{N_{01}}, \ldots \right.$$
$$\left., \sigma^2_{N_{0m}}, \sigma^2_{\tilde{z}_1}, \ldots, \sigma^2_{\tilde{z}_m}, \sigma^2_{dVTEC_\phi}, \sigma^2_{dVTEC_\lambda}, \sigma^2_{DCB}, \sigma^2_{\rho_{amb}}\right) \quad (30)$$

where each σ is the associated standard deviation for each state. Then $\sigma_{N_{0m}}$ may be made a few orders of magnitude larger than the other $\boldsymbol{p}$ states.

IONOSPHERIC PARAMETERIZATION KALMAN FILTER RESULTS

This section presents results that have been produced by applying the ionosonde-GPS SREKF algorithm on ionosonde virtual height data and dual-frequency GPS pseudorange data for quiescent conditions. The data were collected on May 13, 2013, at the High Frequency Active Auroral Research Program (HAARP) at Gakona, Alaska. The virtual heights have been extracted from the HAARP digisonde, developed by the University of Massachusetts at Lowell, and fitted with ARTIST-5 software. GPS data is retrieved from Miami University's Novatel OEM4 receiver at the HAARP facility.

Each measurement batch $k = 1, ..., 20$, at $1s$ intervals, contains the frequency-differenced pseudoranges of 8 GPS satellites, as well as a set of ionosonde virtual heights recorded at the start of the first interval. This particular set spans sounding frequencies between $1.475MHz$ to $5.725MHz$, and is fitted to the refractive ray-tracing model outputs $P'_f(\omega_i, \boldsymbol{p}_k)$ at each k. The initial state standard deviations used are given in Table 1.

State Standard Deviation	Value
σ_{z_1}	$1m$
$\sigma_{z_2}, \ldots, \sigma_{z_m}$	$10^{-3}\ln(m)$
$\sigma_{Z_{0,1}}, \sigma_{Z_{m,m+1}}$	$10^{-3}\ln(m^{-1})$
$\sigma_{N_{01}}, \ldots, \sigma_{N_{0m-1}}$	$10\ln(e^-/m^3)$
$\sigma_{N_{0m}}$	$35\ln(e^-/m^3)$
$\sigma_{\tilde{z}_1}, \ldots, \sigma_{\tilde{z}_m}$	$10^{-3}\ln(m^{-1})$
$\sigma_{d\ln VTEC_\phi}, \sigma_{d\ln VTEC_\lambda}$	$1rad^{-1}$
σ_{DCB}	$10^{-9}s$

Table 1: Initial SREKF standard deviations for each state.

Figures 4 and 5 show a typical fit of the data produced by the SREKF and its final parameterization $\boldsymbol{p}$ at the end of the $20s$ filtering run, at the end of which the increments to $\boldsymbol{p}$ are small. The discrepancies between the model and ionosonde data virtual heights are within a few kilometres for most points on Fig. 4, while data-model differences for the GPS slant TEC values are within 0.2 TECU ($10^{16}e^-/m^2$) on Fig. 5. The TEC latitude and longitude variations are $\frac{d\ln VTEC}{d\phi} = 0.5rad^{-1}$ and $\frac{d\ln VTEC}{d\lambda} = 0.2rad^{-1}$. The final estimated profile is presented in Fig. 6, along with the ARTIST-5 automatic profile from the ionosonde. Note the difference in the top-side profile, due to the additional information from the GPS TEC measurements.

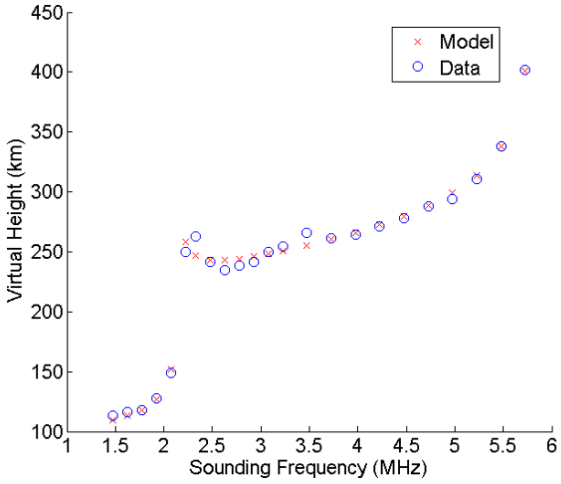

Figure 4: Final Kalman filter fitting of ionosonde virtual heights.

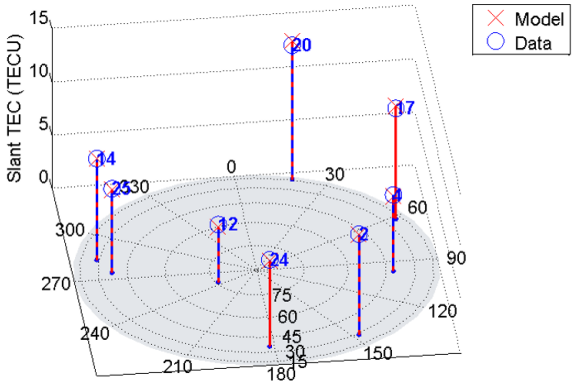

Figure 5: Final Kalman filter fitting of slant TEC for 8 GPS satellites. The satellite locations are shown on an elevation-azimuth polar sky-plot, with PRN numbers in blue.

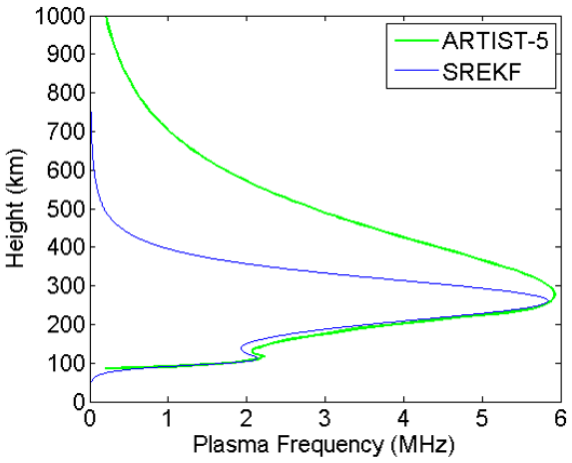

Figure 6: Final Booker profile and ARTIST-5's profile.

Conclusion

Refractive ray-tracing simulations of ionsonde ray paths that implement Hamilton's equations are combined with GPS phase and psuedorange perturbation integrations in a predictive measurement model. The model is used as the measurement model within a fairly standard optimal estimation/Kalman-filtering framework to fuse data from these two sources in order to estimate the parameters that characterize a profile. The particular estimator used is a nonlinear Square Root Information Filter. A test case shows that the filter produces good fits of ionosonde virtual height, to within a few kilometres. Slant total electron content (TEC) for eight satellites have also been fit to within a few tenths of TEC units.

References

[1] Chapman, J.H., and Warren, E.S., "Topside Sounding of the Earth's Ionosphere", *Space Science Reviews*, Vol. 8, No. 5-6, pp. 846-865.

[2] Mitch, R.H., Psiaki, M.L., and Tong, D.M., "Local Ionosphere Model Estimation from Dual-Frequency Global Navigation Satellite System Observables, *Radio Sci.*, Vol. 48, 2013, pp. 671684.

[3] Reinisch, B.W., and Huang, X., "Deducing Topside Profiles and Total Electron Content from Bottomside Ionograms", *Advances in Space Research*, Vol. 27, Issue 1, 2001, pp. 23-30.

[4] Reinisch, B.W., Huang, X., Belehaki, A., Shi, J., Zhang, M., and Ilma, R., "Modeling the IRI Topside Profile Using Scale Heights from Ground-Based Ionosonde Measurements", *Advances in Space Research*, Vol. 34, Issue 9, 2004, pp. 2026-2031.

[5] Stankov, S.M., Jakowski, N., and Heise, S., "A New Method for Reconstruction of the Vertical Electron Density Distribution in the Upper Ionosphere and Plasmasphere", *Journal of Geophysical Research*, Vol. 108, No. A5, 2003, p. 1164.

[6] Stankov, S.M., Stegen, K., Muhtarov, P., Warnant, R., "Local Ionospheric Electron Density Profile Reconstruction in Real Time from Simultaneous Ground-Based GNSS and Ionosonde Measurements", *Advances in Space Research*, Vol. 47, Issue 7, 1 Apr. 2011, pp. 1172-1180.

[7] Jones, R.M., and Stephenson, J.J., "A Versatile Three-Dimensional Ray Tracing Computer Program for Radio Waves in the Ionosphere," *U.S. Dept. of Commerce, OT Report* 75-76, Oct. 1975.

[8] Liu, C.H., and Yeh, K.C., *Theory of Ionospheric Waves*, 1st ed., Academic Press, New York, 1972, pp. 187, 234-240.

[9] Booker, H.G., "Fitting of Multi-Region Ionospheric Profiles of Electron Density by a Single Analytic Function of Height", *J. Atmos. Terr. Phys.*, Vol. 39, 1977, pp. 619624.

[10] Hoque, M.M., and Norbert, J., "Ionospheric Propagation Effects on GNSS Signals and New Correction Approaches", *Global Navigation Satellite Systems - Signal, Theory and Applications*, Prof. Shuanggen Jin, InTech, 2012, pp. 381-404.

[11] Bar-Shalom, Y.X., Li, R., and Kirubarajan, T., *Estimation with Applications to Tracking and Navigation*, John Wiley and Sons, New York, 2001, pp. 371-395.

[12] Bierman, J., *Factorization Methods for Discrete Sequential Estimation*, Dover Publications, New York, 1977, pp. 68-112.

[13] Psiaki, M.L., Theiler, J., Bloch, J., Ryan, S., Dill, R.W., and Warner, R.E., "ALEXIS Spacecraft Attitude Reconstruction with Thermal/Flexible Motions Due to Launch Damage" *Journal of Guidance, Control, and Dynamics)*, Vol. 20, No. 5, Sept.-Oct. 1997, Section II, pp. 1033-1041.

[14] Hatch, R.R., and Knight, J.E., "Method and Apparatus for Smoothing Code Measurements in a Global Positioning System Receiver," *U.S. Patent No. 5,471,217*, Nov. 28, 1995.

[15] Hwang, P.Y., McGraw, G.A., and Bader, J.R., "Enhanced Differential GPS Carrier-Smoothed Code Processing Using Dual-Frequency Measurements," *Navigation*, Vol. 46, No. 2, 1999, pp. 127-137.

[16] "CODE – Center for Orbit Determination in Europe," *University of Bern*, http://www.aiub.unibe.ch/content/research/satellite˙geodesy/code˙˙˙research/index˙eng.html.