Nonlinear Estimation to Assimilate GPS TEC Data into a Regional Ionosphere Model

Mark L. Psiaki
Cornell University, Ithaca, NY 14853-7501, USA

Gary S. Bust
Johns Hopkins University Applied Physics Laboratory, Laurel, MD 20723, USA

and Cathryn N. Mitchell
The University of Bath, BA2 7AY, UK

BIOGRAPHIES

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

Gary S. Bust is a senior scientist in the Geospace and Earth Science Group. He received a Ph.D. in Physics from the University of Texas at Austin. His areas of interest include ionospheric tomographic imaging, data assimilation, and the application of space based observations to ionospheric remote sensing.

Cathryn N. Mitchell is a professor in the Dept. of Electronic and Electrical Engineering. She received a Ph.D. in Physics from The University of Wales Aberystwyth. She researches the development and application of new algorithms in tomography. She also holds an EPSRC advanced fellowship in the effects of the ionized atmosphere on GNSS. She has published over 50 journal papers including invited review papers in the Proceedings of the Royal Society and in American Geophysical Union journals Space Weather and Reviews of Geophysics. Her group has been awarded a number of prizes for their research in both tomography and GPS. Her current academic position is Professor at the University of Bath where she is director of Invert: Centre for Imaging Science.

ABSTRACT

A new method of is being developed to estimate the ionosphere's 3-dimensional electron density distribution based on GPS slant TEC data. The goal of this effort is to develop a generalized parametric ionospheric model that is amenable to data assimilation using powerful nonlinear least-squares batch filtering techniques and related techniques. In addition to assimilating GPS TEC data, this method will eventually be targeted at assimilating additional data types in order to implement true data fusion for ionospheric characterization. The parameterized ionosphere model uses a latitude/longitude bi-quintic spline model to characterize the horizontal variations of parameters of a vertical electron density profile. The result is a truly 3-dimensional electron density distribution. It is parameterized by vertical profile parameter values at latitude/longitude spline nodes and by various latitude and longitude partial derivatives of these parameters at the nodes. This electron density distribution is used in conjunction with quadrature numerical integration to determine slant TEC along line-of-sight paths to tracked GPS satellites. A nonlinear batch estimation algorithm compares the modeled GPS slant TEC values predicted by its current parameter estimates with corresponding measured values. It then updates its parameter estimates to improve its fit to the measurements while balancing a need to use parameters that remain relatively near reasonable *a priori* values, as dictated by an International Reference Ionosphere calculation. A truth-model simulation study shows that the vertical TEC map is observable as part of a latitude/longitude-dependent Chapman profile. The height of peak electron density and the scale height of the Chapman profile are only weakly observable from slant TEC data alone. Tests of this method have also been made with slant TEC data from an array of over 900 dual-frequency GPS receivers distributed over the continental U.S. The method demonstrates an equal or better ability to predict slant TEC at other GPS receivers than that of a traditional thin-shell, fixed-altitude ionosphere data assimilation model like the one used for WAAS.

Preprint from ION GNSS+ 2015

INTRODUCTION

Slant Total Electron Content (TEC) measurements from dual-frequency GPS receivers are regularly used to construct latitude/longitude maps of the ionosphere's Vertical TEC (VTEC) [1] and to reconstruct its full 3-dimensional electron density profile [2]. Existing methods are not so easily adapted to the assimilation of other types of ionospheric data, such as digisonde data, though attempts are being made to do this [3].

Existing methods tend to be based on restrictive assumptions, such as a linear relationship between the model's unknown parameters and slant TEC or the assumption of a thin-shell ionosphere at a known height. The method of Ref. [3] for incorporating ionosonde data is a sort of cascaded method in which the ionosonde provides peak electron density altitudes, but not much more.

The present study is a continuation of efforts to develop a more unified approach to ionospheric data fusion for purposes of estimating the electron density distribution. The first two products in this line of work are Refs. [4] and [5]. The first of these used slant GPS TEC to estimate a local ionosphere 3-dimensional electron density profile. The second effort expanded on that concept by direct fusion of ionosonde and GPS slant TEC data into a local model. The ionosonde data were modeled directly in terms of raw group delay, or virtual height, by use of a ray-tracing calculation.

The present effort is the first of this series that extends the needed smooth electron density profiles to a regional or even a global scale. It develops the needed type of model, and it applies this model to regional ionosphere estimation based on GPS slant TEC data from a network of receivers distributed over the Continental U.S. (CONUS). The extension to handle GPS and ionosonde data fusion using techniques like those of [5] has been left for a future study, but the parameterized electron density model used in the present study is fully capable of being used in such a data fusion application. A long-term goal of this set of efforts is to produce a real-time version of the International Reference Ionosphere (IRI) model or some similar model. It will be based on fused data from global networks of ground-based GPS receivers and ionosondes and on satellite-based radio-occultation slant TEC data.

This paper makes 4 principal contributions to the art of ionospheric electron density estimation. The first contribution is its demonstration of how to combine a latitude/longitude bi-quintic-spline and a vertical profile to produce a regional or global 3-dimensional (3D) electron density profile that can be modified by varying the parameters at its spline nodes. Each bi-quintic spline is a 5^{th} order spline when considered separately in either the latitude or longitude direction. Multiple quantities are modeled by such splines, one spline for each parameter of the model's given vertical electron density profile. An advantage of this type of model is it's smoothness. It is continuous and has continuous 1^{st}- and 2^{nd}-order spatial derivatives. This is the required level of smoothness that will allow it to be used in ray-tracing calculations for the fusion of ionosonde data, as in [5].

This paper's second contribution is a slant TEC model that is based on numerical integration through the splined electron density distribution. This model is nonlinear in the relationship between the unknown, estimated parameters of the spline and its predicted slant TEC.

The third contribution is the definition of a nonlinear estimation problem based on the new splined $N_e(\boldsymbol{r};\boldsymbol{p})$ electron density model, where $\boldsymbol{r}$ is the 3D Cartesian position vector in Earth-Centered, Earth-Fixed (ECEF) coordinates and $\boldsymbol{p}$ is a vector of unknown spline parameters. This estimation problem seeks the $\boldsymbol{p}$ that produces the best fit between the measured and modeled slant TEC values. This paper shows how to incorporate *a priori* information into its batch estimation problem as a means of dealing with the fundamental unobservability of an unknown electron density profile that is a member of an infinite-dimensional function space. That is, $N_e(\boldsymbol{r};\boldsymbol{p})$ is an unknown function of the 3D $\boldsymbol{r}$ vector, and it can never be fully observed based only on a finite set of measurements. This paper also develops a solution algorithm for its nonlinear batch estimation problem. It is based on standard nonlinear least-squares techniques.

The fourth contribution is the processing of real GPS slant TEC data using the new technique. The results are compared to those obtained using a thin-shell, known-altitude model. The comparison considers the two ionospheric data assimilation methods' slant TEC prediction capabilities for stations whose data have not been assimilated.

The remainder of this paper is divided into 7 sections plus a summary and conclusions section. Section II defines a parameterized 3D electron density profile $N_e(\boldsymbol{r};\boldsymbol{p})$ that is based on a latitude/longitude bi-quintic spline of Chapman vertical profile parameters. Section III presents the techniques used for the accurate calculation of slant TEC via numerical integration along GPS ray paths through the $N_e(\boldsymbol{r};\boldsymbol{p})$ profile. It also explains how to compute partial derivative sensitivities of the modeled slant TEC values with respect to the ionosphere model parameters. Section IV develops a method for computing *a priori* bi-quintic spline parameters from the IRI model. Section V defines the batch least-squares estimation problem whose solution will be used to estimate the ionospheric model parameter

vector $\boldsymbol{p}$. It also outlines the solution algorithm. Section VI defines the data set that has been used to develop test cases for the new algorithm. Section VII presents the results of algorithm performance on the test data and on truth-model simulation data. It also compares the performance on real data with that of a thin-shell, known-altitude ionosphere VTEC map estimator. Section VIII discusses potential improvements to this paper's methods. Section IX summarizes this paper's developments and presents its conclusions.

II. BI-QUINTIC-SPLINED GLOBAL CHAPMAN ELECTRON DENSITY PROFILE

A 3D electron density profile of the form $N_e(\boldsymbol{r};\boldsymbol{p})$ can be constructed by combining a vertical profile and a latitude/longitude bi-quintic spline. Suppose that $N_{echap}(h;\boldsymbol{p}_{chap})$ is a Chapman vertical profile in which h is the altitude and the 3-element parameter vector $\boldsymbol{p}_{chap}$ contains the profile's altitude of peak electron density, its scale height, and its VTEC.

This vertical profile, or any vertical profile, can be used to construct a fully 3D electron density distribution if one models its profile parameters as depending on latitude ϕ and longitude λ, i.e., $\boldsymbol{p}_{chap}(\phi,\lambda)$. Suppose that this latitude/longitude "map" of the vertical profile parameters is itself characterized by a vector of parameters $\boldsymbol{p}$ so that its full functional form is $\boldsymbol{p}_{chap}(\phi,\lambda;\boldsymbol{p})$. In this case the full 3D electron density profile becomes

$$N_e(\boldsymbol{r};\boldsymbol{p}) = N_{echap}\{h(\boldsymbol{r});\boldsymbol{p}_{chap}[\phi(\boldsymbol{r}),\lambda(\boldsymbol{r});\boldsymbol{p}]\} \qquad (1)$$

Although a Chapman profile is used in the present study, it would be straightforward to replace $N_{echap}(h;\boldsymbol{p}_{chap})$ with some other vertical profile that had some other parameterization, perhaps one with more elements. This alternate parameterization would also need to be expressed as a latitude/longitude "map", i.e., as a function of ϕ and λ. The functions $\phi(\boldsymbol{r})$, $\lambda(\boldsymbol{r})$, and $h(\boldsymbol{r})$ in Eq. (1) are the standard transformations from Cartesian WGS-84 coordinates to, respectively, latitude, longitude, and altitude as defined relative to the WGS-84 ellipsoid.

The particular form of the Chapman profile parameter map used in the present study takes the form

$$\boldsymbol{p}_{chap}(\phi,\lambda;\boldsymbol{p}) = \exp[\boldsymbol{p}_{logchap}(\phi,\lambda;\boldsymbol{p})] \qquad (2)$$

where the vector function $\boldsymbol{p}_{logchap}(\phi,\lambda;\boldsymbol{p})$ is a latitude/longitude map of the natural logarithms of the 3 Chapman profile parameters. The use of a natural logarithm parameterization and the exponentiation in Eq. (2) combine to preclude the physically impossible situation in which elements of $\boldsymbol{p}_{chap}$ would be non-positive.

The function $\boldsymbol{p}_{logchap}(\phi,\lambda;\boldsymbol{p})$ is modeled using a bi-quintic latitude/longitude spline. Its model takes the form

$$\begin{aligned}\boldsymbol{p}_{logchap}(\phi,\lambda;\boldsymbol{p}) = &\sum_{j=1}^{6}\boldsymbol{p}_{1j}s_{spj}(\phi,\lambda)\\ &+\sum_{i=2}^{M-1}\sum_{j=1}^{9}\boldsymbol{p}_{ij}s_j(\phi,\lambda,\pi_i)\\ &+\sum_{j=1}^{6}\boldsymbol{p}_{Mj}s_{npj}(\phi,\lambda)\end{aligned} \qquad (3)$$

where the $\boldsymbol{p}_{ij}$ vectors are components of the overall bi-quintic spline parameter vector:

$$\boldsymbol{p} = \begin{bmatrix}\boldsymbol{p}_{11}\\ \boldsymbol{p}_{12}\\ \vdots\\ \boldsymbol{p}_{16}\\ \boldsymbol{p}_{21}\\ \boldsymbol{p}_{22}\\ \vdots\\ \boldsymbol{p}_{29}\\ \boldsymbol{p}_{31}\\ \vdots\\ \boldsymbol{p}_{(M-1)9}\\ \boldsymbol{p}_{M1}\\ \boldsymbol{p}_{M2}\\ \vdots\\ \boldsymbol{p}_{M6}\end{bmatrix} \qquad (4)$$

The number of bi-quintic spline nodes is M. Two of them are special nodes at the north and south poles, and the other $M-2$ of them are regular nodes. Each $\boldsymbol{p}_{ij}$ vector has 3 elements because there are 3 parameters in this paper's Chapman vertical profiles. Therefore, the $\boldsymbol{p}$ vector has $3(9M-6)$ elements.

The functions $s_{spj}(\phi,\lambda)$, $s_j(\phi,\lambda,\pi)$, and $s_{npj}(\phi,\lambda)$ are the basis functions of the bi-quintic spline. The 6 $s_{spj}(\phi,\lambda)$ functions are special to the node at the south pole, and the 6 functions $s_{npj}(\phi,\lambda)$ are special to the north pole. The 9 general $s_j(\phi,\lambda,\pi)$ functions apply at all the other nodes, with the parameter vector π defining the latitude and longitude of a general node and the latitude and longitude extent of the influence of the given node's spline functions. Figure 1 shows the 9 general $s_j(\phi,\lambda,\pi)$ functions plotted near a typical node. They have finite support in latitude and longitude, and they are everywhere continuous with continuous first and second latitude and longitude partial derivatives. The special spline functions for the south pole and the north pole have been designed to ensure that the functions and their first and second partial

derivatives are continuous in such a way that the corresponding functions of Cartesian position, $s_{spj}[\phi(\boldsymbol{r}), \lambda(\boldsymbol{r})]$ and $s_{npj}[\phi(\boldsymbol{r}), \lambda(\boldsymbol{r})]$, are everywhere continuous with continuous first and second partial derivatives with respect to $\boldsymbol{r}$. Thus, the function $\boldsymbol{p}_{logchap}[\phi(\boldsymbol{r}), \lambda(\boldsymbol{r}); \boldsymbol{p}]$ exhibits no "belly-button"-type singularities at either of the poles. The 9 regular spline functions also have this property so that $\boldsymbol{p}_{logchap}[\phi(\boldsymbol{r}), \lambda(\boldsymbol{r}); \boldsymbol{p}]$ and its first and second partial derivatives with respect to $\boldsymbol{r}$ are everywhere continuous.

The reason that there are 9 different spline functions for each regular spline node stems from the nature of a bi-quintic spline. A bi-quintic spline of any arbitrary function $a(\phi, \lambda)$ needs to use the a value at each regular node along with the following 8 partial derivatives: $\partial a/\partial\phi$, $\partial^2 a/\partial\phi^2$, $\partial a/\partial\lambda$, $\partial^2 a/\partial\phi\partial\lambda$, $\partial^3 a/\partial\phi^2\partial\lambda$, $\partial^2 a/\partial\lambda^2$, $\partial^3 a/\partial\phi\partial\lambda^2$, and $\partial^4 a/\partial\phi^2\partial\lambda^2$. Each node value of a and the corresponding 8 partial derivatives constitute the spline's coefficients of the 9 basis functions at the given node.

Besides the special nodes at the two poles, the other bi-quintic spline nodes are grouped into different sets that lie on different individual small circles of latitude. On each particular small circle, the choice of the longitudes of the nodes is somewhat arbitrary. An example choice of spline nodes is overlaid on a map of the Earth in Fig. 2. The red dots are the spline nodes. Note how they are arranged along common lines of constant latitude. Their distributions along these lines are more concentrated over CONUS in order to have a finer ability to resolve ionospheric variations in that region for this example spline. There is no particular longitudinal alignment of points between different small circles of latitude. The longitude spacing tends to increase near the poles in order to compensate for the shrinking radii of the corresponding small circles of latitude.

The actual bi-quintic spline calculation of a function $a(\phi, \lambda)$ is based on 7 calculations using 1-dimensional (1D) quintic splines. A 1D quintic spline is fully characterized by its function values and their first and second partial derivatives at the two end nodes of a particular spline interval. Suppose that interval is a longitude interval and the function in question is the scalar $f(\lambda)$. Then its quintic spline formula is

$$\begin{aligned} f(\lambda) = {} & f_a[1 - 10\tau^3(\lambda) + 15\tau^4(\lambda) - 6\tau^5(\lambda)] \\ & + f_b[10\tau^3(\lambda) - 15\tau^4(\lambda) + 6\tau^5(\lambda)] \\ & + \left.\frac{\partial f}{\partial \lambda}\right|_a \Delta\lambda_{ba}[\tau(\lambda) - 6\tau^3(\lambda) + 8\tau^4(\lambda) - 3\tau^5(\lambda)] \\ & + \left.\frac{\partial f}{\partial \lambda}\right|_b \Delta\lambda_{ba}[-4\tau^3(\lambda) + 7\tau^4(\lambda) - 3\tau^5(\lambda)] \\ & + \left.\frac{\partial^2 f}{\partial \lambda^2}\right|_a \Delta\lambda_{ba}^2[0.5\tau^2(\lambda) - 1.5\tau^3(\lambda) \\ & \quad + 1.5\tau^4(\lambda) - 0.5\tau^5(\lambda)] \\ & + \left.\frac{\partial^2 f}{\partial \lambda^2}\right|_b \Delta\lambda_{ba}^2[0.5\tau^3(\lambda) - \tau^4(\lambda) + 0.5\tau^5(\lambda)] \end{aligned} \quad (5)$$

This spline is valid between the two longitude nodes λ_a and λ_b. This spline formula uses the spline interval parameter $\Delta\lambda_{ba} = \lambda_b - \lambda_a$ and the non-dimensional relative position within the interval $\tau(\lambda) = (\lambda - \lambda_a)/\Delta\lambda_{ba}$.

Using the 1D formula in Eq. (5), the bi-quintic spline calculations proceed as follows: Suppose that (ϕ_i, λ_i) and (ϕ_j, λ_j) are the nearest neighboring southwest and southeast bi-quintic spline nodes relative to the point of interest (ϕ, λ). Suppose, also, that (ϕ_k, λ_k) and (ϕ_l, λ_l) are the neighboring northwest and northeast nodes. Then $\phi_i = \phi_j \le \phi \le \phi_k = \phi_l$, $\lambda_i \le \lambda \le \lambda_j$, and $\lambda_k \le \lambda \le \lambda_l$. One uses the values of a_i, $(\partial a/\partial\lambda)_i$, $(\partial^2 a/\partial\lambda^2)_i$, a_j, $(\partial a/\partial\lambda)_j$, and $(\partial^2 a/\partial\lambda^2)_j$ in order to perform 1D spline longitude interpolation along the lower ϕ_i small circle in order to determine $a(\phi_i, \lambda)$. Similar 1D quintic spline calculations are performed in order to determine $\partial a/\partial\phi$ and $\partial^2 a/\partial\phi^2$ at this same point (ϕ_i, λ). Similar quintic spline calculations on the upper ϕ_k small circle are carried out in order to compute $a(\phi_k, \lambda)$ along with $\partial a/\partial\phi$ and $\partial^2 a/\partial\phi^2$ at this same point (ϕ_k, λ). Finally, a 1D quintic spline is calculated between these final two points along the ϕ direction in order to arrive at $a(\phi, \lambda)$.

If one of the neighboring small circles consists of the single node at the south pole or the north pole, then special calculations ensue. Rather than using a longitude quintic spline to compute the values $\partial a/\partial\phi$ and $\partial^2 a/\partial\phi^2$ appropriate to the longitude λ at the pole, a zero-mean once-per-rev sinusoidal function of longitude is used to compute $\partial a/\partial\phi$, and a non-zero-mean twice-per-rev sinusoid is used to compute $\partial^2 a/\partial\phi^2$. The 6 spline parameters that correspond to the 6 nodal functions at the pole consist of the value of a at the pole, the two arbitrary coefficients that determine the $\partial a/\partial\phi$ sinusoid, the two arbitrary coefficients that determine the twice-per-rev part of the $\partial^2 a/\partial\phi^2$ sinusoid, and the arbitrary non-zero offset of the $\partial^2 a/\partial\phi^2$ sinusoid.

This paper's calculations require the first and second partial derivatives with respect to the Cartesian position vector $\boldsymbol{r}$ of the electron density distribution $N_e(\boldsymbol{r}; \boldsymbol{p})$. They also require the first partial derivative with respect to the ionosphere parameter vector $\boldsymbol{p}$ along with the second cross partial derivative with respect to $\boldsymbol{r}$ and $\boldsymbol{p}$. These derivatives can be computed by using the chain rule and taking appropriate derivatives of the sequence of 1D quintic spline calculations that have been described above. The spline nature of $N_e(\boldsymbol{r}; \boldsymbol{p})$ implies that many of the partial derivatives with respect to elements of $\boldsymbol{p}$ will be zero at any given latitude/longitude point. This is true because of the finite support of the spline basis functions, as depicted

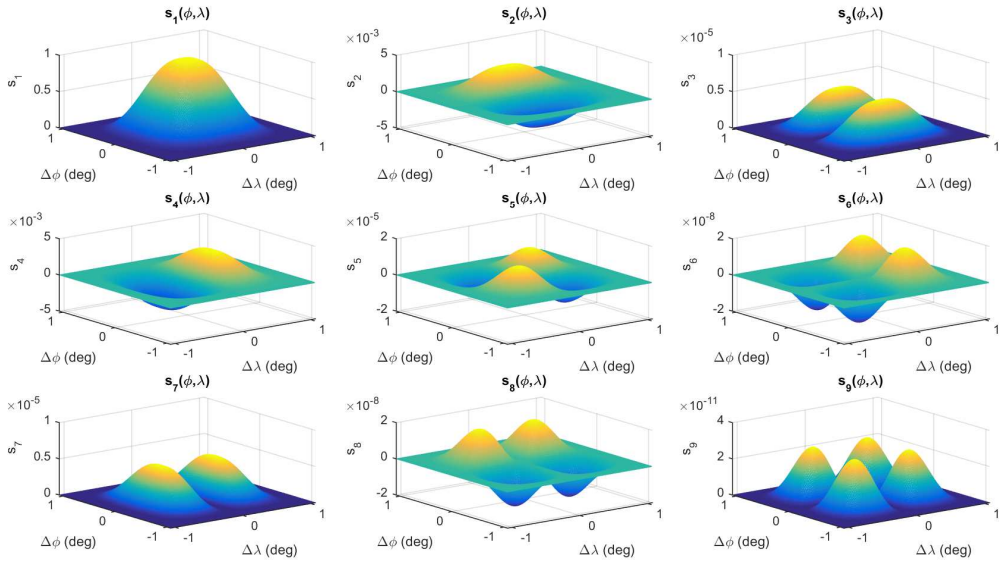

Fig. 1: The nine bi-quintic spline functions of a typical node.

in Fig. 1. Therefore, care should be taken not to waste time calculating elements of any $\boldsymbol{p}$ partial derivatives that are known *a priori* to equal 0.

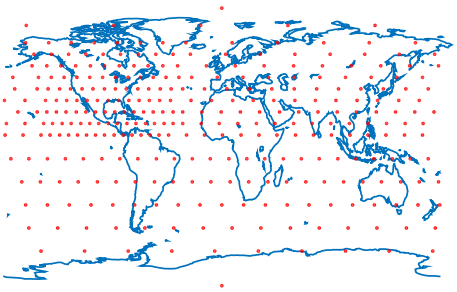

Fig. 2: Example map of possible bi-quintic spline nodes.

III. SLANT TEC MEASUREMENT MODEL

This paper's nonlinear estimation algorithm needs a model by which it can predict each measured slant TEC for a given set of ionosphere parameters. This section explains how to compute that prediction.

A. Numerical Integration to Approximate Slant TEC Model

The slant TEC through the modeled $N_e(\boldsymbol{r};\boldsymbol{p})$ distribution of Section II can be calculated based on knowledge of the GPS receiver location and of the direction vector from that location to the GPS satellite in question. Suppose that the k^{th} slant TEC measurement is made at Cartesian ECEF ground station location $\boldsymbol{r}_{gsk}$, that the ECEF unit direction vector from that ground station to the tracked GPS satellite is $\hat{\boldsymbol{r}}_k$, and that the distance from the ground station to the satellite is ρ_k. If the modeled $N_e(\boldsymbol{r};\boldsymbol{p})$ electron density distribution is correct, then this k^{th} slant TEC measurement should take on the value:

$$\tilde{h}_k(\boldsymbol{p}) = \int_0^{\rho_k} N_e[(\boldsymbol{r}_{gsk}+\rho\hat{\boldsymbol{r}}_k);\boldsymbol{p}]d\rho \qquad (6)$$

This integral can be approximated numerically using quadrature integration. The chosen quadrature formula is based on a cubic spline approximation of the one-dimensional electron density distribution $N_{ek}(\rho;\boldsymbol{p}) = N_e[(\boldsymbol{r}_{gsk}+\rho\hat{\boldsymbol{r}}_k);\boldsymbol{p}]$ and a set of numerical integration grid points along the line-of-sight (LOS) vector $\{\rho_{k0}, ..., \rho_{kL}\}$. Suppose that the electron densities at these grid points are N_{ek0} [$= N_{ek}(\rho_{k0};\boldsymbol{p})$], ..., N_{ekL} [$= N_{ek}(\rho_{kL};\boldsymbol{p})$] and that the corresponding ρ derivatives are N'_{ek0} [$= (dN_{ek}/d\rho)|_{(\rho_{k0};\boldsymbol{p})}$], ... N'_{ekL} [$= (dN_{ek}/d\rho)|_{(\rho_{kL};\boldsymbol{p})}$]. Then the quadrature integration formula used to calculate the slant TEC is

$$\tilde{h}_k(\boldsymbol{p}) \approx \sum_{l=0}^{L-1} \frac{\Delta\rho_{kl}}{2}[N_{ekl} + N_{ek(l+1)} + \frac{\Delta\rho_{kl}}{6}(N'_{ekl} - N'_{ek(l+1)})] \qquad (7)$$

where $\Delta\rho_{kl} = \rho_{k(l+1)} - \rho_{kl}$.

The ranges ρ_{k0}, ..., ρ_{kL} are specially chosen to lie at pre-selected points along the Chapman profile. The bi-quintic spline model of $N_e(\boldsymbol{r};\boldsymbol{p})$ of Section II employs associated spatial functions of the Chapman profile altitude of peak electron density and scale height. Suppose that these functions are defined to be, respectively, $h_{maxN_e}(\boldsymbol{r};\boldsymbol{p})$ and $h_{sh}(\boldsymbol{r};\boldsymbol{p})$. Then it is possible to assign a non-dimensional Chapman altitude to each ECEF location $\boldsymbol{r}$:

$$z(\boldsymbol{r};\boldsymbol{p}) = \frac{h(\boldsymbol{r}) - h_{maxN_e}(\boldsymbol{r};\boldsymbol{p})}{h_{sh}(\boldsymbol{r};\boldsymbol{p})} \tag{8}$$

where $h(\boldsymbol{r})$ is the altitude measured relative to the WGS-84 ellipsoid.

The numerical integration grid points used for Eq. (7) are defined implicitly using a set of pre-specified target non-dimensional Chapman altitudes, z_0, ..., z_L:

$$\begin{aligned} z_l = z[(\boldsymbol{r}_{gsk} + \rho_{kl}\hat{\boldsymbol{r}}_k);\boldsymbol{p}] \\ \text{for } l = 0, \ldots L \end{aligned} \tag{9}$$

It is straightforward to solve each of these implicit equations for its unknown ρ_{kl} value using Newton's method. These iterative calculations can be sped up through careful use of the $\rho_{k(l-1)}$ solution when generating a first Newton guess of ρ_{kl}.

The target non-dimensional Chapman altitudes, z_0, ..., z_L are pre-selected in a way that groups them near the peak of the Chapman electron density profile. A careful arrangement of these points can result in accurate quadrature integration through the varying Chapman profile while economizing on the number of quadrature integration points. The calculations performed in Section VII use just $(L+1)$ = 29 grid points and achieve a relative accuracy better than 1 part in 10,000.

The estimation algorithm of this paper also needs to calculate the Jacobian first partial derivative of its slant TEC model $\tilde{h}_k(\boldsymbol{p})$ with respect to $\boldsymbol{p}$. This is accomplished via analytic partial differentiation of the quadrature integration approximation given in Eq. (7). This differentiation uses the chain rule and accounts for two ways in which the terms in the summation depend on $\boldsymbol{p}$. The first way is through the direct dependence of the function $N_{ek}(\rho_l;\boldsymbol{p})$ on $\boldsymbol{p}$. The second way is through the implicit dependence of the ρ_{kl} node values on $\boldsymbol{p}$. This latter dependence enters through the implicit definition of ρ_{kl} in Eq. (9). Differentiation of this equation with respect to $\boldsymbol{p}$ yields a linear equation for the partial derivative $\partial\rho_{kl}/\partial\boldsymbol{p}$, which can easily be solved to determine this derivative. The full formula of the $\partial\tilde{h}_k/\partial\boldsymbol{p}$ calculation is straightforward to derive. It has been omitted for the sake of brevity.

B. Full Measurement Model with Receiver Biases

The full model for the slant TEC measurement takes the form

$$\begin{aligned} y_k &= \tilde{h}_k(\boldsymbol{p}) + b_{q(k)} + \nu_k \\ &= h_k(\boldsymbol{p},\boldsymbol{b}) + \nu_k \end{aligned} \tag{10}$$

where $q(k)$ is an indexing function that associates the $q(k)^{\text{th}}$ receiver with the k^{th} measurement and where b_q is the inter-channel TEC bias of the q^{th} receiver. The vector $\boldsymbol{b}$ = $[b_1;\ldots;b_Q]$ is the vector that contains the unknown biases of all Q receivers whose data are being used in the ionosphere estimation problem. The quantity ν_k is the slant TEC measurement noise. It is assumed to be a sample from a zero-mean Gaussian distribution with standard deviation $\sigma_{\nu k}$.

Note that multiple k values will typically map to identical $q(k)$ values due to the fact of individual receivers having multiple channels and, therefore, measuring slant TEC to multiple satellites. If any given set of k values produces identical $q(k)$ values, then all of the corresponding $\boldsymbol{r}_{gsk}$ receiver locations in Eqs. (6) and (9) will be identical.

If the entire estimation problem includes a total of K measurements, then these measurements and their corresponding models and errors can be lumped into K-dimensional vectors:

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_K \end{bmatrix}, \quad \boldsymbol{h}(\boldsymbol{p},\boldsymbol{b}) = \begin{bmatrix} h_1(\boldsymbol{p},\boldsymbol{b}) \\ h_2(\boldsymbol{p},\boldsymbol{b}) \\ h_3(\boldsymbol{p},\boldsymbol{b}) \\ \vdots \\ h_K(\boldsymbol{p},\boldsymbol{b}) \end{bmatrix}, \quad \boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \vdots \\ \nu_K \end{bmatrix} \tag{11}$$

The final vector measurement model takes the form

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{p},\boldsymbol{b}) + \boldsymbol{\nu} \tag{12}$$

Its measurement error vector $\boldsymbol{\nu}$ is assumed to be a zero-mean, Gaussian random vector with covariance matrix

$$\begin{aligned} E\{\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}\} &= R \\ &= \begin{bmatrix} \sigma_{\nu 1}^2 & 0 & 0 & \ldots & 0 \\ 0 & \sigma_{\nu 2}^2 & 0 & \ldots & 0 \\ 0 & 0 & \sigma_{\nu 3}^2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \sigma_{\nu K}^2 \end{bmatrix} \end{aligned} \tag{13}$$

IV. A PRIORI PARAMETER ESTIMATES FROM IRI MODEL

Typically the ionosphere model parameter vector $\boldsymbol{p}$ has many elements, perhaps thousands to tens of thousands or more. Such large numbers are necessary in order for

it to allow a sufficiently general parameterization of the $N_e(\boldsymbol{r};\boldsymbol{p})$ electron density distribution, one that has the potential to form a reasonable approximation of the true distribution.

This large number of unknown parameters poses a challenge for any estimation algorithm, the challenge of observability. The slant TEC measurement model in Eq. (12) is said to be observable if there is a unique combination of the vector pair $(\boldsymbol{p},\boldsymbol{b})$ that minimizes the norm squared of this equation's measurement error vector $\boldsymbol{\nu}$. If the minimum is not unique, then there is no practical way to prefer one minimizing $(\boldsymbol{p},\boldsymbol{b})$ combination to another. Given that one combination is likely nearest to the true combination while others are likely far away, a lack of observability makes the measurement model useless for purposes of forming an accurate estimate of $(\boldsymbol{p},\boldsymbol{b})$. As the number of unknowns grows, the challenge of achieving observability grows.

Even if a system is technically observable, its unique optimal estimates of the ionosphere parameter vector $\boldsymbol{p}$ and the receiver bias vector $\boldsymbol{b}$ might be highly inaccurate, thereby making the estimates useless for all practical applications. It is a well known fact of estimation theory that an increase in the number of estimated unknowns will degrade the accuracy of the estimates of all the pre-existing unknowns if all else is equal. Therefore, having a large number of unknown ionosphere parameters in $\boldsymbol{p}$ poses a challenge to the goal of estimating an accurate $N_e(\boldsymbol{r};\boldsymbol{p})$ distribution.

All of the preceding statements assume that $\boldsymbol{p}$ and $\boldsymbol{b}$ must be estimated from scratch based purely in the measured slant TEC data in $\boldsymbol{y}$. If additional information can be gleaned from another source, then the problem of obtaining accurate estimates of $\boldsymbol{p}$ and $\boldsymbol{b}$ becomes less challenging.

Therefore, a method has been developed to obtain a reasonable *a priori* estimate of the ionosphere parameter vector $\boldsymbol{p}$. This estimate is generated using an ionosphere model. The idea is to find a value of $\boldsymbol{p}$ that fits $N_e(\boldsymbol{r};\boldsymbol{p})$ to a modeled $N_e(\boldsymbol{r})$ distribution.

The IRI [6] is the model that has been used in this study to develop an *a priori* $\boldsymbol{p}$ estimate. Note, however, that there is nothing sacred about using the IRI model. One could use any other reasonable model, e.g., the SAMI2 model [7].

A. Calculation of A Priori Ionospheric Parameter Vector

A sequence of operations is used to develop an *a priori* estimate of $\boldsymbol{p}$ from the IRI model. The first operation is to determine the 3 Chapman profile parameters at each of the bi-quintic spline nodes. This involves calculation of the IRI N_e vertical profile at each node followed by calculation of the optimal nonlinear least-squares fit of the 3 Chapman parameters to the IRI profile. The natural logarithms of these fit parameters are used as the elements of the *a priori* $\boldsymbol{p}$ that correspond to bi-quintic spline function values at the spline nodes.

Next, a smoothing process is used to generate reasonable values for the 8 required partial derivatives of each splined quantity at each node (only 5 at the north and south poles). The initial partial derivatives to be calculated are the first and second partial derivatives of each splined quantity with respect to longitude λ. These values are chosen in order to minimize the longitude integral around each small circle of latitude of the square of the third partial derivative with respect to longitude of the periodic quintic-splined quantity. This minimization tends to produce the smoothest possible longitude variations that fit the function values at all of the nodes on the given small circle. This process is repeated for each independent small circle of constant latitude for all the regular spline nodes until all of the required first and second λ partial derivatives have been computed.

A similar process is then applied to each independent great circle of constant longitude (modulo 180$^{\text{o}}$) that passes through one or more spline nodes. The value of a particular splined function is computed at all intersection points between this great circle and the small circles of constant latitude on which all of the spline nodes lie. Values of the first and second ϕ partial derivatives of this function are then chosen in order to minimize the latitudinal integral around the great circle of longitude of the square of the third partial derivative with respect to latitude of the periodic quintic-splined quantity. Again, this minimization tends to produce the smoothest possible latitude variations. The computed first and second ϕ partial derivatives from this procedure are retained for all of the great-circle/small-circle intersection points that correspond to actual bi-quintic spline nodes. This process is repeated for each unique great circle until all of the first and second ϕ partial derivatives have been computed at all of the regular spline points.

This process is also used to compute *a priori* estimates of the periodic ϕ partial derivative parameters that apply at the north and south poles. The first and second ϕ partial derivatives are computed as functions of the great-circle λ samples at the two poles during this procedure. These, in turn, are fit to appropriate sinusoidal models in λ in order to produce the desired elements of the *a priori* $\boldsymbol{p}$ estimate.

The cross partial derivatives with respect to ϕ and λ at

the regular spline nodes are computed in a manner similar to that used to compute the ϕ partial derivatives. Great circles are used, and partial derivatives with respect to ϕ are computed at intersection points between a given great circle and the set of constant-latitude small circles. The difference here is that the splined quantities whose ϕ partial derivatives are to be estimated are quantities that are themselves first or second partial derivatives with respect to λ. Such quantities are computable at the great-circle/small-circle intersection points based on the original small-circle periodic splines with respect to λ.

The resulting estimates are a set of Chapman profile parameter natural logarithms and various of their latitude and longitude partial derivatives at the spline nodes. When lumped into the form of the $\boldsymbol{p}$ ionosphere parameter vector, the result constitutes the *a priori* parameter estimate. Let this estimate be called $\bar{\boldsymbol{p}}$.

The efficacy of the smoothing procedure has been tested by comparing the bi-quintic spline's estimated Chapman parameters with those based on direct fit to the IRI profile. The comparisons have been made at points between the bi-quintic spline nodes. When the spline grid spacing is about 10^{o} in latitude and an equivalent geographic spacing in longitude, the two Chapman parameter maps agree well. Therefore, the smoothing method of fitting the latitude and longitude spline partial derivatives seems to work well.

Another option is to use a thin-shell, constant-altitude *a priori* model of the ionosphere. This option is provided for purposes of generating a comparison ionosphere model fit in Section VII. The generation of the $\bar{\boldsymbol{p}}$ vector in this case is the same for all elements that parameterize the Chapman profile's VTEC natural logarithm. That is, this thin-shell model continues to use the IRI VTEC map to generate its *a priori* VTEC model. The differences lie in the models of peak electron density height and scale height. The computed values from the IRI profile are discarded in favor of known constants that are independent of latitude and longitude. The height of the peak electron density is set to the constant value 350 km. The Chapman scale height is set to 1 km, which corresponds to a very thin shell.

B. Modeled Uncertainty of the A Priori Ionospheric Parameter Vector

The batch filter also needs a covariance matrix that characterizes the *a priori* uncertainty in the estimate $\bar{\boldsymbol{p}}$. A simplified diagonal covariance matrix is used. It starts with assumed *a priori* standard deviations for the Chapman altitude of peak electron density, the Chapman scale height, and the Chapman VTEC. Let these quantities be, respectively, σ_{hmaxNe}, σ_{hsh}, and σ_{VTEC}. These standard deviations are assumed to be independent of latitude and longitude.

At any given node point, the *a priori* standard deviations of the natural logarithms of h_{maxNe}, h_{sh}, and $VTEC$ are computed using a finite-difference-based linearization of the natural logarithm function. The formulas employed are

$$\begin{aligned}
\sigma_{lnhmax} &= \frac{\ln(h_{maxNe} + \gamma\sigma_{hmaxNe}) - \ln(h_{maxNe})}{\gamma} \\
\sigma_{lnhsh} &= \frac{\ln(h_{sh} + \gamma\sigma_{hsh}) - \ln(h_{sh})}{\gamma} \\
\sigma_{lnVTEC} &= \frac{\ln(VTEC + \gamma\sigma_{VTEC}) - \ln(VTEC)}{\gamma}
\end{aligned} \tag{14}$$

where γ is a tuning parameter of the one-sided finite difference calculation. A typical value used in this study is γ = 1.3. The standard deviations given in Eq. (14) are applied directly to the elements of $\bar{\boldsymbol{p}}$ that correspond to node values of the Chapman profile parameter natural logarithms.

Calculation of *a priori* standard deviations for the ϕ and λ partial derivative elements of $\bar{\boldsymbol{p}}$ is somewhat more complicated. For each spline node, the calculation starts by computing the maximum separation of that node's latitude from its two nearest-neighbor small circles of latitude that contain additional nodes. Let this maximum be designated $\Delta\phi$. Similarly, the maximum longitude separation is calculated between the given node and its two nearest neighbors on its own constant latitude small circle. Let this maximum be designated $\Delta\lambda$. These latitude and longitude increments are then used to divide the standard deviations given in Eq. (14) in order to synthesize *a priori* standard deviations for the corresponding partial derivatives. For example, consider the *a priori* standard deviation for the first partial derivative with respect to longitude of the Chapman scale height natural logarithm. It is set to the value $\sigma_{lnhsh}/\Delta\lambda$. Similarly, the value $\sigma_{lnVTEC}/(\Delta\phi^2\Delta\lambda)$ is the chosen *a priori* standard deviation for the third partial derivative of the $VTEC$ natural logarithm twice with respect to ϕ and once with respect to λ. The remaining *a priori* parameter standard deviations are calculated in a similar manner.

The various *a priori* standard deviations for the $\bar{\boldsymbol{p}}$ vector are then assembled into a diagonal *a priori* covariance matrix. It takes the form:

$$E\{(\boldsymbol{p}-\bar{\boldsymbol{p}})(\boldsymbol{p}-\bar{\boldsymbol{p}})^{\text{T}}\} = P_{\bar{p}\bar{p}}$$

$$= \begin{bmatrix} \sigma_{p11}^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_{p12}^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_{p13}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_{pM6}^2 \end{bmatrix} \quad (15)$$

where σ_{pij} is the *a priori* standard deviation of element $\bar{p}_{ij}$, with element indices as defined in Eq. (4).

The chosen values of σ_{hmaxNe}, σ_{hsh}, and σ_{VTEC} are tuning parameters of the batch estimation algorithm. Small values cause the estimator to trust the corresponding elements of $\bar{\boldsymbol{p}}$ and not to change them much in the calculation of its optimal fit. Large values cause the estimator to trust the corresponding entries of $\bar{\boldsymbol{p}}$ less, which opens up the possibility of making larger adjustments to these values during the estimation calculations.

The comparison constant-height, thin-shell ionosphere model has been implemented using the same batch filter that is used for this paper's general model. The only differences lie in the choice of $\bar{\boldsymbol{p}}$ entries corresponding to the peak-electron-density height map and the scale-height map and in the choices of σ_{hmaxNe} and σ_{hsh}. As discussed previously, the modified elements of $\bar{\boldsymbol{p}}$ dictate a constant peak electron density altitude of 350 km and a constant scale height of 1 km at all latitudes and longitudes. In order to maintain these constant values during the batch estimation calculations, the uncertainty standard deviations in σ_{hmaxNe} and σ_{hsh} are set to very small values. This special tuning prevents the batch estimator from making any appreciable adjustments to the altitude or scale height during the solution of its estimation problem. The only permitted adjustments are to its $VTEC$ map.

V. BATCH LEAST-SQUARES ESTIMATION PROBLEM AND SOLUTION ALGORITHM

A. Least-Squares Estimation Problem

The batch filter computes its optimal estimates of the ionosphere parameter vector $\boldsymbol{p}$ and the receiver biases vector $\boldsymbol{b}$ by finding the values of these vectors that minimize the following negative-log-probability density cost function

$$\begin{aligned} J(\boldsymbol{p},\boldsymbol{b}) = {} & \frac{1}{2}[\boldsymbol{y} - \boldsymbol{h}(\boldsymbol{p},\boldsymbol{b})]^{\mathrm{T}} R^{-1} [\boldsymbol{y} - \boldsymbol{h}(\boldsymbol{p},\boldsymbol{b})] \\ & + \frac{1}{2}(\bar{\boldsymbol{p}} - \boldsymbol{p})^{\mathrm{T}} P_{\bar{p}\bar{p}}^{-1} (\bar{\boldsymbol{p}} - \boldsymbol{p}) \\ & + \frac{1}{2\sigma_b^2} \boldsymbol{b}^{\mathrm{T}} \boldsymbol{b} \end{aligned} \quad (16)$$

The third cost term penalizes non-zero receiver bias estimates under the assumptions that their *a priori* values are 0 and that their *a priori* uncertainties have a standard deviation of σ_b.

The minimization of the cost function in Eq. (16) is equivalent to solving for the least-squares solution of the following over-determined system of nonlinear equations

$$\begin{bmatrix} R^{-1/2}\boldsymbol{y} \\ P_{\bar{p}\bar{p}}^{-1/2}\bar{\boldsymbol{p}} \\ 0 \end{bmatrix} = \begin{bmatrix} R^{-1/2}\boldsymbol{h}(\boldsymbol{p},\boldsymbol{b}) \\ P_{\bar{p}\bar{p}}^{-1/2}\boldsymbol{p} \\ \frac{1}{\sigma_b}\boldsymbol{b} \end{bmatrix} + \boldsymbol{\nu}_{tot} \quad (17)$$

where $R^{-1/2}$ and $P_{\bar{p}\bar{p}}^{-1/2}$ are inverses of matrix square roots of, respectively, R and $P_{\bar{p}\bar{p}}$. The requisite matrix square roots can be computed using Cholesky-factorization. The composite equation error vector $\boldsymbol{\nu}_{tot}$ is assumed to be a random sample from a zero-mean, identity-covariance Gaussian distribution.

B. Gauss-Newton Solution Algorithm

The Gauss-Newton method [8] is used to solve the optimal estimation problem that minimizes the nonlinear weighted least-squares cost function in Eq. (16). It is an iterative gradient-based optimization algorithm.

Each iteration of the Gauss-Newton method starts with a guess of the optimal $\boldsymbol{p}$ and $\boldsymbol{b}$ vectors, and it generates improved guesses until no more improvements are possible. The improved guesses are generated after linearizing the over-determined system of equations in Eq. (17) about the current guesses of $\boldsymbol{p}$ and $\boldsymbol{b}$. This over-determined linear system of equations is solved to determine candidate better guesses of the optimal $\boldsymbol{p}$ and $\boldsymbol{b}$. A line search is then performed along the segment from the existing guess in ($\boldsymbol{p}$,$\boldsymbol{b}$) space to the candidate new guess. The search stops short of the candidate new guess if a reduction of the increment is necessary in order to ensure a decrease of the cost function in Eq. (16). This line search procedure guarantees convergence to a local minimum of this cost function.

Some *ad hoc* features have been added to the basic Gauss-Newton method. These features tend to help achieve eventual convergence to the optimal solution. One feature is to optimize only $\boldsymbol{b}$ on the first Gauss-Newton iteration. The biases enter the system of equations in Eq. (17) linearly. Therefore, they can be optimized exactly using some simple linear algebra calculations.

A second *ad hoc* feature restricts the length of the step from the existing guess of ($\boldsymbol{p}$,$\boldsymbol{b}$) to the candidate new guess in such a way that the resulting change in $\boldsymbol{p}$ has a norm which is no larger than the vector norm of the existing $\boldsymbol{p}$ guess. Without this technique, initial $\boldsymbol{p}$ increments are sometimes unreasonably large, and they cause internal calculations to produce ridiculous results that crash the software.

The Gauss-Newton method's convergence is insensitive to the initial guess of $\boldsymbol{b}$ by virtue of the initial optimization of only that unknown. Convergence can be sensitive, however, to the initial guess of $\boldsymbol{p}$. Typically one uses the *a priori* value from the IRI model, $\bar{\boldsymbol{p}}$. If one has a better first guess, perhaps from a previous solution of the problem that uses slightly different tuning parameters or a slightly different set of slant TEC measurements, then one can initialize the algorithm with that better $\boldsymbol{p}$ guess.

C. Cramer-Rao Estimation Error Covariance Lower Bound

The cost function $J(\boldsymbol{p}, \boldsymbol{b})$ has been designed to equal the negative natural logarithm of the Bayesian probability density of $(\boldsymbol{p}, \boldsymbol{b})$ conditioned on the data in $\boldsymbol{y}$ under the diffuse prior assumption of Bayes' postulate. The inverse of the second partial derivative of this function, when evaluated at the truth values of $\boldsymbol{p}$ and $\boldsymbol{b}$, yields the Cramer-Rao lower bound for the error covariance of any estimator. This covariance lower bound is:

$$\begin{bmatrix} P_{pp} & P_{pb} \\ P_{pb}^{\mathrm{T}} & P_{bb} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 J}{\partial \boldsymbol{p}^2} & \frac{\partial^2 J}{\partial \boldsymbol{p} \partial \boldsymbol{b}} \\ \left(\frac{\partial^2 J}{\partial \boldsymbol{p} \partial \boldsymbol{b}}\right)^{\mathrm{T}} & \frac{\partial^2 J}{\partial \boldsymbol{b}^2} \end{bmatrix}^{-1} \tag{18}$$

VI. TEST DATA SET

This paper's new ionosphere estimation algorithm has been tested by applying it to data from a network of almost 1000 CONUS stations. Slant TEC data from this network have been collected for 17 March 2015. Each station returns slant TEC data from about 7 to 8 GPS satellites on average.

The distribution of CONUS ground stations for an example case is shown in Fig. 3. The red dots show the locations of the receiver ground stations whose slant TEC have been used to estimate the ionosphere parameter vector $\boldsymbol{p}$ and the corresponding receiver bias vector $\boldsymbol{b}$. Although this set only applies to one case, most of these ground stations are the same for all cases considered in this paper. The ground stations are concentrated on the west cost, in the northeast, and in the northern part of the midwest. Therefore, one expects better model performance in these regions.

The 6 green diamonds in Fig. 3 give the locations of 6 receivers from which slant TEC data are available but are not used in the batch estimation problem. These 6 stations are used to evaluate the slant TEC prediction accuracy of the $N_e(\boldsymbol{r}, \boldsymbol{p})$ electron density distribution when using an optimal estimate of the $\boldsymbol{p}$ ionosphere parameter vector.

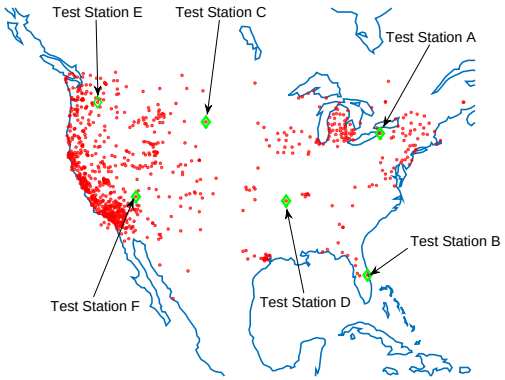

Fig. 3: Map of ground stations used to estimate the electron density profile (red dots) and used to evaluate slant TEC prediction capability of estimate (green diamonds).

These stations' TEC values are compared with those predicted for these stations using the slant TEC measurement model in Eq. (7). After removal of any apparent common-mode bias for the given station, the maximum and root-mean-square (RMS) errors are computed for the given test station. These prediction accuracy metrics are compared among 3 potential ionosphere estimates, the optimal estimate that uses this paper's new method, the optimal thin-shell, known-altitude estimate, and the *a priori* estimate based on the IRI model. Considering their proximity to stations used by the batch estimator, one would expect the best performance for test stations A, E, and F. Worse performance might be expected for stations B, C, and D.

Note that the removal of biases from the 6 test station receivers represents a sort of "cheating". The biases are calculated as being the mean errors between the predicted slant TEC values at the stations and their measured values. This need to rely on measured values means that the de-biased values are not wholly predicted values. Despite this "cheating" such an analysis provides a reasonable measure of whether the slant TEC predictions would be useful for aiding single-frequency GPS navigation. If aiding the navigation of a single-frequency receiver, any bias that was actually part of the model rather than the data would cause an error in the corrected pseudoranges of the aided receiver. Fortunately, this error would be common-mode, which would put it entirely in the computed clock correction rather than in the position estimate.

Data from two different times have been analyzed for 17 March 2015, one from UTC 07:00:00 and the other from UTC 20:00:00. The first corresponds to a local time of 01:00:00 roughly over the middle of CONUS. The second

time corresponds to a local time of 14:00:00. Thus, the first case occurs in the middle of the night, at a time of low expected VTEC. The second case corresponds to the early afternoon, at a time near the expected peak VTEC. Both of these times lie within the 24 period that corresponds to the 2015 St. Patrick's day geomagnetic storm. So, the new algorithm has been presented with interesting estimation cases.

The local nighttime case involves data from Q = 935 independent GPS receivers for a total of K = 7869 slant TEC measurements. The local daytime case has slightly more receivers but fewer measurements: numbers: Q = 937 receivers and K = 6431 TEC measurements.

The *a priori* estimate $\bar{\boldsymbol{p}}$ used by the filter for each case is based on the IRI model for exactly one year earlier than the time of the measurements. This one year offset relieves the need for the IRI to be used in a predictive mode should the present algorithm be implemented in a real-time application. This use of a 1-year offset is reasonable because the IRI model tends to exhibit a strong correlation between its $N_e(\boldsymbol{r})$ distributions when they are separated in time by exactly 1 year.

The various tuning parameters that have been used in the filter are as follows: σ_{nuk} = 1 TECU (10^{16} electrons/m^2), for all k = 1, ..., K; σ_{hmax} = 2 km for the new batch filter under regular operation but σ_{hmax} = 0.0001 km when implementing the thin-shell, known-altitude estimator; σ_{sh} = 1 km for the new batch filter under regular operation but σ_{hmax} = 0.00001 km for the thin-shell, known-altitude estimator; σ_{VTEC} = 5 TECU (5x10^{16} electrons/m^2) for most cases, but it is sometimes lowered to 0.5 TECU or 1 TECU.

The two different nominal bi-quintic spline node spacings have been tried for the batch filter when operating on these data. The primary case uses a nominal spacing of 15^o in latitude and roughly the equivalent geodetic distance between longitude grid points on any given small circle of latitude. A different version of the filter cuts these nominal spacings in half over CONUS, but leaves them nearly the same over the remainder of the globe. All of the data are taken over CONUS, and this refined grid spacing causes roughly a tripling of the number of unknown ionosphere parameters that are actively estimated by the batch filter.

An additional test case has been run using data from a truth-model simulation. The simulation used a $\boldsymbol{p}_{truth}$ ionosphere parameter vector and a non-zero $\boldsymbol{b}_{truth}$ receiver bias vector to generate simulated data. Random noise was added to the simulated data in order to approximate the effects of sensor noise. The $\boldsymbol{p}_{truth}$ vector and the filter's *a priori* $\bar{\boldsymbol{p}}$ vector were deliberately set to significantly different values in order to avoid a sort of "cheating" that would aid the filter's accuracy in a way which would be unrealistic for data from a real receiver network. The method of synthesizing realistic differences was to generate $\boldsymbol{p}_{truth}$ using the IRI model for 23 Dec. 2012 and to generate $\bar{\boldsymbol{p}}$ from the 23 June 2010 IRI model. This 6 month discrepancy in the time of year ensured a significant difference in the two parameter vectors. The goal of the truth-model simulation has been to test the concept in a controlled setting where the "truth" value of $\boldsymbol{p}$ is known so that it can be compared directly to the optimal estimate. If this is a good technique, then the resulting comparison should yield reasonable agreement despite the significant differences between $\boldsymbol{p}_{truth}$ and the batch filter's *a priori* estimate in $\bar{\boldsymbol{p}}$.

The test case used simulated data from Q = 237 GPS receivers at CORS sites that were distributed fairly evenly over the CONUS. The number of slant TEC measurements was K = 1937. The simulated slant TEC measurement accuracy was σ_{nuk} = 0.2 TECU.

The batch filter for the truth-model data used a bi-quintic spline grid with a nominal spacing of 10^o in latitude and roughly the equivalent geodetic distance between longitude grid points on any given small circle of latitude.

VII. ALGORITHM PERFORMANCE

The new ionosphere estimation algorithm has been tested in various ways. Its performance is good, especially when considered relative to the alternatives. The details of how it performed are discussed in the present section.

A. Performance on Truth-Model Simulation Data

The initial test of the new method using data from the truth-model simulation provided confirmation that this method has merit. The differences between the truth and estimated VTEC values over most of the CONUS were 0.5 TECU or less. This is very good and is consistent with computed covariances that have been generated using the Cramer-Rao formula in Eq. (18).

The truth-model simulation demonstrated a degraded ability of the batch filter to estimate the truth Chapman profile's altitude of peak electron density and scale height. Errors in the peak density altitude ranged from 45 to 90 km in the region of high density of ionosphere pierce points of the measurements. The batch filter's Cramer-Rao covariance calculations indicated that these errors should have been smaller, with σ ranging from about 7 to 40 km. Similarly, the scale height estimates were not great. Their errors ranged from about -24 km to 60 km over CONUS. The corresponding Cramer-Rao limits of the scale height

standard deviations in this region ranged from about 9 to 27 km. This is a more reasonable level of consistency. These latter two results indicate that the altitude of peak electron density and the scale height are only weakly observable based solely on slant TEC data from a network of ground-based dual-frequency GPS receivers.

B. Performance on Real Data

The nonlinear least-squares batch estimation algorithm worked well. It tended to reach the minimum of the cost function in Eq. (16) in about 5 major Gauss-Newton iterations after the first iteration that updated only $\boldsymbol{b}$. Each iteration took between 20 minutes and an hour when performing the computations in MATLAB on a 3 GHz laptop computer.

The statistics that characterize the slant TEC prediction capabilities and fit errors are summarized in Table I. These results apply to two cases. One case, called the Local Night case, corresponds to UTC 07:00:00, and results for that case are presented in the 3 left-most columns of the table. The other case is for local day time, for UTC 20:00:00, as previously mentioned. The 3 right-most columns of the table apply to this case. Each entry of the table contains RMS slant TEC errors followed by maximum slant TEC errors after the slash, "/". These error statistics are all given in TECU. The first 6 rows of the table record the slant TEC predictions at the 6 test station locations that are denoted by green diamonds in Fig. 3. The 7$^{\text{th}}$ line aggregates the maximum and RMS errors over all GPS slant TEC measurements for these 6 stations. The 8$^{\text{th}}$ and final line of the table gives the RMS and peak slant TEC errors for the data from the other stations, the ones that have been used by the batch filter to form its ionosphere parameter estimates.

The performance of the new ionospheric estimation filter can be evaluated by comparing its results in the 1$^{\text{st}}$ and 4$^{\text{th}}$ columns of Table I with those of the thin-shell, known-altitude model (2$^{\text{nd}}$ and 5$^{\text{th}}$ columns) and with those of the *a priori* IRI-based model (3$^{\text{rd}}$ and 6$^{\text{th}}$ columns). The new method obviously makes great improvements over the *a priori* IRI-based model: The RMS and maximum errors in the 1$^{\text{st}}$ column are obviously much smaller than those in the 3$^{\text{rd}}$ for the local night case, and the same holds true for the local day comparison between the 4$^{\text{th}}$ and 6$^{\text{th}}$ columns. The improvement of the new method over the thin-shell/known-altitude case is much less dramatic. There is clearly a moderate improvement for the local day case, as one can see by comparing the 4$^{\text{th}}$ and 5$^{\text{th}}$ columns. The improvement for the local night case, however, is marginal – compare the 1$^{\text{st}}$ and 2$^{\text{nd}}$ columns.

Note that all cases in Table I use the *a priori* model VTEC standard deviation σ_{VTEC} = 5 TECU. They also all use a bi-quintic spline grid spacing over CONUS of 15$^{\text{o}}$. This spacing gives roughly the same possible variations as would a 5$^{\text{o}}$ spacing for a simple bi-linear latitude/longitude interpolator. This is true because a bi-linear model has only one independent parameter per node, but the bi-quintic spline model has 9 per node. The spline's nominal grid spacing divided by the square root of the number of parameters per node gives the equivalent lat/long grid spacing for an interpolant with a single parameter per node.

More than 900 of the same receivers apply to the local day and local night cases. These receivers' biases can be compared to check whether their estimates display day/night stability. Bias stability has been studied by a number of authors, and one would expect these receivers to have such stability, perhaps varying a few TECU from day to night due to temperature variations. In fact, variations as large as 21 TECU have been observed. This is less than 10% of the range of biases seen in the 900+ receivers, but this variability is deemed to be larger than is reasonable. What is more, the estimated biases are sensitive to the filter *a priori* VTEC uncertainty tuning parameter σ_{VTEC}. An increase of this value from 0.5 TECU to 5 TECU changes the bias estimates by anywhere from 10 to 34 TECU for a daytime case. Clearly something systematic is occurring, and it implies that the estimated absolute VTEC values are suspect.

One constructive way to analyze the properties of the new ionosphere estimation method is to consider the increments that it estimates to its *a priori* electron density distribution. Such increments are illustrated in Figs. 4 and 5. Figure 4 plots the increment to the Chapman profile's altitude of peak electron density relative to its *a priori* latitude/longitude variations. Figure 5 does the same for the profile's VTEC increment. Both plots also show the ionosphere pierce points of the 6431 ray-paths that are associated with the measurements which have been used by the estimator. In both cases, the increments die off to zero as one moves far away from these pierce points because there is no effect of the distant electron density distribution on the modeled measurements and because the batch filter is designed to revert to its *a priori* model in this situation. In the region of the pierce points, however, the batch filter estimates significant perturbations from the *a priori* Chapman profile parameters.

The most unusual features of these two figures are the large humps in the VTEC increment plot shown in Fig. 5. The largest of these occur in regions that border the region of pierce points. These perturbations are very large, and their distance from the pierce point concentrations makes

TABLE I: Summary of RMS/Maximum Slant TEC Errors (given in TECU)

Case	Local Night New filter	Local Night Thin-shell	Local Night A priori IRI	Local Day New filter	Local Day Thin-shell	Local Day A priori IRI
Station A	3.6/6.3	3.8/6.9	7.0/11.7	2.0/3.2	2.9/5.0	9.4/14.9
Station B	3.2/5.2	3.7/5.7	4.8/7.6	3.5/4.4	3.9/6.0	12.2/25.8
Station C	3.6/5.7	3.8/5.8	5.2/9.0	1.6/2.5	2.3/3.8	9.4/14.9
Station D	5.2/7.1	5.0/6.9	5.6/11.0	2.9/4.4	3.4/4.8	4.9/6.7
Station E	3.3/6.4	3.4/6.4	6.0/10.9	2.3/4.6	2.5/4.9	9.0/14.2
Station F	2.5/4.6	2.8/4.7	5.9/10.8	2.2/3.6	2.5/4.0	7.1/13.2
All 6 Stations	3.6/7.1	3.8/6.9	5.8/11.7	2.5/4.6	3.0/6.0	9.0/25.8
Batch Filter Data Fit	3.5/31.5	3.6/32.0	5.4/–	2.4/26.4	2.8/26.6	9.9/–

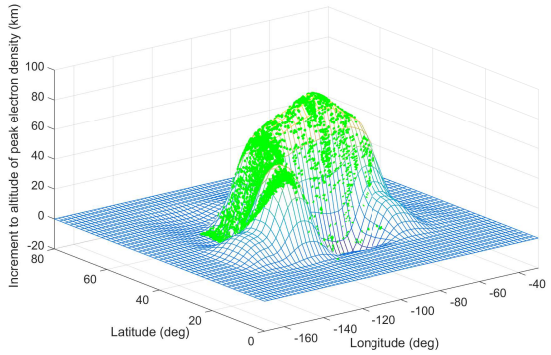

Fig. 4: Increments to *a priori* altitude of maximum electron density as estimated by the new algorithm for a local daytime case.

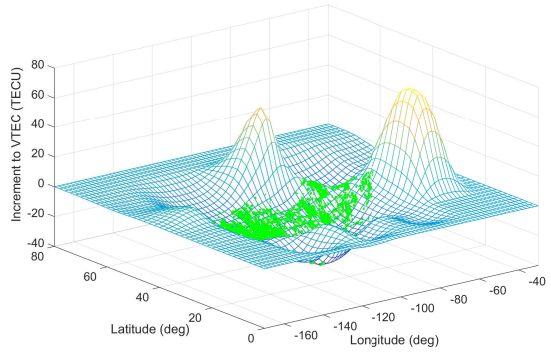

Fig. 5: Increments to *a priori* VTEC as estimated by the new algorithm for a local daytime case.

their large increments suspect. It is conjectured that these large humps are caused by poor VTEC observability at the edges of the pierce points region.

The reason for including the *a priori* penalty terms in the batch filter cost function of Eq. (16) has been to avoid such large unphysical filter anomalies. Obviously the desired result has not been achieved. Therefore, more research needs to be done to develop improved alternate means of adding *a priori* information about the ionosphere's electron density profile. Perhaps a better way of choosing the $P_{\bar{p}\bar{p}}$ *a priori* covariance matrix would alleviate such problems. One idea is to penalize the latitude/longitude curvature of the estimated increments to the *a priori* altitude, scale height, and VTEC profiles. Such penalties would constitute a regularization-type soft constraint on the final electron density profile estimate.

The large humps in Fig. 5 do not detract from the performance improvements found in the 4$^{\text{th}}$ column of Table I relative to the 5$^{\text{th}}$ and 6$^{\text{th}}$ columns. Recall that these three columns and this figure apply to the same case. The table deals mostly with the ability of the new technique to perform a sort of interpolation between measured slant TEC values. The new method can do this well. The humps in the figure have more to do with extrapolation beyond the edges of the data set. The method is not very good at extrapolation in its current form.

VIII. POTENTIAL IMPROVEMENTS AND TESTS

Various strategies hold the potential to improve this new filter. As mentioned in the previous section, it might be beneficial to switch to a penalty on the latitude/longitude curvature of the difference between the estimate and the *a priori* IRI model.

Another planned improvement is to extend this method to assimilate ionosonde and radio-occultation data too. Such data would tend to make the vertical profile parameters, i.e., the altitude of peak electron density and the scale height, more observable. Such data may allow the use of more complicated vertical profiles that can better approximate the profiles that exist in nature.

It might be wise to link the day and nighttime receiver biases. One way to do this would be to implement the estimator as sequential Kalman filter that would process a whole time series of data. The Kalman filter dynamic

model could dictate that receiver biases have limits to their likely drift rates.

It would be helpful to add more physics than is in the IRI model. Perhaps this could be done by using the SAMI2 model of some similar model.

If the number of elements in the $\boldsymbol{p}$ vector becomes too large to perform dense-matrix computations, i.e., on the order of tens of thousands or more, then it may be wise to implement some sort of Ensemble Kalman Filter to perform the estimation calculations. Care must be taken, however, not to preclude the possibility of achieving good estimates as a result of making highly unphysical assumptions in order to implement such a filter. The first author has developed a new type of Ensemble Kalman Filter that may allow more natural assumptions and thereby achieve good results.

Eventually a different vertical profile and parameterization should be tried. A single Chapman profile is too restrictive. It cannot capture the E layer or the D layer. Perhaps one should try a Booker profile or some similar general profile from the curve approximation literature.

It would be good to devise better tests for the fidelity of this new method's electron density profile estimates. *In situ* density measurements from a satellite or density vertical profiles from an incoherent scatter radar would help to evaluate this method's effectiveness. Some or all of the estimator improvements noted above should be implemented prior to taking the trouble of performing such evaluations. In its current form, the present method probably cannot estimate a very accurate 3D electron density profile.

IX. SUMMARY AND CONCLUSIONS

This paper has developed and tested a new method for estimating ionosphere electron density distributions based on GPS slant TEC data. The new method uses a bi-quintic spline to parameterize the latitude/longitude spatial variations of the parameters of a vertical electron density profile, a Chapman profile in the present case. It uses the corresponding 3D electron density distribution to compute modeled values of slant TEC. These values are compared with measured values within a batch nonlinear least-squares filter, and the filter updates the ionosphere parameterization in order to better fit the measured slant TEC values. The output of the filter is a parameterization that provides a data-based model of the ionosphere's 3D electron density distribution.

This new method has been tested using truth-model simulation data and using data from a network of more than 900 dual-frequency GPS receivers distributed over the continental U.S. The batch filter has been able to improve its fit to the slant TEC data by a factor of 1.5 to 3 in comparison to an *a priori* IRI-based model. It has also demonstrated a moderate improvement relative to a fixed-altitude, thin-shell ionosphere model. It does better at predicting the slant TEC at receiver locations other than those where its receivers are located. The best performance improvement has been for a daytime case.

These developments represent and early step in a process that envisions fusing additional data types, such as ionosonde data and radio-occultation data. The goal is to develop accurate real-time or near-real-time data-fusion-based estimates of the ionosphere's electron density profile and other states.

ACKNOWLEDGEMENTS

Mark Psiaki's work on this project has been supported in part by the National Research Council through a Senior Research Associate appointment at the Air Force Research Lab Space Vehicles Directorate, Kirtland AFB, Albuquerque, NM.

REFERENCES

[1] M. Hernández-Pajares, J. Juan, J. Sanz, R. Orus, A. Garcia-Rigo, J. Feltens, A. Komjathy, S. Schaer, and A. Krankowski, "The IGS VTEC Maps: A Reliable Source of Ionospheric Information Since 1998," *Journal of Geodesy*, vol. 83, pp. 263–274, 2009.

[2] G. Bust and C. Mitchell, "History, Current State, and Future Directions of Ionospheric Imaging," *Reviews of Geophysics*, vol. 46, 2008.

[3] C. Cooper, A. Chartier, C. Mitchell, and D. Jackson, "Improving Ionospheric Imaging via the Incorporation of Direct Ionosonde Observations into GPS Tomography," in *Proceedings of the General Assembly and Scientific Symposium (URSI GASS), 2014 XXXIth URSI*. Beijing: IEEE, Aug. 2014.

[4] R. Mitch, M. Psiaki, and D. Tong, "Local Ionosphere Model Estimation From Dual-Frequency Global Navigation Satellite System Observables," *Radio Science*, vol. 48, pp. 671–684, 2013.

[5] K. Chiang and M. Psiaki, "GPS and Ionosonde Data Fusion for Local Ionospheric Parameterization," in *Proceedings of the ION GNSS+ 2014*. Tampa, FL: ION, Sept. 2014, pp. 1163–1172.

[6] D. Bilitza, L.-A. McKinnell, B. Reinisch, and T. Fuller-Rowell, "The International Reference Ionosphere Today and in the Future," *Journal of Geodesy*, vol. 85, pp. 909–920, 2011.

[7] J. Huba, G. Joyce, and J. Fedder, "SAMI2 is Another Model of the Ionosphere (SAMI2): A New Low-Latitude Ionosphere Model," *Journal of Geophysical Research: Space Physics*, vol. 105, pp. 23 035–23 053, 2000.

[8] P. Gill, W. Murray, and M. Wright, *Practical Optimization*. New York: Academic Press, 1981.