Plasmasphere and Upper Ionosphere Contributions and Corrections During the Assimilation of GPS Slant TEC

D. C. Thompson
L. Scherliess
Jan Josef Sojka
Utah State University
R. W. Schunk

Follow this and additional works at: http://digitalcommons.usu.edu/physics_facpub

Part of the Physics Commons

Recommended Citation
Plasmasphere and upper ionosphere contributions and corrections during the assimilation of GPS slant TEC

Donald C. Thompson,1 Ludger Scherliess, Jan J. Sojka, and Robert W. Schunk

Received 6 October 2008; revised 6 February 2009; accepted 23 February 2009; published 29 April 2009.

Total electron content (TEC) measurements from ground stations to Global Positioning System (GPS) satellites provide a rich source of information about the Earth’s ionosphere. These data comprise a significant part of the typical data set used by various data ingestion and assimilation models of the ionosphere. For example, the Utah State University (USU) Global Assimilation of Ionospheric Measurements (GAIM) data assimilation model uses slant TEC, along with various other types of data, to obtain a global reconstruction of the ionosphere. There are presently two different USU GAIM models: the Gauss-Markov Kalman Filter (GAIM-GM), which is operational at the NASA Community Coordinated Modeling Center and the Air Force Weather Agency; and the Full-Physics Kalman Filter (GAIM-FP), which is presently run for scientific studies. TEC is the integrated electron density along the path from the ground to the GPS satellites, which orbit at approximately 20,200 km. The GAIM-FP modeled space ranges up to 30,000 km in altitude, so the entire TEC raypath is contained within the model space. Many ionosphere models do not model this entire region. For example, the GAIM-GM extends up to an altitude of 1400 km. It is necessary to account for the portion of TEC that is not modeled by some means. There are basically two simple techniques that are in common use: (1) correct the measured TEC by using a model of the upper ionosphere and plasmasphere to subtract their contributions, or (2) ignore the effect and assign the full measured TEC to the ionosphere within the assimilation model space. We present the effect of assimilating TEC measurements into the GAIM-GM using both the techniques mentioned above. It is found that derived quantities such as $N_{\text{m}F_2}$ are significantly degraded by ignoring the upper ionospheric contribution to TEC. This degradation is seen at all local times. The effect is most pronounced at night, when the F region densities are small and the upper ionospheric contribution to TEC is relatively large.

Citation: Thompson, D. C., L. Scherliess, J. J. Sojka, and R. W. Schunk (2009), Plasmasphere and upper ionosphere contributions and corrections during the assimilation of GPS slant TEC, Radio Sci., 44, RS0A02, doi:10.1029/2008RS004016.

1. Introduction

In recent years a number of techniques and models have been developed to generate three-dimensional representations of the ionosphere by ingesting or assimilating available data [Scherliess et al., 2004, 2006, 2008; Schunk et al., 2004, 2005a, 2005b; Bust et al., 2004; Khattatov et al., 2005; Mandrake et al., 2005]. TEC measurements from ground stations to Global Positioning System (GPS) satellites provide a rich source of information about the Earth’s ionosphere. GPS ground stations are widely distributed around the globe, and data from thousands of sites are available via the Internet. These data comprise a significant part of the typical data set used by various data ingestion and assimilation models of the ionosphere. Depending on the specific model or technique, other data types may also be used during the construction of the ionosphere specification. Generally, the data are combined with a model of the ionosphere (either parameterized, empirical, or first-principle physics) to generate a fully three-dimensional specification. In this fashion, measurements of integrated slant TEC from scattered ground stations can be used to produce electron density values at any combination of latitude, longitude, and altitude, even over regions some distance removed from available data.
[3] Slant TEC are the integrated electron densities along the paths from ground stations to GPS satellites, which orbit at approximately 20,200 km. Usually the largest contribution to the total is obtained in the ionosphere (particularly the F region). However, contributions from the plasmasphere (protonosphere) can be significant, comprising half or more of the TEC under some conditions [Lunt et al., 1999; Balan et al., 2002; Yizengaw et al., 2008]. The relative contribution of the plasmasphere to TEC is largest during nighttime and solar minimum conditions when F region densities are lowest. While some models contain the entire altitude range covered by GPS TEC [Scherliess et al., 2008] others have only limited altitude coverage, on the order of up to 1500 km [Mandrake et al., 2005; Scherliess et al., 2006]. Since the plasmaspheric contribution to the TEC is not negligible, it is necessary to account for the portion of TEC that is not modeled by some means when using GPS TEC with models that do not automatically account for the plasmaspheric contribution. There are basically two simple techniques that are in common use: (1) ignore the plasmaspheric effect and assign the full measured TEC to the ionosphere within the assimilation model space, or (2) correct the measured TEC by using a model of the plasmasphere to determine its contribution and subtract this value from the data before assimilation. While using technique 1 must bias the ionospheric specification to some degree (with the F region densities being biased too high), it has been claimed that resulting ionospheric parameters such as \(N_e F_2\) are only modestly affected [Mandrake et al., 2005]. The fidelity of the ionospheric specification obtained using technique 2 depends to some degree on the fidelity of the plasmasphere model used to correct the data. However, we will show in this paper that even a very simple representation of the plasmasphere produces a significantly improved \(N_e F_2\) specification compared to the results obtained if the plasmaspheric contribution is ignored. The improvement is seen at all local times, but is most pronounced at night. At night the F region densities decrease significantly while the plasmaspheric densities are relatively stable throughout, hence the plasmaspheric contribution to TEC is expected to be larger at night.

2. Model

[4] There are presently two different USU GAIM models: the Gauss-Markov Kalman Filter (GAIM-GM), which is operational at the NASA Community Coordinated Modeling Center (CCMC) and the Air Force Weather Agency (AFWA); and the Full-Physics Kalman Filter (FPKF), which is presently run for scientific studies [Schunk et al., 2004, 2005a, 2005b; Scherliess et al., 2004, 2006; McDonald et al., 2006; Thompson et al., 2006; Sojka et al., 2007; Jee et al., 2007, 2008]. The FPKF has an altitude range beyond 20,200 km, and so it automatically accounts for the plasmaspheric contribution to slant TEC. The GAIM-GM, on the other hand, only ranges up to approximately 1400 km, and so the plasmaspheric contribution to TEC must be considered with this model. For the work in this paper we will only use the GAIM-GM model.

[5] The GAIM-GM uses the physics-based Ionosphere Forecast Model (IFM) as a background, and solves for deviations from the background using a Kalman filter to assimilate various ionospheric data. The IFM takes account of \(\mathrm{O}^+, \mathrm{O}_2^+, \mathrm{N}_2^+, \mathrm{NO}^+, \text{ and } \mathrm{H}^+\), and covers the E region, F region, and topside up to 1400 km. The IFM is described by Schunk et al. [1997] and validation of the model is given by Zhu et al. [2006]. The GAIM-GM Kalman filter evolves density perturbations and associated errors over time using a statistical Gauss-Markov process. A database of 1107 2-day IFM runs is used to construct the error covariance matrix, which substantially reduces the computational demands of the GAIM-GM. The accuracy of the GAIM-GM model for ionospheric specifications has been shown by Scherliess et al. [2006], Thompson et al. [2006], Sojka et al. [2007], and independently by Decker and McNamara [2007], McNamara et al. [2007, 2008], and McDonald et al. [2006]. The GAIM-GM can assimilate four types of data: slant GPS TEC, in situ electron densities \((N_e)\), electron density profiles (EDPs) from ground-based radars, and optical emissions from reactions that vary approximately as \(N_e^2\) for which the path from emission to detection is optically thin. The GAIM-GM can operate with varying numbers of these data in any combination, including no data. Output from the GAIM-GM is a global three-dimensional time-varying distribution of \(N_e\) from 90 to 1400 km. Reduced parameters, such as slant or vertical TEC, \(N_e F_2\), scale heights, etc., are easily derived from the \(N_e\) distribution.

[6] The GAIM-GM accepts the full measured TEC from the ground station to the GPS satellite. Since the integration path for this data extends in altitude well beyond GAIM-GM model space, the plasmaspheric contribution to the measured TEC must be considered. In the default configuration the GAIM-GM uses an internal representation of the plasmasphere to estimate the plasmaspheric contribution to the measured TEC. This estimate is subtracted from the measured TEC before the data is assimilated, leaving the ionospheric part of the data. The default plasmaspheric representation in the GAIM-GM is remarkably simple. Following the work of Davies [1999] the plasmasphere is assumed to vary exponentially in height with a constant scale height and is constant over time. The simple model is spherically symmetric. Therefore, the model plasmasphere ignores the geomagnetic field-aligned nature of the plasmasphere, and does not account for high-latitude
The location of 201 GPS ground stations in case 0 is about 1.26 times as large as that obtained in case 1. The two resulting ionosphere specifications are then compared to independent data to show the differences in the fidelity of the two configurations.

[8] The study period used in this work is from day 081 through 110, inclusive, in 2004. This period is used for several reasons. It is one of the validation periods used for the GAIM-GM model as presented by Scherliess et al. [2006] and Thompson et al. [2006]. The solar cycle during this period is low, with $F_{10.7}$ flux ranging from 80 to 100, so the plasmappheric contribution to TEC should be significant, especially at night. Also, while geomagnetic conditions for this period are not particularly stormy, there is significant $K_p$ variation (up to $K_p = 6$ for short periods, with typical values around $K_p = 3$) so this period will contain significant day-to-day space weather that the model should reproduce. For this study slant TEC from 201 GPS ground stations and EDPs from 19 ionosondes were assimilated. The global distribution of GPS ground stations and ionosondes is shown in Figure 1. The GPS ground stations are distributed fairly well over all the landmasses around the globe, with higher densities over North America and Europe. The majority of the ionosondes are clustered in North America. The independent data used for comparisons in this study are (1) vertical TEC measured by the TOPEX satellite [Fu et al., 1994] and (2) $N_pF_2$ measured by a dynasonde located at the Bear Lake Observatory (BLO) in northern Utah. The location of BLO is indicated in Figure 1. The process used to extract $N_pF_2$ and other quantities from the BLO ionograms is given by Wright and Pitteway [1998]. TOPEX measures the vertical TEC over the oceans from sea level up to its orbit altitude of 1366 km, so the integrated electron content from the model up to the equivalent altitude is used for comparison. It is well known that the TOPEX TEC data are biased and previous studies [e.g., Orus et al., 2002] have found that this bias is on the order of several TECu. For our work here we have subtracted 4 TECu from the TOPEX TEC values, which results in simultaneously bringing the GAIM-GM into statistical agreement with the TOPEX data and measured $N_pF_2$ over a wide range of geophysical conditions [Scherliess et al., 2006].

3. Method

[7] It is possible to configure the GAIM-GM so that the plasmaspheric correction discussed in the previous section is not performed. In this mode the entire measured slant TEC is assigned to the ionosphere, which necessarily changes the ionospheric specification produced. In this work we run the GAIM-GM twice over a 30-day study period: once in the default configuration where the plasmaspheric correction to slant TEC is considered (which will be identified as case 1), and then in the mode where the plasmasphere is ignored (which will be identified as case 0). The two resulting ionosphere specifications are then compared to independent data to show the differences in the fidelity of the two configurations.

[9] Electron density profiles over the Bear Lake Observatory for a specific period (day 090 at local midnight) for the two cases are shown in Figure 2. In addition, the values of the coincident measurement of the BLO $h_pF_2$ and $N_pF_2$ are identified in Figure 2. Note that the BLO data were not assimilated by the model in either run. As seen in Figure 2, the peak density obtained in case 0 is significantly higher than that obtained in case 1. The value of $N_pF_2$ in case 0 is about 1.26 times as

![Figure 1. The location of 201 GPS ground stations (crosses) and 19 ionosondes (squares) used in this work. The location of the Bear Lake Observatory (BLO) is indicated by the arrow.](image-url)
large as the measured value, while that obtained in case 1 is 1.09 times the measured value. Even though there is variation in each specific result, it is typical that the case 0 values are biased high relative to case 1. The case 0 value of $h_m F_2$ is also higher than that obtained in case 1. The error covariances in the GAIM-GM are generated such that larger densities result when the profile is at a higher altitude, particularly at night. This is the expected result, since recombination rates are lower at higher altitudes, allowing a larger $N_e$ to be maintained. Not all data assimilation models will exhibit this particular behavior, depending on the details of each model's specific construction.

[10] The variation of the model $N_m F_2$ compared to all the available BLO $N_m F_2$ measurements in the study period is shown in Figure 3. Figure 3 shows the results from the IFM (Figure 3, top), case 0 where the plasmaspheric electron content was assigned to the ionosphere (Figure 3, middle), and case 1 where the plasmaspheric electron content was subtracted from the slant TEC data before assimilation (Figure 3, bottom). In each plot the value of the BLO $N_m F_2$ ranges along the logarithmic horizontal axis. The coincident model $N_m F_2$ ranges along the logarithmic vertical axis over the equivalent range. For each of 7302 BLO $N_m F_2$ measurements the coincident model $N_m F_2$ is calculated. The BLO and model values form an ordered pair that locates a bin in Figure 3, and the number of counts in the appropriate bin is incremented. Once all 7302 results have been counted, each distribution of counts is normalized so that the distributions in Figure 3 can be compared to each other. An ideal result would be for each distribution to lie along the dashed line where the BLO and model results are

Figure 2. Electron density profiles produced by the GAIM-GM for case 0 (the solid curve) and case 1 (the dashed curve) at BLO for 2004, day 090, at midnight local time. The coincident measured $N_m F_2$ and $h_m F_2$ are indicated by the vertical and horizontal dotted lines, respectively.

Figure 3. Distributions of GAIM-GM versus BLO $N_m F_2$ for (top) IFM, (middle) case 0, and (bottom) case 1.
equal. Since the models do not match the measurements exactly, and the measurements themselves have some level of noise and uncertainty, the actual distributions are spread away from the ideal. The absolute magnitude of the counts in any given bin is not particularly important. Rather, it is the overall shape of the distributions, their width, biases, and possible off-diagonal blobs that indicate the qualitative fidelity of the results.

\[11\] The distribution of the IFM results shows the largest differences. This is expected since the IFM is driven entirely by geophysical indices (Kp, Ap, F10.7) and does not assimilate data. The range of variation in IFM is considerably less than that exhibited by the measured $N_mF_2$, so the IFM distribution is relatively short and wide. Since the time-dependent climatology of the IFM model tends to give results that are similar, from one day to the next, there is an evident horizontal stratification in the IFM distribution that results from the day-to-day variation of the measured $N_mF_2$ owing to space weather effects not adequately modeled by the IFM. Also, there appears to be a general bias of the IFM result relative to the measurements, with the IFM being too large more often than too small.

\[12\] Assimilation of the globally distributed slant TEC and ionosonde EDPs with the GAIM-GM significantly improves the $N_mF_2$ comparison, as seen in Figure 3 (middle and bottom). In both case 0 and case 1 the horizontal stratification seen in the IFM case has been eliminated. The range of predicted $N_mF_2$ values has also increased (lower minimums and higher maximums) which better matches the range of $N_mF_2$ measured. Also, the spreads of the distributions are relatively more narrow, indicating a better overall fit to the BLO data.

However, the distribution of results seen in case 0 show a consistent bias. In this case the predicted $N_mF_2$ are generally too large. While the effect is noticeable at all densities, the trend is for the bias to become more pronounced at lower densities. This is expected since the plasmaspheric TEC, which is erroneously applied to the ionosphere for the case 0 study, is relatively larger compared to the ionospheric contribution for lower ionospheric densities. The case 1 distribution, while far from perfect, does not show the strong bias seen in case 0. In this case the distribution is nearly centered about the ideal line, indicating that the correction applied to the slant TEC values to account for the plasmaspheric content has the correct average effect.

\[13\] Quantitative descriptions of the model comparisons are given in Tables 1–3. Since the distributions are not Gaussian, the distributions are described in terms of quartiles and deciles in Table 1. In Table 1, the first quartile, second quartile (median), and the third quartile values, the interquartile range (the third quartile minus the first), and the first and ninth decile of the ratio of the model result divided by the measured BLO $N_mF_2$ are given for each case. The values in Table 1 should be interpreted as ratios. For example, the median value for IFM is given as 1.28 in Table 1. Hence, on average, the IFM $N_mF_2$ value is a factor of 1.28 too large compared to the BLO measurement over the study period. Note that while the interquartile spread is improved in both case 0 and case 1 relative to IFM, the median value obtained in case 0 (1.40) is the largest of all three cases. The median value of case 1 is very near the ideal value of unity. Tables 2 and 3 give similar results to Table 1, but select the results for daytime (0800 to 1700 local time) and

<table>
<thead>
<tr>
<th>Case</th>
<th>First Quartile</th>
<th>Second Quartile</th>
<th>Third Quartile</th>
<th>Q3 – Q1*</th>
<th>First Decile</th>
<th>Ninth Decile</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFM</td>
<td>0.85</td>
<td>1.28</td>
<td>1.95</td>
<td>1.10</td>
<td>0.57</td>
<td>2.75</td>
</tr>
<tr>
<td>Case 0</td>
<td>1.13</td>
<td>1.40</td>
<td>1.78</td>
<td>0.66</td>
<td>0.93</td>
<td>2.67</td>
</tr>
<tr>
<td>Case 1</td>
<td>0.81</td>
<td>0.98</td>
<td>1.22</td>
<td>0.40</td>
<td>0.68</td>
<td>1.51</td>
</tr>
</tbody>
</table>

*The interquartile range (the third quartile minus the first).
nighttime (2000 to 0500 local time). Note that the bias (median) of the case 0 result is larger at night (a factor of 1.56) than in the day (1.23), as expected. The case 1 result shows its best agreement at night (1.03), with the daytime median being about 5% low (0.95).

The comparison of the model results to the TOPEX vertical TEC during the study period is shown in Figure 4. Recall that the TOPEX data were obtained globally, over the oceans, and so the comparisons shown in Figure 4 are the accumulated comparisons of the model results to the global TEC variation. Figure 4 shows comparisons between the IFM, case 0, and case 1 model results, as before. The vertical TEC from each model run is obtained by integrating the model result up to the 1336 km TOPEX altitude for each of more than 134,000 TOPEX measurements. Note that the TOPEX data are available in 1-s intervals. For the work presented here these data were averaged into 6-s bins to reduce noise in the data. As in Figure 3 the independent data and model results form an ordered pair, which identify a bin in Figure 4. Each ordered pair is used to increment a counter in the appropriate bin, and once all the data have been counted a distribution of counts is produced. Each distribution is normalized, as before, so that the three cases can be more easily compared. In the TOPEX comparison the axes are linear. The TOPEX TEC is shown along the horizontal axis, and the model result ranges along the vertical axis.

The distribution obtained by running IFM has the largest deviation from the ideal. For values of TEC less than about 10 TECu, which corresponds to nighttime conditions, the IFM results tend to be significantly low. This is not surprising since the nighttime maintenance of the ionosphere in IFM is highly dependent on the empirical drivers it uses. The nighttime densities are particularly sensitive to the neutral winds, which may not be optimal for this study period over much of the globe. In the range from 10 to 30 TECu, which would contain much of the midlatitude daytime results, the IFM does better, with a distribution that is roughly centered about the optimal line. Above 30 TECu, which is associated with the equatorial ionization anomalies, IFM produces results that are consistently too low for this study period.

As was seen with NmF2, the assimilation of the slant TEC and ionosonde EDP data dramatically improves the ionospheric specification. In both case 0 and case 1 the resulting distributions are noticeably narrower and tend to follow the optimal line. However, there is a bias in the case 0 result. The bias can be seen in the small inset within each plot in Figure 4, which shows the distribution of results over a limited range. Note that the peak of the distribution in case 0 (Figure 4, middle) is slightly shifted above the optimal line. This is the direct result of applying the plasmasphere contribution to the ionosphere, almost all of which is below the TOPEX orbital altitude of 1366 km. The effect is that the distribution in case 0 is biased too high by the TEC contained in the plasmasphere, as one might expect. The case 1 result, where the slant TEC data were corrected using the simple plasmaspheric model described above, does not show this bias.

### 5. Conclusions

We have assimilated slant TEC measurements from 201 globally distributed GPS ground stations and EDPs from 19 ionosondes into the Utah State University Gauss-Markov Kalman filter for a 30-day study period. We have done this for two cases. In the first case, the contribution to slant TEC from the plasmasphere was assigned to the GAIM-GM model space (the ionosphere). In the second case the slant TEC data were corrected by subtracting the plasmaspheric contribution obtained from a simple model of the plasmasphere. It is found that predictions of $N_{mF2}$ are significantly degraded by ignoring the upper ionospheric contribution to TEC. This degradation is seen at all local times. The effect is most pronounced at night, when the F region densities are small and the upper ionospheric contribution to TEC is relatively large. On average, failure to correctly account for the plasmasphere results in $N_{mF2}$ values that are 1.4 times too large, a very significant bias. The effect of the predicted TEC is a constant bias, as expected. We find that correcting slant TEC to account for the plasmasphere produces results that are consistent with the observations.

### Table 3. Quartiles and Deciles of the Distributions of the Ratio of Predicted and Measured BLO $N_{mF2}$ for Values During Nighttime*

<table>
<thead>
<tr>
<th>Case</th>
<th>First Quartile</th>
<th>Second Quartile</th>
<th>Third Quartile</th>
<th>Q3 – Q1</th>
<th>First Decile</th>
<th>Ninth Decile</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFM</td>
<td>0.72</td>
<td>1.12</td>
<td>1.53</td>
<td>0.82</td>
<td>0.49</td>
<td>2.03</td>
</tr>
<tr>
<td>Case 0</td>
<td>1.24</td>
<td>1.55</td>
<td>1.89</td>
<td>0.65</td>
<td>1.00</td>
<td>2.32</td>
</tr>
<tr>
<td>Case 1</td>
<td>0.85</td>
<td>1.03</td>
<td>1.26</td>
<td>0.41</td>
<td>0.67</td>
<td>1.51</td>
</tr>
</tbody>
</table>

*Nighttime is 2000–0500 local time.
spheric contribution prior to assimilation simultaneously improves the resulting distributions of both $N_{m}F_2$ (compared to measurements at the Bear Lake Observatory) and TEC (as measured by TOPEX). Even the simple correction used here removes significant biases from the model output, and dramatically improves the model fidelity compared to having no plasmaspheric correction at all. However, there is still significant spreading of both distributions about the optimal results, so there continues to be room for improvement in the model. One avenue available to improve the model is to include a better plasmasphere model. It is likely that a physical model of the plasmasphere, which included proper geomagnetic geometry and space weather dynamics would further improve the fidelity of the assimilation model by reducing the statistical spread of the distributions in addition to removing biases.

References


L. Scherliess, R. W. Schunk, J. J. Sojka, and D. C. Thompson, Center for Atmospheric and Space Sciences, Utah State University, Logan, UT 84322-4405, USA. (thompson@cc.usu.edu)