Spatial variability in the ionosphere measured with GNSS networks

Ragne Emardson,1 Per Jarlemark,1 Jan Johansson,1,2 and Sebastian Schäfer3

Received 15 January 2013; revised 11 September 2013; accepted 4 October 2013; published 28 October 2013.

[1] Traveling ionospheric disturbances (TIDs) appear as medium-scale TIDs at midlatitudes and as polar cap patches at high latitudes. Both can have a negative impact on Global Navigation Satellite Systems (GNSS) measurements, although the amplitude is of tenths of a total electron content unit (TECU), 1 TECU = 1016 el m\(^{-2}\). Due to their spatial extension, they affect GNSS measurements using receivers separated with distances up to ~1000 km. We present statistical measures of the ionospheric spatial variability as functions of time in solar cycle, annual season, and time of day for different geographical locations in Europe. In order to perform this spatial characterization of the ionosphere, we have used archived GPS data from a 13 year period, 1999–2011, covering a complete solar cycle. We find that the ionospheric spatial variability is larger for the northern areas than for the southern areas. This is especially pronounced at solar maximum. For the more northern areas, the ionospheric variability is greater during nighttime than during daytime, while for central Europe, the variability is larger during daytime. At solar maximum, the variability is larger during the months October and November and smaller in June and July.


1. Introduction

[2] A medium-scale traveling ionospheric disturbance (MS-TID) is the signature in the ionosphere of the passage of atmospheric gravity waves, the ions being forced along the field lines by the neutral air winds driven by the pressure wave [e.g., Hocke and Schlegel, 1996]. The generation process is related with meteorological phenomena like neutral winds and the solar terminator [Somsikov, 1995].

[3] A statistical study regarding the occurrence and seasonal variations of MS-TIDs at low and middle latitudes by Kotake et al. [2006] leads to the following picture: The MS-TIDs activity is highest at daytime during the winter season, when the gradient of the neutral temperature near the mesopause is so low, that the atmospheric gravity waves can propagate. This gradient is steep in the summer season, preventing wave propagation. The nighttime activity of the MS-TIDs is also higher during winter. In general, the production of TIDs at midlatitudes increases with the solar activity. Geomagnetic storms can lead to the production of so-called large-scale TIDs due to heating in the thermosphere at high latitudes by the Joule effect [e.g., Jacobson et al., 1995]. This leads to an energy transfer toward lower latitudes in the form of thermospheric waves interacting with ionospheric ions.

[4] At high latitudes, the TIDs are also called polar cap patches due to their location of production in the polar cap and auroral oval. Crowley [1996] defined polar cap patches as plasma structures with a horizontal extent of at least 100 km and a plasma density of at least twice the density of the surrounding background plasma. Such enhancements of the ionospheric plasma density were proposed to be produced at the dayside at auroral and subauroral latitudes [e.g., Weber et al., 1984]. In general, the patches are observed to convect antisunward to higher latitudes in the polar region [e.g., Carlson et al., 2002, 2004] and can reach the nightside auroral oval [e.g., Lockwood and Carlson, 1992]. As the geomagnetic field lines connected to high latitudes reach farther out in the space than field lines at midlatitudes, the polar ionosphere is strongly impacted by geomagnetic perturbations due to increased solar activity [e.g., Moen et al., 2007]. The number and lifetime of polar cap patches are expected to increase during time periods with increased solar activity [Schunk and Sojka, 1987].

[5] It was first observed by Foster [1984], using the Chantika incoherent scatter radar, that the polar cap plasma exhibits seasonal variations. This has been modeled by Schunk and Sojka [1987]. The outcome of these studies is that the plasma densities drawn into the polar cap are much larger and have longer lifetime in winter than in summer. The decreased lifetime in summer was explained by the maintenance of background densities due to solar EUV radiation and the increased recombination rates resulting from increased ion temperatures and differences in the thermosphere composition. The seasonal effect on the polar cap patches in the high-latitude nightside ionosphere has been confirmed by observations and modeling efforts [e.g., Wood and Pryse, 2010].

[6] Free electrons play an important role in affecting Global Navigation Satellite Systems (GNSS) measurements when the signal propagates through the ionosphere [e.g., Hoffman-Wellenhof et al., 1994]. The total integrated electron density
along a propagation path could be on the order of hundreds of total electron content unit (TECU), 1 TECU = 10^{16} \text{ el m}^{-2}. In high-precision GNSS usage, this effect is compensated by using observations at different frequencies. However, several GNSS applications are sensitive to atmospheric spatial variations. Relative GNSS measurements, for example, network Real-Time Kinematic (RTK), are very sensitive to such effects [Emardson et al., 2010]. Thus, despite that MS-TIDs amplitudes are typically less than 1 TECU [Hernandez-Pajares et al., 2006], they can have a negative impact on results obtained from relative GNSS measurements. Due to their spatial extension, they affect GNSS measurements using receivers separated with distances up to \sim 1000 \text{ km}, thus including typical sizes of RTK networks.

Based on historical GPS data, it is possible to characterize the variability of the ionosphere. In this study, we have used 13 years of GPS data from European networks to cover a complete solar cycle. The geographical location covers both high and middle latitudes. In this paper we have chosen to focus on spatial scales from 500 km and downward. Characterization is performed by using observations of the ionospheric delay from three GNSS sites forming a triangle in order to interpolate the ionospheric delay for the site in the middle of the triangle. By then comparing the interpolated time series with those actually measured, the small-scale spatial variability of the ionosphere may be characterized.

2. Methodology

[8] GNSS data are commonly used for estimation of total electron content (TEC) parameters [e.g., Mannucci et al., 1998]. Here we derive measures of the ionospheric spatial variability using ionospheric delay measurements from GNSS. From slant ionospheric delays observed at three sites forming a triangle, we interpolate the expected ionospheric delay for the site in the middle of the triangle. These interpolated values are then compared to those observed. Their difference contains information about the small-scale variability of the electron content distribution, while larger scales are removed.

[9] We calculate the slant ionospheric delays by forming the L4 combination of the phase observations at the two GPS frequencies L1 and L2, \varphi_1 and \varphi_2, respectively, multiplied with the respective wavelength, \lambda. This expression contains information of the ionosphere, I, as well as potential nonionospheric variations from multipath and measurement noise, \mu, receiver inter-frequency L1-L2 bias, \tau_r, satellite inter-frequency L1-L2 bias, \tau_s, and phase ambiguities, n, i.e.,

$$\Theta = \lambda_1 \varphi_1 - \lambda_2 \varphi_2 = I + n + \tau_s + \tau_r + \mu$$ (1)

[10] In our characterization, we only use receivers where the L1-L2 bias variation is considered to be relatively small.

Figure 1. Geographical locations of the sites used for the characterization of the ionosphere. From the three outer sites in each triangle, we interpolate the vertical TEC in the middle site. The difference between this value and the measured value at the middle site contains information about the small-scale variability of the TEC distribution. PTBB (Braunschweig) was used replacing Potsdam during the period 30 March 2009 to 15 February 2011.

Figure 2. Data on solar activity from 1995 to 2012. The different curves show monthly mean sunspot number (red), daily solar flux (blue), and bihourly vertical TEC for latitude 60° north and longitude 15° east (green). Sunspot number data are from http://sidc.oma.be/sunspot-data/, solar flux data are from http://www.ngdc.noaa.gov/nndc/struts/Form?t=102827&s=4&d=8,22,9, and vertical TEC data are from ftp://cddis.gsfc.nasa.gov/pub/gps/products/ionex/. The vertical lines identify the time period used in this study.
This time constant is likely related to the TID travel time over the network. Hence, the average of $\Theta$ over time scales longer than 1 h contains basically no TEC contribution. We therefore claim that no significant component of the TEC variability is removed in our processing.

[13] We refer to Appendix A for a more detailed description of the methodology we have used.

3. Data Set

[14] In order to characterize the variability of the ionosphere, we have used 13 years of GPS data from five geographical areas in Europe. These areas span from central Europe in the south to Svalbard in the north. Figure 1 shows the geographical locations of the sites we use in this study, thus identifying the five geographical regions that we later refer to in this paper. The data set spans from 1 January 1999 to 31 December 2011. This period covers at least one complete solar cycle. Historically, these have a length of about 11 years. However, variations in length exist, and the behavior of the latest solar cycle indicates that it may have been longer than average. Solar maximum was identified around 2001 and solar minimum around 2008. Figure 2 shows the solar activity for the identified period based on three different measurement techniques.

4. Results

[15] We have calculated hourly RMS values of TEC variations. This has been performed over five geographical regions in Europe over a 13 year period. Figure 3 shows an example of TEC variations for one of the regions, namely, middle Sweden, during two separate weeks with high respective low ionospheric variability.

[16] In order to analyze long-term variations during a complete solar cycle, we have derived monthly RMS values of the TEC variability for each region. Figure 4 shows the monthly TEC variability as a function of time for the different regions. The variability is much higher during solar maximum than during solar minimum, especially for northern Europe. The variability is also, in general, larger for the northern regions.
The variability at Svalbard is very high as the short-period data are available for that region. During solar maximum, a strong seasonal variation can be seen. The variability tends to peak during the winter months for all regions, i.e., the variability is largest in the months of October–February and smallest from May to August. For the more northern regions, the variability at the winter periods can be 3 times those of the summer periods. During solar minimum, no seasonal variation is evident.

[17] For a shorter time period, the year 2001, we performed a more detailed study of the statistical distribution of the ionospheric variability. Figure 5 shows the ionospheric variability as a function of time. The four panels in the figure show the RMS, median, 95th percentile, and 99th percentile, respectively. Hence, the first panel in the figure contains the same information shown for the year 2001 in Figure 4. We can see an expected increase for the 95th and 99th percentiles compared to the RMS variations, while much of the structures in the first panel for the RMS are apparent in the third and fourth panels for the 95th and 99th percentiles. The analysis is based on the 1 h generated data as described above, where variations have been already averaged. This means that on a higher sampling basis, which is often used when processing GNSS data, the variability is expected to occasionally reach even higher values than those shown in Figure 5 (fourth panel).

[18] In order to analyze daily variations of the TEC variability, we have grouped the data from each year by the hour of the day. Figure 6 shows the TEC variability as a function of the local time of the day for the five geographical areas.

[19] During solar maximum, the patterns are very different for northern and central Europe. For northern Europe, the variability is at its maximum during nighttime and lower at daytime, while for central Europe, the variability is at its highest at noon. During solar minimum, the diurnal signatures for northern Europe are not that significant. During this period, we can also see local maxima at noon also for this region. As we could see in Figure 4, the variability is much larger in the northern regions than in the more southern regions during solar maximum. The results for the north and middle Sweden areas are very similar over the entire 13 year study period. The daily pattern in the results from Svalbard is different from the other regions. This may be due to the location very far north and thus limited difference between day and night conditions. We cannot, however, exclude that the Svalbard results are affected by instrumentation issues. The data set we used contained, for example, more frequent gaps than those from the other regions.

[20] Figure 7 shows the statistical distribution of the ionospheric variability as a function of the local time of the day. The four panels in the figure show the RMS, median, 95th percentile, and 99th percentile, respectively. Similarly to Figure 5, much of the structures appear in all panels.
5. Conclusions

[21] We have presented statistical measures of the ionospheric spatial variability as functions of time in solar cycle, annual season, and time of day. The analysis has been performed for five different geographical locations in Europe studying spatial scales below 500 km, thus capturing MS-TIDs and polar cap patches which are normally of scales below 1000 km.

[22] In general, the results show that the variability increases with increasing latitude. We presume that polar cap patches dominate in the three most northern regions, while MS-TIDs dominate in central Europe. Hence, the polar cap patches lead to higher ionospheric variability than the MS-TIDs. The different daily behavior of TEC variability at high and middle latitudes is due to the fact that at high latitudes, variations caused by geomagnetic effects dominate, while at middle latitudes, insolation effects are more pronounced. This is seen both from the nighttime maximum for high latitudes (dominated by the magnetotail configuration) and the peak at noon for midlatitudes.

[23] Polar cap patches are more intense during solar maximum than during solar minimum [Schunk and Sojka, 1987]. We see that the seasonal dependence is also much larger during solar maximum than during solar minimum. This is in agreement with the expected occurrence probability of both MS-TIDs and polar cap patches, which have a reduced lifetime during summer. Furthermore, our observations show that the seasonal dependence is much larger for the polar cap patches than for the MS-TIDs.

[24] The observed daily characteristics of the ionospheric variability can be explained by the known activity pattern of midlatitude MS-TIDs and high-latitude polar cap patches. In central Europe, MS-TIDs dominate. Due to insolation, these have a maximum probability of occurrence during daylight, and thus, the variability is higher at that time. In contrast to that, for northern Europe, polar cap patches are the dominating phenomena. These effects lead to high nighttime variability during solar maximum. However, the occurrence of polar cap patches depends much more on geomagnetic and solar activity conditions than the MS-TIDs. Therefore, the variability is much larger at the northern regions during solar maximum (1999–2006). The peak at noon, usually caused by MS-TIDs, appears also at these regions during solar minimum (2007–2010). One can conclude that the effect of MS-TIDs is visible at high latitudes only during solar minimum, while they are dominated by the effect of the polar cap patches during solar maximum.

Appendix A

[25] We derive measures of the ionospheric spatial variability $\sigma_\Delta I$. This measure is based on the deviation between interpolated and true ionospheric electron contents, $\Delta I$, at a center of a triangle. We find $\Delta I$ by forming the L4 combination of the received signal phase and removing nonionospheric contributions.

[26] We model the signal received by a single GNSS receiver. The phase observed by a receiver at frequency ($A$) and a frequency ($B$) can be described by (A1) and (A2), respectively, where $\phi$ is the measured phase in fraction of cycles, $\lambda$ is the signal wavelength, $\rho$ is the true geometrical distance between the receiver and the satellite, $N$ is the integer number of cycles referred to as the ambiguity parameter, and $f$ is the signal frequency. $\delta t'$ and $\delta t''$ represent the satellite and receiver clock errors, respectively; $\ell_1$ is the signal delay in the lower part of the atmosphere, referred to as the troposphere; $\ell_2$ is the signal delay in the ionosphere part of the atmosphere; and $\mu$ is the signal multipath and receiver measurement error. Thus

$$\phi^A = \frac{1}{\lambda_A} \rho + N_A + f^A (\delta t'_A + \delta t''_A) + \frac{1}{\lambda_A} \ell_1 + \frac{1}{\lambda_A} \ell_2 + \mu_A \quad (A1)$$

$$\phi^B = \frac{1}{\lambda_B} \rho + N_B + f^B (\delta t'_B + \delta t''_B) + \frac{1}{\lambda_B} \ell_1 + \frac{1}{\lambda_B} \ell_2 + \mu_B \quad (A2)$$

[27] We can form the combination of (A1) and (A2), using $c_0$ as the speed of light, as

$$\lambda_A \phi^A - \lambda_B \phi^B = \lambda_A^\ast - \lambda_B^\ast + \lambda_A N_A - \lambda_B N_B + c_0 (\delta t'_A - \delta t''_B) + \delta t'_A - \delta t''_A + \lambda_A \mu_A - \lambda_B \mu_B \quad (A3)$$

[28] This combination removes all frequency-independent effects such as geometrical distances and signal delay in the troposphere. A fraction of the ionospheric effect on the two frequencies is left, however. Hence, this combination can be used to derive information about the ionosphere, the total electron content along the line of sight.

[29] By forming the L4 combination of the signals received at the two GPS frequencies L1 and L2, we obtain a measure of the slant ionospheric delay. This estimate, however, contains also potential nonionospheric variations. Equation (A3) can be written as

$$\Theta^I = \Theta^p + \kappa (\lambda_A \phi^A - \lambda_B \phi^B) = I^p + n + \tau_s + r_s + \mu \quad (A4)$$

[30] In equation (A4), $I^p$ is the slant TEC, and $\kappa$ is the scaling factor between L4 measurements and TEC. Hence, this derived L4 combination, $\Theta^I$, contains information of the ionosphere as well as variations from multipath and measurement noise, $\mu$, receiver interfrequency L1-L2 bias, $r_s$, satellite interfrequency L1-L2 bias, $\tau_s$, and phase ambiguities, $n$.

[31] In our analysis, we only use receivers where the L1-L2 bias variation is considered to be relatively small at constant receiver temperatures [Rieck et al., 2003]. The satellite L1-L2 biases are, in general, considered to be constant over several hours. The phase ambiguity offsets are, by definition, constant over satellite passes. Hence, by forming the time average $a\{\}$ of $\Theta^I$ over a satellite passage, we can form a new variable with the mean value removed, containing only contributions from the ionosphere, $I$, and multipath and measurement noise, $\mu$, i.e.,

$$\Theta^I = \Theta^p - a\{\Theta^I\} \quad (A5)$$

where

$$a\{\Theta^I\} = a\{I^p + n + \tau_s + r_s + \mu\} = I^p_m + n + \tau_s + r_s + \mu_m \quad (A6)$$

[32] Characterization is performed by using observations of the slant ionospheric delay from the three sites in a GPS network forming a triangle in order to interpolate the expected ionospheric delay for a site in the middle of the
triangle. Hence, for each site in a triangle, we form the L4 (L1-L2) combination for a specific satellite observation. We use the L4 combination from the three surrounding sites in the triangle to interpolate an expected L4 value at the inner site. The interpolation is performed through bilinear interpolation in the north and east components of plane coordinates. This is essentially the same as fitting a plane surface to the values of the surrounding triangle and using that information to determine the value of the inner site. We compute the difference between the interpolated and measured L4 values. Thus

\[ \Delta \Theta' = \Theta' - \Theta'_{0} \]  \hspace{1cm} (A7)

where

\[ \Theta'_{0} = \sum_{n=1}^{3} a_{n} \Theta'_{n}, \quad \sum_{n=1}^{3} a_{n} = 1 \]  \hspace{1cm} (A8)

\[ a_{n} \] are the weights uniquely given by the bilinearity of the interpolation.\[ a_{1}, a_{2}, a_{3} \] refers to the three surrounding sites.\[ a_{1} = a_{2} = a_{3} = 1/3. \]

We now map the L4 deviations to zenith

\[ \Delta \Theta = \frac{1}{m_{I}(\varepsilon)} \Delta \Theta' \]  \hspace{1cm} (A9)

where \( m_{I}(\varepsilon) \) is the mapping function

\[ m_{I}(\varepsilon) = \frac{R + h}{\sqrt{(R + h)^{2} - R^{2} \cos^{2} \varepsilon}} \]  \hspace{1cm} (A10)

\( \varepsilon \) is the elevation angle of the observation; \( R \) is the radius of the Earth; and \( h \) is the height of the ionosphere, here at 400 km, represented as a thin shell. The interpolation operation and the mapping to zenith result in a \( \Delta \Theta \) containing the sought ionospheric electron content deviation \( \Delta I \).

We now use all available \( \Delta \Theta \) values from all visible satellites during 1 h to form hourly root-mean-square (RMS) values. We use only cycle-slip-free satellite passes that are longer than 1 h and with the satellite observation above 20° in elevation. The duration of the passes we use is typically a couple of hours. Thus

\[ \sigma_{\Delta I} = \sqrt{\frac{\sum (\Delta \Theta)^{2}}{N}} \]  \hspace{1cm} (A11)

[36] This RMS value contains contributions from the ionosphere, \( I \), and multipath and measurement noise, \( \mu \). Since we can assume these contributions to be uncorrelated, we can write

\[ \sigma_{\Delta I}^{2} = \sigma_{\Delta I}^{2} - \sigma_{\mu}^{2} \]  \hspace{1cm} (A12)

[37] We now want to remove the expected contribution from multipath and measurement noise. We use a model presented in Emardson et al. [2010] for frequency, \( \omega \), i.e.,

\[ \sigma_{\mu, \omega}(\varepsilon) = \frac{C_{\mu, \omega}}{\sin \varepsilon} \]  \hspace{1cm} (A13)

with \( C_{\mu,1} = 1.2 \) mm and \( C_{\mu,2} = 1.5 \) mm. To find the multipath contribution, we start with the L4 variance as the sum of L1 and L2 variances, rescale to zenith using the mapping function in (A10) and converted to TEC by multiplying with \( \kappa \). For a representative set of \( K \) elevation angles, \( \varepsilon_{k} \) [Jarlemark et al., 2010], we get

\[ \sigma_{\mu}^{2} = C_{\text{net}} \sum_{k} \kappa^{2} \left[ \sigma_{\mu,1}^{2}(\varepsilon_{k}) + \sigma_{\mu,2}^{2}(\varepsilon_{k}) \right] m_{I}^{2}(\varepsilon_{k}) \]  \hspace{1cm} (A14)

[38] \( C_{\text{net}} = 1 + 3/9 \) account for the four stations in the interpolation (A7) and (A8) with uncorrelated multipath contributions. The central station has weight 1, while we have assumed that the three surrounding stations have weights \( a_{1} = a_{2} = a_{3} = 1/3 \).

[39] Using these values, we obtain

\[ \sigma_{\mu} = 0.036 \text{ TECU} \]  \hspace{1cm} (A15)

[40] Hence, from (A12), we obtain the sought ionospheric spatial variability \( \sigma_{\Delta I} \).

[41] For all five networks studied, the distances between the reference sites are approximately 300 km. We have scaled the results to networks with a mean reference station distance of 300 km using a scaling that is linear with distance, i.e.,

\[ \sigma_{\Delta I,b} = \frac{\sigma_{\Delta I,a}}{d_{b}} \]  \hspace{1cm} (A16)

where \( \sigma_{\Delta I,a} \) and \( \sigma_{\Delta I,b} \) are TEC variabilities from networks with mean distances \( d_{a} \) and \( d_{b} \), respectively. This linear scaling is based on studies where we use different sizes of triangles for interpolation [Emardson et al., 2011].

[42] In order to quantify the spatial filtering effect of our method to obtain TEC variability measures, we performed simulations. We propagated simulated time series of TEC variations through the interpolation scheme described above, separately for the different geographical networks. The TEC variations were of different spatial scales. We found that our method to determine the spatial variability in the ionosphere acts as a spatial high-pass filter on the true electron density variations. Figure A1 shows the spectral response as a function of wavelength for this filter. The wavelengths of the simulated ionospheric variations are normalized to the mean.
reference station distance. The presented curve represents the mean over all the networks and propagation directions. The cutoff frequency is approximately 1.5 which, for our sizes of networks, i.e., some 300 km between the reference stations, is just below 500 km. Hence, the results presented in this paper describe the electron density variability containing wavelengths below 500 km.

[43] Acknowledgments. The work reported in this report has been supported under a contract from the European Space Agency in the frame of the European GNSS Evolutions Programme. The views presented in this report represent solely the opinion of the authors and should be considered as R&D results not necessarily impacting the present EGNOS and Galileo system design. Support was also provided by the Swedish National Land Survey through the Close II project. We also thank reviewers Massimo Materassi and Simon Banville for their comments and corrections.

References


