ELSEVIER

Regularized estimation of TEC from GPS data for certain midlatitude stations and comparison with the IRI model

Feza Arikan [a,*], Orhan Arikan [b], Cemil B. Erol [c]

[a] *Hacettepe University, Department of Electrical and Electronics Engineering, Beytepe, Ankara 06800, Turkey*
[b] *Bilkent University, Department of Electrical and Electronics Engineering, Bilkent, Ankara 06800, Turkey*
[c] *TÜBİTAK - IUEKAE, Kavaklidere, Ankara 06100, Turkey*

Abstract

Regularized estimation of Total Electron Content (Reg-Est) is a novel technique which can combine signals from all the satellites for a given instant and given station and estimate the vertical TEC (VTEC) values for any desired period without missing any important features in the temporal or spatial domain. The preprocessed signals from all the satellites that are received for a certain time period are weighted according to their positions with respect to the local zenith. A two step regularization algorithm combines these signals and provides smooth VTEC estimates for the desired time period which can be as short as half an hour or as long as 24 h. The estimation algorithm is tried on VTEC values obtained from six midlatitude stations for the quiet and disturbed days of October, 2003. Within this period, the same estimation parameter set is used for all stations and time periods. When the regularized estimation results are compared with those from IRI-2001, JPL, CODE, UPC and ESA, best accordance is observed with JPL, UPC and CODE estimates. IRI computations usually provide a better fit for the night values. It is observed that the results from the regularized estimation algorithm are highly accurate in detecting disturbances and irregularities for various time scales and stations.
© 2007 COSPAR. Published by Elsevier Ltd. All rights reserved.

Keywords: Ionosphere; Total Electron Content (TEC); GPS; International Reference Ionosphere (IRI)

1. Introduction

Total Electron Content (TEC), which is defined as the number of free electrons along the ray path above one square meter, is a direct way to investigate the structure of ionosphere and upper atmosphere. In recent years, Global Positioning System (GPS) dual frequency signals are widely used to estimate both regional and global TEC values (Davies and Hartmann, 1997; Hocke and Pavelyev, 2001). TEC values can be estimated from GPS signals using the pseudoranges, phase shifts or a combination of these values recorded at the Earth based receivers. Regularized estimation of TEC (Reg-Est) is a novel technique which

can combine signals from all the satellites for a given instant and given station and estimate the vertical TEC (VTEC) values for any desired period without missing any important feature in temporal or spatial domain (Arikan et al., 2003, 2004). The raw GPS signals for a desired station can be downloaded from the internet site of International GPS Service (IGS) for Geodynamics (http://igs.ens.ign.fr). The downloaded files may include measurement errors or missing data values due to instrumental/hardware problems in either the receiver or the satellite. Thus, the data files are checked for measurement errors and those entries which have error in them are eliminated from the data set. The length of the consecutive missing data over all the satellites is also checked and processed according to the data length as discussed in Section 2. These preprocessed signals from all the satellites that are received for a certain time period are weighted according

* Corresponding author.
 E-mail addresses: arikan@hacettepe.edu.tr (F. Arikan), oarikan@ee.bilkent.edu.tr (O. Arikan), cemil.erol@iltaren.tubitak.gov.tr (C.B. Erol).

to their positions with respect to the local zenith. A two step regularization algorithm combines these signals and provides smooth VTEC estimates for the desired time period which can be as short as half an hour or as long as 24 h. The first step of the regularization includes the minimization of error utilizing a high pass penalty function. This step requires the determination of two regularization parameters which are to be chosen from the error function formed by the L_2 norm of the difference between the estimated and actual VTEC values. An optional weighting function is an added feature to the regularization algorithm which helps to reduce the error in calculation of VTEC due to the inhomogeneous structure of ionosphere in elevation. The second step of regularization includes a sliding window median filter which further reduces the undesired features in the estimated VTEC values. The window length of the sliding median filter is another parameter to be determined.

In (Arikan et al., 2003, 2004; Erol et al., 2002a), and (Erol et al., 2002b), the regularized estimation algorithm is tried on VTEC values obtained from four stations namely Kiruna, Sweden (67.32°N, 20.09°E), Kiev, Ukraine (50.22°N, 30.30°E), Ankara, Turkey (39.53°N, 32.45°E) and M. Dragot, Israel (31.35°N, 35.23°E) for the solar maximum week of April 23–April 28, 2001. The parameter set that is required by the estimation algorithm is chosen optimally using appropriate error functions. For this data set, the chosen robust and optimum parameters can be used for all latitudes and for both quiet and disturbed days for a minimum of a 1 h time period. It is observed that the estimated TEC values are in good accordance with the TEC estimates for the 24 h period. Due to its 30 s time resolution, the regularized VTEC estimates from the developed algorithm are very successful in representation and tracking of sudden temporal variations of the ionosphere, especially for high latitudes and during ionospheric disturbances.

In this study, the regularized estimation technique is applied to the GPS measurements to obtain VTEC values for eight midlatitude stations, namely, Zelenchukskaya, Russia (43.17°N, 41.33°E), Trabzon, Turkey (40.59°N, 39.46°E), Ankara, Turkey (39.53°N, 32.45°E), Istanbul, Turkey (41.06°N, 29.01°E), Gebze, Turkey (40.47°N, 29.27°E) Ohrid, Macedonia (41.07°N, 20.47°E), Sofia, Bulgaria (42.33°N, 23.23°E) and Nicosia, Cyprus (35.08°N, 33.23°E), for the quiet and disturbed days of October, 2003. The optimum parameter set for the estimation algorithm is determined over all the stations and the investigated days of October 2003. It is observed that the parameter set determined in the first group of stations and for the solar maximum week of April 2001 is the same as the estimation parameters determined for the eight stations and investigated days of October 2003. The robustness of the parameter set is a very important indicator of the reliability of the regularization applied in VTEC estimation. When the regularized estimation results are compared with those from IRI-2001, JPL, CODE, UPC and ESA, very good accordance is observed with JPL, CODE

and UPC estimates. IRI computations usually provide a better fit for the night values and for days with high sunspot number. It is observed that the results from the regularized estimation algorithm are highly accurate in detecting disturbances and irregularities for various time scales and stations.

In Section 2, the preprocessing of GPS data and the estimation of the VTEC values from various satellites for any desired time period are discussed. The procedure to choose the parameter set and examples of the estimated TEC values are provided in Section 3.

2. Brief review of regularized estimation of VTEC

This section contains a brief review of the regularized estimation (Reg-Est) method developed in (Arikan et al., 2003, 2004). With Reg-Est, robust and reliable estimates of the total electron content in the zenith direction can be obtained for any desired time duration within the 24 h period for any given GPS receiver station. The algorithm combines the computed VTEC data from all available satellites within the desired time frame.

The input data to the Reg-Est method consists of the computed VTEC values for each satellite which falls above the local horizon limit of 10°. Thus, the data from all the satellites whose local elevation angle is between 10° and 90° at a given time instant are included in the regularized estimation of VTEC. The raw GPS signals for each station and for all satellites are downloaded from the internet site of IGS. The receivers at GPS stations record signals transmitted at two L-band frequencies namely, f_1 at 1575.42 MHz, and f_2 at 1227.60 MHz. The time delay of the two signals are converted to pseudoranges and recorded as P1 and P2 signals. The carrier phase delay measurements on the two L-band signals are also recorded as L1 and L2, respectively. The TEC values can be calculated from the difference between P2 and P1 signals which is called the absolute TEC; the difference between L1 and L2 can be used to compute TEC which is referred to as relative TEC; and it is possible to compute TEC by using both (L1–L2) and (P2–P1) measurements and in some studies the instrumental biases are solved together with TEC estimation using a combination of (P2–P1) and (L1–L2) measurement. Any of the above mentioned methods can be used to compute TEC and they can be used as inputs to the estimation algorithm. The TEC computation methods and their advantages and disadvantages are widely discussed in the literature.

As discussed in (Arikan et al., 2003) and (Arikan et al., 2004) in detail, the input TEC values to the Reg-Est algorithm for this study are chosen as absolute TEC. Absolute TEC provides computationally simple and unambiguous values but they are usually corrupted by noise and multipath signals especially at low elevation angles. It is observed that having a noisier input does not effect the success of Reg-Est due to the fact that Reg-Est provides a robust best linear fit for a wide dynamic range of data from

all the satellites in view at a given time. The first step is to obtain the absolute total electron content on the slant ray path (STEC) from the satellite to the receiver for all satellites for a given station. Usually, the computed slant TEC is projected to the local zenith direction to obtain the vertical TEC (VTEC) through a mapping function, assuming a thin shell model of the ionosphere. The vertical TEC values are then computed for all receiver locations, all satellites over the 10° elevation angle limit within a 30 s period. The satellite and receiver bias pairs, that are obtained from Center for Orbit Determination in Europe (CODE), University of Berne, Switzerland (ftp.unibe.ch/aiub/CODE/2001) are also added to the computed VTEC values. Once VTEC values for all locations, for all satellites and for the desired time period are prepared, a weighting function is utilized to include all the contributions from satellites at an elevation angle higher than 60°. This weighting function also scales down the data from other satellites between 10° and 60° with a Gaussian function. As discussed in detail in (Arikan et al., 2003), with the weighting function, the effect of multipath is reduced, and at the same time, the contribution from the satellites at lower elevation angles are included in the estimation.

Reg-Est assumes an input vector of computed VTEC as

$$\mathbf{x}_m = [x_m(0) \ldots x_m(n) \ldots x_m(N-1)]_N^\mathrm{T} \tag{1}$$

where m denotes the satellite number and ranges from 1 to M (total number of satellites); and N is the total number of samples in the desired time period where $0 \leqslant n \leqslant N-1$. For a 24 h period, with a measurement of every 30 s, $N = 2 \times 60 \times 24 = 2880$ samples. To obtain estimates of TEC which minimize the error between the computed VTEC and the estimated VTEC in the least square sense for the desired time period, a cost function is defined in (Arikan et al., 2004), which includes the L_2 norm error between the estimated and computed VTEC values summed over all satellites plus a high pass penalty filter multiplied by a regularization parameter, μ, as follows:

$$
\begin{aligned}
J_{\mu,k_c}(\mathbf{x}) =& \sum_{m=1}^{M} (\mathbf{x} - \mathbf{x}_m)^\mathrm{T} \mathbf{W}_m (\mathbf{x} - \mathbf{x}_m) \\
& + \mu(\mathbf{x} - a\mathbf{t})^\mathrm{T} \mathbf{H}(k_c)(\mathbf{x} - a\mathbf{t})
\end{aligned} \tag{2}
$$

where $\mathbf{W}_m = \mathrm{diag}(\mathbf{w}_m)$, and $\mathbf{w}_m$ is the weighting according to the elevation angle of the satellite defined in (Arikan et al., 2003). In the above equation, the trend in the data is extracted using a line whose slope is a and $\mathbf{t}$ represents the time vector for the desired sector. As discussed in detail in (Arikan et al., 2003) and (Arikan et al., 2004), $\mathbf{H}(k_c)$ is the high pass penalty function. $\mathbf{H}(k_c)$ is designed to pass all signals up to a cutoff frequency, k_c. The purpose of adding a penalty function to the cost function is to control the smoothness of the data. In order to represent the 24-h cycle of the ionosphere, the high pass penalty function, $\mathbf{H}(k_c)$ was designed as a Toeplitz matrix. For a given parameter set μ, k_c and a, we investigate the values $\mathbf{x}$ which will min-

imize J. Thus, we search for $\mathbf{x}$ which will satisfy $\nabla_{\mathbf{x}} J_{\mu,k_c}(\mathbf{x}) = \mathbf{0}$ and $\partial J/\partial a = 0$.

The optimal solution to the minimization of the cost function in Eq. (2) reduces to the solution of the following linear system

$$\mathbf{A}(\mu, k_c) \begin{bmatrix} \mathbf{x} \\ a \end{bmatrix} = \mathbf{b} \tag{3}$$

where

$$\mathbf{A}(\mu, k_c) = \begin{bmatrix} \sum_{m=1}^{M} \mathbf{W}_m + \mu\mathbf{H}(k_c) & -\mu\mathbf{H}(k_c)\mathbf{t} \\ \mathbf{t}^\mathrm{T}\mathbf{H}(k_c) & -\mathbf{t}^\mathrm{T}\mathbf{H}(k_c)\mathbf{t} \end{bmatrix} \tag{4}$$

and

$$\mathbf{b} = \begin{bmatrix} \sum_{m=1}^{M} \mathbf{W}_m\mathbf{x}_m \\ 0 \end{bmatrix}. \tag{5}$$

Then, the estimate of $\mathbf{x}$, $\tilde{\mathbf{x}}$, is obtained from the above relations as

$$\begin{bmatrix} \tilde{\mathbf{x}}(\mu, k_c) \\ a \end{bmatrix} = \mathbf{A}^{-1}(\mu, k_c)\mathbf{b}. \tag{6}$$

When 24 h period estimates are required, the expression $a\mathbf{t}$ in Eq. (2) is set to zero and the cost function $J_{\mu,k_c}(\mathbf{x})$ is minimized over $\mathbf{x}$ only, by setting $\nabla_{\mathbf{x}} J_{\mu,k_c}(\mathbf{x})$ to $\mathbf{0}$. For this case, $\mathbf{A}(\mu, k_c) = \sum_{m=1}^{M} \mathbf{W}_m + \mu\mathbf{H}(k_c)$, and the estimate of $\mathbf{x}$, $\tilde{\mathbf{x}}(\mu, k_c)$, is obtained by computing $\mathbf{A}^{-1}(\mu, k_c)\mathbf{b}$.

In order to determine the optimum choice for μ and k_c, an error function which will minimize the error between the estimated VTEC, $\tilde{\mathbf{x}}$, and the data from each satellite, $\mathbf{x}_m$, can be defined as follows:

$$e(\mu, k_c) = \sum_{m=1}^{M} \|\mathbf{W}_m(\tilde{\mathbf{x}} - \mathbf{x}_m)\|^2 \tag{7}$$

where $\|\cdot\|$ denotes the L_2 norm.

An optional sliding window median filter can be applied to further reduce the jagged features and irregularities in the estimated VTEC values. For the optimum choice of the median filter length, another error function which computes the square of the L_2 norm between the estimated VTEC and the median filtered VTEC with the filter length N_f as follows:

$$e_f(N_f) = \|\tilde{\mathbf{x}} - \tilde{\mathbf{x}}_{N_f}\|^2 \tag{8}$$

where $\tilde{\mathbf{x}}_{N_f}$ denotes the median filtered $\tilde{\mathbf{x}}$ for the filter length N_f. In the following section, we will demonstrate the optimum choice of μ, k_c and N_f using the above error functions by applying the Reg-Est algorithm on the data for the mid-latitude stations mentioned in Introduction for October 2003.

3. Results

The Reg-Est is applied to GPS data obtained from eight GPS receiver stations, namely, Zelenchukskaya, Trabzon, Ankara, Istanbul, Gebze, Ohrid, Sofia, and Nicosia. All

of these stations are in the same time zone. Selected data period is October 2003. In Fig. 1, Kp, Dst indices and sunspot number for the whole month are provided. As it is observed from Fig. 1, October 10 and 11 correspond to sunspot minimum quiet days. October 27–29 are described as positively disturbed days and October 30 and 31 are listed as negatively disturbed days in the web site of the Ionospheric Dispatch Centre in Europe, (IDCE) (http://www.cbk.waw.pl/rwc/idce.html). On October 10, Kp and Dst indices are around zero and after October 29, a major storm is effective where Kp index reaches extremely high values and Dst index plunges to −400 nT. Reg-Est is tried on both the quiet days and on the extreme storm days and it is observed that very high resolution and accurate VTEC estimates can be obtained without any computational complexity.

In order to estimate the TEC values for the desired time duration, the three parameters of Reg-Est have to be determined optimally, namely the regularization (or smoothness) parameter, μ; the cutoff frequency, k_c; and the sliding window filter length, N_f. The error function provided in Eq. (7) is applied for the eight stations for the month of October 2003. Figs. 2 and 3 are the examples of the error function for various values of μ and k_c, for Istanbul, Gebze, Ankara, Trabzon, Nicosia, Ohrid, Sofia and Zelenchukskaya. When Figs. 2 and 3 are examined together, it is observed that on quiet days, positively dis-

turbed days and negatively disturbed days for any station and any time period, the optimal choice of μ occurs at 0.1 where the error curve starts to increase with a significant slope for all values of the cutoff frequency k_c. Higher values of μ cause smoother TEC estimates and more regularization. Since the penalty function starts to dominate the cost function in Eq. (2), the smoother estimates with high values of μ may not follow the desired features of TEC data. For higher values of k_c, the estimates follow the local trends in the data very well, but since the cutoff frequency increases, more high frequency components are included. Thus, the estimates look more jagged and there may be undesired jumps and irregularities. In order to have a trade off, we choose the lowest possible smoothness parameter and a higher value of cutoff frequency. Thus, for all stations, for all days, and for all time frames, the optimum choice of μ is 0.1 and k_c is 8. The remedy for this choice of regularization parameters is to employ the sliding window median filter to improve the smoothness of the data as mentioned in the previous section.

The optimum value of the sliding window median filter length is determined when the output should follow all the local trends of the data closely, yet the filtered output should be free of the undesired irregularities. As an example, the error function for determining the optimum value of the sliding window median filter length in Eq. (8), is plotted against N_f in Fig. 4 for Zelenchukskaya, Istanbul and
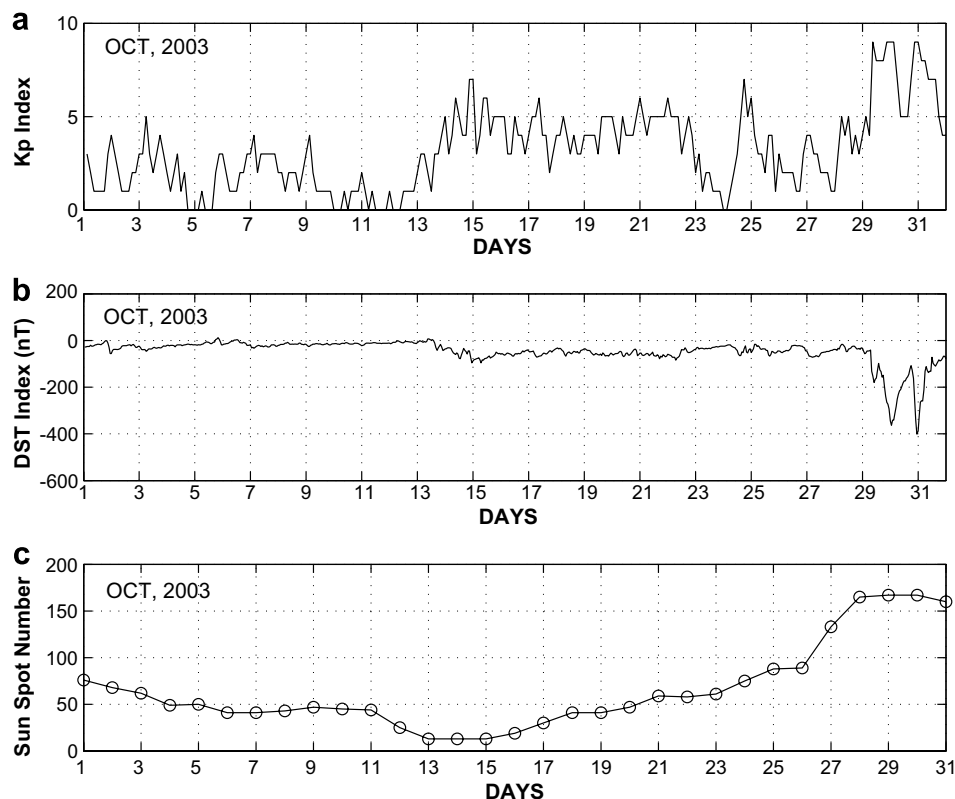

Fig. 1. (a) Kp index (http://swdcdb.kugi.kyoto-u.ac.jp/kp/); (b) Dst index (http://swdcdb.kugi.kyoto-u.ac.jp/dstdir/dst1/f/dstfinal200310.html); (c) Sunspot Numbers (ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SUNSPOT_NUMBERS/2003) for October 2003.
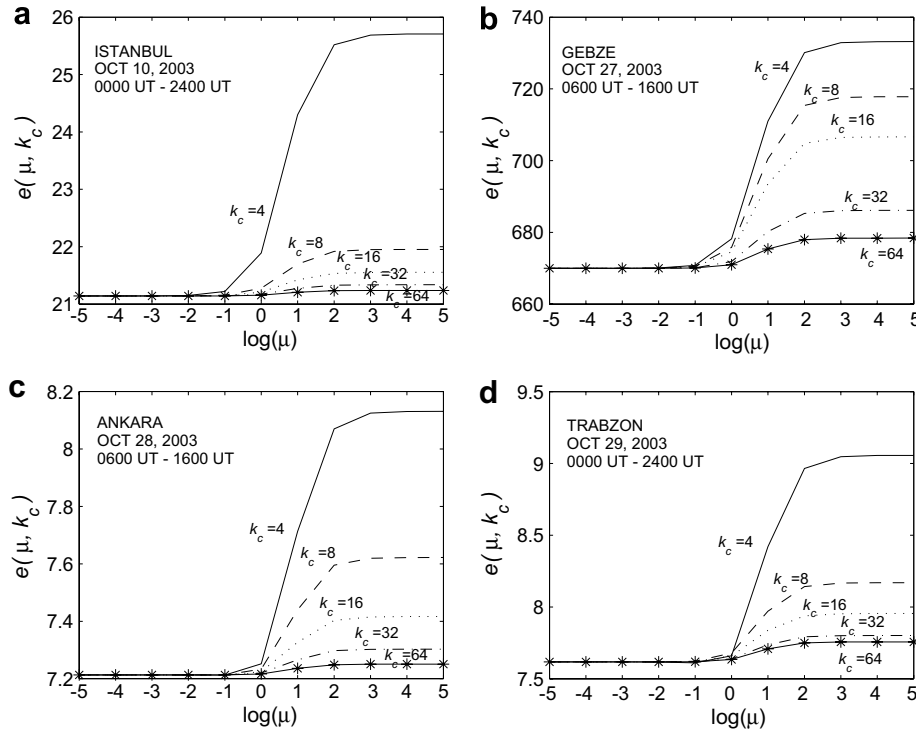
Fig. 2. Error function to determine the optimum values of the regularization parameter, μ, and the cutoff frequency, k_c for (a) Istanbul on October 10, 2003 (0000 UT-2400 UT) (b) Gebze on October 27, 2003, (0600 UT-1600 UT) (c) Ankara on October 28, 2003, (0600 UT-1600 UT) (d) Trabzon on October 29, 2003 (0000 UT-2400 UT).
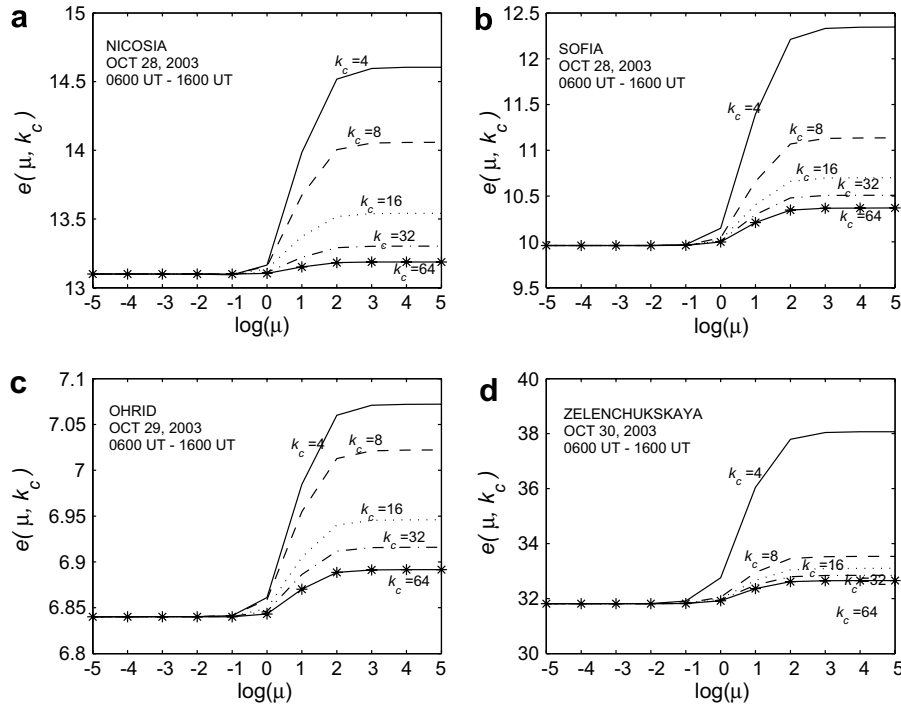

Fig. 3. Error function to determine the optimum values of the regularization parameter, μ, and the cutoff frequency, k_c for (a) Nicosia on October 28, 2003 (0600 UT-1600 UT) (b) Sofia on October 28, 2003, (0600 UT–1600 UT) (c) Ohrid on October 29, 2003, (0600 UT-1600 UT) (d) Zelenchukskaya on October 30, 2003 (0600 UT–1600 UT).

Trabzon both for positively and negatively disturbed days. The optimum choice of median filter length can be obtained at a point where the error function settles to a pla-

teau. For a whole 24 h period, the optimum value of N_f is determined to be 75 samples, corresponding to 37.5 min. This period can be extended to 85 samples or 40 min.
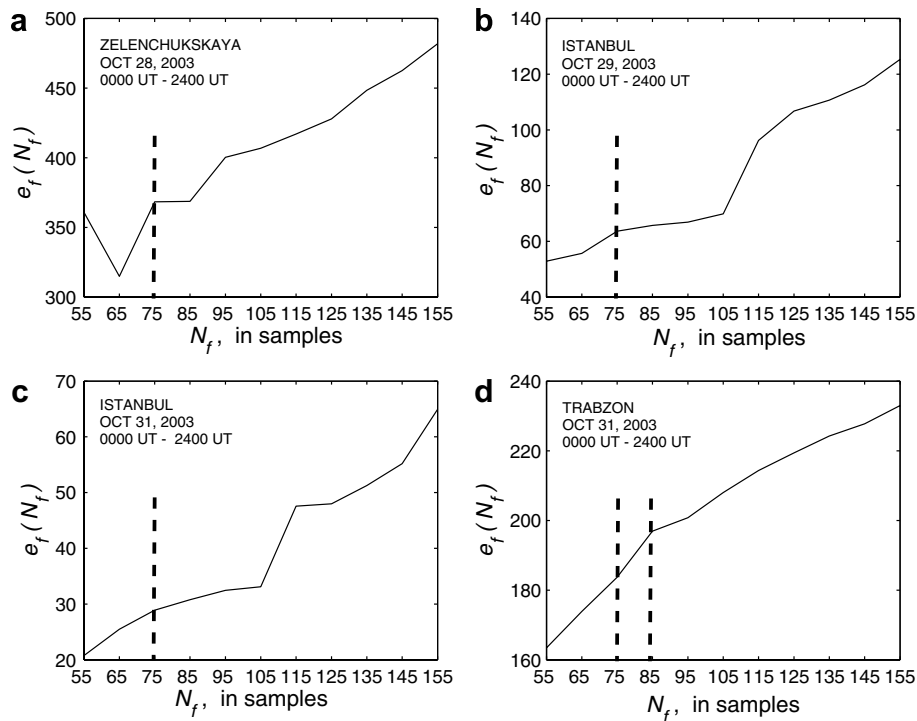
Fig. 4. Error function to determine the optimum value of the sliding window median filter length, N_f (a) Zelenchukskaya on October 28, 2003 (0000 UT-2400 UT); (b) Istanbul on October 29, 2003 (0000 UT-2400 UT); (c) Istanbul on October 31, 2003 (0000 UT-2400 UT) (d) Trabzon on October 31, 2003 (0000 UT-2400 UT).

For shorter time periods, the length of the median filter varies as discussed in (Arikan et al., 2004). For the mentioned stations, in a time period of 0600 UT-1600 UT, the optimum choice of N_f is found to be 15 samples (7.5 min) for various days in October 2003.

The above chosen parameter set is used in the Reg-Est algorithm for the estimation of VTEC. It is remarkable that the optimum values of $\mu = 0.1$ and $k_c = 8$ in this study are the same as those obtained in the previous studies in (Arikan et al., 2003) and (Arikan et al., 2004), for a different station set and another time period. This feature is an indicator of both the success of the regularization and the robustness of the estimation algorithm.

Examples of estimated and median filtered VTEC values are provided in Figs. 5 and 6, for Istanbul, Ankara, Trabzon, Zelenchukskaya, Ohrid, and Nicosia for quiet days, positively disturbed days and negatively disturbed days in October 2003. The VTEC data from the IGS centers are obtained from ftp://cddisa.gsfc.nasa.gov/gps/products/ionex for ESA (The European Space Operations Center (ESOC) of European Space Agency (ESA), Darmstadt, Germany), JPL (Jet Propulsion Laboratory (JPL) Pasadena, California, USA), UPC (Polytechnical University of Catalonia, Barcelona, Spain) and CODE (Center for Orbit Determination in Europe, University of Berne, Switzerland) models. The TEC values from the International Reference Ionosphere (IRI) model (http://nssdc.gsfc.nasa.gov/space/model/models/iri.html) are also included in Figs. 5 and 6 and denoted by plus signs. As it can be observed from Figs. 5 and 6, Reg-Est provides very high

resolution and accurate VTEC estimates with the same parameter set for any station, any disturbance level and any time period. The estimates of VTEC from Reg-Est are in general accordance with those of JPL, CODE and UPC. For high sunspot numbers and positively disturbed days, especially for night time, IRI and ESA estimates are also are of general agreement with the Reg-Est, JPL, CODE and UPC estimates. On October 30 and 31, a major storm lowered the Dst index to −400 nT, yet on those days, the sunspot numbers reached maximum. Thus, having sunspot number as major input, IRI estimated very high VTEC values for all stations whereas all the IGS centers and Reg-Est recorded very low VTEC values due to negative disturbance in the ionosphere.

4. Conclusion

A novel regularization technique Reg-Est, which combines signals from all GPS satellites within the view of a GPS receiver for a given instant is used to estimate the vertical TEC values for any desired time period within the 24 h without missing any important features in temporal or spatial domain. The algorithm is based on the minimization of a cost function including a high pass penalty filter. The developed algorithm is applied to GPS data from eight midlatitude GPS receiver stations in the same time zone for the quiet and disturbed days of October, 2003. It is observed that the Reg-Est is very successful in estimating TEC for both quite and disturbed days. The parameters, that are required by the regularization and smoothing
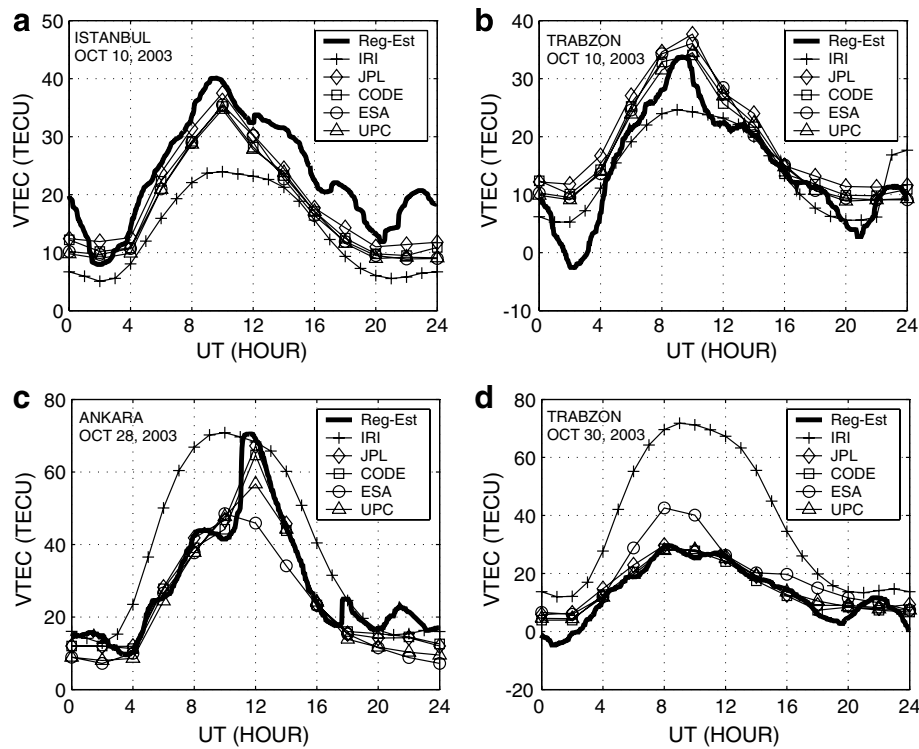
Fig. 5. Comparison of the estimated TEC values; Reg-Est (solid line); IRI model (plus signs connected with a solid line); UPC (triangles connected with a solid line); JPL (diamonds connected with a solid line); ESA (circles connected with a solid line); CODE (squares connected with a solid line) for (a) Istanbul on October 10, 2003, (b) Trabzon on October 10, 2003, (c) Ankara on October 28, 2003, (d) Trabzon on October 30, 2003.
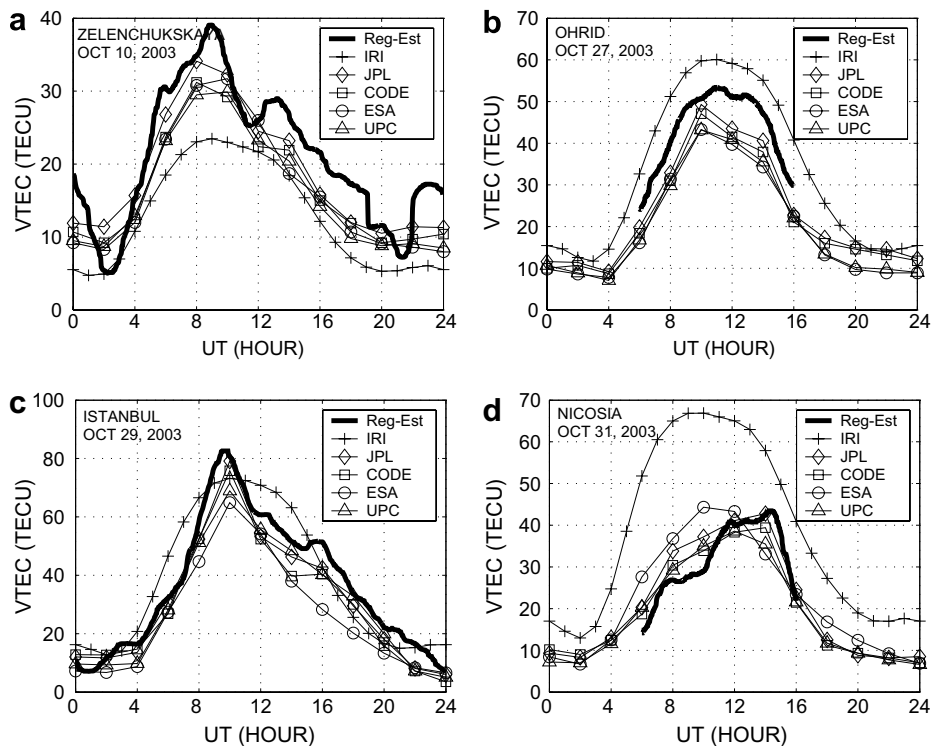

Fig. 6. Comparison of the estimated TEC values; Reg-Est (solid line); IRI model (plus signs connected with a solid line); UPC (triangles connected with a solid line); JPL (diamonds connected with a solid line); ESA (circles connected with a solid line); CODE (squares connected with a solid line) for (a) Zelenchukskaya on October 10, 2003, (b) Ohrid on October 27, 2003, (c) Istanbul on October 29, 2003, (d) Nicosia on October 31, 2003.

procedure, are determined robustly. The same parameter values can be used for all GPS station locations, for all days both quiet and disturbed. When the estimated VTEC values are compared with other available global TEC mapping model outputs and IRI, it is observed that the estimation procedure is very accurate in representing the sharp and sudden temporal variations in the ionosphere compared to the other model outputs.

Acknowledgements

References

Arikan, F., Erol, C.B., Arikan, O. Regularized estimation of vertical total electron content from Global Positioning System data. J. Geophys. Res.-Space Phys. 108 (A12), 1469–1480, 2003.

Arikan, F., Erol, C.B., Arikan, O. Regularized estimation of VTEC from GPS for a desired time. Radio Sci. 39 (6), RS6012/1-10, 2004.

Davies, K., Hartmann, G.K. Studying the ionosphere with the Global Positioning System. Radio Sci. 32 (4), 1695–1703, 1997.

Erol, C.B., Arikan, F., Arikan, O. A new technique for TEC estimation, in: 2002 IEEE International Geoscience and Remote Sensing Symposium and the 24th Canadian Symposium on Remote Sensing IGARSS'02, Toronto, Canada, June 2002a.

Erol, C.B., Arikan, F., Arikan, O. An alternate method for local vertical TEC estimation, in: Proceedings of Progress in Electromagnetics Research Symposium PIERS'2002, Cambridge MA, p. 488, July 2002b.

Hocke, K., Pavelyev, A.G. General aspect of GPS data use for atmospheric science. Adv. Space Res. 27 (6-7), 1313–1320, 2001.