

A NEW APPROACH TO OUTLYING DATA IN ESTIMATION OF VERTICAL TOTAL ELECTRON CONTENT

Ramazan Cuneyt Erenoglu

Original scientific paper

For all scientific studies, for processing analytical data the subject of outliers is of critical importance. The accuracy of Total Electron Content (TEC) models obtained from ground based Global Positioning Systems (GPS) data is strongly affected by some factors such as satellite geometry, cut-off angle, ionospheric shell height, quantity of GPS data etc. Probable outliers within the observables have a significant effect on the quality of the estimated TEC modelling. These should be removed from the TEC time series in order to get more reliable results for further process. Although some analytical methods have been developed for the code-outliers, phase-slips and ionosphere disturbances, significant reduction in the number of measurement could adversely affect the quality of the generated TEC values using geometry-free combinations of those. In this study, we proposed a novel approach for outlier detection to be used at geometry-free combinations in estimation of vertical TEC without reducing the number of code observables. Here the GPS data obtained from 22 permanent stations are used to estimate the VTEC on May 14-15, 2005. The research results from the detailed processing clearly show the efficiency of the method. Our research shows that the proposed approach appears to be a fast, effective and objective method to detect outlying effects which cause improper modelling. Moreover the resulting VTEC maps from the new approach differ reliably with respect to the data including outlying observables. Consequently, the proposed method could provide an alternative to outlier rejection in estimation of VTEC, in which outlying observations are retained in geometry-free combinations.

Keywords: discordant data; GPS; median; outlier; total electron content

Novi pristup vanjskim (outlying) podacima u procjeni vertikalnog ukupnog sadržaja elektrona

Izvorni znanstveni članak

U svim vrstama znanstvenih studija, za obradu analitičkih podataka pitanje outliera je od bitne važnosti. Na točnost Total Electron Content (TEC) modela koja se postiže na temelju podataka Sustava za Globalno Pozicioniranje (GPS) sa zemlje uvelike djeluju čimbenici kao što su geometrija satelita, kut prekida, visina ionosferske ljuske, količina GPS podataka i sl. Vjerojatni outlieri među promatranima imaju značajan učinak na kvalitetu procijenjenog TEC modeliranja. Njih je potrebno ukloniti iz TEC vremenskih serija kako bi se dobili pouzdaniji rezultati za daljnji postupak. Iako su neke analitičke metode razvijene za code-outliers, phase-slips i poremećaje u ionosferi, značajno smanjenje broja mjerenja negativno bi moglo utjecati na kvalitetu nastalih TEC vrijednosti primjenom njihovih kombinacija bez geometrije. U ovom istraživanju, predložili smo novi pristup za otkrivanje outliera koji će se koristiti u kombinacijama bez geometrije u procjeni vertikalnog TEC-a bez smanjenja broja promatrača koda. Ovdje se GPS podaci dobiveni iz 22 stalnih postaja koriste za procjenu VTEC-a na dan 14-15 svibnja 2005. Rezultati istraživanja dobiveni detaljnom obradom jasno pokazuju učinkovitost metode. Naše istraživanje pokazuje da se predloženi pristup pokazao kao brza, učinkovita i objektivna metoda za otkrivanje učinaka outliera koji dovode do nepravilnog modeliranja. Osim toga, VTEC mape dobivene novim pristupom pouzdano se razlikuju u odnosu na podatke, uključujući vanjske promatrače (outlying observables). Slijedom toga, predloženi postupak može biti alternativa odbacivanju outliera u procjeni VTEC-a, u kojoj su vanjska (outlying) zapažanja zadržana u kombinacijama bez geometrije.

Ključne riječi: GPS; medijan; outlier; suprotni podaci; ukupni sadržaj elektrona

1 Introduction

A satellite navigation system with global coverage may be termed a global positioning system (GPS). The GPS is used primarily to get the position coordinates. However, the accuracy requirements of the positioning strongly depend on the desired aim of application. For higher accuracy, it is of significant importance to characterize, reduce or eliminate possible error sources such as orbital errors, satellite and receiver clock errors, tropospheric and ionospheric effects and multipath propagation etc. For example, orbital errors can be mostly eliminated using high-precise achievable with International GPS Service for Geodynamics (IGS) products. Moreover, by using the IGS products one ensures the highest possible consistency within the current International Terrestrial Reference Frame (ITRF) [1]. Clock errors can be removed either by double differencing (DD) or by explicit estimation of receiver and transmitter clock behaviour. One of important error sources is multipath. This is because pseudorange which are subjected to greater multipath dominate the single-frequency ionosphere-free combination. The multipath can be mitigated by using different techniques, such as choosing sites without multipath reflectors, using choke-

ring antennas, a multipath elimination algorithm at the receiver signal processing level [2], a multipath mapping technique surround the GPS antenna [3], a concept of the analysing of the Signal-to-Noise-Ratio (SNR) [4], an approach of multipath template for mitigating multipath [5], a finite impulse response (FIR) filters to extract or eliminate multipath [6], an effective technique based on the use of an adaptive filter [7].

The ionosphere is portion of the Earth's atmosphere and located about between 60 km and 1000 km upon the Earth's surface. Since the ionosphere is ionized by intense solar radiation, principally in the ultraviolet spectrum, the GPS refraction remains a major error source in GPS positioning, especially for real-time applications. GPS microwave signal disruption during penetration of the ionosphere is one of the most significant causes of rectifiable GPS error [8]. The frequency of electromagnetic wave and the free electron density of the ionosphere give an opportunity to estimate the amplitude of the ionospheric effect. The free electron density is quantified by Total Electron Content (TEC) that is the column electron number with a cross section of 1 m^2 along a certain path. The ionosphere is both heterogeneous in structure and non-constant in time, especially being affected by ionospheric storms,

geomagnetic activities and solar flares etc. [9, 10, 11]. The disturbing effect on the electromagnetic signal is mostly in the magnitude from a few meters to tens of meters in the normal solar activity conditions. However, under heavy ionosphere storm, this influence could reach more than 100 meters. Especially, the ionosphere effect has become the biggest error source in GPS positioning when the disable of the Selective Availability (SA). Note that initially the highest quality signal was reserved for military use, and the signal available for civilian use was intentionally degraded. The disadvantage for civilian users was changed with the order of Selective Availability to be turned out at midnight May 1, 2000, improving the precision of civilian GPS from 100 meters to 20 meters. All in all, ionospheric errors are now dangerous troubles for precise positioning.

For high precision GPS positioning, the ionospheric effect must be modelled so that a correction can be made to remove it from the GPS observables. The estimation of the ionospheric effect is significant for all kinds of earth observation applications and space weather studies. There are many currently available ionosphere models in order to define the TEC under various space weather conditions [12, 13, 14, 15, 16, 17, 18, 19]. The dispersive and heterogeneous nature of the ionosphere provides the opportunity to determine the TEC distribution by using dual frequency GPS observables. While some studies were focused on only single GPS station observables [20], the ionosphere models have fitted to global and regional GPS networks in recent years due to the global and regional reasons [17, 18, 19].

The amount of ionospheric delay changes depending on the functions of TEC values in the ionosphere at different levels of solar activity [18]. In order to model the ionosphere variations, GPS observations by dual frequency GPS receivers at sites of a global / regional network are processed with IGS final products. Later, regional ionosphere models are computed and TEC grid mapping is performed [21, 22].

Determining TEC from dual-frequency GPS observations has become an important application of GPS. However there are some varieties of error sources which may affect the estimation of TEC. For example, these are the effect of the differential satellite and receiver instrumental delay biases. These biases are normally estimated simultaneously with TEC. However, the additional estimation of the instrumental biases may constitute an insurmountable burden in some practical applications like real-time estimation of TEC, or the estimation may be difficult or correlated to the ionospheric parameters, particularly in situations where the TEC behaviour may be harder to model. One of the major error sources in GPS positioning is ionospheric refraction which causes signal propagation delays [23]. Ionospheric refraction depends on the Total Electron Content (TEC) in the signal path. TEC can be modelled from the GPS observations. Variations in the TEC depend on the local time of the day, seasons, solar cycle, geographic location of the receiver and earth's magnetic field [20, 24]. TEC values can be converted to Vertical Total Electron Content (VTEC) using different mapping function.

The VTEC values from all satellites and stations were used to obtain VTEC maps. These can be obtained using geometry-free linear combination with differential code bias (DCB) information at single layer [21, 22, 25, 26]. Although the carrier-phase smoothed L_1 and L_2 code observations are simply differenced in the geometry-free linear combination, a probable outlying data that is close to the threshold may arise and reflect on TEC estimation. The calculated VTEC values may also therefore include discordant points at some epochs. If there are some VTEC observations that do not depend on the same data set, the data can be respected as discordant points. Ionospheric mapping requires interpolation (e.g., the methodology referred to as Kriging) of the parameters between the sites. As well known, the Least Squares Estimation (LSE) is extensively used for the solution of Kriging method. The LSE is sensitive to discordant values, which if included in the observation data may distort the estimation of the parameters. In other words, these values can degrade interpolation quality for producing two dimensional VTEC maps. Also in the step of the TEC mapping, one should pay more attention to outlying points that render the estimation of unknowns meaningless. The outlying data should be therefore excluded from the data set. Most of them are based on the LSE method, and so they do not always produce reliable solutions [27, 28, 29], although there exist some methods for analysing the quality of the TEC estimation, see [30]. For removing the effect of outlying data, a novel approach which reflects the quality of the data, can be utilized [31]. In this study, an algorithm has been improved in order to eliminate disturbing effects of the bad data in VTEC maps. To do it, the analysing of VTEC data with a new approach is carried out using epoch by epoch basis for the selected satellite-station combination. We thereby aim to inquire not only the results from the VTEC estimation, but also to robustify the TEC mapping. GPS data obtained from Scripps Orbit and Permanent Array Center (SOPAC) is used to test the proposed approach, and the resulting TEC maps are compared to the ones from the usual method for revealing the benefit of our approach.

This article introduces at first the concept of the total electron content. Secondly, it discusses the strategy on how to estimate total electron content using GPS data. It follows the introduction of proposed method for more reliable regional TEC mapping against outlying data and how the progress in the TEC computation can be performed. These mathematical and theoretical aspects are then applied to the real data from continuously observing GPS stations. In these numerical examples with/without outlying data, it is shown how the proposed approach works and how it is more superior accordingly not applied. Finally, the results are discussed and the conclusions recommended.

2 Computing the total electron content using GPS observables

At a reference station, a dual frequency GPS receiver collects both code and carrier phase observations on L_1 (1575,42 MHz) and L_2 (1227,60 MHz) frequencies. We can formulate an equation for the code observations in

both frequencies that relates the measurements and the various biases:

$$P_i = \rho + c(dt - dT) + d_{\text{orb}} + d_{\text{trop}} + d_{\text{ion}} + b_{P_i} + B_{P_i} + d_{\text{multi}/P_i} + e_{P_i}, \quad (1)$$

where P is the measured pseudorange, ρ is the geometric range between the reference station and the satellite, c is the speed of light, dt and dT the clock errors of the satellite and receiver respectively with respect to GPS time, d_{orb} is the satellite orbit error, d_{trop} is the tropospheric delay, d_{ion} is the ionospheric delay, b_{P_i} is the satellite hardware delay, B_{P_i} is the receiver hardware delay, d_{multi/P_i} is the multipath effect, e_{P_i} is the measurement noise and $i = 1, 2$ [8].

The carrier signal, which has a much higher frequency than the pseudo-random code, is more accurate than using the pseudo-random code alone. Mathematically, we can write the carrier phase observable equation on L_1 and L_2 frequencies if we convert the measured carrier phase in cycles to equivalent distance unit as:

$$\Phi_i = \rho + c(dt - dT) + \lambda_i N_i + d_{\text{orb}} + d_{\text{trop}} + d_{\text{ion}} + b_{\Phi_i} + B_{\Phi_i} + d_{\text{multi}/\Phi_i} + e_{\Phi_i}, \quad (2)$$

where Φ is the carrier phase observation, λ_i is the wavelength of GPS signal on L_i frequency (in meter), N_i is the carrier phase integer ambiguity (in cycle), b_{Φ_i} is the satellite hardware delay, B_{Φ_i} is the receiver hardware delay, d_{multi/Φ_i} is the multipath effect and e_{Φ_i} is the measurement noise ($i = 1, 2$). Note that the carrier phase observation equation is very similar to the given in Eq. (1) for the pseudorange. The main difference is the presence of the ambiguity term [8].

Practically, differencing the dual frequency code observables gives in the following ionosphere measurements:

$$P_{r,GF}^s = P_{r,L_1}^s - P_{r,L_2}^s = -0,105 \cdot TEC_r^s + (DGD_r - DGD^s), \quad (3)$$

where $P_{r,GF}^s$ is the so-called geometry-free combination of the code observations, P_{r,L_1}^s and P_{r,L_2}^s are the L_1 and L_2 P-code observations performed by receiver r on satellite s , DGD_r and DGD^s are the differential group delays for receiver r and satellite s , respectively, and TEC_r^s is the slant TEC along the path from satellite s to receiver r . Using the eq. (3), we can measure slant TEC in TEC Units (or TECU, 1 TECU=1016 electrons/m²). Before January 31, 1994, the code measurements had the precision of almost one meter due to the Anti-spoofing effect. It is clear that these observations include some biases called receiver and satellite differential group delays although they are not ambiguous [8].

Another geometry-free combination is the carrier phase observations on L_1 and L_2 frequencies as:

$$\Phi_{r,GF}^s = \Phi_{r,L_1}^s - \frac{f_{L_1}}{f_{L_2}} \Phi_{r,L_2}^s = -0,552 \cdot TEC_r^s + N_{r,GF}^s, \quad (4)$$

where f_{L_1} and f_{L_2} are the frequency of the L_1 and L_2 carrier phase, respectively, Φ_{r,L_1}^s and Φ_{r,L_2}^s are the L_1 and L_2 carrier phase observables (in cycle) made by receiver r on satellite s , $N_{r,GF}^s$ is an ambiguity of carrier phase. Phase observables have a precision better than one millimeter.

The last geometry-free combination is the code and phase observations.

$$P_{r,GF}^s - \lambda_{L_1} \Phi_{r,GF}^s = (DG_r - DG^s) - \lambda_{L_1} N_{r,GF}^s \quad (5)$$

where λ_{L_1} is the wavelength of the L_1 carrier phase, in meters. This combination technique is applied to solve the ambiguity $N_{r,GF}^s$. Then, $N_{r,GF}^s$ is put into the Eq. (4), and thereby TEC is determined before converting to vertical. It is clear that the TEC information is obtained using the precise carrier phase observations. On the other hand, the ambiguity parameter comes from the combination of code and phase observations [32]. In addition, this combination requires estimation of the receiver and satellite differential delays.

In this study, as is done in practice, the ionosphere can be modelled by means of a polynomial in latitude and local time using the Eq. (3) where the combined biases of receiver and satellite are solved [8]. In computation of the TEC using GPS observations, the most important error source is to determine differential group delay which is obtained with a precision of about a few TECU [33]. After determining these biases, the slant TEC is converted to the vertical using a mapping function [23]. The ionospheric mapping function is the particular relationship between slant and vertical total electron content. That is one of the first assumptions to consider typically when ionospheric corrections are estimated from GPS data.

3 A new method for regional VTEC mapping

By mapping GPS observables collected from ground stations, maps of ionospheric VTEC are produced in real-time. These are produced to test real-time data acquisition, mapping techniques and monitoring facilities [21, 23, 24, 33]. The VTEC mapping provides accurate ionospheric calibrations for navigation systems. Moreover these maps are used to monitor ionospheric weather and to forecast ionospheric storms that often occur responding to activities in solar wind and Earth's magnetosphere as well as thermosphere.

The ionosphere disrupts this approach since the GPS radio signal is slowed by the presence of free electrons, causing an additional time delay and hence an error in the distance to each satellite. The greater the total numbers of electrons on the signal path the greater the time delay. The

GPS system broadcasts on two frequencies and since the ionosphere is a dispersive medium the time delay for a given TEC will depend on the frequency of the signal [30]. This allows the TEC to be measured by examining the differential time delay between the two frequencies.

The objective of TEC mapping is to produce results for atmospheric applications based on Global Navigation Satellite System (GNSS) observations of the current and future GNSS (GPS, GLONASS and Galileo) in association with the ground receiver stations from the European Reference Frame (EUREF) Permanent Network among other European national networks. Atmospheric research determines the zone where GNSS signals are dense enough to get accurate and reliable representations of the ionosphere because atmospheric applications based on GNSS observations such as ionospheric tomography are dependent on the GNSS signal distribution. The resulting maps are produced in two scales: the global and regional maps. The global maps are produced at IPS by using the global ionospheric model. This derived product can be compared with the NASA JPL Global TEC Map, see the map obtained by global ionosphere model in Fig. 1. TEC regional map shows a near real-time ionospheric TEC map produced by combining GPS data with the regional ionospheric model driven by real-time observations from IPS ionosondes distributed throughout the region.

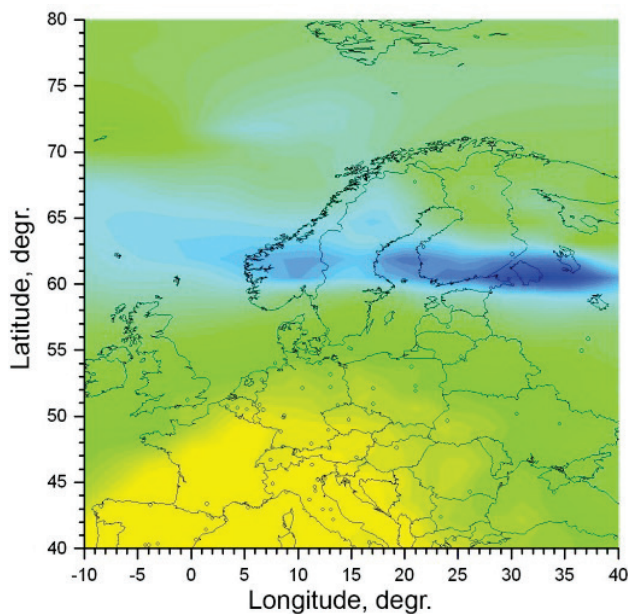


Figure 1 The map obtained by global ionosphere model over the Europe dated on April 17, 2013

For global TEC mapping, the ionosphere working group began the systematic generation of Ionosphere maps in 1990s. To do it, there are five IGS Ionosphere Associate Analysis Centers (IAACs): Center for Orbit Determination in Europe (CODE), European Space Operations Center of ESA (ESOC), Jet Propulsion Laboratory (JPL), Natural Resources Canada (NRCan/EMR) and Technical University of Catalonia (UPC).

The TEC mapping has been performed with different approaches, e.g. Nearest-Neighbour Interpolation, Linear Interpolation, Inverse Distance Weighting, Kriging etc. These methods have lack of robustness against the

contamination of the data set. The mentioned approaches are not effective in the case of different amount of data contamination [33]. Some studies and analysis on the TEC maps should be helpful for users in space weather monitoring navigation, positioning and surveying. In other words, they are very sensitive to discordant data, and produce inaccurate final products, i.e. VTEC maps. In order to deal with this trouble, it is well known in different works in the literature, i.e. various cleaning and filtering algorithms are applied to VTEC data. For example, during observation epoch at the IPP can be estimated for each satellite receiver pair. VTEC estimations are reduced to 3, by taking the median of 10 VTEC. The goal is to avoid discordances along the track, to insure a well distributed set of data (avoid overweight of satellite tracks) and to reduce the computation time. Furthermore, authors [30] developed software to produce global VTEC data automatically from huge number of stations. They proposed to detect the criterion comparing VTEC data from different receivers with respect to the others. According to them, if the discrepancy for the selected receiver is about ± 40 TECU with regard to the others, then it is assumed that it is a discordant point. Besides, the direct observations of VTEC provided by the TOPEX dual-frequency altimeter since 1990s provide a valuable source of independent estimates for GPS TEC maps.

The proposed approach is based on median estimator. Note that the data set includes VTEC solutions estimated on an epoch by epoch basis for the satellite-station combination. To detect outlying data, we used the Median Absolute Deviation (*MAD*) which is a robust measure for the variability of a univariate sample of quantitative data. It can also refer to the population parameter that is estimated by the *MAD* calculated from a sample. For a univariate data set X_1, X_2, \dots, X_n , the *MAD* is computed as the median of the absolute deviations from the data's median [29]:

$$MAD = \text{median}_i (|X_i - \text{median}_j (X_j) |) \tag{6}$$

In this study, the univariate data set includes *VTEC* values estimated using the geometry-free linear combination by simply differencing the carrier-phase smoothed L_1 and L_2 code observations. Finally, each $VTEC_i$ value is compared with the threshold value of a certain (k) times *mad*. k is here taken as 3 [27, 31]. If the absolute value of $VTEC_i$ is bigger than the threshold, it is flagged as a discordant value, and then removed from the data set.

In order to detect outlier in TEC mapping, we developed the following procedure:

- TEC estimation on epoch by epoch for satellite-station combination,
- Apply the approach for outlying points,
- Remove outlying TEC value from the data set,
- Plotting TEC maps.

To do it, Receiver INdependent EXchange (RINEX) files are here processed for code smoothing and DCB estimation by Bernese GPS software V.5. [22]. This information is imposed into geometry-free linear combination in order to estimate VTEC values. Then, in

the second part description process of flagging discordant data is given. Finally, the VTEC mapping is performed.

Using the data pre-processing process before the TEC mapping, some of the values will be marked as bad data and removed from the data set. In other words, one benefit of the proposed approach is that the process of TEC mapping is theoretically robustified using the method with high breakdown point, i.e. median estimator.

4 Case study

The performance of the classical VTEC mapping as well as of the new approach proposed in the previous section, with respect to their ability to detect discordant data, is investigated by a real data. To adapt the new approach to VTEC mapping, we downloaded the GPS data on May 14-15 2005 from Scripps Orbit and Permanent Array Center (SOPAC) web site [34]. The GPS data consists of distributed 22 permanent IGS stations all over Europe. Some of them are randomly located at the south of Central Europe, the others at the middle part of Central Europe. The most important is to obtain a data set with different statistical characteristics before using the proposed approach [35]. For this aim, the main reason for choosing May 14-15, 2005 is different K_p values (Estimated Planetary K-index) for both days [34]. Considering these values, different geomagnetic effects (activities) are already expected from quiet to active for these consecutive days. In other words, KP -index values suddenly change from May 14 to May 15. As a result, we intended to work with the TEC estimation based on a data set adversely affected by signal distortion on GPS data.

The location of the selected stations can be seen in Fig. 2. The data is in RINEX format, 30 second sampling rate and 24 hour period. Elevation mask is limited with 15 degrees in order to reduce multipath effects.

First, using Bernese GPS software V.5.0, we processed RINEX files for code smoothing and DCB estimation [22]. Then, this information is employed in geometry-free linear combination to estimate VTEC values for epoch by epoch basis for the satellite–station combinations. For example, the first 19 lines of the results

for the TEC values for 134th day are shown in Tab. 1. Respectively, the data of 8 columns representing some numerical values is as follows: Observational epoch # (Day of year), station #, satellite #, TEC value, VTEC value, L3, latitude and longitude. We studied geometry-free linear combination of the GPS data according to "station #" and "satellite #" for a systematical approach in VTEC processing as done in previous works. For instance, "station 1" stands for AQUI, "station 2" stands for BOR1, and so on.

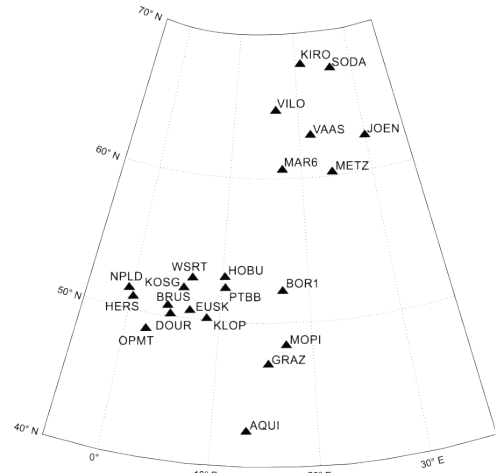


Figure 2 Geographical distributions of the stations used in the experiments

Then we plotted VTEC maps using the usual approach for the region where the GPS data are observed at the stations located. It should be noted that there is a separate map for each observation epoch. As shown in Figs. 3 and 4, we presented the regional maps for 260th and 261th, and 146th and 147th epochs (DOY 134), respectively. Herein, black dots in the figures indicate Ionospheric Pierce Points (IPP) and the maps include a legend bar which indicates VTEC values. In Fig. 3, it seems that they are appropriate to TEC maps including no-disturbing effect.

Table 1 The VTEC results from geometry-free linear combination of the data

Day of year	station #	satellite #	TEC value	VTEC value	L3	latitude	longitude
134	1*	1	16,27611234	6,62426179	0,90320088	49,98501634	4,98043025
134	1	2	22,31673453	13,52641529	1,23840963	40,92539227	19,92028159
134	1	4	16,79154452	7,94999080	0,93180346	46,41633317	21,82726724
134	1	5	13,58004822	12,88575290	0,75358976	43,24533131	14,45445401
134	1	6	20,17158128	12,78155543	1,11936988	38,65574420	9,53124948
134	1	9	19,16739689	14,98526354	1,06364526	40,04994590	15,79939399
134	1	14	19,29956414	12,25103392	1,07097954	43,39848242	7,33659373
134	1	30	14,29271617	12,03217422	0,79313743	43,24470321	10,44068871
134	2	1	10,41489980	5,46228009	0,57794801	55,09484656	-2,82032443
134	2	2	12,32399684	6,18608285	0,68388842	48,70240731	15,05723930
134	2	4	8,92168726	4,20484206	0,49508603	54,71278451	14,44427990
134	2	5	5,02477668	4,73320919	0,27883702	51,12288285	6,31886032
134	2	6	13,93910272	8,04657759	0,77351456	45,69907027	1,14531728

* "station 1" stands for AQUI, "station 2" stands for BOR1, and etc.

As seen from Fig. 4, there are some problematic effects which make the procedure of TEC mapping to fail. It is clearly due to no outlying data obtained VTEC values obtained from geometry-free linear combinations. In other

words, the Unfavourable (disturbed) VTEC maps for 146th and 147th epochs (DOY 134) are due to outlying data. If looking at the legend bar of the maps, one can see some tremendous values which may result from the

ionospheric refraction or the signal propagation delay effects, e.g. 40 or 50. Furthermore, there are mostly improper maps due to these extreme values in the process of VTEC mapping. It is clearly given as an explanation of this matter for the 146th and 147th epochs of DOY 134. The

TEC maps in Fig. 4 are therefore entirely covered in blue. Consequently, the derived TEC data should be carefully analysed using the proposed approach in order to find out outlying data for each observational epoch.

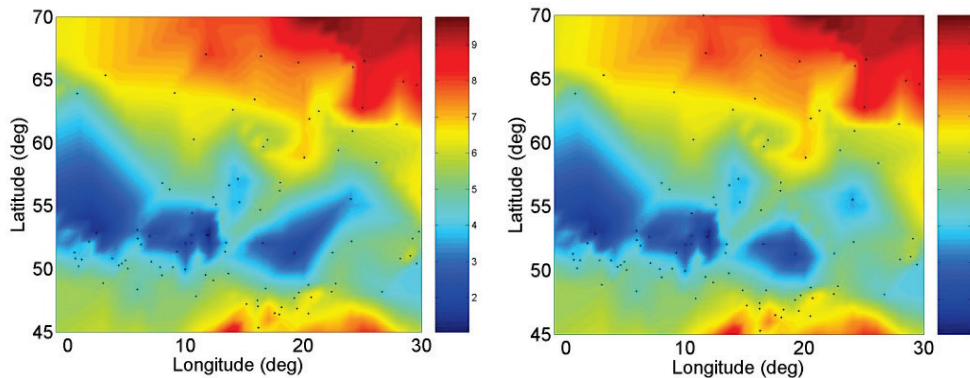


Figure 3 Favourable (good) VTEC maps from DOY 134: 260th and 261th epochs

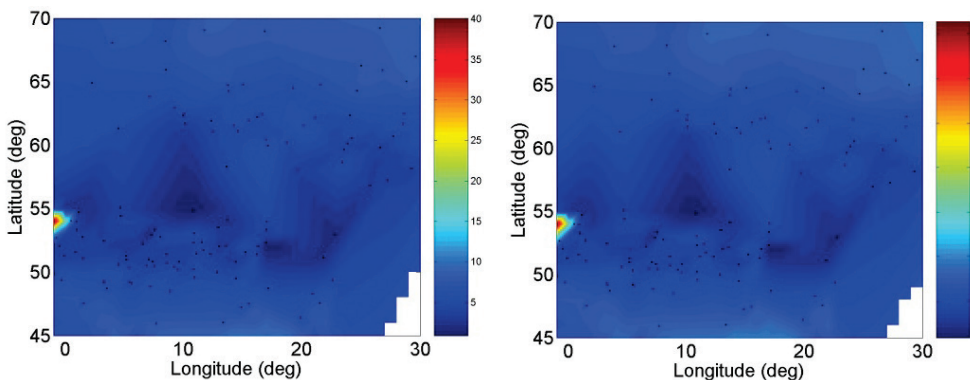


Figure 4 Unfavourable (disturbed) VTEC maps from DOY 134: 146th and 147th epochs

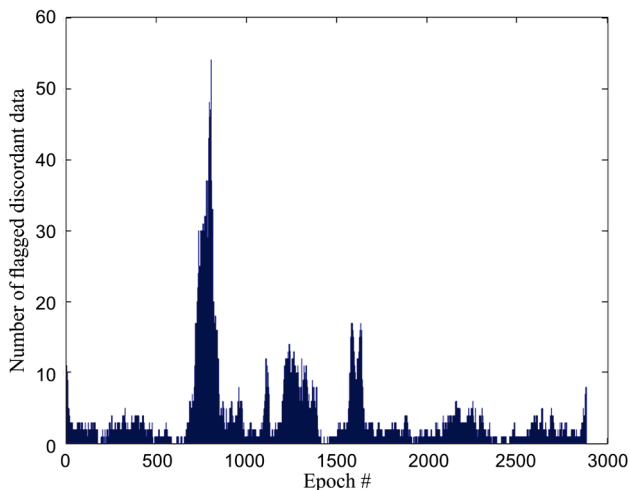


Figure 5 Number of outlying data in the experiments

For dealing with these mapping failures, we applied our procedure given in the previous sections to flag outlying data. First, we processed the 24 hourly data for DOY 134 considering the solutions of all the points in the experiments. If the evaluation strategy was made only with respect to one or a few points, the results arbitrarily could change depending on these points. In other words, the proposed method could give misleading judgement depending on the geographical position and time of some

points according to other ones. The results are shown in Fig. 5. Herein, the vertical and horizontal axes present the number of detected outlying data and epoch #, respectively. From Fig. 5, it is clear that these outlying data resulted from the Ionospheric reflection effects, especially near the 750th observation epoch. These are successfully detected using our approach, and then removed from the data set. It resulted in the maximum outlying effect around 750th observation epoch. The proposed method works well round the related interval as already expected.

We performed the process in three steps according to the rules given previously. An epoch-by-epoch analysis for TEC values is performed. The outlying ones are extracted and finally VTEC maps are re-plotted. We expected that the proposed approach must be successful for the other cases, as well as extreme number of outlying data, i.e. around 750th epoch. For example, we randomly chose 146th and 147th epochs for DOY 134. Figures 6 and 7 show the resulting TEC maps before and after the proposed approach from DOYs 134 and 135, respectively. The mean RMS errors for 10min intervals are about 0,20 TECU and 0,23 TECU for the selected days, respectively. Note that a limited number of results could be given due to the lack of space.

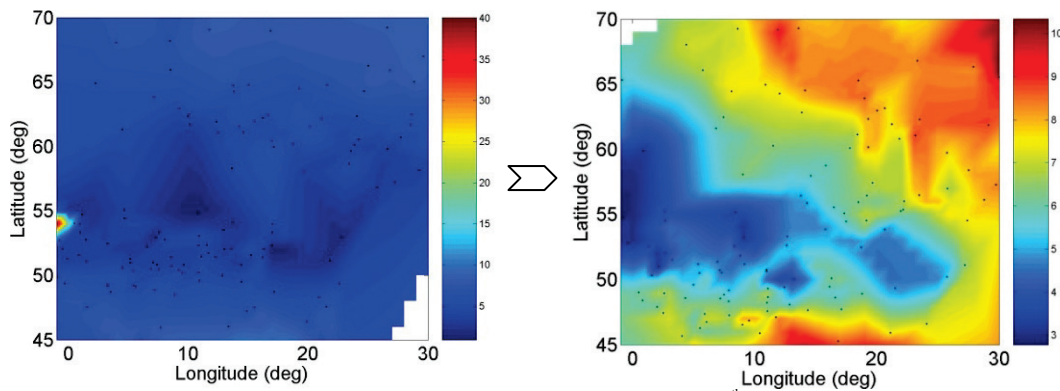


Figure 6 Before (left) and after (right) the proposed approach for 146th epoch from DOY 134

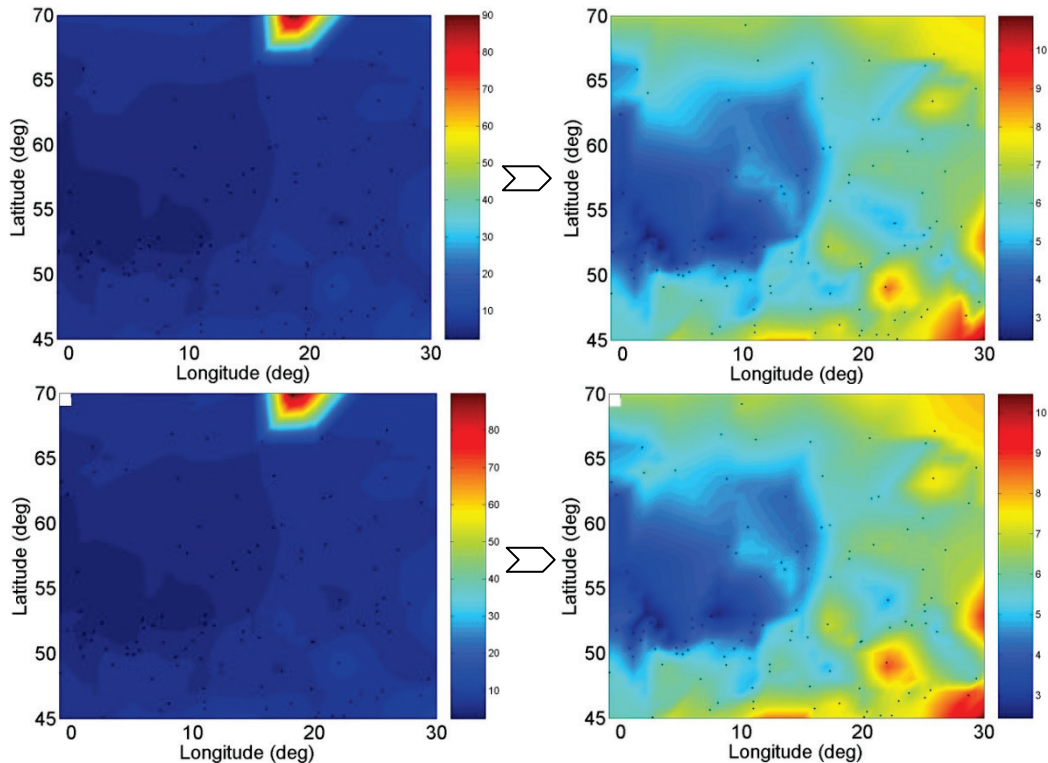


Figure 7 Before (left) and after (right) the proposed approach for 277th and 278th epochs from DOY 135

5 Conclusion

In this study, we developed the algorithm for detecting outlying data in VTEC time series before VTEC mapping. The vertical TEC is obtained that allows favourable (good) VTEC mapping without concerns about outlying data after geometry-free combinations. The input parameter for the using of the proposed method is VTEC that can be easily computed from carrier smoothed-code observations. We worked with the TEC estimation based on a data set obtained from continuous GPS stations for two consecutive days. The positive and negative aspects of the proposed approach are as follows:

- Easy to apply: All you want is to use this approach for TEC estimation; the mentioned procedure in Section 3 can be easily programmed in any technical computing language, e.g. MATLAB.
- Robust TEC estimation: The method is based on the median function, see Eq. (6). Therefore this provides robust solution against the outlying data. It is not affected by extreme values.
- Wide availability: The approach can be used for the production of regional and also global TEC mapping.
- High reliability: The mean RMS errors for 10min intervals are about 0,20 TECU for the selected days.
- Cost effective: Any one can apply the procedure to TEC estimation without any payment of any license.
- Time consuming: The evaluation process may take some time if the data set contains a large number of items.
- Limited treatment: It is not capable of further algebraic treatment.

We performed that successfully analysing the results from the VTEC estimation and also meaningful VTEC mapping without outlying data. Note that all the produced TEC maps were not given due to the lack of space. Finally, we strongly recommend the proposed approach as a tool for finding out outlying data before TEC mapping.

Acknowledgment

The authors are very grateful for the invaluable comments and suggestions of the editor and the reviewers. The GPS data were provided from SOPAC, and ISGI for the geomagnetic data used in this study. In addition, the author would like to thank Dr. N. Arslan for his suggestions to improve this study. Some computations and plots were made using MATLAB (<http://www.mathworks.com/>).

6 References

- [1] Kouba, J. A guide to using international GPS service (IGS) products. 2009. <http://igsb.jpl.nasa.gov/components/usage.html/>. (12.10.2014).
- [2] Townsend, B.; Fenton, P. A Practical Approach to the Reduction of Pseudorange Multipath Errors in a L1 GPS Receiver. // Proceedings of the 7th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS-94) / Salt Lake City, UT, 1994, pp. 143–148.
- [3] Haji, G. A. The Multipath Simulator: A Tool toward Controlling Multipath. // 2nd Symp. on GPS Applications in Space / Hanscom AFB.1990.
- [4] Comp, C.; Axelrad, P. Adaptive SNR-Based Carrier Phase Multipath Mitigation Technique. // IEEE Transactions on Aerospace and Electronic Systems. 34, 1(1997), pp. 264-276. DOI: 10.1109/7.640284
- [5] Bishop, G. J.; Bullet, T. W.; Holland, E. A. GPS Measurements of L-band Scintillation and Total Electron Content in the Northern Polar Cap Ionosphere at Solar Maximum. // Proceedings of the International Beacon Satellite Symposium Edited by Kersley, L. Univ. of Wales / Aberystwyth, U. K. 1994.
- [6] Han, S.; Rizos, C. Multipath Effects on GPS in Mine Environments. // Xth International Congress on the International Society for Mine Surveying Fremantle, Australia, 1997, pp. 447-457.
- [7] Hu, C. W.; Congwei, W.; Chen, Y. Chen; Liu, D. Adaptive Kalman Filtering for Vehicle Navigation. // Journal of Global Positioning Systems. 2, 1(2003), pp. 42-47. DOI: 10.5081/jgps.2.1.42
- [8] Teunissen, P. J. G.; Kleusberg, A. GPS for Geodesy. 2nd enlarged edition, Springer Verlag, 1998. DOI: 10.1007/978-3-642-72011-6
- [9] Klobuchar, J. A. Ionospheric effects on GPS. // GPS World. 2, 4(1991), pp. 48-51.
- [10] Mendillo, M. Storms in the Ionosphere: Patterns and Processes for Total Electron Content. // Rev. Geophys. 44, 4(2006), pp. 1-47. DOI: 10.1029/2005RG000193
- [11] Mendillo, M.; Smith, S.; Coster, A.; Erickson, P.; Baumgardner, J.; Martinis, C. Man-made Space Weather. // Space Weather. 6, 9(2008), pp. 1-7. DOI: 10.1029/2008SW000406
- [12] Komjathy, A. Global Ionospheric Total Electron Content Mapping Using the Global Positioning System. Ph.D. Dissertation, Department of Geodesy and Geomatics Engineering Technical Report No. 188, University of New Brunswick, Fredericton, New Brunswick, Canada, 1997.
- [13] Skone, S. Wide Area Ionosphere Grid Modeling in the Auroral Region. UCGE Reports Number 20123, Ph.D Dissertation, The University of Calgary, Calgary, Alberta, Canada, 1998.
- [14] Jakowski, N.; Sardon, E.; Schlüter, S. GPS Based TEC Observations in Comparison with IRI95 and the European TEC Model NTCM2. // Advances in Space Research. 22(1998), pp. 803-806. DOI: 10.1016/S0273-1177(98)00101-X
- [15] Liao, X. Carrier Phase Based Ionosphere Recovery over A Regional Area GPS Network. UCGE Reports, Number 20143, University of Calgary, Calgary, Alberta, Canada, 2000, pp.120
- [16] Fedrizzi, M.; Langley, R. B.; Komjathy, A.; Santos, M. C.; de Paula, E. R.; Kantor, I. J. The Low-Latitude Ionosphere: Monitoring its Behavior with GPS. // Proceedings of ION GPS 2001, 14th International Technical Meeting of the Satellite Division of the Institute of Navigation / Salt Lake City, UT, 2001.
- [17] Gao, Y.; Liao, X.; Liu, Z. Z. Ionosphere Modeling Using Carrier Smoothed Ionosphere Observations From A Regional GPS Network. // Geomatica. 56, 2(2002), pp. 97-106.
- [18] Arslan, N.; Demirel, H. The Impact of Temporal Ionospheric Gradients in Northern Europe on Relative GPS Positioning. // Journal of Atmospheric and Solar-Terrestrial Physics. 70(2008), pp. 1382-1400. DOI: 10.1016/j.jastp.2008.03.010
- [19] Lejeune, S.; Wautelet, G.; Warnant, R. Ionospheric Effects on Relative Positioning Within a Dense GPS Network. // GPS Solutions. 16, 1(2012). pp. 105-116. DOI: 10.1007/s10291-011-0212-1
- [20] Warnant R.; Pottiaux, E. The Increase of the Ionospheric Activity as Measured by GPS. // Earth Planets Space. 52(2000), pp. 1055-1060. DOI: 10.1186/BF03352330
- [21] Schaer, S. Mapping and Predicting the Earth's Ionosphere Using the Global Positioning System. Ph.D. Thesis, Astronomical Institute, University of Berne, Switzerland, 1999.
- [22] Dach, R.; Hugentobler, U.; Fridez, P.; Meindl, M. Bernese GPS Software Version 5.0. Astronomical Institute, University of Berne, 2007.
- [23] Komjathy, A.; Langley, R. B. An Assessment of Predicted and Measured Ionospheric Total Electron Content Using a Regional GPS Network. // Presented at the Institute of Navigation National Technical Meeting / Santa Monica, Calif., 1996.
- [24] Camargo, P. O.; Monico, J. F. G.; Ferreira, L. D. D. Application of Ionospheric Corrections in the Equatorial Region for L1 GPS Users. // Earth Planets Space. 5, 2(2000), pp. 1083-1089. DOI: 10.1186/BF03352335
- [25] Wild, U. Ionosphere and Satellite Systems: Permanent GPS Tracking Data for Modelling and Monitoring. // Geodätisch-geophysikalische Arbeiten in der Schweiz, Band 48, 1994.
- [26] Sardon, E.; Zarraoa, N. Estimation of Total Electron Content Using GPS Data: How stable are the differential satellite and receiver instrumental biases? // Radio Sci. 32, 5(1997), pp. 1899-1910. DOI: 10.1029/97RS01457
- [27] Koch, K. R. Parameter Estimation and Hypothesis Testing in Linear Models. Springer-Verlag, Berlin, 1999. DOI: 10.1007/978-3-662-03976-2
- [28] Asgari, J.; Harmel, A. Least Squares Spectral Analysis of GPS Derived Ionospheric Data. // In The ION National Technical Meeting, Jan. 24-26 San Diego, California, 2005.
- [29] Lawson, C. L.; Hanson, R. J. Solving Least-Squares Problems. Englewood Cliffs, N. J., Prentice-Hall, Inc. Series in Automatic Computation, 1974.
- [30] Rideout, W.; Coster, A. Automated GPS Processing for Global Total Electron Content Data. // GPS Solution. 10, 3(2006), pp. 219-228. DOI: 10.1007/s10291-006-0029-5
- [31] Rousseeuw, P. J.; Leroy, A. M. Robust Regression and Outlier Detection. New York, John Wiley and Sons, 1987. DOI: 10.1002/0471725382
- [32] Blewitt, G. An Automatic Editing Algorithms for GPS Data. // Geophysical Research Letters. 17, 3(1990), pp. 199-202. DOI: 10.1029/GL017i003p00199
- [33] Liu, J. Y.; Tsai, H. F.; Lin, C. H.; Kamogawa, M.; Chen, Y. I.; Lin, C. H.; Huang, B. S.; Yu, S. B.; Yeh, Y. H.

Coseismic Ionospheric Disturbances Triggered By The Chi-Chi Earthquake. // J. Geophys. Res. 115, A8(2010), pp. 1-12. DOI: 10.1029/2009JA014943

- [34] Scripps Orbit and Permanent Array Center (SOPAC). 2014. URL: <http://sopac.ucsd.edu/>. (16.09.2014).
- [35] Gülal, E.; Dindar, A. A.; Akpınar, B.; Tiryakioğlu, İ.; Aykut, N. O.; Erdoğan, H. Analysis and management of GNSS reference station data. // Tehnički vjesnik-Technical Gazette. 22, 2(2015), pp. 407-414. DOI: 10.17559/TV-20140717125413

Author's address

Ramazan Cuneyt Erenoglu, Assist. Prof. Dr.

Canakkale Onsekiz Mart University
Faculty of Engineering
Department of Geomatics Engineering
17020 Canakkale, Turkey
Tel: +90 (286) 218 00 18
E-mail: ceren@comu.edu.tr