Title page:

Title: Statistical analysis of ionospheric total electron content (TEC): Long-term estimation of extreme TEC in Japan

Author #1: Michi Nishioka, National Institute of Information and Communications Technology (NICT), Tokyo, Japan, nishioka@nict.go.jp

Author #2: Susumu Saito, Electronic Navigation Research Institute (ENRI), National Institute of Maritime, Port and Aviation Technology (MPAT), Tokyo, Japan

Author #3: Chihiro Tao, NICT, Tokyo

Author #4: Daikou Shiota, NICT, Tokyo, Japan

Author #5: Takuya Tsugawa, NICT, Tokyo, Japan

Author #6: Mamoru Ishii, NICT, Tokyo, Japan

Indicate the corresponding author

Michi NISHIOKA (nishioka@nict.go.jp)

Abstract

Ionospheric total electron content (TEC) is one of the key parameters for users of radio-based systems, such as the Global Navigation Satellite System, high-frequency communication systems, and space-based remote sensing systems, since total ionospheric delay is proportional to TEC through the propagation path. It is important to know extreme TEC values in readiness for hazardous ionospheric conditions. The purpose of this study is to estimate extreme TEC values with occurrences of once per year, ten years, and hundred years in Japan. In order to estimate the extreme values of TEC, a cumulative distribution function of daily TEC is derived using 22 years of TEC data from 1997 to 2018. The extreme values corresponding to once per year and ten years are 90 and 110 TECU, respectively, in Tokyo, Japan. On the other hand, the 22-year data set is not sufficient to estimate the once-per-hundred-year value. Thus, we use the 62-year data set of manually scaled ionosonde data for the critical frequency of the F-layer (foF2) at Kokubunji in Tokyo. First, we study the relationship between TEC and foF2 for 22 years and investigate the slab thickness. Then the result is applied to the statistical distribution of foF2 data for 62 years. The result shows that the once-per-100-

year TEC is about 130-190 TECU at Tokyo. The value is also estimated to be 160-230 TECU in Kagoshima and 100-150 TECU in Hokkaido, in the southern and northern parts of Japan, respectively.

Keywords

total electron content (TEC), extreme TEC, long-term ionosonde observation, manually scaled foF2, slab thickness

Main Text

1. Introduction

The ionospheric condition is one of the most important space weather features for users of radio-based systems, such as navigation systems based on the Global Navigation Satellite System (GNSS), high frequency (HF) communication systems, and space-based remote sensing systems. Radio waves propagating in the ionosphere experience a delay in group velocity and advance in phase velocity due to the electrons in the ionosphere. The ionospheric delay is proportional to the ionospheric total electron content (TEC) along the propagation path. The easiest way to correct the ionospheric delay is to utilize broadcast ionospheric delay models based on simple empirical TEC models such as the Klobuchar (Klobuchar 1987) and NeQuick (Hochegger et al. 2000, Radicella and Leitinger 2001) models. The TEC value is determined by many factors, such as solar activity, the season, local time, and geomagnetic activity. There is also latitudinal dependence in TEC variations. TEC variations caused by solar activity, the season, and local time may be estimated using these simple models but those caused by geomagnetic storms and other phenomena cannot be fully removed from these models. Therefore, users of radio-based systems may be affected by positive and/or negative ionospheric storms. During negative ionospheric storms, TEC is greater than or equal to 0 TECU even if the negative storm is extremely severe. On the other hand, extreme TEC values during positive storms are not unknown and should be studied.

For the design and operation of systems that may be impacted by space weather phenomena, it is important to know the possible extent of the impact and how often such events are likely to occur. Thus, it is important to study extreme values related to various space weather phenomena. For users of trans-ionosphere radio-based systems, the extreme TEC value is a key value.

Extreme values of some space weather parameters have been studied. For example, that of the Dst index was investigated using extreme value modeling (Tsubouchi and Omura 2007). Those of the solar flare X-ray flux, speed of coronal mass ejection, Dst index, and proton energy in proton events were studied by Riley (2012) using complementary cumulative distribution functions. More recently, that of short-wave fadeout by a solar flare was examined on the basis of long-term ionosonde observation data (Tao et al., in this issue).

However, extreme TEC values of once per long period of time have not yet been quantitatively estimated. Several countries have prepared documents with space weather benchmarks. The US White House published “Space Weather Phase 1 Benchmarks” in June 2018 (US White House, 2018). Although it lists three factors that cause ionospheric disturbances, such as geomagnetic storms, quantitative benchmarks were not provided because the ionospheric effects of geomagnetic storms on the ionosphere largely differ from event to event and even their mechanism is not completely understood.

Another reason why extreme TEC values have not been fully studied is that only 20 years has passed since the start of fully fledged TEC observations. TEC observations started with measurements of the Faraday rotation or Doppler effect many decades ago (Bauer and Daniels 1959; Evans 1977). Since these observations were conducted by a few transmitters and receivers, it is difficult to study TEC behavior statistically. With the spread of GNSS and its ground-based receivers, the number of TEC observations dramatically increased. Thanks to the GNSS-TEC observation systems, we have learned a lot about TEC behavior during the last 20 years (for example Foster, 2007; Nishioka et al. 2009; Maruyama et al. 2013). The purpose of this study is to estimate extreme values of TEC with their occurrence rates. We investigate the occurrence rates of extreme values of TEC in Japan in the short, mid-, and long term, which are once per year, ten years, and hundred years, respectively.

To evaluate TEC corresponding to an occurrence rate of once per hundred years, 20 years of data is obviously insufficient. Furthermore, solar activity in the last 20 years has on average been moderate, although several intense geomagnetic storms occurred during solar cycle 24. Compared with GNSS-TEC observation, ionosonde observation has a much longer history. This technique was developed in the late 1920s and began to be implemented in the 1940s in order to monitor shortwave propagation (Gladden 1959). In Japan, ionosonde observation began in 1931. After going through various changes, routine ionosonde observation was started by the predecessor of National Institute of Information and Communications Technology (NICT) in 1951 using an automatic system. Ionospheric parameters derived from the long-term ionosonde observation are archived by World Data Center for the Ionosphere at NICT (http://wdc.nict.go.jp/IONO/wdc/). Long-term ionosonde data have been used for various studies such as a study of the long-term trends of the ionosphere (Xu et al., 2004) and for the development of empirical models (Bilitza, 2018; Yue et al., 2006; Maruyama, 2011). As the TEC and the maximum density of the F region derived from ionosonde observation (NmF2) are known to be correlated, NmF2 can be a proxy of TEC. In this study, about 70 years of data of ionospheric parameters derived from the long-term ionosonde observation are used. Although the data period is still shorter than one hundred years, we investigate statistical characteristics of extreme TEC values in order to estimate the ionospheric once-per-hundred-year condition.

The TEC value over Japan depends on the latitude, normally with a larger value in southern Japan. Japan is mainly located in the lower mid-latitude region with a latitudinal range of about 20 degrees. The southern part of Japan is located at the poleward slope of the equatorial ionospheric anomaly (EIA) crest. On the other hand, the northern part is hardly affected by EIA variation and may rather be affected by phenomena originating from the polar region (Cherniak et al., 2015). Therefore, extreme TEC values should also differ among the center, southern, and northern parts of Japan.

Details of the data set used in this study and the analysis method are described in

Sections 2 and 3, respectively. Analysis results are presented in Section 4. In Section 4, the result obtained using about 20 years of TEC data collected in Tokyo, which is almost in the center of Japan, is shown as the first step. Then long-term ionosonde data are analyzed. On the basis of the result, extreme TEC values with probabilities of once per year, ten years, and hundred years are estimated for Tokyo. In the last part of Section 4, the extreme TEC values in southern and northern Japan are also estimated. In Section 5, the results are discussed in comparison with those of case studies of geomagnetic storms in previous papers. Section 6 provides the conclusions of this study.

2. Data Set

In this study, we use TEC data derived from the nationwide GNSS network over Japan, which is called the GNSS Earth Observation Network System (GEONET) and operated by the Geospatial Information Authority of Japan, and ionosonde observation data collected over Tokyo.

GNSS-TEC data derived from GEONET have been archived by NICT since 1997. Using the network data, the slant TEC along the line of sight between the receiver and the satellite was derived from pseudo-range and carrier-phase measurements by dual-frequency GPS receivers (Saito et al., 1998). The instrumental bias of the TEC associated with the inter-frequency bias of the satellite and receiver was obtained by a technique proposed by Otsuka et al. (2002), in which the daily bias values are derived by assuming that hourly averaged TEC values are uniform within the field of view of a given GNSS receiver. The slant TEC is converted to the vertical TEC after removing the instrumental bias. The TEC data from small satellite elevation angles, which is smaller than 35° is neglected to reduce cycle slips and errors due to conversion from slant to vertical TEC. The median value of the vertical TEC whose ionospheric pierce point is located within 100 km from a given location over one hour is derived as an hourly TEC. The largest hourly TEC in a given day is noted as the daily TEC in this paper. The daily TECs of 22 years from 1997 to 2018 are used in this study and studied in Section 4.1.

Ionospheric conditions have been monitored for about 70 years by NICT using ionosondes in Kokubunji, Tokyo (36.7°N, 139.5°E, 26.8°N in Mag.Lat) and other stations. Ionospheric parameters have been manually scaled from ionograms. In order to ensure uniform quality of data, the scalers have discussed and established scaling rules, although automatic scaling tools have been developed in recent years. Thanks to the substantial efforts of the scalers, ionospheric parameters from the 1950's to the present are now available. In this study, the manually scaled critical frequency of the F-layer (foF2), which corresponds to the peak density of the F-layer, is used. In order to study foF2 with the daily TEC, we refer to the maximum foF2 in a given day as the daily foF2. In Section 4.2, a 22-year data set of daily foF2 values from 1997 to 2018 is used. In the same section, a 62-year data set of daily foF2 values from 1957 to 2018 is also used.

3. Method

In order to find extreme values of TEC corresponding to an occurrence frequency of once every certain number of years, the cumulative distribution function (CDF) of daily TEC occurrence is investigated (Riley, 2012; Kataoka, 2020). The CDF of the daily TEC occurrence is a distribution function of daily TEC values that are greater than or equal to a critical TEC. One of the advantages of investigating the CDF instead of a simple occurrence probability is that it is easy to find TEC values with an occurrence frequency of once per long period (Riley, 2012). In other words, the CDF of the daily TEC occurrence provides an occurrence probability of a daily TEC that is greater than or equal to a certain value, while a normal distribution provides the occurrence probability of a daily TEC between two values.

Although a data set of TEC values over 22 years may be sufficient to investigate TEC values with occurrence frequency of once per year and ten years, it would not be sufficient to investigate the TEC value with an occurrence frequency of once per hundred years. To compensate the insufficient TEC data set, we utilized a 62-year data set of foF2 values in order to calculate NmF2 and study a property of the relationship between TEC and NmF2. The relationship between TEC and foF2 is given by the following equation:

$$TEC = S \times NmF2 \qquad (1)$$

where S is the slab thickness. In this study, characteristics of slab thickness are studied using the 22-year data set of TEC and foF2 values. By utilizing the characteristics of the slab thickness and the 62 years of foF2 data, we deduce CDFs of TEC values over 62 years, from which we estimate the TEC value corresponding to occurrence frequency of once per hundred years.

In this paper, two methods are tested to deduce CDFs of TEC values over 62 years. In the former method, which we call Method I, the following four steps are taken to derive the CDF using the 62-year data set of NmF2. For the first step, probability function of slab thickness, P_s, is presumed with the 22-year slab thickness data set. The presumed P_s is used to calculate a probability function of TEC for a given i-th day, P_T^i, with NmF2 observed on the day, $NmF2^i$. In the third step, P_T^i is converted to CDF^i, which is a CDF of TEC for i-th day. Finally, CDF^i is derived with all NmF2 values in 62 years and integrated to deduce CDF of TEC values over 62 years.

Here, in the step one, we assume that slab thickness follows a normal distribution, e.g.,

$S\sim\mathcal{N}(\mu_S,\sigma_S^2)$ where μ_s and σ_S are mean and standard deviation of slab thickness based on 22 years. The probability function of Ps for slab thickness of *s* [km] is described as follows

$$Ps(s)=\frac{1}{\sqrt{2\pi}\sigma_S}\exp\left(-\frac{(s-\mu_S)}{2\sigma_S^2}\right) \qquad (2)$$

One of the problems in estimating extreme TEC value of once-in-hundred-years is that the number of TEC data, or slab thickness data is insufficient compare to a hundred years. Therefore, the normal distribution, $\mathcal{N}(\mu_S,\sigma_S^2)$, cannot reproduce extreme slab thickness. In order to compensate the lack of extreme values with $\mathcal{N}(\mu_S,\sigma_S^2)$, we introduce inflated factor to model the slab thickness. Inflated sigma is described as $\widehat{\sigma_S}$. Inflation factor, $\frac{\widehat{\sigma_S}}{\sigma_S}$, is determined by overbounding a normal distribution to the slab thickness data set.

As the step 2, a probability function of TEC for *i*-th day, P_T^i is calculated on the assumption that NmF2 and slab thickness are independent parameters. The TEC^i follows a normal distribution with mean and standard deviation of $NmF2_i\times\mu_S$ and $NmF2_i\times\sigma_S$, respectively. That is, $TEC^i\sim\mathcal{N}(\mu_T,\sigma_T^2)$ where $\mu_T=NmF2_i\times\mu_S$ and

$\sigma_T = NmF2_i \times \sigma_S$. The distribution of TEC^i for TEC of *t* [TECU] is expressed as the following equation:

$$P_T^i(t) = \frac{1}{\sqrt{2\pi}\sigma_T} \exp\left(-\frac{(t-\mu_t)}{2\sigma_T^2}\right) \qquad (3)$$

Since TEC^i follows normal distribution, CDF of TEC^i, CDF^i, is given using effort function, erf,

$$CDF_i = \int_{TEC}^{\infty} P_T^i\,(t)dt = 1 - \int_{-\infty}^{TEC} P_T^i(t)dt = \frac{1}{2} - \mathrm{erf}\left(\frac{TEC^i}{\sqrt{2}\sigma_T}\right) \qquad (4).$$

In the final step, CDF^i is calculated for each day in the 62 years and added to obtain $CDF,$ that is,

$$CDF = \frac{1}{N}\sum_i CDF^i \qquad (5)$$

where N is the total number of the day in 62 years.

The latter method, which we call Method II, $\mathrm{CDF}_{\mathrm{TEC}}$ of extreme case was deduced by multiplying the extreme slab thickness, which could occur once in ten and hundred years, by the 62-year data set of daily foF2. By assuming that the slab thickness has a normal distribution with mean μ and standard deviation σ, the value corresponding to occurrence of once per ten and hundred years, or 0.03% and 0.003%, are $\mu + 3\sigma$ and

$\mu + 4.2\sigma$, respectively. CDF_{TEC} for the 62 years can be deduced by multiplying the CDF of NmF2 for the 62 years by the extreme values of slab thickness.

Since the slab thickness is known to have seasonal dependence, a single value of the slab thickness is not appropriate for estimating TEC from foF2. In order to estimate P_{ST} in Method I, data set of slab thickness is divided into four seasons, that is, from February to April, from May to July, from August to October, from November to January. Four seasonal P_{ST} are used to estimated CDF_{TEC} in Equations (3), (4) and (5). Three-month data is used to derive P_{ST} in Method I to obtain sufficient number of data for the inflation. On the other hand, monthly data is used to calculate the mean μ and the standard deviation σ in Method II.

4. Results

4.1 Statistical analysis of TEC over 22 years

Figure 1 shows the CDF of the daily TEC occurrence at Tokyo. The occurrence rate is shown on the left axis. The occurrence rate on the left-hand axis of the ordinate is days per hundred years, which is obtained by dividing the occurrence days by the total number of days in 22 years and then multiplying those in 100 years. Therefore, an occurrence rate of one day means an occurrence rate of once per hundred years. The occurrence rate is converted to the occurrence percentage and shown on the right-hand axis of the ordinate. An occurrence probability of 0.3%, which corresponds to a frequency of once per year, is shown as a solid horizontal line. It is found that the daily TEC can reach about 90 TECU with a frequency of once per year. The occurrence probabilities of once per ten years and once per hundred years correspond to 0.03% and 0.003% and are shown with dotted and dashed horizontal lines, respectively. It is found that a daily TEC of more than 100 TECU occurs with a frequency of once per ten years. The TEC values with frequencies of once per year and once per ten years are summarized in Table 1.

On the other hand, the daily once-per-hundred-year TEC value cannot be appropriately estimated from Figure 1 because the distribution is based on only 22 years of data.

The colors in the histograms in Figure 1 represent the classifications based on solar and geomagnetic activity: red, pink, blue, and light blue represent days of high solar activity and high geomagnetic activity (HSHG), high solar activity and low geomagnetic activity (HSLG), low solar activity and high geomagnetic activity (LSHG), and low solar activity and low geomagnetic activity (LSLG), respectively. Solar and geomagnetic activities are respectively defined on the basis of the solar sunspot number (SSN) and disturbance storm-time (DST) index, which are provided as sunspot data from the World Data Center SILSO, Royal Observatory of Belgium, Brussels (http://sidc.be/silso/datafiles) and WDC for Geomagnetism, Kyoto (http://wdc.kugi.kyoto-u.ac.jp/dstdir/index.html), respectively. HS (LS) days are defined as days for which the average daily SSN for the previous 27 days is greater than or equal to (less than) 50. HG (LG) days are defined as days for which the average daily DST of the current day and the previous day is less than or equal to (greater than) -50 nT. It can be seen that a TEC of 60 TECU or larger is most likely to be observed when either the solar activity or the geomagnetic activity is high, while those exceeding 100 TECU are observed only when the solar activity is high.

4.2 Statistical analysis of foF2 over 22 and 62 years

Here, CDFs of the daily foF2 occurrence are studied in order to estimate once-per-hundred-year values. First, a CDF of the daily foF2 occurrence over the same period as in Figure 1, from 1997 to 2018, were examined in comparison with that of the 22 years of TEC data in Figure 1. Figure 2 shows a CDF of the daily foF2 occurrence, that is, the distribution of the daily foF2 that is greater than or equal to some critical foF2. As in Figure 1, the occurrence rate per hundred years is shown on the left-hand axis of the ordinate and the occurrence rate in percentage is shown on the right axis. The occurrence frequencies of once per year, ten years, and hundred years of 0.3%, 0.03%, and 0.003% are shown as solid, dotted, and dashed horizontal lines, respectively. The colors in Figure 2 represent solar and geomagnetic activities similarly to in Figure 1; red, pink, blue, and light blue represent days of HSHG, HSLG, LSHG, and LSLG, respectively. The largest foF2 was about 17.5 MHz. It is found that foF2 was higher than 15 MHz for only HSHG and HSLG days, which is similar to the result in Figure 1.

The same analysis is carried out for the 62-year foF2 data set from 1957 to 2018. The result is shown in Figure 3 in the same format as Figure 2. The maximum observed foF2 is about 18.7 MHz, which is slightly larger than that obtained from the 22-year data set in Figure 2. The maximum foF2 18.7 MHz was observed during geomagnetic storm in November 1960 when DST index reached -333 nT (Cliver and Svalgaard 2004). Moreover, the occurrence rate of daily foF2 values larger than 16.8 MHz in Figure 3, which corresponds to the rightmost bar in the histogram, is about twice of that in Figure 2.

4.3 Estimation of extreme TEC from slab thickness using Method I

As the characteristics of the CDFs of the daily foF2 occurrence are different for the 22- and 62-year data sets, the once-per-hundred-year TEC value cannot be estimated by extrapolating the CDF of the daily TEC occurrence obtained from the 22-year data set. In this sub-section, we estimate the once-per-hundred-year TEC value by using the 62-year foF2 data set with Method I.

The value of foF2 is proportional to the square root of the maximum ionospheric density, NmF2. NmF2 is given by the following equation.

$$NmF2[m^{-3}] = 1.24 \times 10^{10} \times foF2^2\ [MHz] \qquad (5)$$

Figure 4 shows the correlation between daily TEC and NmF2 derived from the daily foF2. All data collected over 22 years are shown in this scatter plot. It can be seen that TEC and NmF2 have a strong correlation. The red line is the least-squares linear approximation of all data. As shown in Equation (1), the slope, which is about 250 km, is equivalent to the thickness of the ionosphere that gives a TEC value with a density of NmF2. This parameter, which is called the ionospheric slab thickness, is used to deduce TEC from NmF2 because of the strong correlation between daily TEC and daily foF2.

In order to derive CDFs of TEC values over 62 years with Method I, distribution of slab thickness is examined. Figure 5 shows distribution of slab thickness from 1997 to 2018. Mean and standard deviation of the distribution is 215 km and 52 km, respectively. The red curve represents the normal distribution with the mean and the standard deviation. The distribution in three months from May to July is shown in Figure 6. The mean and standard deviation of the distribution is 273 km and 45 km. The mean is larger than that in Figure 5, which is one of the seasonal effects. The red curve represents a normal distribution with the mean and the standard deviation. The curve roughly fits the slab thicknesses but does not cover large values such as more than 400 km. Mean values and standard deviations of other seasons, that is, from February to April, from August to October, and November to January, are listed in Table 2. Normal distributions with the mean and the standard deviations for each season listed in Table 2 are applied for P_{ST} in Method I. The result of CDF_{TEC} is shown with black histograms in Figure 7. TEC of once per ten and hundred years, that is, TEC of 0.03% and 0.003% was 35 TECU and 45 TECU, respectively, which are smaller than those can be read in Figure 1. This is because of the assumption of the normal distribution. The normal distribution is inflated using an inflation factor. The blue curves in Figure 6 show the inflated normal distribution with inflation factors. The dashed, solid, dotted, dash-dotted lines are derived with inflation factors of 2.0, 3.0, 4.0, and 5.0, respectively By introducing the inflation factor, large slab thickness can be covered with the distribution. Using the inflated normal distribution with the inflation factor of 3, CDF of TEC are derived as blue histogram in Figure 7. TEC of once per ten and hundred years, that is, TEC of 0.03% and 0.003% was 90 TECU and 130 TECU, respectively.

4.4 Estimation of extreme TEC from slab thickness using Method II

In this sub-section, we estimate the once-per-hundred-year TEC value by using the 62-year foF2 data set with Method II. Here, we calculated the mean and the standard deviation of slab thickness for each month. Figure 8 shows the slab thickness against the day of the year for 22 years from 1997 to 2018. Data are sparser from June to August compared with other months, because foF2 values often cannot be obtained owing to masking by the sporadic E-layer, which often appears in these months. The red polyline is the monthly mean of the slab thickness. The monthly mean slab thickness is about 180 km in winter and 280 km in summer. Blue and red vertical lines indicate the ranges of $\pm 3\sigma$ and $\pm 4.2\sigma$. These ranges are equivalent to probabilities of once per ten and hundred years, respectively, when the estimated slab thickness is assumed to have a normal distribution, that is, occurrence probability of the values larger than average $+3\sigma$ and $+4.2\sigma$ are 0.13% and 0.001%

Here we estimate the daily TEC from the daily NmF2 data, assuming the slab thickness has only seasonal dependence. Figure 9 shows the CDFs of the estimated daily TEC occurrence obtained using the monthly mean slab thickness and observed NmF2 from 1957 to 2018. The black histograms are distributions of the daily TEC estimated with the monthly mean slab thickness, which is shown with a red polyline in Figure 8. The number of days per 100 years and the occurrence rate are shown on the left- and right-hand axes of the ordinate, respectively. The black solid, dotted, and dashed horizontal lines correspond to 0.3% (once a year), 0.03% (once every ten years), and 0.003% (once every hundred years), respectively. The blue histograms in Figure 9 are the distribution of TEC estimated with the average + 3σ slab thickness (upper value of the blue vertical line in Figure 8), which corresponds to a slab thickness with a frequency of once per ten years. According to this histogram, the TEC with a frequency of once per ten years is 130 TECU or more. Furthermore, the red histograms in Figure 9 are derived from the average + 4.2σ slab thickness (upper limit of the red vertical line in Figure 3). This result indicates that TEC values of more than 190 TECU can be observed with a frequency of once per hundred years. These TEC values are summarized in Table 1.

4.4 Latitudinal dependence of extreme TEC

Figures 1–9 are results based on data obtained in Tokyo. Here we estimate extreme TEC values for southern and northern Japan because TEC behavior is expected to be different at different magnetic latitudes. Figure 10 shows the correlations of daily TEC between Tokyo and Kagoshima (31.2ºN, 130.6ºE, 21.7°N in Mag. Lat) and between Tokyo and Hokkaido (45.2ºN, 141.8ºE 36.4°N in Mag. Lat) for 22 years from 1997 to 2018. Basically, the TEC in Tokyo is smaller than that in Kagoshima and larger than that in Hokkaido. The red line represents the linear approximation of these data and reveals that the TECs in Kagoshima and Hokkaido are, on average, 1.2 and 0.8 times that in Tokyo, respectively. From these results, the TEC values with probabilities of once per year, ten years, and hundred years are estimated as 110, 110-160, and 160-230 TECU (70, 70-100, and 100-150 TECU), respectively, in Kagoshima (Hokkaido) as summarized in the second and third rows in Table 1.

5. Discussion

It is important to estimate the occurrence rates of extreme values of TEC in Japan in the short, mid-, and long term, which are once per year, ten years, and hundred years, respectively, in readiness for hazardous ionospheric conditions. “Space Weather Phase 1

Benchmarks", which was published by the USA White House in June 2018, lists three factors that cause ionospheric disturbances: solar flares, proton events, and geomagnetic storms. However, quantitative benchmarks are difficult to derive because the effects of geomagnetic storms largely differ from event to event. Furthermore, the mechanism of ionospheric storms is not yet completely understood. Although the results in this paper are limited to the region around Japan, they are a starting point for evaluating benchmarks in other regions.

In this study, TEC data were extrapolated in two ways: Method I and Method II. Method I assumes the probability distribution of slab thickness as a normal distribution. First, raw σ is used to model the slab thickness with the 22-yead data set. The resulting CDF which is shown with black histograms is Figure 7 underestimates the observed CDFs in Figure 1. The TEC values of once-per-year, for example, was about 90 TECU in Figure 1 while that of black histograms in Figure 7 was less than 30 TECU. By using the inflated normal distribution, the once-per-year TEC value becomes about 80 TECU, which is shown with blue histograms in Figure 7. The value is close to that in Figure 1.

Here, we have to note that the estimated TEC values depend on the inflation factor. TEC of once per hundred year was derived as 130 TECU with an inflation factor of three in this paper. If the inflation factor is assumed as 2, 4, and 5, the once-per-hundred-year TEC is 90, 160, and 200 TECU, respectively. In order to determine the inflation factor, several normal distributions are examined with varying inflation factors. First, we examined a distribution of slab thickness in all seasons. Blue curves in Figure 5 are normal distributions with various inflation factors. It is found that the dashed line (inflation factor =2) among them fits best to the histogram. Inflation factors of 3, 4, and 5 would be too large to overbound the histogram. Next, we examined a distribution of slab thickness in each three months. We have to note that larger inflation factor is necessary to overbound the three-month histogram than to overbound the all-season histogram because the tail of the distribution can be heavier due to less number of data. The inflation factor of 2 is found to be too small for the three-month distribution. For most other seasons, the inflation factor of 2 was too small (Figures not shown), too. Inflation factors of more than three would be the best among the tested inflation factors. Based on these investigations, we concluded that an inflation factor of 2 would be the best for all-season data set among the tested inflation factors while an inflation factor of 3 would be the best for the three-month data set. Since it is clear that slab thickness depends on seasons as shown in Figure 8, we adopted an inflation factor of 3.

Using an inflation factor of 3, extreme value of TEC such as those of once per hundred year would be underestimated with Method I since the distribution does reproduce extreme value of slab thickness. One of the reasons why the inflated normal distribution does not reproduce extreme value of slab thickness is that the number of the 62-year data set is not enough to estimate the extreme value. The resolution of 62-year data set is $1/(365.25 \times 62)$ =0.0044% while the rightmost red bar in Figure 9 shows occurrence rate of approximately 0.005% because of the data lack in some days. This occurrence rate is larger than that of once-in-hundred-year event, $1/(365.25 \times 100) = 0.003\%$.

Another reason why the inflated normal distribution does not reproduce extreme value of slab thickness is that the assumption of the normal distributions P_S are not complete. As shown in Figures 5 and 6, the distribution of slab thickness has long tail, the tail cannot be reproduced by normal distributions even if the σ is inflated. Alternative approach would be to model the distribution in a different way. The distribution in Figures 5 and 6 could be fitted by a sum of two normal functions which centers the core part and the tail parts, so-called double Gaussian, instead of multiplying an inflation factor to the standard deviation, which is left for future studies.

Comparing Method I and Method II, Method II is more conservative than Method I because Method II takes out the extreme slab thickness multiplies it with TEC values. Method I has an advantage in order to grasp the overall distribution while extreme large values are not reproduced. Method II has an advantage in estimating extreme values while overall distribution is not very accurate.

In this study, we estimated extreme TEC values by assuming that the slab thickness has only seasonal dependence. The seasonal dependence of the slab thickness shown in Figure 8 is consistent with the results of previous studies (Jin et al., 2007; Huang et al., 2016). Another factor determining the slab thickness is the dynamics and/or composition change caused by geomagnetic disturbances. According to Stankov and Warnant (2009), the slab thickness is systemically enhanced during geomagnetic disturbances for both positive and negative ionospheric storms. Extreme values of TEC estimated by blue or red histograms in Figure 9 would be recorded during geomagnetic storm conditions.

Extreme positive storms are thought to be caused by a geomagnetic disturbance that induces prompt penetration of the electric field (Tsurutani et al., 2004). The largest reported TEC is about 330 TECU to our knowledge, which was recorded by a GPS receiver onboard the CHAMP satellite at an altitude of about 400 km during the October 2003 Halloween storm (Mannucci et al., 2005). Magnetic latitude where the 330 TECU was observed was about 25°S. Although the observation was in the south hemisphere, the magnetic latitude is similar to that of Tokyo (26.8°N). The TEC value of 330 TECU reported in Mannucci et al. (2005) is much higher than our result of 190 TECU, which is conservatively estimated in Method II.

Before discussing possible reasons for the discrepancy between our result and that reported in Mannucci (2005) , we have to discuss estimation accuracy of the instrumental bias to derive absolute value of TEC. In estimating instrumental bias, we assume that the hourly average of vertical TEC is uniform within an area covered by a receiver; this area approximately corresponds to a surrounding of 1000 km (Otsuka et al., 2002). It is reported that the technique can derive absolute values of TEC with the accuracy of ~3 TECU in the daytime and ~1 TECU in the nighttime, respectively, during quiet and moderated disturbed day. It is also reported the characteristics of temporal and spatial distribution of absolute TEC are consistent with the previous studies during a geomagnetic storm day. Nonetheless, during the geomagnetic disturbed condition, TEC tends to have spatial gradient and large scale travelling ionospheric disturbances (LSTIDs) could appear. The horizontal scale of LSTIDs is more than 2,000 km, which is larger than the assumption of TEC uniformity. Therefore, there is a possibility that the assumption of the TEC uniformity tends to be invalid during severe geomagnetic storm days. Zhang et al. (2009) investigated influences of geomagnetic storms on the estimation of GPS instrumental biases. The bias errors are in order of a few TECU while the errors are different among geomagnetic storms and its duration. Since the order of the errors in estimating instrumental bias is less than ten TEC, we speculate that the error would not reverse the difference between our result (190 TECU) and that in Manucci et al. (330 TECU) while further quantitative investigation would be necessary in order to clarify the estimation errors.

Here we discuss possible reasons for the difference between these values. One possibility is differences in observation opportunities. The characteristics of ionospheric storms are not always similar among geomagnetic storms, with their magnitude varying greatly from event to event. Mannucci et al. (2008) analyzed four intense geomagnetic storms in 2003 including the event for which the extreme value of 330 TECU was observed by the CHAMP satellite. A dramatic increase in TEC was observed in only one event. The observed TEC on the other three storm days was around 100 TECU or less. If the event-to-event difference is too large, 70 years of data might not be enough to estimate TEC values for once-per-hundred-year or once-per-thousand-year events.

Another possibility accounting for the difference between the extreme value of 330 TECU in Mannucci et al. (2005) and our result is the longitude dependence of the ionospheric influence on geomagnetic storms. Immel and Mannucci (2013) analyzed global TEC maps during geomagnetic storms over seven years. Their analysis confirmed that on average the American sector exhibits larger TEC enhancements regardless of the onset UT. Greer et al. (2017) used the Global Ionosphere–Thermosphere Model to carry out an experiment on a geomagnetic storm by modifying the storm arrival UT. The result indicated that the strongest enhancements of TEC during storms are found in the American and Pacific longitude sectors. They suggested that the longitudinal dependences were due to Earth’s asymmetrical geomagnetic topology in the American and Pacific sectors. The difference between our results and that of Mannucci et al. (2005) may originate from the difference between the Japanese and American/Pacific sectors. In order to clarify whether the longitudinal dependence results in the large difference between the results of this study and that of Mannucci et al. (2008), long-term observational data in addition to data over oceans are necessary.

This study focuses on positive ionospheric storms, which may significantly affect GNSS users. On the other hand, the effect of negative storms on space weather users may also be significant, particularly for HF communicators, who may experience blackouts during negative ionospheric storms. In addition, parameters other than TEC, such as maximum usable frequency (MUF) and scintillation indices, should be studied for extreme cases.

6. Summary

In this study, extreme values of TEC with frequencies of once per year, ten years, and hundred years were investigated. The results are summarized as follows:

・The CDF of daily TEC values was studied for a 22-year data set observed in Tokyo in order to estimate TECs with frequencies of once per year and ten years. The obtained once-per-year and once-per-ten-year TECs were 90 and 110 TECU, respectively.

- In order to estimate the once-per-hundred-year TEC value, 62 years of manually scaled ionosonde data were used to augment the insufficient observation period of TEC. The slab thickness was assumed to have only seasonal variation and was used to estimate TEC from 62 years of foF2 data. Two methods are tested. The obtained once-per-hundred-year TEC value was 130-190 TECU in Tokyo.
- Once-per-ten-year TEC value obtained with the two methods are 90 TECU and 130 TECU. The value roughly agrees with 110 TECU which is obtained with 22-year TEC observation.
- Extreme TEC values were also studied for Kagoshima and Hokkaido in southern and northern Japan, respectively. Once-per-hundred-year values of 230 and 150 TECU were obtained, respectively.

Declarations

Ethical approval and consent to participate

Not applicable

Consent for publication

Not applicable

List of abbreviations

EIA: equatorial ionospheric anomaly

EUV: solar extreme ultraviolet (EUV)

foF2: critical frequency of the F-layer

GEONET: GNSS Earth Observation Network System

GNSS: Global Navigation Satellite System

HF: high frequency

HSHG: high solar and high geomagnetic activity

HSLG: high solar and low geomagnetic activity

LSHG: low solar and high geomagnetic activity

LSLG: low solar and low geomagnetic activity

MUF: maximum usable frequency

NICT: National Institute of Information and Communications Technology

NmF2: maximum density of the F2 layer

TEC: total electron content

Availability of data and materials

The TEC data used in this study are archived on NICT’s homepage (https://aer-nc-web.nict.go.jp/GPS/GEONET/). Manually scaled ionosonde parameters are also archived on NICT’s homepage (http://wdc.nict.go.jp/IONO/HP2009/ISDJ/manual_txt.html).

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by MEXT/JSPS KAKENHI Grant 15H05813.

Authors’ contributions

MN conducted the research and has responsibility for the results presented in this paper. SS has supported this analysis and contributed to the discussion. CT, DS, TT, and MI contributed to the discussion as experts of ionosphere and space weather. All authors read and approved the final manuscript.

Acknowledgements

GPS data of GEONET were provided by the Geospatial Information Authority of Japan. MN thanks to ionograms scalers of NICT and forerunner of NICT for valuable manually scaled data.

References

Bauer, Siegfried J., and Fred B. Daniels (1955) Measurements of Ionospheric Electron Content by Lunar Radio Technique, Journal of Geophysical Research, 64:10:1371-1376.

Bilitza, D. (2018) IRI the International Standard for the Ionosphere. Adv. Radio Sci., 16, 1–11, https://doi.org/10.5194/ars-16-1-2018.

Cherniak, I., I. Zakharenkova, and R. J. Redmon (2015) Dynamics of the high-latitude ionospheric irregularities during the 17 March 2015 St. Patrick’s Day storm: Ground-based GPS measurements. Space Weather, 13, 585-597. doi:10.1002/2015SW001237.

Foster, J. C., W. Rideout, B. Sandel, W. T. Forrester, and F. J. Rich (2007) On the relationship of SAPS to storm-enhanced density, J. Atmos. Sol. Terr. Phys. 69:303-313.

Cliver, E. W. and L. Svalgaard (2004) The 1859 solar-terrestrial disturbances and the current limits of extreme space weather activity, Solar Physics. 224, 407–422

Gladden, Sanford C. (1959) A history of vertical-incidence ionosphere sounding at the national bureau of standards, United states department of commerce office of technical services.

Greer, K. R., T. Immel, and A. Ridley (2017) On the variation in the ionospheric response to geomagnetic storms with time of onset. J. Geophys. Res. Space Phys. 122:4512-4525. doi:10.1002/2016JA02345.

Hochegger, G, B. Nava, S.M. Radicella, R. Leitinger (2000) A family of ionospheric models for different uses, Physics and Chemistry of the Earth, Part C: Solar, Terrestrial & Planetary Science, 25:4:307-310

Huang, H., L. Liu, Y. Chen, H. Le, and W. Wan (2016). A global picture of ionospheric slab thickness derived from GIM TEC and COSMIC radio occultation observations. J. Geophys. Res. Space Phys. 121:867-880. doi:10.1002/2015JA021964.

Immel, T. J., and A. J. Mannucci (2013) Ionospheric redistribution during geomagnetic storms. J. Geophys. Res. Space Phys. 118:7928-7939. doi:10.1002/2013JA018919.

Jin, S., J.-H. Cho, and J.-U. Park (2007). Ionospheric slab thickness and its seasonal variations observed by GPS, J. Atmos. Sol. Terr. Phys. 69(15):1864-1870. doi:10.1016/j.jastp.2007.07.008.

Kataoka Ryuho (2020), Extreme geomagnetic activities: a statistical study, Vol 72, 124, https://doi.org/10.1186/s40623-020-01261-8.

Klobuchar, J. A. (1987), Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users, IEEE Transactions on Aerospace and Electronic Systems, Vol AES-23:3 325-331.

Mannucci, A. J., B. T. Tsurutani, B. A. Iijima, A. Komjathy, A. Saito, W. D. Gonzalez, F. L. Guarnieri, J. U. Kozyra, and R. Skoug (2005) Dayside global ionospheric response to the major interplanetary events of October 29 . 30, 2003 Halloween Storm. Geophys. Res. Lett. 32:L12S02. doi:10.1029/2004GL021467.

Mannucci, A. J., B. T. Tsurutani, M. A. Abdu, W. D. Gonzalez, A. Komjathy, E. Echer, B. A. Iijima, G. Crowley, and D. Anderson (2008) Superposed epoch analysis of the dayside ionospheric response to four intense geomagnetic storms. J. Geophys. Res. 113:A00A02, doi:10.1029/2007JA012732.

Maruyama, T (2011) Modified solar flux index for upper atmospheric applications. J. Geophys. Res. 116:A08303. doi:10.1029/2010JA016322.

Maruyama, T., G. Ma, and T. Tsugawa (2013) Storm-induced plasma stream in the low-latitude to midlatitude ionosphere. J. Geophys. Res. Space Phys. 118:5931-5941. doi:10.1002/jgra.50541.

Nishioka, M., A. Saito, and T. Tsugawa (2009) Super-medium-scale traveling ionospheric disturbance observed at midlatitude during the geomagnetic storm on 10 November 2004. J. Geophys. Res. 114:A07310. doi:10.1029/2008JA013581.

Otsuka, Y., T. Ogawa, A. Saito, T. Tsugawa, S. Fukao, and S. Miyazaki (2002) A new technique for mapping of total electron content using GPS network in Japan. Earth Planets Space 54:63-70.

Radicella, S. M., R. Leitinger (2001) The evolution of the DGR approach to model electron density profiles, Advances in Space Research, 27:1:35-40

Riley, P. (2012) On the probability of occurrence of extreme space weather events. Space Weather 10:S02012. doi:10.1029/2011SW000734.

Saito, A., S. Fukao, and S. Miyazaki (1998) High resolution mapping of TEC perturbations with the GSI GPS network over Japan. Geophys. Res. Lett. 25:3079-3082. doi:10.1029/98GL52361.

Stankov, S. M. and R. Warnant (2009) Ionospheric slab thickness. Analysis, modelling and monitoring. Adv. Space Res. 44:1295-1303.

Tao, C., M. Nishioka, T. Tsugawa, S. Saito, D. Shiota, K. Watanabe, and M. Ishii (2020) Statistical analysis of short-wave fadeout for extreme event estimation. Earth Planets Space, submitted to this issue.

Tsubouchi, K. and Y. Omura (2007) Long-term occurrence probabilities of intense geomagnetic storm events. Space Weather 5:S12003. doi: 10.1029/2007SW000329.

Tsurutani, B. T., et al. (2004) Global dayside ionospheric uplift and enhancement associated with interplanetary electric fields. J. Geophys. Res. 109, A08302, doi:10.1029/2003JA010342.

US White House (2018) Space Weather Phase 1 Benchmarks. Space Weather Operations, Research, and Mitigation Subcommittee Committee on Homeland and National Security, National Science & Technology Council. https://www.whitehouse.gov/wp-content/uploads/2018/06/Space-Weather-Phase-1-

Benchmarks-Report.pdf.

Xu, Z.-W., J. Wu, K. Igarashi, H. Kato, and Z.-S. Wu (2004). Long-term ionospheric trends based on ground-based ionosonde observations at Kokubunji, Japan. J. Geophys. Res. 109:A09307. doi:10.1029/2004JA010572.

Yue, X., W. Wan, L. Liu, B. Ning, and B. Zhao (2006). Applying artificial neural network to derive long-term foF2 trends in the Asia/Pacific sector from ionosonde observations. J. Geophys. Res. 111:A10303. doi:10.1029/2005JA01157.

Zhang, W., D. H. Zhang, and Z. Xiao (2009), The influence of geomagnetic storms on the estimation of GPS instrumental biases, Ann. Geophys., 27, 1613–1623, www.ann-geophys.net/27/1613/2009/

Figure legends

Figure 1.

Cumulative distribution function (CDF) of daily TEC occurrence at Tokyo from 1997 to 2018. The occurrence rate, which is the number of days per hundred years, and the occurrence percentage are shown on the left and right axes, respectively. Red, pink, blue, and light blue represent days of HSHG, HSLG, LSHG, and LSLG, respectively. The solid, dotted, and dashed horizontal lines represent occurrence rates of 0.3%, 0.03%, and 0.003%, which correspond to occurrence frequencies of once per year, ten years, and hundred years, respectively.

Figure 2.

CDF of the daily foF2 occurrence from 1997 to 2018 at Kokubunji station, Tokyo. The occurrence rate, which is the number of days per hundred years, and the occurrence rate in percentage are shown on the left- and right-hand axes of the ordinate, respectively. Red, pink, blue, and light blue represent days of HSHG, HSLG, LSHG, and LSLG, respectively. The solid, dotted, and dashed horizontal lines represent occurrence rates of 0.3%, 0.03%, and 0.003%, which correspond to frequencies of once per year, ten years, and hundred years, respectively.

Figure 3.

CDF of the daily foF2 from 1957 to 2018. The plotting format is the same as that of Figure 2.

Figure 4.

Scatter plot of daily TEC and corresponding daily NmF2 from 1997 to 2018. The red line represents a linear fitting to the data points.

Figure 5.

Distribution of slab thickness from 1997 to 2018. The red curve represents normal distribution with the mean and standard deviation. The blue curves represent inflated normal distributions with the mean and the inflated sigma. The inflated sigma is two, three, four, and five times of the original standard deviation for the dashed, solid, dotted, and dash-dotted lines, respectively.

Figure 6.

Distribution of slab thickness during May, June, and July from 1997 to 2018. The red curve represents the normal distribution with the mean and the standard deviation calculated from the data set. The blue curves represent inflated normal distributions with the mean and the inflated sigma. The inflated sigma is two, three, four, and five times of the original standard deviation for the dashed, solid, dotted, and dash-dotted lines, respectively.

Figure 7.

CDFs of the daily TEC occurrence estimated with Method I. The occurrence rate, which is the number of days per hundred years, and the occurrence rate in percentage are shown on the left- and right-hand axes of the ordinate, respectively. The black histograms are derived with the normal distribution of 22-year data set of slab thickness and 62-year data set of daily foF2. The blue histograms are derived with the inflated normal distribution of the slab thickness and daily foF2.

Figure 8.

Slab thickness against day of year: The red polyline is the monthly mean value of slab thickness. Blue and red vertical bars represent $\pm 3\sigma$ and $\pm 4.2\sigma$, respectively.

Figure 9.

CDFs of the daily TEC occurrence estimated with Method II. The occurrence rate, which is the number of days per hundred years, and the occurrence rate in percentage are shown on the left- and right-hand axes of the ordinate, respectively. The black histograms are derived from the average slab thickness shown in Figure 8. The blue and red histograms are derived with slab thicknesses of average$+3\sigma$ and $+4.2\sigma$, which are shown with blue and red vertical lines, respectively. The solid, dotted, and dashed horizontal lines represent occurrence rates of 0.3%, 0.03%, and 0.003%, which correspond to frequencies of once per year, ten years, and hundred years, respectively.

Figure 10.

Correlation of daily TEC between (a)Tokyo and Kagoshima and (b)Tokyo and Hokkaido from 1997 to 2018. The red line represents the linear approximation of each set of data.

Table legends

Table 1.

Estimated TEC of once per one, ten, and hundred years in Tokyo, Kagoshima, and Hokkaido. The unit is in TECU.

Table 2.

Mean and standard deviation of slab thickness in km for four seasons.