

Document downloaded from:

<http://hdl.handle.net/10251/166017>

This paper must be cited as:

Baselga Moreno, S. (2020). A combined estimator using TEC and b-value for large earthquake prediction. *Acta Geodaetica et Geophysica Hungarica*. 55(1):63-82.
<https://doi.org/10.1007/s40328-019-00281-5>



The final publication is available at

<https://doi.org/10.1007/s40328-019-00281-5>

Copyright Akademiai Kiado Zrt.

Additional Information

A combined estimator using TEC and *b*-value for large earthquake prediction

Sergio Baselga

Cartographic Engineering, Geodesy and Photogrammetry Dept., Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain. E-mail address: serbamo@cgf.upv.es. ORCID: 0000-0002-0492-4003.

ABSTRACT

Ionospheric anomalies have been shown to occur a few days before several large earthquakes. The published works normally address examples limited in time (a single event or few of them) or space (a particular geographic area), so that a clear method based on these anomalies which consistently yields the place and magnitude of the forthcoming earthquake, anytime and anywhere on earth, has not been presented so far. The current research is aimed at prediction of large earthquakes, that is with magnitude M_w 7 or higher. It uses as data bank all significant earthquakes occurred worldwide in the period from January 1, 2011 to December 31, 2018. The first purpose of the research is to improve the use of ionospheric anomalies in the form of TEC grids for earthquake prediction. A space-time TEC variation estimator especially designed for earthquake prediction will show the advantages with respect to the use of simple TEC values. Further, taking advantage of the well-known predictive abilities of the Gutenberg-Richter law's *b*-value, a combined estimator based on both TEC anomalies and *b*-values will be designed and shown to improve prediction performance even more.

Keywords: earthquake prediction; total electron content; ionospheric anomaly; *b*-value.

ARTICLE HIGHLIGHTS

- A space-time TEC variation estimator for large earthquake prediction is presented
- A combined estimator using TEC and *b*-values improves prediction even more
- Promising results are obtained for large earthquakes worldwide in years 2011 to 2018

1. Introduction

Earthquake prediction is a field of research that has gained significant attention in recent decades. Generally considered as an impossible task for centuries, the forecast of an earthquake event based on several so-called earthquake precursors seems every day more a matter of future development, rather than impossibility, taking into account the non-negligible degree of success of some of these precursors.

Among the list of possible types of precursors that have shown a modest degree of evidence for particular seismic events, strange animal behavior is normally mentioned: examples include snakes coming out hibernation dens (Buskirk *et al.* 1981) and the relatively recent episode of surprising abandonment of a breeding site by common toads five days before the April 9, 2009 M_w 6.3 L'Aquila earthquake (Grant and Halliday 2010). These abnormal behaviors might be triggered by certain changes in the environment which can be noticed in the land surface, the water, the air and the ionosphere, and have been extensively documented (Grant *et al.* 2011).

This is particularly the case of total electron content (TEC) anomalies observed in the ionosphere prior to several large earthquakes. Pulinets and Boyarchuk (2004) give a comprehensive review of the topic, for which many particular successful cases have been published (e.g. Akhoondzadeh and Saradjian 2011, Dogan *et al.* 2011, Guo *et al.* 2017, Lin 2010 and 2011, Liu *et al.* 2009, Yao *et al.* 2012, Zakharenkova *et al.* 2007a and 2007b). While the complete underlying mechanism may be unknown, a plausible explanation for the observed anomalous TEC values is often suggested: when tectonic stresses build up in the Earth's crust, highly mobile electronic charge carriers are activated causing air ionization, injecting massive amounts of primarily positive ions into the lower atmosphere (Grant *et al.* 2011). For more details on this tectonic stress theory, including the generation of stress-activated electronic charge carriers, ion emission by radon decay and air ionization up into the ionosphere, one can consult Freund *et al.* (2009). Laboratory tests have confirmed the plausibility of the mechanism (e.g. Warwick *et al.* 1982).

Seismic variables, the Gutenberg-Richter law's b -value in particular, have also been demonstrated to be clearly correlated to posterior earthquake events (Asencio-Cortés *et al.* 2018, Borgohain *et al.* 2018, Florido *et al.* 2015 and 2018, Kulhanek *et al.* 2018, Nuannin *et al.* 2005, Reyes *et al.* 2013). As it can be seen in the following subsection 1.2., this b -value is computed from the seismic record only, hence completely independent from TEC measurements, and could be exploited along with ionospheric anomalies to construct a more robust earthquake precursor. This is the main purpose of the present work.

Further, while many successful detection cases have been published, a method producing consistent results in the long run is still missing. Studies should show, in particular, whether or not the type and size of the TEC anomalies observed for an earthquake event can be transferrable to other earthquake event, in a different time and place, and how to distinguish them from ionospheric anomalies of non-seismic origin thus avoiding false positives. Such type of work, addressing both the successes and failures of the application of the selected earthquake precursor to a dataset with a large time span and large geographic domain (worldwide) is still missing.

In the present paper we want to remedy to a certain extent this lack of continuity among the published cases that can be found in the literature, so that all significant earthquakes occurred worldwide from January 1, 2011 to December 31, 2018 are considered in the study. As a note of caution, it is worth mentioning that by significant earthquakes we mean here those above magnitude M_w 6.5 or 7, inasmuch as smaller earthquakes do not easily show their imprint in the ionosphere, as explained in the following subsection.

After summarizing the general consensus in the literature regarding the use of TEC anomalies as earthquake precursors, we describe the input data used in this study and then devise an estimator that consistently forecasts the magnitude of the maximum earthquake to happen in the next 4-5 days for each point and date of study during the years 2011-2018. It is expected that this estimator can be used for prediction of large earthquakes (that is of magnitude M_w 7 or higher) in a future

construction of a successful earthquake prediction real-time system, which should provide public authorities with a reliable location, date and magnitude of significantly large earthquakes well ahead of time (days or at least several hours) to permit taking the corresponding safety measures. The validation of such system, taking into account not only its successes but also its false positives and false negatives will be key to its implementation, which should forecast the large events while avoiding the psychological and economic cost of false alarms.

1.1. TEC anomalies as earthquake precursors

How could we predict a large earthquake from rapid TEC grids available prior to the event? What type of behavior should we look for? What are the fingerprints of TEC values preceding a large earthquake? Much (and sometimes conflicting) has been said in the literature about the features of TEC anomalies serving as earthquake precursors. A significant consensus has been reached in the following points:

- Anomalous TEC values are observed from 5 days to some hours before a seismic shock (Pulinets *et al.* 2003, Pulinets and Boyarchuk 2004, Lin 2010, Liu *et al.* 2006 and 2009, Zakharenkova *et al.* 2007a and 2007b).
- The threshold magnitude at which effects on the ionosphere become observable is $M_w = 5$ in theory (Pulinets and Boyarchuk 2004) although the example cases that can be found in the literature normally study larger magnitudes: $M_w \geq 6.0$ (Guo *et al.* 2017, Lin 2011), or more usually even higher magnitudes, $M_w \geq 7.0$ (Akhoondzadeh and Saradjian 2011, Liu *et al.* 2009), $M_w \geq 8.0$ (Yao *et al.* 2012, Zakharenkova *et al.* 2007b).
- The size of the anomalous area in the ionosphere depends on the magnitude so that the radius of this area in the ionosphere can be obtained (in km) as

$$r = 10^{0.43M_w} \quad (1)$$

called the Dobrovolsky formula (Dobrovolsky *et al.* 1979, Dogan *et al.* 2011, Liu *et al.* 2009, Pulinets *et al.* 2003, Zakharenkova *et al.* 2007b), whereas other works (e.g. Zakharenkova *et al.* 2007a) use the formula

$$r = e^{M_w} \quad (2)$$

The apparent conflict can be resolved recognizing that $10^{0.43} = 2.69153\dots$ which is very close to $e = 2.71728\dots$ In other words, Eqs. (1) and (2) are powers which have bases differing in 1% and the same exponent (M_w). The differences in the application of Eq. (1) or Eq. (2) for defining the size of the anomalous area (for instance $r = 1023$ km or $r = 1097$ km, respectively, for $M_w = 7$) may be regarded as completely negligible for our purposes.

- According to many published works (e.g., Guo *et al.* 2017, Yao *et al.* 2012, Zakharenkova *et al.* 2007a) we will assume that the center of the abnormal area in the ionosphere can be found in the vertical of the epicenter, although it has to be noted that there is no complete consensus in this point (Liu *et al.* 2009 find the center slightly towards the east, Pulinets and Boyarchuk 2004 towards the equator in middle latitudes, while not in high latitudes).
- TEC anomalies can have a positive or negative sign according to some works (e.g. Lin 2011), whereas other studies indicate a reduction of TEC values around 3-5 days before (Liu *et al.* 2009) or a reduction followed by an increase closer to the shock (Guo *et al.* 2017). This feature seems to be not clear enough to make any solid consideration of it.
- In contrast with natural ionospheric storms, which usually last from several hours (normally more than 8) to a few days (Gopinath and Prince 2018, Paul *et al.* 2018), seismo-

ionospheric disturbances usually last only 3-6 h (Akhoondzadeh and Saradjian 2011, Pulinets *et al.* 2003, Pulinets and Boyarchuk 2004). This feature may permit to solve the oftentimes principal cited drawback that "The core technical problem for earthquake prediction is to distinguish ionospheric disturbances associated with earthquakes from those caused by other factors" (Yao *et al.* 2012).

1.2. *b*-values as earthquake precursors

The Gutenberg-Richter law states that the number of events N with magnitude higher or equal to M_w follow the relationship

$$\log_{10}N = a - bM_w \quad (3)$$

with some constants a and b . As explained before, many works have confirmed and extensively used *b*-value's ability to predict earthquakes (e.g. Asencio-Cortés *et al.* 2018, Florido *et al.* 2015 and 2018, Nuannin *et al.* 2005). As we will see, our experiments also confirm this proven fact.

2. Data

2.1. EQ catalog

The United States Geological Survey earthquake catalog (U.S. Geological Survey 2019) is used to retrieve the set of earthquake events of magnitude M_w 4.0 or higher from January 1, 2011 to December 31, 2018, that is, the last eight complete years. For reasons to be explained in the application, the data for some days (15) before and after the period of interest are also considered. As it has been said before, earthquakes of the lowest magnitudes (which means of the order of M_w 5 or less) will be basically considered as non-significant events, while those of the highest magnitudes (of the order of M_w 7 or more) are the ones we want to consistently forecast. Fig. 1 and Table 1 illustrate the distribution of earthquakes in terms of magnitude during the period of study.

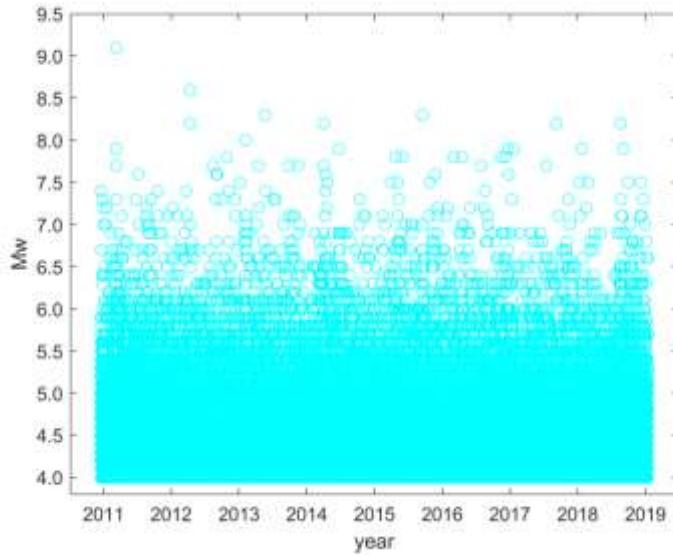


Fig. 1. Earthquake events in terms of their magnitude.

Table 1. Number of earthquake events in the different magnitude intervals (complete years 2011 to 2018).

Min M _w	Max M _w	No. events
4.0	4.4	57911
4.5	4.9	47156
5.0	5.4	10651
5.5	5.9	2753
6.0	6.4	785
6.5	6.9	273
7.0	7.4	82
7.5	7.9	37
8.0	8.4	7
8.5	8.9	1
9.0	9.4	1

2.2. TEC grids

In general, TEC values can be obtained from local probe measurements or remote sounding of the ionosphere by radiowaves thanks to the ionosphere's ability to reflect radiowaves (Pulinets and Boyarchuk 2004). An increasingly popular procedure uses the signals emitted by Global Navigation Satellite Systems (GNSS) like GPS, GLONASS, Galileo or BeiDou as radiowave source for determining TEC values. Given the location of a receiver and a satellite, the received carrier phases of the signals in two (or three) frequencies permit to obtain the slant TEC in the receiver-to-satellite direction, which can be easily converted into vertical TEC. Many works use this approach with a network of GNSS receivers to obtain a set of TEC values for the area of interest (e.g. Dogan 2011, Yao *et al.* 2012, Zakharenkova *et al.* 2007b).

Other works (e.g. Guo *et al.* 2017, Liu *et al.* 2009, Zakharenkova *et al.* 2007a) make use of the TEC values computed from a large number of global International GNSS Service (IGS) stations distributed all over the world and delivered to the user in the form of ionospheric TEC grids (International GNSS Service 2019). In general, there exists a good agreement between the anomalies determined by both approaches, a GNSS network and TEC grids (Şentürk and Çepni 2018a). Among the two existing TEC grid types, 'rapid' and 'final', working with rapid ionospheric TEC grids seems necessary since they offer the same sampling interval (2.5 degrees in latitude and 5 degrees in longitude) with less latency (nominally less than 24 hours instead of 11 days), which is key to a forecast with only few days in advance.

Among the different rapid TEC grid solutions provided by different institutions integrated in the IGS we decided to work with CODE's rapid TEC grids, which for years have been a reliable source of TEC grid solutions due to its long availability and relatively good continuity (e.g. Şentürk and Çepni 2018b). They are available at the NASA site <ftp://cddis.gsfc.nasa.gov/gps/products/ionex> under the corresponding year and day of year folders having file names starting with 'corg'. Their nominal time sampling of 2 h, was however surpassed in practice from November 1, 2014, having the newest files a time sampling of 1 h. These daily files have a latency of several hours (nominally up to 24 h) which for earthquake prediction means that the day prior to an earthquake may not be available for the analysis.

While a dedicated network of continuously operating GNSS reference stations may offer the advantages of a denser coverage of the particular area to be monitored as well as less latency, the use of TEC grids is preferred in this study because of its worldwide coverage and ease of use. It has to be noted, however, that in other applications the use of continuously operating GNSS reference stations may be preferred, especially if aiming at regional or local analyses.

3. Application to the Mw 9.1 Japan earthquake

Now let us take as example the largest earthquake in the period of study: March 11, 2011, UTC 5:46, M_w 9.1 Japan earthquake (epicenter in latitude 38.297° , longitude 142.373°) and try to predict it from the available rapid TEC grids prior to the earthquake. In the present example a bilinear interpolation also known as 4-point bivariate interpolation (Şentürk and Çepni 2018a) has been used to obtain the TEC value above the epicenter from the values of the vertices of the corresponding TEC grid cell. However, more elaborate solutions using different weighting algorithms and different surface interpolation techniques can also be found (Şentürk and Çepni 2019, Şentürk *et al.* 2019).

Fig. 2 shows the vertical TEC values obtained from the TEC grids of the previous days for the latitude and longitude of the epicenter of the posterior event, here indicated with a red diamond on day 11. For the sake of completeness, TEC values for the day of the earthquake and the day before are also shown although it must be taken into account that due to the data latency (up to 24 h) they should not have been available for analysis prior to the earthquake shock.

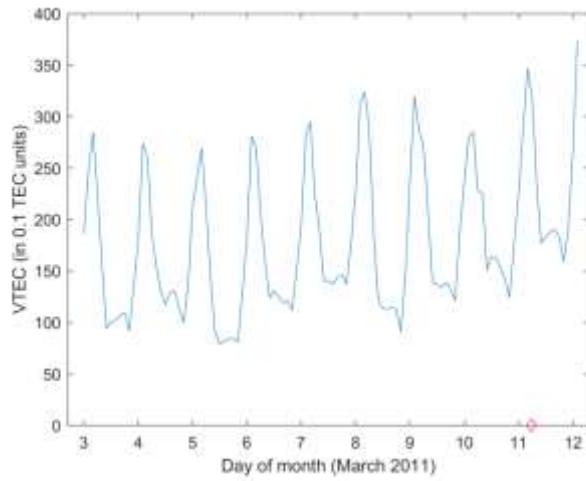


Fig. 2. Vertical TEC values previous to the Japan earthquake (indicated by a red diamond on day 11) at the epicenter (latitude 38.297° , longitude 142.373°).

Fig. 2 is similar to those displayed in Yao *et al.* (2012) for the stations close to the epicenter even though they used some filtering by sliding windows (to compute moving median and standard deviations of 1-10 previous days). Just as Yao *et al.* (2012) indicate, before the day of the earthquake one can only see slightly abnormal high values in days 8 and 9 plus a slightly decreased value in day 5. This slight variations, however, seem to be too subtle to draw solid predictive conclusions. We can also compare these values with the ones in a place detached from the epicenter, e.g. New York city (latitude 40.6° , longitude -74°), depicted in Fig. 3, and see little difference.

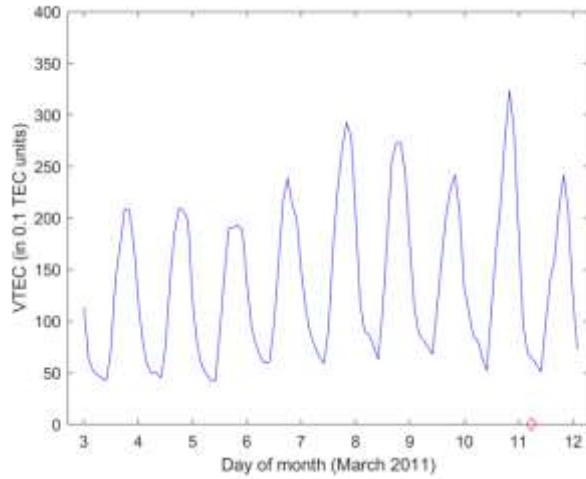


Fig. 3. Vertical TEC values previous to the Japan earthquake (indicated by a red diamond on day 11) at New York city (latitude 40.6° , longitude -74°).

Furthermore, since ionospheric variations can have different origin (mainly space-weather effects, but also seismic activity), it is worth examining the values of solar and geomagnetic activity indicators. The values of the F10.7 (representative of the solar activity), Dst (representative of the severity of magnetic storms) and Kp (global geomagnetic activity) indices, as well as the Interplanetary Magnetic Field (IMF) value, of solar origin, are shown in Fig. 4 as retrieved from the OMNI project web site <https://omniweb.gsfc.nasa.gov/form/dx1.html>.

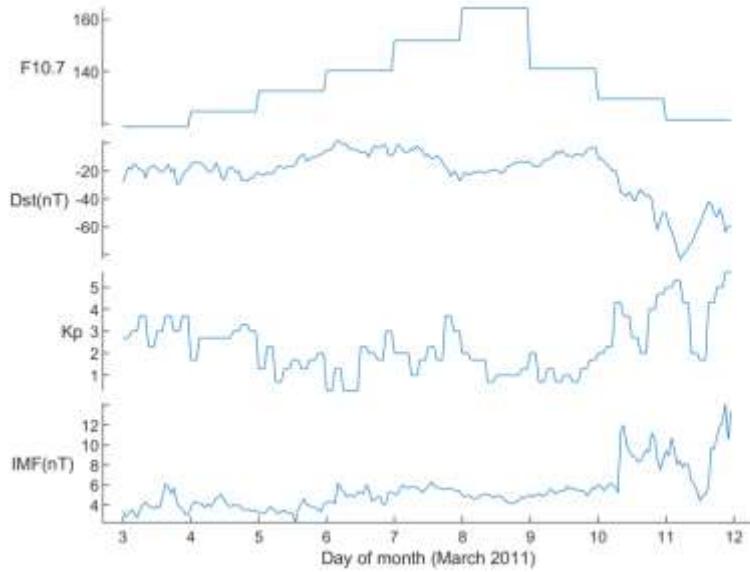


Fig. 4. Solar and geomagnetic activity indicators F10.7, Dst (in nT), Kp and IMF (in nT) previous to the Japan earthquake on day 11.

Apart from the abnormal values in the last day before the earthquake, no significant information can be extracted from these indicators except for the up-to-down trend change in F10.7 index around days 8-9.

We can also see what happens when no significant earthquakes ($M_w < 7$) occur the days after, first, in the case of a geomagnetically calmed period. If we take as example the same place two months after, we obtain the TEC values shown in Fig. 5 which can be compared with the previous Fig. 2 and Fig. 3 recognizing not very significant differences.

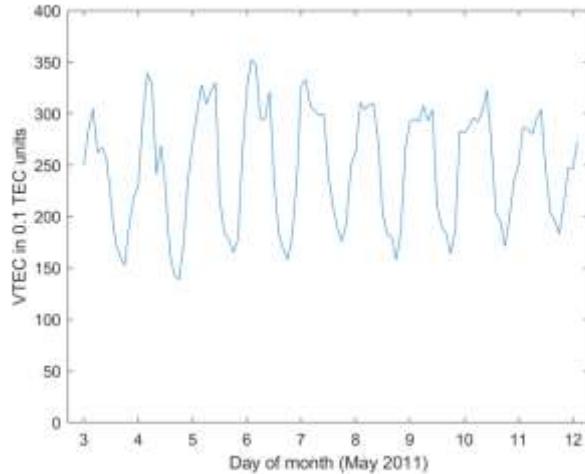


Fig. 5. Vertical TEC values for some days in May 2011 (geomagnetically calm, no significant events) at the epicenter of Japan March 2011 earthquake (latitude 38.297° , longitude 142.373°).

The corresponding indicators of solar and geomagnetic activity for this relatively calm period are also shown in Fig. 6. They show little correlation with the TEC values in Fig. 5, although they all correspond to a time period with no significant earthquakes.

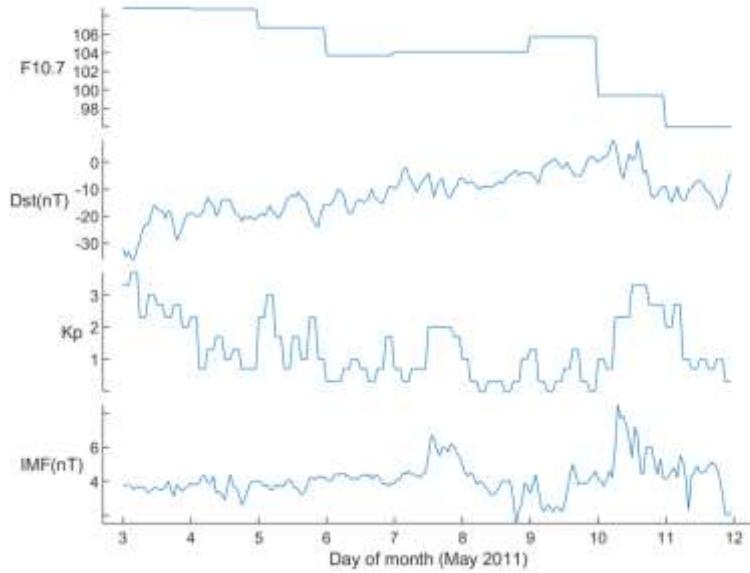


Fig. 6. Solar and geomagnetic activity indicators F10.7, Dst (in nT), Kp and IMF (in nT) for some days in May 2011 (geomagnetically calm, no significant events).

We can also see what happens for a moderately stormy period followed by no significant earthquakes ($M_w < 7$), Fig. 7 and Fig. 8. Not very significant differences are revealed either.

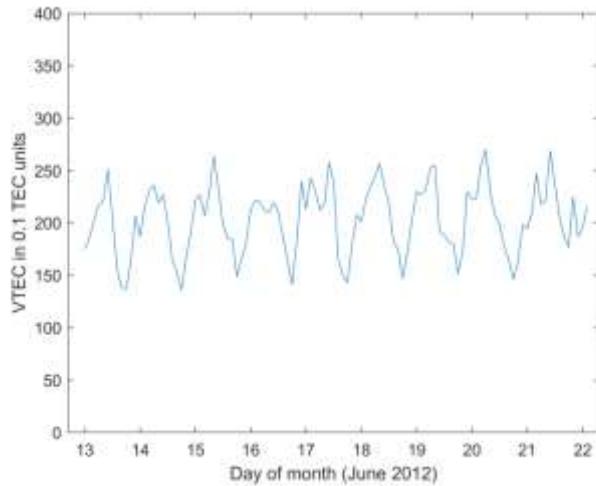


Fig. 7. Vertical TEC values for some days in June 2012 (moderately stormy, no significant events) at the epicenter of Japan March 2011 earthquake (latitude 38.297° , longitude 142.373°).

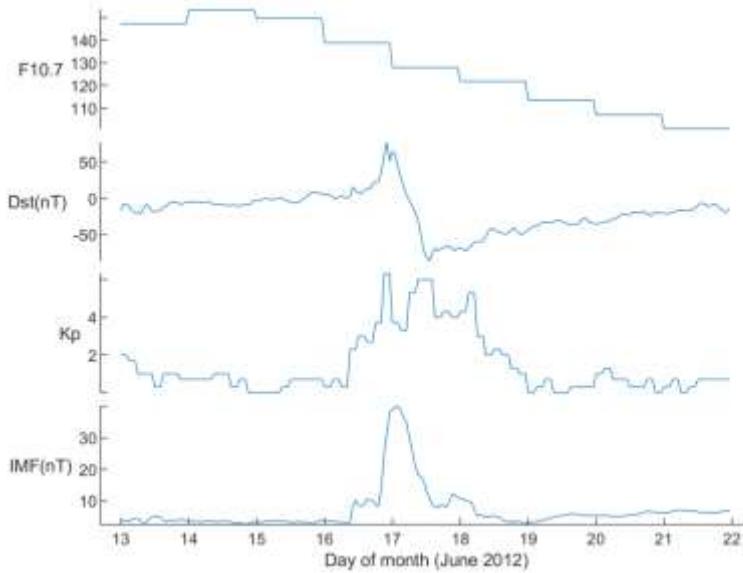


Fig. 8. Solar and geomagnetic activity indicators F10.7, Dst (in nT), Kp and IMF (in nT) for some days in June 2012 (moderately stormy, no significant events).

In summary, examining simple TEC values of the days preceding the M_w 9.1 Japan earthquake we observe so small variations (of similar small amplitude than quiet seismic periods and some of them even correlated with changes in the solar activity index F10.7) that reliably forecasting the occurrence of the earthquake seems not possible. This reinforces the idea that a mere time series of TEC values for a grid cell only is of little value for earthquake forecast. The observed little contrast, which in addition is partly correlated with solar effects, is clearly insufficient for a reliable prediction.

Therefore, we need to amplify the anomalies related to future seismic events only. The key feature to elaborate on is the fact that seismo-ionospheric disturbances usually last only a few hours (Akhoondzadeh and Saradjian 2011, Pulinets et al. 2003, Pulinets and Boyarchuk 2004) whereas ionospheric storms of solar and geomagnetic origin usually last from several hours to a few days (Gopinath and Prince 2018, Paul *et al.* 2018). In the next section we will make direct use of this feature for the construction of our earthquake forecast estimator.

As a note of caution, it should be mentioned that there may indeed be anomalies of geomagnetic origin lasting only a couple of hours, such as these known as "dusk effect" in the literature (e.g. Buonsanto 1999, Kane 2005 and Berényi *et al.* 2018). It is our contention that they can be considered as a relatively rare exception to the general rule, so that separation of solar and geomagnetic ionospheric anomalies from anomalies of seismic origin can generally be made in terms of time duration while a further elaboration of a more satisfactory distinction is left for future research.

4. Method

4.1. Anomaly estimator E

We can significantly improve the contrast by making use of two ideas. First, as said before, the anomalous area in the ionosphere has a considerable extension, given by radius r from Eq. (1), or equivalently Eq. (2), so that it seems appealing to look for anomalies not in a single TEC grid cell but rather in a set of contiguous cells. We may recall that a difference in latitude $\Delta\varphi$ represents a distance d along the meridian given by

$$d = R\Delta\varphi \quad (4)$$

where R is a mean value for the radius of Earth (e.g. 6371 km). Similarly, a difference in longitude $\Delta\lambda$ represents a distance d along the parallel of latitude φ given by

$$d = R \cos \varphi \Delta\lambda \quad (5)$$

Since we are aiming at devising a general procedure to foresee earthquakes of magnitudes M_w higher than 7, we can introduce $M_w = 7$ in Eq. (1), or equivalently Eq. (2), and obtain a value for r that gives the following boundaries for a set of contiguous cells which are expected to have anomalous values (Fig. 9).

$$\Delta\varphi = \frac{r}{R} \quad (6)$$

$$\Delta\lambda = \frac{r}{R \cos \varphi} \quad (7)$$

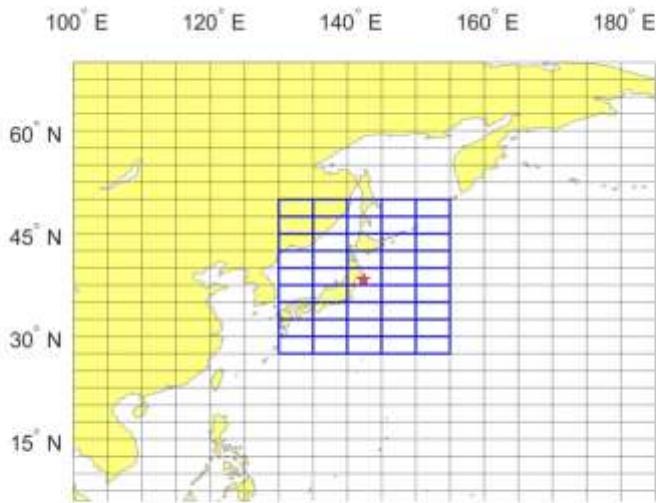


Fig. 9. TEC grid cells contiguous to the epicenter (indicated by a red star) within the boundaries defined by $[\varphi_{\text{epicenter}} - \Delta\varphi, \varphi_{\text{epicenter}} + \Delta\varphi]$ and $[\lambda_{\text{epicenter}} - \Delta\lambda, \lambda_{\text{epicenter}} + \Delta\lambda]$ from Eqs. (6) and (7), highlighted in blue.

Now the question arises how to analyse time variations not for a single cell (as we did e.g. in Fig. 2) but for a set of cells (Fig. 9) and encapsulate the result in a meaningful estimator. Denoting by M the matrix of TEC grid values corresponding to the selected grids (with the rows and columns indicated in Fig. 9), we have in the rectangular matrix M a set of anomalous values (despite there can also be anomalous values outside of this range, e.g. if the magnitude is higher). Matrix M^T times M , (\cdot)^T indicating the transpose of matrix, is a square matrix whose values are also indicative of the observed anomalies. We can propose the use of the determinant of this square matrix bearing in mind that the determinant is proportional to the hypervolume of the hyperparallelepiped defined by the columns of the matrix

$$D = \det(M^T M) \quad (8)$$

The results of computing this estimator D from the selected TEC grid cells (Fig. 9) for the previous days of the earthquake are shown in Fig. 10. As we did in Fig. 2 we have preferred to keep the

original values in the rapid TEC grid files, which are given in 0.1 TEC units, and do not change units for the subsequent analysis. It may be worth noting that for all the computations in the paper we used Matlab R2018b and, in particular, its built-in 'det' function for the computation of Eq. (8).

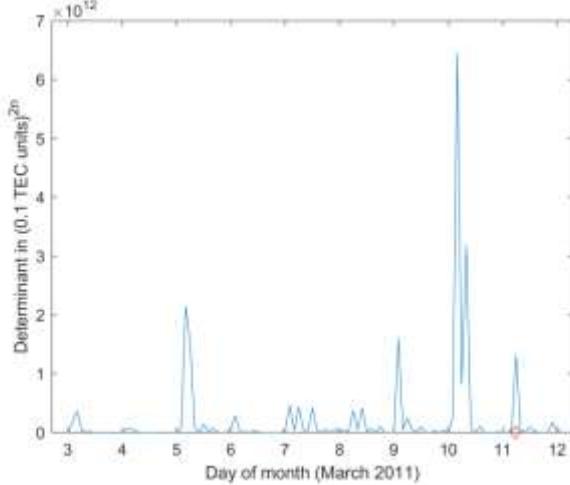


Fig. 10. D values from Eq. (8) previous to the Japan earthquake (indicated by a red diamond on day 11) at the epicenter (latitude 38.297° , longitude 142.373°).

Now we can see a much pronounced contrast, compared to previous Fig. 2. Especially striking are the large values that can be observed in days 5 and 9, apart from the values in day 10 which could have been missed due to the TEC grid latency.

Fig. 11 shows the values of estimator D at a detached place (New York City) for the days preceding the Japan earthquake, Fig. 12 shows the corresponding values at its epicenter (latitude 38.297° , longitude 142.373°) for the quiet period in May and Fig. 13 for the moderately stormy in June 2012. By contrast with the minimum differences in simple TEC values among the days preceding the Japan earthquake at the epicenter, Fig. 2, and those for a different place, Fig. 3, or time, Fig. 5 and Fig. 7, now we can see much more evident differences between D values preceding the earthquake at the epicenter (Fig. 10) and D values for the same time period and a different place (Fig. 11) or the same place and a different time period (geomagnetically calm, Fig. 12 ,or moderately stormy, Fig. 13).

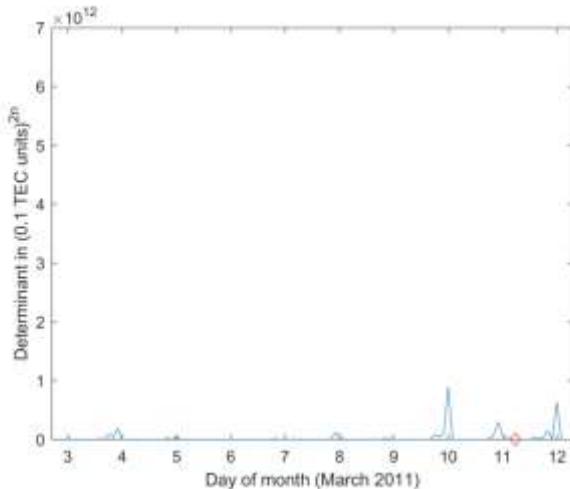


Fig. 11. D values from Eq. (8) previous to the Japan earthquake at New York city (latitude 40.6° , longitude -74°).

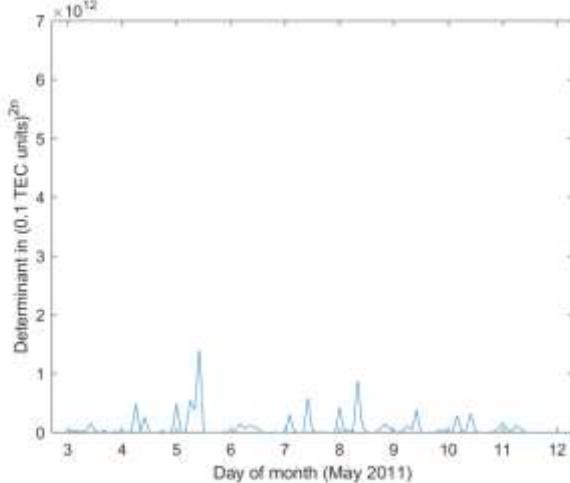


Fig. 12. D values from Eq. (8) for some days in May 2011 (geomagnetically calm, no significant events) at the epicenter of Japan March 2011 earthquake (latitude 38.297° , longitude 142.373°).

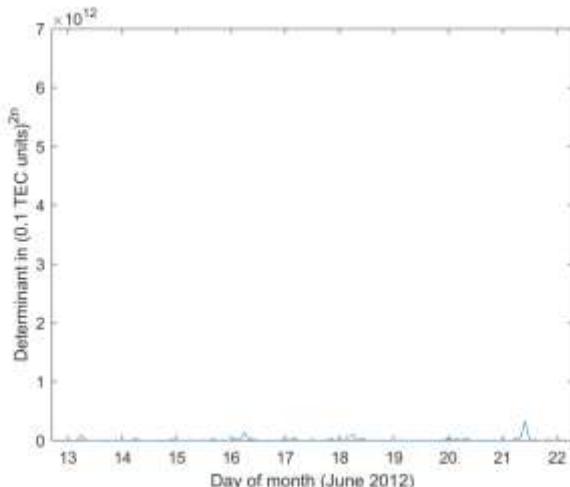


Fig. 13. D values from Eq. (8) for some days in June 2012 (geomagnetically stormy, no significant events) at the epicenter of Japan March 2011 earthquake (latitude 38.297° , longitude 142.373°).

Other places and times yield similar figures. Obviously it is clear that a limited number of trials may serve to disprove an idea but never to soundly demonstrate its validity. We will remedy this limitation in the next section.

Now, estimator D shows spatial variations of TEC values. However, we cannot be satisfied with evaluating spatial variations only, as we have done with this estimator. We want to quantify their time variation or, more specifically, construct an estimator which quantifies space-time variations of TEC values highlighting large variations of short duration (typically less than 4 or 6 hours) since, as said before, they are the imprint of forthcoming earthquakes which permit to tell them apart from other anomalous structures of solar or geomagnetic origin, which last longer. We propose to subtract consecutive D values, bearing in mind that TEC grids are given with a time resolution of hours (1 or 2, as explained in Section 2.2) and we are looking for a rapidly changing phenomenon with a lifespan of few hours, as the one shown in Fig. 10. We can take relative differences between consecutive D values (hence the denominator in the equation) and use an exponent higher than one to amplify even more the differences. We have eventually found that the sum of coefficients of range to the forth power, that is estimator E in Eq. (9), highlights the type of anomalies we are searching for.

$$E = \frac{1}{n-1} \sum_{i=1}^{i=n-1} \left(\frac{\max - \min}{\max + \min} \right)^4 \quad (9)$$

where $\max = \max\{D_i, D_{i+1}\}$ and $\min = \min\{D_i, D_{i+1}\}$ for D values in time i and $i+1$, respectively (n being the total number of matrices M and corresponding D values).

A whole analysis, defined by a 'current day' and a number of previous days whose TEC grids are to be analyzed, can be now represented by a single numeric value E . The next section is dedicated to thoroughly test the usefulness of estimator E along the period of study (January 1, 2011 to December 31, 2018).

Regarding the conception of Eq. (9), our experiments showed that using an exponent of 2 instead of 4, that is using the sum of squared coefficients of range, we obtained similar results although with a smaller contrast between significant and non-significant episodes. The contrast was even less by using the sum of simple coefficients of range.

We want to emphasize, however, that Eq. (9) is simply the best choice that emerged from a vast trial-and-error search in which we tried many ideas including different coefficients of dispersion, variances and higher-order moments (skewness, kurtosis...), standard versus robust measures (e.g. median absolute deviation about the median), autocorrelation, Allan variance, generalized entropy indices, efficiency, amplitude scintillation index, fractal volatility and a long etcetera of little successful approaches with or without prior smoothing by 15-day moving averages (including comparisons with respect to a less affected far-distance exterior area or a limited effort to design and use neural networks under the Deep Learning Matlab toolbox). They all seemed to result little helpful so far, which does not mean that we have exhausted all possibilities and they could not be eventually put to work successfully by other authors using different design choices. Further, new approaches recently proposed, such as the use of cross-wavelet and short-time Fourier transforms (Şentürk *et al.* 2019), may be worth testing in future research.

4.2. Forecasting 2011-2018 worldwide earthquakes

We want to test estimator E , Eq. (9), as earthquake forecast estimator for the large earthquakes ($M_w \geq 7$) occurred worldwide from January 1, 2011 to December 31, 2018, recall section 2.1, esp. Fig. 1 and Table 1, imagining that each sample day represents the present day in the future situation where the methodology will be able to be implemented in a real-time earthquake forecast system.

In order to avoid undesired biases we will repeatedly apply the following testing procedure:

- 1) Compute the corresponding radius of the anomalous area, r , by means of Eq. (1) or (2). Here we use $M_w = 7$ taking into account that our focus is to forecast earthquakes of this magnitude or higher.
- 2) Set the "number of predicting days" and the "number of days ahead in the future". We use here 7 and 4, respectively, meaning that we will use TEC data from 7 days before the date of interest in order to forecast the earthquakes in the near future defined as the date of interest (still lacking the TEC data grids) plus the subsequent 4 days (5 days in total).
- 3) Choose at random one day belonging to the time interval. We want to test the estimator's ability to predict earthquakes in the near future from stable seismic periods only. That is, we want to discard the cases where an earthquake has occurred in the very recent past (i.e. during the days corresponding to the TEC data) since it would be easy to conclude, often successfully, that an earthquake occurs when a large earthquake has happened in the very recent past. In other words, we want to rule out the success related to aftershocks. The condition to be fulfilled for the day randomly selected is that

$$\max(M_w)_{future} - \max(M_w)_{past} \geq 1 \quad (10)$$

In other words, the maximum magnitude of any earthquake occurred anywhere on Earth during the "number of days ahead in the *future*" has to be at least of magnitude higher than 1 than the maximum magnitude of any earthquake occurred anywhere on Earth during the predicting days (*past* days). Otherwise repeat this step 3) until the condition in Eq. (10) is fulfilled. The earthquake with maximum magnitude in the "number of days ahead in the future" becomes the event to be forecast.

4) Given the selected earthquake with epicenter $(\varphi_{epicenter}, \lambda_{epicenter})$ we assume, as explained before, that there are some fingerprints in the data corresponding to the predicting days and predicting area that will lead us to conclude the existence of the earthquake in the future. We form the set of matrices M for space boundaries $[\varphi_{epicenter} - \Delta\varphi, \varphi_{epicenter} + \Delta\varphi], [\lambda_{epicenter} - \Delta\lambda, \lambda_{epicenter} + \Delta\lambda]$ and time limits [present day – number of predicting days, present day], where Eqs. (6) and (7) are used to obtain the increments of latitude and longitude, respectively, along with the r value from step 1).

5) Compute estimator E by means of Eq. (9) for the matrix time series of n matrices M . As we will see next, the value of this estimator E is somehow related to the magnitude of the forthcoming earthquake but this relationship is not straightforward. So, we will also devise first (and then be able to compute) a new estimator yielding the predicted earthquake magnitude by maximizing the correlation with the actual earthquake magnitude.

Since the conclusions for a single day only are not statistically significant, we go back to step 3) and keep on choosing days at random and computing the corresponding estimator E values until we obtain a sufficiently large sample.

To compose Fig. 14, 400 days have been drawn at random. Only a few percentage of them meet the strong requirement in step 3) of stable seismic period prior to the earthquake. This is necessary if we want to rule out aftershock prediction. As it can be seen the E values obtained from predicting days and the M_w values of future earthquakes have a positive dependence. In fact, the analysis of correlation yields a correlation coefficient of 0.56228 with a p -value of 0.023381. Since this p -value is less than the significance level of 0.05, the hypothesis of no correlation between the two variables is rejected.

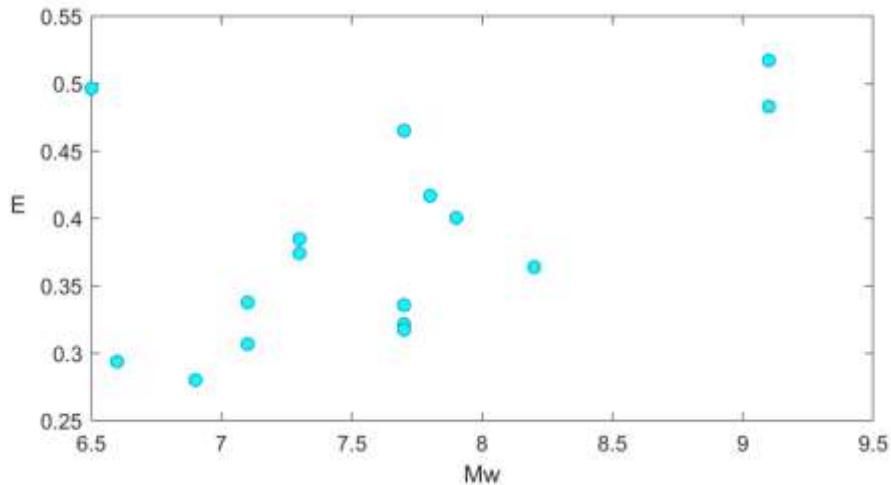


Fig. 14. E values from Eq. (9) using predicting days versus actual magnitude M_w of the future earthquake.

Other random sets of days can be formed and the corresponding analysis be performed yielding similar results. It has to be noted that occasionally TEC grid files contain a column with exactly the same values as an adjoining one (not only for the prescribed latitude boundaries but for the entire latitude values) often not for a single hour but for consecutive hours and days. The origin of this issue is unknown, it might be due to missing information resolved by simply copying values when

the TEC grid file was composed. In any case, this linear dependence of columns produces a rank defect in matrix M and, correspondingly, a null value for the determinant D and, if found in consecutive epochs, unknown values for estimator E . For the moment we have simply discarded the cases when this problem is encountered.

4.2.1. Enhancing prediction by means of b-values

For each sample day we use the predicting days only as data to count the N values corresponding to the intervals of magnitude $[6.5, 6.99]$, $[7.0, 7.49]$, $[7.5, 7.99]$, ... , $[9.0, 9.49]$ and solve the overdetermined system of equations of type Eq. (3) to obtain the b -value. The results for the same data sample used before are shown in Fig. 15.

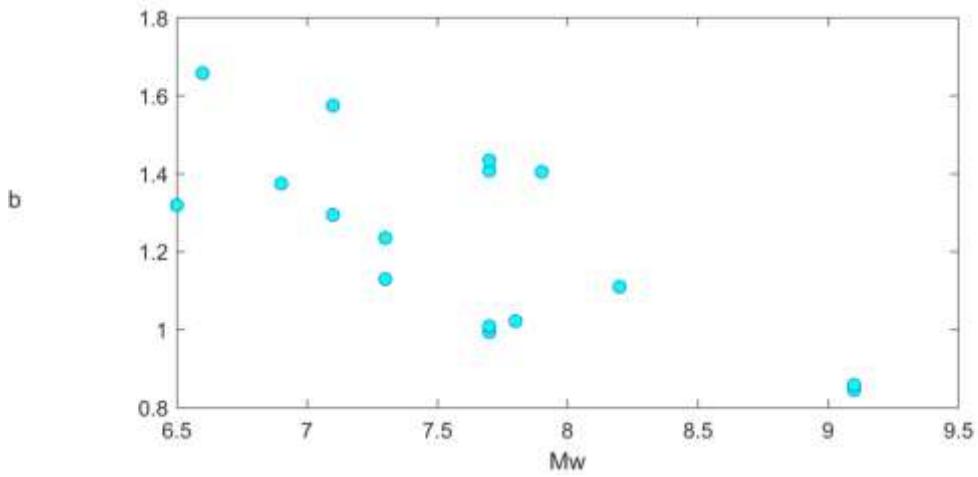


Fig. 15. b -values obtained using predicting days versus actual magnitude M_w of the future earthquake.

As it can be seen there is a clear negative dependence between b -values obtained from predicting days only and the M_w values of future earthquakes. The analysis of correlation yields a correlation coefficient of -0.72832 and a p -value of 0.001377, which is clearly below the 0.05 significance level for the test of no correlation between variables. In other words, the correlation is very much significant, even higher than the correlation between the E values derived from TEC grids and the future earthquake magnitudes.

The key observation here is that this b -value has been computed from the seismic record only, hence completely independent from TEC measurements, which allows us to use both b and E values together to construct a more robust earthquake forecast estimator.

After having realized that there is a positive linear correlation between E values, Eq. (9), and future M_w values, and testing other alternative models with no success, we came to the conclusion that E and M_w should be best related by a linear expression of the type $M_w = c_1 E + c_2$. Since there is an inverse (or negative) correlation between b and M_w values, the best relation between them seems to be an inverse relationship, not necessarily with b to the first power as our simulations indicated, that is an expression of the type $M_w = c_3/b^{c_4} + c_5$. The joint model obtained by simple addition of these contributions happened to work worse in the subsequent global optimization procedure (that is, having a larger standard deviation of differences between predicted and actual magnitude values) than the little-different joint model defined as

$$M_w = \frac{c_1 E}{b^{c_2}} + c_3 \quad (11)$$

with coefficients c_1 , c_2 and c_3 to be obtained by means of a global optimization procedure in order to minimize the differences with respect to actual M_w values. We will not extend our presentation much here, especially since the global optimization problem can be solved equally well by other

different methods, such as Genetic Algorithms (e.g. Shiuly and Roy 2018), and many different tools easily accessible (the Global Optimization Toolbox under Matlab or Solver in Microsoft Excel, for example). Just suffice to say that we opted for the Simulated Annealing Method, which is readily found well documented in the literature (e.g. Pardalos and Romeijn 2002) and has been extensively used in the last years in the field of geosciences (e.g. Abordán and Szabó 2018, Baselga 2018 and 2019, Biswas and Sharma 2017). Also worth mentioning is the fact that in order for the estimator to be as representative as possible we made use of hundreds of randomly selected days. The resulting earthquake forecast model is

$$M_w = \frac{2.670236E}{b^{2.370606}} + 6.8803 \quad (12)$$

Or with some rounding-off which does not significantly affect its performance

$$M_w = \frac{2.7E}{b^{2.4}} + 6.9 \quad (13)$$

which resulted in a standard deviation of differences between predicted and actual magnitude values of 0.445.

Computing E and b values as explained above, predicted M_w values for forthcoming earthquakes can be computed by means of Eq. (13). Including this computation in step number 5 of the testing procedure explained at the beginning of Section 4 we can observe the performance of the prediction: Fig. 16 shows the results for the data sample used in all previous figures.

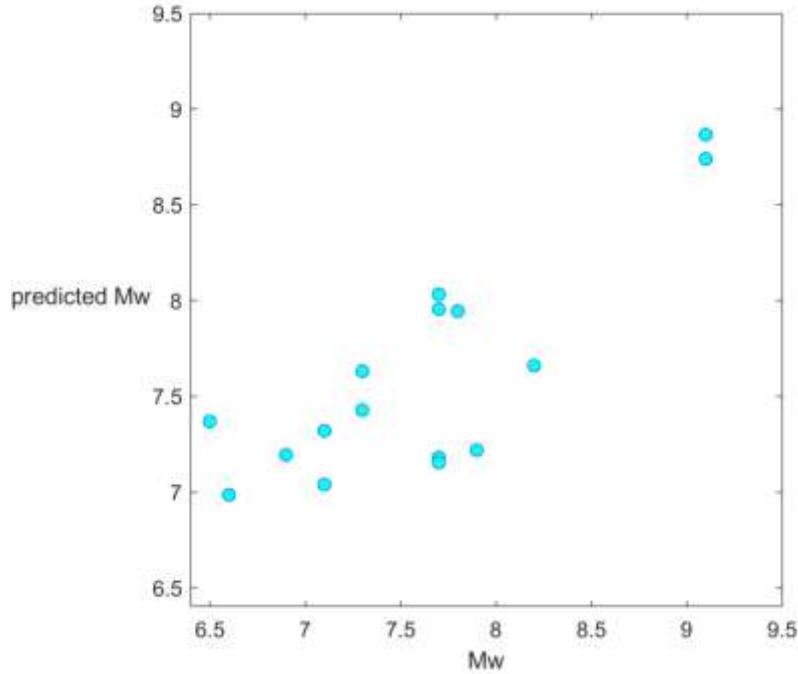


Fig. 16. Predicted magnitude M_w versus actual magnitude M_w of the future earthquake.

The analysis of correlation now yields a correlation coefficient of 0.84754 with a p -value 0.0000034, which is some orders of magnitude less than those previously obtained (Figs. 14 and 15) and represents the minuscule probability of having obtained this result by pure chance assuming there was no real correlation.

5. Conclusions and future work

Just as some previous works show, we have seen that anomalous TEC values may be found in the days prior to an earthquake. As an original contribution we showed that the predictive power can be improved if a suitable space-time TEC variation estimator is used instead of simple TEC values.

After experiencing the well-known high predictive abilities of Gutenberg-Richter law's *b*-values, we showed that earthquake prediction can be further enhanced by developing an original estimator based on both TEC anomalies and *b*-values, Eq. (13). The results seem to be very promising when applied to the large earthquakes (that is of magnitude M_w 7 or higher) occurred worldwide during the period January 1, 2011 to December 31, 2018. Future work will have to address the possibility of designing a real-time earthquake forecast service based on this estimator.

Apart from the expected implementation problems of such a real-time system, one will have to cope with the already experienced problems in the cases where the TEC grids contain duplicated columns in the data (i.e. exactly the same values for neighboring longitudes). In the present work, we decided to simply discard these cases but whether or not it is possible to repair the data and how to do it is a matter for future research.

Acknowledgments

The author is sincerely thankful to the editor and two anonymous reviewers for their detailed comments and suggestions that helped to considerably improve the manuscript.

References

- Abordán A, Szabó NP (2018) Metropolis algorithm driven factor analysis for lithological characterization of shallow marine sediments. *Acta Geod Geophys* 53: 189–199. <https://doi.org/10.1007/s40328-017-0210-z>
- Akhoondzadeh M, Saradjian MR (2011) TEC variations analysis concerning Haiti (January 12, 2010) and Samoa (September 29, 2009) earthquakes. *Adv Space Res* 47(1): 94–104. <https://doi.org/10.1016/j.asr.2010.07.024>
- Asencio-Cortés G, Morales-Esteban A, Shang X, Martínez-Álvarez F (2018) Earthquake prediction in California using regression algorithms and cloud-based big data infrastructure. *Comput Geosci* 115: 198–210. <https://doi.org/10.1016/j.cageo.2017.10.011>
- Baselga S (2018) Fibonacci lattices for the evaluation and optimization of map projections. *Comput Geosci* 117: 1–8. <https://doi.org/10.1016/j.cageo.2018.04.012>
- Baselga S (2019) TestGrids: Evaluating and Optimizing Map Projections. *Journal of Surveying Engineering* 144(3): *in press*.
- Berényi KA, Barta V, Kis Á (2018) Midlatitude ionospheric F2-layer response to eruptive solar events-caused geomagnetic disturbances over Hungary during the maximum of the solar cycle 24: A case study. *Adv Space Res* 61(5): 1230–1243. <https://doi.org/10.1016/j.asr.2017.12.021>
- Biswas A, Sharma SP (2017) Interpretation of self-potential anomaly over 2-D inclined thick sheet structures and analysis of uncertainty using very fast simulated annealing global optimization. *Acta Geod Geophys* 52: 439–455. <https://doi.org/10.1007/s40328-016-0176-2>
- Borgohain JM, Borah K, Biswas R, Bora DK (2018) Seismic b-value anomalies prior to the 3rd January 2016, M_w =6.7 Manipur earthquake of northeast India. *J Asian Earth Sci* 154: 42–48. <https://doi.org/10.1016/j.jseaes.2017.12.013>
- Buonsanto M (1999) Ionospheric Storms – A Review. *Space Sci Rev* 88: 563–601. <https://doi.org/10.1023/A:1005107532631>
- Buskirk RE, Frohlich CL, Latham GV (1981) Unusual animal behavior before earthquakes: a review of possible sensory mechanisms. *Rev Geophys* 19: 247–270. <https://doi.org/10.1029/RG019i002p00247>

- Dobrovolsky IR, Zubkov SI, Myachkin VI (1979) Estimation of the size of earthquake preparation zones. *Pure Appl Geophys* 117: 1025–1044. <https://doi.org/10.1007/BF00876083>
- Dogan U, Ergintav S, Skone S, Arslan N, Oz D (2011) Monitoring of the ionosphere TEC variations during the 17th August 1999 Izmit earthquake using GPS data. *Earth Planets Space* 63(12): 1183–1192. <https://doi.org/10.5047/eps.2011.07.020>
- Florido E, Martínez-Álvarez F, Morales-Esteban A, Reyes J, Aznarte-Mellado JL (2015) Detecting precursory patterns to enhance earthquake prediction in Chile. *Comput Geosci* 76: 112–120. <https://doi.org/10.1016/j.cageo.2014.12.002>
- Florido E, Asencio-Cortés G, Aznarte JL, Rubio-Escudero C, Martínez-Álvarez F (2018) A novel tree-based algorithm to discover seismic patterns in earthquake catalogs. *Comput Geosci* 115: 96–104. <https://doi.org/10.1016/j.cageo.2018.03.005>
- Freund FT, Kulahci, IG, Cyr G, Ling J, Winnick M, Tregloan-Reed J, Freund MM (2009) Air ionization at rock surfaces and pre-earthquake signals. *J Atmos Sol-Terr Phys*, 71(17-18), 1824–1834. <https://doi.org/10.1016/j.jastp.2009.07.013>
- Gopinath S, Prince PR (2018) Nonextensive and distance-based entropy analysis on the influence of sunspot variability in magnetospheric dynamics. *Acta Geod Geophys* 53: 639–659. <https://doi.org/10.1007/s40328-018-0235-y>
- Grant RA, Halliday T (2010). Predicting the unpredictable; evidence of pre-seismic anticipatory behaviour in the common toad. *J. Zool.* 281: 263–271. <https://doi.org/10.1111/j.1469-7998.2010.00700.x>
- Grant RA, Halliday T, Balderer WP, Leuenberger F, Newcomer M, Cyr G, Freund FT (2011) Ground Water Chemistry Changes before Major Earthquakes and Possible Effects on Animals. *Int J Environ Res Public Health* 8: 1936–1956. <https://doi.org/10.3390/ijerph8061936>
- Guo J, Yu H, Li W, Liu X, Kong Q, Zhao C (2017) Total Electron Content Anomalies Before Mw 6.0+ Earthquakes in the Seismic Zone of Southwest China Between 2001 and 2013. *J Test Eval* 45(1): 131–139. <https://doi.org/10.1520/JTE20160032>
- International GNSS Service (2019) IGS products. <https://www.igs.org/products/>, last accessed May 5, 2019.
- Kane RP (2005) Ionospheric foF2 anomalies during some intense geomagnetic storms. *Ann. Geophys* 23: 2487–2499. <https://doi.org/10.5194/angeo-23-2487-2005>
- Kulhanek O, Persson L, Nuannin P (2018) Variations of b-values preceding large earthquakes in the shallow subduction zones of Cocos and Nazca plates. *J South Am Earth Sci* 82: 207–214. <https://doi.org/10.1016/j.jsames.2018.01.005>
- Lin JW (2010) Ionospheric total electron content (TEC) anomalies associated with earthquakes through Karhunen-Loéve Transform (KLT). *Terr Atmos Ocean Sci* 21(2): 253–265. [https://doi.org/10.3319/TAO.2009.06.11.01\(T\)](https://doi.org/10.3319/TAO.2009.06.11.01(T))
- Lin JW (2011) Latitude-Time Total Electron Content Anomalies as Precursors to Japan's Large Earthquakes Associated with Principal Component Analysis. *Int J Geophys*, Article ID 763527, 12 pp. <http://dx.doi.org/10.1155/2011/763527>
- Liu JY, Chen YI, Chuo YJ, Chen CS (2006) A statistical investigation of preearthquake ionospheric anomaly. *J Geophys Res* 111: A05304, 5 pp. <https://doi.org/10.1029/2005JA011333>
- Liu JY, Chen YI, Chen CH, Liu CY, Chen CY, Nishihashi M, Li JZ, Xia YQ, Oyama KI, Hattori K., Lin CH (2009) Seismoionospheric GPS total electron content anomalies observed before the 12 May 2008 Mw7.9 Wenchuan earthquake. *J Geophys Res* 114: A04320, 10 pp. <https://doi.org/10.1029/2008JA013698>
- Nuannin P, Kulhanek O, Persson L (2005) Spatial and temporal b value anomalies preceding the devastating off coast of NW Sumatra earthquake of December 26, 2004. *Geophys Res Lett* 32: L11307, 4 pp. <https://doi.org/10.1029/2005GL022679>
- Pardalos PM, Romeijn HE (Eds.) (2002) *Handbook of Global Optimization* Vol. 1&2, Kluwer Academic, Dordrecht.
- Paul B, De BK, Guha A (2018) Latitudinal variation of F-region ionospheric response during three strongest geomagnetic storms of 2015. *Acta Geod Geophys* 53: 579–606. <https://doi.org/10.1007/s40328-018-0221-4>
- Pulinets SA, Legen'ka AD, Gaivoronskaya TV, Depuev VKh (2003) Main phenomenological features of ionospheric precursors of strong earthquakes. *J Atmos Sol-Terr Phys* 65: 1337–1347. <https://doi.org/10.1016/j.jastp.2003.07.011>

- Pulinets S, Boyarchuk K (2004) Ionospheric Precursors of Earthquakes. Springer-Verlag, Berlin Heidelberg.
- Reyes J, Morales-Esteban A, Martínez-Álvarez F (2013) Neural networks to predict earthquakes in Chile. *Appl Soft Comput* 13: 1314–1328. <https://doi.org/10.1016/j.asoc.2012.10.014>
- Şentürk E, Çepni MS (2018) A statistical analysis of seismo ionospheric TEC anomalies before 63 Mw ≥ 5.0 earthquakes in Turkey during 2003–2016. *Acta Geophys* 66:1495–1507. <https://doi.org/10.1007/s11600-018-0214-2>
- Şentürk E, Çepni MS (2018) Ionospheric temporal variations over the region of Turkey: a study based on long-time TEC observations. *Acta Geod Geophys* 53: 623–637. <https://doi.org/10.1007/s40328-018-0233-0>
- Şentürk E, Çepni MS (2019) Performance of different weighting and surface fitting techniques on station-wise TEC calculation and modified sine weighting supported by the sun effect. *J Spat Sci* 64(2): 209–220. <https://doi.org/10.1080/14498596.2017.1417169>
- Şentürk E, Livaoglu H, Çepni MS (2019) A Comprehensive Analysis of Ionospheric Anomalies before the Mw 7.1 Van Earthquake on 23 October 2011. *J Navig* 72(3): 702–720. <https://doi.org/10.1017/S0373463318000826>
- Shiuly A, Roy N (2018) A generalized VS–N correlation using various regression analysis and genetic algorithm. *Acta Geod Geophys* 53: 479–502. <https://doi.org/10.1007/s40328-018-0220-5>
- U.S. Geological Survey (2019). Earthquake Catalog. <https://earthquake.usgs.gov/earthquakes/search/>, last accessed May 5, 2019.
- Warwick JW, Stoker C, Meyer TR (1982) Radio emission associated with rock fracture: Possible application to the Great Chilean Earthquake of May 22, 1960. *J Geophys Res Solid Earth* 87: 2851–2859. <https://doi.org/10.1029/JB087iB04p02851>
- Yao Y, Chen P, Wu H, Zhang S, Peng W (2012) Analysis of ionospheric anomalies before the 2011 M w 9.0 Japan earthquake. *Chin Sci Bull* 57(5): 500–510. <https://doi.org/10.1007/s11434-011-4851-y>
- Zakharenkova IE, Shagimuratov II, Krankowski A (2007) Features of the ionosphere behavior before the Kythira 2006 earthquake. *Acta Geophys* 55(4): 524–534. <https://doi.org/10.2478/s11600-007-0031-5>
- Zakharenkova IE, Shagimuratov II, Krankowski A, Lagovsky AF (2007) Precursory phenomena observed in the total electron content measurements before great Hokkaido earthquake of September 25, 2003 (M = 8.3). *Stud Geophys Geod* 51(2): 267–278. <https://doi.org/10.1007/s11200-007-0014-7>