J. Space Weather Space Clim. 2021, **11**, 43
© V. Fabbro et al., Published by EDP Sciences 2021
https://doi.org/10.1051/swsc/2021024

JSWSC

Available online at:
www.swsc-journal.org

RESEARCH ARTICLE OPEN ACCESS

GNSS positioning error forecasting in the Arctic: ROTI and Precise Point Positioning error forecasting from solar wind measurements

Vincent Fabbro[1,*], Knut Stanley Jacobsen[2], Yngvild Linnea Andalsvik[2], and Sebastien Rougerie[3]

[1] ONERA/DEMR, Université de Toulouse, 31055 Toulouse, France
[2] Norwegian Mapping Authority, PO 600 Sentrum, 3507 Hønefoss, Norway
[3] CNES, 31400 Toulouse, France

Received 25 September 2020 / Accepted 4 June 2021

Abstract – A model forecasting ionospheric disturbances and its impact on GNSS positioning is proposed, called HAPEE (High lAtitude disturbances Positioning Error Estimator). It allows predicting ROTI index and corresponding Precise Point Positioning (PPP) error in Arctic region (i.e. latitudes > 50°). The model is forecasting for the next hour a probability of a disturbance index or PPP error to exceed a given threshold, from solar wind conditions measured at L1 Lagrange point. Or alternatively, it is forecasting a disturbance index level that is exceeded during the next hour for a given percentage of the time. The ROTI model has been derived from NMA network measurements, considering a database covering the years 2007 up to 2019. It is demonstrated that the statistical variability of the ROTI index is mainly following a lognormal distribution. The proposed model has been tested favorably on measurements performed using measurements from stations of the NMA network that were not used for the model derivation. It is also shown that the statistics of PPP error conditioned by ROTI is following a Laplace distribution. Then a new compound model has been proposed, based on a conditional probability combining ROTI distribution conditioned by solar wind conditions and error distributions conditioned by ROTI index level.

Keywords: Space weather / positioning system / ionosphere (auroral) / ionospheric disturbances / statistics and probability

1 Introduction

High-latitude regions, in the vicinity of the auroral oval, frequently experience ionospheric disturbances which leads to cycle slips, loss of lock and positioning errors. These effects are feared by the community using GNSS applications, notably air navigation that requires a high level of accuracy and integrity. To quantify disturbance effects, different indices are commonly used to characterize the impact of ionospheric fluctuations on GNSS signal degradation in amplitude and in phase. In the polar region, where phase and slant total electron content (STEC) fluctuations are particularly important, σ_ϕ and ROTI (Rate Of Total Electron Content Index) indices are relevant to study such disturbances events (Cherniak et al., 2015; Jacobsen & Andalsvik, 2016). However, their forecasting remains a challenge because our understanding of the causes and consequences of the complex physics involving ionospheric storms is far from complete.

In Space Weather research, numerous magnetic indices exist and the relation between magnetic indices and ionosphere scintillation indices can be researched. Despite the complexity, some authors have investigated the topic. In (Hu et al., 1998), a physical model for ionospheric storm prediction is proposed, based on Dst index to monitor storm onset. The model is proposing short term (in the hours following a storm) predictions of ionospheric storm effects on HF communications.

Fortunately, measurements performed from spacecraft placed at the L1 Lagrange point between the Sun and the Earth allow getting information on the solar wind approximately 1 h upstream before it reaches the Earth. This type of measurement allows proposing new forecasting approaches such as the OVATION (Oval Variation, Assessment, Tracking, Intensity and Online Nowcasting) model (Newell et al., 2010, 2014). The OVATION model forecasts the average characteristic energy of precipitating electrons and ions incoming in ionosphere layers in the polar region, as well as the total energy flux (North or South poles) and total number flux.

*Corresponding author: vincent.fabbro@onera.fr

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Forecasting models of Space Weather impact on ionosphere can also be found in the literature. In (Tsagouri & Belehaki, 2015) the SWIF model (Solar Wind driven auto regression model for Ionospheric short-term Forecast) is presented, producing nowcasting and forecasting of foF2 disturbances for the European region. The SWIF model is an empirical temporal model of ionospheric storms using as proxy the solar wind parameters from ACE spacecraft (Koutroumbas et al., 2008). The model predicts a global occurrence of time disturbances of foF2 due to ionospheric storms.

Prikryl et al. (2012, 2013) are also using as proxy observations at the L1 point from ACE spacecraft to model phase scintillation occurrence at high latitude, from L1 GPS data collected with the Canadian High Arctic Ionospheric Network (CHAIN). They propose a method similar to the McPherron & Siscoe (2004) one, construed by analogy to an approach used in weather meteorology forecasting and based on atmospheric air mass concept.

Lastly, deep learning approaches have also been applied in this context, as for example the work proposed by McGranaghan et al. (2018). They have used CHAIN network scintillation measurements, Solar Wind and Geomagnetic Activity Data, and Particle Precipitation dataset to develop a predictive model for high-latitude ionospheric phase scintillation based on support vector machine (SVM) machine learning algorithm.

To the best of our knowledge, there is no published prediction model of GNSS positioning error at high latitude. In this work, the precise point positioning (PPP) technique has been selected to describe GNSS error.

PPP is a processing strategy for GNSS observations that enables the efficient computation of high-quality coordinates, utilizing undifferenced dual-frequency code and phase observations by using precise satellite orbit and clock products. More detailed descriptions of PPP can be found in e.g. Zumberge et al. (1997) and Kouba & Héroux (2001).

Previous studies by Tiwari et al. (2009) and Moreno et al. (2011) have examined the effects of ionospheric disturbances on PPP calculations at low/equatorial latitudes. Moreno et al. (2011) concluded that the presence of large ROT (Rate Of Change of TEC) can induce a significant degradation of the position estimation.

The Norwegian Mapping Authority (NMA) operates a national network of GNSS receivers, which is used for positioning services and various studies. Researchers at the NMA have previously studied the statistics of ionospheric disturbances and their correlation with GNSS positioning errors in Jacobsen & Dähnn (2014), and published case studies on the effects of space weather events on GNSS in Jacobsen & Schäfer (2012), Andalsvik & Jacobsen (2014) and Jacobsen & Andalsvik (2016).

In this study, a global statistical approach for short term forecasting of ROTI ionospheric disturbances index and associated GNSS PPP error is proposed for high latitudes. It is based on statistical regression of large datasets. Because ionospheric disturbances at high latitude is a highly random process, a statistical approach is the preferred strategy. Moreover, a statistical approach allows introducing the concept of relative risk, associated to the definition of a threshold level in ionospheric disturbance indices, positioning error, or other metrics.

This paper is organized as follows. In Section 2, the NMA database and Solar wind data used for HAPEE model regression are presented. Section 3 presents the ROTI index forecasting model, its solar wind driver, the model regression and its analysis. In Section 4, a validation of the ROTI forecasting is presented. Section 5 details the conditional probability approach used to build the GNSS positioning PPP error statistical model. The conclusion of the work and the prospects are introduced in Section 6.

2 Ionospheric disturbances datasets presentation

2.1 NMA database

The Norwegian Mapping Authority (NMA) has deployed instrumentation, developed software and services to monitor space weather impact on the ionosphere over Norway. NMA is also studying the impact of the ionospheric disturbances on its real-time kinematic (RTK) positioning service and precise point positioning (PPP) techniques (Jacobsen & Andalsvik, 2016).

In this study, ionospheric disturbances are measured by the Rate of TEC Index (ROTI) (Pi et al., 1997). The ROTI index is able to detect ionosphere disturbances and is a robust index, not affected by detrending of the phase measurements needed in the phase scintillation index definition (Ghobadi et al., 2020). While the use of dual-frequency observations allows the correction of ionospheric delay to the first order, higher order terms remain. ROTI main advantage over scintillation indices is that it is calculated based on measurements from standard dual-frequency GNSS receivers sampling at 1 Hz, which have been and still are more common than scintillation receivers. It is noted that ROTI based on 1 Hz measurements will often be dominated by contributions from refractive effects, whereas the scintillation indices calculated based on data with higher sampling rates are designed to measure contributions from diffractive effects (McCaffrey & Jayachandran, 2019). Rino et al. (2019) demonstrates that the phase-scintillation errors are a negligibly small fraction of the stochastic TEC component for most global navigation satellite system operating conditions. ROTI index has been chosen to characterize disturbances and/or rapid variations of TEC, and is related, but not equivalent, to scintillation (Basu et al., 1999; Mushini et al., 2012; Yang & Liu, 2016; Carrano et al., 2019).

In this study, the ROTI values are based on 1 Hz measurements ($\Delta t = \frac{1}{60}$ min), and calculated for time intervals of 5 min ($N = 300$). An elevation cutoff of 5° was used for the input data. However, for the ROTI data used in the processing in this paper, a further elevation angle cutoff of 30° was applied for the ROTI data records. This is done to avoid errors that occur at low elevations, in particular those due to reflections or other types of signal degradation caused by terrain or structures nearby.

The elevation of a ROTI data record is defined as the average elevation angle of the satellite during the time period used for the ROTI calculation.

ROTI is defined as the standard deviation of the Rate Of TEC (ROT) over some time interval. It is calculated as follows, where Ln, λ_n and f_n are the phase measurement, wavelength and frequency for the n-th frequency.

$L_{GF}(i)$ is the geometry-free phase combination at time i:

$$L_{GF}(i) = L1(i) \cdot \lambda_1 - L2(i) \cdot \lambda_2. \quad (1)$$

ROT (in TECU/minute) is calculated as,

$$\mathrm{ROT}(i) = \frac{L_{\mathrm{GF}}(i) - L_{\mathrm{GF}}(i-1)}{\Delta t \cdot 10^{16} \cdot 40.3 \cdot \left(\frac{1}{f_1^2} - \frac{1}{f_2^2}\right)}. \quad (2)$$

TECU (TEC Unit) is defined as 10^{16} electrons per m^2. Δt is the time difference between the epochs, in minutes.

Finally, ROTI, calculated over N epochs is,

$$\mathrm{ROTI}(k) = \sqrt{\frac{1}{N}\sum_{j=k-N}^{k}(\mathrm{ROT}(j) - \overline{\mathrm{ROT}})^2}. \quad (3)$$

$\overline{\mathrm{ROT}}$ is the average of ROT for the interval k ($\overline{\mathrm{ROT}} = \frac{1}{N}\sum_{j=k-N}^{k}\mathrm{ROT}(j)$). ROTI units are TEC unit per minutes.

In addition to basic checks of the input data a particularly important issue is the detection of cycle slips, as undetected cycle slips will lead to erroneous and very large ROTI values. The vast majority of the database used is from a real-time ionosphere monitoring system. In the real-time system, several checks are applied:

- For each signal frequency used, if Doppler measurement is available, the change in phase from the previous to the current timestep is compared to the Doppler. If the difference exceeds a threshold, it is considered as a cycle slip.
- The acceleration of TEC (second order time difference) is compared to a threshold. If the threshold is exceeded, it is considered as a cycle slip.
- The change of the Melbourne–Wübbena linear combination (Melbourne, 1985; Wübbena, 1985) from the previous to the current timestep is compared to a threshold. If the threshold is exceeded, it is considered as a cycle slip. A description of this cycle slip detection method can be found in Liu (2011).

If one or more of these checks trigger, the data is considered to contain a cycle slip. Among these the check using the Melbourne–Wübbena is the best, as the threshold can be set quite low compared to the change incurred in this measure by a cycle slip. However, there are some very special cases of combinations of cycle slips that the Melbourne–Wübbena check cannot detect which may be detected by the other methods. The processing software will consider the time series before and after a cycle slip as separate time series, ensuring that ROTI values are not calculated using data that contains cycle slips. Some parts of the data may be from a post-processing. The post-processing is equivalent to the real-time processing except that a more advanced cycle slip detection method, which is also based on the Melbourne–Wübbena linear combination, is used. A description of that detection method is found in Cai et al. (2013). Regardless of the outcome of the cycle slip detection, the calculated ROT values are also subject to simple checks for outliers or obviously false values (i.e. extremely large values).

A large database composed of ROTI index has been built. The ROTI database corresponds to the period (2007–2019), covering the 24th solar cycle. ROTI data have been recorded every 5 min, with a sample rate of 1 Hz, from 12 stations deployed all over Norway (see stations positions in Fig. 1). For the ROTI model developed in this study, only GPS satellites signals have been considered, leading to 78,990,363 measurements in total. For each ROTI value, the inhomogeneities position is estimated considering the link geometry and assuming an ionosphere layer altitude of 350 km. This generic value is implying that disturbances occurs at the same altitude whatever the solar conditions and close to F region peak level. This point remains questionable and should be studied in further work. For the purpose of the statistical model researched here, this choice is assumed as reasonable. In this paper, all coordinates are converted from geographical to altitude corrected geomagnetic coordinates using AACGM-v2 code (Altitude Adjusted Corrected GeoMagnetic (Shepherd, 2014). ROTI values have been corrected for elevation angle from the theoretical behavior proposed by Carrano et al. (2019) for the ROTI and assuming an isotropic ionosphere. This way, the elevation dependency of ROTI^2 is proportional to $\sec(\theta)$ with θ the zenith angle at the ionosphere inhomogeneity position (i.e. the angle complementary of the elevation angle). A practical reason for the choice of a duration of 5 min for ROTI is that it is the time resolution used in the archive from which we retrieved ROTI data. However, it is generally known that during active conditions, the ionospheric disturbances can vary on timescales shorter than 5 min. If the disturbance levels change during a 5-min interval, the ROTI value computed will represent an average of the disturbance level during that time interval. While this may not be good enough for a scientific study of the exact physics taking place in the ionosphere, we claim that it is sufficiently good for statistical characterization.

2.2 Solar wind data

For the purpose of ionosphere disturbance forecasting, the solar wind and interplanetary magnetic field (IMF) data from a satellite which is stationed in orbit around the Sun–Earth Lagrange Point L1 are a very interesting driver. The OMNIweb site (https://omniweb.gsfc.nasa.gov/) is providing high resolution solar wind data from the OMNI database. These data, which are 1 min averages timeshifted to the bow-shock-nose, have been considered as proxies for ionosphere disturbance forecasting. The Solar Wind data covering all the ground disturbance measurements have been considered, i.e. from 2007 up to 2019.

3 Ionospheric disturbances index model

3.1 Solar wind driver

In previous studies (Fabbro et al., 2019), the two parameters B_z (component of the interplanetary magnetic field perpendicular to the ecliptic, created by waves and disturbances in the solar wind) and p (solar wind pressure) were selected as proxies of the model. We have studied the correlation between ionospheric disturbance indices and these parameters. A global trend of increasing ionosphere ROTI index with increasing pressure and high negative values of B_z was observed, but the influence of other solar wind parameters dependency was missing. Hence, a coupling function that can better represent the interaction between solar wind and Earth magnetosphere has been researched. Newell et al. (2007) have investigated a few of

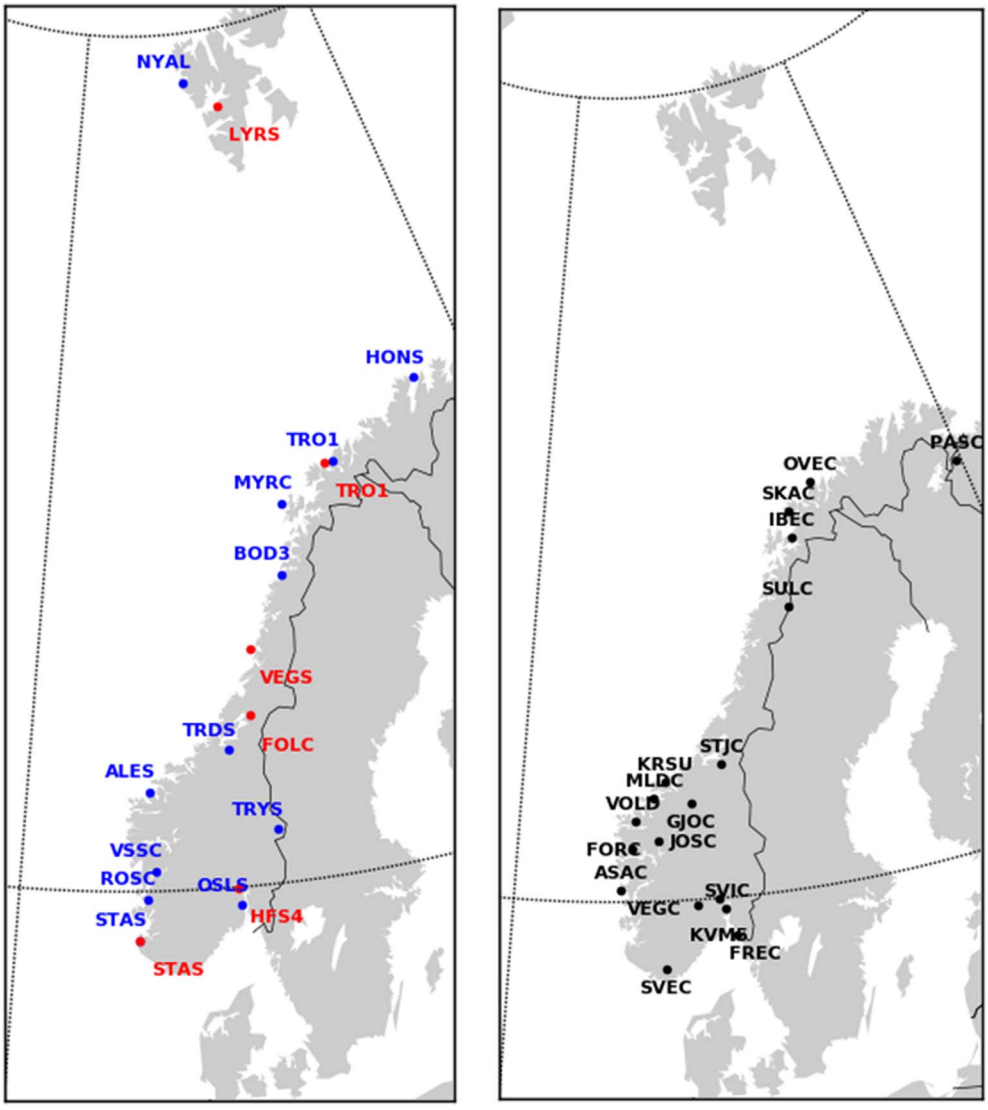


Fig. 1. Geographic locations of NMA stations used, in blue for ROTI model development (left graph), in red geodetic GNSS receivers for PPP positioning regressions (left graph), in black for validation (right graph).

existing coupling functions over a wide variety of magnetospheric activity, and proposed a new one $\frac{d\Phi_{MP}}{dt}$. This coupling function corresponds to the rate of magnetic flux that is opened at the magnetopause. Notably, the $\frac{d\Phi_{MP}}{dt}$ coupling function is used in OVATION-Prime auroral precipitation model (Newell et al., 2014). The definition of this coupling function is,

$$\frac{d\Phi_{MP}}{dt} = v^{\frac{4}{3}} B_T^{\frac{2}{3}} \sin^{\frac{8}{3}}(\theta/2), \tag{4}$$

where the clock angle is $\theta = \tan^{-1}\left(\frac{B_y}{B_z}\right)$ and $\theta \in [0, 2\pi]$, the IMF component transverse to the B_T (in nT) is $B_T = \sqrt{B_y^2 + B_z^2}$, and v is the solar wind velocity in km/s, whereas $\frac{d\Phi_{MP}}{dt}$ is in Wb/s. In Figure 2 a representation of $\frac{d\Phi_{MP}}{dt}$ variability versus the IMF components B_y and B_z is shown. The latter shows the strong variability of this coupling function with B_z sign, whereas the sign of B_y sign has no impact.

In our approach, the coupling function proposed by Newell et al. (2007) has been chosen as proxy and is computed from solar wind velocity and IMF parameters individually averaged over 10 min before the coupling function calculation (considering averaged IMF parameters from $t_0 - 10$ mn up to t_0 for a coupling function at time t_0). This choice of 10 min is based on the time variability of the parameters investigated by Cousins & Shepherd (2010). To study $\frac{d\Phi_{MP}}{dt}$, Newell et al. (2007) compared it with the K_p-index that is a global geomagnetic index, based on 3-h measurements of ground-based magnetometers placed around the world. Considering Figure 4 of Newell et al. (2007), the relation between $\frac{d\Phi_{MP}}{dt}$ and mean K_p-index is close to be linear in the range 1–6 for K_p. From these results, a simple linear regression can be proposed, $K_p \cong \alpha + \beta \frac{d\Phi_{MP}}{dt}$ with $\alpha = 1.1$ and $\beta = 3.74 \times 10^{-4}$. In this paper, this relation is used to compute an equivalent K_p-index to feed a model of Auroral oval boundaries used to be compared with our approach in Section 3.3.

3.2 ROTI model regression

The model has been regressed by a classification of the coupling function values and of the corresponding disturbance indices. To build the regression, 11 intervals of $\frac{d\Phi_{MP}}{dt}$ have been defined in linear scale, with an initial bin [0, 10^3), in Wb/s, followed by 9 intervals of width 2×10^3 Wb/s and a final bin [19×10^3, $+\infty$) Wb/s. Considering the position of each ROTI measure at 350 km and the measurement date, MLat (Magnetic Latitude) and MLT (Magnetic Local Time) are estimated by AACGM-v2 routine (Shepherd, 2014), model recommended by SuperDARN radar community. Then, a classification in terms of MLat and MLT has been proposed, with intervals of 1° of latitude between [50, 82)° for MLat, with two additional intervals [0, 50)° and [82, 90]° for remaining extra points, and intervals of half an hour between [0, 24) h for MLT. This choice leads to 34 intervals in MLat and 48 intervals in MLT. To establish the correspondence between the coupling function at bow-shock-nose and the disturbance indices at ground level, a time delay has to be defined. As we propose a statistical forecasting model, the idea is, from a coupling function value measured at time t_0, to consider all the disturbance indices measurements from t_0 up to $t_0 + \Delta t$. This leads to propose a model able to estimate the probability of a disturbance index to be exceeded in the following period Δt.

After deep analysis, the distributions of ROTI index appears to follow a distribution close to be lognormal. Then, the ROTI PDF (probability distribution function) is modeled by:

$$P(\mathrm{ROTI}) = \frac{1}{\mathrm{ROTI}\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(\mathrm{ROTI}) - \mu)^2}{2\sigma^2}\right). \quad (5)$$

Its mean and variance are respectively $E(\mathrm{ROTI}) = e^{\mu+\frac{\sigma^2}{2}}$ and $\mathrm{var}(\mathrm{ROTI}) = E(\mathrm{ROTI}^2) - E(\mathrm{ROTI})^2 = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$. The parameters μ and σ have been fitted for each combination of coupling function, MLat and MLT intervals of the whole database.

The relevance of the lognormal regression of ROTI variability has been tested. Three examples of ROTI index CDF (cumulative distribution function) are presented in Figure 3, corresponding to different combinations of location (characterized by MLat and MLT bins) and coupling function level. For each example the histogram plotted in blue line is compared to its lognormal CDF regression, plotted in red line. Examples (a) and (b) show a good agreement between measurements and lognormal model. For each graph, the number of point N_{pts} used in the regression is indicated, and the case (c) clearly underline a lack of measurements with $N_{\mathrm{pts}} = 323$.

From such an analysis, the ROTI index has been studied between 0 and 5. Then, the χ^2 goodness of fit test has been applied to check the null hypothesis, between the fitted and measured distributions. The p-value of the χ^2 test indicates if the null hypothesis is true. One of the most commonly used p-value threshold is 0.05. If the calculated p-value turns out to be less than 0.05, the null hypothesis is considered to be false. In other words, if p-value is lower than 0.05, it is very probable that the distribution and the model are different. Then, for each combined bins in MLat, MLT and coupling function, the χ^2 test is computed with ROTI values in the interval [0.1, 5] sampled with 250 values. The p-values have been computed by the Python "scipy.stats.chisquare" function.

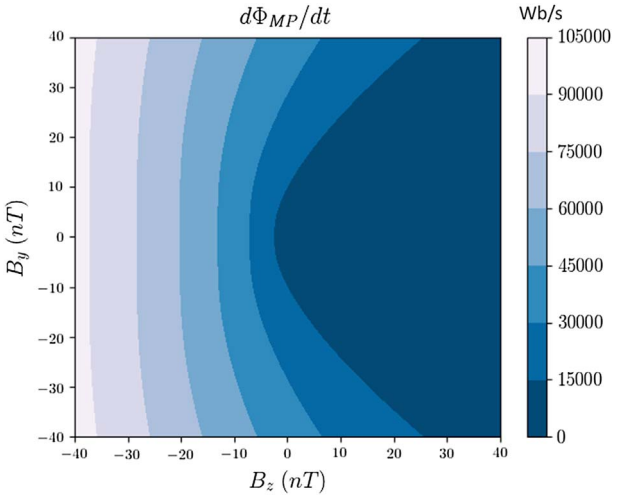


Fig. 2. $\frac{d\Phi_{MP}}{dt}$ (Wb/s) coupling function variations versus B_y and B_z in nT, for v = 850 km/s.

In Figure 3, the p-value of the three cases are provided. The p-value equal to zero in the (c) case indicates that it is very probable that the distribution is not lognormal. The origin of this is presumably the lack of data points to describe the distribution in this case.

The results are reported more globally in Figure 4. To focus on measurements and ignore noise effects, only bins where mean ROTI values are higher than 1 TECu/min have been plotted. This global processing is showing that for ROTI index regression a majority of p-value are higher than the threshold 0.05. Then, it is very probable that the lognormal distribution and the measurements distribution are similar. Moreover, in Figure 4 the areas where ROTI p-value is lower than 0.05 are mainly where the number of points is low (see Fig. 5). So a too low number of points can explain some low p-values, because the fit quality is clearly conditioned by the number of points. This number of points is a function of the solar wind conditions (or coupling function level) and of the magnetic coordinates. As expected, a large number of points is observed for lower coupling function level (higher than 10^5 in most of the areas) and decreases with increasing coupling function level. Note a lower number of points in geomagnetic latitude interval [70, 80], which is located north of mainland Norway and has a sparse measurement coverage, mostly from measurements performed by the Ny–Ålesund station in Svalbard (see Fig. 1).

3.3 ROTI model analysis

The mean and standard deviation of the lognormal distribution derived by this processing are shown in Figures 6 and 7, with respect to different intervals of coupling function. Note that no interpolation has been applied to get these results. For comparison, the auroral oval limits modeled by Feldstein–Starkov model (Starkov, 1994a; Sigernes et al., 2011, 1994b) are reported in the graphs. This model describes the equatorward and poleward limits of the auroral oval (drawn in green full line in Figs. 6 and 7), and the equatorward limit of the diffuse aurora (represented in green dashed line in Figs. 6 and 7). In the

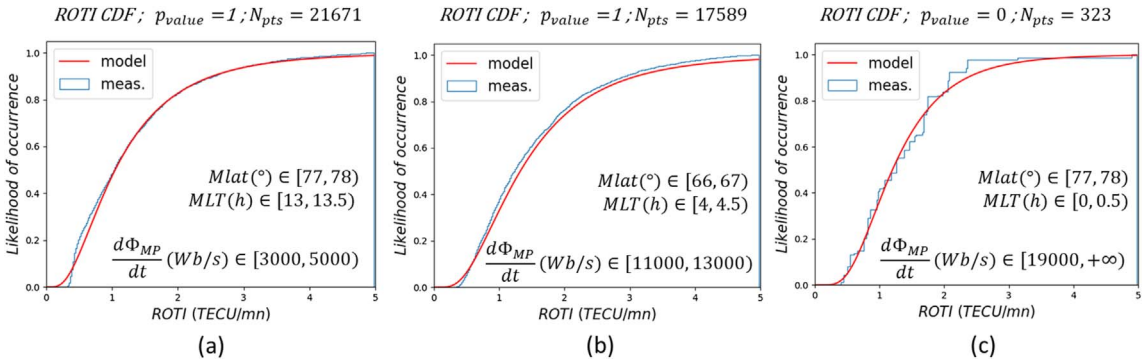


Fig. 3. Examples of ROTI index CDF measured and fitted for different combinations of location and coupling function level.

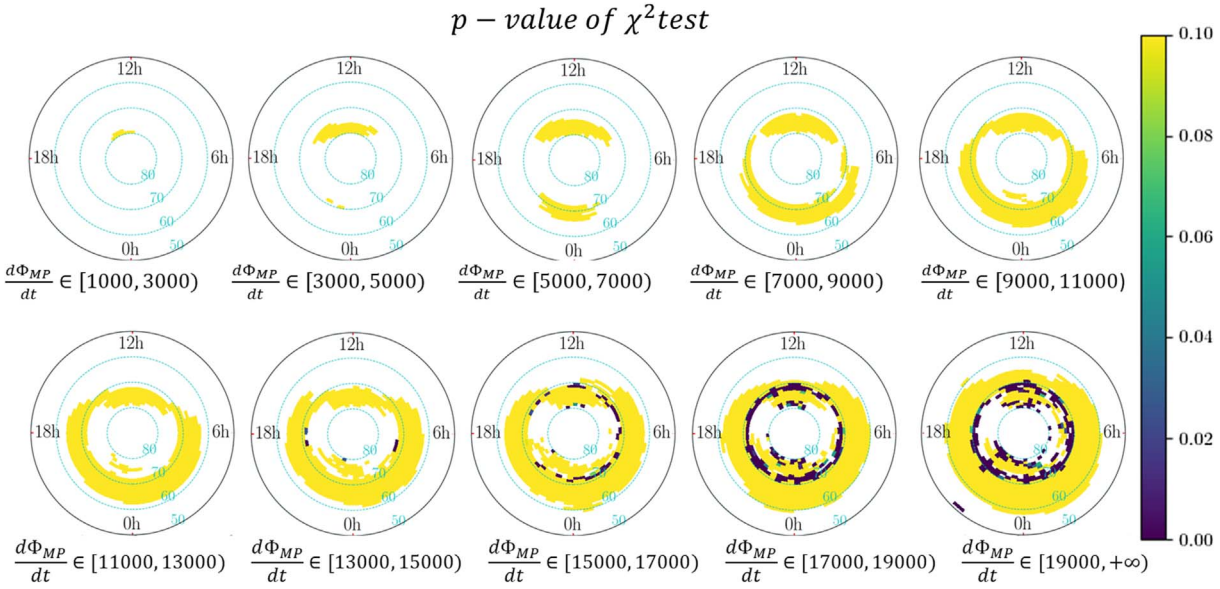


Fig. 4. p-value of χ^2 test for ROTI index modeling versus Mlat and MLT. Each graph corresponds to a different interval of $\frac{d\Phi_{MP}}{dt}$ values in (Wb/s).

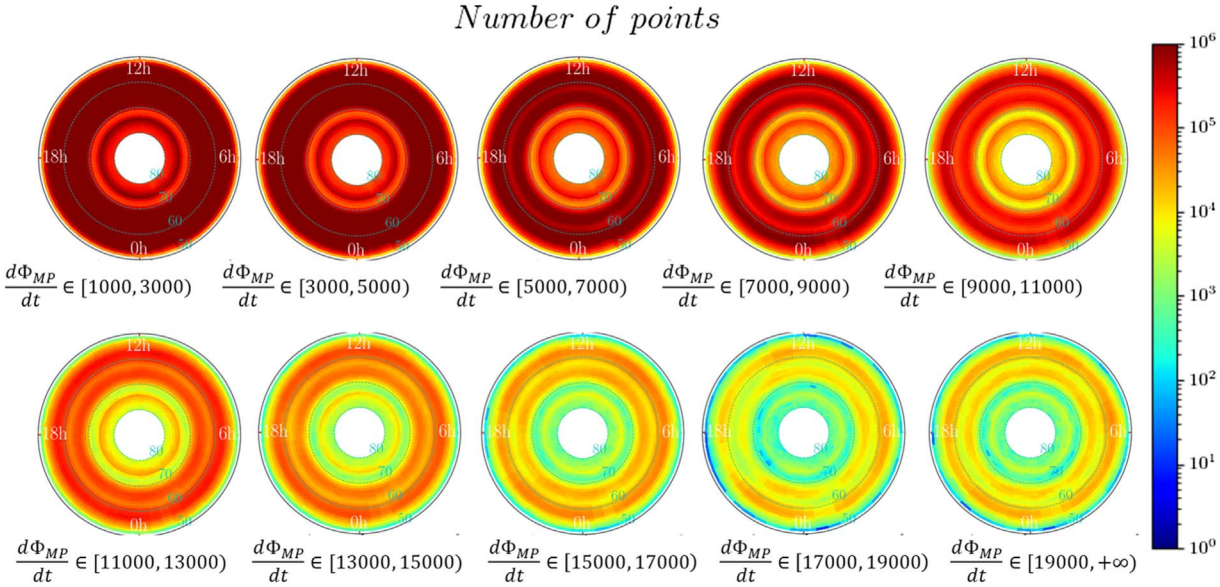


Fig. 5. Number of points of ROTI index versus Mlat and MLT, each graph corresponding to a different interval of $\frac{d\Phi_{MP}}{dt}$ values in (Wb/s).

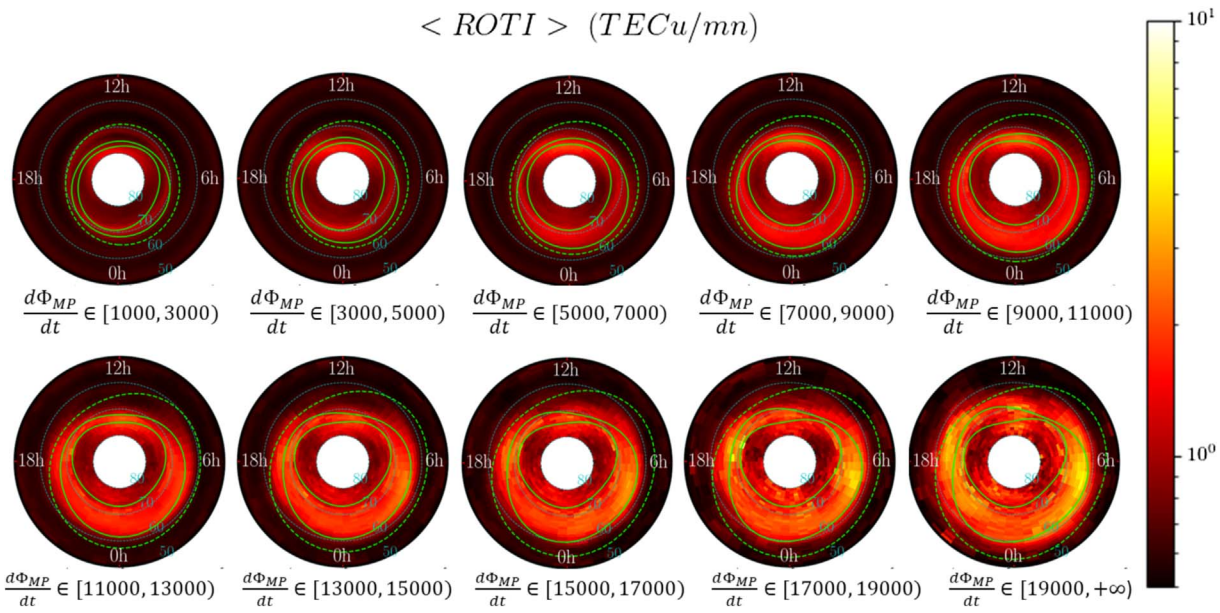


Fig. 6. Mean ROTI value in (TECu/mn) regressed versus geomagnetic latitude and magnetic local time. Each graph corresponds to a different interval of $\frac{d\Phi_{MP}}{dt}$ values in (Wb/s).

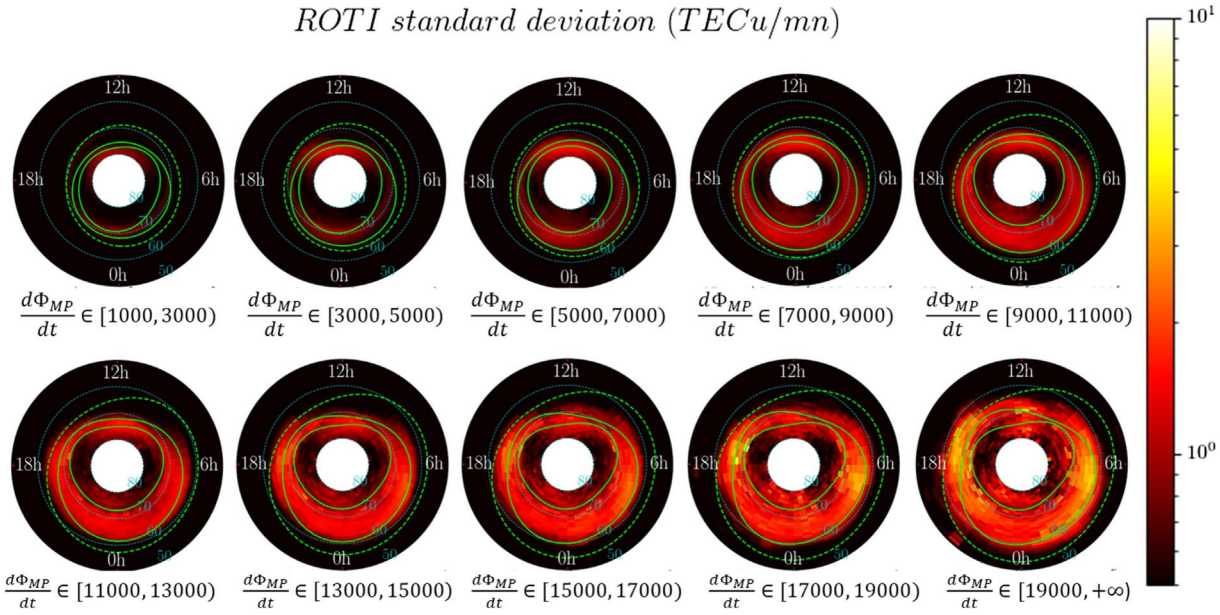


Fig. 7. Standard deviation of ROTI in (TECu/mn) regressed versus geomagnetic latitude and magnetic local time. Each graph corresponds a different interval of $\frac{d\Phi_{MP}}{dt}$ values in (Wb/s).

Feldstein–Starkov model, diffuse aurora is distinguished from the aurora oval because in diffuse aurora the precipitating protons do not contribute significantly to the excitation (Lui et al., 1973). Occasionally subject to controversy in the literature, the aurora limits proposed here must not be considered as an absolute reference, but as an indicator of trend. To feed the Feldstein–Starkov model an equivalent K_p index has been estimated by the simple linear regression proposed in Section 3.1, rounded to an accuracy of 0.3 to follow K_p definition. In graphs Figures 6 and 7, the equivalent K_p is varying between 0.6 and 9, when $\frac{d\Phi_{MP}}{dt}$ is corresponding to intervals [1000, 3000) up to [19,000, $+\infty$) Wb/s.

Observing the variability of the mean and standard deviation of ROTI with increasing solar wind disturbance conditions, it is clear that the typical variations of ionosphere perturbations are retrieved. As the auroral oval described by Feldstein–Starkov model, ROTI disturbance index shows an oval that is wider and more asymmetric with increasing coupling function. Ionospheric disturbances is more intense during night time, and still significant during the day. For low level of perturbation,

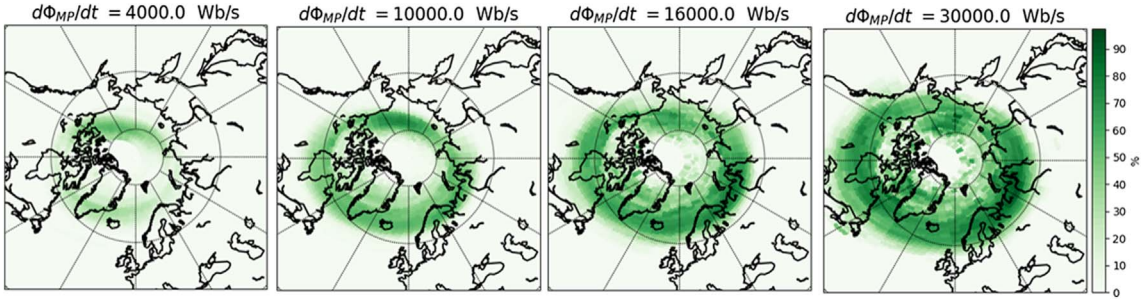


Fig. 8. Examples of probability of occurrence for a ROTI threshold of 1 TECu/mn to be exceeded for different coupling function values, at 0 h, UTC.

i.e. $\frac{d\Phi_{MP}}{dt}$ in [1000, 3000) Wb/s, the auroral oval shape is not observed by ROTI index, but the cusp region can clearly be distinguished on the day side, centered around 12 h and at latitude [75–80]°. Its size and intensity increase with the level of the coupling function. The night side of the auroral oval intensifies and broadens continuously with increasing coupling function. The auroral oval shape is clearly discernible from interval [3000, 7000) Wb/s on mean ROTI (Fig. 6) and [7000, 9000) Wb/s on ROTI standard deviation (Fig. 7). From [7000, 9000) up to [11,000, 13,000) Wb/s, the ROTI standard deviation limits between 60° and 70° of magnetic latitude match fairly well with the aurora limits proposed by Feldstein–Starkov model on the night side. The dissymmetry of the oval proposed in Feldstein–Starkov model is not well reproduced by ROTI index. From [7000, 9000) Wb/s and above levels of coupling function, at the opposite side of the oval corresponding to MLT time 18 h and 6 h, areas where ROTI mean and standard deviation have high values outside the limits proposed by the Feldstein-Starkov model can be noticed.

From $\frac{d\Phi_{MP}}{dt}$ equal to 7000 Wb/s and above, on the dayside the high values of ROTI index (mean or standard deviation > 1) have a larger latitudinal spread than the Feldstein–Starkov oval model, the equatorward limit is exceeded. However, globally the ionosphere disturbances measured by ROTI index happens inside the diffuse aurora limit here described by the Feldstein–Starkov model. Finally, for $\frac{d\Phi_{MP}}{dt}$ in [11,000, 13,000) Wb/s (higher values of coupling function reported in Figs. 6 and 7), two maximum values of ROTI distribution parameters appear close to 19 h and 5 h MLT. These intensifications of ROTI index could be due to the plasma inflow and outflow along R1 and R2 field-aligned currents (FAC) that are in the same area. (A review of magnetosphere-ionosphere current systems can be found in Cowley, 2000). Particle precipitation is associated with changes in plasma density and formation of small-scale plasma structures in the ionosphere, which would cause enhanced ROTI values. The enhancements could also be a result of the combination of the particle precipitation in the FAC current regions and plasma blobs embedded in the ionospheric return convection (van der Meeren et al., 2015; Jin et al., 2017; Jin & Oksavik, 2018).

Assuming a lognormal distribution for ROTI, the cumulative distribution function (CDF) is described by $\Pr(\mathrm{ROTI} \leq T_{\mathrm{ROTI}}) = \frac{1}{2}\mathrm{erfc}\left(-\frac{\ln T_{\mathrm{ROTI}} - \mu_{sw}}{\sigma_{sw}\sqrt{2}}\right)$, where erfc is the complementary error function and T_{ROTI} is the ROTI threshold. The complementary CDF (CCDF) is more interesting for operational applications, describing a percentage of the time that a ROTI value exceeds the ROTI threshold T_{ROTI}. It is known as $\Pr(\mathrm{ROTI} > T_{\mathrm{ROTI}}) = 1 - \frac{1}{2}\mathrm{erfc}\left(-\frac{\ln T_{\mathrm{ROTI}} - \mu_{sw}}{\sigma_{sw}\sqrt{2}}\right)$. Some examples of results obtained with the HAPEE model are shown in Figure 8. The different graphs represent the probability of the ROTI index to exceed the value 1 with regard of different values of the Newell et al. (2007) coupling function. The auroral oval shape appears clearly, its size and thickness growing when the Newell et al. coupling function increases.

4 Validation

A validation of the HAPEE model has been performed from measurements of a new set of ROTI data, measured in 2017 by receivers not used during the HAPEE model regressions. The receivers are part of the NMA GNSS receiver network, and the stations positions are plotted in graph (b) of Figure 1. ROTI data was calculated by NMAs ionosphere monitoring software. Solar wind data was obtained from NASA/GSFC's OMNI data set through OMNIWeb. For every 5 min in the time periods used, the HAPEE model was run to generate ROTI forecasts if valid solar wind data was available at that time. The output was longitude–latitude grids of the probabilities to exceed each of the thresholds during the coming hour. The grids covered the region of 10° West to 40° East, and 55–80° North. The grid resolution was 1 × 1°. For each time and each grid point, the observed ROTI values for IPPs (ionospheric pierce points) within the coming hour and spatially close to that grid point were identified. Spatially close is defined as being within 100 km (great circle distance) of the grid point. For the overall statistics, the data are sorted into bins by two parameters; the forecasted ROTI level (low, medium or high), and the forecasted percent chance of exceeding that level within the next hour (a percentage). The number of data points in each bin is shown in Figure 9. For the Validation, three levels of ROTI have been defined: low corresponding to 0.75 TECU/min; medium corresponding to 1.5 TECU/min and high corresponding to 3.0 TECU/min. The observed percentage of data records that exceeded the ROTI level is shown by the color in Figure 10. For a perfect match, the colors of the observations in the figure (percentage of records exceeding levels) should correspond to the abscissa value describing the percentage forecasted.

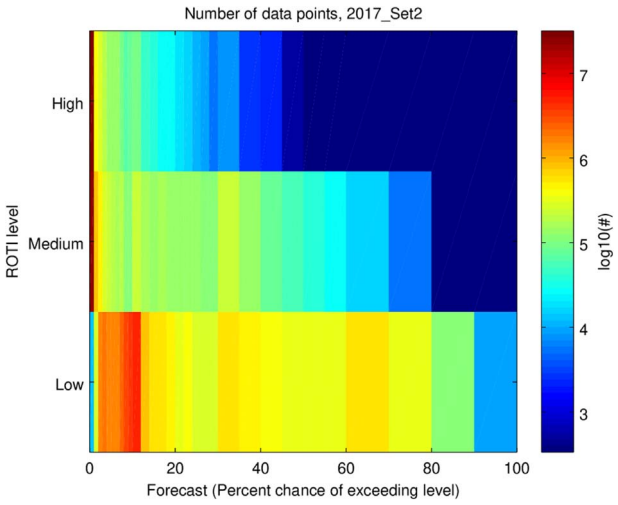


Fig. 9. In color scale, the logarithm of the number of data points for observed percentage of records exceeding the disturbance level, in each bin, versus forecast in abscissa, and low, medium and high thresholds in ordinate for time period of validation.

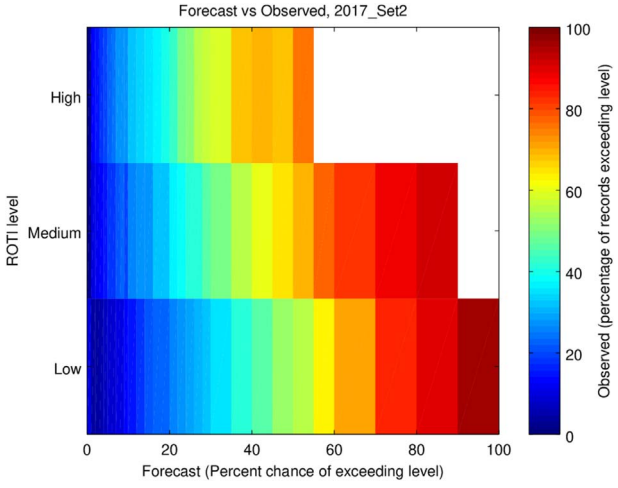


Fig. 10. In color scale, observed percentage of records exceeding the disturbance level, in each bin, versus forecast in abscissa, and Low, Medium and High thresholds in ordinate.

For the low ROTI level, the forecasts matches the observations very well. For the medium ROTI level, there is a shift in the colors on the plot versus the abscissa values, indicating a slight underestimation of the probability of occurrence. The forecasted probability is up to approximately 20 percentage points different from the observed probability. For the high ROTI level the shift is more pronounced, with maximum differences up to approximately 30 percentage points. The shifts in probability at high disturbance levels may indicate an inaccurate characterization of the tails of the distributions for the modeling and/or for the reference data. This is a challenging part which will be further explored in future work. More data are necessary and the future study of the solar cycle 25 are expected to lead to enhancement of the model and of its validation.

5 Positioning error estimation from ROTI

In this section, the distribution of the PPP errors is calculated for different bins of ROTI index. Then, these distributions of the positioning errors are fitted by a canonical distribution function, and their parameters computed for all the ROTI bins.

5.1 Data sources for characterization of ROTI versus PPP accuracy

The work to characterize the connection between ROTI and precise point positioning (PPP) accuracy is based on GNSS data from a set of geodetic receivers for a selection of dates (see Table 1). Figure 1 shows the locations of the receivers. All the receivers are owned and operated by NMA. The receivers record observations at a sample rate of 1 Hz. The days that are included are days that were observed to disturb NMAs RTK positioning systems and/or days with a minimum rating of 2 on NOAAs G-scale for geomagnetic storms (Poppe, 2000). The data has been processed to calculate ROTI (see Sect. 2.1) and PPP coordinates (see Sect. 5.2) at 5 min resolution.

The PPP coordinates were further processed and associated with the ROTI data as described in Section 5.2.

5.2 ROTI versus PPP data processing

We have used the GIPSY software provided by NASAs Jet Propulsion Laboratory (JPL) to compute coordinates. The coordinates were computed with a time resolution of 5 min. Important parameters/models used for the GIPSY PPP solutions are summarized in Table 2.

To determine the position error, a true coordinate was defined for each receiver by estimating a long-term trend and yearly variations. The long-term trend can be seen as an estimate of multi-year drift caused by e.g. continental drift or post-glacial rebound, and was modeled as a linear drift over the entire dataset. The yearly variations were modeled by fitting a 4-segment 5th order spline with periodic boundary conditions to the dataset sorted by day-of-year after removing the long-term trend. The true coordinate is then the sum of the long-term trend and the yearly variations. The PPP coordinate error is the difference between the instantaneous coordinate and the "true" coordinate evaluated at that time.

PPP coordinate errors were transformed from latitude, longitude, height to local North, East, Vertical coordinates. A basic outlier detection was applied to remove position records that were very clearly wrong: All PPP records with a position error greater than 20 m in East/North, or greater than 50 m in Vertical, were removed as outliers. This affected approximately 0.2% of the dataset.

For each timetag and each receiver, ROTI was averaged over all satellites. For each averaged ROTI data record, the position error of the PPP data record closest in time (and closer than 300 s) is selected as the position error for that record.

5.3 ROTI versus PPP accuracy

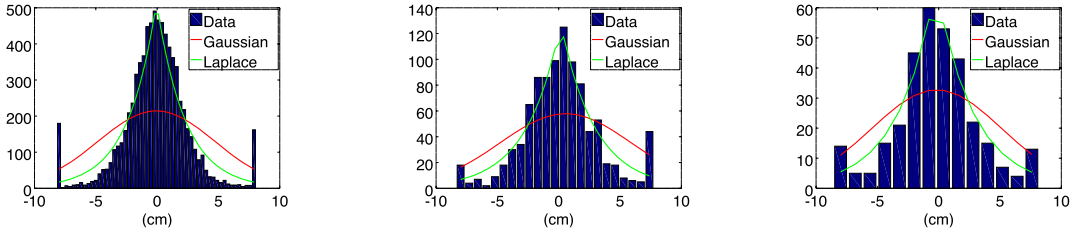
Figure 11 shows the distributions of the Vertical position errors, along with fits of Gaussian and Laplace distribution functions. Figures 12 and 13 show the same for the North and East coordinate components. The histograms count the position

Table 1. List of days included in this study.

Year	Month	Days	G-index
2014	February	15, 18, 19, 20, 27, 28	1, 0, 2, 2, 0, 2
2014	September	12, 24	3, 0
2014	October	14	1
2015	January	7	3
2015	March	17	4
2015	June	22, 23	4, 4
2015	October	7, 8	3, 2
2015	November	7	2
2015	December	20, 21	2, 2
2016	January	20	1
2016	February	16	2
2016	March	6, 7	2, 2
2016	May	8	3
2016	September	25, 26, 27, 28, 29, 30	1, 1, 2, 2, 1
2016	October	1, 2, 3, 4, 13, 14, 25, 26, 27, 28, 29, 30	1, 1, 0, 1, 2, 2, 3, 2, 1, 0, 1, 0
2017	March	1, 2, 27, 28	2, 2, 2, 2
2017	April	20, 22, 23	2, 2, 2
2017	May	27, 28	2, 3
2017	June	16, 17	1, 1
2017	August	22	2
2017	September	7, 8, 9, 15, 16, 27, 28	4, 4, 2, 2, 2, 2, 3
2017	October	11, 12, 13, 14, 15, 16	1, 1, 2, 2, 1, 0
2017	November	7, 8	2, 2

Table 2. Parameters/models used for the GIPSY PPP solution.

GIPSY version:	6.4
Reference frame:	IGS08/IGb08
Antenna phase center (receivers):	Absolute based on IGS standard e.g. igs08_1645.atx
Antenna phase center (transmitters):	Absolute based on IGS standard e.g. igs08_1645.atx
Troposphere mapping function:	VMF1
2nd order ionosphere model:	Not applied
Ocean loading:	FES2004
Ambiguity resolution:	Yes (Bertiger et al., 2010)



(a) $ROTI\ (TECu/mn) \in [0.5, 1.0)$ (b) $ROTI\ (TECu/mn) \in [3., 3.5)$ (c) $ROTI\ (TECu/mn) \in [6, 7)$

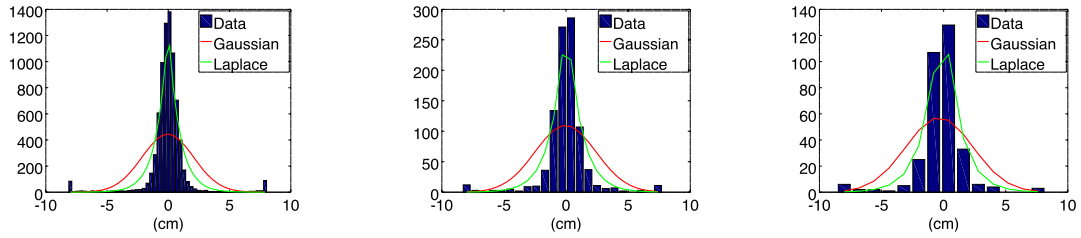
Fig. 11. Histograms showing distributions of PPP vertical coordinate error within specific ROTI bins (TECu/mn), with Gaussian (red line) and Laplace (green line) functions fitted to the data. (a) ROTI (TECu/mn) ∈ [0.5, 1.0). (b) ROTI (TECu/mn) ∈ [3, 3.5). (c) ROTI (TECu/mn) ∈ [6, 7).

errors to the left/right of the *x*-axis in the left/rightmost bar, but the fits have been made to the actual data. For the plots the distribution functions were evaluated at the points of the center of the histogram bars and scaled so that the sum of their values at those points is the same as the sum of the values of the histogram bars.

Table 3 lists the distribution parameters of the Laplace distribution function, being the median and the mean absolute deviation from the median (MAD).

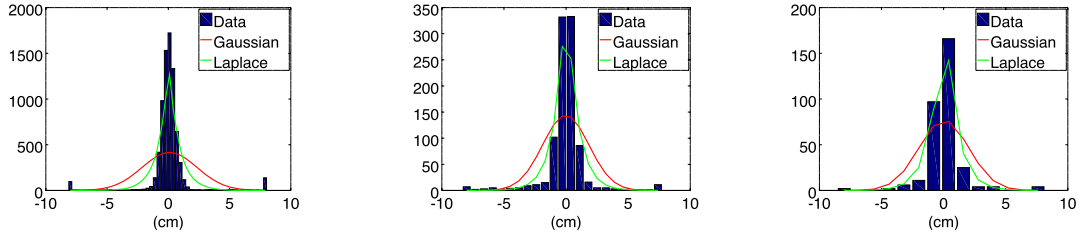
Figures 11–13 show the distributions of the position errors, along with fits of Gaussian and Laplace distribution functions.

That the Laplace distribution is a much better fit than the Gaussian distribution can be shown by comparing the Kolmogorov–Smirnov (KS) test statistic for the two distributions. Figure 14 shows an example of cumulative distribution functions (CDF), for the PPP vertical position error with ROTI in the interval 3–3.5. The blue line is the empirical CDF, while the red line is the Gaussian CDF and the green line is the



(a) $ROTI\ (TECu/mn) \in [0.5, 1.0)$ (b) $ROTI\ (TECu/mn) \in [3., 3.5)$ (c) $ROTI\ (TECu/mn) \in [6, 7)$

Fig. 12. Histograms showing distributions of PPP north coordinate error within specific ROTI bins (TECu/mn), with Gaussian (red line) and Laplace (green line) functions fitted to the data. (a) ROTI (TECu/mn) ∈ [0.5, 1.0). (b) ROTI (TECu/mn) ∈ [3, 3.5). (c) ROTI (TECu/mn) ∈ [6, 7).



(a) $ROTI\ (TECu/mn) \in [0.5, 1.0)$ (b) $ROTI\ (TECu/mn) \in [3., 3.5)$ (c) $ROTI\ (TECu/mn) \in [6, 7)$

Fig. 13. Histograms showing distributions of PPP east coordinate error within specific ROTI bins (TECu/mn), with Gaussian (red line) and Laplace (green line) functions fitted to the data. (a) ROTI (TECu/mn) ∈ [0.5, 1.0). (b) ROTI (TECu/mn) ∈ [3, 3.5). (c) ROTI (TECu/mn) ∈ [6, 7).

Table 3. Fitted Laplace distribution function parameters for PPP error versus ROTI.

ROTI Bin Center (TECU/min)	Median PPP error, North (cm)	MAD PPP error, North (cm)	Median PPP error, East (cm)	MAD PPP, East (cm)	Median PPP error, Vertical (cm)	MAD PPP error, Vertical (cm)
0.25	0.025	0.549	0.038	0.393	0.000	1.274
0.75	0.020	0.594	0.014	0.470	0.047	1.607
1.25	–0.013	0.923	0.042	0.858	–0.052	2.339
1.75	–0.019	0.976	0.014	0.889	–0.081	2.516
2.25	–0.013	1.059	–0.010	0.877	–0.136	2.577
2.75	–0.064	1.148	0.005	0.849	–0.091	2.618
3.25	0.030	1.132	0.012	0.875	0.166	2.790
3.75	–0.106	1.103	0.023	0.787	–0.280	2.744
4.25	0.003	1.166	0.058	0.927	–0.095	2.904
4.75	–0.096	1.373	0.041	0.857	–0.060	3.348
5.50	0.019	1.371	–0.008	0.950	–0.009	3.397
6.50	–0.108	1.272	0.020	0.947	–0.235	3.082
7.50	–0.122	1.353	–0.092	0.876	–0.250	3.331
8.50	0.019	1.166	0.120	0.754	–0.319	3.286
9.50	–0.191	1.514	0.006	1.252	–0.118	3.881
11.00	–0.020	1.472	0.073	1.070	–0.659	3.976
13.00	–0.168	1.687	0.140	0.993	0.063	4.340
15.00	–0.349	1.428	–0.018	0.887	–0.414	3.259
17.00	–0.204	1.296	–0.008	0.986	–0.311	3.470
19.00	–0.039	1.404	–0.009	1.091	–0.497	3.550
22.50	–0.039	1.630	0.157	1.064	–0.387	4.013
27.50	–0.273	1.732	0.101	1.132	–0.057	4.870

Laplace CDF. The KS test statistic is the maximum vertical difference between a distribution CDF and the reference CDF (in this case, the empirical CDF). Lower values are better, with a value of zero being a perfect fit. In this example, the value is 0.182 for the Gaussian distribution and 0.0466 for the Laplace distribution. From these values, we can conclude that the Laplace distribution is a better fit to the data. Figure 15 shows the KS test statistic values for all the bins. Except for the very first ROTI bin, the Laplace distribution is significantly better. Based on the KS test statistics results and manual inspection of the distributions (including more histogram and CDF plots which are not included in this paper), we conclude that the Laplace distribution provide a good fit to the distribution of position errors and that it is a significantly better fit than a

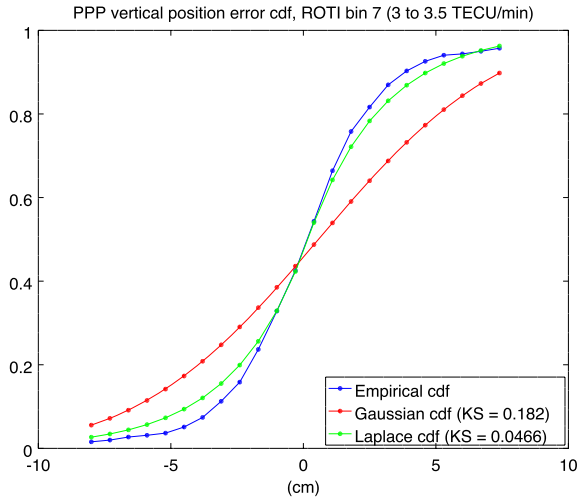


Fig. 14. Cumulative distribution functions for the data (blue line), the Gaussian fit (red line) and the Laplace fit (green line), for the PPP vertical position errors with ROTI in the interval 3–3.5 TECu/min. The Kolmogorov-Smirnov test statistic for the two fitted distributions are in the legend.

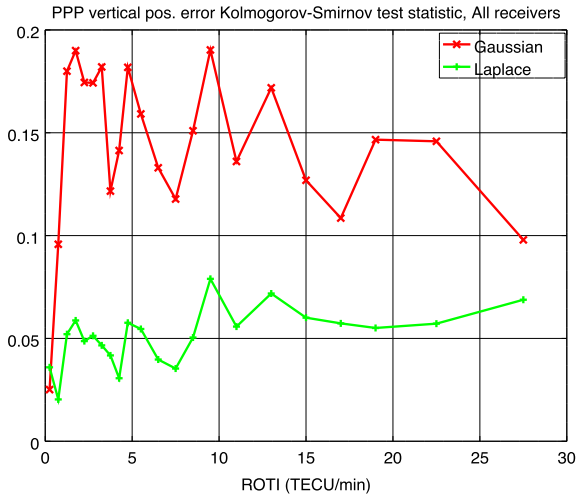


Fig. 15. Kolmogorov-Smirnov test statistic values for PPP vertical error distributions in all ROTI bins (in TECu/min). The red line is for the Gaussian fit compared to the data, and the green line is for the Laplace fit compared to the data.

Gaussian distribution function. Table 3 shows the fitted parameters for the Laplace distribution functions.

5.4 Modeling PPP error from solar wind conditions

For a given solar wind condition (or coupling function), the idea is to couple a model forecasting the probability of a disturbance index $p(\text{ROTI}|\text{sol.wind})$ with a probability of error conditioned by this ROTI value $p(\text{Pos}|\text{ROTI})$. The global compound model giving a probability for a positioning error Pos to exceed a threshold Te is then written as:

$$\Pr(\text{Pos} > \text{Te}|\text{sol.wind}) = \int_{\text{Te}}^{+\infty} p(\text{Pos}|\text{sol.wind})\text{dPos}, \quad (6)$$

with

$$p(\text{Pos}|\text{sol.wind}) = \int_{0}^{+\infty} p(\text{Pos}|\text{ROTI})\, p(\text{ROTI}|\text{sol.wind})\, \text{dROTI}. \quad (7)$$

The problem has been solved supposing the ROTI PDF as lognormal, such as modeled in the climatological part of the model described in Section 3. This way, from a given solar wind condition, the parameters (μ_{sw}, σ_{sw}) are chosen for the corresponding coupling function bin and following a lognormal distribution:

$$p(\text{ROTI}|\text{sol.wind}) = \frac{1}{\text{ROTI}\sigma_{\text{sw}}\sqrt{2\pi}} \exp\left(-\frac{(\ln \text{ROTI} - \mu_{\text{sw}})^2}{2\sigma_{\text{sw}}^2}\right). \quad (8)$$

Then, the probability $p(\text{Pos}|\text{ROTI})$ has to be estimated. The following section presents the regression of this part of the model.

5.4.1 Regression of precise point positioning error

In Section 5.3, the statistical relation between positioning errors and ROTI values has been quantified by regression. PPP North, East and Vertical error distributions have been fitted versus ROTI level. The distribution of errors for each ROTI bin has been fitted by a Laplacian distribution:

$$p(\text{Pos}|\text{ROTI}) = \frac{1}{2b_i} \exp\left(-\frac{|\text{Pos} - m_i|}{b_i}\right), \quad (9)$$

where m_i is the mean and the median of the distribution value, $2b_i^2$ is the variance and b_i the mean absolute deviation (MAD) from the median.

The Laplace distribution parameters for the PPP coordinate errors have been derived per ROTI bin and reported in Section 5.3.

In practice, the positioning error is characterized by the absolute value of the error. Moreover, as far as the PPP positioning errors are concerned, the fitted mean of North, East and Vertical PPP errors have been obtained close to zero (see Sect. 5.3). Then, the PDF considered for the problem becomes $p(|\text{Pos}||\text{ROTI}) = \frac{1}{2b_i} \exp\left(-\frac{|\text{Pos}|}{b_i}\right)$. This canonical distribution is therefore introduced in the global compound model.

5.4.2 Simple case of one link

In this section, the ROTI index is supposed to be estimated from one GNSS link. In continuous form, the global problem to solve is the following conditional probability,

$$\Pr(|\text{Pos}| > \text{Te}|\text{sol.wind}) = \int_{\text{Te}}^{+\infty} \text{dPos} \int_{0}^{+\infty} \text{dROTI}\, p(|\text{Pos}||\text{ROTI})\, p(\text{ROTI}|\text{sol.wind}). \quad (10)$$

Inverting the integrals, introducing the Laplace distribution parameters (whose values depend on the ROTI bin) and the lognormal distribution of ROTI, the formulation becomes,

$$\Pr(|\mathrm{Pos}| > \mathrm{Te}|sol.wind) = \int_0^{+\infty} \frac{\mathrm{dROTI}}{\mathrm{ROTI}\sigma_{\mathrm{SW}}\sqrt{2\pi}} \times \exp\left(-\frac{(\ln \mathrm{ROTI} - \mu_{\mathrm{SW}})^2}{2{\sigma_{\mathrm{SW}}}^2}\right) \times \int_{\mathrm{Te}}^{+\infty} \mathrm{dPos}\frac{1}{2b_i}\exp\left(-\frac{|\mathrm{Pos}|}{b_i}\right). \quad (11)$$

The second integral is $\int_{\mathrm{Te}}^{+\infty} \mathrm{dPos}\frac{1}{2b_i}\exp\left(-\frac{|\mathrm{Pos}|}{b_i}\right) = \Pr(|\mathrm{Pos}| > \mathrm{Te})$ and reminding the Laplace distribution properties, the complementary cumulative distribution function is,

$$\Pr(\mathrm{Pos} > \mathrm{Te}) = \begin{cases} 1 - \frac{1}{2}\exp\left(\frac{\mathrm{Te} - m_i}{b_i}\right) & \text{if } \mathrm{Te} < m_i \\ \frac{1}{2}\exp\left(-\frac{\mathrm{Te} - m_i}{b_i}\right) & \text{if } \mathrm{Te} \geq m_i \end{cases} \quad (12)$$

with in case of PPP positioning errors m_i; 0 (see Sect. 5.3). So $\mathrm{Te} \geq m_i$ can be assumed, and

$$\Pr(|\mathrm{Pos}| > \mathrm{Te}) = \frac{1}{2}\exp\left(-\frac{\mathrm{Te}}{b_i}\right). \quad (13)$$

The global conditional probability becomes,

$$\Pr(|\mathrm{Pos}| > \mathrm{Te}|\mathrm{sol.wind}) = \int_0^{+\infty} \mathrm{dROTI}\frac{1}{\mathrm{ROTI}\sigma_{\mathrm{SW}}\sqrt{2\pi}} \times \exp\left(-\frac{(\ln \mathrm{ROTI} - \mu_{\mathrm{SW}})^2}{2{\sigma_{\mathrm{SW}}}^2}\right)\frac{1}{2}\exp\left(-\frac{\mathrm{Te}}{b_i}\right). \quad (14)$$

Considering that each i bin of ROTI is defined by a minimum and a maximum value of ROTI, $\mathrm{ROTI}^i_{\mathrm{min}}$ and $\mathrm{ROTI}^i_{\mathrm{max}}$, a discretized form is proposed, via the formula,

$$\Pr(|\mathrm{Pos}| > \mathrm{Te}|\mathrm{sol.wind}) = \sum_{i=1}^{N_b} \delta R \frac{\Pr(\mathrm{ROTI} \leq \mathrm{ROTI}^i_{\mathrm{max}}) - \mathrm{PR}(\mathrm{ROTI} \leq \mathrm{ROTI}^i_{\mathrm{min}})}{\delta R} \times \frac{1}{2}\exp\left(-\frac{\mathrm{Te}}{b_i}\right), \quad (15)$$

with the ROTI bin number $i = [1; \ldots; N_b]$ and δR the ROTI interval width. In each bin, the ROTI is following the Laplace distribution with parameters $[m_i, b_i]$. $Pr(\mathrm{ROTI} \leq \mathrm{ROTI}^i_{\mathrm{max}})$ is the CDF of the lognormal distribution of ROTI. It can be written $\Pr(\mathrm{ROTI} \leq \mathrm{ROTI}^i_{\mathrm{max}}) = \frac{1}{2}\mathrm{erfc}\left(-\frac{\ln \mathrm{ROTI}^i_{\mathrm{max}} - \mu_{\mathrm{SW}}}{\sigma_{\mathrm{SW}}\sqrt{2\pi}}\right)$, and leads to,

$$\Pr(|\mathrm{Pos}| > \mathrm{Te}|\mathrm{sol.wind}) = \sum_{i=1}^{N_b}\frac{1}{4}\left[\mathrm{erfc}\left(-\frac{\ln \mathrm{ROTI}^i_{\mathrm{max}} - \mu_{\mathrm{SW}}}{\sigma_{\mathrm{SW}}\sqrt{2\pi}}\right) - \mathrm{erfc}\left(-\frac{\ln \mathrm{ROTI}^i_{\mathrm{min}} - \mu_{\mathrm{SW}}}{\sigma_{\mathrm{SW}}\sqrt{2\pi}}\right)\right]\exp\left(-\frac{\mathrm{Te}}{b_i}\right). \quad (16)$$

The final formulation is relatively easy to use once all the distributions parameters (Laplace for positioning PPP error and lognormal for ROTI) are known. But the problem is more complex because positioning error is commonly estimated from several GNSS links, that can be impacted differently by the ionospheric disturbances. Then, the conditional distribution must consider these different links to estimate a mean ROTI value.

5.4.3 Multiple link case

Several GNSS links have to be taken into account in the processes of estimating the PPP error. As these links are not corresponding to close values of ionosphere pierce point, a main assumption is that all links are considered as not correlated. Then, the average ROTI_N is estimated from the ROTI values measured on N links:

$$\mathrm{ROTI}_N = \frac{1}{N}\sum_{k=1}^{N}\mathrm{ROTI}_k. \quad (17)$$

The number N is of the order of 10, if different satellite constellations are used. We have to estimate the distribution of $N \times \mathrm{ROTI}_N$, sum of N ROTI_k values following lognormal distributions. A second approximation is then assumed: as the GNSS links are corresponding to the same ground station, the ROTI distribution of each link remains the same. The distribution of $N \times \mathrm{ROTI}_N$, estimated by the sum of independent ROTI values following the same distribution has to be derived. As no exact analytical expression is known for statistical characterization of the combination of N lognormal distributions, one will turn to approximation strategies.

From Romeo et al. (2003), a sum of N lognormal random variables of parameters (μ, σ) can be approximated to a new lognormal distribution of parameters (μ_N, σ_N):

$$\begin{cases} \mu_N = \mu + \frac{\sigma^2}{2} + \ln\left(\frac{N}{\sqrt{1 + \frac{e^{\sigma^2}-1)}{N}}}\right); \\ \sigma_N^2 = ln\left(1 + \frac{e^{\sigma^2} - 1}{N}\right), \end{cases} \quad (18)$$

with μ and σ the parameters of the lognormal distributions of the N processes supposed statistically with the same characteristics, μ_N and σ_N are the parameters of the lognormal distribution of $N \times \mathrm{ROTI}_N$. Romeo et al. (2003) show in their paper that this formulation is correct with σ in [0, 1.25].

For a ROTI lognormal distribution, σ parameter is $\sigma^2 = \ln\left(1 + \frac{<\mathrm{ROTI}^2>}{<ROTI>^2}\right)$. From the regression presented in Section 3, the maximum values obtained for σ induce that the formulation proposed by Romeo et al. (2003) is covering the problem.

From such an approach, the typical sum of lognormal distribution can be written $e^{\mu_N + \sigma_N Z}$ where Z is of standard normal distribution. The average is then performed by dividing by N, introducing a normalization $e^{\mu_N + \sigma_N Z - \ln N} = e^{\mu'_N + \sigma_N Z}$. The final characteristics of the average ROTI_N are then:

$$\begin{cases} \mu'_N = \mu + \frac{\sigma^2}{2} + \ln\left(\frac{N}{\sqrt{1 + \frac{e^{\sigma^2}-1}{N}}}\right) - \ln(N) \\ \sigma_N^2 = \ln\left(1 + \frac{e^{\sigma^2} - 1}{N}\right). \end{cases} \quad (19)$$

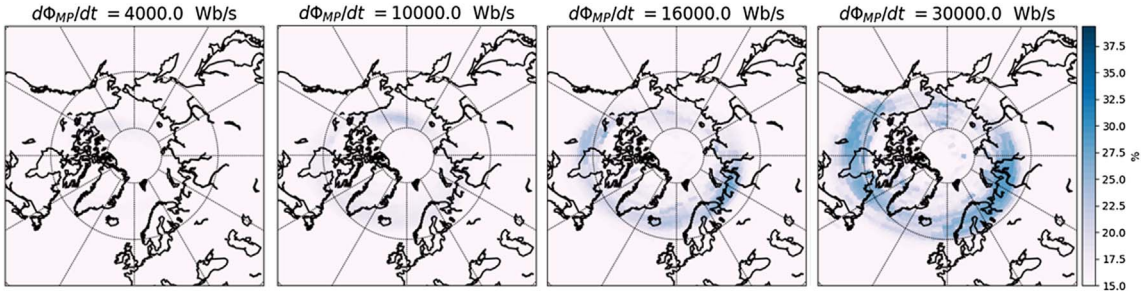


Fig. 16. Probability of East PPP error to exceed 0.5 cm in North Pole area, at 0 h, UTC.

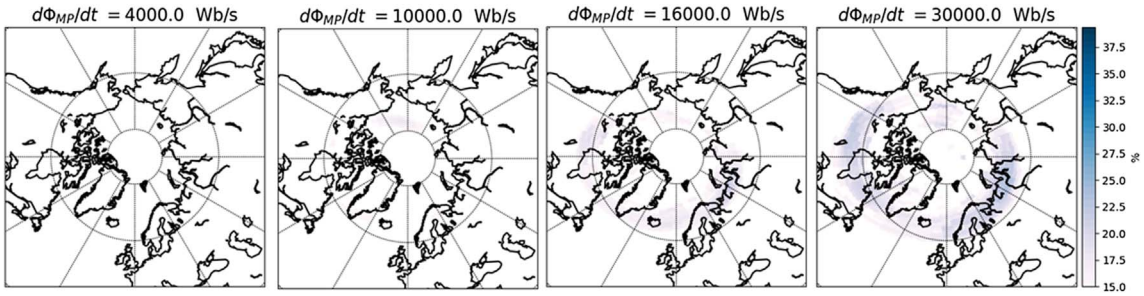


Fig. 17. Probability of North PPP error to exceed 0.75 cm in North Pole area, at 0 h, UTC.

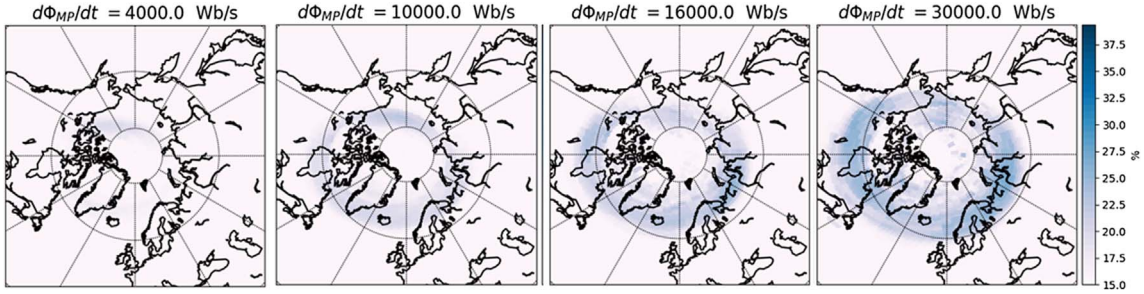


Fig. 18. Probability of Vertical PPP error to exceed 1.5 cm in North Pole area, at 0 h, UTC.

As a conclusion, we propose here an equivalent lognormal distribution of parameters (μ'_N, σ_N) of the ROTI_N stochastic variable, obtained from the mean of N lognormal random processes. Thanks to this approximation, we can now use the same strategy presented in Section 5.4.2 replacing ROTI by ROTI_N in order to estimate the PPP error statistics.

Typical results are shown in Figures 16–18 for East, North and Vertical errors respectively. The four graphs of each figure correspond to different levels of coupling function from 4000 to 30,000 Wb/s. Here the number of GNSS links N has been fixed to 10. The results are maps of the percentage of the time that a given error threshold is exceeded during the hour following 0h UTC, for the arctic region. To illustrate the variability with different thresholds, the values of 0.5, 0.75 and 1.5 cm are considered respectively for East, North and Vertical PPP errors, in Figures 16–18.

This way, in Figure 17 the probability of the East PPP error to exceed threshold 0.5 cm is reported. It shows that for a coupling function of 4000 Wb/s, the threshold of error 0.5 cm is exceeded less than 15% of the time, whereas for 10,000 Wb/s it is exceeded in the cusp close to 20% of the time, and when $\frac{d\Phi_{MP}}{dt}$ is increasing up to 30,000 Wb/s, in the auroral oval a percentage close to 30% of the time is reached. In Figure 17 the same representation is proposed, in the same conditions, but for the probability of North PPP error to exceed 0.75 cm. The Probability to exceed the error threshold of 0.75 cm is globally sparsely reached, with a maximum value close to 25% in the auroral oval for a coupling function of 30,000 Wb/s. Considering an error threshold of 1.5 cm, the Figure 18 shows that the Vertical PPP error is exceeded in the cusp around 20% of the time since $\frac{d\Phi_{MP}}{dt}$ = 4000 Wb/s, and reach a maximum value close to 30% in the auroral oval for $\frac{d\Phi_{MP}}{dt}$ = 30,000 Wb/s.

From PPP error estimation used to establish the correspondence between ROTI and PPP errors, an exercise of validation of error model has been performed. It is based on data from a GNSS ground station named "hamc" and located at the magnetic latitude 65.5° North. The dataset used is from the days listed in Table 1. From solar wind conditions measured at L1 at time t_0, characterized by the Newell coupling function, all the PPP errors in the following hour (up to t_0 + 1 h) have been

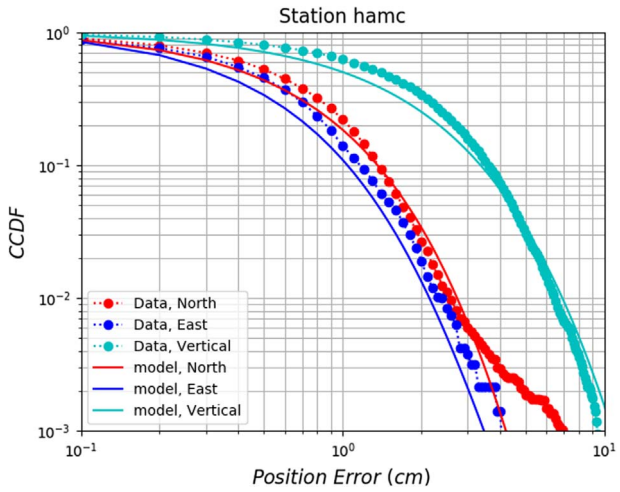


Fig. 19. CCDF of PPP positioning errors, comparison of HAPEE model results (full lines) with direct PPP error (dashed lines), for North, East and Vertical errors in red, blue and cyan respectively.

computed to establish the statistics. All MLT time have been considered. Then, this global statistic of error occurence has been compared to the proposed model. The latter has been applied to estimate $\Pr\left(|\mathrm{Pos}| > \mathrm{Te}|\mathrm{MLT}, \frac{\mathrm{d}\Phi_{\mathrm{MP}}}{\mathrm{d}t}\right)$ at the magnetic latitude corresponding to the station position, using equation (10) and (16) GNSS links. Regular MLT and coupling function intervals have been chosen to compute this probability, with N_h = 48 points used to describe the 24 h of MLT and N_c = 11 the number of coupling function covering the conditions between 0 and 20 $\times$ 10^3 Wb/s. The final result is obtained by the formulation $\Pr\left(|\mathrm{Pos}| > \mathrm{Te}\right) = \sum_{h=1,N_h} \sum_{i=1,N_c} \times \Pr\left(|\mathrm{Pos}| > \mathrm{Te}|\mathrm{MLT}, \frac{\mathrm{d}\Phi_{\mathrm{MP}}}{\mathrm{d}t}\right)/N_c N_h$. In Figure 19 are presented the results of comparison between direct and modeled CCDF, for Horizontal, East and Vertical PPP errors. The comparison shows correct results of compliance between the HAPEE model and the direct PPP estimations, with a clear lack of data to get convergence for probabilities lower than 10^{-2} (in particular for North error). This validation is a check of the model consistency, and the primary perspective of this work is to realize an operational validation exercise of the PPP error model. It necessitates to characterize long term estimation of PPP errors statistics related to solar wind condition.

6 Conclusion

A new forecasting model called HAPEE is proposed, able to forecast during the next hour and at latitudes higher than 50° the distribution of the ROTI index or of the PPP error. The model proxy is the solar wind conditions at the Earth bow-shock-nose (typically measured at the L1 Lagrange point and then time-shifted). These conditions are then reduced to the Newell et al. (2007) coupling function.

The model is combining a lognormal distribution describing ROTI conditioned by solar wind conditions and a Laplace distribution describing PPP error conditioned by ROTI index. The global trend of the model has been checked regarding the mean and standard deviation of the ROTI index obtained versus magnetic local time, magnetic latitude and solar wind conditions. Mean and standard deviation of the ROTI modeled by HAPEE follow the expected behavior with respect to solar wind conditions, i.e. an increase of the auroral oval size and width whereas the solar wind conditions increases.

This study has led to a new global model that can be improved in different points. Further complementary validations by comparisons with other disturbances or PPP error data would be very interesting. An extended version of the model should also be regressed following the same philosophy and using more data, covering different countries and other solar cycle. The statistical approach applied to derive the model is very global and refinements can be imagined. For example, no distinction is made between intense solar wind events such as interplanetary coronal mass ejections (ICMEs) and corotating interaction regions (CIRs) events, associated to high-speed streams (HSSs) from coronal holes. In future work, the type of Space Weather event at the origin of disturbances could be distinguished and the regressions adapted.

From the databases available and following the same approach, fits of other outputs could be proposed to complete the model, such as the S4 and σ_ϕ scintillation indices, or other error parameters.

Both the part of the model that forecasts ROTI based on solar wind parameters, and the part that forecasts PPP error based on ROTI, are empirical models. They are computationally expensive to construct but not computationally expensive to run afterwards. Input solar wind data can be retrieved in near-real-time from e.g. the data service of the Space Weather Prediction Center (SWPC) (currently available at this URL: https://services.swpc.noaa.gov/products/solar-wind/). A proof-of-concept prototype of an operational service has been implemented at NMA, retrieving the solar wind data from SWPC, running the models to generate a forecast map and visualizing this on a webpage. The data retrieval, processing and visualization add less than a minute of latency. The data latency from being measured by the satellite to appearing in the SWPC data file appears to be approximately a few minutes. Thus, forecasts can be made well in advance (typically 30 min to 1 h, depending on solar wind velocity) of the arrival of the solar wind structures. Discussions regarding establishing a full operational service are on-going.

Acknowledgements. We acknowledge use of NASA/GSFC's Space Physics Data Facility's OMNIWeb service, and OMNI data. The editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References

Andalsvik YL, Jacobsen KS. 2014. Observed high-latitude GNSS disturbances during a less-than-minor geomagnetic storm. *Rad Sci* **49(12)**: 1277–1288. https://doi.org/10.1002/2014RS005418.

Basu S, Groves K, Quinn J, Doherty P. 1999. A comparison of TEC fluctuations and scintillations at Ascension Island. *J Atmos Sol-Terr Phys* **61(16)**: 1219–1226. https://doi.org/10.1016/S1364-6826(99)00052-8.

Bertiger W, Desai SD, Haines B, Harvey N, Moore AW, Owen S, Weiss JP. 2010. Single receiver phase ambiguity resolution with GPS data. *J Geodesy* **84(5)**: 327–337. https://doi.org/10.1007/s00190-010-0371-9.

Cai C, Liu Z, Xia P, Dai W. 2013. Cycle slip detection and repair for undifferenced GPS observations under high ionospheric activity. *GPS Solut* **17**: 247–260. https://doi.org/10.1007/s10291-012-0275-7.

Carrano CS, Groves KM, Rino CL. 2019. On the relationship between the rate of change of total electron content index (ROTI), irregularity strength (C k L), and the scintillation index (S4). *J Geophys Res: Space Phys* **124(3)**: 2099–2112. https://doi.org/10.1029/2018JA026353.

Cherniak I, Zakharenkova I, Redmon RJ. 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. https://doi.org/10.1002/2015SW001237.

Cousins EDP, Shepherd SG. 2010. A dynamical model of high-latitude convection derived from SuperDARN plasma drift measurements. *J Geophys Res: Space Phys* **115(A12)**: 1–12. https://doi.org/10.1029/2010JA016017.

Cowley SWH. 2000. Magnetosphere-ionosphere interactions: A tutorial review. In: *Magnetospheric current systems*. Ohtani S-I, Fujii R, Hesse M, Lysak RL (Eds). AGU, Washington, DC, vol. 118, pp. 91–106.

Fabbro V, Jacobsen KS, Rougerie S. 2019. HAPEE, a prediction model of ionospheric scintillation in polar region. In: *3th European Conference on Antennas and Propagation (EuCAP)*, April 2019, Krakow, Poland. https://ieeexplore.ieee.org/document/8740307, pp. 1–5.

Ghobadi H, Spogli L, Alfonsi L, Cesaroni C, Cicone A, Linty N, Romano V, Cafaro M. 2020. Disentangling ionospheric refraction and diffraction effects in GNSS raw phase through fast iterative filtering technique. *GPS Solutions* **24(3)**: 1–13. https://doi.org/10.1007/s10291-020-01001-1.

Hu S, Bhattacharjee A, Hou J, Sun B, Roesler D, Frierdich S, Gibbs N, Whited J. 1998. Ionospheric storm forecast for high-frequency communications. *Rad Sci* **33(Sept–Oct)**: 1413–1428. https://doi.org/10.1029/98RS02219.

Jacobsen KS, Dähnn M. 2014. Statistics of ionospheric disturbances and their correlation with GNSS positioning errors at high latitudes. *J Space Weather Space Clim* **4**: A27. https://doi.org/10.1051/swsc/2014024.

Jacobsen KS, Schäfer S. 2012. Observed effects of a geomagnetic storm on an RTK positioning network at high latitudes. *J Space Weather Space Clim* **2**: A13. https://doi.org/10.1051/swsc/2012013.

Jacobsen KS, Andalsvik YL. 2016. Overview of the 2015 St. Patricks day storm and its consequences for RTK and PPP positioning in Norway. *J Space Weather Space Clim* **6**: A9. https://doi.org/10.1051/swsc/2016004.

Jin Y, Oksavik K. 2018. GPS scintillations and losses of signal lock at high latitudes during the 2015 St. Patrick's day storm. *J Geophys Res: Space Phys* **123(9)**: 7943–7957. https://doi.org/10.1029/2018JA025933.

Jin Y, Moen JI, Oksavik K, Spicher A, Clausen LBN, Miloch WJ. 2017. GPS scintillations associated with cusp dynamics and polar cap patches. *J Space Weather Space Clim* **7**: A23. https://doi.org/10.1051/swsc/2017022.

Kouba J, Héroux P. 2001. Precise point positioning using IGS orbit and clock products. *GPS Solut* **5(2)**: 12–28. https://doi.org/10.1007/PL00012883.

Koutroumbas K, Tsagouri I, Belehaki A. 2008. Time series autoregression technique implemented on-line in DIAS system for ionospheric forecast over Europe. *Ann Geophys* **26(2)**: 371–386. https://doi.org/10.5194/angeo-26-371-2008, https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2010JA015529.

Liu Z. 2011. A new automated cycle slip detection and repair method for a single dual-frequency GPS receiver. J Geod **85**: 171–183. https://doi.org/10.1007/s00190-010-0426-y.

Lui ATY, Perreault P, Akasofu SI, Anger CD. 1973. The diffuse aurora. *Planet Space Sci* **21(5)**: 857–858. https://doi.org/10.1016/0032-0633(73)90102-5.

McCaffrey AM, Jayachandran PT. 2019. Determination of the refractive contribution to GPS phase scintillation. *J Geophys Res: Space Phys* **124(2)**: 1454–1469. https://doi.org/10.1029/2018JA025759.

McGranaghan RM, Mannucci AJ, Wilson B, Mattmann CA, Chadwick R. 2018. New capabilities for prediction of high-latitude ionospheric scintillation: A novel approach with machine learning. *Space Weather* **16(11)**: 1817–1846. https://doi.org/10.1029/2018SW002018.

McPherron RL, Siscoe G. 2004. Probabilistic forecasting of geomagnetic indices using solar wind air mass analysis. *Space Weather* **2(1)**: S01001. https://doi.org/10.1029/2003SW000003.

Melbourne W. 1985. The case for ranging in GPS-based geodetic systems. In: *Proceedings, First International Symposium on Precise Positioning with the Global Positioning System*, Rockville, Maryland, April 15–19, 1985, pp. 373–386. https://ci.nii.ac.jp/naid/10003710300/en/.

Moreno B, Radicella S, Lacy MC, Herraiz M, Rodriguez-Caderot G. 2011. On the effects of the ionospheric disturbances on precise point positioning at equatorial latitudes. *GPS Solut* **15(4)**: 381–390. https://doi.org/10.1007/s10291-010-0197-1.

Mushini SC, Jayachandran PT, Langley RB, MacDougall JW, Pokhotelov D. 2012. Improved amplitude- and phase-scintillation indices derived from wavelet detrended high-latitude GPS data. *GPS Solut* **16(3)**: 363–373. https://doi.org/10.1007/s10291-011-0238-4.

Newell PT, Liou K, Zhang Y, Sotirelis T, Paxton LJ, Mitchell EJ. 2014. OVATION Prime-2013: Extension of auroral precipitation model to higher disturbance levels. *Space Weather* **12**: 368–379. https://doi.org/10.1002/2014SW001056.

Newell PT, Sotirelis T, Liou K, Lee AR, Wing S, Green J, Redmon R. 2010. Predictive ability of four auroral precipitation models as evaluated using Polar UVI global images. *Space Weather* **8(August)**: 1–9. https://doi.org/10.1029/2010SW000604.

Newell PT, Sotirelis T, Liou K, Meng C, Rich FJ. 2007. A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables. *J Geophys Res: Space Phys* **112(A1)**: A01206. https://doi.org/10.1029/2006JA012015.

Pi X, Mannucci AJ, Lindqwister UJ, Ho CM. 1997. Monitoring of global ionospheric irregularities using the Worldwide GPS Network. *Geophys Res Lett* **24(18)**: 2283–2286. https://doi.org/10.1029/97GL02273.

Poppe BB. 2000. New scales help public, technicians understand space weather. *Eos, Trans Am Geophys Union* **81(29)**: 322–328. https://doi.org/10.1029/00EO00247.

Prikryl P, Jayachandran PT, Mushini SC, Richardson IG. 2012. Toward the probabilistic forecasting of high-latitude GPS phase scintillation. *Space Weather* **10(8)**: S08005. https://doi.org/10.1029/2012SW000800.

Prikryl P, Sreeja V, Aquino M, Jayachandran PT. 2013. Probabilistic forecasting of ionospheric scintillation and GNSS receiver signal tracking performance at high latitudes. *Ann Geophys* **56(2)**: R0222. https://doi.org/10.4401/ag-6219.

Rino C, Morton Y, Breitsch B, Carrano C. 2019. Stochastic TEC structure characterization. *J Geophys Res: Space Phys* **124(12)**: 10571–10579. https://doi.org/10.1029/2019JA026958.

Romeo M, Da Costa V, Bardou F. 2003. Broad distribution effects in sums of lognormal random variables. *Eur Phys J B* **32(4)**: 513–525. https://doi.org/10.1140/epjb/e2003-00131-6.

Shepherd SG. 2014. Altitude-adjusted corrected geomagnetic coordinates: Definition and functional approximations. *J Geophys Res: Space Phys* **119(9)**: 7501–7521. https://doi.org/10.1002/2014JA020264.

Sigernes F, Dyrland M, Brekke P, Chernouss S, Lorentzen DA, Oksavik K, Deehr CS. 2011. Two methods to forecast auroral displays. *J Space Weather Space Clim* **1(1)**: A03. https://doi.org/10.1051/swsc/2011003.

Starkov GV. 1994a. Mathematical model of the auroral boundaries. *Geomagn Aeron* **34(3)**: 331–336.

Starkov GV. 1994b. Statistical dependences between the magnetic activity indices. *Geomagn Aeron* **34(1)**: 101–103.

Tiwari R, Bhattacharya S, Purohit PK, Gwal AK. 2009. Effect of TEC variation on GPS precise point at low latitude. *Open Atmos Sci J* **3**: 1–12. https://doi.org/10.2174/1874282300903010001.

Tsagouri I, Belehaki A. 2015. Ionospheric forecasts for the European region for space weather applications. *J Space Weather Space Clim* **5**: A9. https://doi.org/10.1051/swsc/2015010.

van der Meeren C, Oksavik K, Lorentzen DA, Rietveld MT, Clausen LBN. 2015. Severe and localized GNSS scintillation at the poleward edge of the nightside auroral oval during intense substorm aurora. *J Geophys Res: Space Phys* **120(12)**: 10607–10621. https://doi.org/10.1002/2015JA021819.

Wübbena G. 1985. Software developments for geodetic positioning with GPS using TI 4100 code and carrier measurements. In: *Proceedings, First International Symposium on Precise Positioning with the Global Positioning System*, Rockville, Maryland, April 15–19, 1985, pp. 403–412. https://ci.nii.ac.jp/naid/10010660316/en/.

Yang Z, Liu Z. 2016. Correlation between ROTI and Ionospheric Scintillation Indices using Hong Kong low-latitude GPS data. *GPS Solut* **20**: 815–824. https://doi.org/10.1007/s10291-015-0492-y.

Zumberge JF, Heflin MB, Jefferson DC, Watkins MM, Webb FH. 1997. Precise point positioning for the efficient and robust analysis of GPS data from large networks. *J Geophys Res: Solid Earth* **102(B3)**: 5005–5017. https://doi.org/10.1029/96JB03860.

Cite this article as: Fabbro V, Jacobsen KS, Andalsvik YL & Rougerie S 2021. GNSS positioning error forecasting in the Arctic: ROTI and Precise Point Positioning error forecasting from solar wind measurements. *J. Space Weather Space Clim.* **11**, 43. https://doi.org/10.1051/swsc/2021024.