

TECHNICAL ARTICLE

OPEN 3 ACCESS

Estimation of the drift velocity of Equatorial Plasma Bubbles using GNSS and digisonde data

Víctor Navas-Portella^{1,*}, David Altadill¹, Estefania Blanch², Mar Altadill³, Antoni Segarra¹, Víctor de Paula¹, Cristhian Camilo Timoté⁴, and José Miguel Juan⁴

Received 3 July 2024 / Accepted 2 December 2024

Abstract – Equatorial Plasma Bubbles (EPBs) play a crucial role in modulating plasma density and electron content within the equatorial ionosphere. In this work, we present an advanced and more robust version of the method developed by Blanch E et al. (2018, J Space Weather Space Clim., 8, 38-32) for detecting EPBs using data from the Global Navigation Satellite System (GNSS). The enhancements introduced in this version significantly improve the EPB detection process, achieving a notable reduction in the false positive rate compared to the previous approach. These refinements include the application of more rigorous statistical techniques to achieve a more accurate fit for the background Total Electron Content (TEC), leading to better characterization of EPBs through improved estimation of disturbance shapes. Applying the capabilities of this new method in a dense network of GNSS sensors, we have developed an interferometric procedure for estimating EPB drift velocities, including both speed and direction. This procedure provides valuable insights into the dynamic behavior of EPBs in the Caribbean region during 2014. Our analysis reveals a predominant eastward propagation pattern of EPBs, closely aligned with modified dip isolines. Furthermore, by integrating the results from the GNSS-based method with quasi co-located digisondes, we applied a conceptual model to estimate EPB velocities along their drift direction. This model has been tested across different geographical sectors and validated through comparisons with results from other independent studies. This cross-verification confirms the reliability of the methods for capturing EPB characteristics. This approach improves the precision of EPB detection and contributes to a deeper understanding of their spatiotemporal dynamics and behavior, providing a valuable framework for characterizing these phenomena in the equatorial ionosphere.

Keywords: Equatorial plasma bubbles detection / Drift velocity / GNSS and digisonde data

1 Introduction

Equatorial Plasma Bubbles (EPBs) are plasma depletions with respect undisturbed ambient plasma in the ionospheric F region primarily manifesting at low latitudes, especially in equatorial regions. EPBs arise from the nonlinear progression of the generalized Rayleigh-Taylor instability (Dungey, 1956) extending vertically along with the F region (Tsunoda et al., 1982; Kelley, 2009). These disturbances can exhibit density reductions exceeding two orders of magnitude compared to the background levels. EPBs extend along the geomagnetic field lines over several hundred kilometres in the magnetic north-south

direction and have at the bottom of the F region an east-west extension of a few tens of kilometers. These ionospheric irregularities play an important role in the modulation of plasma density and electron content, particularly in the post-sunset periods (Abadi et al., 2014; Bhattacharyya, 2022; Patil et al., 2023). The magnitude and frequency of occurrence of EPBs can be intensified as a result of space weather phenomena, such as geomagnetic storms (Karan et al., 2023). The consequences of EPBs extend beyond their intrinsic ionospheric nature, affecting the reliability and performance of satellite-based communication and navigation systems. The irregularities introduced by EPBs can lead to signal scintillation (Groves et al., 1997) and total electron content (TEC) variations, impacting the accuracy and stability of satellite signals passing through these disturbed regions (Moraes et al., 2018).

¹ Observatori de l'Ebre (OE), Universität Ramon Llull – CSIC, E-43520 Roquetes, Spain

² Institut d'Estudis Espacials de Catalunya (IEEC), E-08860 Castelldefels, Barcelona, Spain

³ Private Researcher, E-43500 Tortosa, Spain

⁴ Universitat Politècnica de Catalunya, UPC Campus Nord, E-08034 Barcelona, Spain

^{*}Corresponding author: vnavas@gencat.cat

EPBs can be identified as depletions in the TEC obtained from the Global Navigation Satellite System (GNSS) data (Haaser et al., 2012; Magdaleno et al., 2012). Significant efforts have been dedicated to develop techniques for identifying TEC depletions through the analysis of time series data from various GNSS quantities (Nishioka et al., 2008; Seemala & Valladares, 2011; Blanch et al., 2018; Mersha et al., 2020; Li et al., 2021). Generally, using specific assumptions and thresholds, these techniques have implemented different algorithms to automatically determine whether a TEC depletion can be associated to an EPB. More recently, machine learning approaches have also been used to detect EPBs from GNSS data (Kapil & Seemala, 2024). The occurrence of EPBs also manifests in the low latitude ionosondes measurements as F trace spreading in the ionograms (Lynn et al., 2011; McNamara et al., 2013) that can be attributed to multiple structured diffused patterns pointing to plasma irregularities (Booker & Wells, 1938; Bhattacharyya, 2022; Patil et al., 2023). The spreading of the F-region traces can be observed around the low-frequency section of the ionogram as diffused echoes of 10 km (or more) over the true range of the main echo. The latter is usually termed as Range Spread F (RSF). The spreading of the F-region traces can also be observed around the critical frequency of the F trace, generally termed Frequency Spread F (FSF). Both RSF and FSF pose difficulties for proper scaling of ionograms (Wakai et al., 1987). Calabia et al. (2024) used a multi-instrumental approach, satellite and ionosondes, to detect and observe the drift of an EPB occurred at Morocco in 2014. It must be mentioned that EPBs can also be detected from spacecraft Earth exploration data. Reddy et al. (2023) measure the Ionospheric Bubble Index (IBI) from magnetometers placed at the Swarm mission. Adkins & England (2023) work with Global-Scale observations of the Limb and Disk (GOLD) data to automatically detect and track

Some important efforts have been made to study the propagation velocity of EPBs, and most of the investigations have focused on individual events or at a regional scale. Haase et al. (2011) detected an EPB propagating at velocities of 85-110 ms⁻¹ over Brazil. Zonal drift velocities of small scale irregularities have also been studied by using GNSS receivers in the Brazilian region (Muella et al., 2013, 2014; Cesaroni et al., 2021). Ji et al. (2015) used the rate of TEC index (ROTI) as the main observable to identify EPBs and computed the detection time delay by cross-correlation of TEC time series from the Hong Kong continuously operating reference station network, thus obtaining velocity module and a mean direction correlating with the zonal drift. Sarudin et al. (2017) used ROTI keograms to investigate the drift velocity of EPBs for the Malaysian realtime kinematic GNSS network. Unnikrishnan et al. (2017) used multistation-single satellite observations to calculate the velocity and the east-west width of EPBs detected by receivers in India. Yu et al. (2018) were able to determine the velocity vectors of EPBs by considering a series of successive airglow all-sky images in Hainan (China), finding values in a range of 98–144 ms⁻¹. Karan et al. (2020) studied EPBs velocities from a geostationary orbit using GOLD data for a range of longitudes 60–25°W. Souza et al. (2021) and Silva et al. (2019) used pairs of very close GNSS receivers in Brazil to estimate EPB velocities by means of the detection time delay in the slant TEC (STEC). Working with two receivers can be enough to determine the module of the drift velocity, but in order to compute the EPB propagation direction one must use at least three receivers. Li et al. (2021) applied an improved rolling barrel algorithm to the TEC to identify EPBs and estimate their drift velocity by using combinations of three triangular receivers in Hong Kong for a particular event occurred in 2018. Christovam et al. (2023) studied several EPBs occurred in Brazil and found velocities of 100–150 ms⁻¹ by using single-frequency TEC keograms and of 75–250 ms⁻¹ by using airglow data. EPBs velocity module and direction are challenging parameters difficult to provide in an automatic way for a large range of longitudes (Magdaleno et al., 2012; Li & Zou, 2016).

This research aims at presenting an improved version in robustness of the method to detect EPBs developed by Blanch et al. (2018), and at expanding the capabilities of this method to derive the two dimensional drift velocity (module and azimuth) of EPBs. Moreover, a method to estimate the velocity and dimension of EPBs along their drift direction is also discussed. The latter uses the combined results of the above EPB detection over a single GNSS receiver and the results of the EPBs detection by quasi co-located ionosondes. The study of the velocity module and direction can help in the understanding of EPB dynamics in different regions.

This article is structured as follows: Section 2 describes the datasets of ionosondes and GNSS receivers used in this work. Section 3 describes the improvements of the new EPB detection method, and the methodology to calculate the EPB two-dimensional drift velocity. Section 4 presents and discusses the results for the improvements of the new EPB detection method and its performance compared to that of Blanch et al. (2018), the methodology applied to compute EPB velocities in a dense network of GNSS receivers in the Caribbean sector, and a conceptual model to estimate the EPB velocity along the drift direction through the integration of data from quasi co-located ionosonde and GNSS receivers. Finally, we expose the conclusions in Section 5. Additionally, some technical details regarding the methods used in this manuscript are presented as appendices.

2 Data

This work uses datasets from equator-to-low latitude GNSS receivers and ionosondes for various purposes. Figure 1 shows a map of all the ionosondes and GNSS receivers used in this work, and Tables 1 and 2 list their coordinates. GNSS data are utilized to obtain the TEC and to identify EPBs, as well as to measure their velocity using a dense network of receivers. Moreover, RSF signatures in the ionograms are used to identify the presence of EPBs passing over the ionosondes. For the purposes of this work, four different equatorial-to-low latitude sectors are evaluated: the American, the Caribbean, the Atlantic, and the Pacific. Analyses on the evaluation of the drift velocity of EPBs have been done for high solar activity periods. Data from the year 2002, at the solar activity maximum of solar cycle 23, are used for the analysis in the Atlantic sector. However, data from the year 2014, at the solar activity maximum of solar cycle 24, are used for the analysis in the American, Caribbean, and Pacific sectors. The latter is to ensure a simultaneous data availability for both GNSS and ionosonde sensors at similar solar activity levels, and to ensure the lowest distance between

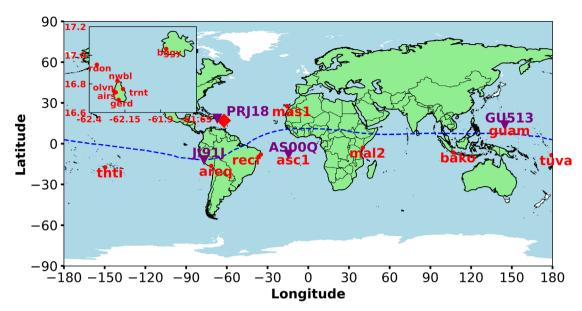


Figure 1. Map indicating the position of the GNSS receivers and ionosondes considered along this work. In the main plot: purple downward-pointing triangles labeled with uppercase letters are the DGSs listed in Table 1, red dots accompanied by lowercase letters correspond to the GNSS receivers listed in Table 2, and the blue dashed line corresponds to the magnetic equator. The inset is a zoomed-map of the red diamond symbol in the Caribbean region and illustrates the dense Caribbean network used to perform the velocity analysis of EPBs. Maps have been generated by means of the GeoPandas library in Python (Jordahl et al., 2020).

Table 1. List of the DGS stations with International Union of Radio Science (URSI) codes, geographical coordinates, cadence and year of the measurements in the studied interval used in Section 4.3.

Station	Sector	URSI Code	Lon (°)	Lat (°)	Cadence (min)	Year
Jicamarca	American	JI91J	-76.80	-12.00	5	2014
Ramey	American	PRJ18	-67.10	18.50	15	2014
Ascension Island	Atlantic	AS00Q	-14.40	-7.95	15	2002
Guam	Pacific	GU513	144.86	13.62	15	2014

Table 2. Location, codes, geographical coordinates, and year of analysis of the different GNSS receivers used along this work. Clovers, spades, and stars denote whether the receivers have been used for the comparison of EPB detection methods performance, for the computation of two-dimensional EPB velocities, or for the computation of EPB velocity along the drift by combining GNSS and DGS data, respectively.

Location (Region)	Code	Lon (°)	Lat (°)	Year
Papeete (Tahiti)	thti 🐥	-149.61	-17.74	2014
Arequipa (Peru)	areq 🜲 🛨	-71.49	-16.36	2014
Recife (Brazil)	recf 🜲	-34.95	-8.00	2014
Ascension Island (Saint Helena)	asc1 ★	-14.33	-7.92	2002
Maspalomas (Spain)	mas1 🜲	-15.63	27.61	2014
Malindi (Kenya)	mal2 🐥	30.09	-1.93	2014
Cibinong (Indonesia)	bako 🖡	106.85	-6.45	2014
Dededo (Guam)	guam 弗 🛨	144.87	13.50	2014
Funafuti (Tuvalu)	tuva 🐥	179.20	-8.47	2014
Redonda Island (Antigua and Barbuda)	rdon ♠ ★	-62.35	16.93	2014
Olveston (Montserrat)	olvn ♠ ★	-62.23	16.75	2014
Air Studios (Montserrat)	airs ♠ ★	-62.21	16.74	2014
North-West Bluff (Montserrat)	nwbl ♠ ★	-62.20	16.82	2014
Geralds Yard (Montserrat)	gerd ♠ ★	-62.19	16.69	2014
Trant's Bay (Montserrat)	trnt ♠ ★	-62.16	16.76	2014
Boggy Peak (Antigua and Barbuda)	bggy ♠ ★	-61.86	17.05	2014

GNSS and ionosonde sensors. This way, GNSS receivers and ionosondes in the Atlantic and Pacific sectors, can be considered to be co-located (see Table 1, GNSS receivers marked with a star in Table 2 and Fig. 1). The GNSS receivers and ionosondes in the American and Caribbean sectors are not strictly co-located. Nevertheless, for the purposes of this work, these pairs of sensors might be considered to be co-located based on the ionospheric correlation distances (Chen et al., 2023; Lv et al., 2024).

2.1 lonosonde data

The ionosonde data used in this study are available at the GIRO portal¹ (last accessed 2024.05.08). A Visualization/Editing Tool, known as SAO Explorer (SAO-X), is provided for working with GIRO ionograms measured by Digisondes (DGSs). SAO-X can be accessed and downloaded at the website of the Center for Atmospheric Research² (last accessed 2024.05.08). Users of SAO-X can connect to the Digital Ionogram Database³ (last accessed 2024.05.08), which is an Internet-accessible archive of DGS ionograms for many locations and periods of time. SAO-X allows for an easy way to extract information from the ionograms recorded by DGSs and in particular of the RSF signatures in ionograms. Figure 2 shows three different ionograms to illustrate the RSF associated with a passing EPB over the Guam station on September 22, 2014. The left panel displays an ionogram under quiet conditions at 8:45 UT, prior to the occurrence of the EPB. The central panel represents the ionogram recorded at 9:45 UT, when the maximum RSF for this event is observed. The right panel corresponds to the ionogram recorded at 11:00 UT, where the RSF is clearly diminished, indicating the end of the EPB over this station.

2.2 GNSS data

For the purposes of this work, GNSS data format of Receiver Independent Exchange (RINEX) and satellite orbit data in Standar Product 3 (SP3) format files are used to estimate the STEC between a GNSS receiver and any of the active satellites (up to 32) of the GPS constellation (Sanz et al., 2013). RINEX and orbit files have data samplings of 30 s and 15 min, respectively. These files are public and available at the Crustal Dynamics Data Information System Data Center.⁴

Table 2 shows the different GNSS receivers used along this work. Depending on which analysis is performed, different receivers are used. Receivers marked with clovers are used for validation of the new detection method in Section 4.1. These receivers have been chosen to be approximately close to the magnetic equator and approximately uniformly distributed in longitude for a year of high solar activity (2014). Although the new EPB detection method is also able to analyze signals from Galileo and Beidou satellites, to ensure an equitable comparison with (Blanch et al., 2018), the analysis has exclusively considered satellites from the GPS constellation. Receivers marked with spades correspond to a dense network of GNSS receivers located at three closely located islands in the

Caribbean are used in Section 4.2 to compute the two-dimensional EPBs velocities. Those receivers marked with stars are used to estimate the velocity of EPBs along the drift direction through the integration of data from quasi-located ionosondes in Section 4.3.

3 Methodology

The methodology section is structured into three different parts. Section 3.1 details the procedure for deriving the time series of the TEC obtained from GNSS signals. Section. 3.2 introduces the new method of detection of EPBs, along with the improvements over the Blanch et al. (2018) version. Section 3.3 describes the method for computing the two-dimensional EPB velocity from the output of the new EPB detection method.

3.1 GNSS signal processing

The ionospheric combination of the dual-frequency carrier-phase and code delays of GPS observations, LI = L1 - L2 and PI = P2 - P1 respectively, have been used to obtain an ionospheric observable related to the STEC (Hernández-Pajares et al., 2006). However, a significant limitation when working with STEC is the requirement for an elevation mask of 50°, as stated by Hernández-Pajares et al. (2006), for the ionospheric pierce point (IPP) to avoid distortions that could lead to false positives. That is why we followed the approach of Blanch et al. (2018), which does not use STEC with a 50° elevation mask but analyzed the standard deviation of the second derivative of the TEC. This way one can retain a substantial amount of data necessary for the analysis. The time series of TEC(t) are obtained from the STEC by means of the obliquity factor:

$$TEC(t) = \sqrt{1 - \left(\frac{R_{\rm E}\cos\epsilon(t)}{R_{\rm E} + H_{\rm ion}}\right)^2} STEC(t), \tag{1}$$

where $\epsilon(t)$ corresponds to the satellite elevation at time t, $R_{\rm E}$ is the Earth radius, and $H_{\rm ion}$ is the altitude at which the ionosphere is considered as a thin layer ($H_{\rm ion} = 350$ km along this work) (Mannucci et al., 1993; Kumar & Singh, 2009).

Similar to previous approaches (Hernández-Pajares et al., 2006; Magdaleno et al., 2012), we estimate the TEC using the geometry-free combination of the ground-based GNSS carrier-phase delays. This approach yields significantly lower noise levels compared to TEC estimation derived from the geometry-free combination of code delays. The TEC estimated from carrier-phase delays is much more precise than the TEC estimated from code delays. However, it is well-established that carrier-phase measurements are affected by arbitrary carrier phase ambiguities leading to discontinuities (cycle-slips) in the TEC estimation and potentially resulting in inaccurate EPB parameter estimates. These cycle-slips can be a consequence of ionospheric irregularities, particularly scintillation, affecting the GNSS signal tracked by the receiver. This involves utilizing TEC data derived by smoothing the unambiguous but noisy code of the GNSS signal. Specifically, a smoothing technique to the TEC derived from the geometry-free combination of the code delays is applied by employing a 5-point running mean. This approach is employed to enhance TEC information and establish connections between TEC estimates obtained from

¹ https://giro.uml.edu/

² https://ulcar.uml.edu/SAO-X/SAO-X.html

³ https://giro.uml.edu/didbase/

⁴ ftp://cddis.gsfc.nasa.gov/pub/gps/data/daily/

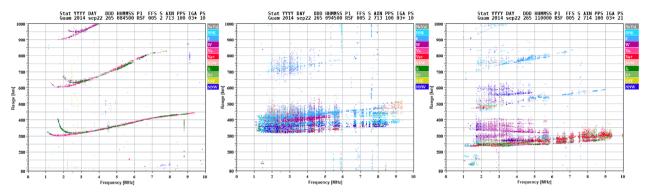


Figure 2. Examples of ionograms recorded at Guam on September 22, 2014. From left to right: before, during and just ending the passing of an EPB over the measuring site.

the phase delays in different arcs. Nevertheless, the noise of the code delays can be about 100 times larger than the noise on carrier-phase delays at low elevation, because of the increased noise due to the multi-path in PI delay. Such noise may be misleadingly interpreted as an additional ionospheric activity. To mitigate this, an elevation mask of 20° for all estimated TEC values derived from code delays are introduced.

Once the TEC(t) time series for one satellite is constructed, it is then ready to be used as input for the new EPB detection method.

3.2 Automatic EPB detection method

Blanch et al. (2018) developed a method to detect the occurrence of EPBs by adapting an existing detector of medium-scale traveling ionospheric disturbances (MSTIDs) (Hernández-Pajares et al., 2006), which analyzed distinct changes in the temporal variations of the second TEC derivative (TEC"(t)). The nominal variations of the TEC can present TEC rates of non-negligible values, which in turn can make the identification of the EPB more difficult by focusing on distinct changes in the variations of the first TEC derivative, TEC'(t). Blanch et al. (2018) method focuses on the standard deviation of the second derivative of the TEC over sliding time windows of a given length of n values:

$$\sigma(\text{TEC}''(t)) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\text{TEC}''(t+i) - \overline{\text{TEC}''} \right)^{2}}, \quad (2)$$

where $\overline{\text{TEC''}}$ corresponds to the mean value of $\overline{\text{TEC''}}(t)$ in the time window. Moreover, analyses comparing $\sigma(\overline{\text{TEC''}}(t))$ with that of $\overline{\text{TEC'}}(t)$ shown that the former remains more stable under quiet intervals than the latter, making easier the identification of depletions on the TEC (Blanch et al., 2018). A window of n=20~(600~s) is used as a first approximation to calculate the $\sigma(\overline{\text{TEC''}}(t))$, similarly to that proposed by Magdaleno et al. (2012). Initial and final times of TEC depletions, T_i and T_f are considered when $\sigma(\overline{\text{TEC''}}(t))$ overcomes and subsequently undergoes a predefined threshold. Although Blanch et al. (2018) analyzed the optimal threshold value for bounding depletions over time, the abrupt termination of the event upon undergoing the threshold may result in missing important portions of the depletion or event splitting. Once the TEC depletion is bounded in time, the points at

the boundaries $\text{TEC}(T_i)$ and $\text{TEC}(T_f)$, and the values of TEC' at these points are used to estimate the background TEC, $\text{TEC}_0(t)$, by fitting a parabolic function. A set of criteria based on $\text{TEC}_0(t)$ is applied to determine whether the TEC depletion is significant and can be associated with an EPB. Note that this approach could lead to anomalous estimations of the background $\text{TEC}_0(t)$ if the TEC(t) is noisy near the EPB boundaries. This, in turn, could result in inaccurate estimation of the depletion observables, potentially leading to false positives or false negatives. By using this method, Blanch et al. (2018) performed a climatological study of the main EPB observables for different International GNSS Service (IGS) receivers.

Some refinements can be performed in order to reduce the ratio of false positives thus conferring more reliability to the automatic detection. The initial improvement proposed in this work is devoted to the method's versatility, and consists in rewriting all those parts of the Blanch et al. (2018) detection method that were previously coded in Matlab (The MathWorks Inc., 2022) to Python (Van Rossum & Drake, 2009). This change of programming language fulfills the FAIR (Findable, Accessible, Interoperable and Re-usable) principles required by the Plasmasphere Ionosphere Thermosphere Integrated Research Environment and Access services: a Network of Research Facilities (PITHIA-NRF) project (Belehaki et al., 2023). Furthermore, some modifications have been performed in the GNSS data-acquisition and post-processing module to allow for multiconstellation analysis by adding signals from Galileo and Beidou constellations to that of GPS.

The second improvement focuses on the inclusion of a relaxation time with the aim of easing the strict imposition of a threshold to define EPB duration used in the previous approach. In order to do this, a Hit Definition Time (HDT) is imposed as a relaxation time to consider the event as finished (Piñal Moctezuma et al., 2019; Navas-Portella, 2020). In the previous version of the method, the final time T_f of the depletion was considered strictly when the time series of $\sigma(\text{TEC}''(t))$ first undergoes a predefined threshold. In this new version of the method, an additional condition is added: the final time is considered as the time when $\sigma(\text{TEC}''(t))$ undergoes a threshold and remains below it for at least the HDT. Consequently, this HDT might alter the EPB final time T_f and the duration. Note that, if the HDT is set too high, the method could consider two or more depletions as one. Conversely, if the HDT is set too low, the

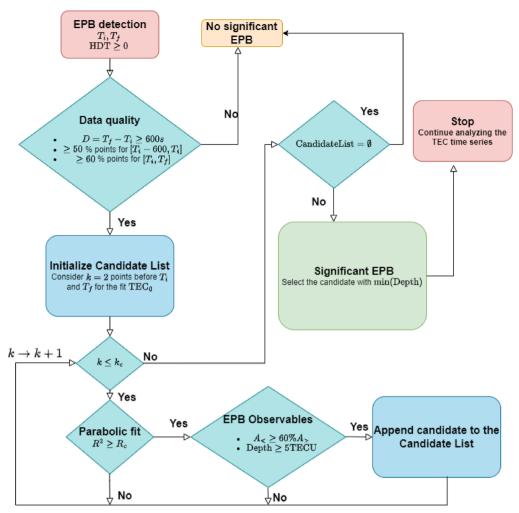


Figure 3. EPB detection algorithm flow diagram. The flow starts when the potential EPB has been bounded in time $[T_i, T_f]$ according to the thresholds for $\sigma(\text{TEC''}(t))$ and the HDT condition. R^2 and R_c^2 correspond to the fit correlation coefficient and its threshold value. $A_>$ and $A_<$ correspond to the positive and negative areas of the integrated TEC in the interval $[T_i, T_f]$. For more details see Appendix A. The flow finishes when the potential EPB has been considered as significant or not, and the procedure would continue by detecting new potential EPBs from the TEC time series.

new method may detect a single depletion as multiple ones. Blanch et al. (2018) method requests a minimum duration of 600 s for EPBs. That is, if the depletion is shorter than 600 s, it cannot be associated to an EPB. To be in agreement with this time, a HDT of 600 s has been used. It has been studied that the distribution of observables shows no significant differences for different values of HDT (Lebyodkin et al., 2013). Various thresholds have been tested, and no significant differences in the main EPB observables have been found.

The third improvement is based on the enhancement of the statistical methods and criteria to obtain a better estimation of the background TEC, $\text{TEC}_0(t)$. It is essential to ensure an accurate estimation of the background TEC in the vicinity of T_i and T_f to determine whether a depletion is significant and can be associated to an EPB. In order to correctly estimate the background $\text{TEC}_0(t)$, a list of candidate fit is considered by taking different number of points in the vicinity of T_0 and T_f . For each set of points, a parabolic fit is performed, resulting in various estimations of the background $\text{TEC}_0(t)$ in the vicinity of the potential EPB.

The second and third improvements are motivated by two objectives: the first one is to reduce the ratio of false positives with respect to Blanch et al. (2018), and the second one is to obtain a more accurate shape of the disturbance caused by an EPB. Reduction of false positives are necessary to make automatic detection more reliable. An accurate shape of the disturbance is important in order to have correctly estimated values of the observables characterizing the EPB.

The new EPB detection method algorithm is shown in the flowchart in Figure 3 and explained in more detail in Appendix A. The most important differences of this algorithm with respect to the Blanch et al. (2018) algorithm are the presence of a HDT to determine the EPB duration and an additional flow for the fit of the background TEC, TEC_0 .

It is important to note that this new method detects the TEC depletions generated by the presence of an EPB and it is not able to uniquely identify an EPB due to the lack of spatial information to bound the EPB. Consequently, multiple TEC depletions originated by the same EPB may be identified from distinct TEC time series of different satellites. The spatial

SIGNIFICANT BUBBLE detected at PRN 4 at station area

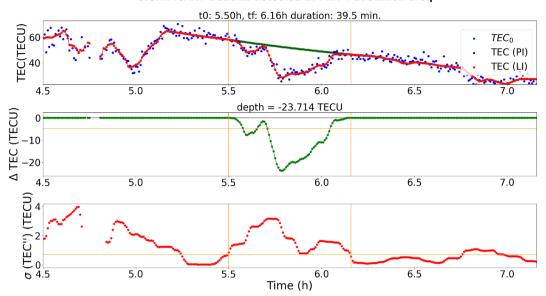


Figure 4. Example of the output plot from the EPB detection method obtained at the Arequipa (areq) receiver in communication with satellite PRN 4 on February 26, 2014. The top panel shows the TEC computed from the code delays (PI) in blue, the TEC obtained from the carrier-phase delays (LI) in red, and the fitted background TEC, TEC₀ during the EPB duration in green. The second panel shows the disturbance Δ TEC(t) in green and the significance threshold for the depth set at -5 TECU in the orange horizontal line. The third panel shows the evolution of the standard deviation of the TEC (in LI) second derivative, σ (TEC"(t)), in red, and the orange horizontal line at 0.714 TECU corresponds to the threshold used to determine the EPB initial and final times T_i and T_f are indicated by orange vertical lines in all the panels.

characteristics of EPBs could be assessed through a dense satellite network. However, such considerations lie beyond the scope of this study. Although different satellites may detect the same EPB, these instances will be treated as distinct EPBs.

Figure 4 shows a typical output plot of the new detection method with the TEC time series, the EPB disturbance curve computed from the fit TEC_0 and the evolution of $\sigma(\text{TEC''}(t))$. This output plot focuses on a particular detection centering it at the center of the plot by also plotting one hour before and after the initial and final times, respectively. The depletion $\Delta \text{TEC}(t)$ is only shown for the particular event. If the method does not identify any EPB for the analyzed day, $\Delta \text{TEC}(t)$ is zero for any value t in the daily time domain.

Let us suppose that a number *b* of EPBs have been detected by a GNSS receiver for a given satellite (sat) along a day. In this situation, the disturbance curve is defined as:

$$\Delta \text{TEC}(t) = \begin{cases} 0 & 0 \leq t < T_i^{(1)} \\ \text{TEC}(t) - \text{TEC}_0^{(1)}(t) & T_i^{(1)} \leq t \leq T_f^{(1)} \\ 0 & T_f^{(1)} < t < T_i^{(l)} \\ \text{TEC}(t) - \text{TEC}_0^{(l)}(t) & T_i^{(l)} \leq t \leq T_f^{(l)} \\ \dots \\ \text{TEC}(t) - \text{TEC}_0^{(b)}(t) & T_i^{(l)} \leq t \leq T_f^{(b)} \\ 0 & T_f^{(b)} < t \leq 86400, \end{cases}$$

$$(3)$$

being l the index running for the b EPBs, and 86400 the total number of seconds in a day. Note that, by construction of the

new detection method, EPBs cannot overlap each other, and there always will be an interval where the value of $\Delta TEC(t)$ will be strictly zero between consecutive EPBs. This method also provide plots of the TEC(t) and $\Delta TEC(t)$ as well as a map with the trajectory of the IPP during the EPB duration.

The end-to-end EPB new detection tool can be found as an application programming interface (API) in the e-Science Center of the PITHIA-NRF project.⁵

3.3 EPB drift velocity calculation

By considering a dense network of GNSS receivers together with the output of the new detection method, we propose here a method to estimate EPB drift velocities. Due to local nature of EPBs, a dense network of GNSS receivers is needed to localize the event and compute its drift velocity. Similar to Souza et al. (2021), our approach is based on finding the time delay between the detection of the same EPB for two different receivers by finding the time lag at which the cross-correlation of the two curves attains its maximum. If this is done with at least three curves, two velocity components can be computed by assuming that the EPB propagates as a planar wave (Wu et al., 2017).

Let us consider a set of disturbance curves $\{\Delta TEC(t)_j\}_{sat}$ for $j=1,\ldots,n_{RCV}$ different GNSS receivers related to the same satellite (sat). If a subset of m ($1 < m \le n_{RCV}$) of these curves exhibit a reasonable cross-correlation among them for a certain time window, one can assume that these m GNSS receivers are identifying the same EPB. Based on this approach, the

⁵ https://esc.pithia.eu/data-collections/DataCollection_OE_EPB_detectionTool/api/

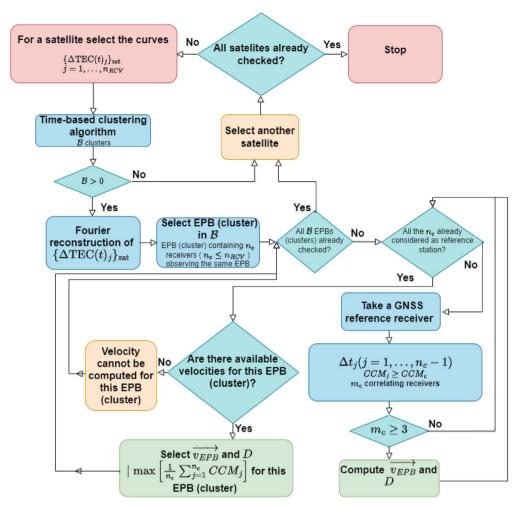


Figure 5. EPB drift velocity calculation flow diagram. The flow starts by considering the depletion curves $\Delta TEC(t)$ computed by the new EPB detection method for a set of satellites at different GNSS receivers. The time-based clustering algorithm is detailed in Appendix C. The flow finishes when all the curves for all the satellites have been checked.

algorithm used to calculate the drift velocity for EPBs is shown in the flowchart in Figure 5 and explained in more detail in Appendix B.

It must be mentioned that network geometry plays a very important role in the determination of the EPB drift velocity. In fact, provided that GNSS data are sampled every 30 s and the distance among GNSS receivers are known, one can estimate the maximum EPB velocity that can be detected by using that network. If a computed velocity exceeds this limit, then it will not be considered as valid.

4 Results and discussion

Section 4.1 presents and discusses the results for the improvements of the new EPB detection method and its performance compared to that of Blanch et al. (2018). The methology to compute EPB velocities based on the output of the new EPB detection method is applied in a dense network of GNSS receivers in the Caribbean sector in Section 4.2. Section 4.3 outlines the conceptual model to estimate the EPB velocity along the drift direction through the integration of data

from quasi co-located ionosonde and GNSS receivers and the results obtained from this model.

4.1 Performance of the new automatic EPB detection method

The goal of this analysis is to compare the differences in performance and EPB observables obtained from the Blanch et al. (2018) and the new version of the detection method. For the GNSS receivers marked with clovers in Table 2, we have been able to identify significant differences in observable distributions and performance.

The method by Blanch et al. (2018) automatically detected 1873 EPBs for 2014 considering all the above GNSS receivers, whereas the new method detected 1724 EPBs. This difference can be attributed to the stricter conditions imposed on the TEC background fit, which might have rejected some EPBs automatically detected by the Blanch et al. (2018) method. The two compared methods provide false positive and false negative detection of EPBs. False negative events (undetected EPBs) are out of the scope of this study and would need truth information of EPBs events. False positives are considered as those cases in

Table 3. Comparison of the automatically detected EPBs by the Blanch et al. (2018) method and the new method for the different GNSS receivers in Table 2. True positive EPBs correspond to the number of events that have been checked for each version of the detection method.

	Blan	Blanch et al. (2018) method			New method		
Receiver	Automatically detected EPBs	True positive EPBs	False positive rate (%)	Automatically detected EPBs	True positive EPBs	False positive rate (%)	
areq	181	169	6.6	192	186	3.1	
bako	147	87	40.8	151	129	14.6	
guam	180	123	31.7	142	107	24.6	
mal2	377	355	5.8	373	362	2.9	
mas1	248	246	0.8	227	222	2.2	
recf	464	415	10.6	332	327	1.5	
thti	72	63	12.5	84	80	4.8	
tuva	204	197	3.4	223	222	0.4	
Total	1873	1655	11.6	1724	1635	5.2	

which an EPB was detected when the irregularity was not really related to an EPB. The main source of false positives is the presence of cycle-slips, which were not previously excluded and fulfill the conditions to be considered as significant EPBs. Under these assumptions, false positives have been identified by visually inspecting all the automatically detected cases for the two detection methods under analysis. Table 3 shows detailed results about the number of events of the automatically detected EPBs and the number of false positives using data of the indicated GNSS receiver. The Blanch et al. (2018) method detects 1655 true EPBs whereas the new method detects 1635 true EPBs. Hence, the measured ratio of false positives for the Blanch et al. (2018) method is $11.6 \pm 0.7\%$, whereas the new method presents a lower ratio of false positives, $5.2 \pm 0.5\%$, both measured with a confidence interval of one standard deviation. Once false positives are removed for each method, the overall number of true positive EPBs for all the receivers is very similar. However, the number of true positives detected at each receiver exhibits some differences due to some effects. The first effect is attributed to the inclusion of more points in the fit, resulting in the identification of significant EPBs that were not detected when only the two points at the boundaries were considered. This can be observed in tuva, bako, areq, and thti receivers, where an increase in the number of true positive EPBs with the new method is found in comparison with its predecessor. Conversely, a second effect linked to the presence of the HDT leads to the merging of events. Some EPBs that were considered separately in the previous method might have merged into a single one due to the presence of the relaxation time HDT. This implies a reduction in the number of true positive EPBs detected by the new method compared to the previous method. As it is depicted in Figure 6, the top plots correspond to two different EPBs detected by the Blanch et al. (2018) method. The time between the ending of the first EPB and the beginning of the second one is of 390 s (6.5 min). Provided that this time is below the HDT of 600 s (10 min) used in the new detection method, these two EPBs will be considered as a single disturbance (bottom plot in Fig. 6). The trade-off between these two effects results in an overall number of true positives detected by the new method, which is slightly lower than that detected with the Blanch et al. (2018) method.

The two-sample Kolmogorov-Smirnov test (Feller, 1948; Press et al., 1996) is used to determine if the distribution of an EPB observable is significantly different. At a 95%

confidence level, this test unveils that the distribution of durations and depths are significantly different (p-values of 7.11×10^{-15} and 1.06×10^{-6} , respectively). However, the distribution of areas and initial times are not significantly altered (p-values of 0.43 and 0.25 respectively). Graphical representation of these distributions are shown in Figure 7.

Figure 7 displays probability density functions (PDFs) on the left and complementary cumulative distribution functions (CCDFs) on the right. The green plots represent the Blanch et al. (2018) method, while the dark blue plots represent the new detection method. These plots show durations, depths, total disturbance areas, and initial times across panels from top to bottom. Note that the CCDF of the initial times is referred to the detected EPBs from the indicated time until the end of the day. As it can be seen in the CCDF of durations in Figure 7 (right panel in the first row), the probability of finding EPBs with a duration between 25 and 125 min is clearly larger for the detected EPBs with the new method. These differences in the duration distribution can be easily explained by the presence of the HDT. Figure 8 shows one example of the same disturbance identified by the Blanch et al. (2018) method (first three panels), and with the current method (second three panels). As it can be seen in the third panel, the $\sigma(TEC''(t))$ undergoes the threshold of $\sigma_{\rm C}({\rm TEC''}(t)) = 0.714$ TECU, and the Blanch et al. (2018) configuration considers that as the ending time t_f of the EPB. However, the $\sigma(\text{TEC}''(t))$ reaches the threshold after 90 s (1.5 min), and the new method considers that as a part of the disturbance because the signal has not remained below the threshold longer than the HDT. Once the $\sigma(\text{TEC}''(t))$ has undergone the threshold and it does not overcome it again in the next HDT, the EPB is considered to be finished. This casuistic is in agreement with the previously explained EPBs merging effect. However, the case illustrated in Figure 8 shows that the extra part of the EPB detected by the new method would not fulfill the conditions imposed by the Blanch et al. (2018) to be significant by itself.

As it can be observed in the CCDF of depths (in absolute value) in Figure 7 (right panel in the second row), the probability of finding EPBs with a maximum depth (in absolute value) between 10 and 50 TECUs is clearly larger for the Blanch et al. (2018) detection method. These differences can be explained by the new conditions in the background TEC fitting procedure. Although the fitting function is still a second-degree polynomial, the inclusion of more points into the fit, together with

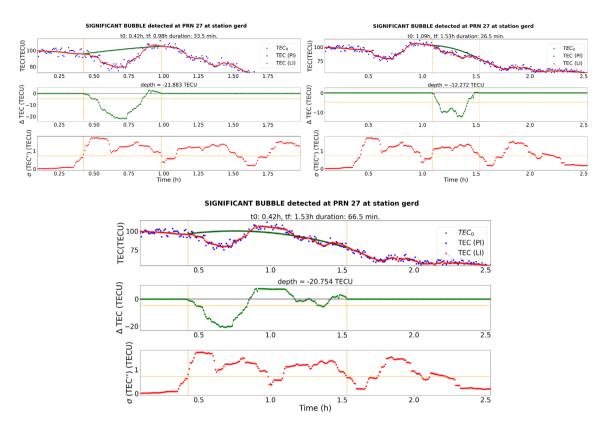


Figure 6. Top panels correspond to the different plots for two different EPBs detected by the Geralds Yard (gerd) receiver in the communication with satellite PRN 27 on February 27, 2014, with Blanch et al. (2018) configuration. The two EPBs are separated by 6.5 min (600 s or 10 min of HDT). Bottom panel corresponds to a unique EPB detected with the new method encompassing the two previously detected EPBs into a single one for the same receiver and date.

the selection of the background TEC fit with the smallest significant depth (in absolute value), has a significant effect on the distribution of depths.

No significant differences are found for the area and initial time distributions (third and fourth rows in Fig. 7). The non-significant difference in areas can be explained by an interplay between the new trend of larger EPB durations and the reduction of the disturbance depth, which has led to a conservation of the total disturbance area. The non-significant difference in initial times indicates that the inclusion of the HDT has a greater impact on the final times than on the initial times.

4.2 EPBs drift velocity from GNSS

The new detection method has been used together with the velocity analysis presented in Section 3.3 to compute the velocity for 28 EPBs detected in 2014 in the Caribbean region with the corresponding dense network of GNSS receivers marked with spades in Table 2. The average velocity of the EPBs is $105 \pm 5 \text{ ms}^{-1}$, with an azimuth of about $75 \pm 7^{\circ}$ from true North. The estimated average size of EPBs is about 302 ± 41 km. Error are given at one standard deviation. Table 4 shows detailed results about the detection of EPBs and their respective drift velocities. Appendix D shows the plots of the TEC disturbance curves involved in the calculation for each EPB listed in Table 4.

Bearing in mind that the new detection method detects the TEC depletions generated by an EPB and it is not able to

spatially discriminate EPBs, the same EPB can be identified by different satellites. By inspecting the figures in Appendix D, it can be observed that some disturbance curves detected by different satellites occurring at the same day and time window are very similar, and thus, they could be associated with the same EPB. This occurs especially in those days in which several events are reported by different satellites. In order to avoid artifacts caused by redundant velocity computation for the same EPB, we have grouped velocity and azimuth results obtained from distinct satellites within overlapping time windows, treating them as a unified event. Table 4 shows the events grouped according to this criterion delimited by horizontal divisions. Each event is then characterized by the mean values of the velocity and azimuth. This post-analysis procedure yields a dataset of 16 EPBs leading to a mean velocity of $100 \pm 5 \text{ ms}^{-1}$ and a mean azimuth of $73 \pm 5^{\circ}$. These results are compatible with those reported by considering all the different events detected by different satellites. The estimated size of the perturbation along its drift obtained with our model ranges from 5 to 550 km (see Table 4). These results are consistent with those reported in the literature (Kelley, 2009; Haase et al., 2011). Moreover, the obtained azimuth direction of propagation closely aligns with the inclination of the modified dip parallels (modip). The modip latitudes, introduced by Rawer (1963), come near the geomagnetic inclination at low latitudes and get closer to the geodetic latitude as latitude increases. These results suggest the EPBs propagate eastward and quasiparallel to modip isolines. We note here that EPBs have been

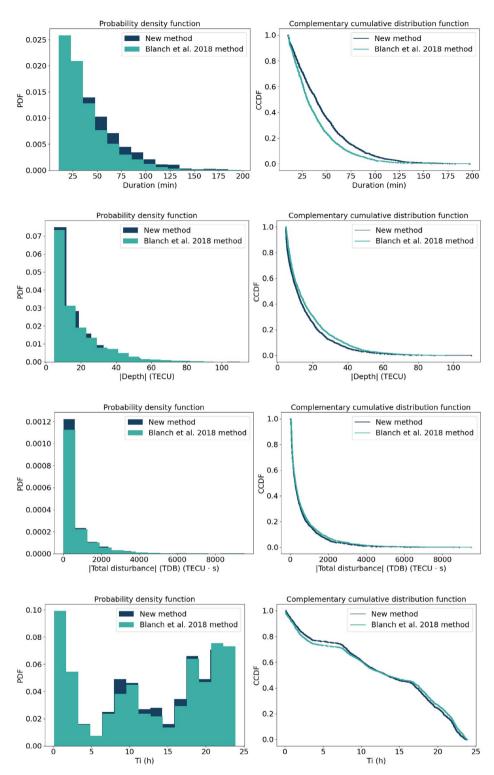
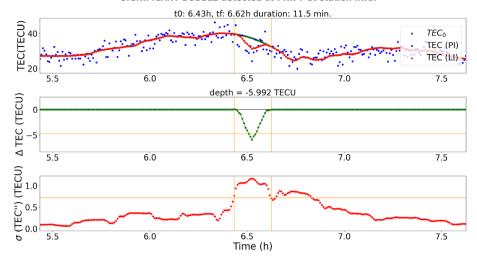


Figure 7. From top to bottom, PDFs (left panels) and CCDFs (right panels) for EPB Duration, total disturbance area, depth, and initial time T_i are shown for the Blanch et al. (2018) method (light blue) and the new detection method (dark blue). Total disturbance area and maximum depths are represented in absolute values for a better reading.

SIGNIFICANT BUBBLE detected at PRN 7 at station nwbl



SIGNIFICANT BUBBLE detected at PRN 7 at station nwbl

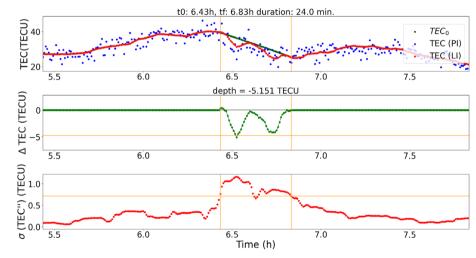


Figure 8. Top three panels correspond to the different plots generated with the Blanch et al. (2018) method for an EPB detected at the North-West Bluff (nwbl) receiver in communication with satellite PRN 7 on January 30, 2014. Three bottom panels correspond to the analogue plots generated with the new method.

found to propagate parallel to the modip isolines at the Indian sector (Unnikrishnan et al., 2017), which is in agreement to what we observe at the Caribbean sector.

4.3 Combining GNSS and digisonde data

We present a conceptual model for estimating EPB velocity along the drift direction through the integration of data from colocated DGSs and GNSS receivers. To do so, we first determine how to detect the passage of an EPB over an ionosonde. For this purpose we look at how the RSF evolves over time in the recorded ionograms (see Fig. 2). Therefore, we consider the EPB starting time detected by a DGS, $T_i^{\rm DGS}$, when the RSF in an ionogram is greater than 80 km for all the scanned frequencies between 3 and 5 MHz, and we consider the ending time of detection of an EPB by a DGS, $T_f^{\rm DGS}$, when the RSF is less than 80 km. Figure 9 shows the evolution of the RSF measured by the ionosonde in Guam for a given day. Three EPB events were

detected over Guam ionosonde, the first one starting at 9:15 UT and ending at 10:45 UT (see ionograms in Fig. 2).

The DGSs are ionosondes that can provide information about the region over which the signal bounces, thanks to their capability to estimate the angle of arrival (Reinisch et al., 2005). These records are known as skymaps (see Fig. 10). The "skymap" is a polar coordinates representation of the ionospheric reflector, where the radius indicates the distance of the bounce point of a radio signal from the vertical (concentric circles or zenith angle), and the angle of the polar diagram indicates the direction from the magnetic north (azimuth) of such a bounce point. In the ideal case that the ionosphere was a perfect reflector, all the bounces would come from the same point or narrow region, and if it did not present any tilt with respect to the surface of the Earth, all the bounces would come from the center of the circle.

The skymap example in Figure 10 (left) shows that most of the points come from a region concentrated within the 5° zenith

Table 4. Complete table with all the events detected by the new detection method for the dense network in the Caribbean, as exposed in Section 2.2. Each EPB is identified for a day of the year (DOY), at an initial time $t_0(h)$ in communication with a satellite from the GNSS constellation with Pseudo-random noise (PRN) code. v_{EPB} corresponds to the velocity module and z_{EPB} to the azimuth angle from the true North, with sub-indexes corresponding to the equatorial plasma bubble (EPB), and D is the estimated front distance in km. Horizontal lines group events that are identified on the same day by different satellites and are coincident in time, and thus, could be considered as the same EPB.

DOY	t_0 (h)	PRN	Duration (min)	$v_{\rm EPB}~({\rm ms}^{-1})$	z _{EPB} (°)	D (km)
030	6.43	07	24.5	47.22	57.61	111.67
057	3.45	01	25.5	104.64	79.14	138.36
057	2.70	11	15.5	102.25	73.43	74.24
057	2.63	07	36.0	100.18	67.39	374.16
057	2.06	11	12.0	167.67	99.34	119.24
057	1.61	13	39.0	53.06	224.54	3.38
058	0.73	03	53.0	101.17	66.49	226.79
058	0.51	19	63.5	179.92	7.65	440.35
058	0.43	27	67.0	110.45	72.69	359.34
063	0.83	03	46.5	95.33	85.15	206.69
063	0.45	19	95.5	109.85	72.24	483.45
063	0.34	23	79.0	137.62	44.18	858.75
063	1.22	27	19.5	95.50	95.04	82.51
289	1.19	21	68.5	141.69	359.46	815.59
289	2.29	25	22.0	95.02	98.21	131.97
307	1.12	25	34.5	102.33	95.56	216.95
311	0.73	18	47.0	114.91	82.99	372.44
311	0.36	25	21.5	111.87	71.02	108.73
311	0.82	25	54.5	98.97	95.35	330.96
322	1.32	22	75.5	104.60	63.99	593.16
322	0.26	25	42.5	92.79	92.13	235.06
325	1.08	22	58.5	73.65	63.10	353.83
332	0.80	22	17.5	113.24	61.23	147.65
334	0.57	22	17.0	108.57	60.95	139.07
334	1.79	22	68.0	106.59	78.29	524.58
336	0.46	18	46.0	87.68	57.79	376.12
336	0.35	22	59.5	101.94	61.40	464.69
336	2.44	22	17.5	83.77	70.89	161.24

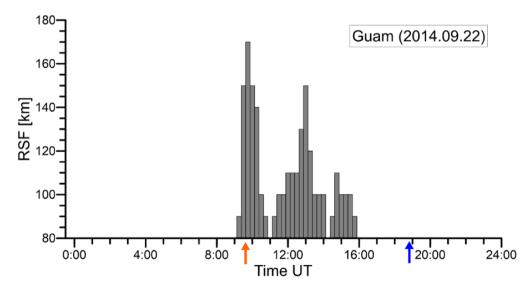


Figure 9. RSF time variation over Guam on September 22, 2014. Orange and blue arrows point to the sunset and sunrise respectively.

angle and slightly tilted in the norh-east sector, indicating that the ionospheric conditions are calm. However, the presence of an EPB causes multiple reflections of the signal in the cavity and bounces in such a way that it appears to come from multiple

directions. This can be observed in the skymap example in Figure 10 (right), where a clearly disturbed ionosphere caused by an EPB passing over the station can be identified. Therefore, DGS ionograms cover an observation volume in such a way that

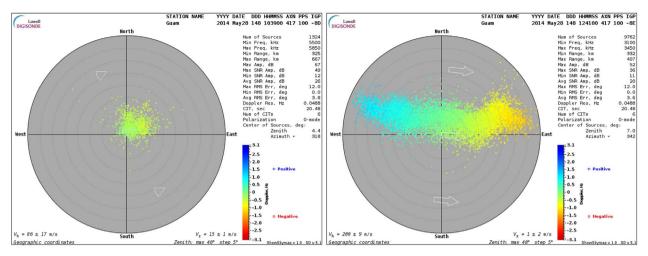


Figure 10. Examples of skymaps recorded by Guam's DGS on May 28, 2014, for quiet ionospheric conditions (left) and disturbed ionospheric conditions (right).

they form a beam with an angle of approximately 35° with respect to the vertical for vertical incidence sounding measurements. Conversely, GNSS receivers specifically cover the observation line corresponding to the IPP trace for a particular satellite. A GNSS receiver close to the location of the DGS will also be able to identify the EPB by analyzing the TEC (see Sect. 3.2). This co-location configuration is shown in Figure 11.

Considering the broader spatial observation of a DGS, it is conceivable that the EPB could be detected earlier by the DGS compared to a co-located GNSS receiver because the GNSS receiver detects the perturbation when it is aligned vertically, bearing in mind that the TEC is used instead of the STEC. Similarly, DGS will stop observing the EPBs later than the co-located GNSS sensor. In fact, experimental results comparing $T_i^{\rm DGS}$ and $T_i^{\rm GNSS}$ and comparing $T_f^{\rm DGS}$ and $T_i^{\rm GNSS}$ confirm this assumption. Figure 12 shows that most of detected EPBs are first observed by DGS than by a co-located GNSS receiver and that DGS finishes observing the EPBs later than a GNSS receiver.

Thus, we postulate that the initial detection time by the DGS, denoted as $T_i^{\rm DGS}$, precedes the initial detection time at the GNSS receiver, denoted as $T_i^{\rm GNSS}$. Taking into account that the EPB covers the distance Δr during the time delay between its initial detection by the DGS at $T_i^{\rm DGS}$ and its initial detection by the GNSS receiver at $T_i^{\rm GNSS}$, the EPB velocity can be readily estimated by:

$$v_{\rm EPB} = \frac{\Delta r}{\Delta t} = \frac{H_{\rm ion} \tan \alpha}{T_i^{\rm GNSS} - T_i^{\rm DGS}}.$$
 (4)

By assuming that the EPB is detected by the DGS at $T_f^{\rm DGS}$, the EPB size can be estimated by:

$$D = v_{\rm EPB} \Big(T_f^{\rm DGS} - T_i^{\rm DGS} \Big) - 2 H_{\rm ion} \tan \alpha. \eqno(5)$$

By considering an angle $\alpha = 35^{\circ}$, and $H_{\rm ion} = 350$ km, a total number of 115 EPB events have been identified across all the investigated sectors for 2014, which are simultaneously identified with the co-located DGSs and GNSS sensors (see

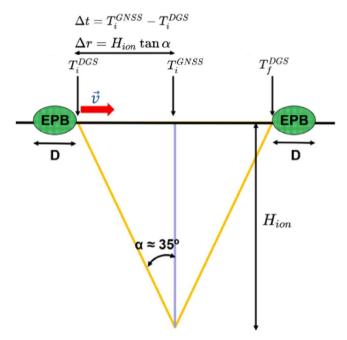


Figure 11. Schematic representation of bubble detection by quasi co-located DGS and a GNSS receiver. Both detectors are placed at the bottom vertex of the triangle. $T_i^{\rm DGS}$ and $T_f^{\rm DGS}$ correspond to the starting and final times at which the EPB is observed by the DGS, whereas $T_i^{\rm GNSS}$ is the starting time at which the EPB is detected by the GNSS receiver. The DGS spans an observation range of $\alpha \simeq 35^\circ$ with respect to the vertical direction, and the altitude of the ionospheric layer, $H_{\rm ion}$, is taken at 350 km.

Fig. 1). Appendix E provides detailed results of the 115 EPB events detected in the four sectors. As stated above, the initial and final times of the detection of an EPB by a DGS $\{T_i^{\text{DGS}}, T_f^{\text{DGS}}\}$ occur when the ionogram starts observing RSF greater than 80 km and stops observing RSF above 80 km. GNSS times $\{T_i^{\text{GNSS}}, T_f^{\text{GNSS}}\}$ were determined by using the new EPB detection method explained in Section 3.2. About the 75% of the detected events exhibited an earlier initial

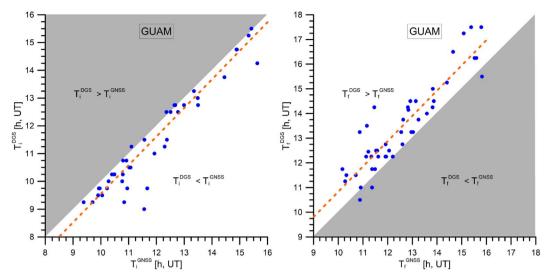


Figure 12. Scatter plots comparing T_i^{DGS} and T_i^{GNSS} (left) and comparing T_f^{DGS} and T_f^{GNSS} (right) of detected EPB events (blue dots) by the colocated DGS and GNSS sensors in Guam for 2014. The grey dashed area indicates the events when GNSS observes EPBs start earling than DGS (left) and when GNSS finishes observing EPBs later than DGS (right). Red-dashed lines correspond to linear regressions: at left for the initial times, $T_i^{\text{DGS}} = 1.03T_i^{\text{GNSS}} - 0.75$ with a determination coefficient $r^2 = 0.89$, and at right for final times, $T_f^{\text{DGS}} = 1.02T_f^{\text{GNSS}} + 0.59$ with $r^2 = 0.83$.

Table 5. Number of events, average time delays τ in initial time detection between DGSs and GNSS receivers, characteristic velocity $\overline{\nu}$, velocity ranges, and reference velocities $\overline{\mathbf{v}_{ref}}$ for the different sectors studied in this work (Sect. 4.2) and other studies in the literature.

Sector	Events	τ (min)	$\overline{v} \; (\text{ms}^{-1})$	$\overline{\mathbf{v}_{\mathrm{ref}}} \; (\mathrm{ms}^{-1})$
Caribbean	6	54	76	50-180
American	25	34	118	60-150
Atlantic	9	22	188	25-200
Pacific	30	26	155	50-200

detection time by the DGS compared to the GNSS receiver $(T_i^{\rm DGS} < T_i^{\rm GNSS})$. The fact that the above condition is not met for all the EPBs detected may be due to the lower sampling frequency of observation by DGSs compared to GNSS receivers. Furthermore, we excluded all the events whose final time detected by the DGS is greater than that detected by the GNSS receiver and those that did not fulfill the positive EPB dimension imposed by equation 5. This implies that 61% of the EPBs conform to the configuration illustrated in Figure 11. It should be noted again that GNSS sensors sample at a higher frequency (30 s) than DGS (5 min at best), see Table 1, introducing an inherent uncertainty to the detected times of the DGS compared to the GNSS. Moreover, we assumed starting and ending times of detection of EPBs by DGS according to a given RSF threshold of 80 km, which might not be accurate enough, posing additional uncertainties. In any case, considering the above detection results meet the indicated velocity constraints, we can calculate an average propagation velocity for each sector by combining the observations of co-located DGS and GNSS, as well as estimate the average size affected by the EPBs. The results of the EPBs velocity estimated by this simple model are reported in Table 5 for different sectors. These results compare well with others obtained by independent methodologies. Souza et al.

(2021) and Vargas et al. (2020) obtained drift velocities of the EPBs between 60 and 150 ms⁻¹ in the region of Brazil. Chapagain et al. (2012) found drift velocities of the EPBs in the Atlantic region between 25 and 170 ms⁻¹, and cite other works where velocities higher than 200 ms⁻¹ are determined in this Atlantic region. Ji et al. (2015) determined a drift velocity of the EPBs in the Pacific region between 50 and 200 ms⁻¹. These reference values are reported in the column $\overline{\mathbf{v}_{ref}}$ in Table 5. Although the characteristic velocity values \overline{v} found by our model are not exactly the same, the results agree reasonable well with those reported in the literature for the four sectors analyzed here. Further refinements could be incorporated into the model to achieve closer alignment with reported values, such as replacing the IPP with the receiver coordinates, integrating a time-varying reflection height, and determining a more precise angle of arrival of the signal to the DGS. Nevertheless, it is important to emphasize that this simple model serves as a qualitative validation of the obtained results for the Caribbean sector presented in Section 4.2, and shows a reasonable agreement with the values reported in the literature for the rest of sectors.

5 Conclusions

This works presents some technical enhancements to the EPB detection method introduced by Blanch et al. (2018). These refinements significantly bolster the method's robustness, leading to a reduction of more than half in the false positive rate compared to its predecessor. Moreover, the statistical techniques for obtaining a more precise fit for the background TEC have yielded an improved estimation of disturbance shape. This new version of the detection method of EPBs made available through the PITHIA e-Science Center⁶ and it complies with

⁶ https://esc.pithia.eu/data-collections/DataCollection_OE_EPB_detectionTool/api/

the FAIR principles required by the PITHIA-NRF project (Belehaki et al., 2023). Leveraging the output from this method, we have devised a method to estimate the velocity of EPBs using data from a dense network of GNSS receivers. The application of this method to the GNSS dataset for the year 2014 in a Caribbean region has unveiled the dynamics of 28 EPBs. The computed mean velocity stands at 100 ms⁻¹, with an average direction of propagation of 75° in azimuth angle. This azimuth direction suggests that EPBs propagate eastward and quasi-parallel to modip isolines in the American-Caribbean sector. The computed average size of the disturbances is about 300 km. These results of velocity (speed and direction of propagation) and size of the perturbation compare qualitatively well with the results of the drift velocity of EPBs obtained in the Indian sector by analyzing the scintillation measurements (Unnikrishnan et al., 2017). The later got values of EPBs propagating eastward, quasi-parallel to the modip isolines in that region, at $80-220 \text{ ms}^{-1}$ and with size of the perturbation along its drift of 50-400 km. However, our model results (see Table 4), indicate that the EPBs speed in the Caribbean region is 50-180 ms⁻¹, and the size of the perturbation along its drift ranges from 5 to 550 km.

Similar results have been obtained for the EPBs size and velocity of propagation along their drift by combining measurements of GNSS receivers with co-located DGS sensors, assuming the conceptual model described in Section 4.3. The dependence between the detection times of EPBs with the respective observation methods (DGS and GNSS) is a direct result of observing the same phenomenon at the same time. Thus, building up a system of near-real-time detection based on these methods independently would allow warning specific users concerned by radio propagation services, i.e., areas monitored by DGS detecting EPBs could warn about deleterious effects on GNSS services uncovered by GNSS receivers and vice-versa. Moreover, knowing the propagation speed of these disturbances, once detected, one could anticipate the arrival of EPBs at neighboring regions and warn about this phenomenon.

Future work can be addressed to design potential services and applications in line with the above discussion. Regarding the new EPB detection method, further refining can be focused on detailed pre-processing of the original observation RINEX files, such as utilizing 1-second sampling intervals or adapting the method for near real-time application. Additionally, more accurate methods for determining the velocity and dimension of EPBs than the one conceptually described in Section 4.3, where the observed reflection height is more accurately considered and the angle of signals received by the DGS is limited.

Acknowledgements

Authors wish to express their gratitude to the International GNSS Service, the Crustal Dynamics Data Information System Data Center, and the National Oceanic and Atmospheric Administration for making the data available. Global Ionospheric Radio Observatory (GIRO) (Reinisch & Galkin, (2011), and GIRO data providers (USAF NEXION Digisonde network, whose Program Manager is Annette Parsons, Jicamarca and Ramey) for making Digisonde data available. This research has been funded by EU Projects PITHIA-NRF (GA 101007599) and T-FORS (GA 101081835), in which V.N-P, D.A., A.S., and V.dP are participating. The editor thanks Claudio Cesaroni and an anonymous reviewer for their assistance in evaluating this paper.

References

- Abadi P, Saito S, Srigutomo W. 2014. Low-latitude scintillation occurrences around the equatorial anomaly crest over Indonesia. *Ann Geophys* **32(1)**: 7–17. https://doi.org/10.5194/angeo-32-7-2014.
- Adkins VJ, England SL. 2023. Automated detection and tracking of equatorial plasma bubbles utilizing global-scale observations of the limb and disk (GOLD) 135.6 nm data. *Earth Space Sci* **10**(10): e2023EA002935. https://doi.org/10.1029/2023EA002935.
- Belehaki A, Häggström I, Kiss T, Galkin I, Tjulin A, PITHIA-NRF Consortium. 2023. *PITHIA-NRF The integration project for an advanced plasmasphere, ionosphere and thermosphere research environment*. https://doi.org/10.5281/zenodo.10145053.
- Bhattacharyya A. 2022. Equatorial plasma bubbles: a review. *Atmosphere* **13(10)**: 1637. https://doi.org/10.3390/atmos13101637.
- Blanch E, Altadill D, Juan J, Camps A, Barbosa J, González-Casado G, Riba J, Sanz J, Vazquez G, Orús-Pérez R. 2018. Improved characterization and modeling of equatorial plasma depletions. *J Space Weather Space Clim* 8: A38. https://doi.org/10.1051/swsc/2018026.
- Booker HG, Wells HW. 1938. Scattering of radio waves by the F-region of the ionosphere. *Terr Mag Atmos Electr* **43(3)**: 249–256. https://doi.org/10.1029/TE043i003p00249.
- Calabia A, Imtiaz N, Altadill D, Yasyukevich Y, Segarra A, Prol FS, Adhikari B, del Peral L, Rodriguez Frias MD, Molina I. 2024. Uncovering the drivers of responsive ionospheric dynamics to severe space weather conditions: a coordinated multi-instrumental approach. *J Geophys Res Space Phys* 129(3): e2023JA031862. https://doi.org/10.1029/2023JA031862.
- Cesaroni C, Spogli L, Franceschi GD, Damaceno JG, Grzesiak M, Vani B, Monico JFG, Romano V, Alfonsi L, Cafaro M. 2021. A measure of ionospheric irregularities: zonal velocity and its implications for L-band scintillation at low-latitudes. *Earth Planet Phys* 5(5): 450–461. https://doi.org/10.26464/epp2021042.
- Chapagain NP, Makela JJ, Meriwether JW, Fisher DJ, Buriti RA, Medeiros AF. 2012. Comparison of nighttime zonal neutral winds and equatorial plasma bubble drift velocities over Brazil. *J Geophys Res Space Phys* **117**: A06309. https://doi.org/10.1029/2012JA017620.
- Chen J, Zhong J, Hao Y, Wan X, Li Q, et al. 2023. Horizontal spatial correlation of the ionospheric day-to-day variations at low latitudes based on GOLD Nmax data. *Space Weather* **21(12)**: e2023SW003627. https://doi.org/10.1029/2023SW003627.
- Christovam AL, Prol FS, Hernández-Pajares M, Camargo PO. 2023. Plasma bubble imaging by single-frequency GNSS measurements. *GPS Solut* **27**(3): 124. https://doi.org/10.1007/s10291-023-01463-z.
- Dungey J. 1956. Convective diffusion in the equatorial F region. *J Atmos Terr Phys* **9**(**5–6**): 304–310. https://doi.org/10.1016/0021-9169(56)90148-9.
- Feller W. 1948. On the Kolmogorov-Smirnov limit theorems for empirical distributions. *Ann Math Stat* **19(2)**: 177–189. https://doi.org/10.1214/aoms/1177730243.
- Groves KM, Basu S, Weber EJ, Smitham M, Kuenzler H, et al. 1997. Equatorial scintillation and systems support. *Radio Sci* **32**(5): 2047–2064. https://doi.org/10.1029/97RS00836.
- Haase JS, Dautermann T, Taylor MJ, Chapagain N, Calais E, Pautet D. 2011. Propagation of plasma bubbles observed in Brazil from GPS and airglow data. *Adv Space Res* **47(10)**: 1758–1776. https://doi.org/10.1016/j.asr.2010.09.025.
- Haaser RA, Earle GD, Heelis RA, Klenzing J, Stoneback R, Coley WR, Burrell AG. 2012. Characteristics of low-latitude ionospheric depletions and enhancements during solar minimum. *J Geophys Res Space Phys* 117, A10305. https://doi.org/10.1029/2012JA017814.

- Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, et al. 2020. Array programming with NumPy. *Nature* **585(7825)**: 357–362. https://doi.org/10.1038/s41586-020-2649-2.
- Hernández-Pajares M, Juan JM, Sanz J. 2006. Medium-scale traveling ionospheric disturbances affecting GPS measurements: spatial and temporal analysis. *J Geophys Res Space Phys* **111(A7)**. https://doi.org/10.1029/2005JA011474.
- Ji S, Chen W, Weng D, Wang Z. 2015. Characteristics of equatorial plasma bubble zonal drift velocity and tilt based on Hong Kong GPS CORS network: from 2001 to 2012. J Geophys Res Space Phys 120(8): 7021–7029. https://doi.org/10.1002/2015JA021493-T.
- Jordahl K, den Bossche JV, Fleischmann M, Wasserman J, McBride J, et al. 2020. Geopandas/geopandas: v0.8.1. https://doi.org/10.5281/ zenodo.3946761.
- Kapil C, Seemala GK. 2024. Machine learning approach for detection of plasma depletions from TEC. Adv Space Res 73(7): 3833–3844. https://doi.org/10.1016/j.asr.2023.04.042.
- Karan DK, Daniell RE, England SL, Martinis CR, Eastes RW, Burns AG, McClintock WE. 2020. First zonal drift velocity measurement of equatorial plasma bubbles (EPBs) from a geostationary orbit using GOLD data. *J Geophys Res Space Phys* 125(9): e2020JA028173. https://doi.org/10.1029/2020JA028173.
- Karan DK, Eastes RW, Daniell RE, Martinis CR, McClintock WE. 2023. GOLD Mission's observation about the geomagnetic storm effects on the nighttime equatorial ionization anomaly (EIA) and equatorial plasma bubbles (EPB) during a solar minimum equinox. Space Weather 21(3): e2022SW003321. https://doi.org/10.1029/ 2022SW003321.
- Kelley MC. 2009. The Earth's ionosphere: plasma physics and electrodynamics. Academic Press. ISBN 978-0-12-088425-4.
- Kumar S, Singh A. 2009. Variation of ionospheric total electron content in Indian low latitude region of the equatorial anomaly during May 2007–April 2008. Adv Space Res 43(10): 1555–1562. https://doi.org/10.1016/j.asr.2009.01.037.
- Lebyodkin MA, Shashkov IV, Lebedkina TA, Mathis K, Dobron P, Chmelik F. 2013. Role of superposition of dislocation avalanches in the statistics of acoustic emission during plastic deformation. *Phys Rev E* 88: 042402. https://doi.org/10.1103/PhysRevE. 88.042402.
- Li D, Zou Y. 2016. Automatic detection of ionospheric TEC depletions by using low-latitude GNSS data. In: 2016 11th International Symposium on Antennas, Propagation and EM Theory (ISAPE), IEEE, pp. 320–322. https://doi.org/10.1109/ISAPE.2016.7833952.
- Li Q, Zhu Y, Wang Z, Fang K. 2021. A method for automatic detection and characterization of plasma bubbles using GPS and BDS data. *Chin J Aeronaut* 34(5): 195–204. https://doi.org/ 10.1016/j.cja.2020.10.014.
- Lv M, Tang Q, Qiao J, Qiao W, Zhou C. 2024. Statistical analysis of ionospheric correlation for shortwave system. *Radio Sci* 59(4): e2023RS007893. https://doi.org/10.1029/2023RS007893.
- Lynn KJW, Otsuka Y, Shiokawa K. 2011. Simultaneous observations at Darwin of equatorial bubbles by ionosonde-based range/time displays and airglow imaging. *Geophys Res Lett* **38**: L23101. https://doi.org/10.1029/2011GL049856.
- Magdaleno S, Herraiz M, Radicella SM. 2012. Ionospheric bubble seeker: a Java application to detect and characterize ionospheric plasma depletion from GPS data. *IEEE Trans Geosci Remote Sens* **50(5)**: 1719–1727. https://doi.org/10.1109/TGRS.2011.2168965.
- Mannucci AJ, Wilson BD, Edwards CD. 1993. A new method for monitoring the Earth's ionospheric total electron content using the GPS global network. In: Proceedings of the 6th International Technical Meeting of the Satellite Division of the Institute of

- Navigation (ION GPS 1993), Salt Lake City, UT, September 22–24, Institute of Navigation, pp. 1323–1332.
- McNamara L, Caton R, Parris R, Pedersen T, Thompson D, Wiens K, Groves K. 2013. Signatures of equatorial plasma bubbles in VHF satellite scintillations and equatorial ionograms. *Radio Sci* **48(2)**: 89–101. https://doi.org/10.1002/rds.20025.
- Mersha MW, Lewi E, Jakowski N, Wilken V, Berdermann J, Kriegel M, Damtie B. 2020.A method for automatic detection of plasma depletions by using GNSS measurements. *Radio Sci* 55(3): e2019RS006978, https://doi.org/10.1029/2019RS006978.
- Moraes ADO, Vani BC, Costa E, Abdu MA, de Paula ER, Sousasantos J, Monico JFG, Forte B, de Siqueira Negreti PM, Shimabukuro MH. 2018.GPS availability and positioning issues when the signal paths are aligned with ionospheric plasma bubbles. *GPS Solut* 22(4): 95. https://doi.org/10.1007/s10291-018-0760-8.
- Muella MT, de Paula ER, Jonah OF. 2014. GPS L1-frequency observations of equatorial scintillations and irregularity zonal velocities. *Surv Geophys* **35**: 335–357. https://doi.org/10.1007/s10712-013-9252-0.
- Muella MT, de Paula ER, Monteiro AA. 2013. Ionospheric scintillation and dynamics of Fresnel-scale irregularities in the inner region of the equatorial ionization anomaly. *Surv Geophys* **34**: 233–251. https://doi.org/10.1007/s10712-012-9212-0.
- Navas-Portella V. 2020. Statistical modelling of avalanche observables: criticality and universality. *PhD Thesis*, Universitat de Barcelona. Available at http://hdl.handle.net/2445/173846.
- Nishioka M, Saito A, Tsugawa T. 2008. Occurrence characteristics of plasma bubble derived from global ground-based GPS receiver networks. J Geophys Res Space Phys 113(A5). https://doi.org/ 10.1029/2007ja012605.
- Oppenheim AV, Schafer RW, Buck JR. 2001. Discrete-time signal processing, 2nd edn. Prentice-Hall Signal Processing Series. Prentice Hall, Upper Saddle River, NJ. ISBN 0137549202.
- Patil AS, Nade DP, Taori A, Pawar RP, Pawar SM, Nikte SS, Pawar SD. 2023. A brief review of equatorial plasma bubbles. *Space Sci Rev* **219**(1): 16. https://doi.org/10.1007/s11214-023-00958-y.
- Piñal Moctezuma F, Delgado Prieto M, Romeral Martínez L. 2019. Performance analysis of acoustic emission hit detection methods using time features. *IEEE Access* **7**: 71119–71130. https://doi.org/10.1109/ACCESS.2019.2919224.
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 1996. Numerical recipes in Fortran 90 the art of parallel scientific computing. Cambridge University Press, Melbourne, Australia. ISBN 978-0-521-57439-6.
- Rawer K. 1963. Propagation of decameter waves (HF band). In: *Meteorological and astronomical influences on radio wave propagation*, Landmark B (Ed.), Pergamon Press, New York, pp. 221–250.
- Reddy SA, Forsyth C, Aruliah A, Smith A, Bortnik J, Aa E, Kataria DO, Lewis G. 2023. Predicting swarm equatorial plasma bubbles via machine learning and shapley values. *J Geophys Res Space Phys* **128(6)**: e2022JA031183. https://doi.org/10.1029/2022JA031183.
- Reinisch B, Huang X, Galkin I, Paznukhov V, Kozlov A. 2005. Recent advances in real-time analysis of ionograms and ionospheric drift measurements with digisondes. *J Atmos Sol-Terr Phys* **67(12)**: 1054–1062. https://doi.org/10.1016/j.jastp.2005.01.009.
- Reinisch BW, Galkin IA. 2011. Global ionospheric radio observatory (GIRO). *Earth Planet Space* **63**: 377–381. https://doi.org/10.5047/eps.2011.03.001.
- Sanz J, Hernandez-Pajares M, Zornoza JMJ. 2013. GNSS data processing: fundamentals and algorithms. ESA TM, European Space Agency. ISBN 9789292218867.

- Sarudin I, Hamid NSA, Abdullah M, Buhari SM. 2017. Investigation of zonal velocity of equatorial plasma bubbles (EPBs) by using GPS data. *J Phys Conf Ser* **852(1)**: 012014. https://doi.org/10.1088/1742-6596/852/1/012014.
- Seemala GK, Valladares CE. 2011. Statistics of total electron content depletions observed over the South American continent for the year 2008. *Radio Sci* 46: RS5019. https://doi.org/10.1029/ 2011RS004722.
- Silva RP, Souza JR, Sobral JHA, Denardini CM, Borba GL, Santos MAF. 2019. Ionospheric plasma bubble zonal drift derived from total electron content measurements. *Radio Sci* **54**(7): 580–589. https://doi.org/10.1029/2018RS006727.
- Smith SW. 2003. Digital signal processing: a practical guide for engineers and scientists. Demystifying technology series. Newnes Amsterdam, Amsterdam. ISBN 9780750674447.
- Souza ALC, Camargo PO, Muella MTAH, Tardelli-Coelho F. 2021. Drift velocity estimation of ionospheric bubbles using GNSS observations. *Radio Sci* 56(8): e2020RS007220. https://doi.org/10.1029/2020RS007220.
- The MathWorks Inc. 2022. MATLAB version: 9.13.0 (R2022b). Available at https://www.mathworks.com.
- Tsunoda RT, Livingston RC, McClure J, Hanson W. 1982. Equatorial plasma bubbles: vertically elongated wedges from the

- bottomside F layer. *J Geophys Res Space Phys* **87(A11)**: 9171–9180. https://doi.org/10.1029/JA087iA11p09171.
- Unnikrishnan K, Sreekumar H, Choudhary RK, Ashna VM, Ambili KM, Shreedevi PR, Rao PB. 2017. A study on the evolution of plasma bubbles using the single station-multisatellite and multistation-single satellite techniques. *J Geophys Res Space Phys* **122** (3): 3678–3688. https://doi.org/10.1002/2016JA023503.
- Van Rossum G, Drake FL. 2009. *Python 3 reference manual*. CreateSpace, Scotts Valley, CA.
- Vargas F, Brum C, Terra P, Gobbi D. 2020. Mean zonal drift velocities of plasma bubbles estimated from keograms of nightglow all-sky images from the Brazilian sector. *Atmosphere* 11(1): 69. https://doi.org/10.3390/atmos11010069.
- Wakai N, Ohyama H, Koizumi T. 1987. *Manual of ionogram scaling*. Radio Research Laboratory, Ministry of Posts and Telecommunications, Japan.
- Wu K, Xu J, Wang W, Sun L, Liu X, Yuan W. 2017. Interesting equatorial plasma bubbles observed by all-sky imagers in the equatorial region of China. *J Geophys Res Space Phys* 122(10): 10596–10611. https://doi.org/10.1002/2017JA024561.
- Yu T, Li M, Xia C, Zuo X, Liu Z, Zhao B. 2018. A new method for deriving equatorial plasma bubble velocity by tracing OI 630 nm all-sky images. *J Geophys Res Space Phys* 123(11): 9619–9633. https://doi.org/10.1029/2018JA025332.

Cite this article as: Navas-Portella V, Altadill D, Blanch E, Altadill M, Segarra A, et al. 2025. Estimation of the drift velocity of Equatorial Plasma Bubbles using GNSS and digisonde data. *J. Space Weather Space Clim.* 15, 2. https://doi.org/10.1051/swsc/2024038.

Appendix A

Automatic new EPB detection method algorithm

The new EPB detection method works automatically analyzing the time series of $\sigma(\text{TEC}''(t))$ and TEC(t). In accordance with various predefined conditions, the method identifies the EPBs and their primary observables through the execution of the following algorithm:

- 1. **EPB Initial and final time definition**: Starting from the beginning of the time series, a potential bubble is considered once $\sigma(\text{TEC}''(t))$ overcomes a threshold of 0.714 TECU (Blanch et al., 2018). This first-time T_i in which threshold is overcome determines the initial time of the potential bubble. The bubble finishes at T_f once σ (TEC''(t)) has undergone the threshold and remains below it for more than a hit definition time (HDT). This HDT has been set at 600 seconds in order to be in agreement with the value used for the calculation of $\sigma(\text{TEC}''(t))$, which ensures reasonable time scales to determine EPBs.
- Data quality checks: After temporally bounding the potential EPB, the subsequent requirements are enforced to ensure appropriate data quality:
 - (a) Minimal EPB duration $D = T_f T_i \ge 600 \text{ s.}$
 - (b) Availability of at least 50% of TEC(t) data points of the immediate 600 seconds before T_i . If this condition is not satisfied, it may lead to anomalous estimations of the TEC background.
 - (c) Availability of at least the 60% of TEC(t) points between the initial and final times.

- If all the conditions are met, proceed to step 3. Otherwise, this potential bubble is considered as non-significant. In this case, continue analyzing the time series by coming back to step 1.
- 3. Create a bubble candidate list based on the different estimations of the background TEC: Once T_i and T_f are defined, a reconstruction of the background TEC, $\text{TEC}_0(t)$, is performed to determine the main statistics (area and depth of the depletion) that will indicate whether the potential bubble is significant or not. Create the list of candidate fits that will consider those fits of the background TEC, $\text{TEC}_0(t)$, leading to significant EPBs.
- 4. Determine the set of points and its weights to perform the para**bolic fit**: The first candidate fit will consider the k = 2 immediate points of the TEC(t) before T_i and after T_f . Note that the total number of points might not be necessarily 2k due to the existence of data gaps. If the number of points k is lower or equal than a predefined threshold value k_c , proceed to step 5. Otherwise, go to step 7. Note that k_c is related the time considered before T_i and after T_f . In this work, it has been used $k_c = 10$, that corresponds to 5 min before and after the candidate EPB. However, this might not be true if there are undefined values in the surrounding background TEC. In this case of undefined values, only points that are at maximum 600 s before or after the candidate EPB will be considered. This is in agreement with the HDT and bubble duration in order to avoid EPB overlapping. If the number of points considered before the initial time $k_{t < T_i}$ is equal to the number of points after the final time $k_{t>T_{\ell}}$, all points in the fit will receive the same weight. However, if $k_{t < T_i} \neq k_{t > T_t}$, weights will be assigned to the points included in the fit to account for the difference in the number of points on either side.

6. Compute the main EPB observables and determine EPB significance for this candidate fit: The total area *A* of the perturbation can be computed as:

$$A = \int_{T_i}^{T_f} (\text{TEC}(t) - \text{TEC}_0(t)) dt = \int_{T_i}^{T_f} \Delta \text{TEC}(t) dt = A_> + A_<,$$
(A1)

where $\Delta TEC(t)$ is the disturbance curve, and $A_{>}$ and $A_{<}$ correspond to the positive and negative areas respectively. The bubble depth is defined as:

Depth =
$$|\min \Delta TEC(t)|$$
. (A2)

Once these values are computed, the following physics-based criteria (Blanch et al., 2018) will determine if the candidate can be considered as a significant EPB:

- (a) The positive area A_> must be lower than 40% of the negative area A_<. This will reduce the possibility to identify the fluctuating pattern of traveling ionospheric disturbance (TID) activity as EPBs
- (b) The depth must overcome a threshold of at least 5 TECUs.

If both conditions are met, the potential EPB is considered to be significant with respect to this candidate fit of the TEC background. Append this EPB observables to the candidate list. If these conditions are not met, increase the number of points to fit the background TEC $k \rightarrow k+1$ and go back to step 4.

7. If the candidate list is not empty, from all the significant candidates select the one with the minimum depth. This selection is performed to prevent parabolic fits that may produce artifacts due to minor TEC fluctuations in the vicinity of the EPB. If the candidate list is empty, then the EPB will be definitely considered to be non significant.

Appendix B

EPB drift velocity calculation algorithm

Let us expose the step-by-step methodology to compute the drift velocity of an EPB with a dense network of GNSS receivers:

- 1. Select a set of curves $\{\Delta TEC(t)_j\}_{sat}$ for a satellite (sat) in the list of available satellites. Recall that $j=1,\ldots,n$ where n is the number of GNSS receivers in the network. Each curve in this set has a number b_j of detected EPBs, each of them exhibiting non-zero values of $\Delta TEC(t)$ in the intervals $T_i^{(l)}, T_j^{(l)}$ ($l=1,\ldots,b_j$). If b_j equals zero, remove this GNSS receiver for the velocity analysis referred to this satellite.
- 2. Time-based EPBs clustering: To ascertain whether the same EPB is identified by various of the n receivers, a criterion is required to group the curves. However, criteria used for this clustering must be based on temporal information, given the absence of information about the spatial degrees of freedom of the EPB. According to the time-based clustering algorithm explained in Appendix C, clusters can be formed in order to state that an EPB has been observed by n_c ($n_c \le n$) different GNSS receivers. From the clustering algorithm, a number \mathcal{B} of clusters can be formed. Note that the letter \mathcal{B} refers to the EPBs that are detected by this satellite for a certain number of

- GNSS receivers. If no clusters can be formed, select another satellite and come back to step 1. If $\mathcal{B} > 0$, proceed to step 3.
- 3. Reconstruction of ΔTEC(t) via Fourier analysis is done in order to refine the time sampling from 30 s to 1 s. The original data sampling rate (30 s) is rather coarse and results in a rough estimate of the time delays used for the calculation. Pre-processing consists of Discrete Fourier Transform (DFT) interpolation and high-pass filtering. First, the raw data are DFT interpolated to increase the sampling rate (upsampling). This procedure consists of calculating the DFT spectrum of the raw data, and then performing the inverse DFT calculation using the obtained spectral amplitudes and specifying the sine and cosine spectral components with a finer time step (higher sampling rate) than the raw data. Note that in the preprocessing, the real DFT (Smith, 2003) calculation is used, although the algorithm is equivalent to spectral zero-padding interpolation made with complex DFT calculation (Oppenheim et al., 2001).
- 4. Select a cluster from \mathcal{B} that has not been already analized. If all the clusters have been analyzed, select another satellite and come back to step 1. Take a reference GNSS receiver from that cluster and compute the cross-correlation functions between this reference curve and the rest of GNSS receivers involved in the cluster. Determine the time delays Δt_j ($j=1,\ldots,n_c$) by finding the lag at which the cross-correlation attains its maximum value, CCM. Those receivers for which CCM² does not reach 0.75 will also be excluded.

After computing the maximum cross-correlation between curves, a subset of m_c ($m_c \leq n_c$) curves will have enough correlation to compute the velocity. If $m_c \geq 3$, go to the next step. If $m_c < 3$, velocity cannot be computed for that reference receiver. If this is the case, select another reference GNSS receiver and start this step again. The initial time of the EPB detected at the reference GNSS receiver will state a reference time $T_0^{\rm ref}$. At this reference time, each GNSS receiver has the record of the IPP coordinates and velocities of the satellite trace. Even if the maximum cross-correlation is good enough, if any of these records has undefined values for the reference receiver, EPB velocity cannot be computed. If this is the case, select another reference GNSS receiver and start this step again. If all the n_c receivers have been set as references, go to step 7.

5. For each EPB cluster and a reference receiver, consider only those receivers exhibiting a significant value of the maximum cross-correlation of 0.75. The time delay Δt_j with respect to the reference receiver can be written as:

$$\Delta t_j = \vec{\mathbf{s}} \cdot \Delta \vec{\mathbf{r}_j},\tag{B1}$$

for $j = 1, ..., m_c$, where $\vec{\mathbf{s}}$ is the slowness vector. $\Delta \vec{\mathbf{r_j}} = (\Delta x_j, \Delta y_j)$ is the relative position vector with N-E components of the *j*-th IPP with respect to the one from the reference receiver at T_{ref} . The EPB propagating at velocity $\mathbf{v_{\vec{ePB}}}$, which is an unknown variable, is related to the velocity by means of the following expression:

$$\mathbf{v}_{\mathbf{EPB}}^{\mathbf{J}} = \frac{\vec{\mathbf{s}}}{\mathbf{s}^2}.$$
 (B2)

If there are more than three GNSS receivers involved, this system of equations is over-determined and can be solved by least squares method. At this point, it must be noticed that data are weighted according to their squared maximum correlation coefficient CCM_i^2 . The azimuth z_{EPB} from true north can be easily obtained by considering the velocity components:

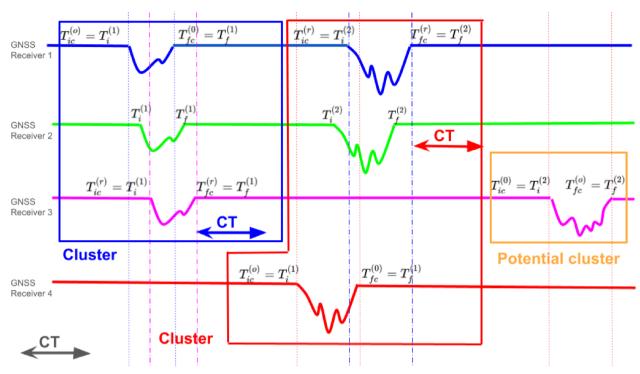


Figure C1. Schematic representation of the time-based clustering algorithm for four different GNSS receivers. Vertical dotted lines correspond to original cluster times $\{T_{ic}^{(o)}, T_{fc}^{(o)}\}$, whereas vertical dash-dot lines correspond to reference cluster times $\{T_{ic}^{(r)}, T_{fc}^{(r)}\}$.

$$z_{\text{EPB}} = \tan\left(\frac{v_x}{v_y}\right). \tag{B3}$$

The size D of the perturbation can be estimated as:

$$D = \left(|\mathbf{v}_{\mathbf{EPB}}| - \frac{\mathbf{v}_{\mathbf{EPB}} \cdot \mathbf{v}_{\mathbf{sat}}}{|\mathbf{v}_{\mathbf{EPB}}|} \right) \left(T_f^{\text{ref}} - T_i^{\text{ref}} \right), \quad (B4)$$

where $\overrightarrow{\mathbf{v_{sat}}}$ is the satellite velocity vector with N-E components, and T_i^{ref} and T_i^{ref} are the initial and final EPB times taken from the reference GNSS receiver. At this step, the velocity module, the azimuth and the distance of the EPB have been computed for a reference GNSS receiver.

- 6. Select another GNSS receiver in the cluster that has not been considered as a reference before. Come back to step 4 and repeat the next steps. If all the involved GNSS receivers have been considered as reference receivers, proceed to the following step.
- 7. For those reference receivers in which a result can be determined, the one exhibiting the highest average cross-correlation $\langle \text{CCM}^2 \rangle = \left[1/n_c \sum_{j=1}^{n_c} \text{CCM}_j^2 \right]$ is chosen. If there are more clusters to analyze, select another one and come back to step 3.

Appendix C

Time-based clustering algorithm

Figure C1 is merely an auxiliary sketch to illustrate the algorithm, and does not encompass all the diverse cases that the algorithm might encounter. Let us present the time-based clustering algorithm:

- 1. Let us consider the set of curves $\{\Delta TEC(t)_j\}_{sat}$ (j = 1, ..., n), being *n* the number of GNSS receivers, as well as the different EPBs in these curves sorted by initial times $\{T_i^{(l)}\}_{\text{sat}}$ $(l = 1, ..., b_j)$, being b_i the number of EPBs detected by the j-th receiver. By taking into account the first EPB detection time for all the GNSS receivers, $T_{ic}^{(o)} = \min\{T_i^{(l)}\}_{\text{sat}}$, and its corresponding final time $T_{fc}^{(o)}$. Note that the subscripts ic and fc refer to initial and final cluster time, respectively. These times will define the first EPB potential cluster times $\{T_{ic}^{(o)}, T_{fc}^{(o)}\}$, where the superscript (o) refers to original cluster times. The reference cluster times $\{T_{ic}^{(r)}, T_{fc}^{(r)}\}$, will be identical to the original cluster times. inal cluster times in this first iteration. These reference cluster times will be denoted by the superscript (r). Both original and reference times are associated to a specific GNSS receiver. It is important to differentiate between original cluster times (superscript (o)) and reference cluster times (superscript (r)) because the latter might change depending on the inclusion of EPBs in the cluster. The initial step of the algorithm in Figure C1 involves considering $\{T_{ic}^{(o)}, T_{fc}^{(o)}\}$ and $\{T_{ic}^{(r)}, T_{fc}^{(r)}\}$ corresponding to the first EPB detected by GNSS Receiver 1 (vertical blue dotted lines indicate original cluster times).
- 2. By taking into account the next EPB in the set sorted by ascending T_i and its corresponding initial and final times. This would correspond to the first EPB detected by the GNSS Receiver 2 in Figure C1 (green curve). If this initial time occurs within a predefined Clustering Time (CT) with respect to the $T_{ic}^{(o)}$, then this EPB will be considered to be in the same cluster as the first one. In this study, and in agreement with the time parameters used in the new detection method, CT has been set to 600 s. Hence, by construction, it is not possible to consider two EPBs detected at the same GNSS receiver in the same cluster. The inclusion of this EPB in the cluster will replace the reference cluster initial time $T_{ic}^{(r)}$, and will also replace the reference cluster final time as long as the final time of this EPB is larger than the value of the current $T_{fc}^{(r)}$. For the situation depicted

- in Figure C1, both reference cluster initial and final times $\{T_{ic}^{(r)}, T_{fc}^{(r)}\}$ will be replaced by $T_i^{(1)}$ and $T_f^{(1)}$ from the first EPB in Receiver 2. Note that this might not be always the case.
- 3. To include another EPB into the cluster, let us check the next EPB in the sorted list and, similarly to the previous step, check whether its initial time is within CT with respect to the reference cluster time $T_{ic}^{(r)}$, and 2CT with respect to the original reference cluster time $T_{ic}^{(p)}$. In Figure C1 this corresponds to the first EPB in the GNSS Receiver 3. Additionally, it has to be checked if its final time is in the interval \pm CT with respect to the final reference cluster time $T_{ic}^{(r)}$. If both conditions are met, then include this EPB in the cluster and redefine the cluster reference times according to the previous step.
- 4. The method proceeds iteratively by checking the remaining EPBs in the list and adding them to the cluster if the previous conditions are iteratively fulfilled. If the next EPB in the list does not fulfill these time conditions (as it occurs for the first EPB in GNSS Receiver 4) or there are not at least three EPBs in the cluster, then the cluster cannot be formed. Then, consider this not included EPB as the beginning of a new potential cluster and go to the first step of the procedure to check if the remaining EPBs in the list can be part of this new cluster.

Appendix D

Velocity results

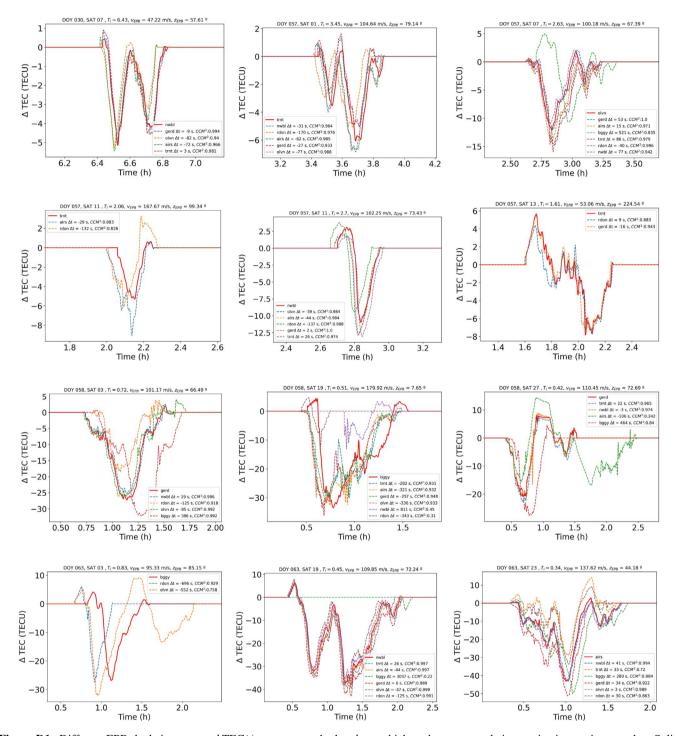


Figure D1. Different EPB depletion curves $\Delta TEC(t)$ to compute the lag Δt at which each cross-correlation attains its maximum value. Solid red lines correspond to the reference curve for each case. Panels from left to right follow the same order as the table in Table 4.

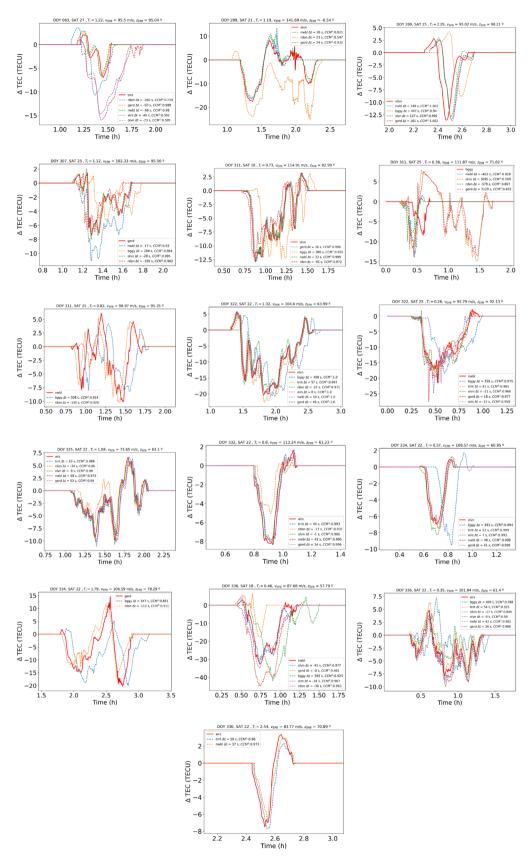


Figure D2. Different EPB depletion curves $\Delta TEC(t)$ to compute the lag Δt at which each cross-correlation attains its maximum value. Solid red lines correspond to the reference curve for each case. Panels from left to right follow the same order as the table in Table 4.

Appendix E

EPB events combining GNSS and DGS data

Table E1. Initial and final detection times of EPBs registered by co-located GNSS receivers and DGSs for the Pacific, Atlantic, American, and Caribbean sectors.

Sector	Year	DOY	$T_i^{\rm GNSS}$ (h)	$T_f^{ m GNSS}$ (h)	T_i^{DGS} (h)	$T_f^{\rm DGS}$ (h)
Pacific	2014	078	10.80	11.49	10.75	11.75
Pacific	2014	078	11.97	13.03	12.25	13.25
Pacific	2014	078	13.35	13.82	13.25	14.25
Pacific	2014	078	15.42	15.79	15.50	17.50
Pacific	2014	079	10.40	12.55	10.25	13.25
Pacific	2014	079	13.02	14.40	13.25	15.25
Pacific	2014	080	10.83	11.62	11.25	12.25
Pacific	2014	080	12.51	13.00	12.50	13.25
Pacific	2014	081	10.49	10.91	10.25	11.00
Pacific	2014	081	12.65	12.92	12.75	14.50
Pacific	2014	081	14.92	15.38	15.25	17.50
Pacific	2014	082	10.79	11.58	10.25	12.50
Pacific	2014	082	13.50	13.83	12.75	15.00
Pacific	2014	083	10.28	11.94	10.00	12.25
Pacific	2014	083	11.88	12.84	12.25	14.15
Pacific	2014	083	14.88	15.59	14.75	16.25
Pacific	2014	100	11.02	12.06	10.50	12.50
Pacific	2014	100	12.78	13.59	12.50	14.00
Pacific	2014	101	10.77	11.13	10.00	12.25
Pacific	2014	101	11.72	12.96	12.25	13.25
Pacific	2014	105	11.93	12.60	11.00	12.75
Pacific	2014	105	12.68	13.14	12.75	14.50
Pacific	2014	106	12.38	12.82	11.50	14.25
Pacific	2014	126	10.77	11.22	10.00	12.45
Pacific	2014	126	14.90	15.08	14.75	17.25
Pacific	2014	128	9.90	10.33	9.50	11.50
Pacific	2014	148	12.30	12.49	11.25	18.75
Pacific	2014	216	10.95	11.34	9.75	12.25
Pacific	2014	219	10.92	11.16	10.75	13.50
Pacific	2014	219	14.45	14.65	13.75	16.50
Pacific	2014	239	9.93	10.17	9.75	11.75
Pacific	2014	261	9.38	10.28	9.25	11.25
Pacific	2014	261	11.57	11.92	11.50	12.75
Pacific	2014	261	13.49	13.87	13.00	14.50
Pacific	2014	261	15.32	15.52	15.25	16.25
Pacific	2014	263	10.05	10.88	9.50	10.50
Pacific	2014	263	11.07	11.37	10.50	11.75
Pacific	2014	265	9.70	11.37	9.25	11.00
Pacific	2014	265	11.10	11.46	11.25	14.25
Pacific	2014	266	9.96	10.87	9.75	13.25
Pacific	2014	283	11.67	12.21	9.75	12.25
Pacific	2014	283	12.36	12.63	12.50	13.75
Pacific	2014	287	10.22	10.71	9.75	11.50
Pacific	2014	287	15.63	15.82	14.25	15.50
Pacific	2014	289	10.84	11.53	9.25	12.50
Pacific	2014	289	12.99	13.25	12.75	13.75
Pacific	2014	308	11.56	11.91	9.00	12.25
Atlantic	2002	054	22.29	22.89	22.00	23.25
Atlantic	2002	055	24.27	24.58	24.08	25.17
Atlantic	2002	055	21.41	21.88	21.42	21.92
Atlantic	2002	059	24.10	25.07	24.17	25.25
Atlantic	2002	059	21.68	22.58	21.75	22.25
Atlantic	2002	060	21.35	22.32	21.50	22.83
Atlantic	2002	060	22.54	23.37	23.00	24.08
Atlantic	2002	079	20.90	22.65	21.08	23.33

(Continued on next page)

Table E1. (Continued)

Sector	Year	DOY	$T_i^{\rm GNSS}$ (h)	$T_f^{ m GNSS}$ (h)	$T_i^{\rm DGS}$ (h)	$T_f^{\rm DGS}$ (h)
Atlantic	2002	080	22.51	24.44	22.53	24.50
Atlantic	2002	082	27.98	28.28	26.42	29.50
Atlantic	2002	105	21.92	22.47	21.50	22.00
Atlantic	2002	106	22.72	23.18	21.25	23.75
Atlantic	2002	124	23.22	23.67	23.00	23.75
Atlantic	2002	284	22.64	23.43	22.00	23.25
Atlantic	2002	285	22.62	22.87	23.75	24.75
Atlantic	2002	286	21.42	22.43	22.00	22.50
American	2014	009	0.83	1.57	0.75	1.75
American	2014	012	3.20	3.38	3.00	4.50
American	2014	015	1.31	1.94	0.75	2.25
American	2014	032	0.78	1.35	1.00	1.75
American	2014	032	1.42	2.08	1.75	3.75
American	2014	034	1.64	2.31	1.25	2.75
American	2014	035	2.85	3.18	2.75	4.50
American	2014	036	1.97	2.72	1.25	2.25
American	2014	054	1.77	2.23	1.50	2.25
American	2014	056	1.95	3.21	1.00	2.75
American	2014	056	2.87	3.54	2.75	4.50
American	2014	057	2.99	3.75	1.50	3.50
American	2014	057	4.38	5.25	4.25	4.50
American	2014	058	4.13	4.89	4.00	5.25
American	2014	100	2.78	3.42	2.25	3.50
American	2014	100	3.26	3.69	3.50	5.00
American	2014	101	2.53	3.23	0.75	2.50
American	2014	102	3.11	3.46	1.75	3.25
American	2014	261	2.42 3.94	2.84 4.14	2.00 3.25	3.50 5.25
American	2014	262	3.94 2.81		3.23 1.00	3.23
American	2014	263	3.29	3.04 3.48	0.50	3.50
American American	2014 2014	264 265	0.75	1.18	0.75	1.50
American	2014	265	1.33	1.71	1.50	2.75
American	2014	267	1.79	2.27	1.00	3.75
American	2014	283	1.43	2.10	1.00	2.00
American	2014	283	3.42	3.67	2.00	3.50
American	2014	283	6.25	6.46	6.00	7.75
American	2014	289	2.88	3.07	1.00	4.50
American	2014	307	1.32	2.24	0.75	3.75
American	2014	308	2.47	2.73	1.25	2.75
American	2014	310	0.72	1.52	1.00	1.75
American	2014	310	3.18	4.16	2.00	5.25
American	2014	312	1.06	1.54	0.25	3.00
American	2014	329	0.24	1.82	0.50	4.00
American	2014	330	0.28	2.38	0.50	7.00
American	2014	331	0.33	1.53	0.25	7.25
American	2014	332	1.77	2.45	0.75	5.50
American	2014	333	1.97	2.52	0.75	6.25
American	2014	334	3.01	3.98	0.75	7.00
American	2014	335	2.18	2.37	0.75	7.25
American	2014	352	0.72	1.14	0.75	8.00
American	2014	353	1.38	1.66	1.00	8.00
American	2014	354	1.67	2.19	1.00	8.00
American	2014	355	0.77	2.04	0.50	8.00
American	2014	357	1.27	2.01	1.25	8.25
Caribbean	2014	057	1.61	2.26	1.75	2.75
Caribbean	2014	063	0.83	1.61	0.75	1.75
Caribbean	2014	289	1.19	2.33	1.00	2.00
Caribbean	2014	322	1.32	2.58	1.00	2.00
Caribbean	2014	334	1.79	2.93	1.50	2.25
Caribbean	2014	336	2.44	2.73	2.00	0.40