ADVERTIMENT. La consulta d’aquesta tesi queda condicionada a l’acceptació de les següents condicions d'ús: La difusió d’aquesta tesi per mitjà del servei TDX (www.tesisenxarxa.net) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d’investigació i docència. No s’autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d’un lloc aliè al servei TDX. No s’autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tesisenred.net) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you’re accepting the following use conditions: Spreading this thesis by the TDX (www.tesisenxarxa.net) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it’s obliged to indicate the name of the author

Ph.D. Thesis

Doctoral Program in Aerospace Science & Technology

Contributions to ionospheric determination with Global Positioning System: solar flare detection and prediction of global maps of Total Electron Content

Alberto García-Rigo[1]

Advisors:
Prof. Manuel Hernández Pajares[1]
Prof. Enric Monte Moreno[2]

[1] Research group of Astronomy and Geomatics (gAGE)
Dept. of Applied Mathematics IV
Universitat Politècnica de Catalunya (UPC), Spain

[2] Dept. of Signal Theory and Communications
Universitat Politècnica de Catalunya (UPC), Spain

April 12, 2012

Acknowledgements

First of all, I want to express my gratitude to my advisors, Prof. Manuel Hernández Pajares and Prof. Enric Monte Moreno. They have given me full support and have spent hours and hours of discussion. I admire their wisdom, experience and enthusiasm. It was also nice to join them in some of their cafés ilustrados!

I sincerely want to thank Prof. Andrzej Krankowski and Dr. Pawel Wielgosz for their support and close collaboration during the last years. Thanks as well for their hospitality during my visit to the University of Warmia and Mazury in Olsztyn, Poland. Their encouragement to carry on to complete this dissertation was really important!

The author is also grateful to IGS for providing ionosphere data products, to BKG for NTRIP real time datastreams, to NASA/CNES for JASON data, to NOAA's Space Weather Prediction Center for GOES data and to USC Space Sciences Center for SOHO/SEM EUV data. Apart from that, this thesis work has partially been funded by the Spanish Ministry of Science, Technology and Innovation and the European Social Fund under the ESP2004-05682-C02-01 and the CTM2010-21312-C03-02 projects.

I am truly grateful to the whole gAGE group: Manuel, Miguel, Jaume, Adrià, Àngel, Angela, Dago, Pedro, Pere and Raül, with whom I have shared not only work but also very pleasant times, talks, laughter during these fantastic years. They have helped me immensely to get to this point! Formeu part dels meus millors records d'aquests anys! I que perduri l'amistad en el temps! Mil gràcies!

Thanks as well to all my friends outside gAGE! Thanks for joining me in the adventure of life. Gràcies a tots i totes! Una forta abraçada per l'Aleix, els Martins i el Zeus, la Noe, el Marco, la Navete i la Nora, el Carles, l'Alicia, el Martí i el Giuseppe, al Ferru i al David, a la Núria, al Bernard, al Guillem i al Joan, a la Milla, al Jordi, a la Teresa i al Nil, al Bern i al Llopis, a l'Arantxa, al Rogent, a la Isa, al Raúl, a la Patricia i al Jose, a Jesús, a Estela y a Isabel, a l'Andreu, a la Ona i a l'Aram, a l'Ausiàs i a la Maria, a l'Anna, al Jordi i a la Laia.

Bardzo dziękuję moi przyjaciele! Ania, Basia, Carolina, Michael and Miłosz. Tak daleko, tak blisko.

Todas las gracias a mi familia (que ha ido creciendo con el tiempo). A mis padres, Jesús y Magdalena, a mis hermanos Iñaki y Marta (y sus respectivas parejas, Laia y Nello), a mis abuelos Jaume y María, Encarna y Ramón, a mis tíos José Ramón y Malen, a mis primas Rebeca y Úrsula, a Gerardo, a Úrsula y a Gerardo Ramón, a Carlos, a Malena y a Carlos, a Andreu, Marisa y Andreu, a Jeronia y Biel, a Pilar, Roy y Becky (thanks for your time to correct this thesis!), Owen y Rodrigo, y un largo etcétera. Més recentment, a Rafael i Tilín, al Carlos, la Lali i a l'Èlia, a tota la família del Rafael i tota la de la Tilín. Clara, Júlia, els vostres somriures han estat el millor motor durant aquests anys. Som una petita però grandíssima família!! I ara tindrem més temps per viure junts noves i fantàstiques aventures! Todos vosotros habéis sido mis mejores acompañantes y me habéis dado el mejor apoyo desde pequeñito hasta ahora! Y siempre habéis tenido una paciencia infinita, no solo durante esta tesis!! Tots vosaltres heu estat els meus millors acompanyants i m'heu donat el millor recolçament des de ben petit fins ara! I sempre heu tingut una paciència infinita, no només durant aquesta tesi!!

Contents

List of Figures

List of Tables

Introduction

The work in this thesis is focused on two research topics. The first is the detection of solar flares in real-time by monitoring ionospheric Total Electron Content (TEC) in the recent context of the MONITOR project, funded by the European Space Agency (ESA). The second is the prediction of Global Ionospheric TEC Maps (GIMs) in the frame of the support of the International Global Navigation Satellite Systems (GNSS) Service (IGS) Ionosphere Working Group (IGS Iono-WG) to SMOS ESA mission. These topics are in line with the characterisation and monitorisation of ionospheric spatial and temporal processes. Both research studies are based on processing Global Positioning System (GPS) data, which has proven its suitability for ionospheric sounding in the past.

Regarding the detection of solar flares, their related sudden increase of radiation can produce significant ionospheric perturbations. If the solar flare is facing the Earth and is powerful enough, these perturbations will affect the whole sunlit area of the ionosphere. In this regard, a world-wide network of GNSS dual-frequency receivers can be used to detect such perturbations. This is of great interest since solar flares can be a precursor of the arrival of an increased number of charged particles, especially in the case of a Coronal Mass Ejection (CME). These charged particles can produce outstanding problems in today's technology. A side contribution of this thesis is an early warning of the increase of charged particles, due to the time lags between photons and charged particles. Also, we would remark that the detection of solar flares and the generation of early warnings are based on processing the available GNSS observations.

Regarding the prediction of TEC, we have implemented an approach based on considering the state of the ionosphere at a given moment as an image, where each image corresponds to a UPC GIM with a temporal resolution of two hours. The image is represented by its Discrete Cosine Transform (DCT) coefficients and the predicted TEC is constructed using the forecasted coefficients by means of linear regression. Predicted TEC values can be of great interest for the scientific community. A recent example is its applicability as a background model in the generation of real-time GIMs for the IGS Iono-WG.

It is important to mention that the research on detection of solar flares has enabled the publication of two articles in peer-reviewed journals ([*García-Rigo et al.* (2008a)] and [*Hernández-Pajares et al.* (2011)]). Additionally, a recent article on this topic was under review in another peer-reviewed journal at the moment this dissertation was issued ([*Hernández-Pajares et al.* (2012)]). As mentioned before, this has enabled the author to actively participate in the MONITOR project, funded by the European Space Research and Technology Centre (ESTEC) of the European Space Agency (ESA).

The author of this thesis has given support to gAGE/UPC on tasks related to the IGS Iono-WG. His responsibilities have been related to the generation and distribution of UPC products and IGS combined products since 2005. In this context, the author contributed to a peer-reviewed article (see [*Hernández-Pajares et al.* (2009)]). In addition, the author has developed the UPC Predicted product (predicted Global Ionospheric Maps for two days ahead) from scratch. The work carried out in TEC prediction has led to the publication of another article in a peer-reviewed journal ([*García-Rigo et al.* (2011)]).

The UPC Predicted product, which is made available through CDDIS FTP site, is generated automatically on a daily basis and combined with CODE (Center for Orbit Determination in Europe, Astronomical Institute, University of Berne, Switzerland) and ESOC (European Space Operations Center of ESA, Darmstadt, Germany) Predicted products to generate the combined IGS Predicted product. Furthermore, the UPC Predicted product has recently been used in the generation of the UPC real-time ionospheric products ([*Orús et al.* (2010a)]).

Also, the obtained results have been presented in a few Congress Proceedings ([*García-Rigo et al.* (2010a)] and [*García-Rigo et al.* (2007c)]) as well as in multiple oral and poster presentations ([*García-Rigo et al.* (2010b)], [*García-Rigo et al.* (2009a)], [*García-Rigo et al.* (2008b)], [*García-Rigo et al.* (2007b)] and [*García-Rigo et al.* (2007d)], among others). These presentations were given at several international scientific congresses such as the European Geosciences Union (EGU) General Assembly, the Geodesy for Planet Earth (IAG) and the Beacon Satellite Symposium (BSS), among others.

It is also worth mentioning that the developed software is designed to run in a Linux Operating System (OS) and is mainly implemented in C-shell script and gawk. In addition, certain specific parts are also written in Fortran, C and Matlab programming languages.

Thesis breakdown

The two research topics in this thesis work are covered in two independent chapters:

- Chapter 1, *GNSS detection and monitoring of solar flares.* This covers the explanation of the method that has been developed for the detection of solar flares from GNSS signals.

- Chapter 2, *GNSS Prediction of ionospheric Total Electron Content.* This covers the explanation of the method that has been developed for global TEC forecasting.

Afterwards, general conclusions are provided as well as two appendices. The first of these appendices covers the GPS fundamentals and how ionosphere information can be derived from GNSS observables. In this regard, I want to express my gratitude to Dr. Angela Aragon-Angel, also member of the group of Astronomy and GEomatics at UPC, for sharing associated information ([*Aragon-Angel* (2010)]). The second appendix is named *SISTED outputs* since it specifies the format of the outputs of the solar flare detector (SISTED for Sunlit Ionosphere Sudden TEC Enhancement Detector), including the three output messages and the multiple plots being generated.

Chapter 1

GNSS detection and monitoring of solar flares

1.1 Introduction

A solar flare can be one of the most violent events that take place on the Sun's surface. They are generated near sunspots and are characterised by the emission of radiation in all electromagnetic ranges. In addition, they are typically accompanied by the ejection of charged particles. The radiation produced by a solar flare is across the whole spectrum, but especially important in the ultraviolet UV and X-ray bands, and reaches the Earth in about 8 minutes. In contrast, the ejected particles (a mix of electrons, protons and heavy nuclei) typically take one/two days to reach the Earth, following the Interplanetary Magnetic Field (IMF) lines. In this way, the detection of solar flare radiation can be used to generate warning messages prior to an arrival of charged particles. This is of great interest since preventive actions could be taken against potentially harmful consequences.

In more detail, the solar flare radiation produces a sudden increase of ionization in the sunlit atmosphere and thus, of the ionospheric Total Electron Content (TEC; [*Mendillo* (1974)]). This over-ionization occurs before the possible arrival of particles and is dependent of the Solar-Zenith Angle (SZA; [*Tsurutani et al.* (2009)]). As the ionization process is mainly given in the ionosphere, a way to detect that a solar flare facing the Earth has happened is by monitoring the ionospheric TEC variations (looking for any positive curvature that is simultaneously given in the whole or most part of the sunlit ionosphere). And this can be done in a simple, precise and economical way by using the Global Positioning System (GPS) system and GPS data from a world-wide network of GPS permanent receivers, such as the International Global Navigation Satellite Systems (GNSS) Service (IGS) network or the IGS-IP (IP for Internet Protocol) network in real time.

The increase of TEC is maximised near the subsolar point but noticeable in most of or the whole sunlit ionosphere. Note that even dawn and dusk regions will be affected by the TEC enhancement as they are still irradiated by the Sun (see [*Zhang et al.* (2004)]). On the contrary, no variations due to solar flares are expected in the night-side ionosphere region.

The occurrence of solar flares is highly correlated with solar cycle activity. This correlation results in an increase of solar flares during maximum solar activity periods. During low solar activity periods, the number of solar flares decreases abruptly and only very low powerful flares, or subflares, are triggered.

1.2 Motivation

In this work we have designed a system to detect solar flares. The motivation for this system is due to the economic and social impact of the solar flares and the associated particles ejection. Moreover, the use of GNSS leads to a robust system which is easy to implement.

Some of the main problems related to solar flares and the associated particles ejection can be summarised as follows (see, for instance, [*Brekke* (2004)] and [*Kappenman* (2012)]):

- Serious risks for astronauts in case of space walks (especially dramatic outside of the Earth's atmosphere). Also, hazards for aircraft passengers travelling through the poles.

- Mobile telephone grid damages. Note that they are synchronised with the GNSS time scale.

- Satellite damages including solar panels degradation (silicon-based solar cells) and computer and memory upsets and failures.

- Increase of satellites air-drag since the Earth's atmosphere expands. This requires increase operational efforts, e.g. by the necessity to prepare and perform additonal orbit correction manoeuvres (including increased fuel usage on the spacecraft and thus lifetime reduction), especially in the case of Low-Earth Orbit (LEO) satellites. For instance, it caused an early reentry of SkyLab in 1979.

- Important interferences in communications and power grid outages (such as the massive power outage at Quebec in 1989[1]).

---

[1] The corresponding magnetic storm shut down Hydro-Quebec power grids, the corporation that gives service to Canada's Quebec province. The service was restored after nine hours for 83 percent of the system. The estimated cost of the incident was 6 billion dollars ([*OECD* (2011)]).

- Aurorae[2] can probably be triggered due to the increase of charged particles that can be released after a solar flare.

In addition to such problems, in the case of GNSS users, the increase of radiation caused by a solar flare facing the Earth can be responsible for the interruption of GNSS services when a significant component in L-band appears, since the receivers may lose the tracking of the GNSS satellites (this was especially important in the case of the flare triggered on 6th December, 2006[3]). The possibility of new events with these characteristics is a major concern as it might affect the reliance of GNSS systems (continuity, availability, accuracy and integrity).

The GNSS positioning accuracy can also be threatened by powerful flares. In fact, the related increase of radiation can produce abrupt variations of more than 20 TEC units (TECUs) in the Total Electron Content (this was the case of the solar flare triggered on 28th October, 2003). Therefore, the ionospheric contribution to the pseudorange and phase GNSS observables would be greater. That would especially affect single-frequency GNSS users.

Finally, particles ejected from the Sun may be accelerated to near-relativistic velocities in certain cases. This is of great importance since the time gap between the detection of the increase of radiation can be very close to the arrival of particles. Thus, there would be little time for warning and prevention (this occurred in the cases of the solar flares triggered on 28th October, 2003 and 20th January, 2005; see, for instance, [*Bieber et al.* (2005)] and [*Masson et al.* (2009)]).

1.3 State-of-the-Art of solar flare detection

Recently, several satellite-based missions have been equipped with instruments to observe and study the solar characteristics. In particular, the effects of solar flares facing the Earth. This is the case of the Geostationary Operational Environment Satellites (GOES), the SOlar and Heliospheric Observatory (SOHO), the Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) or the more recent Solar Dynamics Observatory (SDO). Thanks

---

[2] Aurorae are dazzling lights that are mainly visible at high latitudes. Photons are emitted by atoms, molecules and ions after they are excited by charged solar particles, which enter the Earth through the poles. This happens in the case of magnetic reconnection between the Interplanetary Magnetic Field (IMF) lines and Earth's magnetic field lines.

[3] This event caused the loss of carrier lock in GPS receivers across the entire sunlit hemisphere. In addition, the U.S. Wide Area Augmentation System (WAAS) stopped providing corrections for 15 minutes ([*Last* (2010)]).

to these missions, images and records on solar flare occurrences are provided online for public access.

Furthermore, the GNSS are also useful to detect solar flares in an indirect way. GNSS data can be processed to determine whether an increase of ionization has happened in the Earth's ionosphere as a consequence of a powerful solar flare. This is possible by monitoring the variations of the TEC in the sunlit ionosphere region. For this purpose, we can use the large amount of data of GPS dual-frequency GNSS receivers distributed world-wide to study the ionospheric behaviour with a high spatial and temporal resolution.

There are many scientific papers studying the ionospheric response to particular solar flares using GNSS data, especially the one that occurred on 14th July, 2000 or the one which occurred on 28th October, 2003 (to cite some of them: [*Zhang et al.* (2002)], [*Woods et al.* (2004)] and [*Tsurutani et al.* (2005)]; see also Section 1.5.4). In addition, the concept of using GNSS to obtain the global ionospheric response to solar flares was firstly addressed by [*Afraimovich et al.* (2000)] by developing the GLOBal DETector (GLOBDET).

The goal of this research is to provide a simple and efficient way of detecting as many powerful solar flares as possible in post-processing (looking for solar flares backwards) and in real-time. In this way, solar flare warning messages can be broadcast in order to give early warning of the harmful consequences explained in the previous Section 1.2. The technique for detection that has been developed within this research work is called SISTED for Sunlit Ionosphere Sudden TEC Enhancement Detector, which is part of the so-called solar flare monitoring method (see Section 1.4). In brief, the system is fed with GPS data obtained from a selected set of dual-frequency receivers and it is based on detecting sudden ionospheric TEC increases that must happen simultaneously in most of the sunlit ionosphere region. This can be done by monitoring the GPS ionospheric combination L_I derived from phase observables.

The main benefit of the SISTED technique as compared with other detection techniques is that it does not require data from external providers outside the GNSS community. In addition, it can run in real-time operation and could provide value added data for GNSS users in the near future.

The solar flare monitoring method and SISTED can also be a good starting point to better understand how the ionosphere behaves in the occurrence of solar flares (from the flare onset until the end of the recombination phase).

In the following sections, a detailed review of the approach to detect solar flares is given. In this regard, Section 1.4 will address the description of the technique, Section 1.5 on performance analysis will include evaluation criteria to validate the SISTED detections and the results that have been obtained for more than a half solar cycle. Finally, Section 1.6 will present the conclusions.

1.4 The solar flares monitoring method

Methods to monitor the high-frequency variations of the ionospheric TEC have been developed based on processing dual-frequency GPS data, such as in the case of [*Hernández-Pajares et al.* (1997)] (using sidereal day differences) or [*Afraimovich et al.* (2000)]. As mentioned before, the ionosphere is a dispersive medium and so, it affects the L_1 and L_2 GPS signals differently. Taking advantage of this fact, the GPS geometry-free combination L_I is constructed (see Appendix A). L_I can be used to estimate the ionospheric TEC that is affecting the GPS rays between each receiver-satellite pair.

In this research, variations of the GPS geometry-free combination have been obtained by applying a second finite difference in time (also named Laplacian; referred to as L_I drift rate or $\Delta^2 L_I$; see [*Hildebrand et al.* (1987)]). The Laplacian is used to be more sensitive to any abrupt change of TEC. Consequently, the second finite difference in time of the ionospheric vertical Total Electron Content (referred to as VTEC drift rate, or $\Delta^2 V$) could also be derived.

The ionospheric monitoring method is suitable for the detection of ionospheric variations, both fast and slow at local and global scale. For instance, it can be used to identify the perturbations induced by Travelling Ionospheric Disturbances (TIDs), scintillation, solar flares, eclipses and earthquakes, among other examples.

In the following sections, the different steps that are considered to detect solar flares facing the Earth are explained. These include: gathering data; processing them to get information on ionospheric high frequency variability; classifying them according to the SZA; and, finally, applying SISTED to decide whether a solar flare has most probably occurred or not.

Note that there are three modes of operation of the SISTED detector: the post-processing, the high-rate and the real-time. In the post-processing mode of operation, 30 seconds data are used with latencies of about one day. In the high-rate mode of operation, 1 Hz data are used with latencies of about two hours. In the real-time mode of operation, 1 Hz data from NTRIP (for Networked Transport of RTCM via Internet Protocol) real-time datastreams are gathered (see Section 1.4.2 for additional details) with latencies of a few seconds.

1.4.1 General overview

As a first overview, Figure 1.1 contains a box diagram including the main steps carried out within the solar flares monitoring approach, from gathering input data to sending user warnings in case of detection. It must be remarked that the data flow is almost equivalent for the post-processing, high-rate and real-time modes of operation.

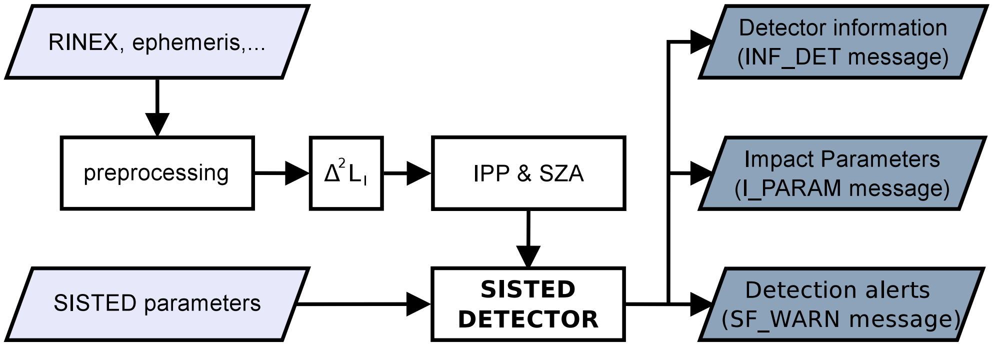


Figure 1.1: *Diagram showing the main steps of the solar flares monitoring method. SISTED stands for Sunlit Ionosphere Sudden TEC Enhancement Detector, IPP for Ionospheric Pierce Point, SZA for Solar-Zenith Angle and* $\Delta^2 L_I$ *corresponds to the second finite difference in time of the GPS Ionospheric carrier phase combination* L_I.

1.4.2 Input data

For the post-processing mode of operation, input data gathered from the IGS network of dual-frequency GNSS receivers are used. In order to be sensitive to any solar flare facing the Earth at any time of the day, data at 30 seconds sampling rate period are used from about 120-140 GPS receivers distributed world-wide (an example is depicted in Figure 1.2). One receiver for each cell of 10^o in longitude range and 10^o in latitude range is selected, if any receiver is available. In this way, the method achieves a high spatial and temporal resolution without undue computational burden. These data are publicly available and can be gathered from GARNER FTP site or that of the Crustal Dynamics Data Information System (CDDIS; at ftp://garner.ucsd.edu and ftp://cddis.gsfc.nasa.gov/, respectively). Note that there is scarcity of GNSS receivers in equatorial latitudes, in particular. As a consequence, a low number of GNSS measurements will be retrieved near the subsolar point at any epoch[4].

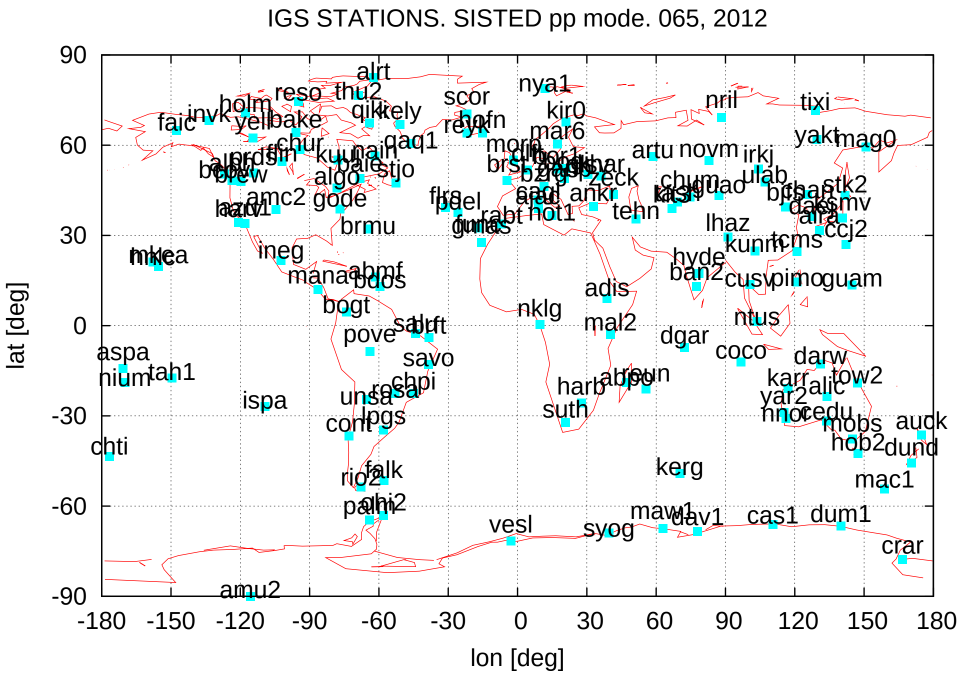


Figure 1.2: *Example of distribution of the International GNSS Service (IGS) receivers being used as input data for the solar flare monitoring method in post-processing mode, for the* 5^{th} *May of year 2012.*

For the high-rate mode of operation, 1 Hz IGS data are gathered from all the available stations at CDDIS FTP site (through ftp://cddis.gsfc.nasa.gov/gps/data/highrate). For instance, when looking at the first 15 minutes RINEX file for each station and day, the maximum number of simultaneous IGS high-rate stations was 56 in 2003, and 136 in 2011). This mode of operation is only used in specific cases since it is very demanding in terms of CPU usage and storage requirements.

For the real-time mode, the method is fed with dual-frequency GPS data streams that are distributed via NTRIP broadcasters (also named casters), such as the ones hosted by IGS (at http://igs-ip.net) and the European Reference Frame (EUREF; at http://euref-ip.net). At the moment, about 90 receivers are monitored globally though not regularly over the entire Earth due to the scarcity of data. Note that data are obtained at 1 Hz sampling rate but a down-sampling is applied to work with 30 seconds values since this is enough for the detection of powerful solar flares.

---

[4] In this work, epoch is taken as a synonym of the given time.

In this research work, algorithms to select and retrieve raw data streams from multiple NTRIP casters have been implemented. The streams are gathered automaticaly and are reformatted to an internal format. For this purpose, public software tools distributed by the Bundesamt für Kartographie und Geodäsie (BKG; available at http://igs.bkg.bund.de/ntrip/index) have been used as well as the teqc software to read binary GPS RAW data, which is distributed by UNAVCO (available at http://facility.unavco.org).

1.4.3 Data pre-processing

Following any of the above-mentioned data acquisition modes, a pre-processing step is applied. It includes several sub-steps such as the ones listed below (for further details, see [*Valls-Moreno* (2008)]).

- Determination of the receiver and satellite positions at transmitting time.
- Detection of cycle-slips (in this way, the continuous carrier phase arc number is determined).
- Correction of potential leap milliseconds.
- Correct antenna phase centre offsets.
- Calculation of the ionospheric code and carrier-phase combination, P_I and L_I respectively.
- Reformatting data to an internal line format for easier processing.

1.4.4 L_I processing

In the following step, the second finite difference, or Laplacian, of the GPS ionospheric carrier-phase combination L_I (denoted as $\Delta^2 L_I$) is calculated for each given receiver-satellite pair and for each epoch (see Equation 1.1). The Laplacian operator can be considered as a discrete-time analogous to a second derivative. Therefore, the offset and the linear trend of the time series are cancelled. In addition, as a linear filter, the Laplacian operator attenuates low frequencies. In this way, a detrending is applied to the measurements, i.e. the common trends and the low frequency ionospheric variations (including the carrier-phase ambiguity) are filtered. The Laplacian operator nullifies constant values and an increase of constant slope. On the other hand, as it is a high-pass filter, it might increase the high frequency noise components. The detrending is done independently for each continuous carrier phase arc.

Therefore, the obtained values will consist mainly of medium and high frequency ionospheric variations, such as solar flares related over-ionization, TIDs or scintillation. It is important to remark that using the Laplacian is acceptable since the measurements are not affected by high frequency noise. As a typical example, Figure 1.3 shows the lack of high frequency components.

In Figure 1.3, for example, the effect of a solar flare in L_I and $\Delta^2 L_I$ is shown for all the satellites tracked by a particular GNSS receiver. In this case, the sudden increase of TEC is related to the GOES X17.2 flare on 28th October, 2003. The over-ionization is clearly observed from station *bahr* for the five satellites in-view with elevation above 30^o (see Section 1.4.5). In addition, this is a typical example where it can be seen that high frequency noise is absent.

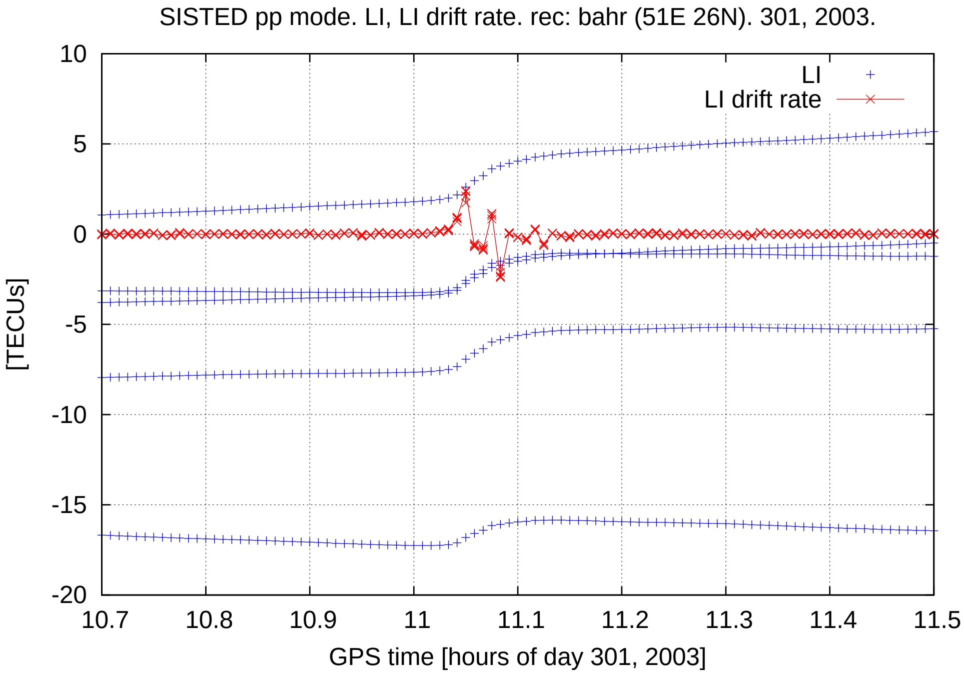


Figure 1.3: *L_I and its second finite difference in time ($\Delta^2 L_I$ or L_I drift rate) at 30 seconds sampling rate in function of time. Data obtained from the GPS signals gathered by the IGS receiver bahr during the GOES X-class flare reported on 28th October, 2003 and considering a masking angle of 30 degrees. The station longitude, latitude and mean Solar-Zenith angle are 51E, 26N and 56 degrees, respectively.*

To calculate each $\Delta^2 L_I$ value, we use three consecutive samples of L_I at 30 seconds sampling rate period in a continuous carrier phase arc of data.

As can be seen in Equation 1.1, each epoch $\Delta^2 L_I$ value is calculated from the current and the previous two samples (sample n and samples $n-1$ and $n-2$, respectively).

$$\begin{aligned}\Delta^2 L_I^n = \Delta L_I^n - \Delta L_I^{n-1} &= (L_I^n - L_I^{n-1}) + (L_I^{n-1} - L_I^{n-2}) \\ &= L_I^n - 2L_I^{n-1} + L_I^{n-2}\end{aligned} \qquad (1.1)$$

The Laplacian filter introduces a systematic delay of one sample. In other words, each $\Delta^2 L_I$ value corresponds to the drift rate of L_I at the previous sample. When there is a solar flare increase, the beginning of such increase is always between the previous and the processed epoch. In this way, in post-processing and real-time modes of operation, we are assuming a maximum latency of 30 seconds in the start time of an increase of ionization.

As an additional remark, the units are transformed from metres of L_I to TEC units or TECUs (1 TECU $= 10^{16} e^-/m^2$). Apart from that, remember that L_I is directly proportional to the ionospheric slant Total Electron Content (STEC; see Section A.2.1).

1.4.5 Ionospheric Pierce Points and Solar-Zenith Angle determination

The Ionospheric Pierce Points (IPPs) are the points where the GPS rays cross the ionosphere if we assume that the ionosphere can be approximately modelled as a single thin layer. In this research work, this is assumed considering a thin layer at a fixed height of $h_{iono} = 450Km$ (the same being used to generate IGS GIMs; see, for instance, [*Hernández-Pajares et al.* (2009)]). IPPs are calculated for each GPS receiver-satellite pair and at all epochs being processed.

In this context, it is important to set a masking angle in elevation (of 30^o in this work, although elevation masks as low as 15^o can be used as well). In case of low satellite elevation, the thin single layer mapping function, M (see Equation 1.3 and [*Orús-Pérez* (2005)]), could not be a good approximation and its use can produce large VTEC errors. In addition, the increased multipath at such elevations would also contribute to a noise increase (see, for instance, [*Ramos-Bosch* (2008)]).

From this point, the Solar-Zenith Angle (the angle between the Sun and the zenith of the IPP regarding to the Earth's geocentre; abbreviated as SZA or χ) can be determined. A diagram showing how to geometrically derive the SZA is depicted in Figure 1.4.

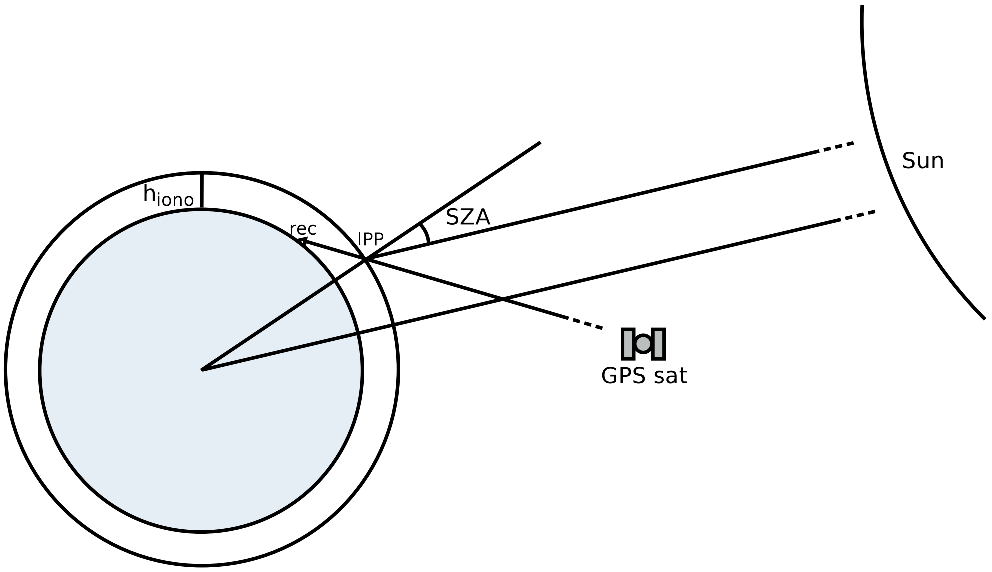


Figure 1.4: *Example of geometry diagram for a given GPS receiver-satellite pair, being SZA the Solar-Zenith Angle, IPP the Ionospheric Pierce Point, rec the GNSS receiver and sat the GPS satellite. For this purpose, a thin single layer ionospheric model at h_{iono} height is assumed.*

In case that a solar flare facing the Earth occurs, it is expected that most receiver-satellite rays in the daylight ionosphere will be affected by a simultaneous sudden over-ionization, especially near the subsolar point (i.e., $\chi = 0^o$). In contrast, the rays in the night-side ionosphere should not experience any significant change (see Table 1.1). In fact, solar flares produce a ionospheric over-ionization that is dependent on the SZA (see, for instance, [*Tsurutani et al.* (2009)]).

In Figure 1.5, the second VTEC drift rate values ($\Delta^2 V$) for the selected GPS receiver-satellite pairs with elevation above 30^o, are plotted against the SZA at 11h02m30s GPS time on 28^{th} October, 2003. In the sunlit hemisphere, this relationship approximately follows a cosine function with its maximum at the subsolar point location.

It is remarkable that the increase of TEC is still noticeable outside the sunlit boundary region as radiation can penetrate the ionosphere a longer path (depending on the adopted h_{iono} height; see, for instance, [*Zhang et al.* (2004)]). Most of the night-side $\Delta^2 L_I$ values (and so $\Delta^2 V$ values) are close to zero as they are not affected by the over-ionization. Nevertheless, some of them may still present higher or lower values due to scintillation.

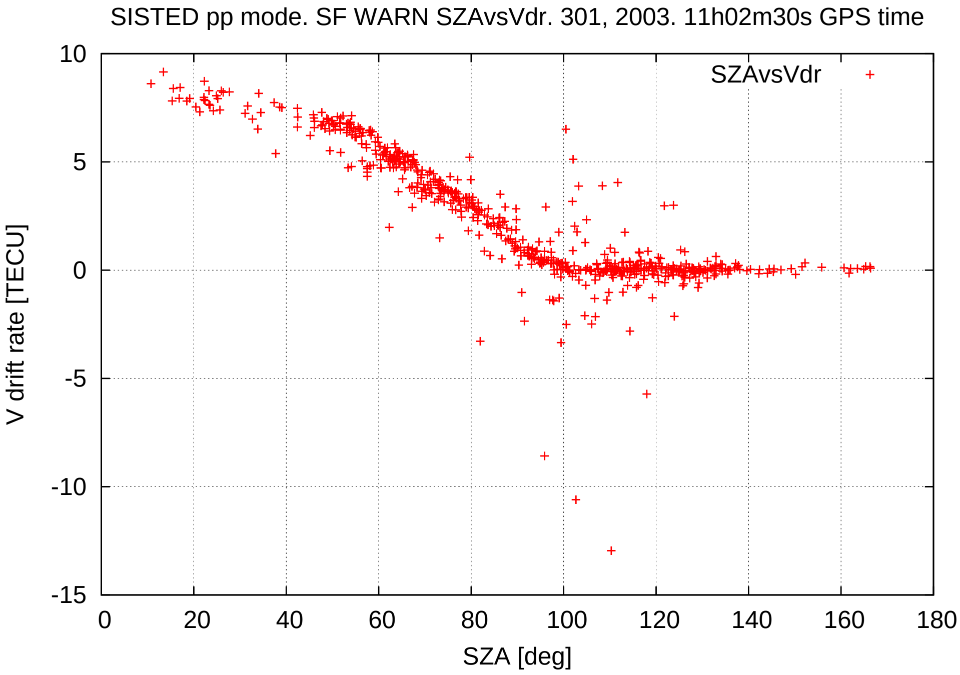


Figure 1.5: *Dependence of the VTEC drift rate ($\Delta^2 V$ or Vdr) on the Solar-Zenith angle (SZA, at the IPP locations) for the X-class solar flare, precursor of the Halloween storms (at 11h02m30s GPS time). Note that a few measurements with SZA between 80 and 120 degrees are affected by scintillation and thus, the corresponding $\Delta^2 V$ values are not close to zero.*

This dependence can also be seen in Figure 1.6 for the GOES X17.2 flare that occurred on 28th October, 2003. In it, the L_I variation and its second finite difference $\Delta^2 L_I$ in time have been depicted for GPS measurements in the sunlit ionosphere region r_1 and in the night-side ionosphere region r_2 (see also Table 1.1).

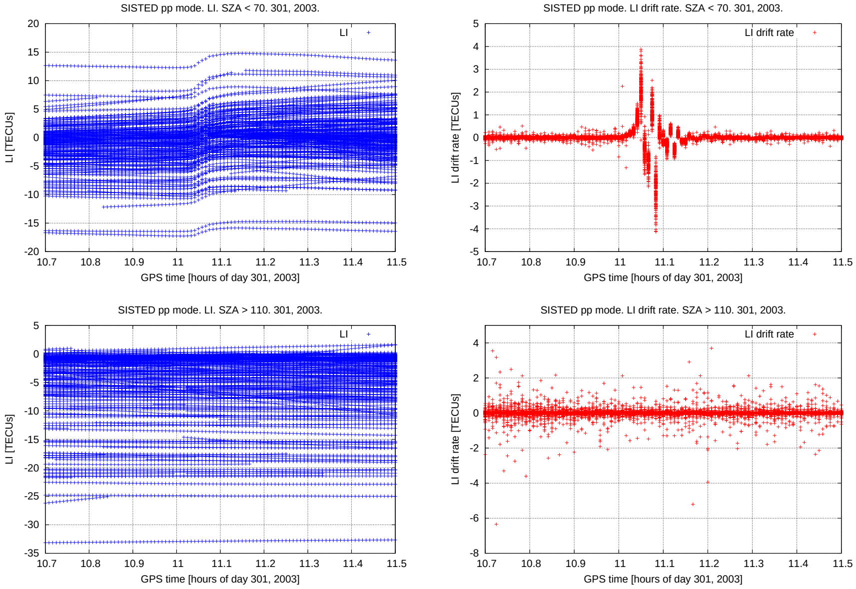


Figure 1.6: *GPS ionospheric combination L_I evolution and its second finite difference in time ($\Delta^2 L_I$) as function of GPS time (left-hand and right-hand plots, respectively) during the X17.2 solar flare on 28th October, 2003. A 30 seconds sampling rate and a masking angle of 30^o have been considered. A sudden increase of TEC is simultaneously given for the sunlit ionosphere SZA region (top plots) that does not affect the night-side ones (bottom plots).*

1.4.6 The Sunlit Ionosphere Sudden TEC Enhancement Detector (SISTED)

The SISTED detection technique, which has been developed from scratch in this research, monitors the variations of the ionospheric TEC in order to detect any sudden positive increase of ionization affecting the whole sunlit ionosphere. Such simultaneous sudden TEC enhancement would be caused by the increase of radiation related to a solar flare facing the Earth. In case of detection, SISTED is designed to send a warning message automatically to all interested users.

In brief, the obtained results for more than a half solar cycle show that the SISTED technique is sensitive to most of the powerful solar flares, whose related increase of radiation reaches the Earth. In addition, an important number of less powerful solar flares of M-class, and even C-class, can also be detected (see Section 1.5.4 for the details). It is also worth mentioning that a few B-class flares are also detected using SISTED. Their related ionospheric TEC response is very low but still detectable in very specific cases. Note that this should be further assessed.

Regarding its implementation, the main steps carried out within the SISTED detector are shown in Figure 1.7. From a global perspective, SISTED analyses whether an abrupt over-ionization is affecting each receiver-satellite pair observation. After analysing all available receiver-satellite pairs, it calculates the percentage of simultaneously affected observations (this percentage is also named Impact Parameter) for each ionosphere region. A certain threshold determines whether the Impact Parameter in the sunlit ionosphere region is high enough to be interpreted as a solar flare signature.

In more detail, the detection technique steps can be described as follows:

- SISTED is fed with the output data from the previous step (step IPP & SZA in the diagram in Figure 1.1). In this way, $\Delta^2 L_I$ and the SZA, among other parameters, are available for all the GPS receiver-satellite pairs and at all epochs being processed. Then, independently for each epoch, SISTED is executed.

- For each GPS receiver-satellite pair, we evaluate if the ray elevation is above a certain masking elevation angle (of 30^o in this research). This is important since a thin single layer model of the ionosphere has been adopted. Applying the corresponding mapping function M to transform between slant and vertical components can be a bad approximation for low satellite elevations (see Section 1.4.5). Apart from that, the carrier phase multipath (typically much lower than 1 cm) is neglected.

$$\Delta^2 L_I|_{thres} = \alpha_i * M * \Delta^2 V|_{thres}, \tag{1.2}$$

$$M = \frac{1}{\sqrt{1-(\frac{R_e * cos(E)}{R_e+h_{iono}})^2}}, \tag{1.3}$$

$$\tag{1.4}$$

where $\alpha_i = 0.105$ [m/TECU] is the conversion factor, M is the mapping function assuming a thin single layer model of the ionosphere

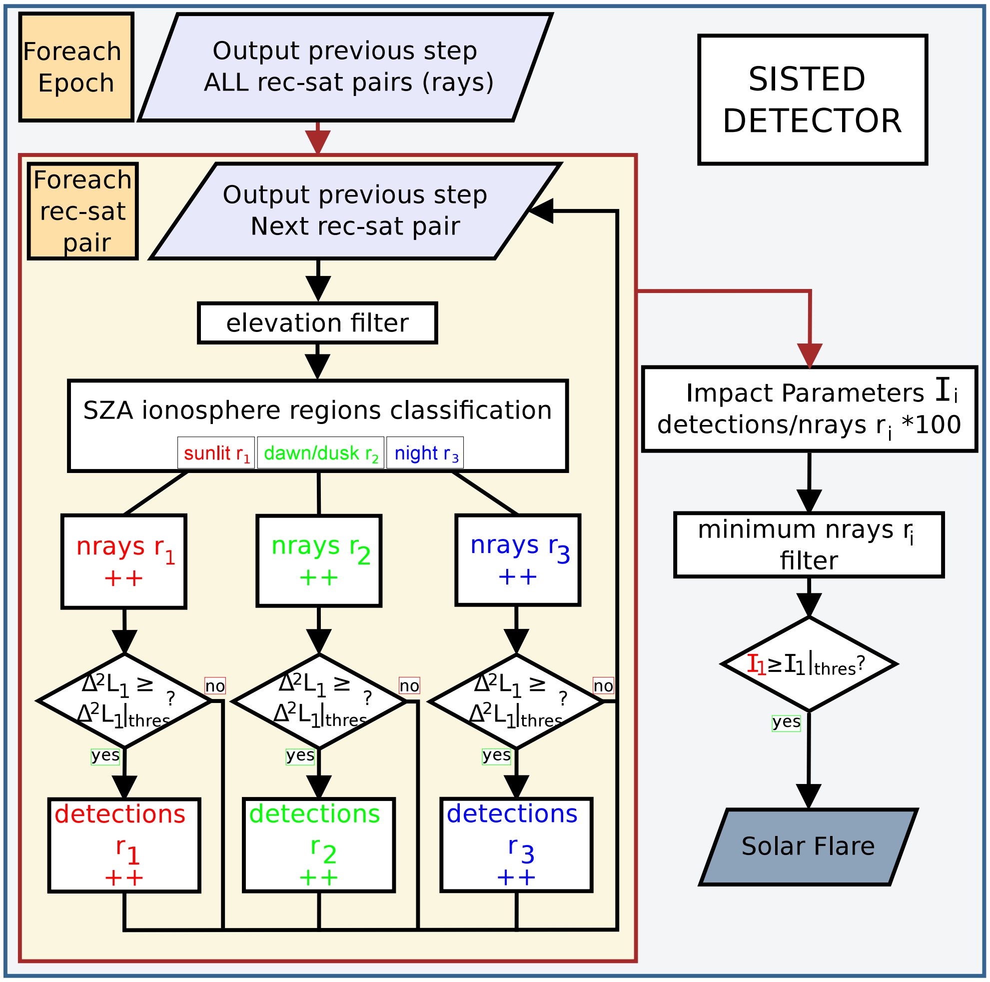


Figure 1.7: *SISTED detector steps. The abbreviation rec-sat pair corresponds to receiver-satellite pair (or ray).*

at a fixed height $h_{iono} = 450Km$, R_e is an approximation of the Earth radius and E is the elevation.

- Then, the observations are classified according to the SZA that is given at the corresponding IPP location (in this work, at the ionosphere height $h_{iono} = 450km$). In this way, three independent ionosphere SZA regions (r_i, where $i = 1, ..., 3$) are distinguished, as specified in Table 1.1. The boundaries correspond to values close to the solar terminator at the considered ionosphere height ($\cong 75.6^o$) and the last point being irradiated by the Sun ($\cong 110.93^o$). This last point is the intersection between the considered thin single layer ionosphere and the prolongation of the vector Sun-solar terminator at Earth's surface. It should be remarked that in such a case the radiation passes through a larger proportion of ionosphere than in the sunlit region. The SZA regions can also be observed in Figure 1.8 for a specific solar flare event.

Table 1.1: *Boundaries and approximate coverage for the three considered ionosphere regions in function of the Solar-Zenith angle, SZA or* χ*.*

Region Identifier	Coverage	SZA band
r_1	Sunlit ionosphere	$0^o \leq \chi < 70^o$
r_2	Dawn and dusk ionosphere	$70^o \leq \chi \leq 110^o$
r_3	Night-side ionosphere	$110^o < \chi \leq 180^o$

In order to track the number of GPS rays (receiver-satellite pairs) for each ionosphere region r_i, three counters are defined at each epoch n. In a generic way, we will refer to such counters as $nRays^n_{r_i}$.

- The value of the second finite difference, or Laplacian, of L_I ($\Delta^2 L_I$) is compared against the threshold to consider abrupt over-ionization, $\Delta^2 L_I|_{thres}$. In other words, this threshold corresponds to the minimum value to consider that a significant increase in $\Delta^2 L_I$ is given that might be related to a solar flare event.

  In case $\Delta^2 L_I$ is higher or equal to the threshold $\Delta^2 L_I|_{thres}$, the number of GPS rays (receiver-satellite pairs) detecting a tentative solar flare is incremented in the corresponding ionosphere region r_i at the epoch being processed (see left bottom in Figure 1.7). In a generic way, we will refer to such counters as $nRaysDet^n_{r_i}$.

  Note that the threshold is derived from its vertical component, the VTEC drift rate, labelled $\Delta^2 V|_{thres}$. In this way, the threshold does not depend on the GPS receiver-satellite ray inclination. Note that this threshold has been adjusted as explained in Section 1.5.3.

- After processing all available receiver-satellite pairs at a given epoch, the percentage of receiver-satellite pairs detecting a tentative solar flare is calculated for each ionosphere region r_i. As mentioned before, these percentages are named Impact Parameters (referred to as $I_1/I_2/I_3$ for the sunlight/dawn&dusk/night ionosphere regions) and they are useful

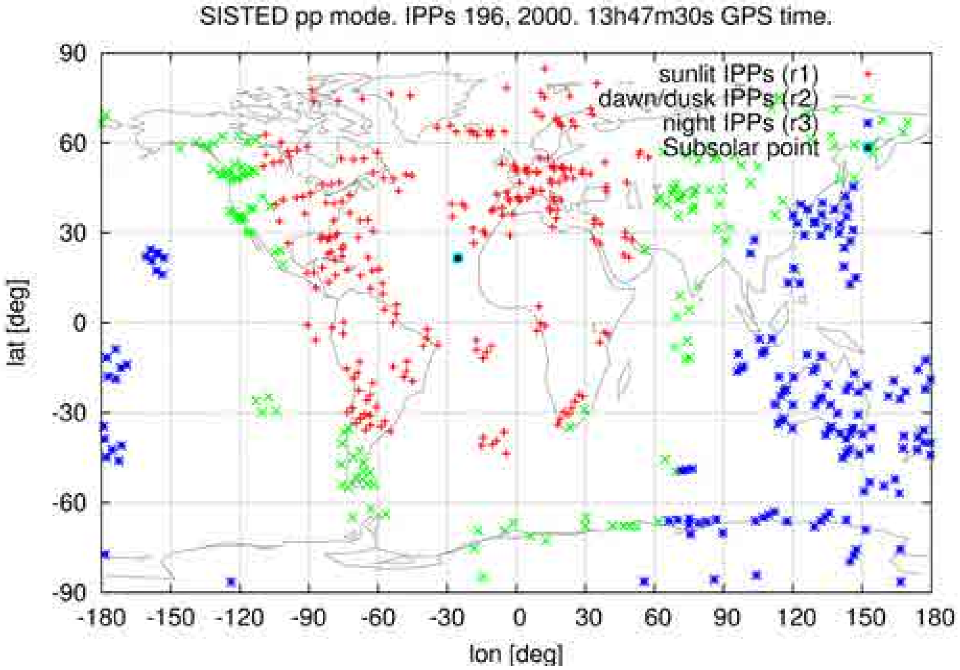


Figure 1.8: *Distribution of the IPPs according to the SZA regions at 13h47m30s GPS time for the day 196, 2000. The IPPs in the sunlit ionosphere region r_1 are marked in red, the ones in the dawn and dusk region r_2 in green and the ones in the night-side region r_3 in blue.*

since they give information on the simultaneity of the detection in a whole ionosphere region.

Equation 1.5 shows the way to derive the Impact Parameter for a generic ionosphere SZA region r_i at epoch n. And in Figure 1.9, an example of the temporal evolution of the Impact Parameters is depicted for the 31 days in August, 2011, showing detections compatible with GOES X- and M-class solar flares.

$$I_i^n = \frac{nRaysDet_{r_i}^n}{nRays_{r_i}^n} * 100, 1 \leq i \leq 3 \tag{1.5}$$

- To decide whether a solar flare has most probably occurred or not, a minimum number of rays must be given for each ionosphere region r_i.

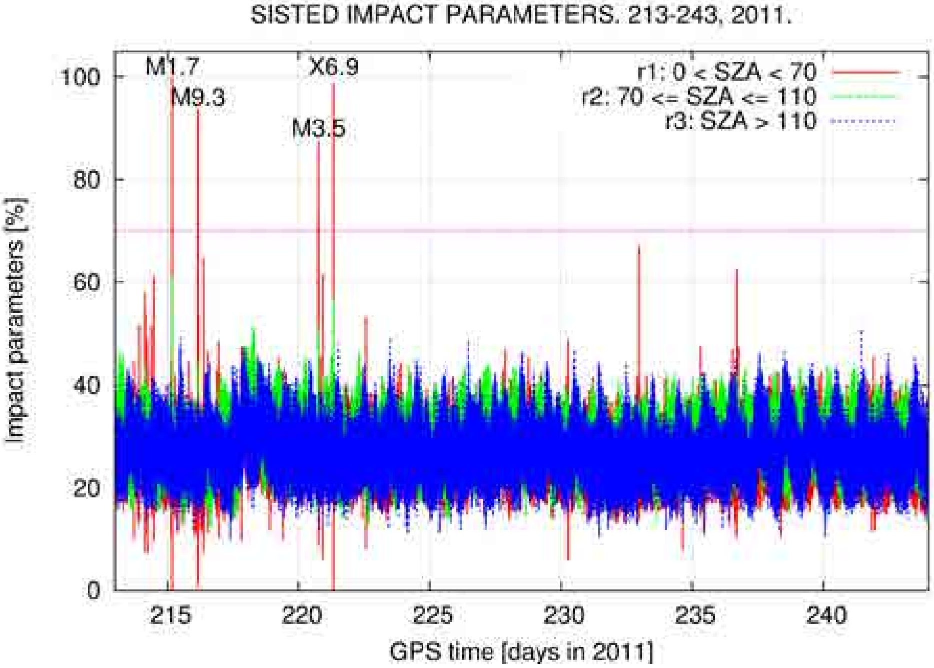


Figure 1.9: *Temporal evolution of the SISTED Impact Parameters obtained for the 31 days in August, 2011 considering* $\Delta^2 V|_{thres} = 0.1$ *and* $I_1|_{thres} = 0.70$*. The sunlit ionosphere region values are marked in red, the dawn and dusk values in green and the night-side ones in blue. The SISTED detections correspond to GOES X- and M-class solar flares and are marked for their identification.*

In addition, a solar flare warning will only be triggered if the Impact Parameter in the sunlit ionosphere region I_1 is equal or greater than a certain threshold of detection, labelled as $I_1|_{thres}$. This threshold fixes the minimum percentage of satellite-receiver pairs detecting an over-ionization over $\Delta^2 V|_{thres}$ to assume that there is a general increase of ionization in the whole or most part of the sunlit hemisphere at the epoch being processed.

- Finally, each time there is a detection, SISTED sends warning alerts via e-mail to the interested users and generates additional plots as specified in Appendix B.

It is worth mentioning that another SISTED condition was considered. This condition was that a minimum ratio between I_1 value in the sunlit ionosphere region r_1 and I_3 value in the night side region r_3 had to be given at the epoch being processed, since the increase of radiation related to a solar flare does not affect the night-side ionosphere. Nevertheless, the results were not improved and the condition was rejected (refer to Section 1.5.3 for the details).

Note that I_2 for dawn and dusk ionosphere region is not taken into account to determine whether a solar flare has occurred or not. Nevertheless, this information is processed and stored. In this way, it could be used in the future to study the ionospheric response at such regions when solar flares facing the Earth occur. This parameter could also give extra confidence on the SISTED detections (see Figures 1.19 and 1.21), though this should be further investigated.

It is important to remark that a solar flare warning would not be triggered just due to other ionospheric perturbations, such as TIDs and scintillation (see [*García-Rigo et al.* (2008a)]). The contribution by such other perturbations would be filtered out since these phenomena cause disturbances in a local scale, not in most of the sunlit ionosphere region. As these effects may only affect a few satellites in specific local areas, the overall results in the whole sunlit ionosphere region (obtained from a global network of IGS receivers) would not be altered. In addition, the solar-azimuthal dependence will not be given. Furthermore, TIDs are also not detected due to the implicit filtering of the second consecutive differences of L_I, taken at 30 seconds interval (too small when compared with the targeted TID periods of many hundreds of seconds).

As a final remark, since we are processing $\Delta^2 L_I$ measurements every single epoch in an independent way (without taking into account previous or next $\Delta^2 L_I$ values), we may detect a concrete solar flare more than once due to the fact that the enhancement pattern in L_I may not be a steep continuous gradient. This also applies for tentative false detections. For instance, Table 1.10 in Section 1.5.4 shows that the ionospheric response given for the so-called La Bastille day flare is detected for as much as 18 different epochs very close in time. For the period being analysed and for the selected parameters, C-class flares are usually detected at a single epoch or no more than three times. If there are a high number of detections close in time, it is very likely that a powerful flare has been detected by SISTED. In any case, a SISTED warning is triggered for every epoch detecting a sunlit generalised $\Delta^2 L_I$ increase. Also in this context, we are assuming that the VTEC recombination process after an overionization is a slow process.

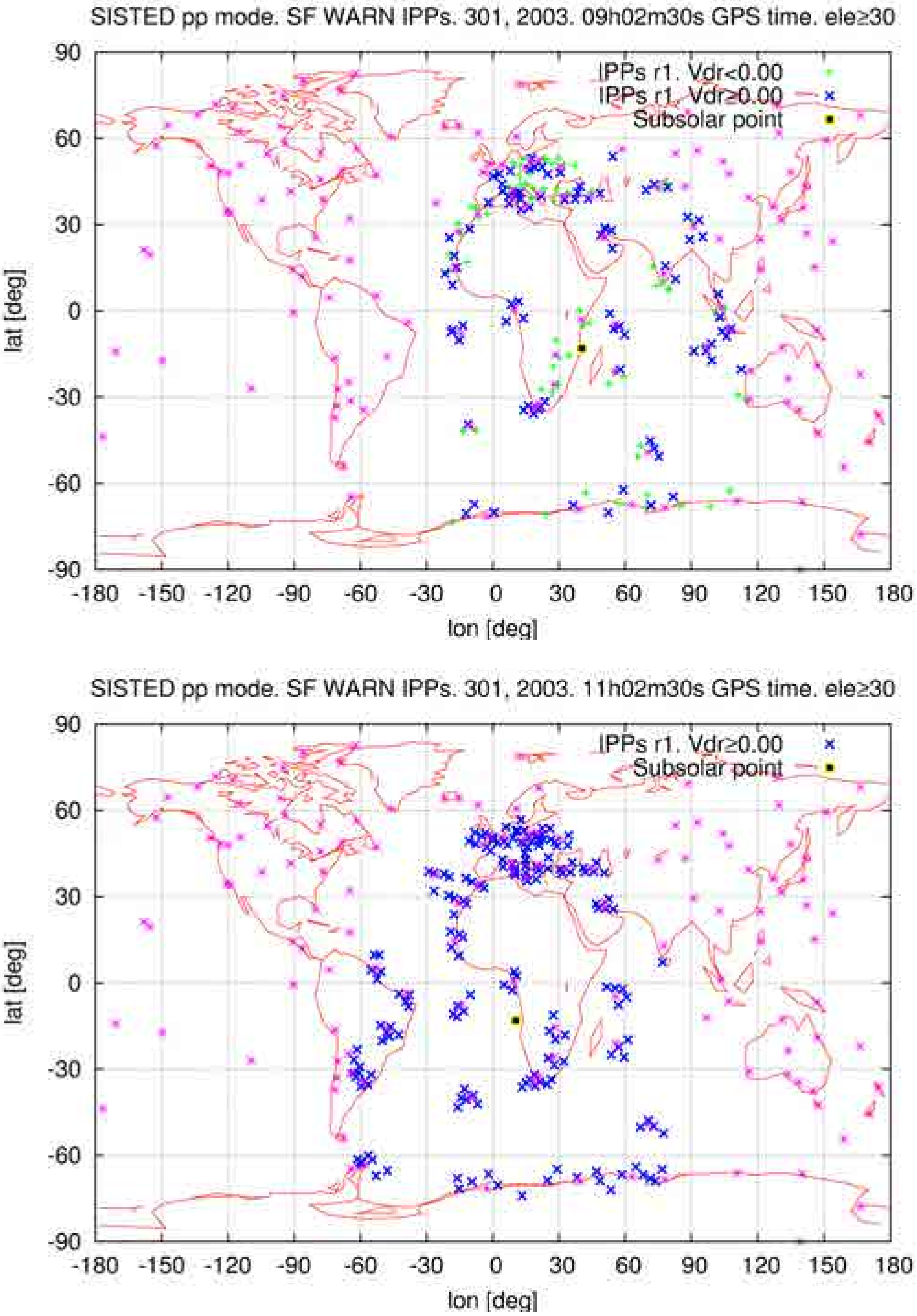


Figure 1.10: *Distribution of the IPPs according to the SZA regions at 09h02m30s and 11h02m30s GPS time for the 28th October, 2003. The sunlit ionosphere region IPPs with over-ionization (positive VTEC drift rate, labelled Vdr) are marked in blue. The IPPs without over-ionization are marked in green. The location of the GPS receivers is marked in magenta and the Subsolar point in yellow. Note that at 11h02m30s GPS time (lower plot), all $\Delta^2 V$ values were simultaneously positive due to the solar flare.*

1.5 Performance analysis

1.5.1 Datasets and external sources for SISTED evaluation

Two separate datasets will be distinguished: the training and the test dataset. The training set will be used to adjust the SISTED parameters. The test set will be used to assess the performance of the detector with independent data.

The performance will be given in terms of the percentage of correctly detected X-class solar flares in the test dataset when keeping the percentage of false alarms below a certain maximum. Note that we are focusing on the detection of the most powerful X-class flares according to GOES classification (see Table 1.2).

In order to validate a SISTED detection, we take into account GOES records and the optical flares in H-alpha reported in the Edited Solar Events Lists (FLA records). In addition, the SOHO-SEM EUV data and the GNSS Solar FLare Activity Indicator (GSFLAI) data are also used in specific cases for visual comparison. As a final remark, the performance evaluation is carried out automatically.

The following subsections include a brief description of the external data sources being used for SISTED performance evaluation.

GOES

GOES satellite series ([*Hill et al.* (2005)]) consist on satellites equipped with solar X-ray Sensors (XRS) to monitor the solar radiation in two X-ray bands (0.1 - 0.8 nm and 0.05 - 0.4 nm). Data since September, 1996 from GOES08 to GOES15 satellites are publicly available through the FTP site of the National Oceanic and Atmospheric Administration (NOAA) National Geophysical Data Center (NGDC) at ftp://ftp.ngdc.noaa.gov. Data are also part of the Edited Solar Events Lists files at NOAA's Space Weather Prediction Center (SWPC) FTP site (at http://www.swpc.noaa.gov; denoted as XRA events).

An X-ray event is reported when four consecutive X-ray values at 0.1 - 0.8 nm band satisfy the following conditions (ftp://ftp.ngdc.noaa.gov):

1. *All 4 values are above the B1 threshold*
2. *All 4 values are strictly increasing*
3. *The last value is greater than 1.4 times the value which occurred 3 minutes earlier*

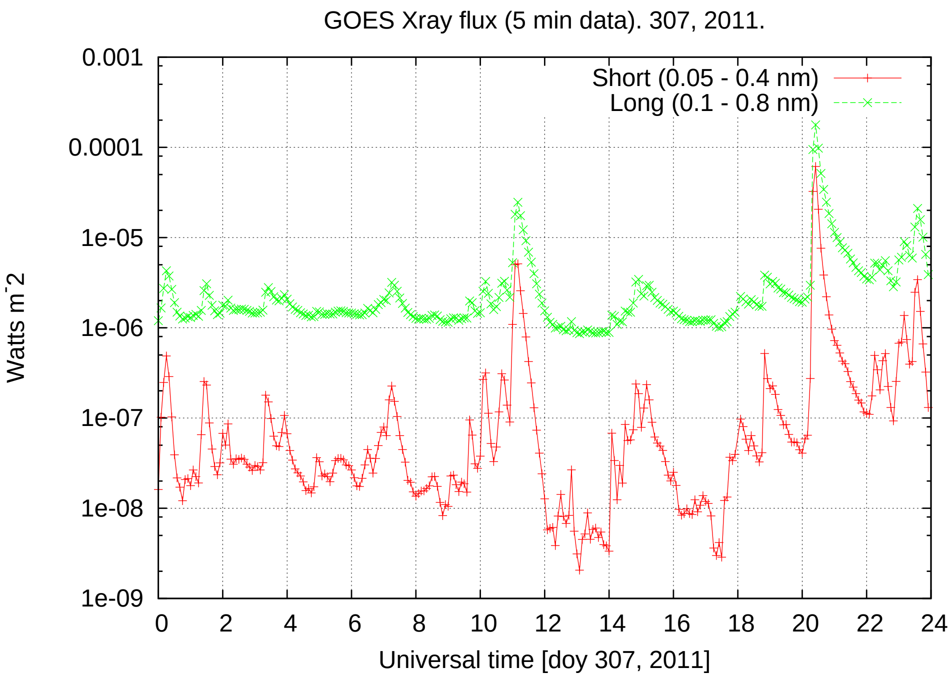


Figure 1.11: *GOES15 records for 0.05 - 0.4 nm and 0.1 - 0.8 nm bands on* 3^{rd} *November, 2011. Note that an X1.9 flare was reported with its maximum peak at 20:27 UTC time.*

Additionally, it is of great importance that GOES has established a classification of solar flares according to their level of irradiance. This classification, which is shown in Table 1.2, is widely used in the scientific community as a reference. The X-class flares are the most powerful and the A-class flares are the least powerful in terms of their irradiance.

Table 1.2: *GOES X-ray classification according to the peak burst intensity (I) measured in the 0.1 - 0.8 nm band (from ftp://ftp.ngdc.noaa.gov).*

Class	W/m^2	$Ergs/cm^2/s$
B	$I < 10E^{-6}$	$I < 10E^{-3}$
C	$10E^{-6} \leq I < 10E^{-5}$	$10E^{-3} \leq I < 10E^{-2}$
M	$10E^{-5} \leq I < 10E^{-4}$	$10E^{-2} \leq I < 10E^{-1}$
X	$I \geq 10E^{-4}$	$I \geq 10E^{-1}$

The comparison of SISTED results with GOES data is automatically performed considering the events where the *begin* and *end* times are well defined (nor missing neither uncertain). In this context, it must be taken into account that GOES uses Coordinated Universal Time (UTC) time instead of GPS time (at the moment, this implies a 15 seconds gap). In addition, GOES data are made available at 1 minute sampling resolution. Therefore, 30 additional seconds of uncertainty are given that must be taken into account.

Finally, data in the Edited Solar Events Lists are updated every half an hour. It must be remarked that these records are compiled from preliminary reports and may be incorrect or incomplete so that changes may be introduced in the following three days.

SOHO Solar EUV Monitor

Another external source that provides evidence of a solar flare increase of radiation is the SOHO Solar EUV Monitor instrument (SOHO/SEM; [*Judge et al.* (1998)]). In this case, information in the 26 -34 nm band, which corresponds to the Extreme Ultra-Violet band (EUV), is measured. As mentioned before, this information is relevant since solar flares' increase of radiation is mainly produced in X-rays and EUV bands.

It is important to mention that the SEM EUV measurements are provided in UTC time and that, in this work, the 15 seconds average values are being used. These are the highest rate records that are publicly available through the SOHO/SEM public site (at http://www.usc.edu/dept/space_science/semdatafolder/). Note that the values correspond to the middle point of the sample and may not always be given at exact regular intervals of 15 seconds. Regarding latency, SEM EUV data become available typically after several days (not in real time, at least for public access).

In this work, data retrieved by the SOHO/SEM instrument have also been used to validate the SISTED results. This has been done by visual comparison for the selected X-class solar flares in Section 1.5.4 as well as those that were neither validated with GOES records nor with the optical flares in H-alpha reported in the Edited Solar Events Lists for such days. As an example, the 26 - 34 nm flux signature and its second finite difference for the flare preceding the Halloween storm events are depicted in Figure 1.12.

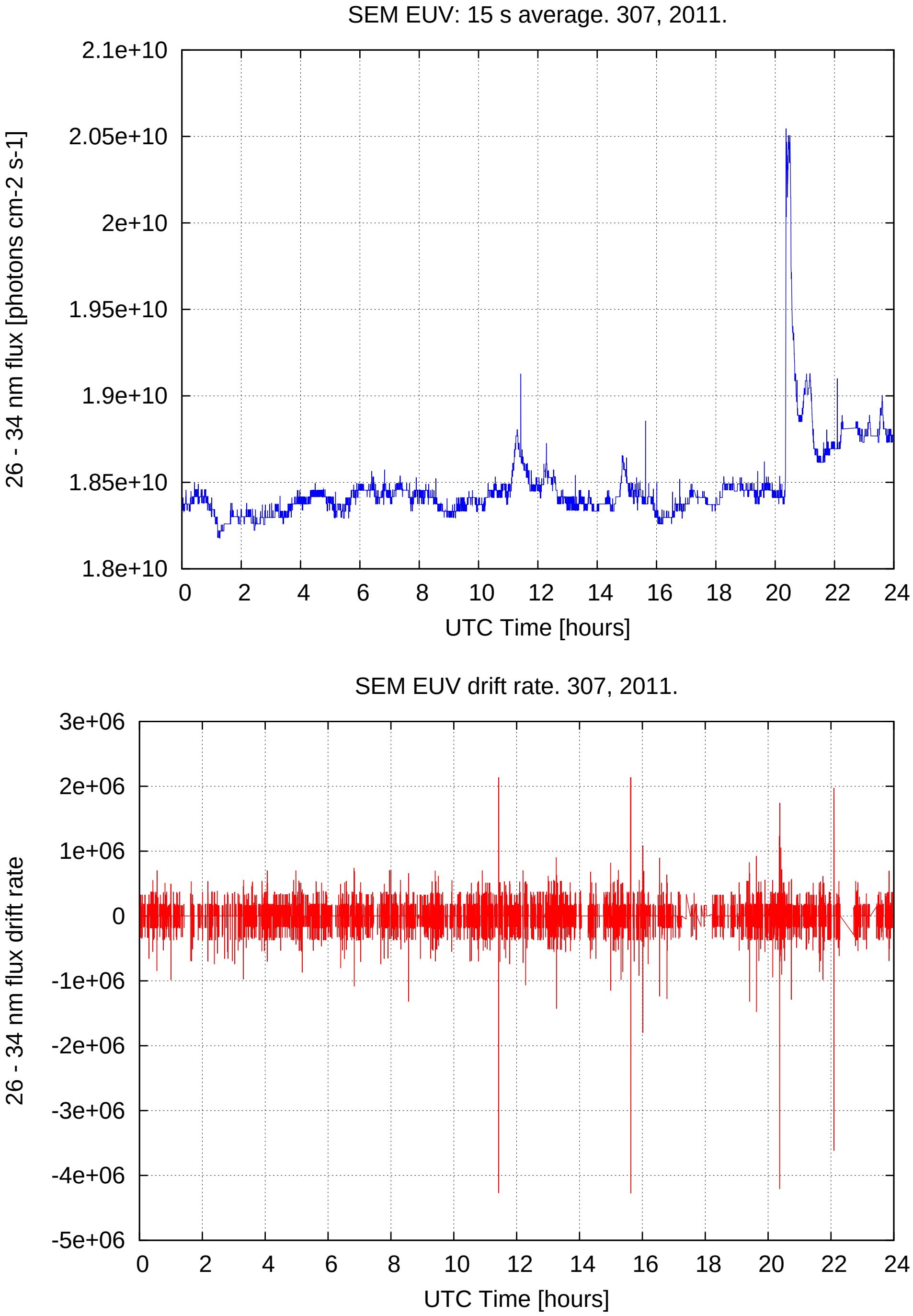


Figure 1.12: *15 seconds average records of the SEM-EUV 26 - 34 nm flux on* 3^{rd} *November, 2011. GOES X1.9, M2.5, C5.4 and C3.8 flares were reported with their maximum peaks at 20:27, 11:11, 22:18 and 14:54 UTC time, respectively (see previous Section 1.5.1). Note that there are a few short data gaps, especially at the end of the day.*

Other sources

1. **Optical flares detected in H-alpha** Information on these events is available in NOAA's Edited Solar Events Lists (http://www.swpc.noaa.gov) and they are labelled as FLA records. These records are mainly provided by the following reporting observatories: Learmonth (labelled LEA; located in Australia), Holloman AFB (HOL; USA), Ramey AFB (RAM; USA) and San Vito (SVI; Italy) and consist of observations taken from photographic, electronic or visual data. Solar H-alpha Flare events are published since 1980. These measurements distinguish between subflares and powerful flares in function of *the corrected area of the flare in heliospheric square degrees at maximum brightness, observed in the H-alpha line* (656.3 nm; this information is provided at http://www.swpc.noaa.gov/ftpdir/indices/events/README).

2. **GNSS Solar FLAre Indicator (GSFLAI)** It is given by the solar-azimuthal gradient of the VTEC rate measured in a sparse global network of dual-frequency GPS receivers. This index has also been created by gAGE/UPC as a potential proxy of the EUV solar flux rate and can be very useful to evaluate the tentative false detections, as well as the SISTED behaviour at high-rate mode, since both methods are fed with GPS data. This index could be complementary to SISTED.

3. **NASA's Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI)** A flare list is generated based on the observing summary count rate in the 6 to 12 keV energy band. Currently, data from RHESSI, which are available at ftp://hesperia.gsfc.nasa.gov/, are not used to validate the results. Note that this source is not taken into account in this research.

4. **NASA's Solar Dynamics Observatory (SDO)** It was launched on 11^{th} February, 2010 and provides solar flare images in multiple energy bands. SDO data can be accessed at http://sdo.gsfc.nasa.gov. Note that this source is not taken into account in this research.

1.5.2 Selected periods

Initially, data on years 2001 and 2005 have been considered as training dataset (also referred to as TRN) to adjust the parameters of the detector. As stated in the previous section, this has enabled to adjust several parameters of the SISTED technique so that most of the 39 X-class solar flares that were recorded by GOES satellites during the 730 day period can be detected, while keeping tentative false detections lower than a certain maximum. In 2001, solar activity was high (solar cycle maximum conditions), and quiet/disturbed geomagnetic conditions were given. Also in year 2005, solar activity was high but during solar cycle minimum conditions and quiet/disturbed geomagnetic conditions were also given.

The main reason to use post-processing data (among a better characterisation of the proposed technique) is that the availability of real-time streams was very limited in the past. Nonetheless, the adjusted parameters are also used in real-time mode of operation since the processing chain is equivalent (see diagram in Figure 1.1) and the sampling rate is also maintained. In fact, the main differences between both modes are the different GPS receivers being considered and the way input data are encapsulated.

In addition, a second independent period has been selected to evaluate the performance of the detector in post-processing mode when running against unseen data, i.e. data not taken into account in the training dataset. This period is named test dataset (also referred to as TEST) and includes every single day in the following 7 years: 1999, 2000, 2002, 2003, 2004, 2006 and 2007. In this way, we evaluate SISTED to correctly detect the X-class flares facing the Earth reported by GOES without incurring in a high false alarm level. Note that more than a half solar cycle is evaluated and that the data retrieved from 2557 days are processed. In addition, both maximum and minimum solar activity conditions are being evaluated, including both quiet and disturbed ionosphere as in the training dataset.

1.5.3 SISTED parameters adjustment

The main SISTED parameters or thresholds that have been adjusted are:

1. The minimum increase in $\Delta^2 V$ per receiver-satellite ray crossing the ionosphere (named $\Delta^2 V|_{thres}$ threshold)

2. The minimum value of the sunlit Impact Parameter to consider an overall and simultaneous over-ionization caused by a solar flare facing the Earth (named $I_1|_{thres}$ threshold)

3. The ratio between sunlit and night-side Impact Parameters ($I_1/I_3|_{thres}$ threshold; finally not considered)

As mentioned before, these thresholds have been adjusted through a training process using data from years 2001 and 2005. In addition, they are selected as a compromise between the number of detected GOES X-class flares (also XRA records) and the number of tentative false alarms for the training dataset. This has been done by evaluating the results by means of the Receiver Operational Characteristics (ROC) curves (see, for instance, [*Duda, Hart et al.* (2001)] and [*Swets* (1996)]). The use of ROC is justified because it gives the performance (i.e. correct detections, and false alarms) as a function of a threshold.

For such a purpose, the percentage of detected GOES X-class flares is considered as the true positive rate (y-axis parameter; *sensitivity*) and the percentage of detections that have not been validated is interpreted as the false positive rate (x-axis parameter; *1-specifity*).

Regarding false detections, they are considered when a SISTED detection is neither validated with GOES records nor with the optical flares in H-alpha (see Section 1.5.1). Since we are only evaluating whether a SISTED detection is false or not, the comparison with optical flares will not be necessary if the detection is already compatible with GOES records. Note that a SISTED detection may be compatible with multiple optical flares.

In this context, the number of false detections can be considered an upper bound (pessimistic scenario). This is because we are comparing events obtained from indirect ionospheric response with events obtained at the wavelengths corresponding to XRA/FLA records derived from direct measurements at different bands, which are not covering 100% of events. For example, in Section 1.5.4 a tentative false detection was considered correct. In addition, other sources of flare events and ionospheric monitoring (such as RHESSI or GSFLAI, respectively), may bring light to a part of such false detections. This reaffirms considering the false detections rate as an upper bound of the real value.

In the following subsections, the way to adjust the three main SISTED parameters is shown. For this purpose, two tests have been carried out. In each test, ROC curves are generated for multiple values of two (out of the three adjustable) SISTED parameters.

SISTED $\Delta^2 V|_{thres}$ parameter adjustment

The first test is carried out to fix the first of the above mentioned parameters, the threshold $\Delta^2 V|_{thres}$. In this case, SISTED is executed for the training days considering a certain range of values for the first two adjustable parameters: 0.00 to 0.35 $TECUs$ for the minimum $\Delta^2 V$ threshold in 0.01 or 0.05 steps (note that three samples are taken into account at a sampling rate period of 30 s; see the legend in such figures) and 0 to 100 % for $I_1|_{thres}$ in 2 % steps (these points are remarked in each curve; see also Figure 1.17). Note that the condition on the ratio between sunlit and night-side Impact

Parameters is left aside in this first test (the corresponding threshold is set to 0.0). The ROC results of this test are depicted in Figures 1.13 and 1.14.

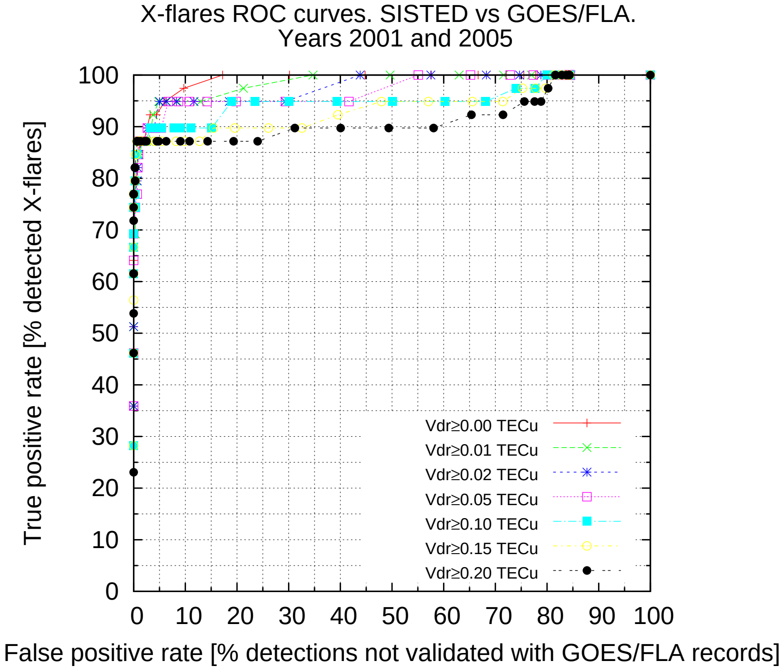


Figure 1.13: *SISTED ROC curves for the detection of GOES X-class flares in years 2001 and 2005. Each curve corresponds to a different* $\Delta^2 V|_{thres}$ *threshold (Vdr in the legend) to consider sudden VTEC increase. In addition, each curve is built taking into account steps of 2 % of the minimum value of the sunlit Impact Parameter* $I_1|_{thres}$ *(ranging from 0 % to 100 %).*

The results show a good performance of the solar flare detector that can be optimised when considering the minimum $\Delta^2 V|_{thres}$ threshold parameter equal to zero (red curve). The criterion for selecting the optimal threshold was maximising the area below the ROC curve[5] (values are given in Table 1.3). The reason for this criterion is that the area of the ROC allows for a comparison of performance that does not depend on the Impact Parameter detection threshold. Therefore, the bigger the area the better the performance because for a given rate of false positives, the ROC with a bigger area will give a higher true positive rate. In this regard, we are assuming that the statistical expectation is concave and monotonous.

---

[5] In order to calculate the ROC area, it was necessary to close each ROC curve at 100 % detections when considering 100% false detections, i.e. to the upper point of the figure.

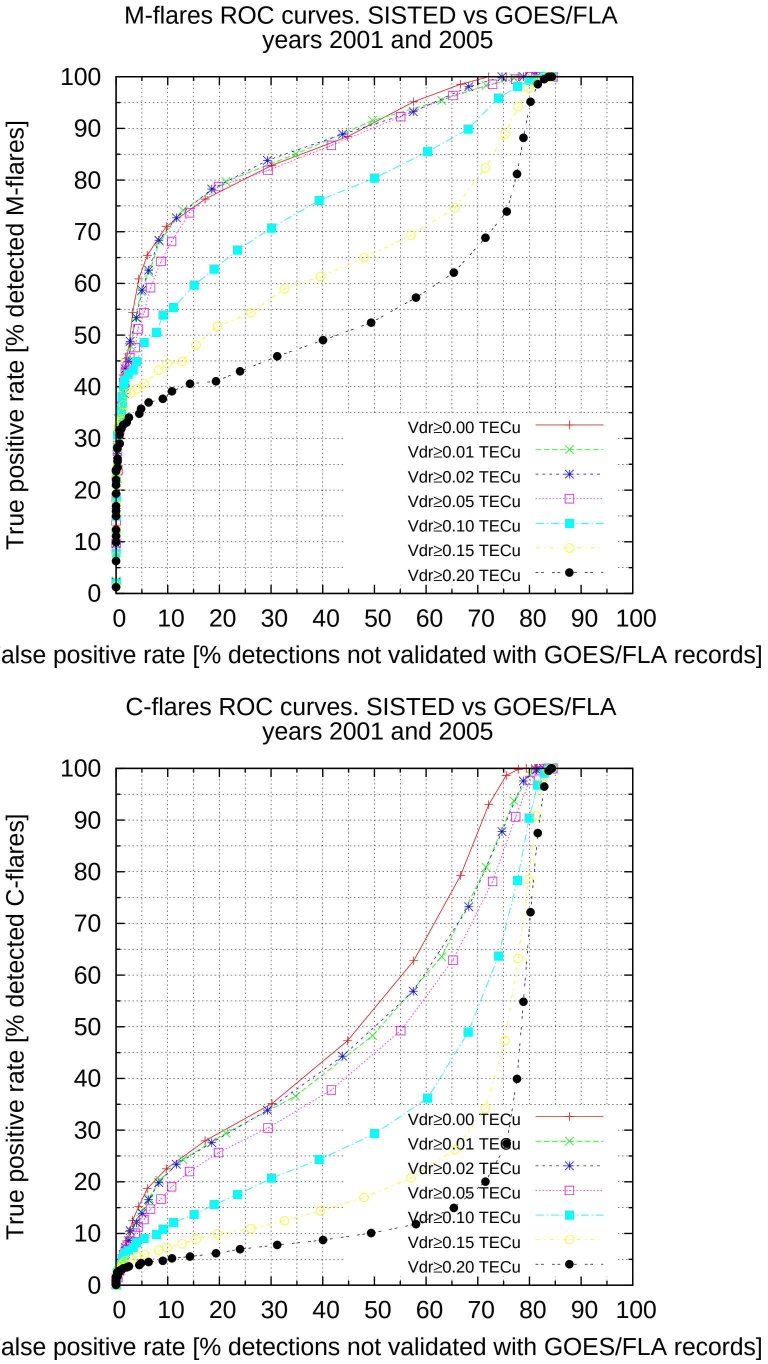


Figure 1.14: *SISTED ROC curves for the detection of GOES M-class (top) and C-class flares (bottom) in years 2001 and 2005. Each curve corresponds to a different* $\Delta^2 V|_{thres}$ *threshold (Vdr in the legend) to consider sudden VTEC increase. In addition, each curve is built taking into account steps of 2 % of the minimum value of the sunlit Impact Parameter* $I_1|_{thres}$ *(ranging from 0 % to 100 %; remarked points).*

Table 1.3: *ROC area for each of the ROC curves being compared in order to fix the* $\Delta^2 V|_{thres}$ *threshold. SISTED results are obtained for the training period so that all days in years 2001 and 2005 are processed. Note that the area below each ROC curve is given as a percentage and that the condition on the minimum ratio between day and night is not applied in this first test.*

Period	$\mathbf{I_1\|_{thres}}$	$\mathbf{I_1/I_3\|_{thres}}$	$\mathbf{\Delta^2 V}$	**ROC area**		
				X	**M**	**C**
2001/2005	$0 \leq I_1\|_{thres} \leq 100\%$	-	**0.00**	**99.1%**	**87.8%**	**59.3%**
			0.01	98.5%	87.4%	56.3%
			0.02	97.7%	87.6%	56.5%
			0.05	97.0%	86.6%	53.4%
			0.10	95.2%	79.2%	43.4%
			0.15	93.6%	69.7%	34.8%
			0.20	91.7%	61.0%	28.9%

Note that the ROC results are meaningful from 2 % of false alarms, where the ROC curves separate and the areas of the curves for the thresholds 0.00-0.05 are clearly better than the others. Before that value, the sampling variability is not uniform and a change in the classification of one sample is reflected in a jump on the number of detections. This gives rise to the overlap of the curves at such region.

The detection is done by majority vote and, therefore, the I_1 threshold can be very low. In this context, it can be remembered that SISTED in post-processing mode requires at least 50 GPS receiver-satellite pairs for each ionospheric region taken into account. Nonetheless, the number of rays is, in general, much higher (see, for instance, Figure 1.23).

SISTED $I_1/I_3|_{thres}$ parameter adjustment

In the second test, the purpose is to adjust the minimum ratio between day and night Impact Parameters ($I_1/I_3|_{thres}$). ROC results are given in Figures 1.15 and 1.16. For this test, $\Delta^2 V|_{thres}$ is fixed to 0.0 TECU from the results of the first test. $I_1/I_3|_{thres}$ values being checked range between 0.00 and 3.00 in 0.25 steps (see the legend in both figures). $I_1|_{thres}$ values range between 0 and 100 % in 2 % steps, as in the first test. To further understand the scenarios being evaluated, if $I_1/I_3|_{thres} = 1.5$, the Impact Parameter in the sunlit ionosphere region r_1 will have to be 50% higher than in the night-side ionosphere region r_3 in order to consider solar flare detection. If $I_1/I_3|_{thres} = 1.0$, then I_1 will have to be equal or higher than I_3.

Note that ratios above 2.5 do not produce any acceptable detection and thus, curves have a ROC area equivalent to 50% (since we are closing each curve at 100% flare detections for 100% false detections, as mentioned before). This can be observed in Table 1.4.

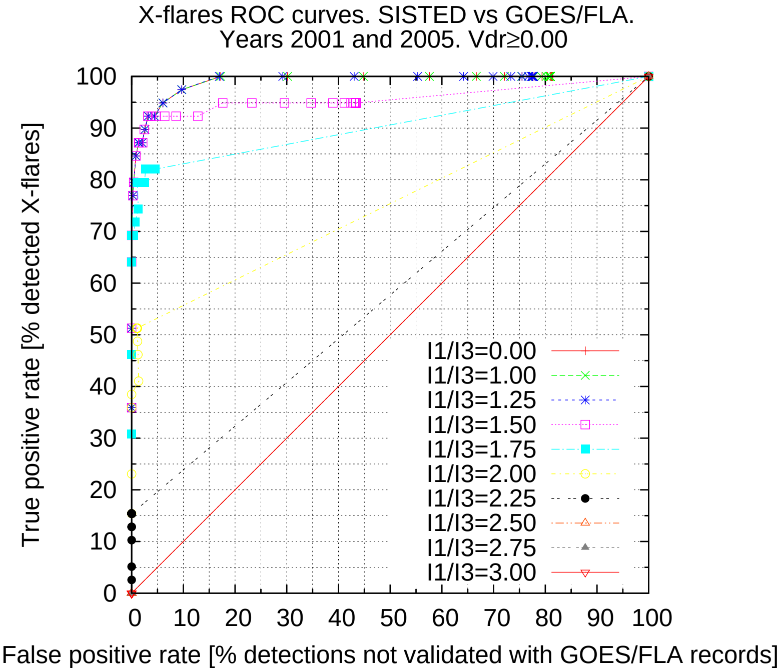


Figure 1.15: *SISTED ROC curves for the detection of GOES X-class flares in years 2001 and 2005. Each curve corresponds to a different* $I_1/I_3|_{thres}$ *threshold (see the legend) to consider sudden VTEC increase. In addition, each curve is built taking into account steps of 2 % of the minimum value of the sunlit Impact Parameter (ranging from 0 % to 100 %; emphasised points). In this case, the* $\Delta^2 V|_{thres}$ *threshold is fixed to 0.0 from the first test.*

The results confirm that this condition is not required for solar flare detection, as it was advanced in Section 1.4.6. From Figures 1.15 and 1.16, this condition is discarded since there is no improvement when considering the two complete years in the training set.

SISTED $I_1|_{thres}$ parameter adjustment

Finally, $I_1|_{thres}$ parameter can be adjusted from the first test results since the condition on $I_1/I_3|_{thres}$ is not required. The selection of the optimal value will depend on the user needs. If the user accepts a 7% of false alarms, SISTED will probably enable the detection of the 95% of GOES X-class flares. If the

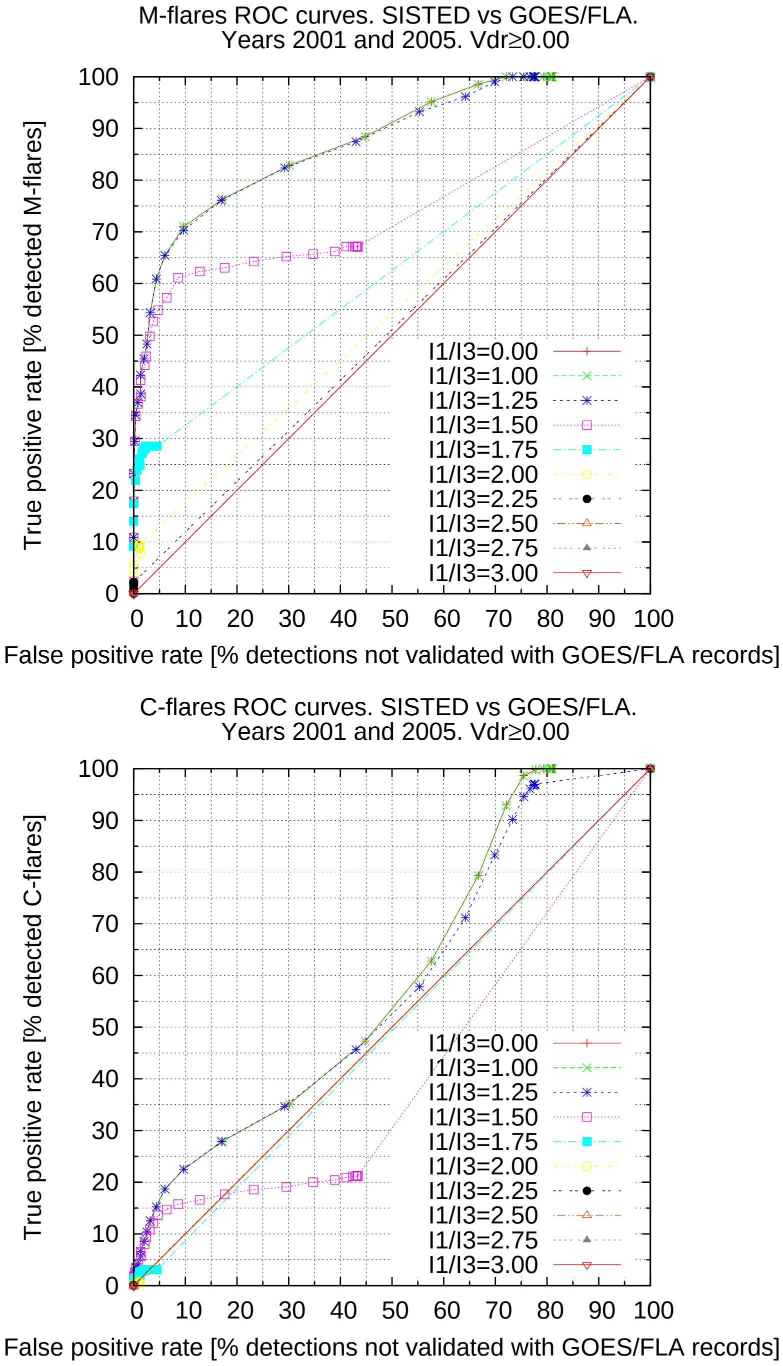


Figure 1.16: *SISTED ROC curves for the detection of GOES M-class (top) and C-class flares (bottom) in years 2001 and 2005. Each curve corresponds to a different* $I_1/I_3|_{thres}$ *threshold (see the legend) to consider sudden VTEC increase. In addition, each curve is built taking into account steps of 2 % of the minimum value of the sunlit Impact Parameter (ranging from 0 % to 100 %; emphasised points). In this case, the* $\Delta^2 V|_{thres}$ *threshold is fixed to 0.0 from the first test.*

Table 1.4: *ROC area for each of the ROC curves being compared in order to fix the* $I_1/I_3|_{thres}$ *threshold. Results are obtained for the training period so that all days in years 2001 and 2005 are processed. Note that the area below each ROC curve is given as a percentage and that the* $\Delta^2 V|_{thres}$ *is fixed to 0.0 from the first test.*

Period	$\mathbf{I_1\|_{thres}}$	$\mathbf{I_1/I_3\|_{thres}}$	$\mathbf{\Delta^2 V}$	**ROC area**		
				X	**M**	**C**
2001/2005	$0 \leq I_1\|_{thres} \leq 100\%$	**0.00**	**0.00**	**99.1%**	**87.8%**	**59.3%**
		1.00		99.1%	87.8%	59.3%
		1.25		99.1%	87.5%	58.2%
		1.50		95.7%	74.1%	41.8%
		1.75		90.5%	62.5%	49.3%
		2.00		75.3%	54.2%	49.8%
		2.25		57.7%	51.1%	50.0%
		2.50		50.0%	50.0%	50.0%
		2.75		50.0%	50.0%	50.0%
		3.00		50.0%	50.0%	50.0%

user accepts a 4% of false alarms, SISTED will detect almost 93% of the GOES X-class flares. In any case, note that these results are obtained from the training set and will have to be confirmed in the test set (see Section 1.5.4).

In Figure 1.17, a zoom of the selected ROC curve for the detection of X-class flares is depicted. Note that one of every two steps of the minimum value of the sunlit Impact Parameter $I_1|_{thres}$ is labelled.

For instance, considering a minimum percentage in the sunlit region $I_1|_{thres}$ of 74 %, SISTED could detect almost 95 % of the 39 GOES X-class flares occurring in the analysed interval with a false positive rate of about 7 % (and also 65 % of the 414 GOES M-class flares and 18 % of the 2679 GOES C-class flares) when considering a $\Delta^2 V|_{thres} = 0.00$ TECU.

In this way, we are accepting that the increase of TEC may not be observable for all the receiver-satellite rays in the sunlit ionosphere region r_1. Although the solar flare increase of radiation could be equivalent for the nearby IPPs, the presence of local perturbations like scintillation could be masking the results. In principle, this will only happen for a few receiver-satellite rays so that this would not be a problem as long as the sunlit Impact Parameter I_1 (i.e., taking into account the overall number of receiver-satellite rays in the sunlit ionosphere region r_1) is over the selected threshold $I_1|_{thres}$.

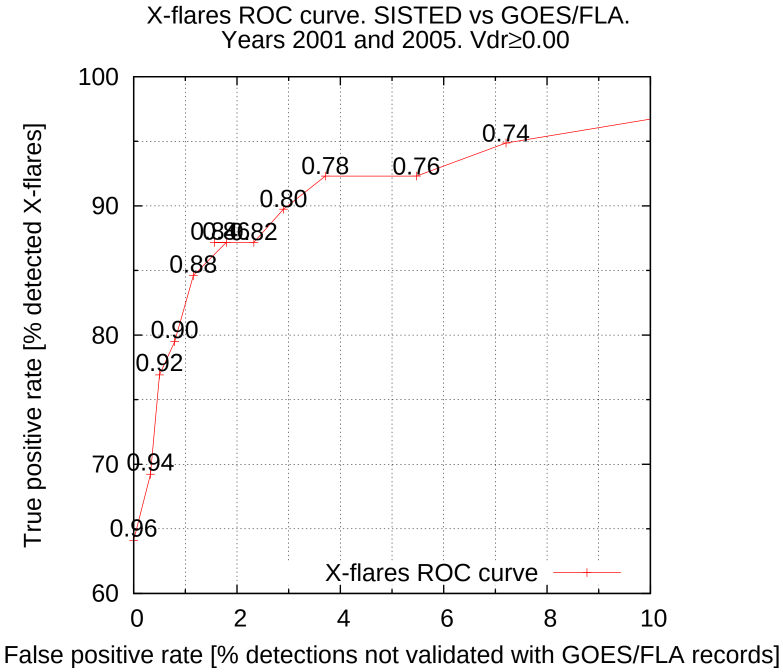


Figure 1.17: *Zoom of the selected SISTED ROC curves for the detection of GOES X-class flares. The curve is obtained when considering the following threshold:* $\Delta^2 V|_{thres} = 0.00$ *to consider sudden VTEC increase. The steps of 2 % of the minimum value of the sunlit Impact Parameter* $I_1|_{thres}$ *(ranging from 0 % to 100 %) are partially labelled.*

1.5.4 Results

Training dataset

The overall results for the training dataset, when considering the parameters being selected and accepting less than 7% / 4% of tentative false alarms, are given in Tables 1.5 / 1.6. These values are obtained from Figure 1.17, considering a minimum desirable detection of more than 90 % of X-class flares and selecting the parameter that gives a lower number of tentative false alarms for the same percentage of detection.

From Table 1.5, it can be emphasised that 92.8% of the SISTED detections are directly correlated with X, M and C-class flares reported by GOES or Optical flares in H-alpha (FLA records). In addition, almost 95% of the 39 X-class flares recorded by GOES in the analysed period were detected using SISTED. Regarding less powerful flares, 65.7% of the M-class flares occurred

Table 1.5: *Validated/Total SISTED detections and the corresponding percentage comparing with GOES X-ray events (XRA) and Optical flares observed in H-alpha (FLA) from the Edited Solar Events Lists. Results are obtained for the training (TRN) dataset considering* $I_1|_{thres} = 0.74$. *Note that these results are obtained for the selected SISTED parameters accepting a percentage of tentative false alarms lower than 7%. Remember that each solar flare may be detected for multiple epochs.*

	Year	SISTEDvsXRA||FLA	GOES XRA		
			X-class	M-class	C-class
	2001	1413/1541	19/21	211/311	399/2101
val./det.	2005	506/527	18/18	61/103	102/578
	TRN	**1919/2068**	**37/39**	272/414	501/2679
percent.%	TRN	**92.8%**	**94.9%**	65.7%	18.7%

Table 1.6: *Validated/Total SISTED detections and the corresponding percentage comparing with GOES X-ray events (XRA) and Optical flares observed in H-alpha (FLA) from the Edited Solar Events Lists. Results are obtained for the training (TRN) dataset considering* $I_1|_{thres} = 0.78$. *Note that these results are obtained for the selected SISTED parameters accepting a percentage of tentative false alarms lower than 4%. Remember that each solar flare may be detected for multiple epochs.*

	Year	SISTEDvsXRA||FLA	GOES XRA		
			X-class	M-class	C-class
	2001	948/989	19/21	167/311	263/2101
val./det.	2005	376/386	17/18	58/103	74/578
	TRN	**1324/1375**	**36/39**	225/414	337/2679
percent.%	TRN	**96.3%**	**92.3%**	54.4%	12.6%

within the test dataset period have been correctly detected, as well as 18.7% of the C-class flares. It is also worth mentioning that a few B-class flares are also detected using SISTED. Their related TEC ionospheric response is very low but still detectable in very specific cases.

From Table 1.6, increasing the $I_1|_{thres}$ to 78%, a much lower number of false detections is given (only 3.7%) but, in contrast, the percentage of X-class flares being detected is reduced (36 out of 39, which corresponds to the 92.3%). Regarding less powerful flares, 54.4% of the M-class flares occurred within the test dataset period have been correctly detected, as well as 12.6% of the C-class flares. In this regard, note that SISTED is not specifically designed for detecting such flares or subflares. The main concern is the correct detection of the highest number of powerful X-class flares affecting Earth's ionosphere.

All in all, both above-mentioned $I_1|_{thres}$ values (74 and 78%) will be considered in order to evaluate the test dataset results.

Test dataset

As mentioned before, all days in years 1999, 2000, 2002, 2003, 2004, 2006 and 2007 have been processed searching for solar flares backwards. In this way, more than a half solar cycle is covered in the test dataset. In addition, the selected thresholds are $\Delta^2 V|_{thres} = 0.00$ TECU, $I_1/I_3|_{thres} = 0.0$ and $I_1|_{thres} = 0.74$ or $I_1|_{thres} = 0.78$ (accepting less than 7% or 4% of tentative false alarms; see Section 1.5.3). Note that results for the training dataset were already given in the previous subsection.

The overall SISTED test results are presented in Table 1.7. In it, information on the number of detections is given as well as the correlation between SISTED and GOES reported events for each of the years being analysed within the test dataset. Such comparison is also included distinguishing between X-class, M-class and C-class solar flares. In this regard, it must be emphasised again that SISTED has been designed to enable the detection of the powerful X-class flares and not the M- and C-class ones.

From Table 1.7, it can be emphasised that 93.4% of the SISTED detections are directly correlated with X, M and C-class flares reported by GOES or Optical flares in H-alpha from the Edited Solar Events Lists (FLA records). In addition, 94.2% of the X-class flares recorded in the analysed period were detected using SISTED. It is remarkable that results are compatible with the values obtained for the training dataset in Table 1.5. Regarding less powerful flares, 62.4% of the M-class flares occurred within the test dataset period have been correctly detected, as well as 16.6% of the C-class flares. In this case, four X-class flares were neither validated with GOES nor by Optical flares.

From Table 1.8, 96.5% of the SISTED detections are validated and 63 out of 69 X-class flares are correctly detected. Regarding less powerful flares, 69% of the M-class flares occurred within the test dataset period have been correctly detected, as well as 14.3% of the C-class flares.

Table 1.7: *Validated/Total SISTED detections and the corresponding percentage comparing with GOES X-ray events (XRA) and Optical flares observed in H-alpha (FLA) from the Edited Solar Events Lists. Results are obtained for the test dataset considering* $\Delta^2 V|_{thres} = 0.74$. *Remember that SISTED results from years 2001 and 2005 were already used as training set to adjust the detector parameters.*

	Year	SISTEDvsXRA||FLA	GOES XRA		
			X-class	M-class	C-class
val./det.	1999	883/982	4/4	115/170	330/1854
	2000	1222/1309	16/17	137/215	426/2262
	2002	970/1032	11/12	129/219	375/2319
	2003	693/742	18/20	91/160	170/1316
	2004	569/590	12/12	78/122	145/913
	2006	111/114	4/4	9/10	24/150
	2007	48/49	0/0	6/10	9/73
	TEST	4496/4818	65/69	565/906	1479/8887
percent.%	TEST	**93.4%**	**94.2%**	62.4%	16.6%

For both $I_1|_{thres}$ options in consideration, we can conclude that the adjusted SISTED parameters enable the detection of most X-class flares in the test period.

Undetected GOES X-class flares

The missed detections for the training and test datasets are listed in Table 1.9. In this case, the SISTED parameter $I_1|_{thres}$ has been set to 74% and thus, the number of undetected X-class flares is 2 out of 39 in the training period and 4 out of 69 in the test period (6 out of 108).

The SISTED results for such periods are depicted in Figure 1.18.

14^{th} July, 2000 analysis

This solar flare occurred on 14^{th} July, 2000 (day of year 196) and its GOES classification is X5.7. This flare preceded the well-known La Bastille storm. In addition, the background X-ray flux was at the level of a C2.9 flare (see the Daily Solar Data file for year 2000, provided by NOAA at http://www.swpc.noaa.gov/ftpdir/indices/).

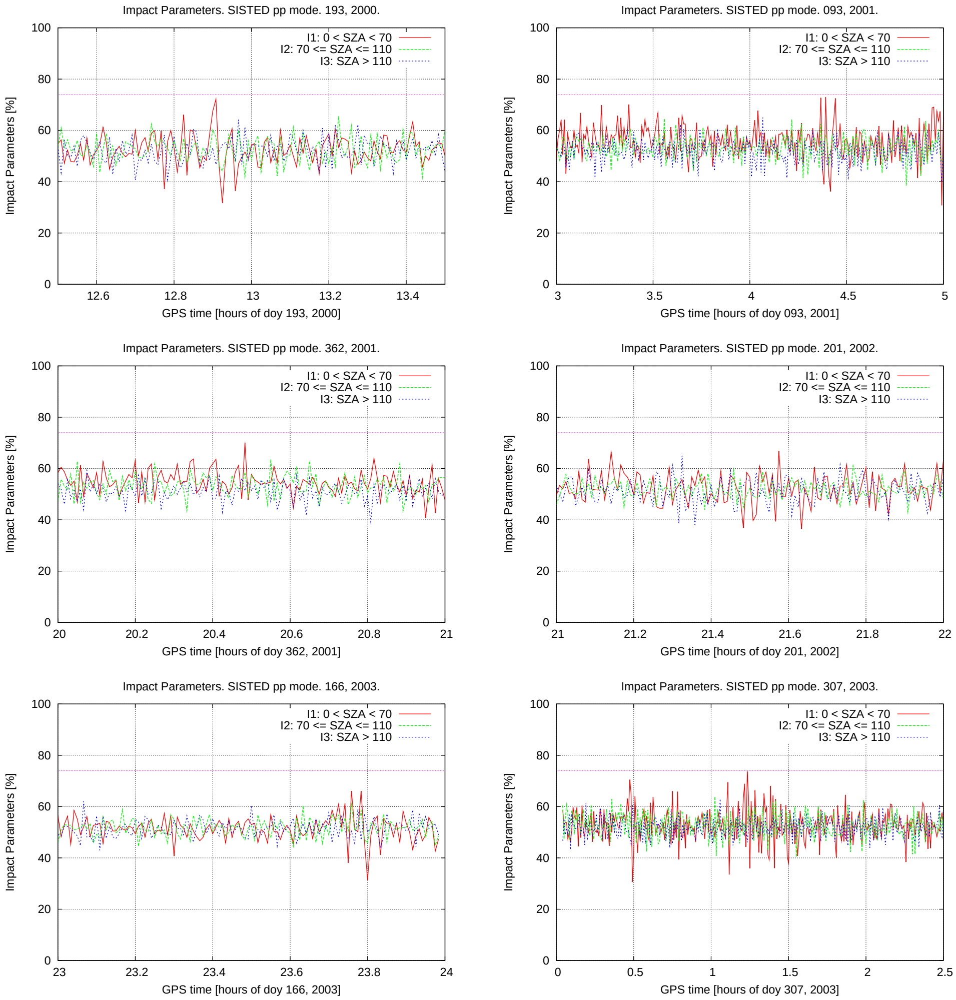
Table 1.8: *Validated/Total SISTED detections and the corresponding percentage comparing with GOES X-ray events (XRA) and Optical flares observed in H-alpha (FLA) from the Edited Solar Events Lists. Results are obtained for the test (TEST) dataset considering $\Delta^2 V|_{thres} = 0.78$. Note that SISTED results from years 2001 and 2005 have been used as training set to adjust the detector parameters.*

	Year	SISTEDvsXRA||FLA	GOES XRA		
			X-class	M-class	C-class
val./det.	1999	596/615	4/4	104/170	227/1854
	2000	831/864	16/17	121/215	281/2262
	2002	660/682	10/12	118/219	251/2319
	2003	494/519	17/20	80/160	114/1316
	2004	423/436	12/12	72/122	108/913
	2006	77/78	4/4	7/10	18/150
	2007	35/35	0/0	6/10	6/73
	TEST	3116/3229	63/69	508/736	1005/7033
percent.%	TEST	**96.5%**	**91.3%**	69%	14.3%

Table 1.9: *List of SISTED undetected GOES X-class flares in the period between 1999 and 2007 (training and test datasets). Begin, Max and End fields are in UTC time.*

YY	MM	DD	Begin	Max	End	Class
00	07	11	1212	1310	1335	X 1.0
01	04	03	0325	0357	0455	X 1.2
01	12	28	2002	2045	2132	X 3.4
02	07	20	2104	2130	2154	X 3.3
03	06	15	2325	2356	0025	X 1.3
03	11	03	0109	0130	0145	X 2.7

When considering $I_1|_{thres} = 0.74$, SISTED triggers alarms for 27 epochs (considering a 30s sampling rate), which correspond to five clearly distinguishable periods (see Table 1.10 and Figure 1.19). Due to the amount of detections between 10h07m30s and 10h32m00s GPS time, a powerful solar flare did most probably occur (as it was the case). In fact, the sunlit ionosphere region Impact Parameter I_1 reaches 100 % of the sunlit rays at 10h25m00s GPS time. In addition, the values of the ionosphere dawn/dusk



Impact Parameters. SISTED pp mode. 193, 2000.

Impact Parameters. SISTED pp mode. 093, 2001.

Impact Parameters. SISTED pp mode. 362, 2001.

Impact Parameters. SISTED pp mode. 201, 2002.

Impact Parameters. SISTED pp mode. 166, 2003.

Impact Parameters. SISTED pp mode. 307, 2003.

Figure 1.18: *SISTED Impact Parameters obtained for the SISTED undetected GOES X-class flares. The values of the sunlit ionosphere region Impact Parameter I_1 are marked in red, the values of the dawn and dusk Impact Parameter I_2 in green and the night-side ones, I_3, in blue.*

region Impact Parameter I_2 are also very high (for instance, I_2 reaches 78.2 % of a total of 230 receiver-satellite rays at 10h27m00s GPS time).

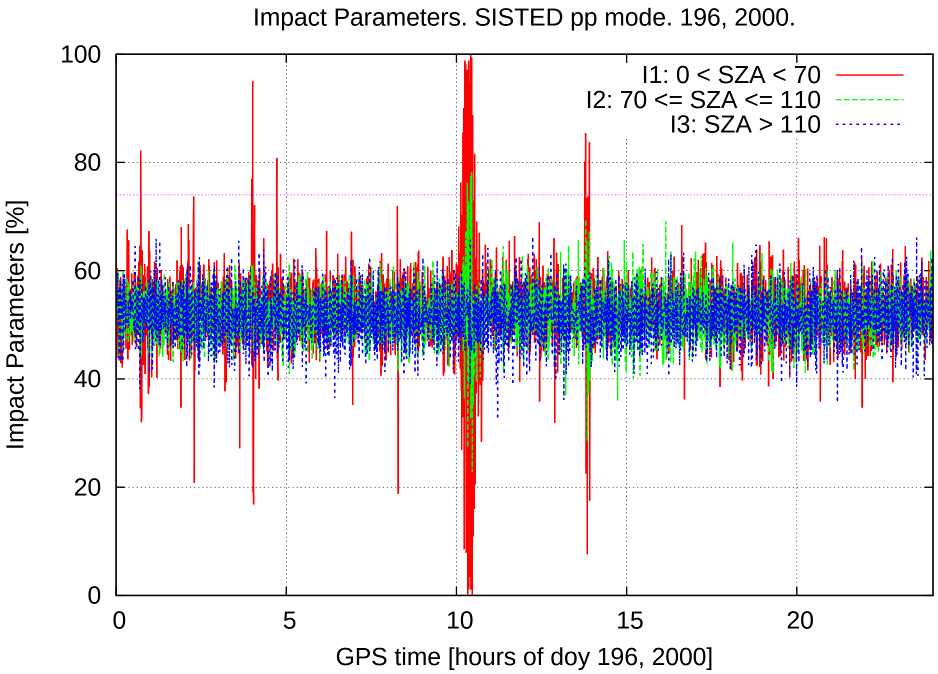


Figure 1.19: *SISTED Impact Parameters obtained on 14th July, 2000 when considering* $I_1|_{thres} = 0.74$. *The values of the sunlit ionosphere region Impact Parameter* I_1 *are marked in red, the values of the dawn and dusk Impact Parameter* I_2 *in green and the night-side ones,* I_3, *in blue.*

Table 1.10: *Simplified SISTED Warning messages on 14th July, 2000. Time is the GPS time expressed in hours of day and HHMMSS corresponds to the two-digit hour, month and second,* I_i *are the Impact Parameters in parts per one for each ionosphere SZA region* r_i. *See Table B.3 for format details.*

Type	YY	Doy	Time	I_1	I_2	I_3	HHMMSS
SF_WARN	00	196	0.7250000000	0.822	0.545	0.470	004330
SF_WARN	00	196	3.9833333333	0.770	0.506	0.543	035900
SF_WARN	00	196	4.0166666667	0.950	0.611	0.569	040100
SF_WARN	00	196	4.7250000000	0.808	0.509	0.514	044330
SF_WARN	00	196	10.1250000000	0.755	0.519	0.453	100730
SF_WARN	00	196	10.1333333333	0.768	0.596	0.516	100800
SF_WARN	00	196	10.1916666667	0.857	0.614	0.573	101130
SF_WARN	00	196	10.2083333333	0.891	0.626	0.486	101230
SF_WARN	00	196	10.2166666667	0.897	0.606	0.548	101300

Table 1.10: *Simplified SISTED Warning messages on* 14^{th} *July, 2000. Time is the GPS time expressed in hours of day and HHMMSS corresponds to the two-digit hour, month and second,* I_i *are the Impact Parameters in parts per one for each ionosphere SZA region* r_i*. See Table B.3 for format details.*

Type	YY	Doy	Time	I_1	I_2	I_3	HHMMSS
SF_WARN	00	196	10.2416666667	0.822	0.594	0.539	101430
SF_WARN	00	196	10.2500000000	0.988	0.609	0.452	101500
SF_WARN	00	196	10.2750000000	0.982	0.673	0.530	101630
SF_WARN	00	196	10.2916666667	0.818	0.570	0.549	101730
SF_WARN	00	196	10.3166666667	0.971	0.762	0.525	101900
SF_WARN	00	196	10.3416666667	0.916	0.579	0.539	102030
SF_WARN	00	196	10.3583333333	0.988	0.728	0.460	102130
SF_WARN	00	196	10.3750000000	0.946	0.618	0.482	102230
SF_WARN	00	196	10.4000000000	0.869	0.618	0.389	102400
SF_WARN	00	196	10.4083333333	0.976	0.737	0.657	102430
SF_WARN	00	196	10.4166666667	1.000	0.775	0.508	102500
SF_WARN	00	196	10.4500000000	0.994	0.782	0.452	102700
SF_WARN	00	196	10.4750000000	0.884	0.364	0.508	102830
SF_WARN	00	196	10.5333333333	0.812	0.542	0.490	103200
SF_WARN	00	196	13.7666666667	0.801	0.615	0.605	134600
SF_WARN	00	196	13.7916666667	0.851	0.630	0.496	134730
SF_WARN	00	196	13.8000000000	0.850	0.695	0.580	134800
SF_WARN	00	196	13.9083333333	0.838	0.673	0.506	135430

In Tables 1.11 and 1.12, GOES X, M and C-class flares as well as the optical flares in H-alpha are shown for 14^{th} July, 2000 (reported in the Edited Solar Events Lists). Comparing the obtained results, the SISTED detections correspond to the most powerful solar flares being recorded for such day, which are an X5.7 and an M3.7 flares.

It is worth mentioning that SEM-EUV data were not available for this period.

Regarding Table 1.13, SISTED tentative false alarms may be correlated with the optical flare event numbered 840 in Table 1.12. The start time of such event was not properly determined (B0430 means before 0430 UT) and the quality of the observing conditions was rather poor (Q=2).

In order to verify that the flare is not a false alarm, the dependency with the SZA is shown in Figure 1.20. It can be seen that there is a significant linear trend following the expected cosine function with its maximum at the subsolar point location. In addition, plots on L_I and $\Delta^2 L_I$ show clearly the

Table 1.11: *List of GOES X, M and C-class flares on 14th July, 2000 (from NOAA's Edited Solar Events Lists). Begin, Max and End fields are in UTC time.*

YY	MM	DD	Begin	Max	End	Class
00	07	14	00:20	00:23	00:26	C 4.9
00	07	14	00:39	00:45	00:50	M 1.5
00	07	14	02:14	02:17	02:20	C 6.1
00	07	14	06:52	06:57	07:05	C 7.1
00	07	14	09:51	09:55	09:59	C 5.9
00	**07**	**14**	**10:03**	**10:24**	**10:43**	**X 5.7**
00	07	14	13:44	13:52	14:00	M 3.7

Table 1.12: *List of optical flares in H-alpha on 14th July, 2000 (from NOAA's Edited Solar Events Lists) coinciding with SISTED detections not validated with GOES events. LEA corresponds to Learmonth Observatory in Australia and G08 corresponds to GOES-08 spacecraft records. Note that the reported flare is considered as a Subflare (area <= 2.0 square degrees) with faint brightness and several eruptive centres. Begin, Max and End fields are in UTC time. Obs is the reporting observatory, Q the quality and Loc the location. For further details on the formatting, refer to http://www.swpc.noaa.gov/ftpdir/indices/events/README.*

YY	MM	DD	Begin	Max	End	Obs	Q	Type	Loc
00	07	14	0430	0447	0520	LEA	2	FLA	N20W02

Table 1.13: *List of SISTED tentative false alarms on 14th July, 2000.*

Type	YY	Doy	Time	I_1	I_2	I_3	HHMMSS
SF_WARN	00	196	3.9833333333	0.770	0.506	0.543	035900
SF_WARN	00	196	4.0166666667	0.950	0.611	0.569	040100

simultaneous VTEC increase and the typical over-ionization pattern (not included). Therefore, this detection, as an example, is most probably not a false alarm.

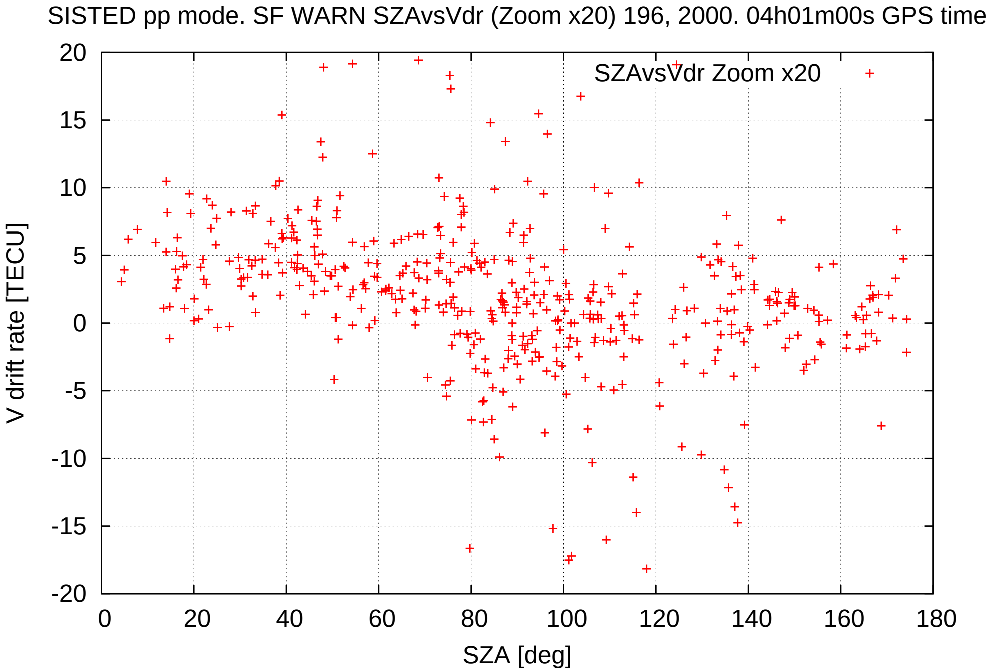


Figure 1.20: *Dependence of the VTEC drift rate ($\Delta^2 V$) on the Solar-Zenith angle (SZA, at the IPP locations) for a SISTED detection on 14th July, 2000 (zoom at 04h01m00s GPS time). Note that a few measurements with SZA between 80 and 140 degrees are affected by local perturbations like scintillation and thus, the corresponding $\Delta^2 V$ values are not close to zero.*

28^{th} October, 2003 analysis

One of the largest Solar Flare events detected on GNSS measurements, from the point of view of TEC increase, was recorded preceding the so called Halloween storm (on day 28^{th} October, 2003). The GOES classification of this solar flare was X17.2. In addition, the background X-ray flux was at the level of a C3.2 flare (see the Daily Solar Data file for 2003 provided by NOAA at http://www.swpc.noaa.gov/ftpdir/indices/). In addition, the VTEC enhancements reached up to more than 20 TECUs.

Table 1.14: *Simplified SISTED Warning messages on 28^{th} October, 2003. Time is the GPS time expressed in hours of day and HHMMSS corresponds to the two-digit hour, month and second, I_i are the Impact Parameters in parts per one for each ionosphere SZA region r_i. See Table B.3 for format details.*

Type	**YY**	**Doy**	**Time**	I_1	I_2	I_3	**HHMMSS**
SF_WARN	03	301	0.0916666667	0.746	0.570	0.481	000530
SF_WARN	03	301	0.7333333333	0.848	0.622	0.514	004400
SF_WARN	03	301	0.9833333333	0.904	0.591	0.531	005900
SF_WARN	03	301	1.5000000000	0.775	0.527	0.541	013000
SF_WARN	03	301	1.5083333333	0.836	0.660	0.488	013030
SF_WARN	03	301	1.5500000000	0.781	0.581	0.477	013300
SF_WARN	03	301	5.1750000000	0.836	0.716	0.524	051030
SF_WARN	03	301	7.9166666667	0.846	0.576	0.521	075500
SF_WARN	03	301	8.4500000000	0.859	0.589	0.480	082700
SF_WARN	03	301	9.9083333333	0.953	0.704	0.531	095430
SF_WARN	03	301	10.3333333333	0.927	0.695	0.487	102000
SF_WARN	03	301	10.3500000000	0.766	0.604	0.467	102100
SF_WARN	03	301	10.3833333333	0.856	0.563	0.530	102300
SF_WARN	03	301	10.6166666667	0.833	0.619	0.497	103700
SF_WARN	03	301	10.9916666667	0.791	0.628	0.532	105930
SF_WARN	03	301	11.0083333333	0.815	0.611	0.534	110030
SF_WARN	03	301	11.0166666667	0.963	0.660	0.550	110100
SF_WARN	03	301	11.0250000000	1.000	0.767	0.502	110130
SF_WARN	03	301	11.0333333333	0.994	0.854	0.447	110200
SF_WARN	03	301	11.0416666667	1.000	0.842	0.560	110230
SF_WARN	03	301	11.0500000000	0.993	0.894	0.530	110300
SF_WARN	03	301	11.0750000000	1.000	0.835	0.527	110430
SF_WARN	03	301	11.0916666667	0.857	0.358	0.485	110530
SF_WARN	03	301	11.1166666667	1.000	0.781	0.502	110700
SF_WARN	03	301	11.1333333333	0.969	0.606	0.490	110800

Table 1.14: *Simplified SISTED Warning messages on 28th October, 2003. Time is the GPS time expressed in hours of day and HHMMSS corresponds to the two-digit hour, month and second,* I_i *are the Impact Parameters in parts per one for each ionosphere SZA region* r_i*. See Table B.3 for format details.*

Type	YY	Doy	Time	I_1	I_2	I_3	HHMMSS
SF_WARN	03	301	11.1583333333	0.794	0.453	0.482	110930
SF_WARN	03	301	11.1916666667	0.969	0.652	0.551	111130
SF_WARN	03	301	11.2166666667	0.888	0.510	0.537	111300
SF_WARN	03	301	11.2250000000	0.888	0.661	0.507	111330
SF_WARN	03	301	11.2500000000	0.808	0.605	0.509	111500
SF_WARN	03	301	11.3666666667	0.898	0.637	0.536	112200
SF_WARN	03	301	11.5500000000	0.750	0.533	0.513	113300
SF_WARN	03	301	11.5583333333	0.747	0.558	0.546	113330
SF_WARN	03	301	12.4583333333	0.771	0.600	0.482	122730
SF_WARN	03	301	14.9500000000	0.848	0.596	0.494	145700
SF_WARN	03	301	14.9750000000	0.744	0.566	0.464	145830
SF_WARN	03	301	16.2166666667	0.830	0.527	0.510	161300
SF_WARN	03	301	16.2333333333	0.754	0.543	0.465	161400
SF_WARN	03	301	16.2500000000	0.856	0.524	0.494	161500
SF_WARN	03	301	16.2583333333	0.812	0.654	0.494	161530
SF_WARN	03	301	16.4833333333	0.777	0.557	0.494	162900
SF_WARN	03	301	20.6416666667	0.786	0.569	0.505	203830

Considering the adjustable parameters fixed to $\Delta^2 V|_{thres} = 0.0$, $I_1|_{thres} = 0.74$ and $I_1/I_3|_{thres} = 0.0$, the solar flare detector triggers alarms for 42 epochs in post-processing mode, which correspond to multiple clearly distinguishable periods. These detections have been validated with GOES X- and C-class flares as well as with several optical flares observed in H-alpha (see Tables 1.15 and 1.16). It must be remarked that GOES provides a very wide start-end interval of the X17.2 flare and that the instrument became saturated ([*Tsurutani et al.* (2009)]).

Due to the number of detections between 11h01m00s and 11h22m00s GPS time, a powerful solar flare most probably occurred (as it was the case). Note that the percentage of sunlit receiver-satellite rays detecting a simultaneous TEC increase over the threshold, I_1, reaches 100 % at several epochs (11h01m30s, 11h02m30s, 11h04m30s and 11h07m00s GPS time). It is also remarkable that the values of the Impact Parameter at the dawn/dusk region I_2 are also very high (for instance, it reaches 89.4 % at 11h03m00s GPS time).

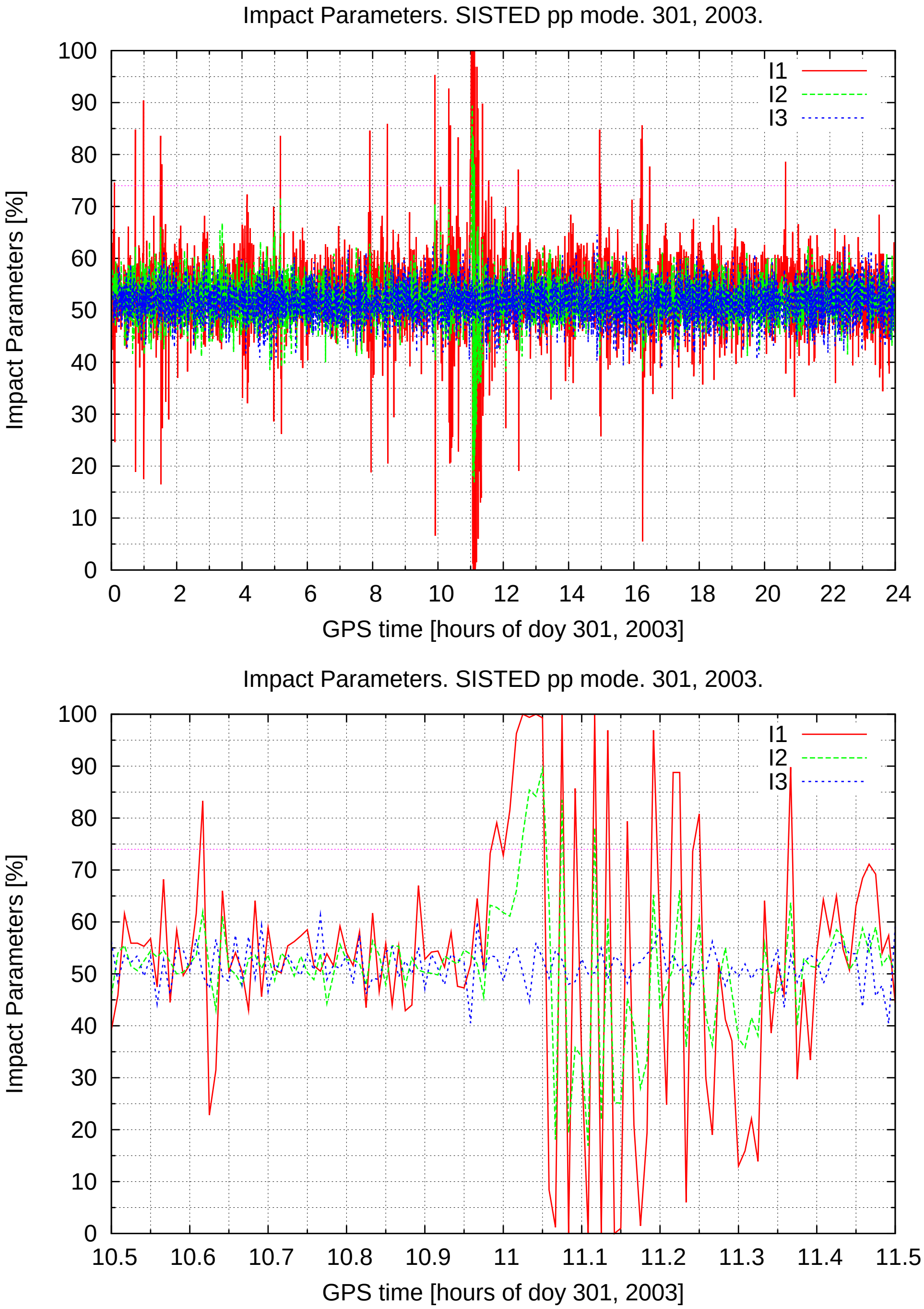


Figure 1.21: *SISTED Impact Parameters obtained on 28th October, 2003. The values of the sunlit ionosphere region Impact Parameter* I_1 *are marked in red, the values of the dawn and dusk Impact Parameter* I_2 *in green and the night-side ones,* I_3*, in blue.*

Comparing with the results from GOES and the Edited Solar Events Lists (by NOAA), the SISTED detections between 09h54m30s and 11h22m00s GPS time would correspond to the most powerful solar flare being recorded for such day. SISTED was also able to detect less powerful class flares such as the C6.7 that occurred at around 01h00m00s UTC time.

Table 1.15: *List of GOES X, M and C-class flares on* 28^{th} *October, 2003 (from NOAA's Edited Solar Events Lists). Begin, Max and End fields are in UTC time.*

YY	MM	DD	Begin	Max	End	Class
03	10	28	00:41	00:45	00:48	C 5.3
03	10	28	00:56	00:59	01:02	C 6.7
03	10	28	01:27	01:33	01:45	C 7.5
03	10	28	05:07	05:11	05:14	C 7.7
03	10	28	08:35	08:39	08:44	C 8.7
03	**10**	**28**	**09:51**	**11:10**	**11:24**	**X 17.2**
03	10	27	2345	23:52	00:09	C 6.8

Table 1.16: *List of optical flares in H-alpha on* 28^{th} *October, 2003 (from NOAA's Edited Solar Events Lists) coinciding with SISTED detections not validated with GOES events. LEA corresponds to Learmonth Observatory in Australia and G08 corresponds to GOES-08 spacecraft records. Note that the reported flare is considered as a Subflare (area* $\leq=$ 2.0 *square degrees) with faint brightness and with several eruptive centres. Begin, Max and End fields are in UTC time. Obs is the reporting observatory, Q the quality and Loc the location. For further details on the formatting, refer to http://www.swpc.noaa.gov/ftpdir/indices/events/README.*

YY	MM	DD	Begin	Max	End	Obs	Q	Type	Loc
00	07	14	04:30	04:47	05:20	LEA	2	FLA	N20W02

In this case, SEM-EUV records were available and its flux rate evolution is depicted in Figure 1.22. The X17.2 flare is clearly visible as well as their related subsequent perturbations. Nonetheless, no SISTED false alarms were reported in spite of the arrival of relativistic electrons, which seriously affected SEM-EUV records.

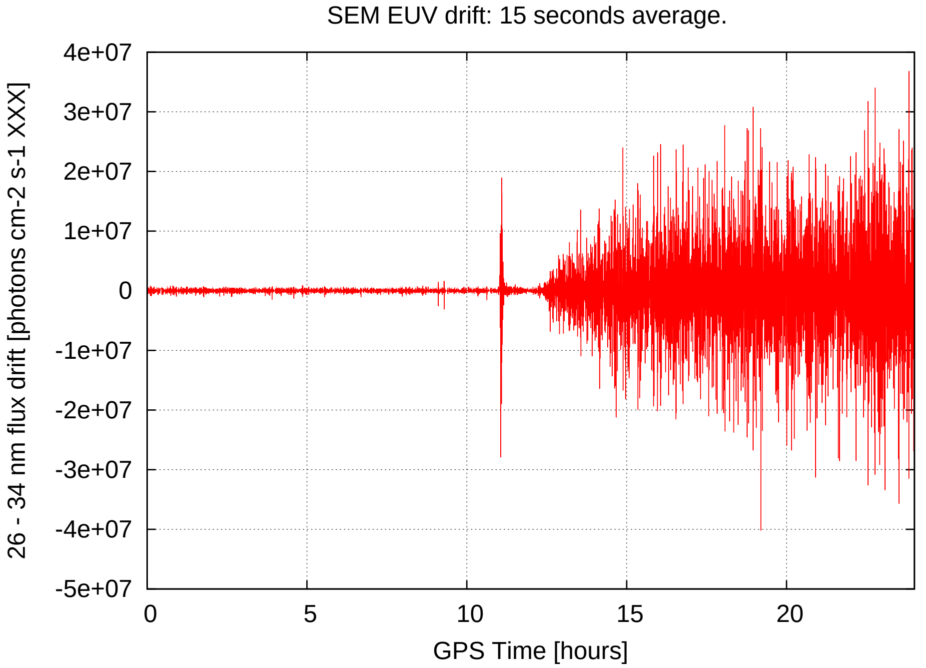


Figure 1.22: *15 seconds average records of the SEM-EUV 26 - 34 nm flux on* 28th *October, 2003. Note that GOES X17.2 flare was reported with its maximum peak at 1110 UTC time.*

The number of rays being processed is of about 600 at any epoch throughout the day. Plotting a zoom of the number of rays for the time period close to the X17.2 flare is interesting. This is because a few receivers lost signal lock for certain satellites. These disruptions were only affecting the receivers in the sunlit area and thus, may probably be associated with the increase of radiation related to the X17.2 flare.

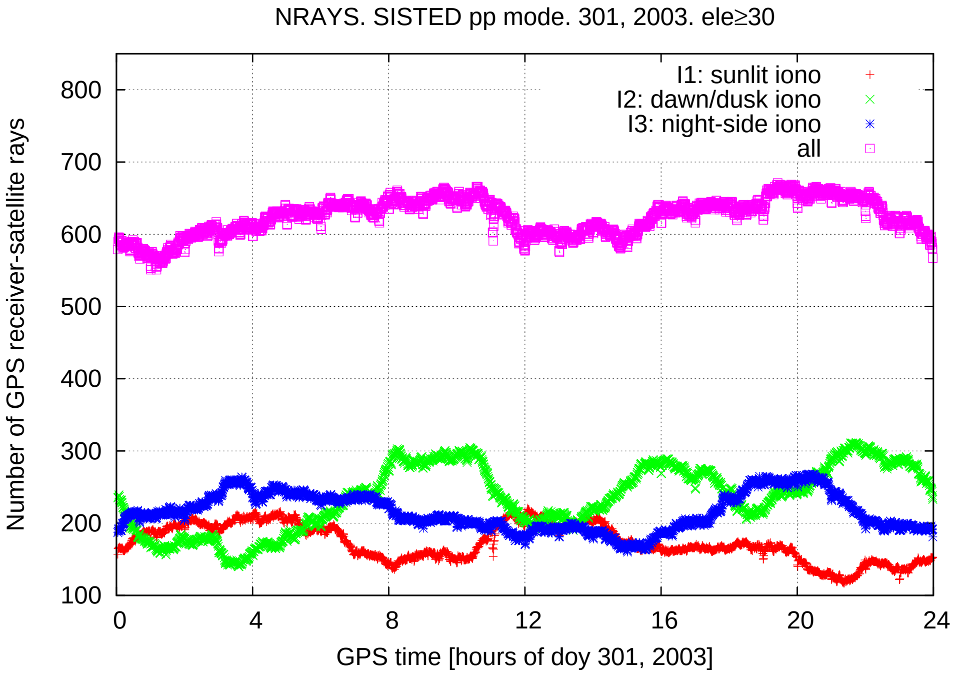


Figure 1.23: *SISTED number of rays obtained on 28th October, 2003. The rays in the sunlit ionosphere region at the IPP location are marked in red, the rays in the dawn and dusk region in green and the rays in the night-side region in blue.*

1.6 Conclusions

In this chapter, the solar flare monitoring method has been presented showing its good performance to detect the most powerful X-class flares. One can select SISTED threshold of solar flare detectability as a compromise between the percentage of detected X-class flares and the number of tentative false alarms. In any case, the number of tentative false alarms is a pessimistic threshold. In fact, tentative alarms for certain days have been studied and discarded.

Detecting solar flares is a must from the technological and scientific points of view in the years to come, since we are approaching a solar cycle maximum. In this way, SISTED can be a very useful tool to prevent the harmful consequences of the solar flares' related particles. In addition, GNSS users could be informed on the ionosphere behaviour and mitigation techniques could be applied to avoid the solar flares' direct and indirect effects on positioning accuracy (that can be especially important in case of potential ionospheric storm warning).

In this sense, a service based on SISTED could be deployed in the near future. Its goal would be to warn GNSS users of powerful solar flares facing the Earth that produce significant perturbations in the sunlit ionosphere region. This would be especially useful for GNSS users with integrity requirements, such as in Civil Aviation.

Chapter 2

GNSS Prediction of ionospheric Total Electron Content

2.1 Introduction

In this chapter, a new approach used to predict global VTEC maps from GPS data is presented for a very demanding operational scenario. Predicting TEC is possible thanks to the GPS capability for ionospheric sounding when using, at least, two frequencies. In fact, GPS enabled having an unprecedented combination of accuracy, temporal and spatial resolution, and availability (low cost) that led to important progress in fields such as ionospheric tomography (see review in [*Bust and Mitchell* (2008)]). This was useful to enhance precise GNSS navigation ([*Hernández-Pajares et al.* (2000)]) and generate reliable global VTEC maps in a continuous operational way in the context of the International GNSS Service (IGS; [*Dow et al.* (2009)]) Ionosphere Working Group (IGS Iono-WG; see [*Feltens* (2003)], [*Feltens and Schaer* (1998)] and [*Hernández-Pajares et al.* (2009)]).

Regarding the structure of this chapter, the Motivation and State-of-the-Art of VTEC prediction are included at first (Sections 2.2 and 2.3, respectively). Secondly, in Section 2.4, the UPC prediction approach is explained in more detail. Afterwards, the Section 2.5 includes how the performance analysis is carried out. And finally, before the conclusions, the results for two days ahead forecast are presented for three periods (in 2004, 2006 and 2011) and validated against UPC Final IGS products (designated as $UPCG$) and JASON external data (Sections from 2.5.5). Finally, before the conclusions, a first performance comparison is also included between UPC and the other preliminary IGS forecast products. Note that part of the contents of this chapter were published in a recent peer-reviewed article ([*García-Rigo et al.* (2011)]).

2.2 Motivation

Developing short-term ionospheric predictions can be of great importance for certain scientific and technological applications. For instance, the European Space Agency (ESA) indicated great interest for such predictions for the Soil Moisture and Ocean Salinity (SMOS; [*Sivestrin et al.* (2001)]) mission. Ionospheric products provided by the IGS Iono-WG were required, including global VTEC predictions for two days ahead ([*Krankowski et al.* (2008)]).

Apart from that, predictions are being used by the authors as background model to generate preliminary real-time global VTEC maps (see details in [*Orús et al.* (2010a)]). Among many other applications, this could be used to achieve sub-metre accuracies for mass-market single-frequency receivers. This would improve the quality of current GNSS applications, such as in automobiles, road mapping, location-based advertising, personal navigation or logistics.

2.3 State-of-the-Art of TEC prediction

In the past and recent years, methods for ionospheric forecasting have been developed to predict ionospheric parameters such as the critical frequency of the F2 layer (f_0F2), the maximum usable frequency at a distance of $3000km$ of the F2-layer ($M(3000)F2$) or the TEC itself. These methods are based on neural networks ([*Cander et al.* (1998)], [*Francis et al.* (2000)], [*Tulunay et al.* (2006)]), auto-correlation and auto-covariance procedures ([*Muhtarov et al.* (1999)], [*Dick et al.* (1999)], [*Stanislawska et al.* (2001)]), linear regression ([*Muhtarov et al.* (2001)], [*Krankowski et al.* (2005)]), among others.

In terms of global VTEC forecasting using GNSS, the first IGS Associate Analysis Center (IAAC) to generate its own ionospheric predicted product was the Center for Orbit Determination in Europe (CODE). Their model, which is based on the extrapolation of the Spherical Harmonic coefficients by collocation, is explained in detail in [*Schaer et al.* (1999)]. Their one and two days ahead VTEC maps (named $C1PG$ and $C2PG$, respectively) have been released for public access for years via the FTP server of the Astronomical Institute of the University of Bern (AIUB; http://www.aiub.unibe.ch). More recently, the ESA's Space Operations Centre (ESOC) and the Technical University of Catalonia (UPC) have created their own two days ahead forecast product in the frame of the IGS Iono-WG (named $E2PG$ and $U2PG$, respectively) to fulfill the requests of the SMOS mission. These products have been released for public access through the FTP site of the Crustal Dy-

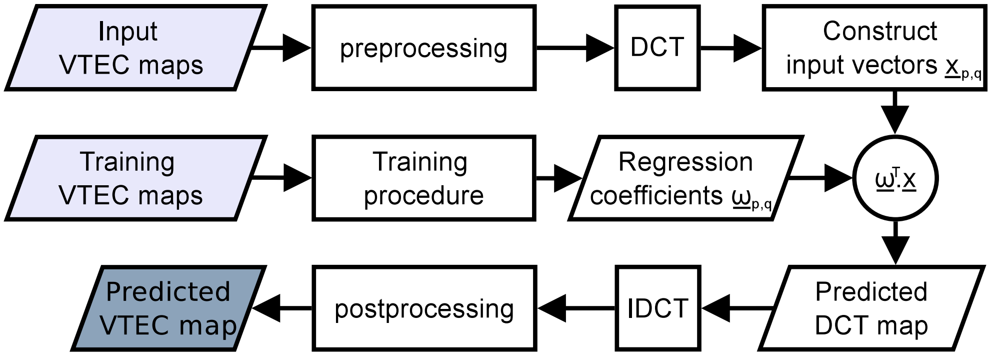


Figure 2.1: *Diagram showing the main steps of the UPC prediction approach. DCT stands for Discrete Cosine Transform, IDCT for Inverse DCT and superindex T denotes transpose.*

namics Data Information System (CDDIS; ftp://cddis.gsfc.nasa.gov/) since September, 2009. Furthermore, preliminary combined IGS Predicted products for one day and two days ahead have also been generated by ESA-ESOC since the end of December, 2009. These products are produced by combining up the three individual preliminary VTEC predicted products and are designated as *I1PG* and *I2PG* for one day and two days ahead forecast, respectively.

2.4 Implemented approach

A model for short-term VTEC predictions was developed by the author to generate a two days ahead forecast product in the frame of the IGS Iono-WG and to fulfill the demanding SMOS mission requirements. In brief, the forecast model is based on applying linear regression to a temporal window of VTEC maps in the Discrete Cosine Transform domain (for a summary of the DCT, see [*Ahmed et al.* (1974)] or [*Oppenheim et al.* (2010)]; for the justification of the selected prediction method, see Section 2.4.2). Figure 2.1 shows a simplified diagram with the main steps that are required to compute the UPC Predicted product.

The UPC Final/Rapid IGS products are the input data for the prediction model. These products are encapsulated in IONosphere map EXchange format (IONEX; see [*Schaer et al.* (1998)]) and contain 13 global VTEC maps at a temporal resolution of two hours, from 00UT of the corresponding day to 00UT of the following day. Each VTEC map is a two-dimensional map with spatial grid points every $2.5^o/5^o$ in the latitude/longitude range assuming a thin shell layer model of the ionosphere at a height of 450 *km*. Thus, the map contains $MxN = 5112$ independent VTEC samples, where $M = 71$, the number of grid points in the latitudinal/vertical direction (the number of rows), and $N = 72$, the number of grid points in the local-time/horizontal direction (the number of columns). From now on, the VTEC value at row m and column n will be referred to as $V_{m,n}$. Since the DCT is used as an image processing tool, the standard nomenclature in image processing has been followed. Therefore, the ordinate direction is also denoted as vertical direction, which corresponds to the latitude direction in the IONEX data and after converting to a Sun reference frame. In the same way, the abscissa direction is denoted as horizontal direction, which corresponds to the longitude in the IONEX data and the local-time after the conversion to a Sun reference frame.

Using the DCT is mainly justified by two arguments. The first, as mentioned before, is that the DCT is the standard transform used in image coding systems such as JPEG. Since the implemented approach is inspired in image processing methods, it was decided to use it as default. Another argument is related to the assumptions of the problem. We have rectangular images, which originate from spherical maps. Therefore, there has to be continuity at the boundaries. Both the DCT and the Discrete Sinus transform (DST) assume a periodical signal, which is compatible with the fact that the problem has circular symmetry. Nevertheless the DCT assumes that the signals have an even symmetry, while the DST assumes that they have an odd symmetry. The fact that the sphere is cut in order to generate the maps means that the boundaries of the map should be continuous and, therefore, a change in sign would be erroneous. The DST will impose a change in sign at the boundaries due to the odd symmetry, while the DCT maintains the sign. Then, the use of the DCT is compatible with the physical model.

In daily computations, the last seven consecutive UPC Final IGS products ($UPCG$) are used as the input dataset for the model (prediction window; 1^{st} row in Figure 2.1). If any of these products is not available or was not generated, the corresponding UPC Rapid IGS product ($UPRG$) is retrieved instead. Note that UPC Rapid IGS products are generated with a one-day latency while the UPC Final IGS products are internally generated with a five- to six-day latency. This is an acceptable procedure because UPC Rapid IGS VTEC maps are 5 - 7 % worse than the final ones, as stated in [*Hernández-Pajares et al.* (2009)]. For the two-days ahead prediction on day T, the input dataset includes the UPC products from day $T-8$ to $T-2$ (being $T-1$ the day before day T). Note that the two days ahead prediction for day T is calculated on day $T-1$ with data until day $T-2$.

Regardless, because 00UT VTEC maps are calculated twice for adjacent products, a total of $U = 85$ input VTEC maps are considered (one map

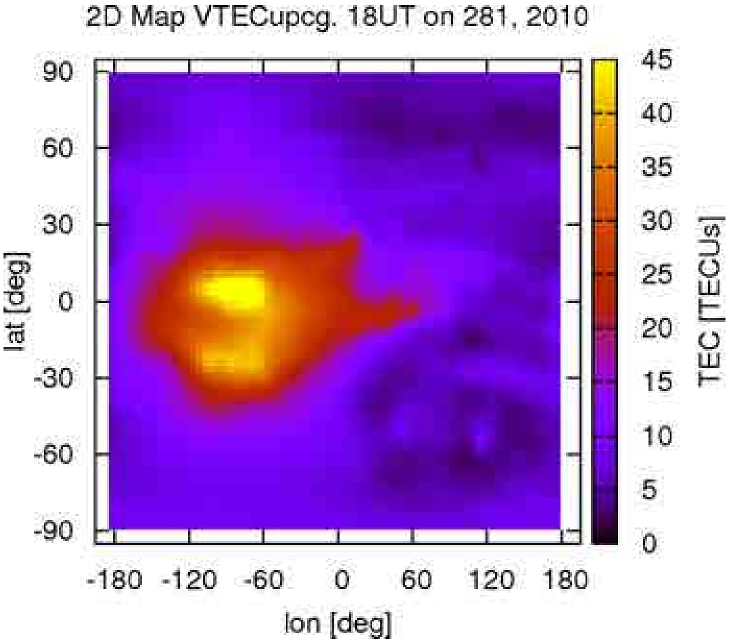


Figure 2.2: *Representation of the UPC Final VTEC map in geomagnetic longitude/latitude range at 18UT on day 281, 2010.*

every two hours for seven days, taking into account the fact that 00UT maps are not accounted for twice; $U = 12 * 7 + 1$ is the length of the prediction window).

The preprocessing step in Figure 2.1 includes the transformation of the input data to a sun-fixed reference frame (local-time and latitude, where the electron content distribution is much more stationary). Afterwards, each VTEC map is represented by a set of parameters of a generative model in the spatial frequency domain, in this case the two-dimensional DCT. The basis functions of the DCT represent the latitude, longitude and crossed latitude/longitude spatial frequency components of the original VTEC map. In other words, the DCT provides information on the vertical, horizontal and diagonal patterns of the input VTEC grid image. The output values of the DCT of a given map are referred to as DCT coefficients, following the notation $C_{p,q}$, where p and q correspond to the vertical and horizontal discrete frequencies, respectively. Thus, the DCT coefficient $C_{0,0}$ is related to the continuous component (i.e., the mean VTEC value of the map), and the DCT coefficients correspond to higher frequency components as the p and q indices increase.

Note that the DCT allows basic physical properties of the VTEC to be modeled. Specifically, this includes the smooth spatial distribution of the

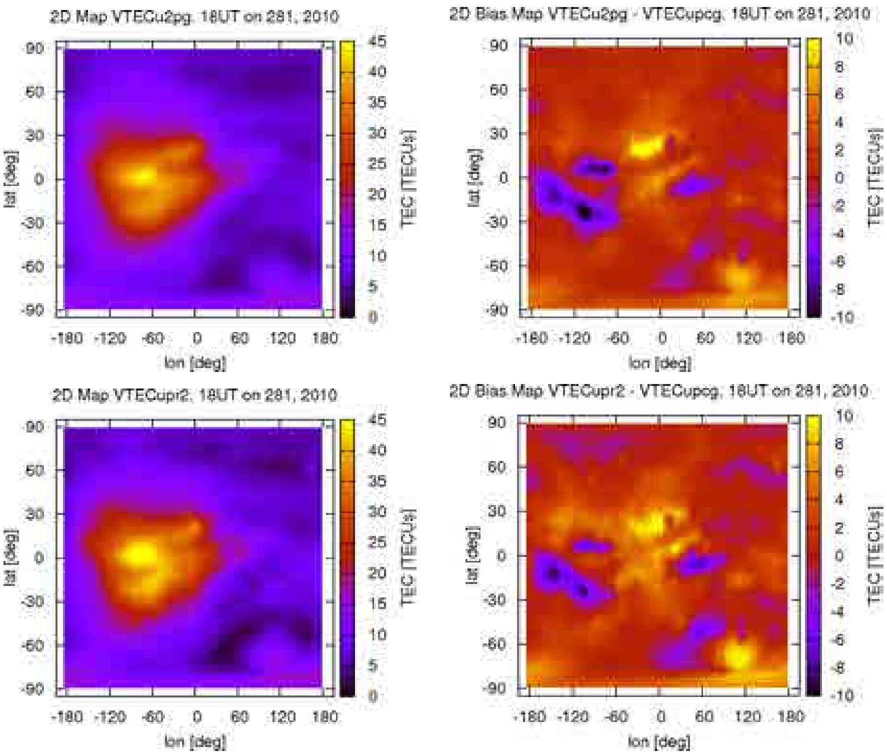


Figure 2.3: *Representation of the UPC Predicted/Time-invariant VTEC map in geomagnetic longitude/latitude range at 18UT on day 281, 2010 (upper/bottom left), and the differences with respect to the corresponding UPC Final VTEC map (upper/bottom right). Note that the UPC Final VTEC map for this day is depicted in Figure 2.2. The same range of colors are used but for different scales: 0 to 45 TECUs for the left-hand plots and -10 to 10 TECUs for the right-hand plots.*

VTEC (i.e., the fact that the absolute value of high frequency components is smaller than that of low frequency components) and the fact that energy is more concentrated along each frequency axis. Therefore, the original VTEC map in the spatial domain can be generated with a subset of the DCT coefficients preserving the main trends of its spatial distribution. In this work, a triangular subset considering 2556 DCT basis functions was selected to reduce the noise produced by high frequency coefficients and to smooth the spatial variation of the VTEC. Thus, the index of the vertical discrete frequencies p ranges between 0 and $P-1=70$, and the index of the horizontal discrete frequencies q ranges between 0 and $Q-p=71-p$. In Equation 2.2, the mathematical formula of computing the DCT coefficients of a given map is shown. The map would be represented by the VTEC values $V_{m,n}$.

$$C_{p,q} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} V_{m,n} \cos\frac{\pi(2m+1)p}{2M} \cos\frac{\pi(2n+1)q}{2N}, \quad 0 \leq p \leq P-1, 0 \leq q \leq Q-p \tag{2.1}$$

It is important to note that the future values of each DCT coefficient, rather than the spatial VTEC values, are predicted. The predicted DCT coefficients can then be used to reconstruct the predicted VTEC map by applying the Inverse Discrete Cosine Transform (IDCT; see diagram in Figure 2.1) and a post-processing, where the change in the kind of coordinates is applied. This approach is more robust, as the evolution of the VTEC for a specific grid point may not only be correlated with its previous values, but also may be highly correlated with the time and spatial evolution of the physical processes affecting the whole VTEC map. In fact, the transformation implies that the spatial information of the whole map/image is implicit in each DCT coefficient.

The relationship between past and future values of each DCT coefficient is obtained by linear regression. The input vectors $\underline{x}_{p,q}$ in the diagram in Figure 2.1 are built with the last $U = 85$ values of each DCT coefficient and include a constant term to model the offset. Additional terms, for potential further improvement in future research, could be included with information such as the 10.7 cm Solar Flux, the Sunspot number, the Kp geomagnetic index or the ionospheric parameter Global Electron Content (GEC; [*Afraimovich et al.* (2006)]). On the other hand, the corresponding linear regression coefficients are computed via a training process and are named $\underline{w}_{p,q}$. The model yields a high number of parameters and the inputs could be collinear and, therefore, the problem might be ill-conditioned. As a result, ridge regression is used, which is a modification of the linear regression where a regularization term $\lambda \cdot \underline{\underline{I}}$ has been added to the covariance matrix

([*Hastie et al.* (2001)]), where λ is a scalar below one and $\underline{\underline{I}}$ is the identity matrix.

2.4.1 Training and application of the model

The following training process is executed each day (2^{nd} row in Figure 2.1) taking approximately 45 minutes of computer processing in a standard PC under Linux OS. First, the training data obtained from 366 days of UPC Final IGS products are transformed to a sun-fixed reference frame. For the prediction on day T, the training dataset includes the UPC IGS products from day $T - 9 - 365$ to $T - 9$. Secondly, all VTEC maps are converted into the DCT frequency domain and one-dimensional one-year time series are generated for each DCT coefficient. A sliding window moving one VTEC map per step determines the number of training subsets of length $U = 85$ VTEC maps. Afterwards, the coefficients $\underline{w}_{p,q}$ are computed as shown in Algorithm 1. These regression coefficients model the relationship between the training subsets and their corresponding reference values $\underline{r}_{p,q}$ (known "predictions"). These reference values are obtained from the VTEC map that is Δt_{maps} maps ahead from the last input VTEC map. Thus, t_{maps} and Δt_{maps} are integers that refer to a VTEC map or snapshot. In order to make the notation clearer and distinguish the indices from the time, which is a continuous variable, the subindex *maps*, which referes to VTEC maps, will be used.

Note that the linear regression coefficients are calculated independently for each DCT coefficient and for each of the VTEC maps to be predicted. The training process has to be configured in 13 independent ways because the predicted product includes 13 VTEC maps as it is encapsulated in IONEX format. Thus, the regression coefficients are calculated considering that the distance Δt_{maps} between the last VTEC map in each training subset and the reference can be 12 to 24 maps ahead. Note that for the two days ahead forecast on day T, the prediction model is executed using input products until day $T - 2$ (see the third paragraph in this Section 2.4). Nevertheless, the last input VTEC map is the one at 00 UT on day $T - 1$. Therefore, the prediction for 00 UT on day T is 12 maps ahead and the prediction for 00 UT on day $T + 1$ is 24 maps ahead.

Finally, each predicted value is computed by performing a dot product between the coefficients $\underline{w}_{p,q}$ and the sequence of DCT coefficients (input vector $\underline{x}_{p,q}$), as can be seen from Equation 2.3 and in the diagram in Figure 2.1.

Algorithm 1 Pseudocode used to compute the regression coefficients $\underline{w}_{p,q}$

```
for $i \leftarrow 1, nTrnSubsets$ do
    $X_{p,q}(1,i) \leftarrow 1$                                        ▷ Offset
    for $k \leftarrow iFirstMap$ to $iLastMap$ do
        $X_{p,q}(1+k,i) \leftarrow C_{p,q}[k]$                         ▷ Values assignment
    end for
    $r_{p,q}(i) = C_{p,q}[iLastMap + \Delta t_{maps}]$    ▷ Reference is $\Delta t_{maps}$ maps ahead
end for
$\underline{w}_{p,q} \leftarrow (\underline{\underline{X}}_{p,q} \cdot \underline{\underline{X}}^{T}_{p,q} + \lambda \cdot \underline{\underline{I}})^{-1} \cdot \underline{\underline{X}}_{p,q} . \underline{r}_{p,q}$    ▷ Pseudo inverse to get the regression coefficients
```

$$\widehat{C_{p,q}}[t_{maps} + \Delta t_{maps}] = \omega_{p,q}[0] + \sum_{u=1}^{U} \omega_{p,q}[u] \cdot C_{p,q}[t_{maps} - u + 1] \qquad U = 85; 12 \leq \Delta t_{maps} \leq 24 \tag{2.2}$$

The resulting UPC Predicted product is automatically generated on a daily basis and is designated as $U2PG$, which is compatible with the IONEX naming convention. Note that the predicted product for day T uses information until day $T-2$ and it is computed and made publicly available on day $T-1$ before midday (following SMOS mission request).

2.4.2 Justification of the prediction approach

During the design of the prediction system, different configurations have been explored before selecting the final architecture. The selection criterion was based on building sequential models, adding levels of complexity to a simple approach. Since the selection of the optimal configuration is a problem of combinatorial optimisation, which has an algorithmic complexity given by the binomial coefficient $\binom{n}{k}$, the following heuristics were considered in order to design the final system:

- **Basic features for the prediction**: we could have chosen the use of raw data, the representation of the raw data by means of picture elements (pixels), Spherical Harmonics (SH; [*Schaer et al.* (1999)]) or the representation by means of the DCT.

  Nevertheless, the use of raw data was unfeasible, due to the fact that it requires an extremely high dimensionality and also because it does not convey the natural information about the problem. By natural information we mean the relevant information that depends on the underlying physical processes. These processes evolve in both space and time and are responsible of the characterisation of the ionosphere at any given location.

Working directly with pixels (for instance, those associated with IONEX files) is not convenient for prediction (though we are assuming in all the cases that the analysis is done in a quite stationary reference frame, such as the one associated to local-time/latitude coordinates) due to the high number of terms involved and the prediction uncertainties, which makes a moothed product more desirable to the user.

An alternative is the use of Spherical Harmonics. But, since we wanted to explore the implementation of a simple image processing technique in GNSS ionospheric prediction, we decided to take as a natural representation of the problem a 2D flat image (in this work, a VTEC map). In addition, we decided to use the DCT instead of Spherical Harmonics because the Final IGS Predicted product is based on a combination of different forecasting products, one of which is already based on Spherical Harmonics (see [*Schaer et al.* (1999)]). As it is known, the square error of a linear combination of prediction methods can be decomposed as the addition of a bias term plus a variance term. Therefore, we decided to use a forecasting method as *orthogonal* to the rest as possible. In this way, the final error of the combined product will be smaller due to two effects: (1) the mean on the variance diminishes because of the aggregation of different methods with a difference variance, and (2) the systematic error (bias) will be different, if the underlying features of the model taken into account for the forecasting are different. And more importantly, in our preliminary experiments we compared systematically whether to use DCT or not, and we found a significant difference in favor of using it.

- **Memory of the forecast system**: the forecasting systems are based on using past information to give an estimate of future values. In this sense, we explored multiple window lengths of input data from the past. The selected length of the prediction and training windows used for the forecasting were justified by statistical performance (see section 2.4).

  The length of the prediction window, U, was tested for values between 13 and 481 (1 to 40 daily UPC IGS products from the past as input dataset) for three independent weeks (in 2004, 2010 and 2011 with or without disturbed ionospheric conditions, respectively) providing a consistent optimum value of 7 days (81 VTEC maps). As the performance curve is convex, the cutoff at 481 is justified because of the monotonous decrease in performance and because the degradation is so severe when U tends to 481 that higher values were not considered.

In the same way, the length of the training past window was tested for values up until 4393 (366 daily UPC IGS products). The results for the three independent weeks showed a better performance when selecting at least 100 days. In addition, the performance considering 366 days of past data was slightly better and thus, this was the adopted value for the research.

By observational evidence, partially supported by first principles, such selection would also be justified. Note that the ionization state, and thus the ionospheric VTEC, at a given moment depends on multiple periodic components. Among others, some of them are (see [*Hernández-Pajares et al.* (2009)]):

- The solar cycle (11-year period)
  tem Annual, semiannual and seasonal periods
- Three solar synodic rotations (81-day period)
- One solar synodic rotation (27-day period)
- The lunar semimonthly tide (14-day period)
- Planetary wave type oscillations at the Equatorial Ionospheric Anomalies (ex. 2-day, 5-day and 9-day periods)
- The diurnal, semidiurnal and terdiurnal tide effects (1, 0.5, 0.33-day periods, respectively)

Therefore, the selected lengths for the past windows would be compatible with our present knowledge.

- **The forecast method**: Apart from that, we had to decide the forecasting method. For this purpose, different forecasting techniques were tested. As a matter of fact, testing such methods was done in parallel with the evaluation of the memory of the forecast system by statistical performance.

  We decided to use a linear regression model for different reasons. The possible options were: decision trees, neural networks (multilayer perceptron) and Support Vector Machines (SVM).

  Decision trees were first dismissed because the approximation mechanism is based on a sequential set of decisions, which has to be carried out in parallel for each dimension. Therefore, if there is a high dimensionality manifold, the tree can grow to be extremely large and end up being a lookup table that could not be generalised. We did a first prototype with the primitive function on Matlab, and we found this effect. Note that if there is a high dimensionality manifold, which is diagonal along different dimensions, a decision tree might not be able to model the manifold properly. A linear regression model or a multilayer neural network would be the natural structure to model it.

The multilayer perceptron neural networks were also dismissed due to the high dimensionality of the input vectors, which are vectors of length equal to the dimensionality of the DCT representation times the length of the window with past samples. With such a high dimensionality vector we were limited in the selection of the training algorithm, which made unfeasible the use of a Levenberg-Marquardt kind of algorithm ([*Bishop* (1995)]). Finally, we tried with different versions of conjugate gradient learning algorithms. The result was that the performance was slightly lower than in the case of linear regression. Nevertheless, the training time and the forecasting time were much bigger (two orders of magnitude for the former, and one order of magnitude for the latter).

SVM were also dismissed due to the fact that the Gram matrix was too big, and there were problems in the optimisation phase related to the SVM algorithm.

Therefore, we selected as forecasting algorithm a linear regression.

All in all, the final algorithm that we have used in our experiments is based on a DCT representation of part of the spectrum of each VTEC map, followed by a linear regression using VTEC maps from the past (as described in Section 2.4).

2.5 Performance Analysis

2.5.1 Evaluated products

As mentioned before, the final, rapid and preliminary predicted IGS products for the different IAACs and their combined products were used for performance evaluation. And an alternative approach, termed the "frozen" or time-invariant ionosphere method in which no prediction model is applied (just keeping constant the VTEC values in local-time / latitude reference system), was also considered for comparative purposes. In this last case, the raw prediction relies directly on the VTEC values provided by the Final/Rapid IGS product produced two days before. The Rapid product for $T-2$, which is made available at CDDIS on day $T-1$, is directly considered as the Time-invariant ionosphere product for day T (no prediction being applied; ionosphere is considered frozen for two days). Note that this internal product is designated as $UPR2$. Also, Time-invariant products for CODE, ESA, JPL and IGS are generated internally (named $COR2$, $ESR2$, $JPR2$ and $IGR2$, respectively). In cases where Rapid IGS products are not generated, the corresponding Final IGS products are used instead (for instance, the Time-invariant product for UPC would be designated $UPC2$).

2.5.2 Reference products

UPC Final IGS products

As mentioned before, the UPC post-processed maps corresponding to the UPC Final IGS product ($UPCG$) are used as reference data. Additionally, the combined Final and Rapid IGS products ($IGSG$ and $IGRG$, respectively), are taken into account since they provide the most accurate ionospheric VTEC values at a global scale when using GNSS data. The suitability and reliability of the above-mentioned products in representing ionospheric VTEC have been demonstrated in previous years ([*Hernández-Pajares et al.* (2009)]). In this way, IGS products are widely used in the scientific community. Nevertheless, it is important to note that combined Final and Rapid IGS products are not totally independent because UPC Final and Rapid VTEC maps are used to generate these products. Note that, in general, the combined VTEC maps have a better performance than the individual products because systematic errors are partially filtered thanks to the combination process (see [*Orús et al.* (2007)]).

JASON altimeter data

Additionally, an external source of global VTEC measurements is recorded with the dual-frequency altimeter instrument on board the JASON-1 spacecraft, which was launched on December 7, 2001. The VTEC observables provided by these altimetry data over the oceans between latitudes of 66^oN and 66^oS (this restriction is given through the inclination of the JASON orbit), where no permanent GNSS receivers can be placed. In this way, JASON provides independent reference data that can be used to evaluate the performance of GNSS-derived VTEC maps over the oceans except in the polar regions.

JASON data have been used to validate Final IGS products ([*Hernández-Pajares et al.* (2009)]). In this respect, two considerations have to be taken into account. First, JASON VTEC measurements are very accurate but are affected by a certain offset with respect to the IGS products ([*Azpilicueta et al.* (2008)], [*Hernández-Pajares* (2004)]). In this regard, it should be noted that the GPS constellation orbits the earth at an altitude of $\cong 20200km$. Therefore, GPS VTEC observations for receiver-satellite pairs include the plasmaspheric contribution, in contrast to the case of JASON that orbits at an altitude of $\cong 1300km$. Second, this is a pessimistic scenario because JASON direct measurements are compared with interpolated values at the same exact location derived from the nearby grid points of the VTEC map being validated. And it is likely that even the VTEC at these grid points had to be interpolated from insufficient real data due to the lack of permanent GNSS stations over the oceans.

As a final remark, JASON VTEC data are made publicly available from year 2003 through the FTP site of the National Aeronautics and Space Administration-Jet Propulsion Laboratory (NASA-JPL) Physical Oceanography Distributed Active Archive Center (PO.DAAC; http://podaac.jpl.nasa.gov).

2.5.3 Methods for performance evaluation

As a first validation of the results, the UPC Predicted product is compared to the UPC Final IGS product, which includes the most accurate post-processed VTEC data that UPC provides. As mentioned before, a long time series of UPC IGS products is used as input dataset for the training process. Therefore, the model is specifically designed to predict the UPC Final VTEC values as accurately as possible.

Table 2.1 shows the bias, standard deviation, Root Mean Square (RMS) and the minimum and maximum values of the differences between the predicted UPC VTEC maps and the UPC Final IGS VTEC maps for the periods being investigated. Its main purpose is to determine whether there is an offset or a deviation that is systematically affecting all of the values of the VTEC maps. Histograms on a semi-logarithmic scale are also provided (Figures 2.4 and 2.5) as they reflect the behaviour of the model at the tails of the data distribution, which correspond to the areas with the worst prediction performance (i.e., details of the extreme values are emphasised and an idea of how likely they are is given).

As a second validation, a test against the VTEC records provided by the JASON altimeter instrument is used to verify that the predicted product is consistent with an external VTEC data source. The purpose of this test is to show that the prediction performs well in modeling the real VTEC and not only the VTEC of the corresponding UPC Final IGS product. This test is also sensitive to the quality of the input data used in the prediction model.

In this case, Table 2.2 shows the bias, standard deviation, RMS, the minimum and maximum values, and the number of comparisons of the differences between the UPC Final/Predicted/Time-invariant products and the JASON VTEC determinations. Furthermore, plots of the bias and standard deviation in terms of the geomagnetic latitude are provided (Figures 2.6 and 2.7) to quantify and show where the prediction systematically over- or underestimates VTEC. Finally, these plots can be useful for determining whether the peaks of the ionospheric anomaly have been correctly determined. For instance, this can be of great interest to identify whether the prediction model is providing excessively smoothed predicted VTEC values.

Finally, a comparison between the preliminary IGS predicted products and the JASON VTEC values is performed for the different IAACs. This test is of interest to performance comparisons between the UPC Predicted product and the predicted products produced by the other IAACs (not to comparisons between prediction methods). It is also useful to first evaluate the two days ahead combined IGS Predicted product ($I2PG$).

For this purpose, boxplots of the differences between the predicted products and JASON VTEC values are depicted (Figure 2.8). The boxplots allow the distribution of the results to be summarised for the different prediction methods in a simplified and compact way by observing the quartiles (see labels on the rightmost boxplot in Figure 2.8) as well as the smallest and largest daily standard deviation observations. Remember that the lower quartile (designated as $Q1$) is the threshold for the lowest 25% values in the range of results, $Q2$ corresponds to the median and $Q3$ is used to indicate the highest 25% of data. Note that, in this work, the boxplots were computed following the Tukey method ([*Hoaglin* (1983)]).

2.5.4 Selected periods

One month of data has been analysed for each of the years 2004 and 2006, as well as 115 days in 2010. In this way, representative datasets have been selected to cover conditions of both low and mid-high solar and geomagnetic activity. Although data from the maximum activity period of the 23^{rd} Solar Cycle has not been considered, the effects derived from certain solar flares and geomagnetic storms are present during all the three selected periods. Nevertheless, it should be stated that the correct predictability of impulsive events is not the current concern of the UPC prediction model and can be the subject of further research.

In 2004, the month of August was analysed, i.e., from day of year 214 to 244. This period corresponds to medium solar cycle conditions. Nevertheless, the geomagnetic activity was strong at the end of the month due to a solar storm and two X-class flares were recorded by GOES ([*Hill et al.* (2005)]) on days 226 and 231. In fact, the Kp index ranged from 0 to 7 and its maximum value at 00 - 03UT on day 244. At that time, the IGS did not provide Rapid products so the Time-invariant product has been generated using the corresponding UPC Final IGS products ($UPC2$).

A second period is selected (December, 2006) for the minimum solar cycle conditions. In this case, the Kp index ranged from 0 to 8.3 and its maximum value at 00 - 03UT on day 349. In addition, a geomagnetic storm occurred on day 348, and three X-class flares were recorded by GOES on days 340, 347 and 348. Also in this case, Time-invariant VTEC maps are generated using UPC Final IGS products.

The period in 2010 covers 115 days, from day of year 184 to 355. Here, performance comparisons could be made with the other IAACs predicted products for the exact same period. During this time interval, a geomagnetic storm happened on days 215-216, but no powerful X-class solar flares were recorded by GOES. The Kp index ranged from 0 to 6.7, and its maximum at 00 - 03UT on day 216.

Note that the selected days in 2010 are not consecutive, as days for which rapid/predicted/time-invariant products of each IAAC centre and the combined IGS products were not available were discarded. Additionally, days in which there were a low number of JASON observations (in this work, less than 10,000 observations) were also discarded. Furthermore, it is also required that the combined IGS Predicted product for two days ahead ($I2PG$) was generated using the three individual predicted products from CODE, ESA and UPC. In this regard, the first day in which all three products were combined was day of year 184 in 2010. Taking these factors into account, a total of 115 days were selected between days 184 and 355 in 2010.

2.5.5 Performance against UPC Final IGS product

In Table 2.1, the differences between the UPC Predicted/Time-invariant products (the test data) and the UPC Final IGS product (the reference data) are compared for the three selected periods. The bias, standard deviation, RMS and minimum and maximum values are included expressed in TEC units or TECUs (1 TECU is equivalent to $10^{16} el/m^2$), as well as the number of comparisons.

In this table, the UPC prediction model systematically provides better results than in the Time-invariant product in a solar-fixed reference frame.

Period	Product	BIAS	Std. Dev.	RMS	min	max	#Comp
$214 - 244, 2004$	UPC Predicted	0.23	2.37	2.38	-23.1	18.2	2,060,136
	UPC Frozen	-0.25	2.63	2.64	-25.5	27.0	2,060,136
$335 - 365, 2006$	UPC Predicted	0.07	2.27	2.27	-46.9	24.2	2,060,136
	UPC Frozen	0.11	2.69	2.69	-50.4	53.8	2,060,136
$184 - 355, 2010$	UPC Predicted	-0.27	2.49	2.50	-34.4	21.3	7,642,440
	UPC Frozen	-0.28	2.89	2.90	-35.6	31.9	7,642,440

Table 2.1: *Statistics of the differences between the UPC Predicted/Time-invariant VTEC data and the UPC Final IGS reference data, including bias, standard deviation, RMS, minimum and maximum differences for the three selected periods in 2004, 2006 and 2010 in [TECU].*

In fact, the RMS of the differences between the UPC Time-invariant product and the UPC Final IGS product ($UPR2/UPC2$ vs. $UPCG$) is 11%, 18% and 16% worse than the RMS of the differences for UPC Predicted product (for 2004, 2006 and 2010 periods, respectively). Regarding the bias over all grid points in each interval, a systematic under- or overestimation cannot be confirmed. The UPC prediction model tends to overestimate VTEC in the 2004 interval while it underestimates VTEC in the 2010 interval.

It is also interesting to look at the histograms of the above-mentioned comparisons. In Figures 2.4 and 2.5, the histograms of the differences between the UPC Predicted/Time-invariant products and the UPC Final IGS products ($U2PG$ vs. $UPCG$ and $UPR2/UPC2$ vs. $UPCG$, respectively) are plotted. For the selected periods, the results for the UPC Predicted product show lower residual values than for the UPC Time-invariant ionosphere. Both the negative and positive tails of the distributions finish at lower values for the UPC Predicted product (see also Table 2.1). A significant improvement is obtained looking at the relative differences between the maximum values of the UPC Predicted product and the Time-invariant product. The Time-invariant maximum values are 48%, 122% and 50% higher than the Predicted ones (in 2004, 2006 and 2010 periods, respectively), while the minimum values also experience higher values but in a lower proportion.

These results are caused by the disturbed conditions affecting the three selected periods. Although such effects cannot be predicted, the UPC Predicted product is more robust against them than the Time-invariant product. This is because the disturbed VTEC values are directly used to generate the Time-invariant product two days afterwards (for $T + 2$). In contrast, the

UPC Predicted product is generated taking into account a set of seven days of VTEC files as input dataset so that the impact two days afterwards is lower (see Section 2.4).

The histograms in 2004 interval (upper plot in Figure 2.4) show an important asymmetry affecting the negative values (underestimation of VTEC). This is due to the increase in VTEC related to the storm that occurred at the end of August, 2004, which was not predicted. The histograms in 2006 interval (lower plot in Figure 2.4) show that the UPC Predicted product histogram ($U2PG$ vs. $UPCG$) is asymmetrical with a higher number of samples on the negative side due to the underestimation of VTEC for the disturbed period on day 348. On the other hand, the Time-invariant histogram ($UPC2$ vs. $UPCG$) is symmetrical because the underestimation on day 348 is compensated by the corresponding overestimation on day 350 (as the UPC Time-invariant product is generated from the disturbed UPC Final IGS product). Finally, for the 115 days in 2010 (Figure 2.5), the histograms show a sudden decrease in the positive tail related to the days not taken into account (remember that the selected days may not be consecutive) that are close to the most disturbed days in the period being considered.

2.5.6 Performance against JASON data

For the external validation with JASON, bias, standard deviation, RMS and the number of comparisons are shown in Table 2.2. In this table, the combined Final IGS product and the UPC Final IGS product against the JASON VTEC values are shown as reference ($IGSG$ vs. JASON and $UPCG$ vs. JASON, respectively). The standard deviation, which is not affected by the JASON offset (see Section 2.5.2), and RMS results obtained in cases where the UPC prediction model is applied are systematically better than for the time-invariant ionosphere approach.

On the one hand, the standard deviation results of the differences between UPC Predicted products and the JASON reference data ($U2PG$ vs. JASON) are 27%, 33% and 28% worse than that of UPC Final IGS product ($UPCG$ vs. JASON) for the 2004, 2006 and 2010 periods, respectively. On the other hand, the standard deviation results for the UPC Time-invariant products ($UPR2/UPC2$ vs. JASON) are 32%, 50% and 38% worse than that of the UPC Final IGS product ($UPCG$ vs. JASON) for the three periods. In this way, the percentage of improvement, which is the decrease in the variability with respect to JASON, is 5%, 17% and 10% for the three periods. Apart from that, the bias results show that all IGS VTEC values are underestimated with respect to the JASON ones. In this regard, the biases on JASON VTEC (reported as an excess of few TECUs by several

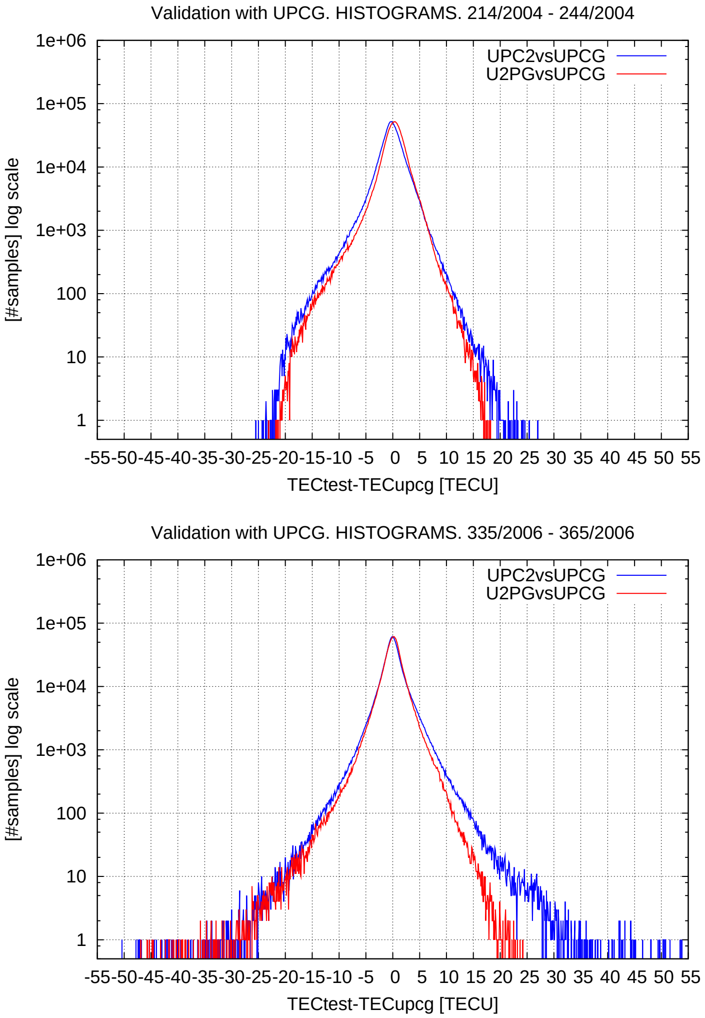


Figure 2.4: *Histogram on a semi-logarithmic scale of the differences between UPC predicted/Time-invariant VTEC values and the UPC Final IGS VTEC values (U2PG vs. UPCG and UPC2 vs. UPCG, respectively) for the periods from days 214-244, 2004 (upper plot) and 335-365, 2006 (lower plot).*

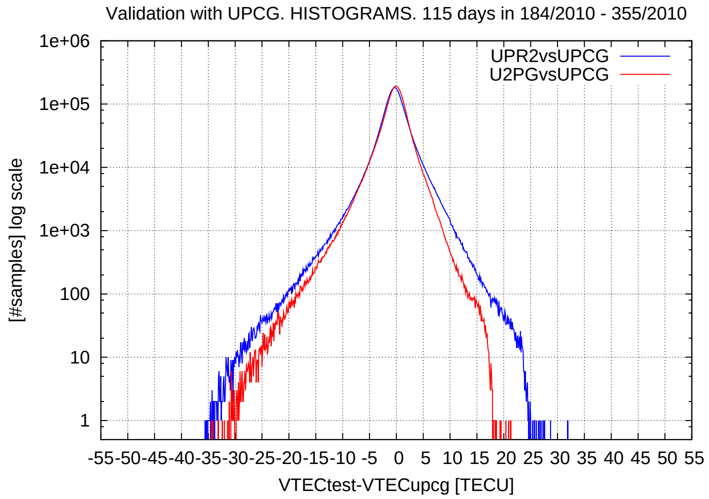


Figure 2.5: *Histogram of the differences between the UPC predicted/Time-invariant VTEC values and the UPC Final IGS VTEC values (U2PG vs. UPCG and UPR2 vs. UPCG, respectively) on a semi-logarithmic scale, shown for the 115 day period from days 184-355, 2010.*

authors) and the potential biases in each GPS VTEC map can be larger than the plasmaspheric component (which maximum is at low geomagnetic latitude) during solar minimum and mid conditions. Also note that the plasmaspheric Electron Content (EC) component should basically affect GPS measurements due to the orbit height of JASON altimeter (the differential EC would be given between 1300 and 20200km of height). The signature of this component is clearly seen when the average deviation between JASON VTEC and GPS VTEC is represented in terms of the latitude, showing an excess of few TECUs at the equator with respect to high latitude regions ([*Hernández-Pajares et al.* (2009)]).

Latitudinal behaviour

The dependence of model performance on the geomagnetic latitude is studied for the different IAACs as well as for the IGS combination for the period in 2010. The latitudinal behaviour of the VTEC predictions can be analysed; for example, this can be accomplished by plotting the bias and standard deviation values in multiple geomagnetic latitude bins (of two degrees in this work) when comparing these values with external JASON data.

In the upper plot in Figure 2.6, the bias of the differences between the predicted products and the JASON data is depicted. In it, the plasmas-

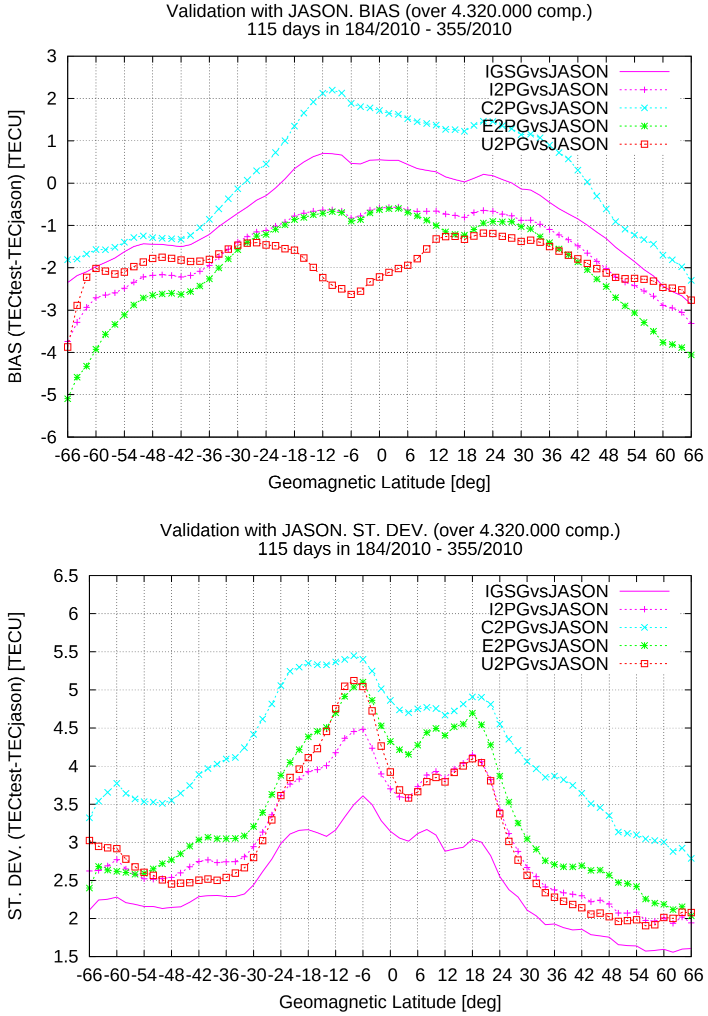


Figure 2.6: *Bias (top) and standard deviation (bottom) of the differences between the IGS, CODE, ESA and UPC two days ahead predicted products and the JASON data (I2PG vs. JASON, C2PG vs. JASON, E2PG vs. JASON and U2PG vs. JASON; dotted lines). The assessed values were in two degrees geomagnetic latitude bins for the 115 day period from days 184-355, 2010. The differences between the combined Final IGS product and JASON data are also included as reference (IGSG vs. JASON; continuous line).*

Period	Product	BIAS	Std. Dev.	RMS	#Comp
$214 - 244, 2004$	IGS Final	-0.43	3.22	3.25	1,167,252
	UPC Final	-1.29	3.14	3.40	1,167,252
	UPC Predicted	-0.97	4.00	4.12	1,167,252
	UPC Frozen	-1.54	4.16	4.43	1,167,252
$335 - 365, 2006$	IGS Final	-1.65	2.83	3.28	1,249,145
	UPC Final	-2.02	2.78	3.44	1,249,145
	UPC Predicted	-1.97	3.70	4.19	1,249,145
	UPC Frozen	-1.87	4.16	4.56	1,249,145
$184 - 355, 2010$	IGS Final	-0.57	2.77	2.83	4,341,359
	UPC Final	-1.30	2.61	2.92	4,341,325
	UPC Predicted	-1.83	3.34	3.81	4,341,359
	UPC Frozen	-1.83	3.61	4.05	4,341,359

Table 2.2: *Statistics of the differences between the UPC Predicted/Time-invariant VTEC data and the JASON reference data, including bias, standard deviation and RMS for the three selected periods in 2004, 2006 and 2010 in [TECU].*

pheric component of the VTEC clearly affects GNSS measures at low latitudes ([*Pierrard and Stegen* (2008)], [*Hernández-Pajares et al.* (2009)]) and thus produces the typical (inverted) U shape in terms of the latitude ([*Aragon-Angel* (2010)]). Note that the bias of the differences between UPC Predicted product and JASON data ($U2PG$ vs. JASON) experiences a decrease in the equatorial region centred at -6 degrees. This VTEC underestimation is also present in the bias of the differences obtained from the UPC Rapid IGS products ($UPRG$ vs. JASON; not plotted). In this regard, note that UPC Rapid IGS products are used as the input dataset for the UPC prediction model to generate the UPC predicted VTEC maps (see Section 2.4). This issue will have to be further investigated in the future.

In the lower plot in Figure 2.6, the standard deviation of the differences are plotted showing a clear correlation with the effect of the Appleton-Hartree equatorial anomaly. In particular, higher standard deviation values are obtained when modelling the anomaly in the Southern Hemisphere.

First of all, the behaviour of the differences between the combined IGS Predicted product and the JASON data ($I2PG$ vs. JASON) is quite similar (in shape, but at a higher level) to the differences between the combined IGS Final product and the JASON data ($IGSG$ vs. JASON; used as a point of reference). This shows the importance of the combination process as the performance of the combined IGS Predicted product is better than the individual predicted products. Additionally, the UPC Predicted product shows a good overall standard deviation performance compared to the results obtained from CODE and ESA Predicted products. Nevertheless, VTEC is still underestimated when modeling the Southern Hemisphere peak. This feature seems to be related to the decrease in the bias that can be seen in the upper plot in Figure 2.6.

In the upper plot in Figure 2.7, the standard deviation in terms of the latitude band is shown for the differences between UPC Final/Predicted/Time-invariant product and the JASON data ($UPCG/U2PG/UPR2$ vs. JASON). The UPC Predicted product systematically yields better results than for the UPC Time-invariant ionosphere. In the lower plot in Figure 2.7, the combined IGS Predicted product ($I2PG$ vs. JASON) behaviour is shown and yields to slightly better results than for the IGS Time-invariant ionosphere.

Boxplot analysis

The boxplot is a non-parametric representation that summarises graphically the important statistics of a sample. This allows to compare easily different distributions by comparing the position of specific quartiles (for a more detailed description of this tool see the last paragraph in Section 2.5.3). Figure 2.8 shows the boxplots for the GNSS VTEC map sources under consideration, i.e., IGS as well as the IAACs providing predicted products, against JASON data. For each of these products, three types of data are plotted for comparison: the IGS and IAACs official Rapid products (on the left hand side for each case; indicated with a dotted green line), the preliminary predicted products (in the centre; a continuous red line) and the time-invariant products (on the right; a dotted blue line). Note that the rapid products are being used rather than the final ones to show the variability of the most recent VTEC maps used to generate the predictions (at least in the UPC case).

The results show that there is a strong dependence on the performance of the corresponding rapid products. Applying a prediction model leads to better results than does the use of the time-invariant ionosphere for all of the individual products that are taken into account. In addition, applying a prediction model generates an increase in the data dispersion which is, in general, lower than the dispersion observed in the time-invariant ionosphere VTEC maps.

In the case of the differences between the UPC Predicted product and JASON reference data ($U2PG$ vs. JASON), the obtained standard deviation is lower than that observed for the other IAACs for the analysed dataset

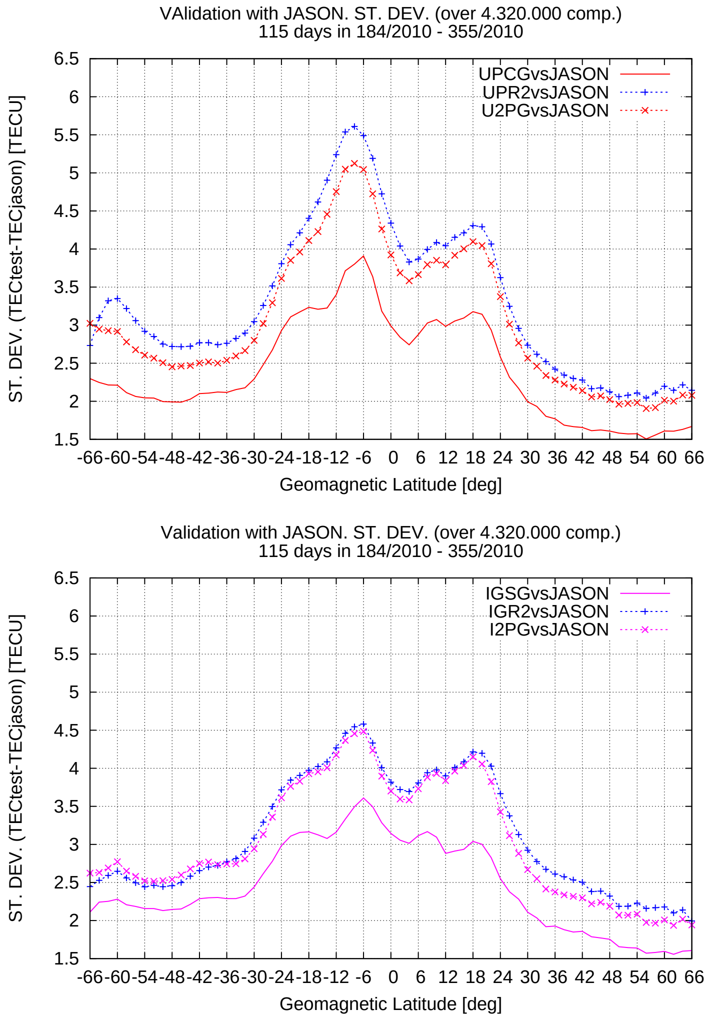


Figure 2.7: *Standard deviation of the differences between the UPC and IGS Final/Predicted/Time-invariant products and the JASON data (UPCG/U2PG/UPR2 vs. JASON in the upper plot and IGSG/I2PG/IGR2 vs. JASON in the lower plot). The assessed values were in two degrees geomagnetic latitude bins for the 115 day period from days 184-355, 2010.*

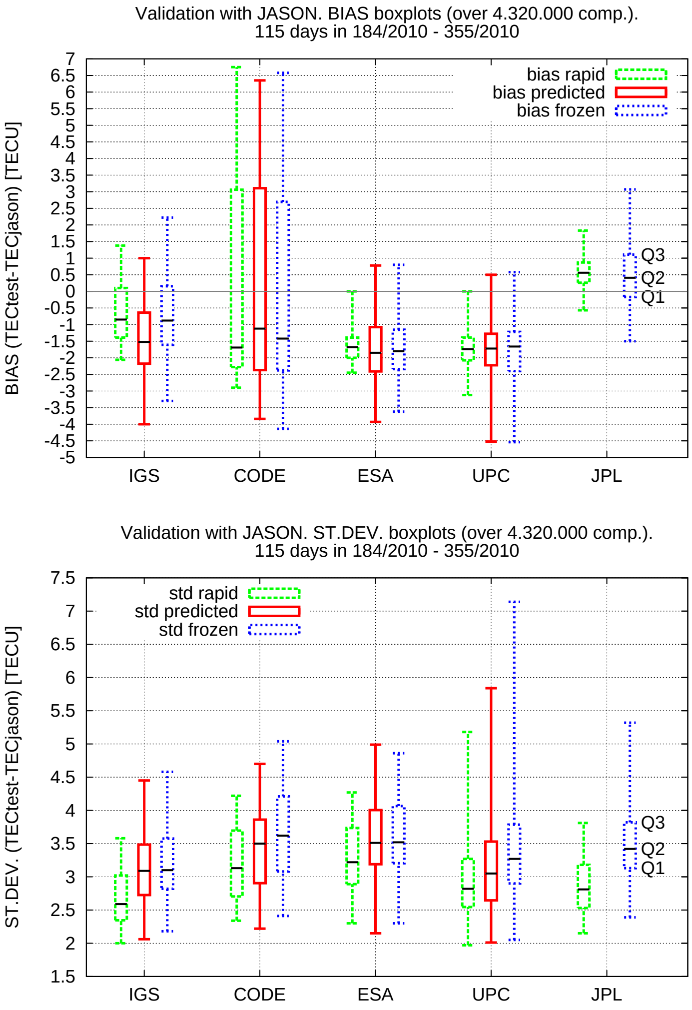


Figure 2.8: *Bias (top) and standard deviation (bottom) boxplots of the differences between the rapid/predicted/time-invariant test data and the JASON reference data for the 115 day period from days 184-355, 2010. The results for the rapid/predicted/time-invariant test data are plotted in the leftmost/centre/rightmost boxplots for each test source on the x-axis. Test data include CODE, ESA, UPC, JPL and combined IGS rapid products, preliminary predicted products, and time-invariant products (generated internally), if available (note that JPL does not currently provide predicted products).*

of 115 days in 2010. Note that the standard deviation values for the differences between UPC Predicted product and JASON ($U2PG$ vs. JASON) and the differences obtained from a CODE Predicted product for two days ahead forecast ($C2PG$ vs. JASON) are clearly better than that observed for the corresponding time-invariant VTEC values (i.e., $UPR2$ vs. JASON and $COR2$ vs. JASON). It should also be noted that the standard deviation boxplot displays a slightly lower median ($Q2$ quartile) for the UPC Predicted product than for the combined IGS Predicted product. Regarding the bias boxplots, which are depicted in the upper plot in Figure 2.8, the differences for the UPC Predicted product (U2PG vs. JASON) and the ESA Predicted product (E2PG vs. JASON) indicate a similar performance compared to the corresponding Time-invariant products ($UPR2$ vs. JASON and $ESR2$ vs. JASON). In the case of the CODE Predicted products, the bias results show a larger dispersion but a lower median.

Regarding the combined forecast product, the differences between the combined IGS Predicted product and the JASON data ($I2PG$ vs. JASON) indicate a slightly better performance compared to the corresponding Time-invariant product ($IGR2$ vs. JASON) based on the standard deviation values.

Using IGS Time-invariant prediction ($IGR2$) may lead to better results than considering the combined IGS Predicted product. This could be related to the weighting scheme that has been applied during the combination process (see [*Orús et al.* (2007)]). For the period under consideration, this may be related to the bias performance of the JPL products as there is no predicted product provided by JPL. Note that the JPL Rapid IGS product ($JPRG$) and the Time-invariant product ($JPR2$) have also been added for reference. In this regard, it seems that the future availability of a JPL predicted product could further improve the combined product.

2.6 Conclusions

The use of representative datasets has demonstrated that the UPC prediction model performs well, particularly when results are compared with those obtained by other IAACs. It is shown that applying the prediction model leads to better results than the use of time-invariant ionosphere for two days ahead predictions. This conclusion can be extended to the prediction models of CODE and ESA as well. Nevertheless, the predicted products still have the potential for further improvements to achieve better results. The results obtained in this thesis work suggest that the inclusion of a potential future predicted product from JPL could increase the accuracy of the combined IGS predicted product. In addition, it might be possible to apply a prediction model directly to the IGS final/rapid products, although this may deviate from the philosophy of IGS of combining several independent products.

Last but not least, the future multi-frequency/multi-constellation GNSS scenario, as well as the deployment of more permanent GNSS stations distributed worldwide within the IGS framework, can enable the generation of improved UPC Final/Rapid VTEC maps. Consequently, better predicted products could be released in case these products are used as input data for the prediction models, as is the case of the UPC prediction model.

Conclusions

In the first chapter, we present a summary of the conclusions related to the new solar flare detector. The main conclusions are (see Section 1.6):

- A new GNSS solar flare monitoring method has been presented showing its good performance to detect the most powerful X-class flares
- The detector was able to detect as many as 93% of the X-class flares reported by GOES when considering a percentage of *tentative* false detections below 4% (95% accepting a 7% of false detections)
- SISTED can be a very useful tool to prevent the harmful consequences of the solar flares' related particles
- SISTED automatically generates warning messages that could be sent to GNSS users
- A potential service based on SISTED could be deployed in the near future, which would be especially useful for providing potential ionospheric storm warnings to GNSS users with integrity requirements (such as in Civil Aviation).

In addition, in the second chapter, we have developed a new method for the prediction of VTEC, which has resulted in a new UPC prediction product. This product is used to generate the combined IGS Predicted product (together with CODE and ESA products). The main conclusions are (see Section 2.6):

- The UPC prediction model for two days ahead performs well, particularly when the results are compared with those obtained by other IAACs
- Applying the prediction model leads to better results than the use of time-invariant ionosphere for two days ahead predictions

- For improved predictions, especially for medium and long-term predictions, physics-based models should be envisaged
- The future multi-frequency/multi-constellation GNSS scenario, as well as the deployment of more permanent GNSS stations, can lead to better predicted products.

Appendix A

The Global Positioning System

A.1 Introduction

Nowadays, there is an increased interest in Global Navigation Satellite Systems (GNSS) with the modernisation of the NAVigation Satellite Timing and Ranging Global Positioning System (NAVSTAR GPS). Moreover there is a renewed impulse to the Russian GLObal NAvigation Satellite System (GLONASS) and the deployment of new GNSS such as the European Global Navigation Satellite System (GALILEO) or the Chinese GNSS System (COMPASS). In the last years, the potential users with access to GNSS and the number of comercial applications based on GNSS have grown significantly, in fields such as precise navigation and timing, and ionospheric and tropospheric sounding, among others.

Descriptions of the GPS system can be found in [*Wells* (1987)], [*Seeber* (1992)], [*Hoffmann-Wellenhof et al.* (2004a)] and [*Parkinson and Spilker* (1996)], among others.

A.1.1 GPS system description

The GPS system is the most well-known GNSS. It is comprehensively described in literature, for example, in [*Wells* (1987)], [*Seeber* (1992)], [*Hoffmann-Wellenhof et al.* (2004a)] and [*Parkinson and Spilker* (1996)].

Regarding GPS architecture, one can distinguish between three main segments:

- *Space segment:* Consists of the GPS satellite constellation which is comprised of 24 satellites plus some on-orbit spares (Full Operational Capability, FOC), evenly distributed within 6 orbital planes with an inclination to the equator of the Earth of 55^o and equally spaced 60^o (see Figure A.1). Their orbits are near-circular with a semi-major axis of about 4.1 times the Earth radii (nominal orbits of 20200km with respect the surface of the Earth). The orbital period is approximately 12 sidereal hours. This configuration guarantees a global 24-hour coverage with, at least, four satellites in view, which is the minimum number of satellites required to solve the position of a GPS receiver.

- *Control segment:* Consists of a network of ground stations, whose functions are:
    - Control and keep the status and configuration of the satellite constellation.
    - Predict satellites ephemeris and on-board clock evolution by solving the inverse problem with the directly gathered GPS measurements.
    - Keep the GPS time scale.
    - Periodically update the navigation message broadcast via the satellites.
- *User segment:* Consists on the GPS receivers of all the users. The GPS receivers gather the GPS signal from the satellites and solve the navigation equations in order to obtain its own coordinates and clock error.

A.1.2 GPS signal

Each GPS satellite continuously broadcasts a set of binary codes by coherently modulating the phase of the carrier of the transmitted radio signal. These codes, which are unique to the broadcasting satellite, are pseudorandom and mutually orthogonal. The codes are also known as Pseudo Random Noise (PRN) codes and there are more codes than potential number of satellites, so each code identifies the satellite. They are used for ranging and for transmitting almanac and timing information. The mutual orthogonality property of the codes enables the receiver to isolate the received signals broadcast by a given satellite from all others by cross-correlation techniques, and to process in parallel the signals from all satellites in view of the receiver. The GPS satellites broadcast ranging codes on a pair of phase coherent L-band carriers, the $L1$ carrier at a frequency of 1575.4 MHz and $L2$ at 1227.6 MHz:

$$\begin{aligned} f1 &= 154 \cdot 10.23\ MHz = 1575.42\ MHz \\ f2 &= 120 \cdot 10.23\ MHz = 1227.60\ MHz \end{aligned}$$

which correspond to approximate wavelengths of 19 cm (λ_1) and 24 cm (λ_2). These include an encrypted precision (P) code with a chip rate of 10.23 MHz on both carriers and the clear access or Coarse/Acquisition (C/A) code at 1.023 MHz on the $L1$ carrier. The dual carriers are needed primarily

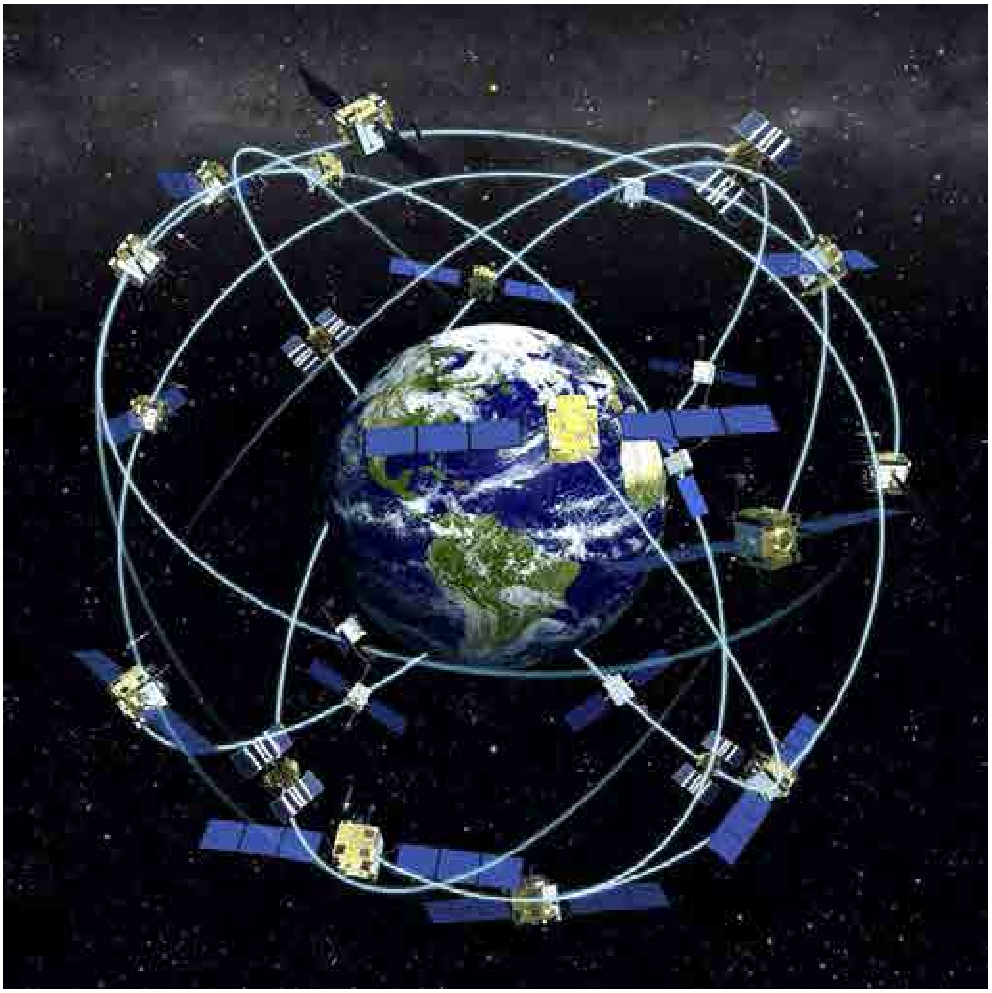


Figure A.1: *The space segment includes a satellite constellation of at least 24 Medium Earth Orbiting (MEO) satellites (image taken from http://www.defenseindustrydaily.com).*

to eliminate (or determine) the refraction effect from the ionosphere. For a microwave in the ionosphere, the refractivity is very nearly proportional to the local electron density and inversely proportional to the square of the carrier frequency. Therefore, the range and phase information received separately from the two carriers can be applied in concert to nearly completely decouple the ionospheric refraction effect by using this dispersive property of the ionospheric plasma. Newer versions of the GPS satellites forseen for the modernisation of GPS, **GPS III**, will have an additional carrier at 1176.45 MHz ($L5$) and a C/A-like code also on $L2$. This will significantly improve receiver tracking operations using clear access ranging codes and increase the accuracy of the ionosphere calibration.

The resulting transmitted signal corresponds to expression A.1, its schematic interpretation being shown in Figure A.2.

$$\begin{aligned} S_{GPS}(t) = \; & A_c \cdot C(t) \cdot D(t) \cdot \sin(2\pi f_1 + \phi_c) + \\ & + A_p \cdot P(t) \cdot D(t) \cdot \cos(2\pi f_1 + \phi_{p1}) + \\ & + A_p \cdot P(t) \cdot D(t) \cdot \sin(2\pi f_2 + \phi_{p2}) \end{aligned} \tag{A.1}$$

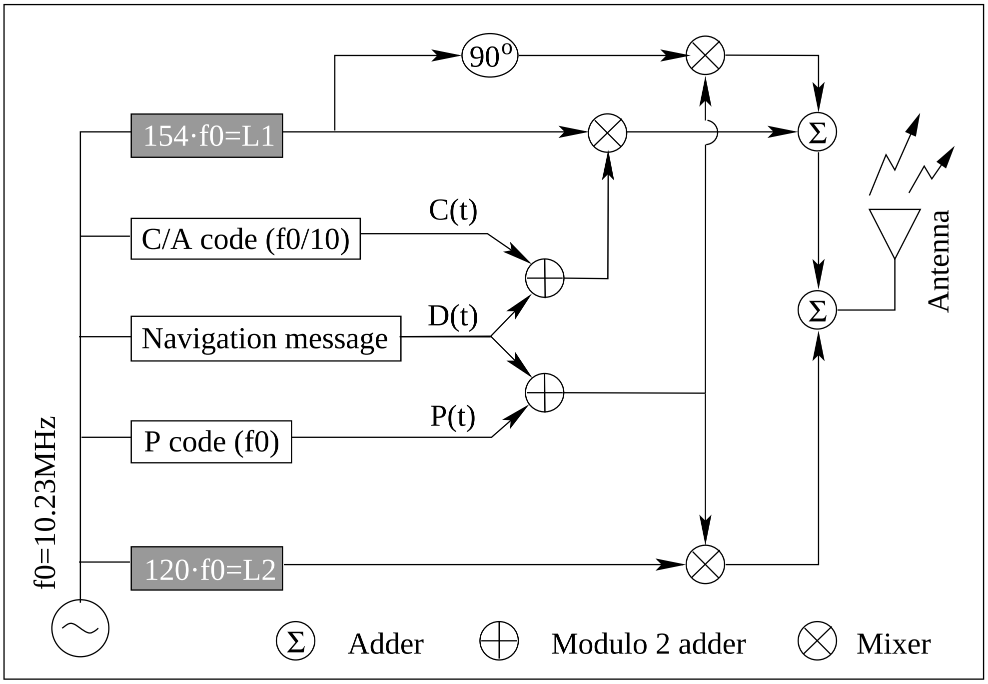


Figure A.2: *Diagram of the GPS signal structure. Source [García-Fernández (2004)].*

In order to limit the precision of the position of civilian users calculated by means of GPS, two modifications were applied to the GPS signal by the Department of Defense of Unites States:

1. Selective availability (SA) which consists of the intentional manipulation of the satellite clocks. The ephemerides contained in the navigation message are also modified. The resulting effect is an increase in the positioning error from 10m up to more than 100m approximately. On 1st May, 2000, the SA was disconnected.

2. Anti/Spoofing (AS) which consists of the encryption of the precision P-code (into the Y-code) so non-authorised receivers are unable to use

it, forcing these users to rely solely in the worse C/A code on f_1, and in an indirect and noisier estimate of a code on f_2.

A.2 GPS observables

In order to obtain the GPS observables that will be used for processing, the GPS receivers correlate (compare) the incoming signal with an internally generated copy. They basically measure the time or phase differences between both signals. If the differences are obtained from the PRN C/A or P codes (time differences) one will obtain the Code Pseudorange. Otherwise, by obtaining phase differences of the carrier frequency one will obtain Phase pseudorange. Pseudoranges are "ranges" because they are an estimation of the geometric distance between the satellite and the receiver (traveling time multiplied by the light speed gives apparent distance), and are "pseudo" because this range is not the actual geometric range since it is affected by a set of errors and delays. The Code Pseudoranges (expressed in units of length) can be modeled for both frequencies as:

$$\begin{aligned} P1_i^j &= \rho_i^j + c(dt_i - dt^j) + I_{1,i}^j + T_i^j + rel_i^j + K1_i^j + M_{P1,i}^j + \varepsilon_{P1,i}^j \\ P2_i^j &= \rho_i^j + c(dt_i - dt^j) + I_{2,i}^j + T_i^j + rel_i^j + K2_i^j + M_{P2,i}^j + \varepsilon_{P2,i}^j \end{aligned} \tag{A.2}$$

where,

- ρ_i^j is the geometric range between the satellite j and the receiver i at emission and reception time, respectively ($\sim$20000 km).
- c is the speed of light (299792458 m/s is the standard in GPS system).
- dt_i is the offset of the receiver i from GPS time (<300 km).
- dt^j is the offset of the satellite j from GPS time (<300 km).
- rel_i^j is the relativistic effect (<13m).
- T_i^j is the tropospheric effect (2-10m).
- $I_{k,i}^j$ is the ionospheric effect, which can be expressed in first order (more than 99.9% of the total effect) as: $I_{k,i}^j = \alpha_k \cdot STEC$ being $\alpha_k = 40.3/f_k^2$ (2-50 m) and STEC, the slant Total Electron Content i.e. the line-of-sight integrated electron density.
- K_i^j is the satellite and receiver instrumental delays, also called Total Group Delay or TGD (<2m).

- $M^j_{P,i}$ is the effect of multipath (<15m).

- $\epsilon^j_{P,i}$ is the thermal noise and other unmodeled sources of errors (3m).

Similarly, the Phase Pseudoranges, expressed in units of length as well, can be modeled as:

$$\begin{aligned} L1^j_i &= \rho^j_i + c(dt_i - dt^j) - I^j_{1,i} + T^j_i + rel^j_i + B1^j_i + w_{L1} + m^j_{L1,i} + \varepsilon^j_{L1,i} \\ L2^j_i &= \rho^j_i + c(dt_i - dt^j) - I^j_{2,i} + T^j_i + rel^j_i + B2^j_i + w_{L2} + m^j_{L2,i} + \varepsilon^j_{L2,i} \end{aligned} \tag{A.3}$$

where, apart from the notation introduced in the previous expressions, one can find:

- w_L is a term due to the relative rotation of the transmiting and receiving antennas. Known as wind-up, the direct consequence of this effect is that spinning the antenna is understood by the receiver as an apparent variation of distance between satellite and receiver (<20 cm).

- $B1^j_i$ and $B2^j_i$ are the ambiguity terms, including the carrier-phase instrumental delays (∼20000 km), this value is kept constant while the receiver keeps track of the GPS satellite.

- $m^j_{L,i}$ is the effect of multipath. This effect is much smaller than pseudorange multipath (<2 cm).

- $\epsilon^j_{L,i}$ is the thermal noise and other unmodeled sources of errors. Also much smaller than pseudorange $\epsilon^j_{P,i}$ (<1 cm).

These terms are summarised in Table A.1.

The term Bn^j_i in the phase pseudoranges is defined as $Bn^j_i = b_i + b^j + \lambda N^j_i$, thus including both instrumental delays and integer ambiguity term. Phase processing consists basically on integrating the Doppler effect of the incoming GPS signal. Nevertheless it is not possible to measure the number of cycles between the GPS satellite and the receiver at the instance of first observation (this unknown quantity given by the Doppler integration process is the integer ambiguity). As a consequence, when the receiver loses visibility with the GPS satellite (for instance due to a building or vegetation) and relocks afterwards, the phase observable shows discontinuities known as cycle slips (an example of cycle slips in GPS signal is given in Figure A.3).

The terms corresponding to clock bias (with respect to the GPS time scale) present large errors unless they are properly accounted for. In the case of the GPS receiver, since it commonly uses a simple quartz clock to generate the replica of the GPS signal, a larger clock bias with respect to the

Geometric distance	ρ_i^j	$\simeq 20000km$
Receiver clock offset	dt_i	$< 300km$
Satellite clock offset	dt^j	$< 300km$
Ionospheric delay	I_i^j	1 to 50m
Tropospheric delay	T_i^j	1 to 20m
Relativistic effect	rel_i^j	$\simeq 10m$
Code Multipath effect	M	$0m - 3m$
Phase Multipath effect	m	$0cm - 5cm$
Code Noise	ε_P	$\simeq 3m(C/A)$ $\simeq 30cm(P)$
Phase Noise	ε_L	$\simeq 3mm$

Table A.1: *Main contributions to Pseudoranges. The multipath errors are difficult to quantify since it highly depends on the environment. Additionally, the a priori broadcasted satellite position and clocks can cause an extra error of 2 m approximately.*

GPS time scale occurs. Therefore, four GPS satellites, at least, are going to be needed to estimate this bias along with the 3D position coordinates of the GPS receiver. In the case of the GPS satellite, satellite clock bias can be mostly corrected with the data included in the navigation message (D(t)). Further details of each pseudorange contribution can be found in [*Hoffmann-Wellenhof et al.* (2004a)] and [*Parkinson and Spilker* (1996)].

Using these basic observables, it is possible to linearly combine them:

Free ionospheric combination (LC and PC): Taking into account that the ionospheric delay depends on the square frequency, it is possible to remove its effect by constructing this combination as follows:

$$Pc = \frac{f_1^2 \cdot P_1 - f_2^2 \cdot P_2}{f_1^2 - f_2^2} \qquad Lc = \frac{f_1^2 \cdot L_1 - f_2^2 \cdot L_2}{f_1^2 - f_2^2} \tag{A.4}$$

Obtaining:

$$Pc_i^j = \rho_i^j + c(dt_i - dt^j) + rel_i^j + T_i^j + M_{Pc,i}^j + \epsilon_{Pc,i}^j$$
$$Lc_i^j = \rho_i^j + c(dt_i - dt^j) + rel_i^j + T_i^j + Bc_i^j + w_{Lc} + m_{Lc,i}^j + \epsilon_{Lc,i}^j$$

This combination is basically used for navigation purposes in receivers that are able to process both GPS frequencies.

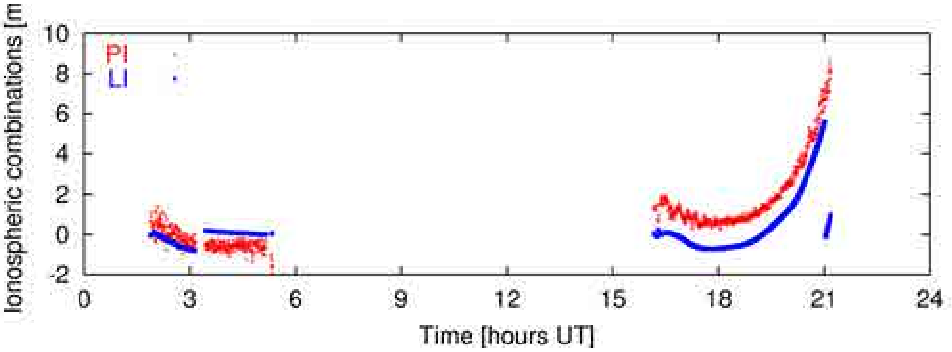


Figure A.3: *Cycle slips in GPS signal using Ionospheric combination of observables (defined later, it contains essentially the ionospheric delay and the instrumental delays and the additional phase ambiguity term in the case of LI). Note that the phase observable is typically reinitialised whenever a cycle slip or a new signal lock takes place. It can be seen as well that, although being ambiguous, the phase observable is much more precise than the code observable.*

Narrow and Wide lane combinations (P_n and L_δ, respectively): The wide lane combination is used for cycle-slips detection since it provides with an effective long wavelength of λ_δ=86.2 cm, which becomes very useful for this purpose. The L_δ and the corresponding combination for the code are constructed as follows:

$$P_n = \frac{f_1\cdot P_1+f_2\cdot P_2}{f_1+f_2} \qquad L\delta = \frac{f_1\cdot L_1-f_2\cdot L_2}{f_1-f_2} \tag{A.5}$$

Obtaining:

$$P_{n}{}_i^j = \rho_i^j + c(dt_i - dt^j) + rel_i^j + T_i^j + \alpha_\delta I_i^j + M^j_{P\delta,i} + \epsilon^j_{P\delta,i}$$
$$L\delta_i^j = \rho_i^j + c(dt_i - dt^j) + rel_i^j + T_i^j + \alpha_\delta I_i^j + B\delta_i^j + m^j_{L\delta,i} + \epsilon^j_{Lc,i}$$

Ionospheric (or geometric free) combination (L_I and P_I): It cancels all terms that do not depend on frequency such as geometric range, troposphere and so on, leaving the ionospheric contribution, instrumental biases and wind-up among others. This combination is constructed as follows:

$$P_I = P_2 - P_1 \qquad L_I = L_1 - L_2 \tag{A.6}$$

Obtaining:

$$P_{I}{}_i^j = \alpha_I I_i^j + K_{I}{}_i^j + M^j_{P_I,i} + \epsilon^j_{P_I,i}$$

$$L_{I}{}_i^j = \alpha_I I_i^j + B\delta_i^j - B2_i^j + m_{LI,i}^j + w_{L_I} + \epsilon_{L_I,i}^j$$

In order to respect the sign convention, the order of the observables is changed since the ionosphere causes a delay in the code and an advance in the phase in the same absolute amount. In ionospheric sounding the information given by the ionospheric (or geometric free) observable becomes essential, therefore the next section offers a deeper insight to this combination and the effect of ionosphere to GPS signals.

A.2.1 Ionosphere and GPS: L_I and P_I

The STEC plays a key role in determining the group and phase delay (I_{group} and I_{phase}, respectively) caused by the ionosphere to electromagnetic signals. A relationship between these quantities can be established (see [*Davies* (1990)] or [*Aragon-Angel* (2010)] for the details):

$$I_{group} = \frac{40.3}{f^2}\int_{Tx}^{Rx} N_e \, ds_0 \qquad I_{phase} = -\frac{40.3}{f^2}\int_{Tx}^{Rx} N_e \, ds_0 \tag{A.7}$$

Therefore, the ionospheric effect is equal in value for phase and group but of opposite sign, that is an advance in phase and a delay in pseudorange.

At this point, the Slant Total Electron Content or STEC is defined as the integral of the electron density along the signal path, that is:

$$STEC = \int_{Tx}^{Rx} N_e \, ds_0 \tag{A.8}$$

$$I_{group} = \tfrac{40.3}{f^2} \cdot STEC \qquad I_{phase} = -\tfrac{40.3}{f^2} \cdot STEC \tag{A.9}$$

being f the frequency, expressed in Hz, the STEC in $electrons/m^2$ and the ionospheric delay I expressed in units of metres of ionospheric delay.

According to each pseudorange observable (Definitions A.2 and A.3) and the definition of the ionospheric combination (Equation A.6), P_I and L_I observables can be modeled as:

$$\begin{aligned} P_I &= \alpha_I \cdot STEC + K_I + M_{P_I} + \varepsilon_{P_I} \\ L_I &= \alpha_I \cdot STEC + k_I + \lambda_1 \cdot N_1 - \lambda_2 \cdot N_2 + m_{L_I} + w_{L_I} + \varepsilon_{L_I} \simeq \\ &\simeq \alpha_I \cdot STEC + b_I \end{aligned} \tag{A.10}$$

where

$$\alpha_I = \alpha_2 - \alpha_1 = \frac{40.3}{f_2^2} - \frac{40.3}{f_1^2} \simeq 1.05 \frac{m_{L_I}}{10^{17} electron/m^2} = 0.105 \frac{m_{LI}}{TECU} \quad (A.11)$$

and b_I contains the contribution of both the instrumental delays and phase ambiguities. As in [*Blewit* (1989)], the terms due to noise, multipath and higher-order ionospheric terms (whose error is typically less than one centimetre) are not explicitly shown in the L_I expression, since the remaining terms are orders of magnitude larger.

Appendix B
SISTED outputs format

Output messages

The SISTED detector provides several output messages that are stored in one unique file per day. Every line added to this file in real-time, high-rate or post-processing modes will correspond to one of the three existing messages: DET_INF, I_PARAM and SF_WARN.

The message labelled as DET_INF includes the detector general information and the values being used for its adjustable parameters (see section 1.5.3). The message named I_PARAM provides the Impact parameters for each ionosphere SZA region r_i at the sampling rate time resolution and, finally, the SF_WARN message informs on SISTED detections. In this last case, the system does not only generate this message but also informs potential users by sending a warning mail automatically. For details on the format of each message, see Tables B.1, B.2 and B.3 below.

The three above-mentioned SISTED messages are encapsulated in one single file per day. The naming convention of such file is as follows: sisted.pp.messages.YYDOY, where YY and DOY correspond to the two-digit year and three-digit day of year, respectively.

DET_INF message

- **Purpose**
  This message is used to inform or update the general information of the detector as well as its main configurable parameters. Note that this message is self-explanatory.

- **Example**
  DET_INF 11 238 Vdrift|thres=0.00 I1|thres=0.74 I1/I3|thres=0.00 r1r2|szabound=70 r2r3|szabound=110 ele|thres=30 nrays_r1|thres=50 nrays_r2|thres=50 nrays_r3|thres=50

- **Update interval**
  One message at the beginning of SISTED execution (including each time the detector is restarted) and one after the change of the day being processed.

- **Description of fields**

Table B.1: *Format description of the fields included in the SISTED DET_INF message.*

#	Parameter	Meaning	Format	Examples
1	<type>	The following parameters describe the technique that is being used to detect solar flares	7 Characters, A7	DET_INF
2	<detector>	Name of the detection technique	7 Characters, A7	SISTED
3	<version>	Version of the program (including pp/hr/rt label to distinguish between post-processing/high-rate/real-time modes of operation)	Integer, I2.2. 3 Characters, A3	24.pp
4	<year>	Year	Integer, I2.2	99, 02, 11
5	<doy>	Day of Year	Integer, I3.3	001, 204, 366
6	<Vdrift$\|_{thres}$ >	Minimum $\Delta^2 V$ increase to consider a significant TEC increase	8 Characters, A8. Float, f4.2	Vdr|thres=0.10
7	$< I_1\|_{thres} >$	Minimum percentage of IPPs (in parts per one) detecting an increase to consider a generalized overionization in the sunlit ionosphere region r_1	8 Characters, A8. Float, f4.2	I1|thres=0.74
8	$< I_1/I_3\|_{thres} >$	Minimum ratio between the percentage of IPPs detecting an increase in the sunlit ionosphere region r_1 and night-side region r_3	12 Characters, A12. Float, f4.2	I1/I3|thres>=1.50
9	$< r_1 r_2\|_{szabound} >$	SZA value delimiting sunlit ionosphere region r_1 and dawn/dusk region r_2 (in degrees)	14 Charaters, A14 Integer, I3	r1r2|szabound=70
10	$< r_2 r_3\|_{szabound} >$	SZA value delimiting dawn/dusk ionosphere region r_2 and night-side region r_3 (in degrees)	14 Charater, A14 Integer, I3.3	r2r3|szabound=110
11	<ele$\|_{thres}$ >	Minimum receiver-satellite elevation	5 Characters, A5. Integer, I2	ele|thres=30
12	<nrays_r1$\|_{thres}$ >	Minimum number of receiver-satellite rays for the sunlit ionosphere region r_1	9 Characters, A9. Integer, I3	nrays_r1|thres=35, 50
13	<nrays_r2$\|_{thres}$ >	Minimum number of receiver-satellite rays for the dawn/dusk ionosphere region r_2	9 Characters, A9. Integer, I3	nrays_r2|thres=35, 50
14	<nrays_r3$\|_{thres}$ >	Minimum number of receiver-satellite rays for the night-side ionosphere region r_3	9 Characters, A9. Integer, I3	nrays_r3|thres=35, 50

I_PARAM message

- **Purpose**
  This message provides information about the ionospheric response to TEC variations, observed in $\Delta^2 V$, for each of the three SZA regions r_i. It is the input of the SISTED detector decision logic.

- **Example**
  I_PARAM 11 238 0.0833333333 209 61 0.291 241 64 0.265 182 52 0.285

- **Update interval**
  One message for each processed epoch (30 seconds in post-processing mode and one second in high-rate/real-time mode of operation).

- **Description of fields**

Table B.2: *Format description of the fields included in the SISTED I_PARAM message.*

#	Parameter	Meaning	Format	Examples
1	<type>	The following parameters describe the Impact Parameters for the different ionospheric regions as output of the detection technique	7 Characters, A7	I_PARAM
2	<year>	Year	Integer, I2.2	99, 02, 11
3	<doy>	Day of Year	Integer, I3.3	001, 204, 366
4	<thours>	Hours of day (GPS Time)	Float, f13.10	15.9333333333
5	<nrays_r1>	Number of GPS rays in the sunlit ionosphere region r_1	Integer, I3	130, 145
6	<nraysdet_r1>	Number of GPS rays in the sunlit ionosphere region r_1 detecting a TEC enhancement over $\Delta^2 V\|_{thres}$	Integer, I3	10,130
7	$< I_1 >$	Impact parameter, i.e. percentage (in parts per one) of rays affected by a TEC enhancement over $\Delta^2 V\|_{thres}$ in the sunlit ionosphere region r_1	Float. f5.3	0.118
8	<nrays_r2>	Number of GPS rays in the dawn/dusk ionosphere region r_2	Integer, I3	130, 145
9	<nraysdet_r2>	Number of GPS rays in the dawn/dusk ionosphere region r_2 detecting a TEC enhancement over $\Delta^2 V\|_{thres}$	Integer, I3	10,130

Table B.2: *Format description of the fields included in the SISTED I_PARAM message.*

#	Parameter	Meaning	Format	Examples
10	$< I_2 >$	Impact parameter, i.e. percentage (in parts per one) of rays affected by a TEC enhancement over $\Delta^2 V\|_{thres}$ in the ionosphere dawn/dusk region r_2	Float, f5.3	0.118
11	<nrays_r3>	Number of GPS rays in the night-side ionosphere region r_3	Integer, I3	130, 145
12	<nraysdet_r3>	Number of GPS rays in the night-side ionosphere region r_3 detecting a TEC enhancement over $\Delta^2 V\|_{thres}$ threshold	Integer, I3	10,130
13	$< I_3 >$	Impact parameter, i.e. percentage (in parts per one) of rays affected by a TEC enhancement over $\Delta^2 V\|_{thres}$ in the night-side ionosphere region I_3	Float, f5.3	0.118

SF_WARN message

- **Purpose**
  This message is used to warn of the existence of a SISTED solar flare detection. In this case, a warning will also be triggered, currently implemented as an automatic e-mail that is distributed internally.

- **Example**
  SF_WARN 03 301 11.0250000000 193 193 1.000 251 162 0.645 197 55 0.279 031028 110130

- **Update interval**
  One message for each epoch a SISTED detection arises.

- **Description of fields**

Table B.3: *Format description of the fields included in the SISTED SF_ WARN message. Note that the fields 2 to 13 are coincident with the ones described in I_ PARAM message. For this reason, only the new fields are described. It must be remarked that the additional fields are added since they are useful for the subsequent validation process.*

#	Parameter	Meaning	Format	Examples
1	<type>	The following parameters describe the information on the solar flare detection.	7 Characters, A7	SF_WARN
(. . .)	(. . .)	(. . .)	(. . .)	(. . .)
14	<yy><mm><dd>	Two-digit Year, Month and Day of the detection epoch (GPS time)	Integer, 3 x I2.2	031028,110410
15	<hh>mm><ss>	Two-digit Hour. Minute and Second of the detection epoch (GPS time)	Integer, 3 x I2.2	051030, 225900

Output plots

Several plots are automatically generated and made available through the SISTED FTP site. For instance, these plots show the evolution of the Impact parameters, the number of GPS rays taken into account or the IGS stations being processed. These plots are updated every 30 seconds (in post-processing mode) or one second (in high-rate/real-time modes) and are closed day by day. In case of SISTED detection, additional plots are also generated. Being YY the two-digit year, DOY the three-digit day of year, HH the two-digit hour, MM the two-digit month, SS the two-digit second and ZZ the two-digit value of the zoom being applied (00, 10 or 20 scale factors of y-axis component), detailed information on the main SISTED plots is included in Table B.4.

Table B.4: *Format description of the main SISTED plots including time of file generation, update period and file naming convention (D=Daily; SR=Sampling Rate; SF=Solar Flare event; MO=Mode of operation).*

Plot description	File Gen.	Update	File Naming
Impact Parameters evolution	D	SR	sisted.MO.I_PARAM.YYDOY.gif
Number of GPS rays	D	SR	sisted.MO.NRAYS.YYDOY.gif
IGS stations map	D	D	sisted.MO.STATIONS.YYDOY.gif
Sunlit IPPs det. positive $\Delta^2 V$	SF	SF	sisted.MO.SF_WARN.IPPs.YYDOY.HHhMMmSSs.gif
SZA against $\Delta^2 V$	SF	SF	sisted.MO.SF_WARN.SZAvsd2V.YYDOY.HHhMMmSSs.gif

Acronyms list

AIUB	Astronomical Institute of the University of Bern
BKG	Bundesamt für Kartographie und Geodäsie
BSS	Beacon Satellite Symposium
C/A	Coarse/Acquisition
CDDIS	Crustal Dynamics Data Information System
CME	Coronal Mass Ejection
CNES	National Center for Space Studies
CODE	Centre for Orbit Determination in Europe
COMPASS	Compass/Beidou: Big Dipper (constellation) in Chinese. The Chinese GNSS system
DCB	Differential Code Biases
DCT	Discrete Cosine Transform
EC	Electron Content
EGU	European Geosciences Union
ESA	European Space Agency
ESOC	European Space Operations Centre
ESTEC	European Space Research and Technology Centre
EUREF	European Reference Frame
EUV	Extreme Ultra Violet

FOC	Full Operational Capability
FTP	File Transfer Protocol
gAGE	Research group of Astronomy and Geomatics
GALILEO	European Global Navigation Satellite System
GIM	Global Ionospheric Maps
GLOBDET	GLOBal DETector
GLONASS	GLObal NAvigation Satellite System
GNSS	Global Navigation Satellite System
GOES	Geostationary Operational Environment Satellites
GPS	Global Positioning System
GSFLAI	GNSS Solar FLare Activity Indicator
HTML	Hyper Text Markup Language
IAAC	IGS Associate Analysis Center
IAG	Geodesy for Planet Earth
IDCT	Inverse Discrete Cosine Transform
IGS	International GNSS Service
IGS Iono-WG	IGS Ionosphere Working Group
IMF	Interplanetary Magnetic Field
IP	Internet Protocol
IPP	Ionospheric Pierce Point
IONEX	IONosphere map EXchange format
JPL	Jet Propulsion Laboratory
LC	Carrier Phase ionosphere-free combination
MATLAB	MATrix LABoratory
MEO	Medium Earth Orbit

MSTIDs	Medium Scale Travelling Ionospheric Disturbances
NASA	National Aeronautics and Space Administration
NAVSTAR	Navigation System with Time And Ranging
NGDC	National Geophysical Data Center
NOAA	National Oceanic and Atmospheric Administration
NTRIP	Networked Transport of RTCM via Internet Protocol
OS	Operative System
PC	Code ionosphere-free combination
PODAAC	Physical Oceanography Distributed Active Archive Center
RHESSI	Reuven Ramaty High Energy Solar Spectroscopic Imager
RINEX	Receiver INdependent EXchange format
RMS	Root Mean Square
ROC	Receiver Operational Characteristics
RTCM	Radio Technical Commission for Maritime
RTK	Real-Time Kinematics
Rx	Receiver
S/A	Selective Availability
SDO	Solar Dynamics Observatory
SEM	Solar EUV Monitor
SF	Solar Flare
SISTED	Sunlit Ionosphere Sudden TEC Enhancement Detector
SOHO	SOlar and Heliospheric Observatory
SMOS	Soil Moisture and Ocean Salinity
STEC	Slant Total Electron Content
SVM	Support Vector Machines

SWPC	Space Weather Prediction Center
SZA	Solar Zenith Angle
TEC	Total Electron Content
TECU	Total Electron Content Unit
TGD	Total Group Delay
TID	Travelling Ionospheric Disturbance
Tx	Transmitter
UPC	Technical University of Catalonia
USC	University of Southern California
UTC	Universal Time Coordinated
UV	Ultra Violet
VTEC	Vertical Total Electron Content
WAAS	Wide Area Augmentation System
WARN	Warning
WARTK	Wide Area Real-Time Kinematics
XRS	X-Ray Sensors

Bibliography

[*Afraimovich et al.* (2006)] Afraimovich, E. L., E. I. Astafyeva, and I. V. Zhivetiev (2006), Solar activity and global electron content, *Doklady Earth Sciences, 409A*, N6, 921-924.

[*Afraimovich et al.* (2000)] Afraimovich, E. L., E.A. Kosogorov, and L.A. Loenovich. (2000), The use of the international GPS network as the global detector (GLOBDET) simultaneously observing sudden ionospheric disturbances, Letter in: *Earth Planets Space*, 52: 1077-1082.

[*Ahmed et al.* (1974)] Ahmed, N., T. Natarajan, and K.R. Rao (1974), Discrete Cosine Transform, *IEEE Trans. Computers*, C-23, 90-93.

[*Aragon-Angel* (2010)] Aragon-Angel, A. (2010), Contributions to ionospheric electron density retrieval, *Ph.D. dissertation*. Doctoral Program in Aerospace Science & Technology, Universitat Politècnica de Catalunya, Barcelona, Spain.

[*Azpilicueta et al.* (2008)] Azpilicueta F, and C. Brunini (2008), Analysis of the bias between TOPEX and GPS vTEC determinations, *Journal of Geodesy, 83*, Issue 2, 121-127.

[*Bieber et al.* (2005)] Bieber, J.W., J. Clem, P. Evenson, R. Pyle, D. Ruffolo, et al. (2005), Relativistic solar neutrons and protons on 28 October 2003, *Geophysical Research Letters*, *32*, L03S02, doi:10.1029/2004GL021492.

[*Bilitza et al.* (2001)] Bilitza, D. (2001), International Reference Ionosphere 2000, *Radio Science*, *36*, 261-275, doi:10.1029/2000RS002432.

[*Bishop* (1995)] Bishop, C.M. (1995), Neural Networks for Pattern Recognition, *Oxford University Press*, Oxford.

[*Blewit* (1989)] Blewitt, G. (1989), Carrier phase ambiguity resolution for the global positioning system applied to geodetic baselines up to 2000 km, *J. Geophys. Res.*, *94(B8)*, 10187-10203.

[*Brekke* (2004)] Brekke, P. (2004), Space Weather Effects, *First European Space Weather Week (ESWW)*, ESA-ESTEC, Noordwijk.

[*Bust and Mitchell* (2008)] Bust G.S., and C. N. Mitchell (2008), History, current state, and future directions of ionospheric imaging, *Reviews of Geophysics*, *46*, RG1003, 1-23.

[*Cander et al.* (1998)] Cander, L. R., M. M. Milosavljevic, S. S. Stankovic, and S. Tomasevic (1998), Ionospheric forecasting technique by artificial neural network, *Electron. Lett.*, *34(6)*, 1573-l574.

[*Davies* (1990)] Davies, K. (1990), Ionospheric Radio, *Peter Peregrinus*, London.

[*Dick et al.* (1999)] Dick, M.I., M.F. Levy, Lj. R. Cander, I. Kutiev, and P. Muhtarov (1999), Short-term ionospheric forecasting over Europe, *IEEE Publication*, *461*, 105.

[*Dow et al.* (2009)] Dow, J.M., R.E. Neilan, and C. Rizos (2009), The International GNSS Service in a changing landscape of Global Navigation Satellite Systems, *Journal of Geodesy, 83*, 191-198.

[*Duda, Hart et al.* (2001)] Duda, R.O., P.E. Hart, D.G. Stork (2001), Pattern Classification, 2nd Ed., *John Wiley & Sons*.

[*Feltens* (2003)] Feltens, J. (2003), The activities of the ionosphere working group of the International GPS Service (IGS), *GPS Solutions*, *7(1)*, 41-46, doi:0.1007/s10291-003-0051-9.

[*Feltens et al.* (1999)] Feltens, J., Dow, J., Martin-Mur, T., Romero, I., and Garcia-Martinez, C. (1999), Routine production of ionosphere TEC maps at ESOC, in Proceedings of the *IGS Analysis Centers Workshop*, La Jolla, California.

[*Feltens and Schaer* (1998)] Feltens, J. and Schaer, S. (1998), IGS products for the ionosphere, in IGS Analysis Centers Workshop, Darmstadt, Germany.

[*Francis et al.* (2000)] Francis, N. M., P. S. Cannon, A. G. Brown, and D. S. Broomhead (2000), Nonlinear prediction of the ionospheric parameter on hourly, daily, and monthly timescales, *J. Geophys. Res.*, *105(A6a)*, 12839-12849.

[*García-Fernández* (2004)] García Fernández, M. (2004), Contributions to the 3D ionospheric sounding with GPS data, *Ph.D. dissertation.* Doctoral Program in Aerospace Science & Technology, Universitat Politècnica de Catalunya, Barcelona, Spain.

[*García-Rigo et al.* (2011)] García-Rigo, A., E. Monte, M. Hernández-Pajares, J. M. Juan, J. Sanz, A. Aragon-Angel, and D. Salazar (2011), Global prediction of the vertical total electron content of the ionosphere based on GPS data, *Radio Sci.*, *46*, RS0D25, doi:10.1029/2010RS004643.

[*García-Rigo et al.* (2010a)] García-Rigo, A., E. Monte , M. Hernández-Pajares, J.M. Juan, J. Sanz and R. Orús (2010), UPC VTEC Forecast Model Based On IGS GIMs, Proceedings of the *Beacon Satellite Symposium*, June 2010, Barcelona, Spain.

[*García-Rigo et al.* (2010b)] García-Rigo, A., E. Monte , M. Hernández-Pajares, J.M. Juan, J. Sanz and R. Orús (2010), UPC VTEC Forecast Model Based On IGS GIMs, Oral presentation in *Beacon Satellite Symposium*, June 2010, Barcelona, Spain.

[*García-Rigo et al.* (2009a)] García-Rigo, A., M. Hernández-Pajares, E. Monte, J.M. Juan, J. Sanz, A. Krankowski, and P. Wielgosz (2009), Assessment of UPC model for ionosphere VTEC prediction, Poster in *Geodesy for Planet Earth (IAG)*, August-September 2009, Buenos Aires, Argentina.

[*García-Rigo et al.* (2009b)] García-Rigo, A., E. Monte, M. Hernández-Pajares, J.M. Juan, J. Sanz, A. Krankowski and P. Wielgosz (2009), Prediction of Global Ionospheric TEC maps: First results on a UPC forecast product, Poster in *European General Assembly (EGU)*, April 2009, Vienna, Austria.

[*García-Rigo et al.* (2008a)] García-Rigo, A., M. Hernández-Pajares, J.M. Juan and J. Sanz (2008), Real Time Ionospheric TEC monitoring method applied to detect Solar Flares, *Geophysical Research Abstracts*, *10*, SRef-ID 1607-7962/gra/EGU2008-A-09358, EGU General Assembly 2008.

[*García-Rigo et al.* (2008b)] García-Rigo, A., M. Hernández-Pajares, J.M. Juan and J. Sanz (2008), Real Time Ionospheric TEC monitoring method applied to detect Solar Flares, Poster presentation in *EGU General Assembly*, April 2008, Vienna, Austria.

[*García-Rigo et al.* (2007a)] García-Rigo, A. , M. Hernández-Pajares, J.M. Juan and J. Sanz (2007), Solar flare detection system based on global positioning system data: First results, *Advances in Space Research*, *39*, 889-895, doi:10.1016/j.asr.2006.09.031.

[*García-Rigo et al.* (2007b)] García-Rigo, A., M. Hernández-Pajares, J. M. Juan, and J. Sanz (2007), Solar Flare detection system applied to the X65 flare on 6th December, 2006, Poster in *Beacon Satellite Symposium*, June 2007, Boston, United States of America.

[*García-Rigo et al.* (2007c)] García-Rigo, A., M. Hernández-Pajares, J.M. Juan, and J. Sanz (2007), Solar flare detection system using Global Positioning System data, Proceedings of the *7th Geomatic Week*, February 2007, Barcelona, Spain.

[*García-Rigo et al.* (2007d)] García-Rigo, A., M. Hernández-Pajares, J.M. Juan, and J. Sanz (2007), Solar flare detection system using Global Positioning System data, Oral presentation in *7th Geomatic Week*, February 2007, Barcelona, Spain.

[*Hastie et al.* (2001)] Hastie, T., R. Tibshirani, and J. Friedman (2001), The elements of statistical learning: data mining, inference and prediction, *Springer*.

[*Hernández-Pajares et al.* (2012)] M. Hernández-Pajares, A. García-Rigo, J.M. Juan, J. Sanz, E. Monte and A. Aragon-Angel (2012), High resolution measurement of Solar Flare EUV 2 photons flux rate from sparse global GPS networks, *Space Weather*, under review process.

[*Hernández-Pajares et al.* (2011)] Hernández-Pajares, M., J.M. Juan, J. Sanz, A. Aragón-Àngel, A. García-Rigo, et al. (2011), The ionosphere: effects, GPS modeling and the benefits for space geodetic techniques, *Journal of Geodesy*, *published on-line*, 1-21, doi:10.1007/s00190-011-0508-5.

[*Hernández-Pajares et al.* (2009)] Hernández-Pajares, M., J.M. Juan, J. Sanz, A. García-Rigo, J. Feltens, A. Komjathy, S.C. Schaer, and A. Krankowski (2009), The IGS VTEC maps: a reliable source of ionospheic information since 1998, Special IGS Issue of *Journal of Geodesy, 83:3-4*, 263-275.

[*Hernández-Pajares et al.* (2007)] Hernández-Pajares, M., J.M. Juan, J. Sanz, A. García-Rigo, R. Chen, X. Li, J. Talaya, E. Bosch, T. Schueler, E. Schueler, S. Soley and M. Reche (2007), WARTK-EGAL: WARTK Based on EGNOS and Galileo: Technical Feasibility Study, invited presentation at *Growing Galileo Conference & Exhibition*, European GNSS Supervisory Authority, November 2007, Brussels, Belgium.

[*Hernández-Pajares* (2004)] Hernández-Pajares, M. (2004), IGS Ionosphere WG status report: performance of IGS Ionosphere TEC Maps -Position Paper-, presented at *IGS Technical Meeting*, Bern, Switzerland.

[*Hernández-Pajares et al.* (2001)] Hernández-Pajares, M., J.M. Juan, and J. Sanz (2001), GPS data processing: code and phase. Algorithms, Techniques and Recipes, *CPET UPC*, Barcelona, Spain.

[*Hernández-Pajares et al.* (2000)] Hernández-Pajares, M., J.M. Juan, J. Sanz, and O.L. Colombo (2000), Application of ionospheric tomography to real-time GPS carrier-phase ambiguities resolution, at scales of 400-1000 km and with high geomagnetic activity, *Geophysical Research Letters*, *27:13*, 2009-2012.

[*Hernández-Pajares et al.* (1997)] Hernández-Pajares, M., J.M. Juan, and J. Sanz (1997), High resolution TEC monitoring method using permanent ground GPS receivers, *Geophys. Res. Lett.*, *24(13)*.

[*Hildebrand et al.* (1987)] Hildebrand, F.B. (1987), Introduction to numerical analysis, 2nd Ed., *Dover* books on mathematics, New York.

[*Hill et al.* (2005)] Hill, S. M., et al. (2005), The NOAA Goes-12 Solar X-Ray Imager (SXI): 1. Instrument, operations, and data, *Sol. Phys., 226*, 255-281.

[*Hoaglin* (1983)] Hoaglin, D., F. Mosteller, and J. Tukey (1983). Understanding Robust and Exploratory Data Analysis. *Wiley*, New York.

[*Hoffmann-Wellenhof et al.* (2004a)] Hoffmann-Wellenhof, B., H. Lichtenegger, J. Collins (2004), GPS: Theory and Practice, 4th Edition, *Springer-Verlag*, Wien New York.

[*Judge et al.* (1998)] Judge, D.L., D.R. McMullin, H.S. Ogawa, D. Hovestadt, B. Klecker, et al. (1998), First Solar EUV Irradiances Obtained from SOHO by the CELIAS/SEM, *Solar Physics*, *177(1-2)*, 161-173.

[*Kappenman* (2012)] Kappenman, J. (2012). A perfect storm of planetary proportions, *IEEE Spectrum*, 26-31, February 2012.

[*Klobuchar et al.* (1986)] Klobuchar, J.A. (1986), Design and Characteristics of the GPS Ionospheric Time delay Algorithm for Single Frequency Users, *IEEE Plans 1986 Position, Location and navigation Symposium*, Las Vegas, USA.

[*Krankowski et al.* (2008)] Krankowski, A., M. Hernández-Pajares (2008), Status of the IGS ionosphere products and future developments, presented at *IGS Analysis Center Workshop* 2008, Miami Beach, USA.

[*Krankowski et al.* (2005)] Krankowski, A., W. Kosek, L.W. Baran, and W. Popiński (2005), Wavelet analysis and forecasting of VTEC obtained with GPS observations over European latitudes, *J. Atmos. Solar-Terr. Phys., 67*, 1147-1156.

[*Last* (2010)] Last, D. (2010), GNSS - The Present Imperfect, *Inside GNSS*, May 2010, GNSS Forum, 5, 3, 60-64 (Invited).

[*Masson et al.* (2009)] Masson, S., K.L. Klein, R. Butikofer, E.O. Fluckiger, V. Kurt, et al. (2009), Acceleration of Relativistic Protons during the 20 January 2005 Flare and CME, Proceedings of the 31st *ICRC, ŁÓDŹ*.

[*Mendillo* (1974)] Mendillo, M., J.A. Klobuchar, R.B. Fritz, A.V. Da Rosa, L. Kersley, et al. (1974), Behavior of the ionospheric F region during the great solar flare of August 7, 1972, *J. Geophys. Res.*, *79*, 665-672.

[*Muhtarov et al.* (2001)] Muhtarov, P., I. Kutiev, Lj. R. Cander, B. Zolesi, G. de Franceschi, M. Levy, and M. Dick (2001), European ionospheric forecast and mapping, *Physics and Chemistry of the Earth*, Part C: Solar, Terrestrial and Planetary Science, *26*, Issue 5.

[*Muhtarov et al.* (1999)] Muhtarov, P., and I. Kutiev (1999), Autocorrelation method for temporal interpolation and short-term prediction of ionospheric data, *Radio Sci.*, *34(2)*, 459-464, doi:10.1029/1998RS900020.

[*Nava et al.* (2008)] Nava, B., P. Coïsson, and S. M. Radicella (2008), A New version of the NeQuick ionosphere electron density model, *J. Atmos. Sol. Terr. Phys.*, 1856-1862, doi:10.1016/j.jastp.2008.01.015.

[*Oppenheim et al.* (2010)] Oppenheim, A. V., and R. W. Schafer (2010), Discrete-Time Signal Processing, 3rd Ed., *Prentice Hall Signal Processing*.

[*OECD* (2011)] OECD (2011), Future Global Shocks: Improving Risk Governance, OECD Reviews of Risk Managelement Policies, *OECD Publishing*, http://www.oecd-ilibrary.org/governance/future-global-shocks_9789264114586-en

[*Orús et al.* (2010a)] Orús, R., M. Hernández-Pajares, J.M. Juan, J. Sanz, A. Aragón-Angel, and A. García-Rigo (2010), Real Time Application of TOMION Model, Proceedings of the *Beacon Satellite Symposium*, Barcelona, Spain.

[*Orús et al.* (2010b)] Orús, R., M. Hernández-Pajares, J.M. Juan, J. Sanz, A. Aragón-Angel, and A. García-Rigo (2010), Real Time Application of TOMION Model, Poster in *Beacon Satellite Symposium*, Barcelona, Spain.

[*Orús et al.* (2007)] Orús, R., L.R. Cander, M. Hernández-Pajares (2007), Testing regional vertical total electron content maps over Europe during the 17-21 January 2005 sudden space weather event, *Radio Science*, *42*, RS3004, doi:10.1029/2006RS003515.

[*Orús-Pérez* (2005)] Orús Pérez, R. (2005), Contributions on the improvement, assessment and application of the Global Ionospheric VTEC Maps computed with GPS data, *Ph.D. dissertation.* Doctoral Program in Aerospace Science & Technology, Universitat Politècnica de Catalunya, Barcelona, Spain.

[*Parkinson and Spilker* (1996)] Parkinson, B.W., Spilker Jr., J.J. (1996), Global Positioning System: Theory and applications. Volumes I & II, *American Institute for Aeronautics and Astronautics Inc.*

[*Pierrard and Stegen* (2008)] Pierrard, V., K. Stegen (2008), A three-dimensional dynamic kinetic model of the plasmasphere, *J. Geophys. Res.*, 113, A10209.

[*Ramos-Bosch* (2008)] Ramos Bosch, P. (2008), Improvements in autonomous GPS navigation of Low Earth Orbit satellites, *Ph.D. dissertation.* Doctoral Program in Aerospace Science & Technology, Universitat Politècnica de Catalunya, Barcelona, Spain.

[*Schaer et al.* (1998)] Schaer, S., W. Gurtner, and J. Feltens (1998), IONEX: The IONosphere Map EXchange Format Version 1, February 25, 1998, in Proceedings of the 1998 *IGS Analysis Centers Workshop*, ESOC, Darmstadt, Germany, 233-247.

[*Schaer et al.* (1999)] Schaer, S. (1999), Mapping and Predicting the Earth's Ionosphere Using the Global Positioning System, *Ph.D. dissertation.* Astronomical Institute, University of Berne, Switzerland.

[*Seeber* (1992)] Seeber, G. (1992), Satellite Geodesy, *Walter de Gruyter Inc.*

[*Sivestrin et al.* (2001)] Sivestrin, P., M. Berger, Y. Kerr, and J. Font (2001), ESA's second earth explorer opportunity mission: The soil moisture and ocean salinity mission SMOS, *IEEE Geosci. Remote Sensing Newslett.*, *118*, 11-14.

[*Stanislawska et al.* (2001)] Stanislawska, I., and Z. Zbyszynski (2001), Forecasting of the ionospheric quiet and disturbed f_0F2 values at a single location, *Radio Sci.*, *36*, 1065-1071, doi:10.1029/1999RS002242.

[*Swets* (1996)] Swets, J. A. (1996), Signal Detection Theory and ROC Analysis in Psychology and Diagnostics : Collected Papers. *Lawrence Erlbaum Associates*, Mahwah, NJ.

[*Tsurutani et al.* (2009)] Tsurutani, B.T., O.P. Verkhoglyadova, A.J. Mannucci, G.S. Lakhina, G. Li, et al. (2009), A brief review of solar flare effects on the ionosphere, *Radio Sci.*, *44*, RS0A17, doi:10.1029/2008RS004029.

[*Tsurutani et al.* (2005)] Tsurutani, B.T., Judge, D.L., Guarnieri, F.L., Gangopadhyay, P., Jones, et al. (2005), The October 28, 2003 extreme EUV solar flare and resultant extreme ionospheric effects: Comparison to other Halloween events and the Bastille Day event, *Geophys. Res. Lett.*, *32*, 3 L03S09, doi:10.1029/2004GL021475.

[*Tulunay et al.* (2006)] Tulunay, E., E. T. Senalp, S. M. Radicella, and Y. Tulunay (2006), Forecasting total electron content maps by neural network technique, *Radio Sci.*, *41*, RS4016, doi:10.1029/2005RS003285.

[*Valls-Moreno* (2008)] Valls-Moreno, A. (2008), Wide Area Real Time Kinematic (WARTK): Usage of RTCM format, and real-time implementation. *Master thesis.* Research group of Astronomy and GEomatics, Universitat Politècnica de Catalunya, Barcelona, Spain.

[*Wells* (1987)] Wells D. (1987), Guide to GPS positioning, *Canadian GPS associates*, Frederickton.

[*Woods et al.* (2004)] Woods, T.N., F.G. Eparvier, J. Fontenla, J. Harder, G. Kopp, et al. (2004), Solar irradiance variability during the October 2003 solar storm period, *Geophys. Res. Lett.*, 31, L10802, doi:10.1029/2004GL019571.

[*Zhang et al.* (2004)] Zhang, D.H., Xiao, Z., Liu, J., Liu, S., Gong, J. (2003), Sunlit boundary ionospheric response to the great flare on Oct. 28, 2003, *Chinese Science Bulletin*, *49(15)*, 1570-1574.

[*Zhang et al.* (2002)] Zhang, D.H., Z. Xiao, K. Igarashi, and G.Y. Ma. (2002), GPS-derived ionospheric total electron content response to a solar flare that occurred on 14 July 2000, *Radio Sci.*, 37(5), 1086, doi:10.1029/2001RS002542.