



Visual exploration and analysis of ionospheric scintillation monitoring data: The ISMR Query Tool



Bruno César Vani^{a,*}, Milton Hirokazu Shimabukuro^b, João Francisco Galera Monico^c

^a Graduate Program in Cartographic Sciences, Faculdade de Ciências e Tecnologia, UNESP – Universidade Estadual Paulista, Rua Roberto Simonsen, 305, 19060-900 P. Prudente, SP, Brazil

^b Department of Mathematics and Computing, Faculdade de Ciências e Tecnologia, UNESP – Universidade Estadual Paulista, Rua Roberto Simonsen, 305, 19060-900, P. Prudente, SP, Brazil

^c Department of Cartography, Faculdade de Ciências e Tecnologia, UNESP – Universidade Estadual Paulista, Rua Roberto Simonsen, 305, 19060-900, P. Prudente, SP, Brazil

ARTICLE INFO

Article history:

Received 5 October 2015

Received in revised form

20 July 2016

Accepted 25 August 2016

Available online 26 August 2016

Keywords:

Ionospheric scintillation

GNSS

Visual exploration

ABSTRACT

Ionospheric Scintillations are rapid variations on the phase and/or amplitude of a radio signal as it passes through ionospheric plasma irregularities. The ionosphere is a specific layer of the Earth's atmosphere located approximately between 50 km and 1000 km above the Earth's surface. As Global Navigation Satellite Systems (GNSS) – such as GPS, Galileo, BDS and GLONASS – use radio signals, these variations degrade their positioning service quality. Due to its location, Brazil is one of the places most affected by scintillation in the world. For that reason, ionosphere monitoring stations have been deployed over Brazilian territory since 2011 through cooperative projects between several institutions in Europe and Brazil. Such monitoring stations compose a network that generates a large amount of monitoring data everyday. GNSS receivers deployed at these stations – named Ionospheric Scintillation Monitor Receivers (ISMR) – provide scintillation indices and related signal metrics for available satellites dedicated to satellite-based navigation and positioning services. With this monitoring infrastructure, more than ten million observation values are generated and stored every day. Extracting the relevant information from this huge amount of data was a hard process and required the expertise of computer and geoscience scientists. This paper describes the concepts, design and aspects related to the implementation of the software that has been supporting research on ISMR data – the so-called ISMR Query Tool. Usability and other aspects are also presented via examples of application. This web based software has been designed and developed aiming to ensure insights over the huge amount of ISMR data that is fetched every day on an integrated platform. The software applies and adapts time series mining and information visualization techniques to extend the possibilities of exploring and analyzing ISMR data. The software is available to the scientific community through the World Wide Web, therefore constituting an analysis infrastructure that complements the monitoring one, providing support for researching ionospheric scintillation in the GNSS context. Interested researchers can access the functionalities without cost at <http://is-cigala-calibra.fct.unesp.br/>, under online request to the Space Geodesy Study Group from UNESP – Univ Estadual Paulista at Presidente Prudente.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The ionosphere is the upper layer of the atmosphere located approximately between 50 km and 1000 km above the Earth's surface. Ionospheric scintillations are rapid changes in phase and/or amplitude of a radio signal as it propagates through small-scale plasma density irregularities in the ionosphere (Conker et al., 2003). The scintillation can degrade signals from Global

Navigation Satellite Systems (GNSS), such as the American GPS, the Russian GLONASS, the Chinese BDS and the European Galileo, among other related systems. The occurrence and intensity of scintillations vary according to several factors, such as solar activity, the epoch of the year, geographic location and local time. Since the Brazilian region is one of those most affected by scintillations – mainly at periods of high solar activity – monitoring stations have been established over Brazilian territory since February 2011 configuring a data monitoring network infrastructure – the so-called CIGALA/CALIBRA Network. Monitoring stations from the network were deployed in the context of the Projects CIGALA and CALIBRA funded by the European Commission (EC) in the

* Corresponding author.

E-mail address: brunovani22@gmail.com (B.C. Vani).

framework of the FP7-GALILEO-2009-GSA and FP7-GALILEO-2011-GSA-1a, respectively. Each monitoring station has a GNSS receiver that is characterized as an 'Ionospheric Scintillation Monitor Receiver' (ISMR). Such specific kinds of receiver provide scintillation indices and additional signal metrics that characterize the scintillation occurrence, the so-called ISMR data. In the current monitoring infrastructure, a large amount of data is generated by the monitoring receivers. Samples of more than sixty parameters and indices are obtained every minute, totaling more than ten millions of observations every day.

In 2011, with the deployment of the first monitoring station in the city of Presidente Prudente, São Paulo State, Brazil, in the absence of a computational resource to support access to the ISMR data, users had to download several files and seek the parameter and indices of interest manually. Such tasks were very time consuming and increased error susceptibility. The software previously available to deal with ISMR data was designed to work only with hourly files provided by a single monitoring station and it was not possible users to obtain a wider overview of the scintillation scenario, both at temporal and spatial domains. In this paper, we describe the design and implementation of a web based software for visual exploration and analysis of the ISMR data, the so-called "ISMR Query Tool", as well as the related theoretical concepts. The software constitutes an analysis infrastructure to complement the monitoring one, and aims to extend the possibilities of analysis of ISMR data by applying information visualization and time series mining techniques, deriving benefits from a sense-making loop (Keim et al., 2008). Experiments conducted with the support of our approach are also presented, highlighting the capabilities of the data visualization and time series mining techniques available in the system.

Since it was released in 2011, the ISMR Query Tool has been used to support tasks on several topics related to the Ionosphere, mainly in the context of GNSS. The user's segment of ISMR data is mainly characterized by two groups: professional users of GNSS Positioning and researchers on GNSS and related topics. The former is mainly characterized by applications on surveying and navigation under several domains, like Agriculture and Offshore. Several problematic issues occur due to ionospheric scintillations during the realization of their activities. Such phenomena can degrade the quality of the positioning service, causing undesirable situations. The ISMR data supports this group of users by providing information about the occurrence and intensity of scintillation events, they can therefore plan strategies to deal with this issue. The latter is characterized by several topics of research related to GNSS and Ionosphere. We highlight the research themes related to scintillation mitigation on GNSS positioning (see Aquino et al. (2009)) and research under Space Weather domain.

At the time of writing this paper, July 2016, more than 30,000 outputs had been provided by the ISMR Query Tool in the previous five years, emphasizing its contribution to the scientific community by making the process of accessing data and extracting useful information more efficient. Such outputs include plots, maps, reports and interactive visualizations describing the occurrence of ionospheric scintillations.

Data visualization and data mining techniques make a significant contribution in dealing with large amounts of data. In the scope of this project, information visualization and time series mining techniques were used and adapted whenever necessary. Visualization takes advantage of the abilities of humans to find structural patterns from images, while data mining refers to the process of seeking and discovering knowledge from large amounts of data. It is desirable that the user interacts with the tasks performed by the system by choosing configuration parameters and checking results presented through visual inspection. For such purposes, the user can drive the process by choosing deeper or

general analysis through their own perspectives. With these features, the system acts as a knowledge discovery tool (Shneiderman, 2001; Müller and Schumann, 2003; Han and Kamber, 2006). Related work can be found in Darbeheshti et al. (2013), where the design of a web application to visualize data of the Earth's gravity field is presented; at Zhang et al. (2007), where an approach to visualization and analysis of large spatial data sets is presented; and at Románach et al. (2015), in which a cross-platform application designed to help users explore gridded data is presented.

The rest of this paper is organized as follows. Section 2 contextualizes GNSS and ionospheric scintillation, as well as the monitoring infrastructure. Section 3 presents the implementation of the ISMR Query Tool, showing its components, the information visualization and time series mining techniques applied, its user interfaces, and its results through examples of applications. Conclusions are presented in the Section 4. The current version of our software is available by authorized login at <http://is-cigala-calibra.fct.unesp.br/>.

2. GNSS and ionospheric scintillations

GNSS is a general name given to satellite-based navigation systems. In general, these systems are characterized by constellations of satellites that allow users to obtain positioning through range measurements between the antenna of the satellite and the user receiver. Several systematic errors affect GNSS observations, such as satellite clock errors, atmospheric refraction, multipath of signals, hardware errors and geodynamic effects (Seeber, 2003; Monico, 2008).

As mentioned in the Introduction section, GNSS signals are affected by ionospheric scintillations. Among other aspects, amplitude scintillations can drop the signal intensity below a receiver's lock threshold, and phase scintillations may cause cycle slips and lead the receiver to lose track of one or more satellites. In critical circumstances, e.g. when a minimum of four satellites are not successfully tracked by the receiver, the positioning service can be lost. Parameters indicating these variations can be computed with data collected by GNSS monitoring receivers. Amplitude scintillations are mainly quantified by the S4 index, while phase scintillations are quantified by σ_ϕ (Sigma-phi) indices (Van Dierendonck et al., 1993; Datta-Barua et al., 2003).

The occurrence and the intensity of scintillations are related to several factors, such as signal frequency, geographic location, local time, the epoch of the year and the period of the solar cycle (approximately every 11 years). Due to its location around the geomagnetic equator, Brazil is one of world regions most affected by scintillations. Such conditions motivate the deployment of the monitoring network presented in Section 2.1. In Brazil, large plasma depletions (regions with depleted density of the ionosphere), also called plasma bubbles, are generated around the magnetic equator after sunset. The bubbles can vary in extension and drift velocity. Rezende et al. (2007) have shown that, in Brazil, the plasma bubbles can reach an east-west size of around 480 km normally moving in an east-west direction with a velocity of about 150 m/s at the pre-midnight local time. Small-scale plasma irregularities (in the order of few hundred meters) embedded in these large plasma bubbles lead to scintillations in the signal. Plasma bubbles chiefly occur between September and March during solar peaks. Such a dynamic scenario leads to difficulties in dealing with scintillations, since the range of measurements made by the GNSS receiver from different satellites can be degraded differently according to how the signal propagation path is being affected by the irregularities in the ionosphere that cause scintillations (Sreeja et al., 2011).

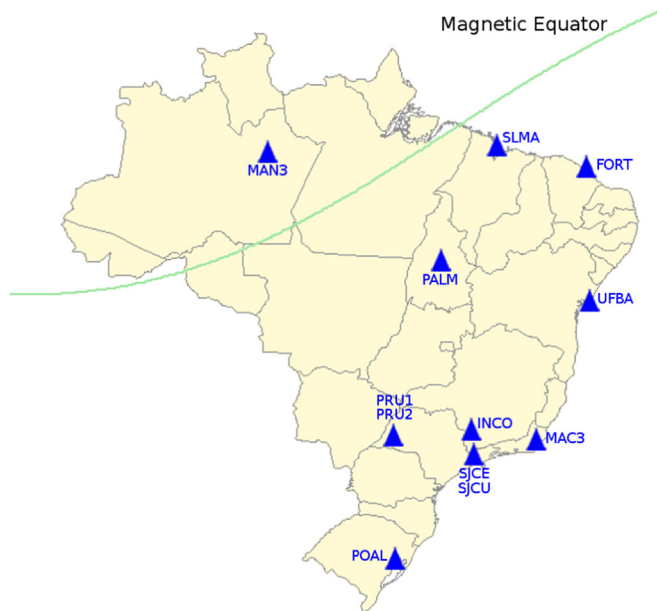


Fig. 1. CIGALA/CALIBRA network: monitoring stations locations over Brazilian territory. Status in June/2016.

2.1. The CIGALA/CALIBRA monitoring network and the ISMR data

Monitoring receivers that provide scintillation indices like S_4 and σ_ϕ , as well as other signal metrics, have been deployed over Brazilian territory since February 2011, mainly in the context of the international projects CIGALA (*Concept for Ionospheric Scintillation Mitigation for Professional GNSS in Latin America*) and CALIBRA (*Countering GNSS high Accuracy applications Limitations due to Ionospheric disturbances in BRAZIL*). Both projects were designed in the context of the development of approaches to mitigate the effects of ionospheric scintillation on GNSS positioning and navigation.

Fig. 1 shows the distribution of these monitoring stations over Brazilian territory, composing what is called the CIGALA/CALIBRA network. More than ten million monitoring observations are obtained from these stations every day, making it difficult to extract specifically relevant information.

The monitoring stations comprise a GNSS monitoring receiver (ISMR), a GNSS antenna and a computer for storing the data and managing the receiver configuration. Hourly raw files, containing all the high rate (50 Hz) data obtained by the receiver, and hourly ISMR files, containing 62 ionosphere monitoring indices computed every minute, are preliminarily stored at this local computer. After that, both files are sent to a central repository located at the “Faculdade de Ciências e Tecnologia, UNESP - Universidade Estadual Paulista, Presidente Prudente”, state of São Paulo, Brazil. The parameters provided by the those receivers – built by Septentrio – from the CIGALA/CALIBRA Network that describe the scintillation occurrence during the track of a GNSS satellites include: time labels, satellite identification, indicators of direction (azimuth and elevation angle of the satellites), time of synchronism between satellites and receiver, phase spectral slope (p), phase spectral strength (T), absolute and differential total electron content at the slant trajectory of the signal (TEC and $dTEC$, respectively), averaged signal-to-noise ratio, average and standard deviation of the code and carrier divergence, and finally, indices that describe the scintillation levels, such as the widely used S_4 (for amplitude scintillations) and σ_ϕ (for phase scintillations). The parameters are available for each frequency used by the GNSS satellites, totaling

62 parameters recorded each minute.

All measurements are time labeled for each full minute. These samples come from observations obtained for each satellite tracked by each receiver during the period of one minute, but being collected at a high rate (50 Hz). Considering the status of the network with twelve fixed monitoring stations over Brazilian territory (as presented in Fig. 1), and considering also the current satellite constellations available (GPS, GLONASS, Galileo, BDS and also other augmentation systems SBAS), more than ten million observations are collected every day. This scenario motivated the development of this research, since it was very difficult to extract specifically relevant information from such a large amount of data. For instance, the simple task of trying to find a period of high scintillation occurrence on a specific monitoring station during a week was very difficult due to the necessity of manipulating hundreds of hourly ISMR files. In this research, data visualization and data mining techniques were first investigated and later implemented to support users of ISMR data, resulting in the implementation of the ISMR Query Tool. Outcomes of the investigation and the implementation aspects are presented in Section 3.

3. Results: the ISMR Query Tool

The designed and implemented integrated software platform – the ISMR Query Tool – aims to aggregate information from all monitoring stations and provide mechanisms of analysis for the ISMR data stored at our central facilities servers. The general objective is to apply data visualization and data mining techniques that support exploratory data analysis and knowledge discovery from the monitoring data. The users group of this software is comprised of the users of GNSS for professional applications, like surveying and navigation, and by researchers from many areas, mainly in the context of GNSS and/or space and weather. Several versions of the software have been released since October 2011 contributing to several research projects applying ISMR data.

For many of the monitoring parameters, like the S_4 index, there are different empirical/theoretical threshold values defined by the scientific community. Similarly, for convenience of analysis, it is common not to include satellites tracked under low elevation angles, such as 5° or 15° . Taking account of such peculiarities when dealing with ISMR data, we followed the aspects highlighted by Shneiderman (2001) when considering the development of tools to integrate data mining and information visualization. The main aspect is to allow users to specify what they are seeking according to their own perspectives. They are therefore able to test their hypothesis and to develop a deeper understanding of their data. The other aspect followed is related to the limits that might be stated: where operations are very complex, users may be led to avoid using them. It is therefore important to consider the human nature of the analyst by establishing an acceptable level of interaction between the user and the tool. Following such aspects, we combined data mining and data visualization to design a discovery tool applied to analyzing the ionosphere monitoring data – the ISMR Query Tool. The System Components of the tool are presented in Section 3.1. Information visualization and time series mining approaches employed by the tool are presented in Section 3.2. Interfaces available to users are described in Section 3.3, followed by examples of application in Section 3.4.

3.1. System components

The monitoring data contain data generated every minute, therefore characterizing the time oriented data. The discrete label time is recorded at each full minute, and the ionosphere

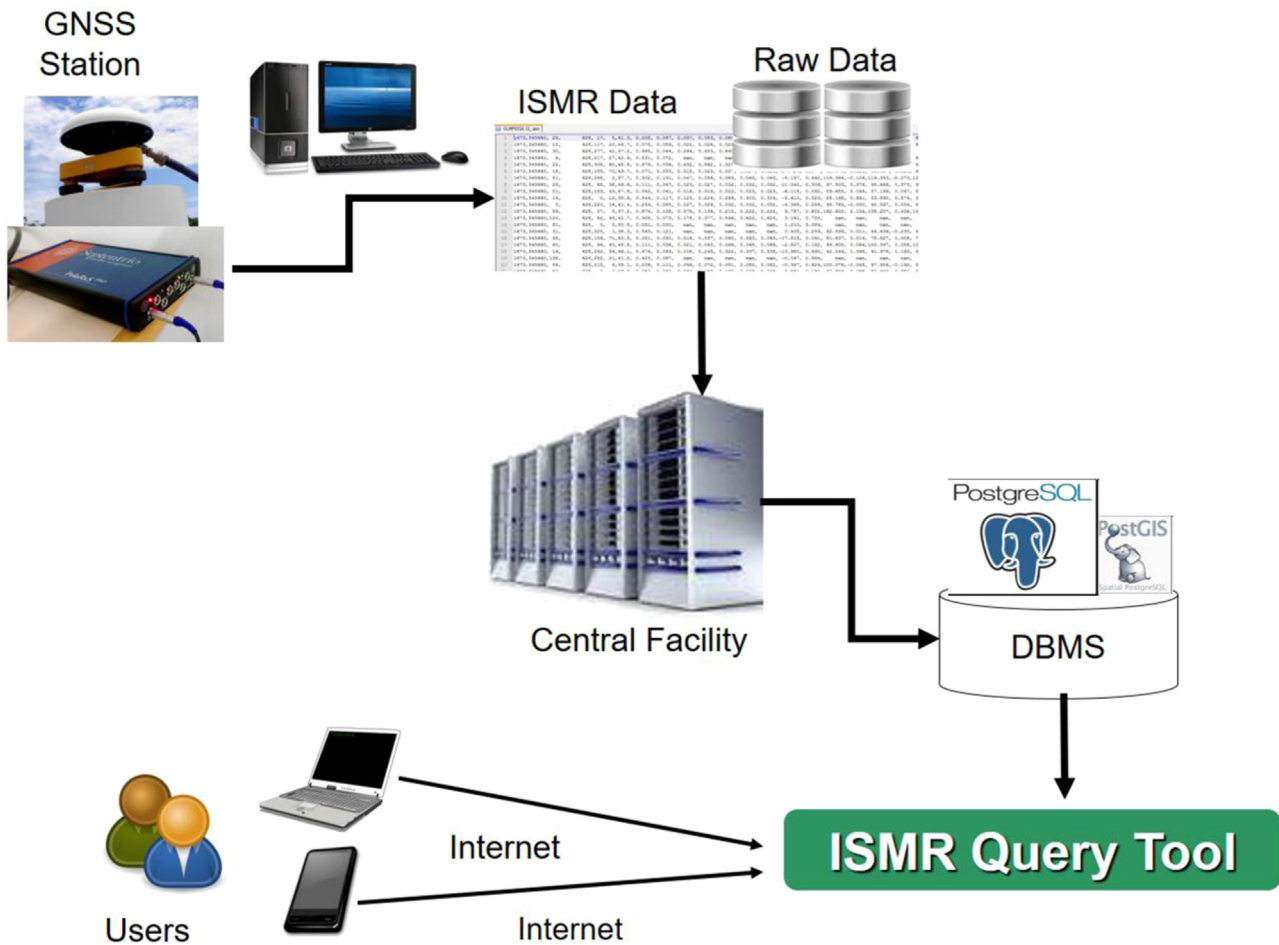


Fig. 2. System components and their logical sequence.

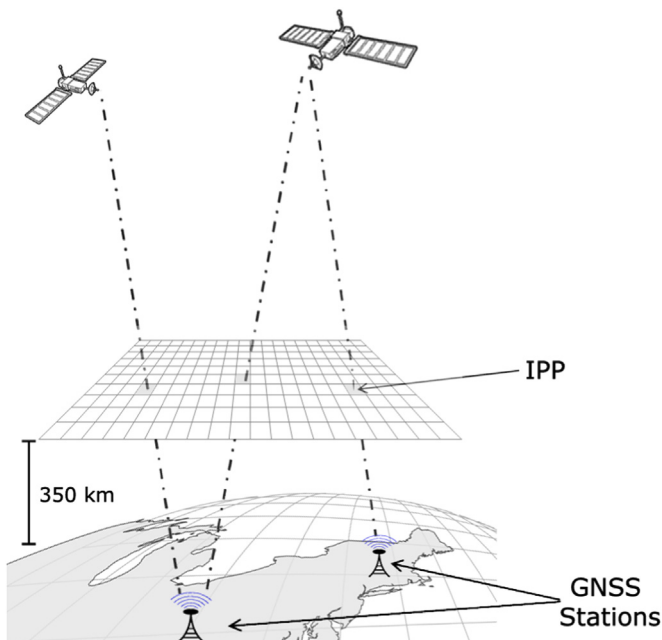


Fig. 3. Ionospheric Pierce Point (IPP) representation considering a regular grid at some altitude of the Ionosphere (e.g. 350 km). Based on Rezende et al. (2007).

monitoring data contain a subset composed of 62 monitoring variables (as mentioned on Section 2.1). Among the observations, we also have spatial attributes, such as the positions of stations.

These features evidenced the need for a spatial database modeling, having time as one attribute. This was modeled through the relational Database Management System (DBMS) PostgreSQL with the spatial extension PostGIS. We noticed that, at the current stage of development, the ISMR Query Tool does not include topological operation functionalities over geographic features. Therefore, this sort of resource is a subject for future work.

The relationship between the system components is illustrated in Fig. 2. The monitoring records obtained by the GNSS antenna and receiver are stored preliminarily on a local computer connected to each station, including raw data and ISMR data with hourly files. Following this step, the both hourly files are sent to a central facility located at UNESP. From the repository at UNESP, data are automatically imported to the DBMS in order to constitute the database for the ISMR Query Tool, which is made available to the users via the Web.

3.2. Information visualization and time series mining techniques employed at the ISMR Query Tool

In this section we present concepts regarding information visualization and the time series mining techniques selected to support the exploration and analysis of ISMR data in an integrated software platform. We focused on approaches applicable to temporal and spatial data. Regarding the temporal domain, we considered the time label recorded with each set of parameters, constituting the sample time series obtained every minute. Concerning the spatial domain, we considered the geographic location of the monitoring receivers, as well as the signal propagation path,

i.e., the projection of the trajectory of the signal arising from the satellite and reaching to the receiver on the ground (see Fig. 3). We chose techniques that are applied to the temporal and spatial domains of the monitoring data, considering the domains either individually or concurrently.

The first technique we describe is the Calendar View (Van Wijk and Van Selow, 1999), which was selected to allow the visualization of the time series of scintillation monitoring data through a general perspective. In this approach, as a first step, clusters are formed by using an appropriate cluster algorithm. After that, the clusters are displayed over a calendar where the daily data values are represented. Each day of the calendar is colored according to the cluster it belongs to. User interaction is introduced to ensure effective data exploration. The use of a calendar based visualization allows one to check large amounts of univariate time series data (Müller and Schumann, 2003). It can also be seen that the user benefits from an interface metaphor by using a calendar based representation. In the ISMR Query Tool, Calendar Views are applied mainly to multi-annual analysis of scintillation indexes, allowing one to get a wide overview of scintillation occurrence over yearly scales.

Next to the Calendar View, the traditional scatterplot (or xy-plot) was chosen to allow one to analyze time series of monitoring attributes, with the time attribute being represented along the x-axis. The effectiveness of this visualization approach can be seen in the benefits for exploratory data analysis and the contribution regarding human ability to judge information of attributes over a delimited space, as stated in Reimann et al. (2011) and Ward et al. (2010). The scatter plots allow one to identify duration and peaks of scintillation, as well as visually identify the most affected satellites.

In order to allow simultaneous graphical or visual representation of multiple time series, allowing the visualization of a quantity of scintillation monitoring information at the same time, we selected the approach of the Horizon Charts (Few, 2008). This kind of representation combines position and colors aiming to increase human perception in analyzing data. One can also highlight the effective use of horizontal lines allowing the representation of several time series in the same visualization scheme; time is represented on the x-axis and several attributes can be represented on the y coordinate axis. Horizon charts allow the users to analyze several time series of monitoring attributes simultaneously.

The Symbolic Aggregation Approximation (SAX) approach (Lin et al., 2003) was chosen in order to allow the dimensionality and numerosity reduction of the ISMR data. With this approach, as the first step, a time series C of length n is reduced from its original length n to w dimensions through the Piecewise Aggregation Approximation (PAA) approach. At the PAA step, the time series is divided into w equal-sized segments; values belonging to each segment are approximated and replaced by a single coefficient (their average). After that, the PAA representative coefficients are mapped to symbols. Breakpoints (that act as range limits) are used to divide the distribution of the coefficients into a subset of α equiprobable regions. The α parameter delimits the desired size of the alphabet used to map the coefficients from PAA. For instance, with $\alpha = 3$, the alphabet of symbols $\{a, b, c\}$ is used to map the coefficients according to the breakpoint values. The SAX approach is mainly applied on classification tasks (such as labeling the scintillation level as weak, moderate or strong according to the value of S_4), allowing the user to perform queries having benefits of dimensionality and numerosity reduction.

Aiming to consider both spatial and temporal domains, we emphasize the concept of Ionospheric Pierce Point (IPP) which can be used to represent ISMR data on maps. Firstly, a regular bi-dimensional grid of an arbitrary resolution (for instance, $2 \times 2^\circ$ of geographical latitude and longitude) is established over the Earth's

surface in the Ionosphere, for instance at an altitude of 350 km. Afterwards, one can use the projection of the signal propagation path on this grid (by supposing a straight propagation line) and represent the desired monitoring attribute on this crossed cell. This scheme is depicted in Fig. 3.

As GNSS satellites are continually orbiting through its specific orbital plane, the projected IPP changes over time. Therefore, in order to consider the both spatial and temporal attributes, we can project the IPP trace over time on the map or use aggregation functions, like mean or standard deviation, in order to accumulate information observed over time at the same IPP. In the ISMR Query Tool, the IPP maps allow the users to produce maps considering the spatial and temporal domains of any of the monitoring attributes.

3.3. Interfaces available to the users

Information visualization and mining modules were implemented in such way that the satellites most affected by scintillations can be found in time and space perspectives by analyzing the appropriate monitoring indices available in the database. One of the most important features is the possibility of applying and combining filters with Boolean operators at any of the available monitoring indices. Consequently, the user has the benefit of a sense-making loop as defined by Keim et al. (2008), generalizing and specifying queries according to his/her own interest. For instance, the user can test different thresholds applicable for the S_4 index, choose different elevation cut-offs for specific satellites, establish criteria for outlier detection, and so on.

Fig. 4 presents one example of a filtering operation using the ISMR Query Tool, available for users with basic knowledge of GNSS and ionospheric scintillation. On the left, the user can choose a period of interest (such as a few hours or one week), one or more monitoring stations, full constellations (like GPS and GLONASS) or few specific satellites and apply as many Boolean filters as he/she desires. On the right, the user can select specific output options. In this case, for the scatter plot output, options like aggregation configuration for the time axis and output parameters like size, labels and color schemes are available.

3.4. Examples of application

Using scatter plot representation, one can easily find a subset of data featured by specific scintillation scenario, for instance a behavior of daily peaks of high S_4 , and identify most affected satellites by analyzing different indexes. One example of the result is presented in three steps in Fig. 5. Firstly, in step (a), we show the S_4 index considering all the GPS satellites tracked by a receiver located at Presidente Prudente (PRU2 station, see Fig. 1) for the whole month of September 2012, by applying an elevation cut-off filter of 15° . Secondly, in step (b), we restrict this query for a period of about ten days and use a color identification plus geometric icon resource enabling individual identification of the most affected satellites; considering the GPS, each satellite is uniquely identified by its PRN number that ranges from 1 to 32. It can be seen that each satellite can be distinctly affected by the scintillations since its signal propagates through different regions on the ionosphere. Finally, in step (c), we compare two monitoring indices in order to demonstrate the capability on visualization other parameters – the S_4 and the Sigma-phi are shown restricting the analysis for only the three most affected satellites (visually identified on the second step).

Besides the example shown in Fig. 5, the scatter plot graphs can also be used to compare indices tracked at the same time by different stations. Exemplifying, one can analyze the correspondence of indices considering the same satellites tracked by neighbor

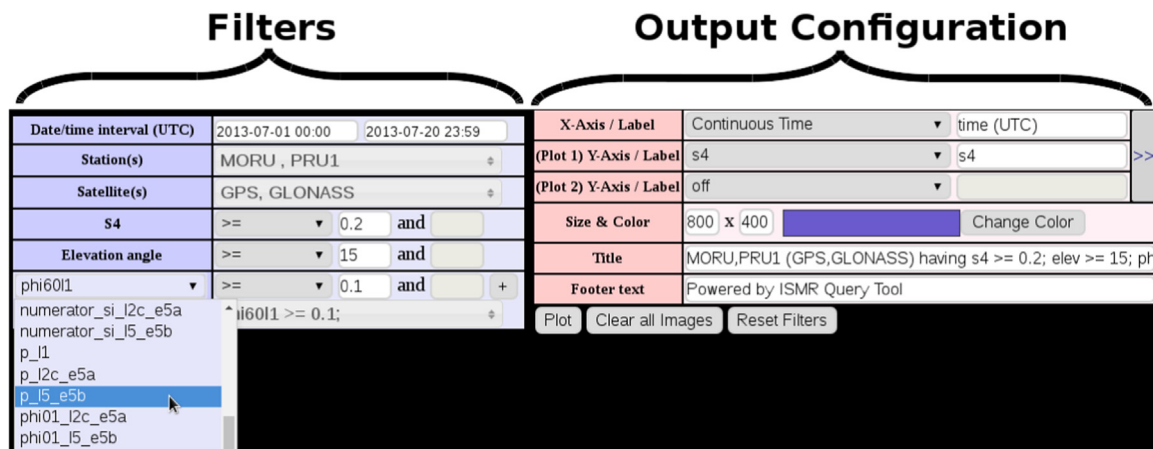


Fig. 4. Example of filters based interface (left) and output configuration (right) for the scatter plots at the ISMR Query Tool.

receivers and make inferences about the motion of the irregularities that cause scintillations. Another application of the scatter plot representation is to identify the most affected time of day along a continuous subset of time (like a month). This allows identification of the most susceptible time of scintillation occurrences at different regions covered by the network of monitoring stations.

In Fig. 6 we present an example of the Calendar View approach. This technique was applied to generate calendars mainly for the averaged daily value of the S4 index; calendars are automatically generated for each single station and also for the average considering all the stations, as presented in the Fig. 6 for the years of 2011, 2012 and 2013. Results allow one to quickly obtain a general overview of scintillation behavior over the years – with predominant occurrence, in general, from September to March. One can also analyze the individual calendar view for single stations and observe peculiar aspects of them considering the different station locations.

The spatial location of scintillation occurrence can be observed in the map presented in Fig. 7. This map shows the standard deviation of the S4 index by using the Ionospheric Pierce Point (IPP) representation over a period of fifteen days (the first half of December 2013) considering seven monitoring stations. The standard deviation is used as an aggregation function on the IPP in order to allow one to note the regions where this monitoring index presented higher variation, therefore indicating the regions of greatest susceptibility for the occurrence of scintillations. As expected, the results demonstrated consonance with the aspects highlighted in the literature: these regions lie mainly about 15° below or above the geomagnetic equator (gray dashed line). The small gray triangles identify the location of monitoring stations.

Besides the example presented in Fig. 7, the IPP maps can also be applied to analyze the spatial coverage of the network over the period, the repeatability of patterns of indices by using threshold values, the spatial behavior of scintillations and the geometry of the satellites.

Fig. 8 presents one example of the use of horizon charts where the selected attributes share the time axis of one day. Eleven attributes were selected for the satellite Galileo PRN 20 tracked by the FORT station (located at Fortaleza, state of Ceará) on January 23, 2014 from 12:00 to 00:00 (UTC). The mouse interaction allows the user to obtain details about the time series of each attribute. We demonstrate such features by clicking to evidence data observed at 09:50 p.m., that is highlighted by a vertical rule. The station, the name of the attribute, the satellite identification and the value of the attribute are then displayed.

In summary, the horizon charts can also be applied to any kind

of comparison among time series: one can compare attributes from different stations tracked by different satellites on different days. One example of this application is the detection of sequential patterns on data before or after the occurrence of scintillations.

For convenience of analysis, thresholds can be applied to classify the scintillation level with labels (such as weak, moderate or strong scintillation according to the S4 value). The International Telecommunication Union (ITU) Recommendation P.531-11 (“Ionospheric propagation data and prediction methods required for the design of satellite services and systems”) indicates the labels ‘weak scintillation’ for S4 less than 0.3, ‘moderate scintillation’ for S4 between 0.3 and 0.6, and ‘strong scintillation’ for S4 greater than 0.6. In Fig. 9 we present one example of our approach to analyzing the ISMR data through the SAX time series mining approach by applying such classification levels. The SAX technique was adapted and applied over the S4 index considering all GPS satellites tracked in the first half of December 2012 at the PRU1 station (located in the city of Presidente Prudente, state of São Paulo). First, the time series of each GPS satellite was reduced through the PAA approach. We chose windows of 15 min and applied the concept of sliding windows with an overlap of 5 min. In order to map the PAA coefficients of each time series into symbols according to the SAX approach, breakpoints were chosen according to the ITU classification thresholds for the scintillation levels: weak, moderate or strong scintillations. PAA coefficients of each time series were mapped to a symbol by using $\alpha = 3$, with symbol ‘a’ for weak scintillation, ‘b’ for moderate scintillation, and ‘c’ for strong scintillation.

For the visualization scheme, symbols are mapped to colors according to the corresponding intensity levels of the color scale (color scale by Harrower and Brewer (2003)). The steps of an interactive visualization session are presented in Fig. 9. The ‘Summary’ section provides an overview of the results, displaying the absolute and relative number of windows classified with each symbol with support of vertical bars. The detailed results are presented for each satellite time series through interactive horizon bar segments. The satellite identification is presented on the left side of each time series. The satellites are vertically ordered by a specific chosen parameter. The value of chosen parameter is displayed at the end of the time series. Available sort options are the satellite time tracked or the absolute (or relative) number of occurrences of a given symbol, like ‘a’, ‘b’ or ‘c’ for this example case. Mouse interaction provides details for a selected satellite or specific duration of a window.

We emphasize that in this visualization we chose not to match the temporal axis for all satellites since the time range over which each satellite is tracked by a receiver would lead to several gaps on

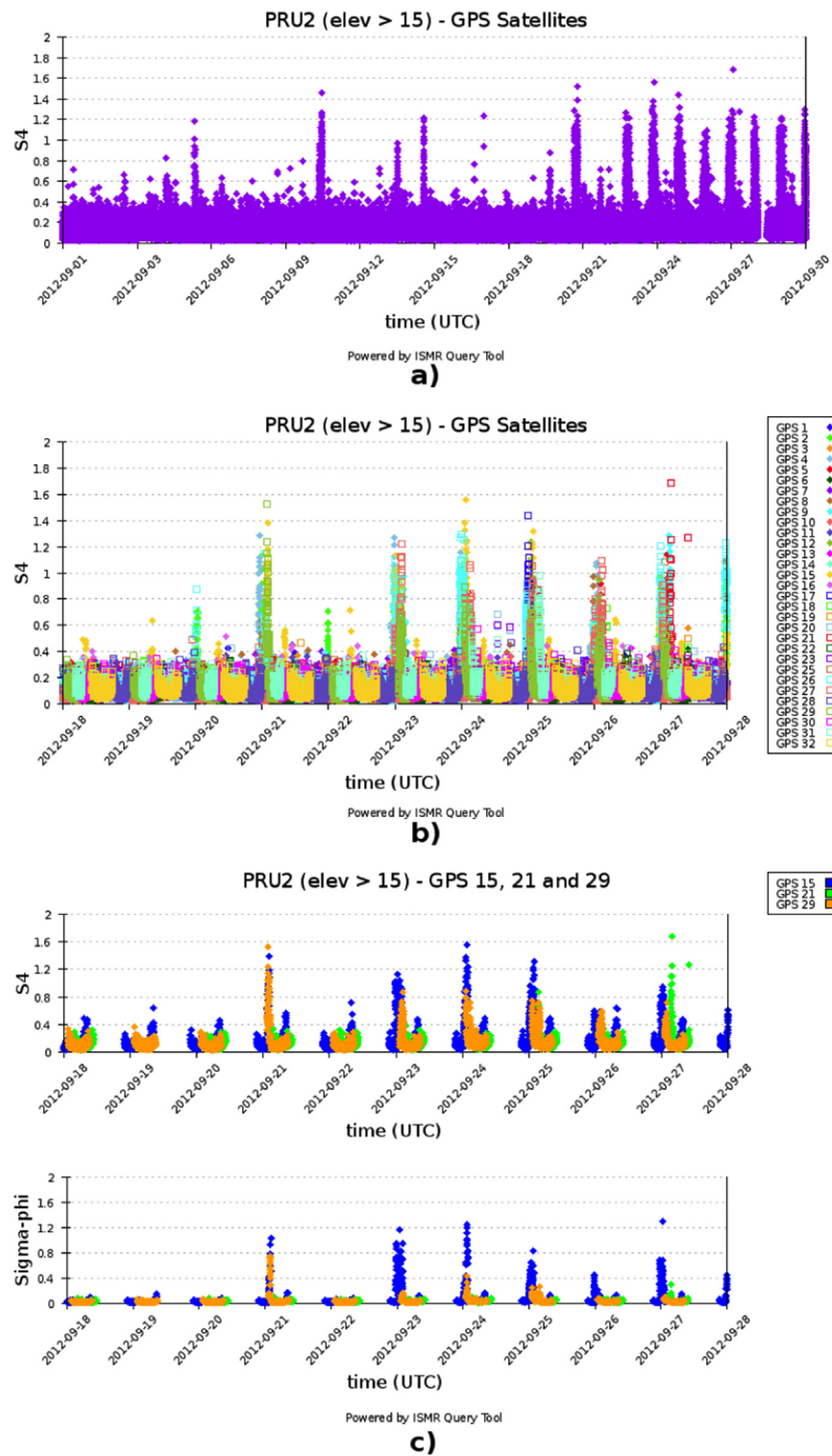


Fig. 5. Different scatter plot configurations allowing general to specific perspective: a) S4 for all GPS satellites during September/2012; b) Restricting for a period of 10 days and identifying the satellites by colors; c) Restricting for just three satellites and visualizing S4 and Sigma-phi at the same plot.

the time axis. In conclusion, such a scheme allows one to analyze and, principally, to compare ISMR data more effectively, having the benefits of the SAX approach capabilities, like dimensionality and numerosity reduction.

Another example of the application of the SAX approach is the comparison between constellations over the time. One comparison example is presented in Fig. 10. Using the same classification for weak, moderate or strong scintillations, we show the results for the

classification of S4 values considering all satellites from GPS, GLO-NASS and Galileo constellations individually. The relative number of windows classified as moderate or strong scintillations are presented. Data collected during the whole year of 2013 were used (with an elevation cut-off of 20°). One can see that the results are quite similar, therefore allowing one to conclude that any of the constellations can be individually applied to characterize the scintillation scenario through a long time series analysis (e.g. one year).

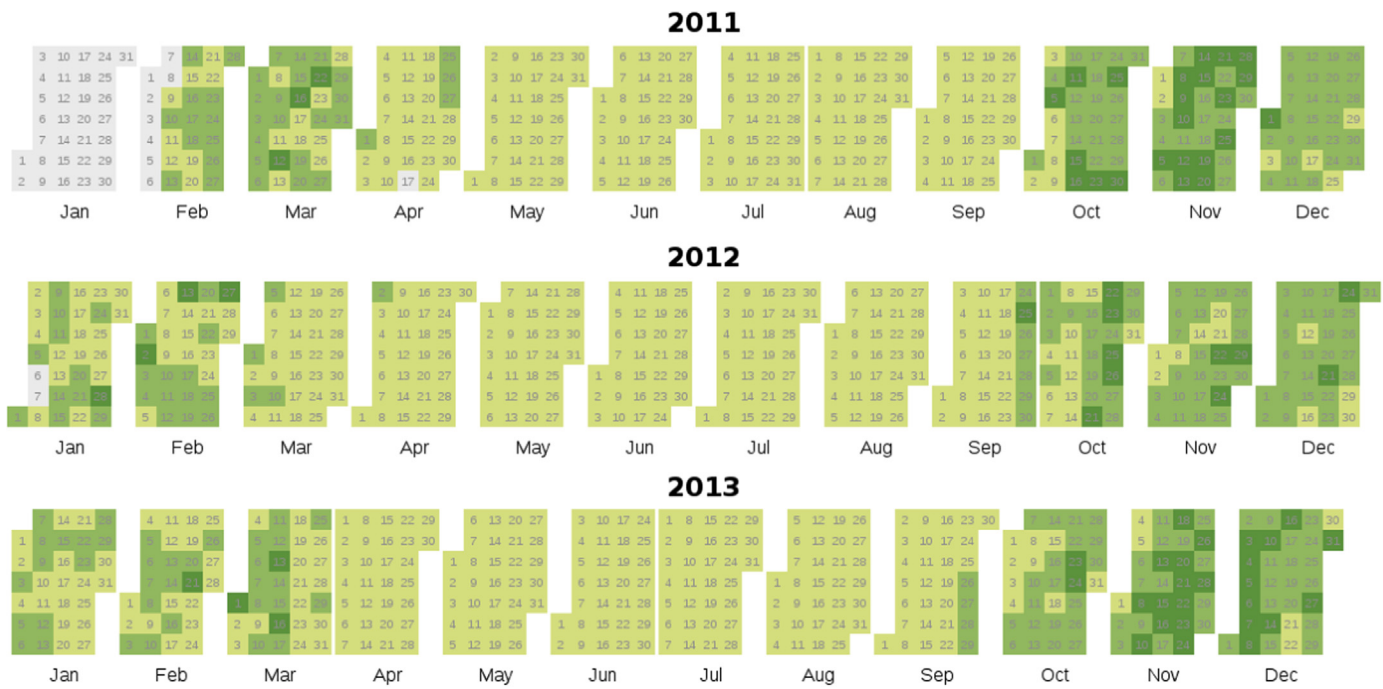


Fig. 6. Calendar views for the averaged daily S4 index considering all the monitoring stations (gray cells indicate absence of data).

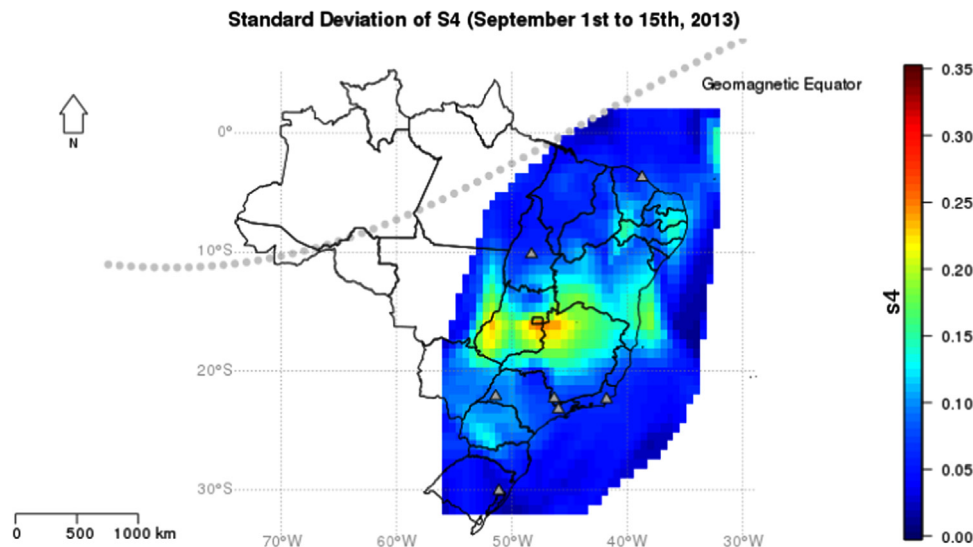


Fig. 7. Standard deviation of the S4 index with IPP representation. GPS data from the first half of December 2013 collected by seven stations (grid resolution of $2 \times 2^\circ$; elevation cut-off 30°).

4. Conclusions

Aiming to extend the possibilities of exploratory analysis of ISMR data, techniques of data visualization and data mining were investigated and, then, those selected according to their suitability for research on ionospheric scintillation impacts to satellite-based navigation and positioning were implemented and the software ISMR Query Tool was made available.

The software resources support several kinds of analysis and experiments in the context of data from the CIGALA/CALIBRA Network, like the evaluation of temporal and spatial behaviors of ionospheric scintillations in Brazil. Among the available resources, it can be stated that the main contribution is the possibility of rapid identification of the presence, intensity and duration of scintillations from a large amount of data. Such identification can contribute in different scenarios, mainly in fields related to GNSS

survey and atmospheric studies. All results presented in this paper are reproducible online at the ISMR Query Tool web page.

The availability of this software on the Web makes a significant contribution to research related to the ionosphere, not only in the context of GNSS, but also in the context covered by other areas, such as Aeronomy and Space Weather Climatology. The users' segment of the ISMR Query Tool also includes professional users, mainly from offshore surveying applications; in such cases, the ISMR Query Tool supports the development of reports about scintillations that could interrupt survey applications with GNSS. Its implementation as a web application was an appropriate decision due to the availability of its updated version to our collaborators distributed at several sites; thousands of queries were registered annually from associated researchers in Brazil and other countries. During the five subsequent years since the first released version of the software (2011), more than 30,000 outputs (such as

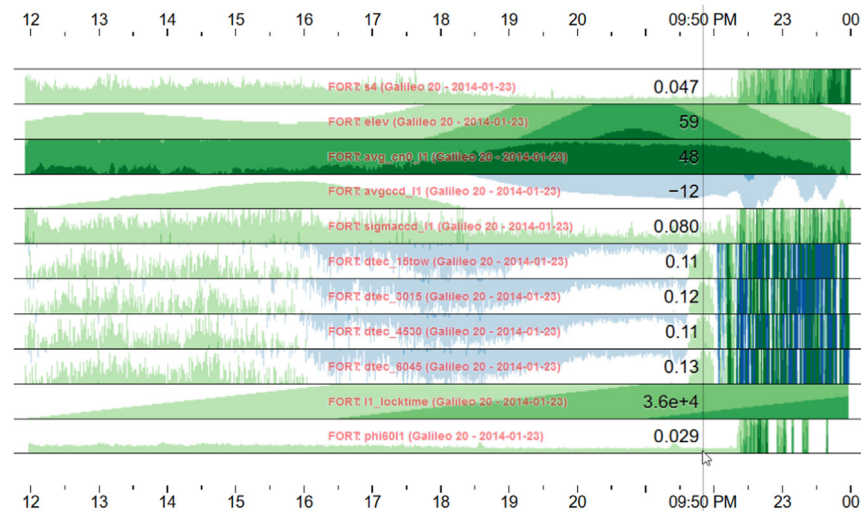


Fig. 8. Horizon chart based visualization applied to eleven monitoring parameters for a satellite tracked by the FORT station (green means positive values and blue means negative values). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

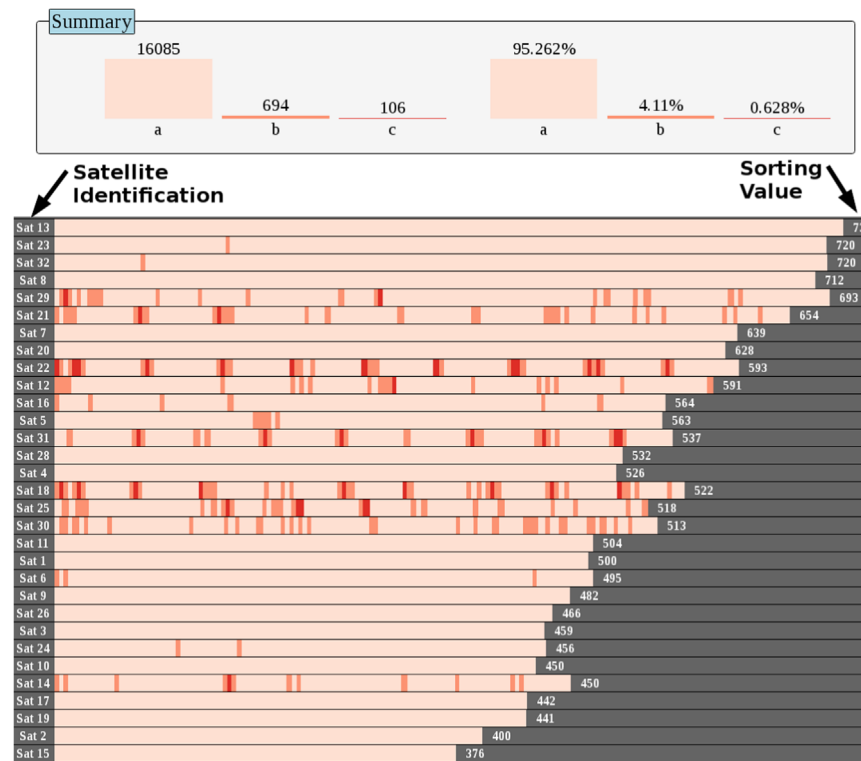


Fig. 9. Visualization for the SAX approach output applied on GPS data from PRU1 receiver tracked during the first half of September 2012.

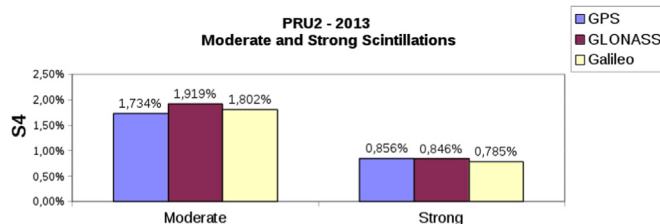


Fig. 10. Moderate and strong scintillations: comparison between GPS, GLONASS and Galileo consider all data tracked by the PRU2 receiver during the whole year of 2013.

plots and maps) were generated by the software from people from more than twenty different countries. This underlines its contribution to professional users and to the scientific community.

As future work, we have in mind the implementation of other data visualization and data mining techniques applicable for spatial and temporal data. Multiple and linked view approaches can also be investigated in order to assist the data exploration and analysis from receivers at different locations. Design and implementation of a module of geospatial data analysis would be desirable at the ISMR Query Tool. Such a module could provide density analysis and analysis of spatial point patterns in order to

support inferences about ionospheric plasma bubbles based on scintillation indices, as well as their variations in space and time. Improvements on the interactive visualization schemes would also be desirable in order to make the sense-making loop more attractive to the user. We are also targeting the use of prediction techniques applicable to the both the spatial and temporal domain, since the fundamental infrastructure is currently organized and operational. The development of a scintillation forecast service is under investigation in ongoing research. The scintillation prediction tasks include the measurement of the S4 level in near real time as well as alert emissions and correction transmissions for GNSS users in real time. The usage of predictive data mining approaches, such as Artificial Neural Networks and Support Vector Machines are under investigation. We emphasize that the prediction of scintillation phenomena has significant importance to GNSS users, since the prediction can be applied with strategies to mitigate scintillation effects on GNSS positioning, as well as to contribute to field operations on GNSS, like navigation and real time survey.

Finally, functionality implemented in the ISMR Query Tool is available, via an authorized login, at <http://is-cigala-calibra.fct.unesp.br/>. The login request can be made on the same website using an online request to the Space Geodesy Study Group from UNESP – Univ Estadual Paulista. This computational software constitutes of an Ionosphere Monitoring service available online for research purposes.

Acknowledgments

The authors thank the Coordination for the Improvement of Higher Level Personnel (CAPES) for funding the masters research related to this paper. The authors also thank partners of the projects CIGALA and CALIBRA that contributed to this research. Both projects were funded by the European Commission (EC) in the framework of the FP7-GALILEO-2009-GSA and FP7-GALILEO-2011-GSA-1a, respectively. We also thank FAPESP Project Number 06/04008-2 and associated collaborators.

References

- Aquino, M., et al., 2009. Improving the GNSS positioning stochastic model in the presence of ionospheric scintillation. *J. Geod.* 83 (10), 953–966.
- Conker, R.S., et al., 2003. Modeling the effects of ionospheric scintillation on GPS/Satellite-Based Augmentation System availability. *Radio Sci.* 38 (1), 1–1.
- Darbeheshti, N., et al., 2013. The ANU GRACE visualisation web portal. *Comput. Geosci.* 52, 227–233.
- Datta-Barua, S. et al., 2003. Ionospheric scintillation effects on single and dual frequency GPS positioning. In: *Proceedings of ION GPS/GNSS*, pp. 336–346.
- Few, S., 2008. Time on the horizon. *Vis. Bus. Intell. Newsl.*, 1–7.
- Han, J., Kamber, M., 2006. *Data Mining: Concepts and Techniques*, 2nd ed. Morgan Kaufmann, San Francisco, CA, 770 p.
- Harrower, M., Brewer, C.A., 2003. ColorBrewer.org: an online tool for selecting colour schemes for maps. *Cartogr. J.* 40 (1), 27–37.
- Keim, D., et al., 2008. Visual analytics: definition, process and challenges. In: Kerren, A., et al. (Eds.), *Information Visualization: Human-centered Issues and Perspectives*. Springer Science & Business Media, Berlin, pp. 154–175.
- Lin, J. et al., 2003. A symbolic representation of time series, with implications for streaming algorithms. In: *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, pp. 2–11.
- Monico, J.F.G., 2008. *Posicionamento pelo GNSS: descrição, Fundamentos e Aplicações*, 2nd ed. Editora UNESP, São Paulo.
- Müller, W. and Schumann, H., 2003. Visualization methods for time-dependent data – an overview. In: *Proceedings of the 2003 Winter Simulation Conference*, Vol. 1, pp. 737–745.
- Reimann, C., et al., 2011. *Statistical Data Analysis Explained: Applied Environmental Statistics with R*. John Wiley & Sons, Chichester, 343 p.
- Rezende, L.F.C., et al., 2007. Mapping and survey of plasma bubbles over Brazilian territory. *J. Navig.* 60 (01), 69–81.
- Romañach, S.S., et al., 2015. EverVIEW: a visualization platform for hydrologic and Earth science gridded data. *Comput. Geosci.* 76, 88–95.
- Seeber, G., 2003. *Satellite Geodesy: Foundations, Methods, and Applications*. Walter de Gruyter.
- Shneiderman, B., 2001. Inventing discovery tools: combining information visualization with data mining. In: Jantke, Klaus P., Shinohara, Ayumi (Eds.), *Discovery Science*. Springer, Berlin Heidelberg, pp. 17–28.
- Sreeja, V., et al., 2011. Impact of ionospheric scintillation on GNSS receiver tracking performance over Latin America: introducing the concept of tracking jitter variance maps. *Space Weather* 9, 10.
- Van Dierendonck, A.J., Klobuchar, J., Hua, Q. 1993, September. Ionospheric scintillation monitoring using commercial single frequency C/A code receivers. In *proceedings of ION GPS*, Vol. 93, pp. 1333–1342.
- Van Wijk, J.J., Van Selow, E.R., 1999. Cluster and calendar based visualization of time series data. In: *Info Vis' 99 Proceedings of the 1999 IEEE Symposium on Information Visualization*, pp. 4–9.
- Ward, M.O., et al., 2010. *Interactive Data Visualization: Foundations, Techniques, and Applications*. CRC Press, Natick.
- Zhang, L., et al., 2007. Visualization of large spatial data in networking environments. *Comput. Geosci.* 33 (9), 1130–1139.