IONOSPHERIC EFFECT ON GLOBAL NAVIGATION SATELLITE SYSTEM PERFORMANCE - A REVIEW

Sujan Kumar Saraswati
Regional Officer, CNS
International Civil Aviation Organisation (ICAO)
Asia Pacific (APAC) Office
252/1, Vibhavadi Rangsit Road, Chatuchak
Bangkok-10900, Thailand
Email : ssaraswati@bangkok.icao.int

Abstract

Satellite-based Communication, Navigation, Surveillance and Air Traffic Management (CNS/ATM) systems have been adopted by the international aviation community to meet the challenges of continuously rising global air traffic. Global Navigation Satellite System (GNSS) has been prescribed for the provision of navigation services to the air traffic. Performance of GNSS depends upon many factors including refraction and scintillation effects of ionosphere on the satellite signals. Severity of this effect depends largely upon the geographical location of the navigation receiver, the time of the day, the month of the year and the year on the 11-year solar cycle. This paper studies ionosphere mechanism and discusses its effect on the performance of GNSS. The paper also predicts the performance during solar cycle maximum expected in 2013-2014 [Jensen, Anna B.O., Mitchell, Cathryn 2011]. It touches upon the preparations going on to face the adverse effects of increased solar activity during the cycle maximum and recommends coordination/cooperation between various relevant organizations/institutions in the processes of data collection and development of a model.

Introduction

To meet the requirement of enhancing airspace capacity, continuous efforts are being made to upgrade aeronautical CNS/ATM capabilities homogeneously all over the world. The Future Air Navigation Systems (FANS) Committee, established by International Civil Aviation Organization (ICAO), recommended extensive usage of satellite technology and computer based systems for enhancing the global air traffic handling capabilities. Global Navigation Satellite System (GNSS) was envisaged in the FANS committee recommendation for providing uniform global navigation, including over oceanic and remote airspace where provision of this service by ground based systems is not feasible. The first package of GNSS Standards and Recommended Practices (SARPs) developed by ICAO was published in Annex 10 to become applicable on 01 November, 2001 [Iatsouk, Victor 2004].

Development of the first core component of GNSS started way back in 1973 and the system became fully operational in 1994 [Lammertsma, P.F 2005] with more systems following shortly afer that. In all these systems, the position information of the receiver is calculated by measuring distances (called pseudo-ranges) from four or more satellites, and then resolving the matrix for the un-known receiver-position information (both lateral and vertical). The distance is measured in terms of the time taken by the signal to travel from the satellite to the receiver and hence any error in measuring the transit time from satellite to the receiver reflects on the accuracy of the distance calculated which in turn affects the accuracy of the positional information of the receiver.

There are many factors contributing to the measurement error of pseudo-range. Prior to its removal on 01 May, 2000, Selective Availability (SA) was the largest contributor of error in the US GPS performance. But now, in the absence of SA, the ionospheric effect on the satellite signals can be considered as the largest source of error

Paper Code : V63 N4/753-2011. Manuscript received on 25 Apr 2011. Reviewed and accepted as a Full Length Contributed Paper on 20 Sep 2011

[Klobuchar, John A. 1991]. This effect, particularly during the solar maximum period has become one of the most important parameters which challenge the use of GNSS for applications requiring a very high degree of accuracy and integrity, as in aviation.

Error introduced because of ionospheric effect could vary from less than a meter to more than hundred meters [Klobuchar, John A. 1991] and this error depends upon many factors, including geographical position of the receiver, time of day, day of year and the year on the solar cycle. If a receiver uses signals from satellites which are more or less evenly spaced out in the horizontal, then the lateral error resulting from ionospheric effect on signal from each satellite tends to cancel each other and hence the overall error gets reduced. But for the elevation information, since all the satellites are located above the surface of the earth, there is no cancellation of error caused by the ionospheric effect and hence the error remains un-cancelled and is higher compared to the lateral error. Whereas the delay is minimum for the signals received from the satellites in the zenith (vertically above), the delay in the signals received from satellites located below 5° angle in elevation could be as high as three times and it can be around 300 nanoseconds contributing an error of 100 meters (uncorrected) approximately. Moreover, under certain conditions, amplitude fading and scintillation effects can be additional factors affecting GNSS performance

Discussion

Ionospheric Effect

The upper part of the earth's atmosphere contains Ionosphere, a thick layer of ionized gases located from about 50 km to about 1000 km above the surface of the earth. This ionized gas, containing free electrons and the positive ions is called plasma. For most parts, these free electrons are the valence electrons, stripped from upper atmosphere atoms and molecules by the extreme ultraviolet light being continuously emitted by the sun. On the night side of the earth, the electrons and ionized atoms/molecules tend to combine in the absence of sun's ultraviolet light. This ionization and recombination, along with its interaction with earth's magnetic field, governs the density of electrons in the ionosphere at a particular location, at a particular time. Ionospheric behavior varies depending upon the geographical location of the observer on or near the earth's surface. This behavior can be generalized into auroral, mid-latitude and equatorial areas [Datta-Barua, S., Dehel, T., Klobuchar, J.A. 2003]. The geographical band 10 to 15 degree north and south of magnetic equator is called the equatorial anomaly region and the next one beyond equatorial anomaly region is called the mid-latitude region. It has been observed that the regions of highest ionospheric delay are located on an average between +/- 15° to +/- 20° either side of earth's magnetic equator and the delay here reaches about 50 nanoseconds (roughly 15 meters in range error) [Klobuchar, John A. 1991). It may be added that the magnetic equator is not aligned with the geographic equator (because the axis of earth's magnetic field is not aligned with the earth's geographical axis) and it passes through places in India, Brazil and etc., so the effect is expected to be more prominent in these countries. In addition to the ultraviolet rays, earth's ionosphere is also affected by the solar wind and its associated magnetic field. Though the shell created by earth's magnetic field (magnetosphere) tends to deflect solar winds, the interplanetary magnetic field which is associated with the magnetic wind can cause disturbances in the geomagnetic field. When this happens, particles of the solar wind enter the geomagnetic field and cause an increase in ionization of ionosphere. Solar wind contains charged particles, mostly electrons and protons with energies between 10 and 100 kilo-electron-volts, travelling at an average speed of 450 kilometers per second (varying between 200 to 900 kilometers per second depending on the solar activity) [Jensen, Anna B.O., Mitchell, Cathryn 2011]. Also, sudden eruptions of the sun, such as solar flares and coronal mass ejections (CME) cause increased ionization, thereby creating larger ionospheric variability. Solar flares are usually associated with strong magnetic fields that manifest themselves in sunspots - relatively cooler areas that appear as dark blemishes on the sun's surface [Langley, Richard B. 2000]. As seen from the earth, these sunspots rotate over the surface of sun, with the sun's rotation period of about 27 days (Furgo 2008). CMEs are huge bubbles of gas threaded with magnetic field lines that are ejected from the sun. Frequency of CMEs varies with the sunspot cycle. In a study carried out in Brazil during the solar minimum, approximately one CME per week was observed, which increased to two or three CMEs per day near the solar maximum [Langley, Richard B. 2000].

Signals from global navigation satellites have to transit through the ionospheric layer on their way to the receivers located on or near the surface of the earth. This transition through the ionosphere causes an error in the form of a transition delay or an addition to the transit time (which reflects on the pseudo-range calculation) caused by the presence of the free electrons. In addition to causing variation to transit speed, Ionospheric effect also causes scintillation causing rapid changes in the phase and amplitude of the transiting signal. Sometimes the satellite signals are completely lost because of scintillation.

In addition to the ionosphere, in the lower parts of earth's atmosphere that is in the troposphere and the stratosphere, where the atoms and molecules are electrically neutral the satellite signals suffer additional refractions because of pressure, temperature and humidity. This effect is however, not being discussed in this paper, as its influence on GNSS performance is much less significant as compared to the ionospheric effect.

Signal

Propagation speed of radio signal in a medium is affected by its refractive index 'n', which is defined as the ratio of speed 'c' in vacuum to the speed 'v' in the medium (1).

$$n = c/v \tag{1}$$

Speed of un-modulated electromagnetic wave propagating through a medium is called the phase speed. When the medium is dispersive, the phase speed becomes dependent on the frequency of the electromagnetic wave. A modulated carrier can be interpreted as the superposition of a group of un-modulated carriers, each having its own speed through the medium. This differential in speed for each frequency leads to a different speed of the modulated carrier through the medium and is called the group speed (v_g). Corresponding to the phase refractive index 'n', the group refractive index 'n_g' can be calculated using following formula [Langley, Richard B. 2000] (2).

$$n_g = c/v_g \tag{2}$$

It can be shown that

$$n_g = n + f\,(dn/df) \tag{3}$$

where 'dn/df' in (3) describes how the phase refractive index changes with frequency.

Free electrons present in the ionosphere contribute towards the ionospheric delay, caused by a change in the refractive index along the signal path, which is dependent on the composition of the medium. For the path between the GNSS satellite (approximately 20,000 km from the earth surface) to the upper limit of ionosphere (about 1000 km above earth's surface), the change in the refractive index is usually very small and can be ignored. The satellite signal obviously is affected by the integration of free electron density along the whole path. Signal delay 'd' in meters because of first order ionospheric effect can be estimated by equation [Jensen, Anna B.O., Mitchell, Cathryn 2011] (4).

$$d = (40.3/f^2)\,TEC \tag{4}$$

where 'f' is the signal frequency (1.57542×10^9 Hz for L1 in GPS), 40.3 is a constant calculated on the basis of values of electron charge, electron mass and permittivity of free space and TEC is the Total Electron Content which is defined below.

Total Electron Content (TEC): TEC is the integration of number of free electrons along the signal path in a cross section of one square meter [Jensen, Anna B.O., Mitchell, Cathryn 2011; Langley, Richard B. 2000; Klobuchar John A. 1991]. TEC is measured in TEC Units (TECU), where one TECU is 10^{16} electrons per square meter.

Large number of models and methods for estimating ionospheric signal delay have been developed, the most widely used being Klobuchar Model. Coefficients for Klobuchar model, determined by the GPS control segment, are distributed as part of the GPS navigation message [Klobuchar John A. 1991]. GPS receivers, using Klobuchar model and the coefficients received from the navigational message, estimate the ionospheric delay and apply corresponding correction in the calculation of positional information. The coefficients are updated at less than 10-day intervals. The correction [Klobuchar] algorithm was originally designed to reduce the group delay error by approximately fifty percent in root-mean-square (rms) terms, but in actual application the reduction in the error is found to be much better.

Ionospheric delay is a function of frequency also. So if pseudo-range measurements from more than one frequency are available, these can be used for enhanced modeling of the ionospheric effect. Receivers using dual or more frequencies can correct ionospheric range error through an appropriate combination of pseudo-ranges observed on different frequencies [De Paula, Eurico R. et.al. 2009]. Single frequency receivers, on the other hand either have to live with the reduced measurement accuracy or employ a model for the correction of ionpheric range errors. The delay calculated using pseudorange measurements on two frequencies is given by expression (5) [Jensen, Anna B.O., Mitchell, Cathryn 2011].

$$d_1 = \left[(f_2)^2 / \{ (f_2)^2 - (f_1)^2 \} \right] (P_1 - P_2) \qquad (5)$$

where P is the pseudo-range and 'f' denotes frequency. The estimations given by these models are first order only. It may be mentioned here that delay calculation using measurements on two frequencies does not need the TEC and hence it becomes absolutely autonomous, whereas the delay calculation using one frequency only needs to use TEC, which is required to be provided from an external source. It is for this reason that a receiver using two frequencies on board an aircraft can autonomously calculate the ionospheric delay and apply the corrections accordingly.

Scintillation

Scintillation occurs when the GNSS satellite signal travels through small-scale irregularities (bubble) in electron density in the ionosphere, typically in the evening and nighttime and in equatorial region [Datta-Barua, S., Dehel, T., Klobuchar, J.A 2003]. These bubbles of rarefied plasma regions could vary in size from a few centimeters to many kilometers [De Paula, Eurico R. et. al. 2009]. These irregularities in electron densities can diffract the signal, leading to rapid fluctuations in signal intensity, known as amplitude scintillation. Below 30db carrier/noise ratio, the signal reaches a critical point with the possibility of loss of lock [De Paula, Eurico R. et. al. 2009; Dubey, Smita., Wahi, Rashmi., Gwal, A.K. 2005] due to scintillation. The 11-year solar cycle, local season of the year and geomagnetic location of the receiver all play a role in defining the degree of scintillation activities.

Figure 1 describes how the GNSS operations are affected by the scintillation phenomenon. If the scintillation patches cover a large portion of the sky, as indicated in Fig.1, there is a probability that the GNSS receiver will lose a significant number of satellites and this may even result in discontinuation of navigation services.

Solar Maxima

Since ionospheric effect is mainly driven by solar activities, its effect on the density of free electrons follows a daily cycle with the largest TEC values occurring in early afternoon local time and the lowest activity just before the sun rise. The level of activity also depends upon the geographical location, the highest electron density being in the equatorial region and lowest in the high latitude regions.

Solar activity and the level of emissions from the sun are closely related to the number of sunspots. The number of sunspots generally follows a cycle of about 11 years, but cycles of 9 and 13 years have also been observed. During the last couple of years (2007 to 2009), we have seen a period of low numbers of sunspots; in fact there were many days in a row when no sunspot was visible. During the next three to four years, the number of sunspots is expected to increase [Jensen, Anna B.O., Mitchell, Cathryn 2011] and this is expected to be followed by a period of low sunspots most probably during 2019 to 2020 again. It has been observed that the frequency of solar wind, solar flares and emissions also increases with the increasing number of sunspots. All these add up to a higher number of free electrons in the ionosphere and a larger variability of ionospheric delay during the period of higher sunspot activity during solar maximum.

Analysis of the data collected in Brazil during December 2005 to January 2006 (solar minimum period), indicated that the scintillation effect on GNSS navigation even in the equatorial region was not very severe during the solar minimum [Jiwon seo et. al. 2007]. Under strong solar activity during solar maximum period, scintillation is expected to be more frequent and scintillation patches may cover larger portion of the sky resulting in frequent loss of more than two satellites simultaneously leading to non-availability of service. Data collected on Ascension Island (South Atlantic) in March 2001 (solar maximum period) was analyzed and it was found that four out of seven satellites visible at the location were affected by scintillation for 36 minutes, the services however were not affected much, because deep fading of individual satellite signals did not coincide. It is hence predicted that if a receiver re-acquires a lost satellite very quickly (satellite re-acquisition time typically is about 1.5 seconds) before it loses another satellite, loss of lock can be avoided and the navigation services can continue. WAAS receiver usually re-acquires phase lock within 2 seconds during strong scintillation, but the maximum time for re-acquisition could go as high as 20 seconds [Jiwon seo et. al. 2007]. Through simulation study, it has been proved that there would be no case of simultaneous loss of three satellites if every channel is reacquired in 2 seconds. So, in addition to the severity of the scintillation, the availability of GNSS will also depend upon the reacquisition time of the receivers used.

There are many other effects of solar cycle like increased drag on spacecraft, altering their orbits [Fugro 2008], but these issues are not being discussed in this paper.

Last Solar Maxima

Sunspots have been observed since at least 1600 AD with regular daily sunspot counts beginning in 1749 [Langley, Richard, B. 2000]. Current solar cycle is referred to as Cycle 24 (traditionally 1755 - 1766 solar cycle is counted as Solar Cycle 1), which began on 24 January 2008 [Jensen, Anna B.O., Mitchell, Cathryn 2011]. As the solar cycle is progressing since then, the average number of daily sunspots is rising, creating conditions favorable for solar flares and CMEs. During the last solar cycle (23), the GNSS community was aware of what could happen, and therefore many events were observed and analyzed. Among the most significant ones that were observed was the sequence of solar storms during October/November 2003 resulting from CME on 29 October 2003 at 20:43 UTC, which impacted earth's magnetic field at 16:20 UTC on 30 October. It produced a great geomagnetic storm lasting for many hours [Jensen, Anna B.O., Mitchell, Cathryn 2011]. It was reported that the Wide Area Augmentation System (WAAS) vertical error limit of 50 meters was exceeded for a period of about 11 hours on 30 October, 2003.

Prediction for the Next Solar High

Because of a very long period of low solar activity from 2007 to 2009, it is expected that there will be a sudden outburst of activity and a very large cycle maximum will be observed. There are however other predictions also, which claim that there will not be any increase in solar activity for many years [Jensen, Anna B.O., Mitchell, Cathryn 2011]. With the general increase in the number of sunspots during 2010, it looks like that we are well into the solar cycle number 24. It is predicted (by NASA) that the maximum of current solar cycle will be lower than the maximum of the last cycle encountered in 2001. The maximum sunspot number in this solar cycle is expected to be about 59 in June/July 2013. Fig.2 provides information on the sunspots for Cycle 23 and predictions for Cycle 24.

Consequences for GNSS Users

Signal delay because of ionospheric effect has been modeled and these models have been able to provide satisfactory GNSS service for most of the applications. But during increased ionospheric activities, the users may experience high residual ionospheric effect, which may ultimately reduce positioning, navigation and timing performance of GNSS. During enhanced ionospheric or geomagnetic storm activity caused by the sudden eruptions in the sun, increased variability in the ionospheric activity is expected leading to signal delay and scintillation peaking during Solar Max period in 2013-14. For the current solar cycle, the civil aviation community is much better prepared compared to the last cycle. GNSS software and receiver technology have been improved to better resist the challenges expected during increased solar activities. For example, the user of WAAS and EGNOS have correction and integrity information available (which were not there during last cycle), which can be of great help in identifying time epochs when positioning and navigation solutions might not be trustworthy. However, it should be possible for the receivers to detect the disruption and display a warning message when SBAS service is disrupted. In addition to the improvement in the warning systems, a number of internet sites are now providing current and predicted ionospheric environment. Multimode receivers, using two or more autonomous core GNSS constellations or receivers using two or more frequencies, will stand to benefit during this period.

Getting Ready

As has been mentioned, satellite navigation, including core satellite systems and various augmentations have been included in Annex 10 Volume 1 SARPs. For all these systems ionospheric prediction models are required to be created to provide ionospheric correction data.

It has been observed that many countries in the Asia Pacific region have taken up collection of ionospheric data for study. Ionospheric data being collected by different organizations (academic and non-academic) in different locations can be used for characterizing the ionosphere. Need for coordination amongst these different agencies engaged in the study of ionosphere has been felt at different levels.

Because of different ionospheric conditions existing in different parts of the globe, a single model may not provide the best possible solution. As mentioned earlier, ionospheric effect is most severe in the lower latitudes and hence a common model may be useful for all the Countries which lie in this region. A coordinated effort between the countries including Japan, Thailand, India, Australia, which lie in this region is likely to provide a very useful solution. Many organizations are coming together to cooperate with each other in collecting data during the solar max period and contribute towards the development of a suitable model for the region.

In the ICAO Asia Pacific region, the requirement for coordination in the study of ionosphere was felt by ICAO Asia Pacific Air Navigation Planning and Implementation Regional Group (APANPIRG). The Group adopted a number of conclusions promoting cooperation between the countries in the region in characterizing ionosphere. Under APANPIRG directives, Focal Contact points to coordinate efforts in this direction have been identified and a workshop to develop the modalities for ionospheric data collection was organized in May 2011 with the ultimate objective of developing a model which is more suitable than the available existing ones.

Recommendations

There are many organizations and academic institutions located in the lower latitude region, which are engaged in the study of ionospheric effect on the propagation of electromagnetic signal. Since the electromagnetic signals, including the GNSS signals, radio signals in HF band etc. are likely to be affected by the increased ionospheric activities during the solar maximum period, it will be more beneficial for all such organizations to pool their resources for the development of a comprehensive predication model and other mitigation techniques for the overall benefit of the region/countries. Pooling of archival data to carry out simulation studies will also help in augmenting information on the subject.

References

1. Datta-Barua, S., Dehel, T. and Klobuchar, J.A., "Ionospheric Scintillation Effects on Single and Dual Frequency GPS Positioning", Proceedings of ION GPS/GNSS 2003, Portland, OR, September 2003, pp.336-346.

2. De Paula, R., Muella, M.T.A.H., de Rezende, L.F.C., Kantor, I.J., Correa, Keila S., de Siquera, P.M., Dutra Anderson and Guedes Cedrique., "Ionospheric Effects over GNSS", Galileo Networking Meeting, 24-25, November 2009, INPE - Sao Jose dos Campos - Brazil.

3. Dubey Smita., Wahi Rashmi and Gwal, A.K., "Effect of Ionospheric Scintillation on GPS Receiver at Equatorial Anomaly Region Bhopal", International Union of Radio Science, XXVIIIth General Assembly, 23 - 29, October, 2005.

4. Iatsouk Victor., "Development of Standards for Aeronautical Satellite Navigation System", Acta Astronautica, 54, 2004, pp.961-963.

5. Jensen Anna, B.O. and Mitchell Cathryn., "GNSS and the Ionosphere, What's in Store for the Next Solar Maximum?", GPS World, February, 2011, pp.42-48.

6. Jiwon Seo., Todd Walter., Edward Marks., Tsung-Yu Chiou and Per Enge., "Ionospheric Scintillation Effects on GPS Receivers During Solar Minimum and Maximum", International Beacon Satellite Symposium 2007, 11 - 15, June, 2007, Boston, MA, USA.

7. Klobuchar John, A., "Ionospheric Effects on GPS", Innovation, GPS World, April, 1991.

8. Lammertsma, P.F., "Satellite Navigation", Institute of Information and Computing Sciences, Utrecht University, February, 2, 2005.

9. Langley, Richard B., "GPS, the Ionosphere, and the Solar Maximum", Innovation, GPS World, July, 2000, pp.44-49.

10. "Solar Cycle 24 and its Impact on GNSS Positioning", Fugro, November, 2008.

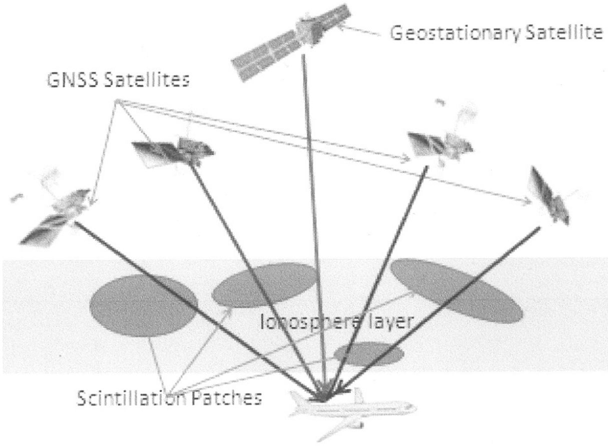

Fig.1 Scintillation Patches

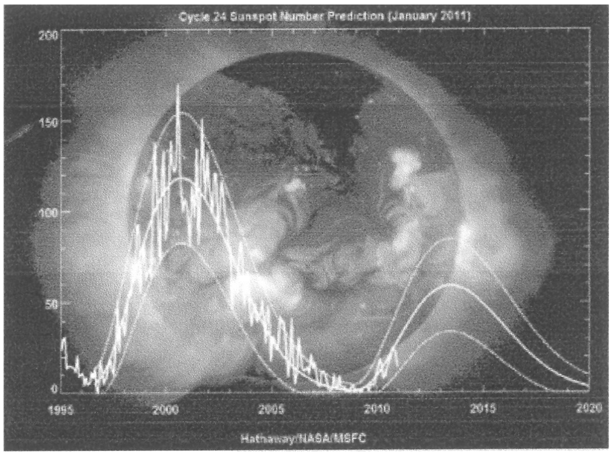

Fig.2 Sunspots for Cycle 23 and 24
(from NASA Marshall Space Flight Center)