# VALIDATION OF F2 LAYER PEAK HEIGHT AND DENSITY OF REAL-TIME INTERNATIONAL REFERENCE IONOSPHERE

BY

ARTEM MIKHAYLOVICH VESNIN
B.S. IRKUTSK STATE UNIVERSITY, RUSSIA (2009)

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE
DEPARTMENT OF PHYSICS AND APPLIED PHYSICS
UNIVERSITY OF MASSACHUSETTS LOWELL

**Signature of Author:**_____**Date:**_____

**Signature of Thesis Supervisor:**_____
**Name Typed: Prof. Ivan Galkin**

**Signature of Thesis Supervisor:**_____
**Name Typed: Prof. Paul Song**

**Signatures of Other Thesis Committee Members:**
**Committee Member Signature:** _____
**Name Typed: Prof. Timothy Cook**

**Committee Member Signature:** _____
**Name Typed: Prof. Silas Laycock**

# VALIDATION OF F2 LAYER PEAK HEIGHT AND DENSITY OF REAL-TIME INTERNATIONAL REFERENCE IONOSPHERE

BY

ARTEM MIKHAYLOVICH VESNIN

ABSTRACT OF A THESIS SUBMITTED TO THE FACULTY OF THE
DEPARTMENT OF PHYSICS AND APPLIED PHYSICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE
UNIVERSITY OF MASSACHUSETTS LOWELL
2014

Thesis Supervisor: Ivan Galkin, Ph.D.
Associate Professor, Space Science Laboratory, Department of Physics and Applied Physics, University of Massachusetts Lowell

Thesis Supervisor: Paul Song, Ph.D.
Professor Physics & Atmospheric Science, Space Science Laboratory, Department of Physics and Applied Physics, University of Massachusetts Lowell

## ABSTRACT

Relying on the well-proven formalism of the International Reference Ionosphere (IRI) model, an IRI Real-Time Assimilative Mapping (IRTAM) extension is being developed at the Space Science Laboratory (SSL) University of Massachusetts Lowell. The IRTAM uses measurements from the Global Ionosphere Radio Observatory (GIRO) and knowledge about ionospheric climatology to now-cast global ionospheric weather. This thesis presents the results of the validation of current assimilative technique based on “morphing”. This technique morphs the climatology empirical model, IRI, with the current ionospheric conditions data by introducing corrections for expansion coefficients rather than in the domain maps. Global, time-dependent mapping of two main ionospheric characteristics, F2 layer peak density and peak height, has been implemented in IRTAM. The validation is successful if IRTAM can (1) routinely and accurately reproduce the modeled characteristic, peak density and peak height, and (2) adequately represent response to space weather events in the Sun-Earth system. To satisfy the first criterion, numerous statistics were collected over an 11–year time span, which includes 13 million of individual data points. Overall, IRTAM reduces the standard deviation of the IRI by half, although sometimes the improvement factor, which is measured by the ratio between IRI and IRTAM average deviations from the observations, can be greater. To satisfy the second criterion, known and well-studied space weather events were used. Independent Global Positioning System (GPS) total electron content (TEC) and Radio

Plasma Imager (RPI) data were employed to support the IRTAM simulations. Two events were chosen for the phenomenological study: Sudden Stratospheric Warming of January 2013 and interplanetary shock that hit Earth's magnetosphere at 1830 UT on 7 November 2004. We check IRTAM's ability to reproduce major ionospheric effects caused by these events. It is shown that the IRTAM reproduces major ionospheric effects. What is more interesting is that the IRTAM, being a global empirical assimilative model, is able to capture features that would have been missed if only local data were considered. Inspired by the interplanetary shock study, a detailed analysis of the ionospheric effects during the 7-8 November 2004 storm recovery phase was conducted. During the recovery phase of the storm, a large F2 layer peak density decrease is observed. We used simultaneous RPI measurement of the plasma density in the plasmasphere, to argue that the observed plasma density depletion in the ionosphere may be linked to the refilling of the plasmasphere. Conducted studies confirm that IRTAM is a new tool for modeling and data analysis with potential for ionospheric and space weather research.

## ACKNOWLEDGMENTS

I would like to thank the following: Alexander Kozlov for help in creation of the GAMBIT database and general guidelines on program implementation of IRTAM, Dr. Bodo Reinisch and Dr. Xueqin Huang for useful inputs to the thesis, Dr. Dieter Bilitza for the description of the IRI model, LDI staff for ensuring the availability of GIRO measurements, Steve Mendonca for help with designing the IRTAM webpage, personnel of ionosonde observatories for making data available online in GIRO and for maintaining instrument in continuous operation (full list of the institutions responsible for ionosonde operations can be found at the bottom of the IRTAM web page), Dr. Larisa Goncharenko for information on SSW and MIT Haystack Observatory staff for providing the Madrigal data and for the accumulation of GNSS TEC measurements.

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF ILLUSTRATIONS

# I. INTRODUCTION

## 1.1 Ionosphere and Space Weather

"Timely, accurate, and reliable space environment observations, specifications, and forecasts" is the goal of the United States National Space Weather Program [*NSWP*, 2012]. The practical need to mitigate the adverse effects of space weather on human activities (see, for example *Stanislawska et al.*, [2010]) has created a critical need to understand and predict space weather effects in the Earth's upper atmosphere and ionosphere. In particular, obtaining the actual, real-time, worldwide knowledge of ionospheric plasma conditions is an important and challenging research task that will be the subject of this investigation.

The ionosphere is a partially ionized layer that extends from 80-90 kilometers to more than one thousand kilometers[1]. The ionosphere, as part of the solar-terrestrial system, plays a big role in space weather due to its sensitivity to various processes in the heliosphere. It's important to many practical areas such as Global Navigation Satellite

[1] Top boundary of the ionosphere is latitude-dependent and is affected by many dynamic processes.

System (GNSS) precision navigation. GNSS is the common name for global navigations system such as Global Positioning System (GPS). Plasma flow from the Sun, called the solar wind, impacts the Earth starting with the Earth's outer magnetosphere and via the perturbations of underlying plasma regions its effects are seen in the ionosphere. The ionosphere can be thought of as a global plasma observatory, which images the footprints of the processes going on in different parts of the solar-terrestrial system.

The two main characteristics of the ionosphere (critical frequency, *foF2*[1], and peak height, *hmF2*) define maximum of electron density profile shown in Figure 1. The *foF2* values vary from 1 to 20 MHz and *hmF2* from 200 to 500 km.

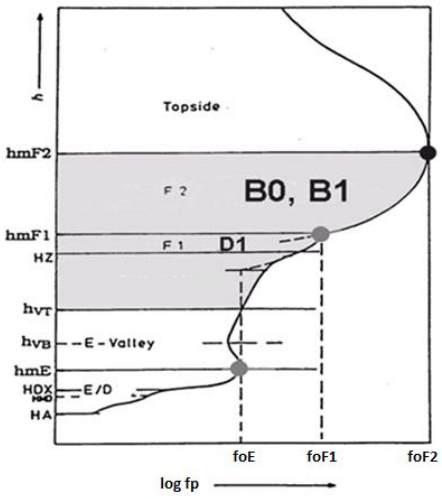



Figure 1. The plasma frequency (electron density) profile of IRI model. Adapted from *Bilitza* [1990].

[1] The critical frequency of an ionospheric layer is defined as the frequency of a vertically transmitted ordinary wave signal that begins to penetrate the layer while half of its energy still reflects and returns to the Earth surface to be detected by ionosonde. The critical frequency is proportional to square root of peak electron density of the layer, NmF2. $NmF2 = 0.0124 \bullet (foF2)^2$

Both characteristics are observed by ionosondes indirectly via a process of extracting traces of ionospheric reflections of the ionosonde signal and inverting them into electron density profile. The process of deriving *foF2* and *hmF2* from ionograms is well established and validated against independent observations by incoherent scatter radars [*Chen et al.*, 1994], which justifies our use of these characteristics to update the model.

Both *foF2* and *hmF2* are strongly affected by space weather events, and these events could be tracked using *foF2* and *hmF2* timelines. Moreover, *foF2* and *hmF2* together with slab thickness, *B0*, define the major part of total electron content (TEC) stored in the ionosphere. The value of knowledge of the current ionospheric state is hard to overestimate. This information can be used for various academic and practical purposes. Real-time global ionospheric data are most interesting to the aviation and telecommunications communities. From the perspective of low frequency communications and over-horizon radar systems, which use the ionosphere as a reflecting layer, *foF2* and *hmF2* define maximum usable frequency and maximum range. For the GNSS systems total electron content in the ionosphere is proportional to the phase delay between two signals on different carrier frequencies. Accurate maps of the TEC are required in order to make navigation precise. The accurate maps of the ionospheric characteristics can be used as inner boundary conditions for the numerical simulations of the magnetosphere.

## 1.2 Statement of Problem

The real-time, hour-by-hour space weather specification and forecast requires reliable global observations and accurate data-driven predictive models. Reliable, assimilative ionospheric models are the main tools used to estimate the current state of the ionosphere and make forecasts. Currently, global ionospheric models are of particular interest. In order to be a global ionospheric model, the model should give specific characteristics of the ionosphere for any location and time. In this case, time means the season and solar cycle phase, not just the time of day.

The empirical International Reference ionosphere (IRI) model [*Bilitza et al.*, 2010] was chosen as basis for IRTAM. Empirical models capture an average "essence" of large amounts of observed data. Temporal and spatial compressed expansion creates an averaged, smoothed representation that can be used to estimate the "typical" state of ionosphere where or when sensor data are not available. Thus, empirical models represent the ionospheric climate. IRI is the example of this class of models. Computational demand of the IRI formalism is light, which allows efficient implementations of the 4D data assimilation (4DDA) approach that involves past observations in assimilation.

The IRI model fares exceptionally well among a great number of ionospheric models currently in use [*AIAA*, 2011], which are too numerous to enumerate and analyze critically for our task. I will limit consideration to two models as the basis for further

discussion. These two models are the empirical climatology IRI model and a physics-based assimilative Utah State University Global Assimilation of Ionospheric Measurements (USU-GAIM) model [*Scherliess* et al., 2006]. Both models are top performers of the National Science Foundation (NSF) Coupling, Energetics, and Dynamics of Atmospheric Regions (CEDAR) Electrodynamics Thermosphere Ionosphere (ETI) Challenge http://ccmc.gsfc.nasa.gov/support/CETI2010/. This research program's goal was to evaluate the current state of thermosphere/ionosphere models on a suite of test examples. The CEDAR Challenge study shows a surprising result: in the current state of our space weather understanding, performance of the best assimilative theoretical model is not much different from an empirical model that describes an average climatology of the ionosphere. Even though theoretical models seem to be more comprehensive and reliable due to their design and use of underlying physical laws, the Earth's ionosphere, as a physical system of high complexity subjected to various forces from above and below, allows the empirical models to be as good as theoretical ones or even more accurate.

Typically, the IRI user community selects the IRI for applications in the role of a basic, background model, relying on its proven performance. The IRI has more than 40 years of successful usage [*ITU*, 2009]. Another recent in-depth IRI validation campaign included data from 250 worldwide ionosondes obtained over 7 solar cycles (77 years) [*Damboldt and Suessmann*, 2011]. The data from a single station might not be continuously available during 77 years. The observational dataset counted an impressive

1,569,708 values of *foF2* and 1,393,561 values of *hmF2* measurements. The median-to-median comparison of these two characteristics was performed using IRI. The results were better than expected: *foF2* error was 0.01 MHz (σ=0.78 MHz); *hmF2* error was 1.51 km (σ=25 km) [*Damboldt and Suessmann*, 2011]. Thus IRI is a proven, excellent source of accurate averaged climatology information [*Bilitza*, 2006].

Continuous work to revise and improve the IRI is required to keep it up to date and provide trustworthy representation of the ionosphere. A step further in the development of the enhancement of the IRI model is to build algorithms for assimilation of near real-time data. That would allow such an "assimilative IRI" in real-time to characterize phenomena far from the average/quiet conditions, with spatial and temporal resolutions that were not preserved in the original model due to averaging. The IRTAM model is an effort that addresses real-time IRI task. The IRTAM, however, preserves the original expansion formalism in IRI that has been optimized for ionospheric modeling.

Building a real-time global assimilative model with real-time and prediction capabilities is a monumental task, which takes a huge effort and years to research and to take a model into operation. Many models were created during history of ionospheric exploration, and the best ones are listed in *AIAA American National Standard* [2011]. By using an existing climatology model as a basis, we can concentrate on work which is more promising for research and practical needs, namely real-time modeling. Also, building an empirical model requires a huge database of observations. Such a database is critical in order to make expansions properly and obtain coefficients that reasonably

represent the global state of the ionosphere. To estimate the size of such a database, the original IRI coefficients were obtained by expanding data from about 150 ionosondes [*Jones et al.,* 1969]. Currently, the biggest database of ionosondes measurements that provides free online access to the observations is the Global Ionospheric Radio Observatory (GIRO) [*Reinisch and Galkin*, 2011]. The GIRO provides access to the data from more 40 locations throughout the globe. This is enough to make corrections to the climatology coefficients, but it is not enough to obtain new coefficients, which would only be as accurate as the original ones. The IRI model, more precisely its mathematical formalism for *foF2* characteristics, was chosen as basis for the IRTAM due to several reasons. One of them is the performance and accuracy of climatology representation of the ionosphere. The second reason is that the basis functions' explicit form is known. The procedure to calculate maps from the coefficients and expand maps back to the coefficients is also established. This is important, since the IRTAM routinely manipulates the expansion functions to find corrections for the climatology coefficients set.

Two different approaches were applied to bring assimilation of the real time data to life [*Galkin et al.*, 2012; *Huang et al.*, 2012]. For this thesis, I have explored one promising scenario based on a concept of "morphing" an empirical model based on historic data into agreement with the current data [*Galkin et al.,* 2012]. The validation of the model is a necessary step in the model development to ensure its reliability, robustness and accuracy for users. Although the IRI itself has a long history of validation and improvement, the IRTAM should be treated as a new model, and, thus, must be

tested separately. The validation of the IRTAM should reveal to which extent it is accurate for different geographic and geomagnetic regions, different times of the day and seasons, geomagnetic quiet and disturbed time. The validation will show what features were unresolved with the climatology model due to averaging and now appear in the IRTAM. Sensitivity of the IRTAM to data noise will also be examined. Although the IRTAM is designed to be a global ionospheric model, the IRTAM maps cannot be trusted equally for all locations. The IRTAM maps for locations spatially close to the locations of the GIRO sites, contributing observational data to the IRTAM, are expected to have better agreements with a real ionospheric state and the location spatially far – less agreement.

## 1.3 Components of the IRTAM

The IRTAM, having an empirical model as a basis, uses some kind of basis function set to fit observational data. Naturally, in the problems those involve spherical geometry this set of functions appears to be spherical harmonics. The IRTAM relies on a well-proven IRI formalism for *foF2* mapping. This formalism is similar to spherical harmonics but is modified in such a way, that it takes into account the asymmetry of the Earth's magnetic field. The IRI coefficients of *foF2* mapping are compressed representations of the 2D global time dependent distribution. Each set of coefficients gives the spatial and temporal variation of the characteristic based on *ITU* [2009]

formalism, which was firstly proposed by *Jones et al.* [1969]. The diurnal behavior of any characteristic is given by expansion into Fourier series. The *foF2* for any particular location is given by the harmonic series of 6th order:

$$f_o\mathrm{F2}(\lambda,\theta,\tau) = A_0(\lambda,\theta) + \sum_{i=1}^{6}\left(A_{2i-1}(\lambda,\theta)\cos i\tau + A_{2i}(\lambda,\theta)\sin i\tau\right) \qquad (1)$$

where $A_i$ are:

$$A_i(\lambda,\theta) = \sum_{j=0}^{J(0)} C_{ij0}\,P_{j0}(\lambda,\theta) + \sum_{k=1}^{8}\sum_{j=0}^{J(k)}\left(C_{ij\,2k-1}\cos k\theta + C_{ij\,2k}\sin k\theta\right)P_{jk}(\lambda,\theta) \qquad (2)$$

where $\lambda$ is geographical latitude, $\theta$ is geographical longitude, $C_{ijk}$ are coefficients of spatial expansion and $P_{jk}(\lambda,\theta)$ represent the geographical variations and include Earth's magnetic field geometry. $P_{jk}(\lambda,\theta)$ are the critical point in understanding the IRI formalism for *foF2*. The compact form of spatial expansion looks as follows:

$$A_i(\lambda,\theta) = \sum_{n=0}^{N=75} U_{in}\,G_n(\lambda,\theta) \qquad (3)$$

where $G_n(\lambda,\theta)$ are geographical functions and $U_{in}$ are coefficients of the expansion [*Jones et al.*, 1969]. For full set of IRI geographical functions, see *Jones et al.* [1969]. IRI formalism for *foF2* assumes 988 $C_{ijk}$ coefficients (13 diurnal * 76 spatial = 988 coefficients). These coefficients give a one day distribution of the *foF2* for all locations and local times. As it was described earlier in this thesis, the IRTAM requires background climatology coefficients and only introduces corrections for them. The *foF2* climatology coefficients used in current thesis are the Union Radio-Scientifique

Internationale (URSI) coefficients obtained by *Rush et al.* [1989]. The direct mapping of the *hmF2* was not introduced in the IRI, thus there were no expansion coefficients existing for *hmF2* mapping. The IRTAM uses the same formalism for both *foF2* and *hmF2* mapping; the *hmF2* coefficients are discussed later**.**

Keeping $G_n(\lambda,\theta)$ aside for a while, I consider how the IRI represents solar activity and seasonal variations of the ionospheric characteristics. The IRI reflects variations due to solar activity by introducing two sets of coefficients (2 * 988 = 1976 numbers). One set is for low solar activity, and another is for high solar activity. By "solar activity" the IRI implies the 12-month running mean of the sunspot number, $Rz_{12}$, which was found to have the best correlation with ionospheric characteristics. Low solar activity corresponds to $Rz_{12}=0$, and high solar activity corresponds to $Rz_{12}=100$. When the actual sun spot number is available for given time, the actual coefficients are found by interpolation between "low" and "high" coefficients sets. Seasonal variations are provided since two independent 988 coefficients sets (low/high solar activity) are given for every month of the year. These 24 (12 month * 2 solar activity levels) sets have not changed since they were introduced (for URSI option by *Rush et al.* [1989]) and don't vary from year to year.

It is known that the Earth magnetic field is not constant, and since the dynamics of the ionosphere depend on Earth's magnetic field, it is expected that coefficients should evolve over time for better representation of the ionospheric characteristics. Actually,

coefficients do not vary and the reason why IRI representation is still accurate is in the design of the geographical functions. To understand why it is not necessary for coefficients to be time dependent, consider $P_{jk}(\lambda,\theta)$ term of the formalism. $P_{jk}(\lambda,\theta)$ are designed as follows:

$$P_{jk}(\lambda,\theta) = \left(\sin\chi(\lambda,\theta)\right)^j \left(\cos\lambda\right)^k \tag{4}$$

where $\chi(\lambda,\theta)$ is the modified dip latitude introduced by *Rawer* [1963] for optimal global F-peak mapping. The choice of $\chi(\lambda,\theta)$ as parameter instead of geographical latitude showed better performance [*Jones et al.*, 1969] and $\chi(\lambda,\theta)$ has the explicit form:

$$\chi(\lambda,\theta) = \arctan\frac{I(\lambda,\theta)}{\sqrt{\cos\lambda}} \tag{5}$$

$P_{jk}(\lambda,\theta)$ suppress the magnetic field dependence at poles where solar zenith angle dependence is dominant. $I(\lambda,\theta)$ is the true magnetic dip angle, or magnetic inclination, for a particular location at height 350 km. Magnetic inclination is defined as an angle between magnetic field and horizon. The Earth's magnetic equator corresponds to zero inclination. Due to the $P_{jk}(\lambda,\theta)$ design the basis functions are not orthogonal, and the spatial expansion cannot be called "spherical harmonics". Nevertheless, this term is one of the most important ideas leading to the outstanding IRI performance. The Earth's magnetic field is not constant over time and it is different from dipole magnetic field. One can see complex structure of Earth's magnetic field on Figure 2. Due to $P_{jk}(\lambda,\theta)$

term carrying magnetic field dependence in it all formalism and coefficients stay the same. What changes and helps keep track of the magnetic field is the magnetic inclination, which makes $\chi(\lambda,\theta)$ and, hence, $P_{jk}(\lambda,\theta)$ time dependent, while the formalism is preserved. Magnetic inclination is obtained using International Geomagnetic Reference Field (IGRF) model [*Finlay et al.*, 2010].

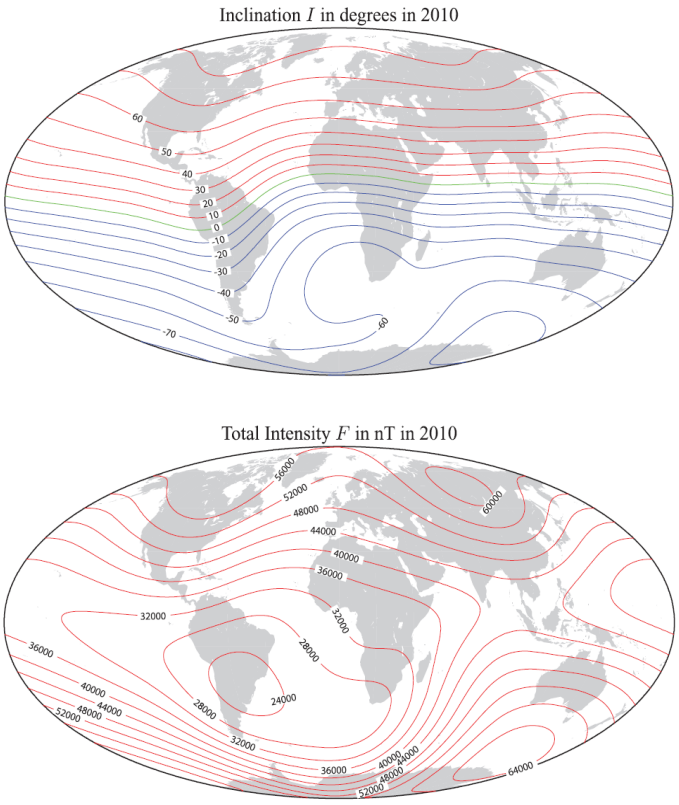



Figure 2. Contours of magnetic inclination *I* (left, units are degrees) and total intensity of the magnetic field *F* (right, units are nT) at the Earth's surface in 2010. Data obtained from model IGRF 2010. Green line shows the magnetic equator. Adapted from *Finlay et al.* [2010].

Another side of such an approach is that expansion functions have some physical meaning versus the case of a pure spherical harmonics expansion. Figure 3 maps some of geographical functions on the Earth surface.

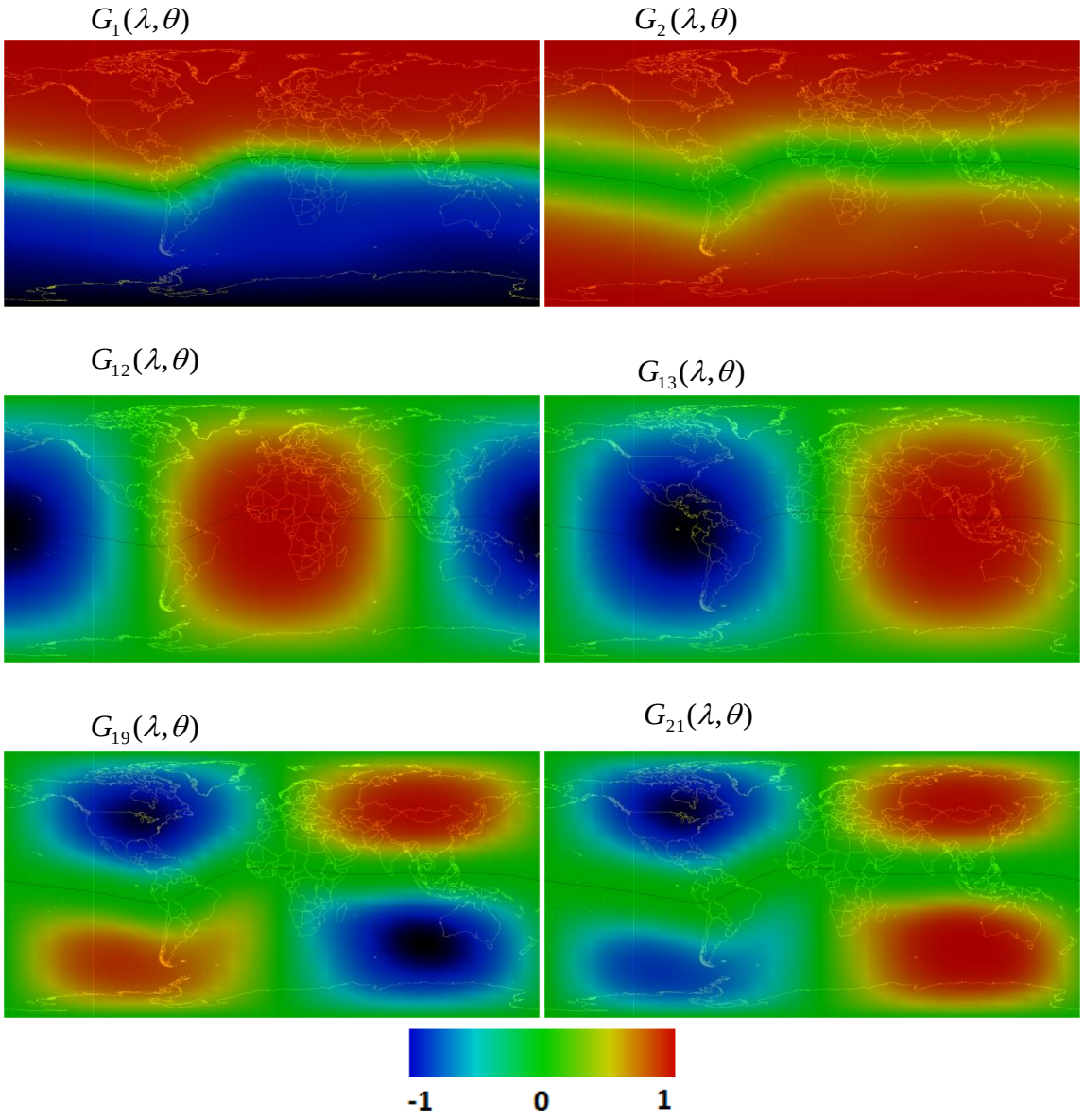



Figure 3. Some geographic functions. The Black line indicates to magnetic equator. Contour lines of the continents are shown on the background.

Figure 3 shows that geographic functions represent the geometry of the Earth's magnetic field as well as latitude and longitude variations. Some geographic functions could be linked to the particular physical features observed in the ionosphere. For example, $G_1(\lambda,\theta)$ and $G_2(\lambda,\theta)$ can be attributed to the main trend shown by the equatorial ionization anomaly. $G_1(\lambda,\theta)$ is symmetric around magnetic equator and $G_2(\lambda,\theta)$ is asymmetric. Other functions, $G_{12}(\lambda,\theta)$ and $G_{13}(\lambda,\theta)$, give semidiurnal variations. Higher order longitudinal variations could be linked to four-cell structure driven from an underlying neutral atmosphere [*Immel et al.*, 2006].

The assimilative part of the IRTAM demands data. The global option requires data to have good spatial coverage and the real-time option requires observational data be available continuously and in the near real-time regime. The GIRO satisfies both requirements. Spatial coverage of sensors is shown on the Figure 4. The SSL offers an excellent opportunity to use ionospheric data streams from GIRO (http://giro.uml.edu) for any research need. The IRTAM makes it possible to represent the weather variability in real-time and globally by assimilating observed data provided by the GIRO. Figure 4 shows locations of the ionosondes contributing data to the GIRO over the course of several decades.

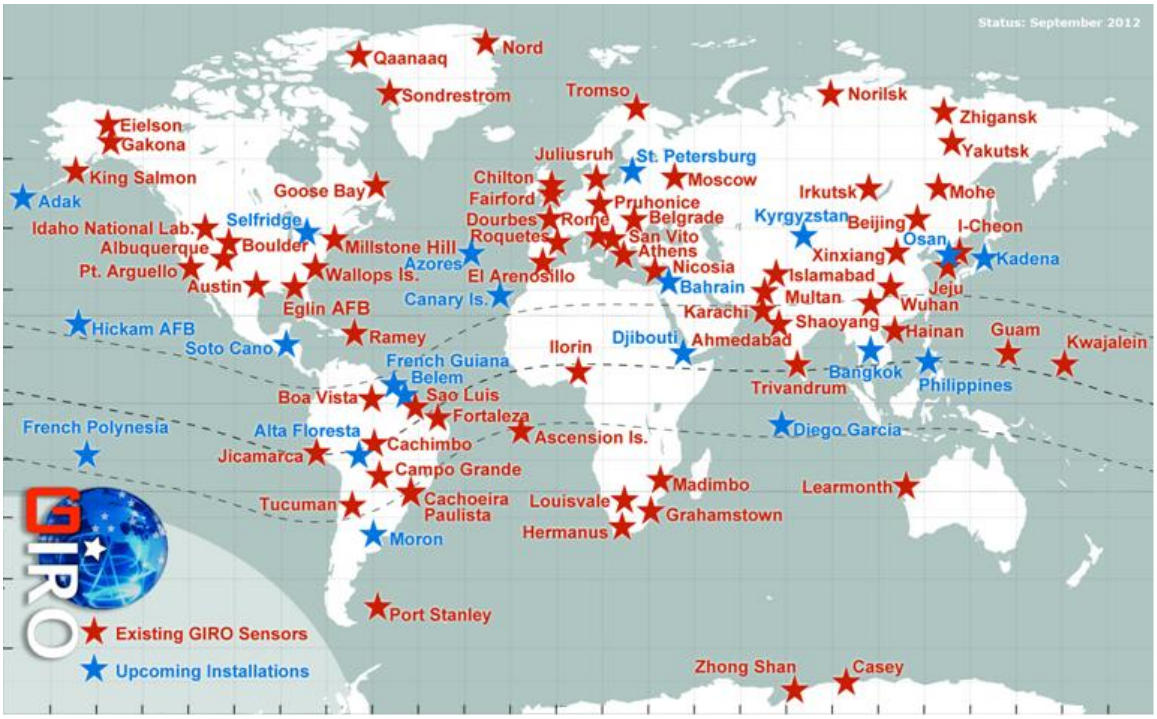



Figure 4. GIRO observatories (ionosondes) locations. Dashed curve shows the Earth magnetic equator line. Retrieved from http://digisonde.com/.

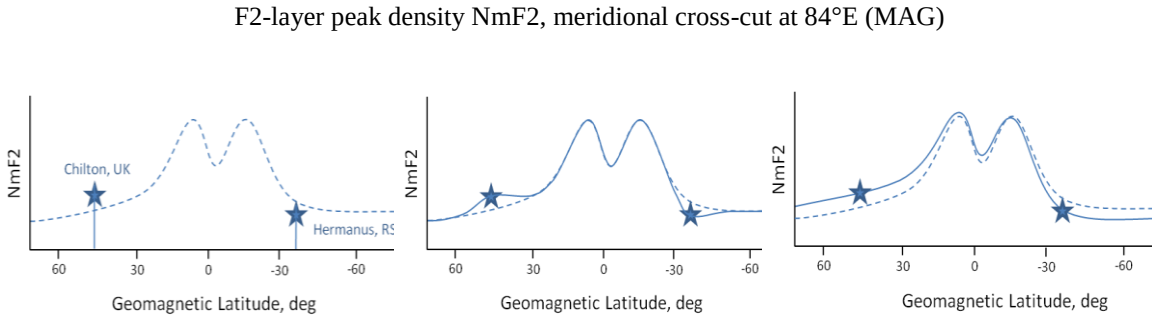
Currently, the GIRO data are the only observations fed into the IRTAM. Other independent measurements, for example GNSS TEC measurements, are used to validate assimilation results. In the future, data from other sources will be assimilated along with the GIRO data.

## 1.4 Description of the Assimilative Technique

The IRTAM task is to adjust the coefficients to match the data points while preserving the formalism of expansion functions. The assimilation approach I studied in

this thesis was developed by *Galkin et al.* [2012] and is called Non-linear Error Compensation Technique Associative Restoration (NECTAR). The NECTAR determines corrections to the climatology coefficients via direct expansion of the (observation – IRI) differences, $\varepsilon_{IRI}$. The NECTAR method introduces corrections to individual basis functions in order to bring the model closer to data, i.e., working in the coefficients, spectral, domain rather than in domain of resulting values. In order to accomplish robust and reasonable expansion of the differences, spatial interpolation of $\varepsilon_{IRI}$ to locations where the measurements are not available is done. Once this $\varepsilon_{IRI}$ map is available, it is expanded spatially using the same IRI formalism, thus obtaining 988 coefficients that represent $\varepsilon_{IRI}$. Finally, the new 988 IRTAM coefficients are obtained as the sum of the climatology IRI coefficients and the coefficients representing $\varepsilon_{IRI}$. This approach finds the best solutions without applying an iterative procedure.

The NECTAR accuracy is mainly determined by the assumptions made for interpolation. The Hopfield artificial neural network is used to expand the sparse $\varepsilon_{IRI}$ data, that are available only for station locations, to new locations [*Galkin et al.*, 2012]. This type of neural network is self-organized, and for one particular location, the value of interpolated $\varepsilon_{IRI}$ is contributed by a number of neighbors. To illustrate the morphing based assimilation capability of IRTAM technique, Figure 5 shows *foF2* profile in the meridional cross section. One can see the equatorial ionization anomaly in the electron density distribution as two peaks around the equator crossing.




(a) IRI Climatology (monthly average)

(b) Assimilation by spot-correction (instant data only, no inertia)

(c) Assimilation by morphing (24 hours of data)

Figure 5. Illustration of different correction methods. Stars indicate measured *NmF2* (*foF2*) at two stations, dashed line indicates IRI representation and solid line indicate assimilation result. Adapted from *Galkin et al.* [2012].

The middle panel illustrates the "spot correction" assimilation that changes the value of *foF2* in a small vicinity of the station locations. The NECTAR calculates corrections individually to each harmonic of the diurnal expansion, making sure the corrections do not introduce artificial sharp gradients in any direction. Thus, corrections are introduced not only in the vicinity of observational sites, but could extend spatially far from those locations. In general, IRTAM is a project that addresses real-time IRI task, and NECTAR is one possible approach to this task. Further in this thesis, when I refer IRTAM, I assume the use of NECTAR approach.

Figure 6 shows IRTAM's robustness to a data noise. To build each map, IRTAM uses a 24 hour sliding window history of observations rather than single measurement.

Single or even several erroneous data points do not affect significantly the accuracy of the IRTAM representation.

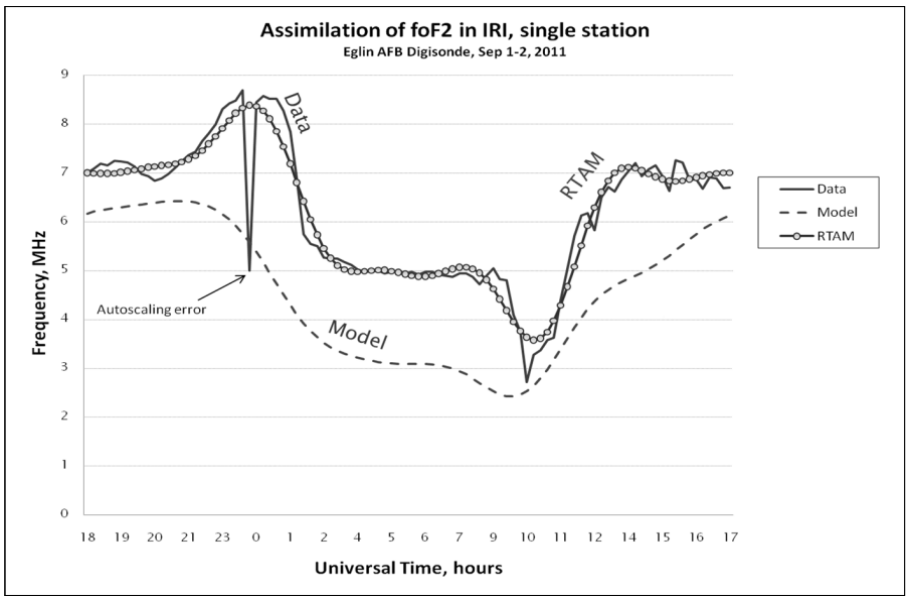



Figure 6. Illustration of IRTAM robustness to autoscaling errors. Solid line: *foF2* measurements at site Eglin (30.5°N, 86.5°W) from 1800 UT on 1 September 2011 to 1800 UT on 2 September 2011. Dashed line: IRI representation of *foF2* for September 2011. Circles: IRTAM representation. Adapted from *Galkin et al.* [2012].

Nevertheless, continuous long-lasting noise is frequently observed in the ionospheric data. Noise includes both noisy data due to the ionospheric and non ionospheric reasons. The first kind of noise source is usually linked to distortions of the ionospheric plasma, and the second is man-made interference of the sounding signal. Noise leads to the measurements that are hard to interpret and may cause autoscaling errors of *foF2* and *hmF2* deduction. The assimilative technique should be examined for robustness when such data are fed into the IRTAM.

## II. Methodology

Before providing details of validation, I want to discuss several general issues that have been resolved during the work on the thesis. These issues are not directly related to the validation, but they are worth mentioning beforehand. So far, only the *foF2* parameter was discussed. The IRTAM, as it was introduced by *Galkin et. al.* [2012], included only *foF2* mapping, the *hmF2* mapping in the IRTAM was introduced later by myself. The IRTAM uses same formalism for *hmF2* and *foF2*. This is different from original IRI mapping for *hmF2*, and the IRI does not assume independent expansion coefficients for *hmF2* mapping. In the original IRI, maps of *hmF2* characteristic are derived using *foF2* maps and an empirical relationship between *foF2* and *hmF2*. The formalism IRI model uses for *foF2* mapping was developed, and the first sets of the expansion coefficients were obtained before IRI even existed [*Jones and Gallet et al.*, 1962; *Jones et al.*, 1969]. Later on, the new coefficients for *foF2* were introduced by *Rush et al.* [1989], which the IRTAM currently uses as a climatology coefficients. Originally, the model application was seen mostly in a radio communications, especially long-wave communications, which use the ionosphere as a reflecting layer. Moreover, *hmF2* deduction from raw measurements joint with a lot more complication than *foF2* deduction, and this will be discussed in details below. Only models for *foF2* and propagation factor, *M3000F2*, and

no model for *hmF2* were introduced originally. The *M3000F2* characteristic is important parameter in long wave communications. The *M3000F2* is maximal usable frequency at 3000 km normalized by *foF2*. The maximal usable frequency is the frequency for which signal transmitted from the ground reflects from the ionosphere and comes back to the ground at distance 3000 km from transmitter. Later, it was found that *hmF2* and *foF2*, *M3000F2* are related by empirical expressions. The most advanced versions of them are introduced by *Dudeney* [1975] and *Bilitza et al.* [1979]. I will not list these expressions, since they are not of particular interest for the current thesis, but they have to be mentioned as they were used in recent work which addressed direct *hmF2* mapping [*Brunini et al.*, 2013]. The latter paper describes calculation of *hmF2* expansion coefficients using the same basis as for *foF2* mapping. First, *foF2* and *M3000F2* expansion coefficients were used to calculate 2D global maps of corresponding characteristic for each month and low/high level of solar activity. Then, these two maps were transformed into *hmF2* map, using relation between *hmF2* and *foF2, M3000F2* given by *Bilitza et al.* [1979]. Finally, the same formalism as for *foF2* were used to fit the *hmF2* maps, and hence to obtain coefficients. *Brunini et al.* [2013] presents two different approaches for *hmF2* mapping (see http://www.fcaglp.unlp.edu.ar/~fconte): 1) standard IRI formalism *foF2*, and; 2) spherical harmonics of $15^{th}$ order. The expansion with spherical harmonics implies new coefficients set for every 1 UT hour. So "24 hour", global characteristic behavior is described by 256 coefficients * 24 hours = 6144 total coefficients. I have "24 hour" in quotes, since coefficients are given for every 1 UT hour,

so if anyone is to calculate characteristic value for UT time that does not equal to integer number of hours, an interpolation between two coefficients set should be applied. As in case of the IRI formalism for *foF2*, 6144 coefficients for spherical harmonics expansion are given for high and low solar activity and every month of a year. In my opinion, the spherical harmonics approach is too massive. Since very high-order harmonics are included, there is not much real physical use of such an approach. Potentially, that approach is more accurate, but less stable.

There were three possibilities for the IRTAM *hmF2* mapping: 1) to use the IRI-way, i.e. to calculate *hmF2* as a product of *foF2* and *M3000F2* maps; 2) to use spherical harmonics formalism, and; 3) to use the same formalism as for *foF2*, new coefficients for *hmF2* and implement direct *hmF2* mapping. The first approach implies that *M3000F2* mapping has to be implemented, and it includes extra uncertainties introduced by *foF2+M3000F2* to *hmF2* transformation procedure. And this work has been done already by *Brunini et al.* [2013]. The second approach can be unstable and is not really time-dependent. It was decided to implement direct mapping of *hmF2*, using the *foF2* formalism.

Another issue that had to be resolved before performing the validation is observation jump discontinuity at the end points of 24-hour sliding window. For example, the jump in data could be as big as few MHz for *foF2*. Discontinuity appears due to different ionospheric state at the assimilation time and 24 hours before. The example of typical data discontinuity is shown on Figure 7.

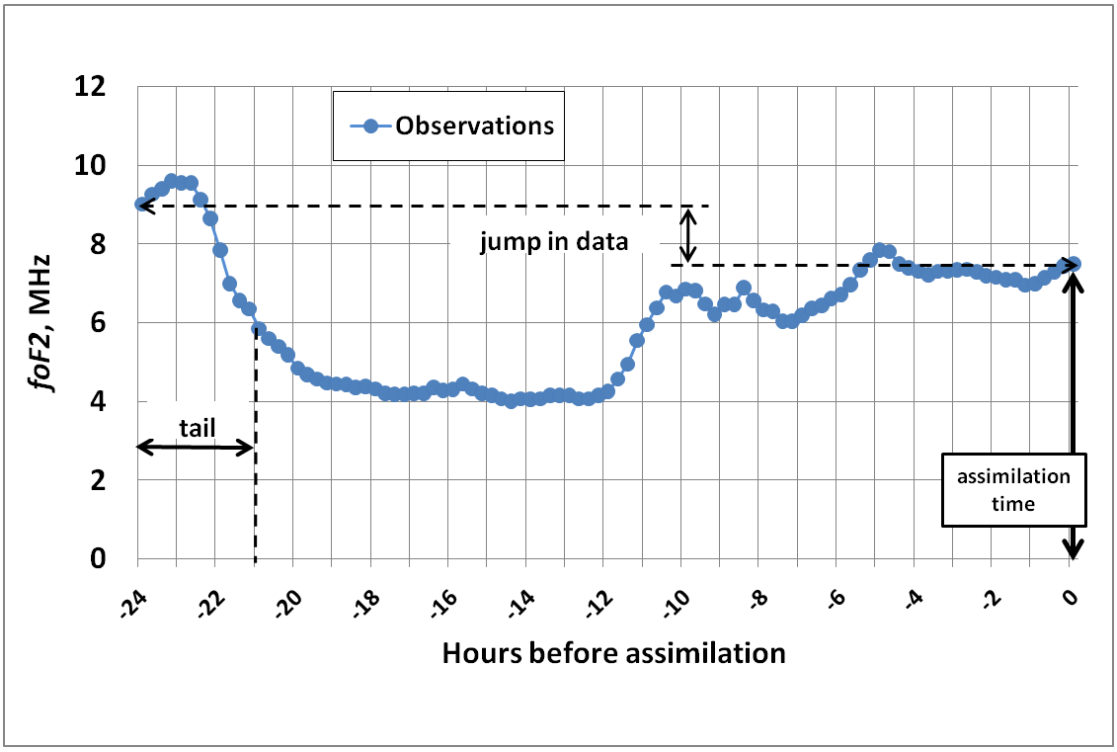



Figure 7. Observational data discontinuity at the end points of 24 hour sliding window. Horizontal axis shows hours before assimilation. Meaning of tail will be described further.

The horizontal axis shows the hours before assimilation. For example, value -4 corresponds to the data point that is 4 hours before assimilation time. Zero at horizontal axis corresponds to assimilation time. As a reminder, IRTAM uses 24 hour sliding window of measurements to perform assimilation. New coefficients are prescribed to the assimilation time (0 at horizontal axis). Nevertheless, the coefficients represent not only a characteristic at assimilation time, but 24 hour characteristic behavior. The model representation close to the both ends of 24 hour window appears to be worse due to discontinuity drawn by observations. To reproduce the time variation of a characteristic

the IRTAM uses a $6^{th}$ order Fourier series. The Fourier series describes periodic variation. In the case of IRTAM the period is equal to 24 hours. If we consider data at the time of the assimilation time and 24 hours before assimilation time (-24 on horizontal axis), the series representation will be inaccurate here. The modeled curve is continuous, and the observations have a discontinuity. The Fourier series gives same numeric value for both points (at time of assimilation and 24 hours ago), which is different from real observations. Another issue that prevents the IRTAM from accurate representation at the assimilation time is the Gibbs phenomenon. To estimate how strong discontinuity affects the IRTAM's accuracy, we perform assimilation and then collect values of (observations – model representation) for every time in 24-hour window. Then 24-hour window is shifted one time step further and the procedure is repeated. Large dataset of 13 million individual data points were used for this analysis. For each time the data were averaged by dividing by total number of data points. As the result of these calculations we end up with single average model deviation from observations for any time in 24-hour window. For the time step equals 15 minutes, the number of data points in the window is 96. These 96 values were then normalized to unity, so the largest one has amplitude of 1. Figure 8 plots the results of described analysis.

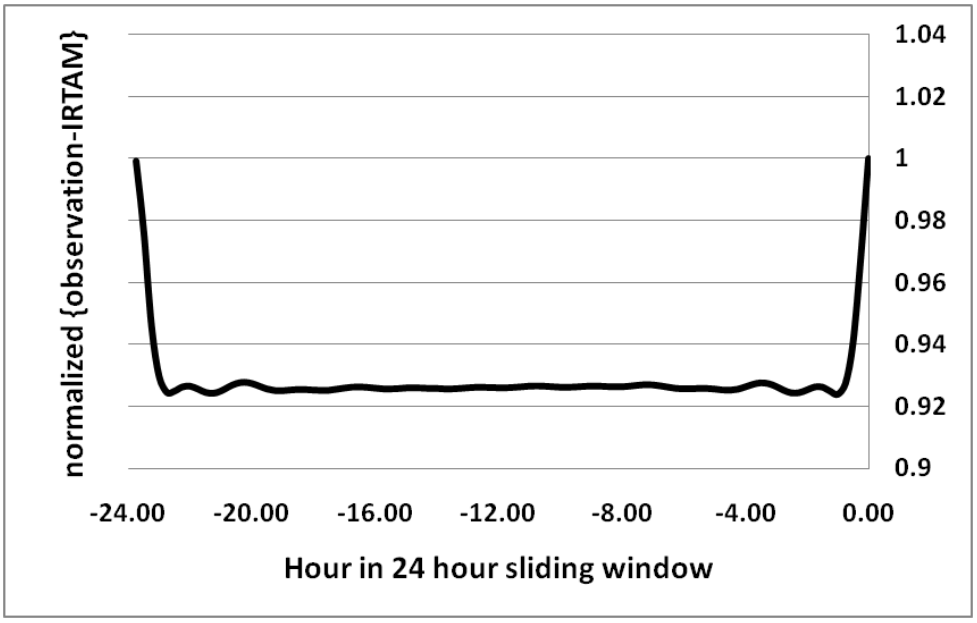



Figure 8. Average normalized $\varepsilon_{IRTAM}$ for different points within 24-hour sliding window. Plot is based on 13 million individual data points. Horizontal axis shows number of hours before assimilation. Oscillations close to the ends arise from the Gibbs phenomenon. More uniform distribution of $\varepsilon_{IRTAM}$ is better.

As expected, the accuracy is better for a data point close to the middle of 24-hour sliding window than for data points near both ends of sliding window. Nevertheless, a user specifies the time of assimilation and is most interested in accurate representation of an ionospheric characteristic at this particular time. So, the IRTAM should be able to describe observations at the time of assimilation as accurate as possible. The points 24 hour ago are least important for the user and can be sacrificed if it is resulted in better ionospheric characteristic representation at assimilation time.

When using the IRTAM in a retrospective regime, the state of the ionosphere at a particular time in the past is of the interest. Hence, the issue of discontinuity could be

avoided by taking 24 hour sliding window such that time of interest appears in the middle of the sliding window (assimilation time is still attributed to the recent time of the sliding window). The observational data are available for any time in the past, and the sliding window can be chosen arbitrarily. In this case, the time of interest is far from the discontinuity and modeled value is not affected by discontinuity. If modeling is performed for now-time, there is no an option of choosing a 24 hour window. In order to avoid undesirable errors due to discontinuity in general for any regime, I came up with the special procedure. Among all data points in 24 hour window, the most recent data point has the highest priority and the oldest data point has the lowest priority. As it was mentioned above, old data points, or tail, of 24-hour sliding window can be sacrificed in order to make representation of the characteristic at assimilation more accurate. The tail was chosen to consist of the 3-hours of the oldest data points. In the case of 15 minutes cadency, which is usual, 3 hours are equivalent to 12 data points. The tail is substituted by specially designed data such that it removes discontinuity. The procedure is as follows:

1. Observations from assimilation time back to 21 hour from assimilation time are preserved.
2. IRI modeled values of the characteristic for times when observations were dropped are calculated. Hence, there are three hours of the IRI modeled data.

3. Offset and linear shrink/expand whole three hour sequence of IRI data is done in order to match end points of the preserved 21 hours of observations. After the match is complete, data are continuous.

The results show that such an approach removes the problem of data discontinuity, and the data representation at assimilation time is as accurate as for any time in 24-hour sliding window.

At this point I will start explaining the validation of the IRTAM. The method used to conduct the validation study is generally applied for ionospheric models. Validation consists of two major parts, where the model should: 1) routinely accurate represent modeled characteristics in the absence of extreme space weather events, and; 2) adequately reproduce major ionospheric effects of extreme events. In the text below, the first test is referred as a *statistical study* and the second as a *case study*. The IRTAM maps two ionospheric characteristics, peak density recalculated to critical frequency, *foF2*, and peak height, *hmF2*. The test of model accuracy only makes sense if there is some reference value to compare with. In the case of the IRTAM, such reference points are observations provided by different sources, including ionosonde measurements from the GIRO and GNSS receivers' measurements. Only GIRO data are used to run the IRTAM, and others are used to validate IRTAM maps.

Since ionospheric characteristics have high variability it is impossible to say what values are correct without a reference point. Such reference point is, of course, *foF2* and

*hmF2* observations provided by the GIRO. To perform statistical study, I analyzed (observation – model) differences, $\varepsilon$:

$$\varepsilon = O - M \tag{6}$$

where $O$ is a observation at one of the GIRO sites for some time moment, $M$ is a modeled value of ionospheric characteristics for the same location and time. The smaller the value of $\varepsilon$ is the better a model performs. In this sense, $\varepsilon$ numerical expresses the model's accuracy. By subtracting modeled value from observation, I remove the main trend from the data and leave the part that is of the interest, i.e. model deviations from the observations. I avoid using the term "error" for $\varepsilon$, and use "deviations" or "differences" instead since, as we have learned above, the observations are not error free. With observation as a reference, I have also compared the IRTAM with the IRI. The IRI modeled values' deviation from corresponding observation gives a sense how good the climatology representation is. The IRI model value comparison with IRTAM model value shows how much ionospheric characteristics could be improved due to assimilation of the observations. Finally, the IRTAM modeled values' deviation from corresponding observations depicts how good the chosen formalism can potentially represent the ionospheric weather. The smaller the $\varepsilon_{IRTAM}$ values are, the better the formalism is.

To conduct a statistical study, we created database. The database is called Global Adaptive Modeling of Bottomside Ionosphere Timelines (GAMBIT) and contains IRTAM for the entire history of measurements contained in the GIRO. Maps are stored in compact form of 988 coefficients. The coefficients format is completely consistent with

the IRI formalism, and anyone can run the IRI code using new set of coefficients to get updated *foF2* or *hmF2* maps. The general comparison procedure is as follows:

1. Download GIRO observations.
2. Download IRTAM coefficients.
3. Calculate IRTAM maps.
4. Compare IRTAM values and observations.

This procedure is repeated for every available observation. The resulting data are then averaged for particular geographical location and/or over particular time interval in order to draw conclusions about particular features. The same set of rules and procedures were applied for both characteristics when conducting the validation. The statistical study has been conducted using observations available from October 2003 to September 2014 inclusively. Most of the stations submit measurements to the GIRO every 15 minutes, which gives 96 measurements per day. On average, 40 stations contribute to the IRTAM mapping. This gives approximately 13 million individual data points for 11 years. The time interval of 11 years is not an arbitrary value. The solar cycle is approximately 11 years long, so it is critical to have data for at least one solar cycle to avoid bias in the results. Thus, I can make sure that analysis covers all major time components of ionospheric variability: diurnal variations (periodicity – 24 hours); seasonal variations (periodicity – 12 months), and; variations due to solar activity (periodicity – 11 years). Statistical studies show how accurate a model is in average and how efficient the IRTAM is as an ionospheric model. This test reveals such general properties of the model such as

mean value and standard deviation of $\varepsilon_n$, which are commonly used in any area, as well as properties that are specific to ionospheric models. Ionosphere specific parameters show $\varepsilon_{IRTAM}$ dependence on geographic and geomagnetic coordinates, time of the day, season and solar activity phase.

When the accuracy dependence of geographical coordinates was studied, the ionosphere was divided into three regions: low latitude, mid latitude and high latitude. This classification naturally occurs due to ionospheric dynamic dependence on the Earth's magnetic field, solar zenith angle and effects of the solar wind interaction with the Earth's magnetosphere. Solar ionizing radiation gives rise to the major part of the ionospheric plasma and heats the neutral atmosphere, which interact with plasma via collisions. At low latitudes ionization in the ionosphere is larger since during day time ionosphere is subjected to higher level of solar radiation. To explain magnetic field importance, consider major ionospheric structure know as equatorial ionization anomaly. The equatorial ionization anomaly is observed in *foF2* and is lower ionization (lower *foF2*) just at magnetic equator than to the north and south from the magnetic equator. A plasma slab moves up in vertical direction at the magnetic equator due to high pressure caused by heating due to high level of solar radiation. Closer to equator the larger the pressure is. As the plasma slab moves up, it finds itself in lower surrounding pressure, so it tends to expand. This is the place where the magnetic field comes into the picture. The magnetic field does not allow charged particles move across field lines, and forces particles to gyrate around the field line (given collisions are negligible). Plasma can move

across field lines if an external agent, such as pressure due to the neutral atmosphere, is acting on it. Plasma motion along magnetic field is different and is not restricted by the magnetic field. At the magnetic equator, the magnetic field lines are parallel to the Earth's surface, and the plasma of the ionospheric F layer, where collisions are negligible, is confined into the magnetic field line. When the plasma slab is uplifted, it prefers to expand along field line. Hence, it expands from the magnetic equator to higher south and north latitudes. This continuous process of uplifting and expansion results in lower plasma peak density at magnetic equator and higher plasma peak density off equator (see Figure 5). While peak plasma density, or *foF2,* is the largest off the equator, the peak height, *hmF2,* is still the largest at the magnetic equator. In addition to the complex morphology, low-latitude region also draws high variability in ionospheric parameters. For example, sunset and sunrise at low latitudes happen at fast rates. Fast changes in the solar ionizing radiation level causes instabilities in ionospheric plasma. The plasma is perturbed which result in noise, and ionospheric characteristics are difficult to interpret from measurements. It will be shown that the *foF2* at low latitudes shows the largest discrepancies from the climatology average than at the other regions.

At high latitudes, where magnetic field lines are nearly perpendicular to the Earth's surface, plasma is free to move in a vertical direction. Variations of the ionospheric parameters at high latitudes originate from interaction of the solar wind and the Earth's magnetosphere. This is a region is more sensitive to changes of solar wind, which causes plasma convection and particle precipitation. For example, aurora glow is caused by

energetic particles precipitation from the magnetosphere to the ionosphere. At high latitude the solar zenith angle is changing gradually. For example, at the polar cap latitudes during polar day the ionosphere is subjected to solar radiation 24 hours a day. Thus, there is not much variability due to a change in solar radiation when compared to the low-latitudes. The mid-latitude ionosphere is an intermediate region between two extremes. The mid-latitude ionosphere corresponds to the closed magnetic field lines, which have both footprints on the Earth, and variability of the ionospheric parameters is less here, on average. During extreme space weather events, all regions are affected, although through different mechanisms.

The IRTAM was tested during high and low geomagnetic activity times. The geomagnetic activity is understood by the meaning of Disturbed Storm-Time (DST). The time lines of DST index were scanned to find quiet (low activity) and disturbed (high activity) time intervals. The DST indices show variations of Earth magnetic field and were retrieved from ftp://ftp.ngdc.noaa.gov/STP/GEOMAGNETIC_DATA/INDICES/DST/. The DST index is linked to variations of the Earth's ring current which circles the Earth from 3 to 8 Earth radii [*Wanliss and Showalter*, 2006]. The index is derived from the horizontal Earth's magnetic field measurements at magnetic equator. At the Earth's surface, horizontal components of the ring current magnetic field and the Earth's magnetic field are opposite to each other. A change in ring current causes changes in the measured magnetic field. During storm time, enhanced ring current leads to smaller measured value of the magnetic field compare to quiet time. Deviation of Earth's

magnetic field from its quiet time values directly corresponds to the DST index, which has nanoTesla (nT) as units. According to *Loewe and Prölss* [1997], the DST index less than -50 nT corresponds to moderate or stronger storm (more negative DST more severe storm is). The threshold of -50 nT was chosen to distinguish quiet and disturbed time intervals, for the IRTAM examination.

As it was pointed out in the introduction, the IRI coefficients depend on sun spot number. To illustrate importance of the correct sunspot number, recall that variations due to solar activity are provided by two sets of 988 coefficients for each month: one for low solar activity and another for high solar activity. When actual (observed) sunspot number is available, the coefficients that would represent ionospheric characteristics are found by interpolation between the low/high coefficients. It is worth to point out, that $Rz_{12}$ are not available immediately in the real-time. The $Rz_{12}$ is 12-month running mean of sunspot number, so precise (observed) value of $Rz_{12}$ for a particular time can be available only 6 months later. This means calculating an IRI model value for *foF2* for now-time requires that I always use a predicted value of $Rz_{12}$. Extrapolated sunspot number is frequently used and brings some uncertainties in ionospheric characteristics representation. The discrepancies due to wrong extrapolated sunspot number were most noticeable for the solar minimum of $24^{th}$ solar cycle. The solar minimum lasted much longer than expected. The extrapolated sunspot number has significantly overestimated later observed (real) sunspot numbers. The IRI predictions of the *foF2* and *hmF2* were inaccurate during the

long-lasting solar minimum, which was shown by *Cherniak et al.* [2013]. The IRI coefficients do not change with time, thus IRI can be used to model ionospheric characteristics for any time in future or past. The IRTAM performance is expected to converge to IRI performance, if coefficients are used to predict ionospheric characteristic after assimilation time. However, it has been shown, that the IRTAM coefficients give better results than the IRI coefficients. The last statement is true for the case if sunspot number is inaccurate. This feature was found during testing the IRTAM forecast capability. The forecast means that coefficients are calculated for some time, $t_{assim}$, using observations available up to this time, then same coefficients are used to represent observational data obtained later, at $t_{test}$, those were not used in the assimilation, and did not contribute to the IRTAM coefficients. This test can only be performed in the retrospective regime, when $t_{assim}$ and $t_{test}$ can be arbitrary chosen. The IRTAM accuracy is expected to converge to IRI accuracy as time passed since assimilation gets bigger. At the time of assimilation IRTAM accuracy is twofold better than IRI accuracy which will be shown in the next section. In other words, bigger $t_{test} - t_{assim}$, the closer IRTAM accuracy should be to IRI accuracy. The results show that IRTAM accuracy converges to the constant value if $t_{test} - t_{assim} > 4hours$. In some case IRTAM accuracy is better than IRI accuracy, given sunspot number is inaccurate. This occurs since the sunspot number does not drive IRTAM anymore. All variations due to solar activity are contained in observational data. The sunspot number still is important to predict ionospheric

parameters far from the GIRO sites' locations. The example of such behavior will be shown in the next section.

The IRTAM would not be called a global model if it only would be able to represent the ionospheric characteristics close to GIRO sites locations, i.e. locations where observations were obtained. In order to prove that IRTAM is able to represent the ionospheric characteristics outside the GIRO sites locations, I performed a spatial limits test. This test is designed to reveal how far from observation locations IRTAM maps can still be trusted. I excluded observations from some GIRO sites from the assimilation. Excluded data were used as if they were independent measurements. I called such excluded sites *control sites*. Particular choice of control sites is explained in the next section together with the results of this test. Even though there are no observations directly from control site location assimilated into IRTAM, the IRTAM representation is expected to be better than the climatology representation. This is expected due to observations from control site's neighbors were assimilated. The numerical values for the spatial limits of the IRTAM maps were established.

The second test major test, the case studies, depicts the IRTAM performance during particular events Solar-Terrestrial system, and is designed to show how accurate the IRTAM is as the ionospheric weather model. Term event usually correspond to the unusual behavior of ionospheric characteristics caused by some external drivers. As you might remember from the previous section, the ionospheric dynamics is driven by both the Earth's magnetosphere and neutral atmosphere. In other words, changes in

ionospheric characteristics can be driven from above and below. For case studies it is convenient to consider (IRTAM (weather) – IRI (climate)) maps, since they depict *foF2* and *hmF2* behavior relative to the climatology background. Such maps directly show enhancement or depletion of peak plasma density (*foF2* maps) and uplift or compression of the ionospheric F layer (*hmF2* maps). Since the maps are time dependent, it is easy to estimate the amplitude of the ionospheric F layer perturbations. In the text below, such maps will be referred as *ΔfoF2* and *ΔhmF2* maps. Ionospheric characteristics vary significantly over the globe. For example, *foF2* can be several times larger at low latitudes than on high latitudes for same local time. The perturbation of 1 MHz can be only 10% of quiet time *foF2* in low latitude ionosphere, while in high latitude ionosphere same 1 MHz perturbation can reach 50% of *foF2* quiet time value. Thus, sometimes it is more illustrative to consider perturbation relative to the absolute value of the same characteristic at particular location. These relative or percentage delta maps are constructed as ((IRTAM – IRI) / IRI)*100%.

I chose the following three cases: 1) interplanetary shock of 7 November 2004 [*Zong et al.*, 2010]; 2) plasmasphere refilling observations during the recovery phase of 7-8 November 2004 geomagnetic storm, and 3) Sudden Stratospheric Warming (SSW) of January 2013 [*Goncharenko et al.*, 2013]. The interplanetary shock, that passed the Earth on 7 November 2004, is the first event chosen for case study. An interplanetary shock is a common name for sudden change in a solar wind density and velocity and the interplanetary magnetic field polarization and strength. For an event to be shock three

aforementioned parameters should exhibit jump in their amplitudes simultaneously. Although both decrease and increase jumps can be called a shock, usually jump increase of the parameters associated with more energy input in the Earth's magnetosphere and causes more significant perturbations. The sudden jump of the solar wind parameters is observed when front of Coronal Mass Ejection (CME) from the Sun passes by the observer. The IRTAM was able to capture the major effects appeared due to this event. All effects are supported by observations. In this and in the third case studies I considered only the ionospheric phenomenology reported by earlier studies and the IRTAM ability to reproduce it. The physical explanation of the observed phenomena was previously reported by *Zong et al.* [2010] and is not revised in this thesis.

The second case study is different from the previous one, since I believe that I discovered a phenomenon not previously reported. Both the ionospheric phenomenology and physical mechanism responsible for observed features are presented for this case. Without going deep into other details of 7-8 November 2004 geomagnetic storm, I concentrated on *foF2* behavior during the recovery phase of aforementioned storm. It will be shown, that *foF2* was significantly decreased at middle latitudes on the day side during recovery phase of the storm. This behavior has been explained by the plasmasphere refilling process. The detailed explanation of this phenomenon is provided in the next section for the purpose of consistency. The plasmasphere is the region of the Earth's magnetosphere that extends from boundary of the ionosphere up to L-shell 4. L-shell describes the family of the dipole field lines that cross the plane containing the magnetic

equator at a certain distance from the dipole center. If the Earth's magnetic field is approximated by the dipole filed, the path along field line is described by expression:

$$r = R_E \cdot L\cos(\lambda) \qquad (7)$$

where $R_E$ Earth's radius, $\lambda$ magnetic latitude, $L$ is L-shell. For example, for L = 2 filed lines will cross equatorial plane at 2 Earth's radii from the center or 1 Earth radius from the surface of the Earth. Figure 9 illustrates concept of L-shell in meridian cross section.

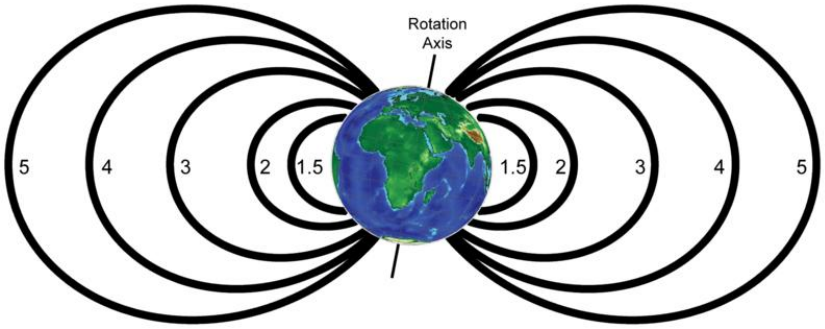



Figure 9. Meridional cross section of L-shells is shown. Retrieved from en.wikipedia.org.

The concept of L-shell is important in space physics, since plasma particles are confined into the field and the regions of the magnetosphere have geometry similar to magnetic field geometry. By considering L-shells, it is easy to understand which regions of the magnetosphere are mapped to which geomagnetic latitudes near the Earth surface. Thus, the ionosphere at particular latitudes can be linked to the particular regions of the magnetosphere. Using equation describing path along particular L-shell, it is easy to calculate to which latitudes an L-shell is mapped. For example, L-shell 4 is mapped to

60° of geomagnetic latitude, and L-shell 2 is mapped to 45° of geomagnetic latitude, which is middle latitude region of the ionosphere.

Plasma in the plasmasphere originates from the ionosphere, and it co-rotates with the Earth. The ionospheric plasma "pumped" into the plasmasphere on the day side and on the night side plasmaspheric plasma is "dumped" back into the ionosphere. In quiet geomagnetic conditions plasmapause, boundary of the plasmasphere, is usually located at L = 4. The plasmapause is defined as topological boundary between co-rotating field lines of the plasmasphere and convective field lines coming from the magnetospheres' tail due to reconnection. From the observation perspective, the plasmapause corresponds to the sharpest gradient of the plasma density in the equatorial plane [*Goldstein et al.*, 2003]. During the storm, time enhanced plasma convection from the tail disturbs the equilibrium between co-rotating and convective regions, and some of outer co-rotating plasmaspheric field lines become convective ones. Equivalently, the plasmapause moves to the smaller L-shells. The plasma from the L-shells beyond plasmapause is swept away, which is supported by observations [*Reinisch et al.*, 2004]. At the recovery phase of the storm plasma convection from the tail reduces up to its quiet time state. Thus, the plasmapause also moves to its quiet time position. The equilibrium between co-rotating field lines and convective ones restores. Restored co-rotating field lines (between storm time and quiet time plasmapause L-shells) are only partially filled with plasma compared to the quiet time. At this point, refilling of the plasmasphere with the ionospheric plasma starts until the plasmasphere reaches its quiet time state. *Reinisch et al.* [2004] showed

that during another storm started on 31 March 2001 plasmapause moved from L > 3.6 to the L~2.2. Plasma density during storm dropped to 30% percent of its quiet time density and it took more than one day to refill plasmasphere to 100%.

To support features observed in IRTAM maps, I analyzed data from Radio Plasma Imager (RPI) instrument [*Reinisch et al.*, 2000] on Imager for Magnetopause-to-Aurora Global Exploration (IMAGE) spacecraft. The IMAGE is orbiting at near polar orbit, and the orbital period is approximately 14 hours. Part of the orbit passes through plasmasphere, where the RPI record plasmagrams. A plasmagram is the amplitude of the sounding radio waves plotted on the range-frequency grid [*Reinisch et al.*, 2001a]. A plethora of useful information is derived from plasmagrams. For the current study, I was only interested in the plasma density profiles of the plasmasphere. The plasma density profiles were derived from the traces of the signal propagating along the magnetic field line. Hence, a plasma density profile can be prescribed to the specific L-shell. Another important attribute if the plasmagram is local time (LT), more precisely magnetic local time (MLT). MLT has direct analogy to the LT where all point along a particular geographical meridian has the same LT. In case of MLT all point along geomagnetic meridian have the same MLT. The results of this analysis are described in the next section.

The third case study is ionospheric effect of January 2013 SSW. SSW is large-scale meteorological event that originates from neutral atmosphere and affects the ionosphere. Although strength of SSW varies from year to year, it may cause significant

ionospheric variability. The number of SSWs occurs during observational history and the ionospheric effects of SSW are too numerous to study. I restricted the current study to the January 2013 SSW effects on the low/mid-latitude ionosphere. The SSW in 2013 was a major event, and showed pronounced discrepancy from the quiet time. Two independent observational data sets were used: GPS TEC measurements and GIRO measurements. Only the latter was used to run IRTAM. GPS TEC measurements were acquired from Madrigal data base of Haystack Observatory, http://madrigal.haystack.mit.edu/madrigal/. Although they are not exactly proportional to each other, *foF2* and TEC have positive correlation. This kind of comparison is reasonable, and *ΔfoF2* should show the same features as ΔTEC. The IRTAM was able to reproduce major features shown by GNSS TEC measurements; however one effect was missing due to limited number of GIRO sites. A detailed comparison and discussion of this case is given in the end of the next section.

# III. RESULTS

## 3.1 Statistics

Before trying to apply the IRTAM to particular space weather events, I have conducted a comprehensive statistical study. These statistics are a necessary piece of information for IRTAM users, since they are interested in range of applicability and trustworthiness of the models. First, I calculated average value of absolute $\varepsilon$, $\bar{\varepsilon}_M$, using the expression:

$$\bar{\varepsilon}_M = \frac{\sum_{n=1}^{N} |O_n - M_n|}{N} \tag{8}$$

where $O_n$ is an observation, $M_n$ is a model value, and *N* is number of data points. Since there are two models, and sometimes it is necessary to refer both of them at the same time, I used both $\bar{\varepsilon}_{IRTAM}$ for the IRTAM and $\bar{\varepsilon}_{IRI}$ for the IRI. Both $\bar{\varepsilon}_{IRTAM}$ and $\bar{\varepsilon}_{IRI}$, along with the improvement factor, which is the ratio between $\bar{\varepsilon}_{IRTAM}$ and $\bar{\varepsilon}_{IRI}$, are listed in Table 1. I examined the IRTAM in three different regimes: 1) using only manually scaled measurements (by human) available for 11 years (~ 1 million data points); 2) using all measurements available, including auto-scaled ones (by software), for 11 years (~13

million data points), and; 3) using only measurements made during geomagnetic disturbed time. On the next page, I explain the reasons why these regimes were used.

Table 1. Average values of absolute $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$.

| *foF2* | IRI error, MHz | IRTAM error, MHz | Impro-vement |
|---|---|---|---|
| Only manual scaling for 11 years | 0.69 | 0.31 | 2.25 |
| All measurements for 11 years | 0.76 | 0.38 | 2.00 |
| Disturbed time (DST < -50 nT) | 1.05 | 0.49 | 2.12 |
| *hmF2* | IRI error, km | IRTAM error, km | Impro-vement |
| Only manual scaling for 11 years | 27.0 | 15.0 | 1.82 |
| All measurements for 11 years | 33.0 | 18.0 | 1.83 |
| Disturbed time (DST < -50 nT) | 40.0 | 23.0 | 1.74 |

*Manually scaled measurements*: From description of the IRTAM it is known, that the IRTAM is robust to auto-scaling errors, but obviously auto-scaling errors contribute to $\bar{\varepsilon}_{IRTAM}$. If auto-scaling measurements are taken into account, $\bar{\varepsilon}_{IRTAM}$ increases and the IRTAM accuracy will decrease, which does not reflect real IRTAM performance. The best way to estimated pure IRTAM accuracy is to exclude all auto-scaled measurements and leave only manually scaled measurements when $\bar{\varepsilon}_{IRTAM}$ is calculated. The manually scaled data are assumed to be error free, while auto-scaled data can contain errors. Unfortunately, manually scaled data are not continuous and do not have global coverage

to run both assimilation and validation using purely manually scaled data. Hence auto-scaled data are still assimilated. For *foF2* and *hmF2,* I have obtained average errors equal to 0.31 MHz and 15 km respectively. The improvement factor is 2.25 for *foF2* and 1.82 for *hmF2*. The numbers given above can be thought as lower limit for $\bar{\varepsilon}_{IRTAM}$ at current stage of the IRTAM development.

*All measurements*: All available measurements represent more significant statistical set, which includes 13 million data points. Manually scaled data are available only for some locations, and timelines of such data, except few locations, are not continuous. Hence, users will run the IRTAM using dominantly auto-scaled data. To estimate the real IRTAM performance, I calculated $\bar{\varepsilon}_{M}$ using all data available for 11 years. For *foF2* $\bar{\varepsilon}_{IRTAM}$ is 0.38 MHz and $\bar{\varepsilon}_{IRI}$ is 0.76 MHz. The IRTAM shows an improvement of 2x over the IRI. For *hmF2* $\bar{\varepsilon}_{IRTAM}$ is 33.0 km and $\bar{\varepsilon}_{IRI}$ is 18.0 km, which gives improvement factor equals to 1.83. The $\bar{\varepsilon}_{M}$ values of both models are large for this regime than for the regime when only manually scaled data were used. As noted above, auto-scaling errors contributed into the average error. One can see contribution due to auto-scaling error by comparing the first and the second rows for each characteristic. The discrepancy between the first and the second row can give only lower limit of noisy data contribution to the $\bar{\varepsilon}_{IRTAM}$ . Only auto-scaled errors were excluded and only from $\bar{\varepsilon}_{IRTAM}$ calculation, and the other types of errors were not inspected.

*Disturbed time*: The IRTAM is designed to be an ionospheric weather model. This is the reason for testing the IRTAM during geomagnetic disturbed time. The third row for each characteristic shows $\bar{\varepsilon}_M$ as it was calculated using only measurements made during disturbed time. Both manually and auto-scaled measurements were used. Geomagnetic activity of moderate level or more severe was understood by disturbed time. This is equivalent to DST index less than -50 nT. The results show that absolute values of $\bar{\varepsilon}_M$ are larger, but their ratio, i.e. improvement factor, stays almost the same when compared to the improvement factor from the second row. For the *foF2* improvement factor is even larger than the improvement factor during quiet time (shown in the second row). This is reasonable since the IRTAM is designed to capture perturbations of the ionospheric characteristics by data assimilation, while the IRI does not have this option.

To show what the $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ distributions are, I plotted $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ histograms for *foF2* and *hmF2*. Each $\varepsilon_{IRTAM}$ or $\varepsilon_{IRI}$ was prescribed to the corresponding bin, and total number of $\varepsilon_{IRTAM}$ or $\varepsilon_{IRI}$ at particular bin was counted. Then the resulting distribution was normalized to unity and multiplied by hundred to see which percentage of $\varepsilon_{IRTAM}$ or $\varepsilon_{IRI}$ falls into which bin. Figure 10 shows the histograms for both models and both characteristics.

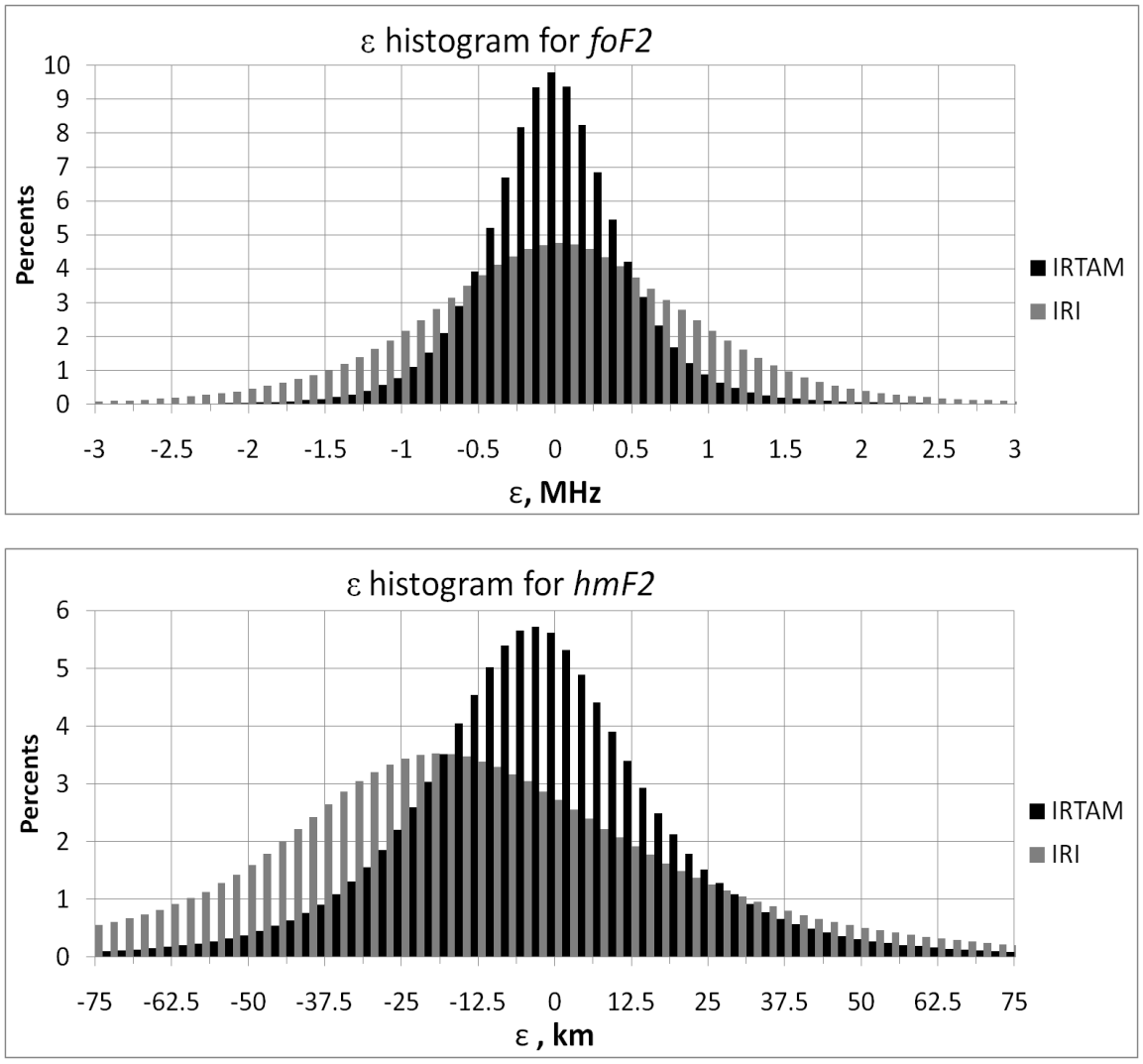



Figure 10. $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ histograms for *foF2* (top) and *hmF2* (bottom).

The top panel of Figure 10 shows *foF2* $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ histograms. Both IRTAM and IRI histograms are zero centered Gaussian-shaped distributions. Zero-centered distribution for the IRI model again indicated that the IRI is good climatology model and URSI coefficients are reliable and accurate. The bottom panel of Figure 10 shows *hmF2*

histograms. Both histograms are Gaussian-shaped distributions, but the IRI histogram is not zero centered. Although I am primarily interested in the IRTAM results, which show almost zero-centered distribution, this issue is worth discussing in detail. The IRI is the basis for the IRTAM, thus by improving IRI, I will also improve IRTAM. A detailed discussion of the bias drawn by IRI *hmF2* histogram is given in the next section. The mean values and standard deviations of $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ are summarized in Table 2.

Table 2. Mean and standard deviation for *foF2* and *hmF2*.

| | *foF2*, MHz | *hmF2*, km |
|---|---|---|
| IRTAM mean | 0.01 | -1.7 |
| IRTAM standard deviation | 0.56 | -14.3 |
| IRI mean | 0.01 | 20.3 |
| IRI standard deviation | 1.05 | 41.6 |

The mean value of $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ distributions for *foF2* are both 0.01 MHz. They can be counted equal zero since 0.01 MHz is several times smaller than the ionosonde's resolution for critical frequency measurements, which is usually in range from 0.025 MHz to 0.05 MHz. The mean values of $\varepsilon_{IRTAM}$ distribution for *hmF2* is -1.7 km, which means that IRTAM corrected the IRI mapping to minimize bias. These results are consistent with the results of *Damboldt and Suessmann* [2011]. The standard deviation is higher for the present study, since *Damboldt and Suessmann* [2011] performed median-to-median comparison, and I compared individual data points.

## 3.2 Geographical and Local Time Distribution

The reasons why ionospheric characteristics vary with geographical and geomagnetic coordinates are discussed in the previous section. To examine latitudinal and longitudinal dependence of the models' accuracy, I used the same expression for $\bar{\varepsilon}_M$ with a small correction. Here the data are averaged not over the entire dataset, but over a particular site location, i.e. to a particular ionosonde. Since sites are distributed all over the globe, I can estimate accuracy at various locations $\bar{\varepsilon}_M\big|_{site}$ for both IRTAM and IRI models. After $\bar{\varepsilon}_M\big|_{site}$ have been calculated at all sites, they could be plotted as a function of the geographic or geomagnetic coordinate. Figure 12 shows $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ of *foF2* dependence on geographical latitude.

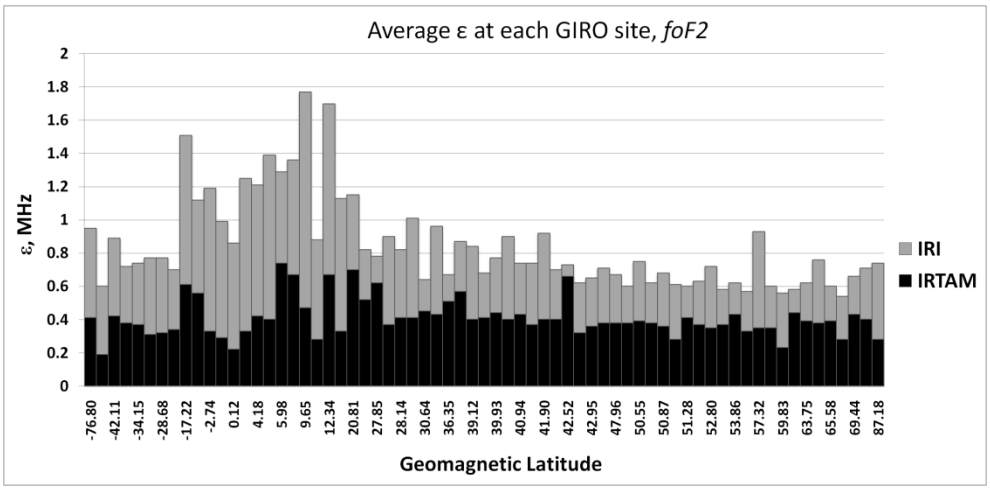



Figure 11. Dependence of average (over particular location) $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ on geomagnetic latitude for *foF2*. Each bin coresponds to indvidual GIRO site. Negative values on horizontal axis correspond to sothern hemisphere. Horizontal axis is not uniform, and values correspond to obsercational sites coordinates.

For the IRI model, bigger errors seem to group at low latitudes. This is not surprising, since the equatorial ionosphere is very dynamic and draws high variability in *foF2*. Moreover, IRI shows higher deviations from the observations for the southern hemisphere. In the time that the IRI coefficients were created, there were not many observations from southern hemisphere [*Jones et al.*, 1969]. We see that the IRTAM almost corrects both issues. We see that IRTAM data are more uniformly distributed with latitude, although there are some locations with higher average error. I examined some of them, and the result is that high deviations are mostly due to noisy measurements.

Longitudinal $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ dependence plotted in Figure 12 does not show significant patterns.

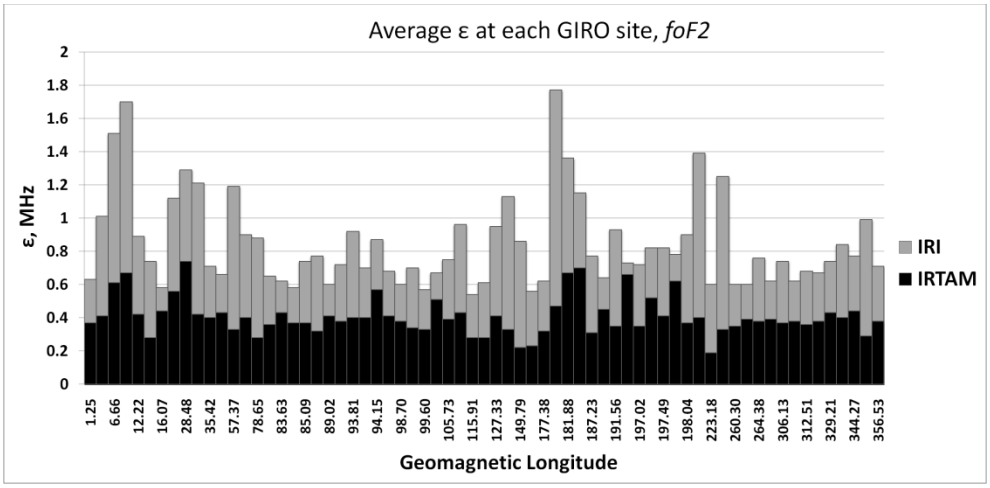



Figure 12. Same as Figure 11, but for dependence on geomagnetic longitude.

It could appear that there are two clusters of high deviations. It does not reflect longitudinal variations and is due to sites from equatorial regions and the southern

hemispheres which happen to have these longitudes. When latitudinal dependence is under investigation, data could be average over longitude. To find longitudinal variations, it is critical to have data from similar latitudes and different longitudes. Currently, the GIRO sites spatial coverage does not allow performing accurate studies on the longitudinal variations.

Figure 13 shows IRTAM *hmF2* mapping accuracy dependence on geographical latitude.

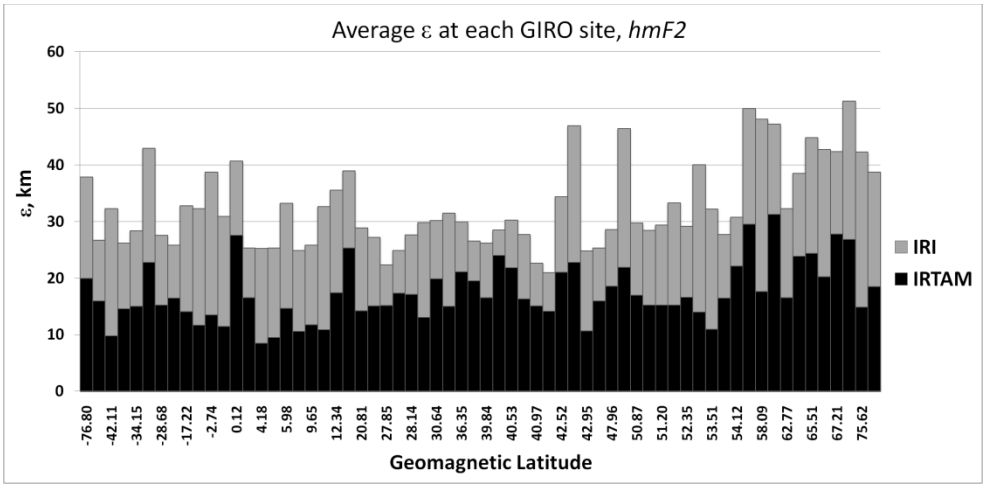



Figure 13. Same as Figure 11, but for *hmF2*.

In this case $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ are not as uniform as for *foF2*. This behavior led me to the discussion of the *hmF2* mapping uncertainties, which will be given below. The main feature depicted by Figure 13 is that *hmF2* modeling becomes less accurate at high latitudes, which is opposite to the *foF2*. The reasons for this are described in the previous section, and observed the picture is basically results from magnetic field geometry. At

high latitudes, the field lines are nearly perpendicular to the Earth's surface, which allows plasma to move in vertical direction more freely than at low latitudes, where the field lines are nearly parallel to Earth's surface.

Furthermore, I studied IRTAM mapping accuracy dependence on local time. The ionosphere was divided on three regions: low, middle and high latitude ionosphere. Geomagnetic latitude was used to attribute a particular site to a particular region. Low latitude region contains all site with geomagnetic latitude less than $20°$. Sites with geomagnetic latitudes greater than $60°$ are prescribed to high latitude region. All other sites belong to middle latitude region. Each site was attributed to one of three regions. Each individual $\varepsilon_M$ for any region was attributed to a particular local time. Local time for each data point was calculated using expression:

$$LT = UT + 1hour\frac{longitude}{15°} \tag{9}$$

where *LT* is local time, *UT* is universal time, *longitude* is the geographical longitude of an ionosonde location, appears from total longitude degrees divided by the number of hours in one day $360°/24 = 15°$. $\bar{\varepsilon}_M$ is calculated for each LT and latitude region. All observations available for 11 years were used. Figure 14 shows resulting plots for these three regions.

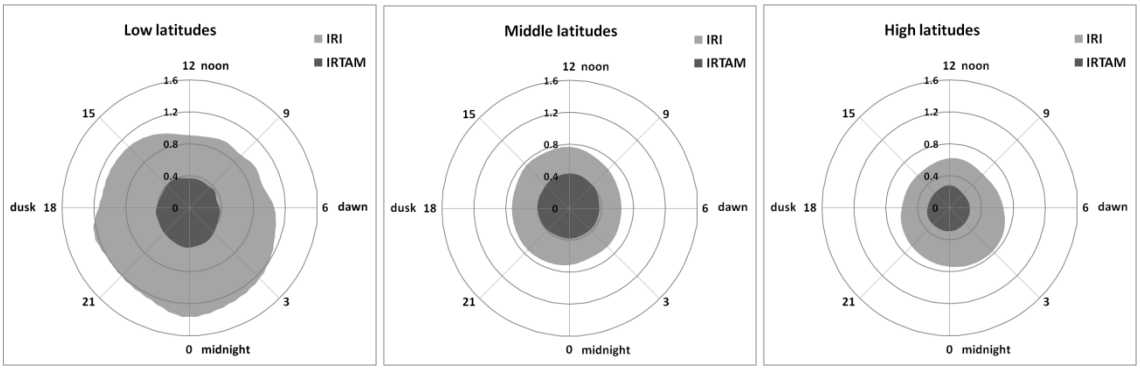



Figure 14. Dependence of *foF2* $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ on local time for three latitude regions. Low latitudes (left), middle latitudes (middle) and high latitudes (right).

Such polar plots are commonly used in Space Physics. Numbers around the circle show the local time: 12 is local noon, 0 is midnight, 6 is dawn and 18 is dusk. $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ for each local time are shown as a distance from the origin. The smaller the distance from the origin, the better model accuracy is. The left plot of Figure 14 shows that the equatorial ionosphere is a very dynamic region. Ionospheric conditions there can be far from the climatology average, when no extreme space weather events occur. For equatorial ionosphere, the improvement factor is the largest. The ionospheric characteristics are further from the climatology average around dusk and midnight time. The IRTAM deviation from the observations is largest during the night time, but still the IRTAM representation is ~4 times better than the climatology representation. For other two regions, there is no significant IRTAM accuracy dependence on local time, and the improvement is about 2 times.

Further the *hmF2* representation accuracy was examined. Same plots as on Figure 14, but for *hmF2* are shown on Figure 15.

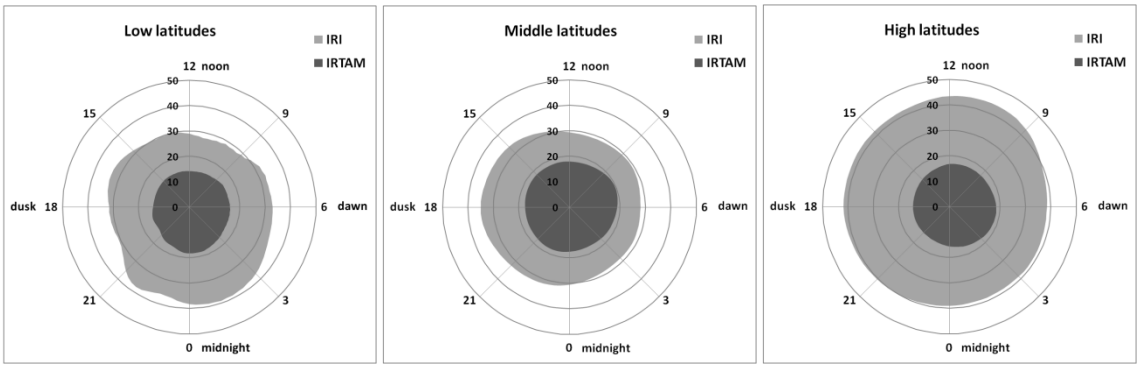



Figure 15. Same as Figure 14, but for *hmF2*.

The left panel shows that $\varepsilon_{IRI}$ is not as uniform with local time at low-latitudes as in the other two regions. The highest deviations appear at night time, and IRTAM did not show significant dependence on local time. The mid-latitude ionosphere shows small local time dependence for both IRI and IRTAM, although IRTAM average deviations from the observations are bigger than for the low latitude ionosphere. As expected, the accuracy of the modeling is lowest in the higher latitudes due to the topology of the magnetic field. There is no significant dependence of LT and the improvement factor is the highest for high latitudes.

## 3.3 Seasonal and Solar Cycle Variations

The seasonal variations of the ionospheric characteristics appear due to different level of solar ionizing radiation for particular geographical location over the course of the year. To study seasonal variations of the IRTAM modeling, it is illustrative to divide a year on four parts: summer solstice, winter solstice, spring equinox and fall equinox. For each season $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ were collected independently. The IRTAM did not show significant discrepancy between these four periods. The average deviation for *foF2* is 0.38±0.025 MHz, and for *hmF2* the average deviation is 18.0±1 km. The uncertainties (0.025 MHz and 1 km) are in the range of *foF2* and *hmF2* measurements uncertainties. Thus, the dependence of the IRTAM modeling on season is negligible.

Variations due to solar activity appear due to different level of solar ionizing radiation over the course of the solar cycle. The sunspot number varies even more significantly than the radiation level, from near zero at solar minimum to more than hundred at solar maximum. Sunspots are the sources of the fast solar wind, and the fast solar wind cause perturbation of the ionospheric characteristics. At solar cycle maximum, larger sunspots number result in higher probability that fast solar wind buffets the Earth's magnetosphere. Hence, during solar maximum the deviations from climatology are larger than during solar minimum. To study solar cycle variations $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ were collected for each data point over 11 year time interval from 2003 to 2014. Then I averaged this data set using median filter with size equals to one month. Averaged $\varepsilon_{IRTAM}$

and $\varepsilon_{IRI}$ for *foF2* along with averaged DST are presented in Figure 16. The *hmF2* version of the Figure 16 is not shown here, since it depicts the same pattern.

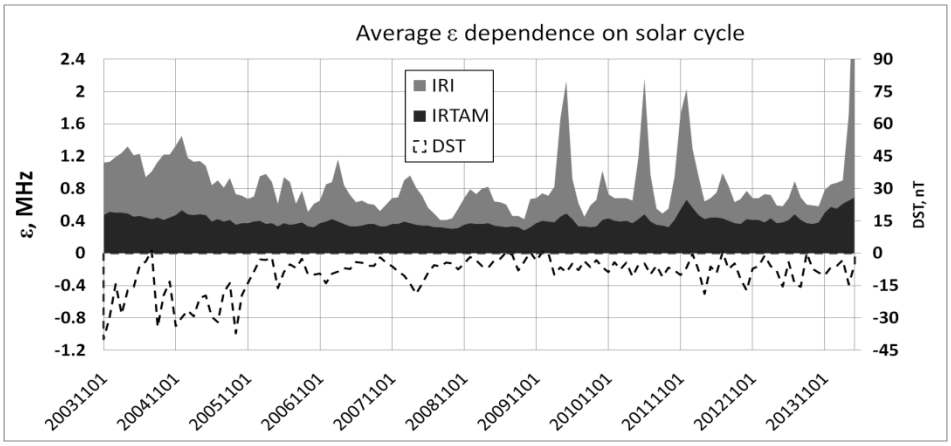



Figure 16. $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ dependence on solar cycle phase. Only *foF2* data are shown. Dashed line represents DST index. Time interval equals to 11 years. The DST and $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ are averaged with median filter 30 days long.

The IRI is more sensitive to space weather events, and the IRTAM performs rather well during disturbed time, which is seen from the 2003-2004 data. During high geomagnetic activity the climatology representation is less accurate, which is obvious from the comparison of DST index and $\varepsilon_{IRI}$ . This is reasonable and was explained above. There are three major spikes shown by IRI during 2010-2011 years that do not correspond to high geomagnetic activity, and average DST stays below -10 nT. This means that the source of the high deviations from climate is not in geomagnetic activity. The study of these spikes is not directly related to the IRTAM validation and is left for future research.

## 3.4 Forecast Capability

Currently, the IRTAM does not have forecasting option, but it is imperative to research its forecast capabilities. In the previous section, I explained how the forecast capability can be examined by using retrospective data. Figure 17 shows the IRTAM forecast capability together with backcast accuracy. The backcast accuracy refers to the IRTAM ability to model all data points equally accurate in 24-hour sliding window. As a reminder, the special procedure was developed in order to avoid inaccuracy due to observational data discontinuity at the ends of a 24 hour sliding window. I provided the backcast result to show that the IRTAM accuracy is uniform for all points in the sliding window, and the problem observed in Figure 8 is resolved.

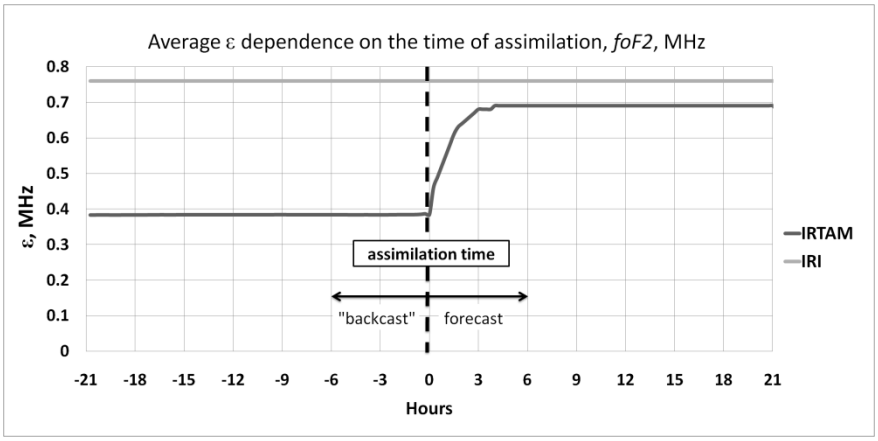



Figure 17. Illustration of the IRTAM forecast capability (positive horizontal axis) and backcast accuracy (negative horizontal axis). Numbers on the horizontal axis show number hours after/before assimilation. Plot is based on 2010-2013 data. Value of $\varepsilon_{IRI}$ is given for reference and does not depend on time.

First, consider the parts of the curves in Figure 17 corresponding to negative values on horizontal axis and showing the IRTAM backcast accuracy. The negative hours mean that a data point was taken within the 24 hours sliding window, and absolute value shows how many hours before assimilation time it appears. The negative part of the curve is horizontal, which means that, in average, all data points are modelled equally accurate. Now, consider the part of the plot corresponding to the positive hours. The number of hours on horizontal axis equals to the number of hours passed since the assimilation time. For example, a value at +3 hours means that coefficients were calculated using data obtained three hours ago. The IRTAM accuracy is higher as time of interest is closer to assimilation time. The behavior of $\bar{\varepsilon}_{IRTAM}$ is most interesting after 4 hours after assimilation time. Although $\bar{\varepsilon}_{IRTAM}$ takes constant value around 0.69 MHz, it is not equal to $\bar{\varepsilon}_{IRI}$, which is 0.76 MHz. The reason for that is in inaccurate extrapolated sunspots number used to drive the IRI. It is important to note that only data for the years 2010-2013 were used to build Figure 17. Measured sunspot number was available only for 2010, and for 2011-2013 sunspot numbers were extrapolated, thus were not precise. Since sunspots numbers are not available immediately in real-time, the delay can be more than a year. The IRTAM avoids the problem of precise sunspots number by using real-time observations that has solar activity "built in".

## 3.5 Spatial Limit of Assimilative Maps

The previous subsection discussed prediction in time. This subsection will explain prediction in space. So far, I examined the IRTAM using only the same data that were assimilated. Observations that went to assimilation and those used to calculate $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ are exactly the same observations. Thus, all studies completed above can be considered as a test on the formalism ability to reproduce observations in various conditions.

In this subsection the IRTAM accuracy is examined outside the observational sites' locations. The ionosphere is a medium which displays high variability, but it does not change randomly. In absence of extreme space weather events, which are rare, the time lines of an ionospheric characteristic, *foF2* or *hmF2*, at two spatially close locations should not be much different from each other. To understand how far from observational sites locations the IRTAM maps are still reliable, I assigned some sites to be control sites. Control sites do not participate in the assimilation, they are only used to calculate $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$. Hence, we can estimate spatial limits of the IRTAM maps. The IRTAM uses the same formalism to represent both *foF2* and *hmF2* characteristics. Since spatial limits are tested, I used only *foF2* observations in this study. The *foF2* characteristic was preferred over *hmF2*, because *foF2* measurements are more precise, which is illustrated in Figure 10, and is discussed in the next section.

Mid-latitude ionosphere was examined first. I chose the North American sector to perform this study, since it has a somewhat spatially uniform distribution of observational sites. The sites with coordinates in range from $15°N$ to $45°N$ by geographical latitude and from $65°W$ to $130°W$ by geographical longitude were selected. Although geographical latitudes are low for some sites, the geometry of the magnetic field at this region is such that all sites except the PRJ18 (see below) are attributed to the mid-latitude region. The site PRJ18 is located in the intermediate region between low and middle latitudes, so caution should be taken when considering this particular site data. The high latitudes sites were also excluded from the analysis. Data available from 1 January 2013 to 30 September 2014 were considered for this study. The sites were checked for the number of the measurements available for every day in this time interval. Only the days with number of the observations greater than 24 per day, which were distributed uniformly over a day, were counted. The data availability is shown in Figure 18.

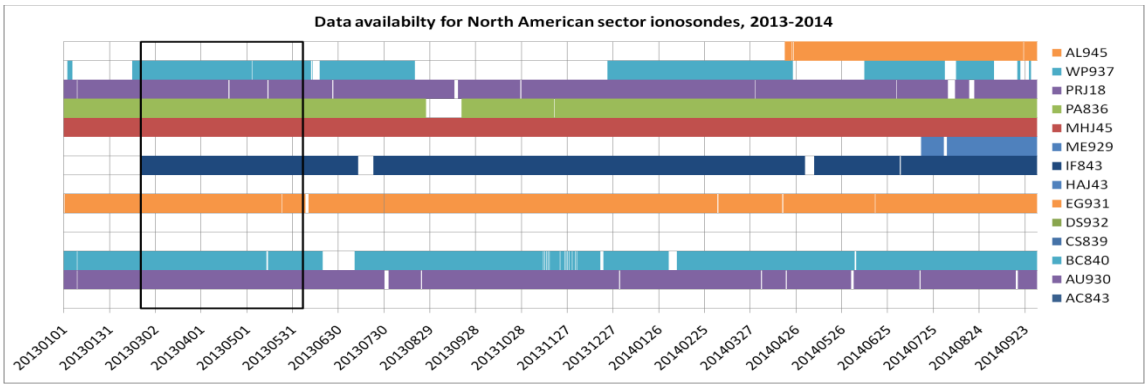



Figure 18. Data availability plot for sector from 15°N to 45°N and from 65°W to 130°W. All sites ever were installed are shown. Each horizontal stripe corresponds to individual site shown on the left. Black rectangle denote the time period chosen for further study.

The horizontal axis of Figure 18 shows time (from 0000 UT on 1 January 2013 to 2359 UT on 30 September 2014). Each horizontal row corresponds to an individual site, which are labeled on the right. A filled portion of the row denotes that observations were available for given time, and a blank space denotes that observations were not available. Sites HAJ43, DS932, CS839 were not operating during the given period of time. To exclude possible bias due to patchy data availability, I chose time interval from 2013.02.19 to 2013.06.11, when data from eight sites were continuously available. The sites are PA836 (Pt. Arguello 34.8°N, 120.5°W), IF843 (Idaho National Lab 43.8°N, 112.7°W), BC840 (Boulder 40.0°N, 105.3°W), AU930 (Austin 30.4°N, 97.7°W), EG931 (Eglin 30.5°N, 86.5°W), PRJ18 (Ramey 18.5°N, 67.1°W), WP937 (Wallops Island 37.9°N, 75.5°W), MHJ45 (Millstone Hill 42.6°N, 71.6°W). The geographical locations of the sites are shown in Figure 19.

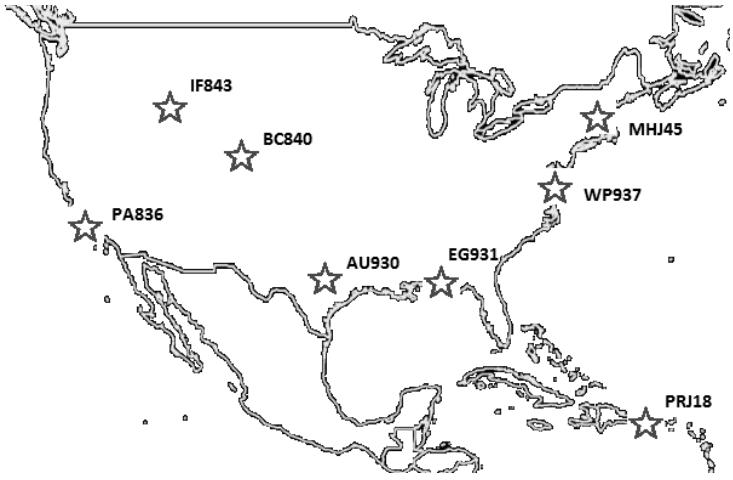



Figure 19. Geographical locations of the sites used for spatial limits study in North American sector are shown. Each site location is shown by star and labeled appropriately.

I chose several test configurations:

1. Internal test: observations from PA836, IF843, PRJ18 and MHJ45 sites were assimilated, BC840, AU930, EG931 and WP937 sites were assigned control sites.
2. External test: opposite of the internal test
3. Advanced external: only observations from AU930 and EG931 sites took part in the assimilation, BC840, PA836, IF843, PRJ18, WP937 and MHJ45 were assigned control sites
4. Higher latitude test: observations from PA836, AU930, EG931 and PRJ18 sites took part in the assimilation, IF843, MHJ45, BC840 and WP937 were assigned control sites
5. Lower latitudes test: opposite of the high latitude test

Accuracy of the mapping was estimated using the same procedure as earlier in this section. Average IRTAM deviation from the observations at each site, $\bar{\varepsilon}_{IRTAM}\big|_{site}$, were calculated. The corresponding $\bar{\varepsilon}_{IRI}\big|_{site}$ for each site was taken as a upper limit, so it is expected that $\bar{\varepsilon}_{IRTAM}\big|_{site}$ will never be greater than $\bar{\varepsilon}_{IRI}\big|_{site}$. As a lower limit, I provided the $\bar{\varepsilon}_{IRTAM}\big|_{site}$ if data from all eight sites are assimilated. The improvement factors were calculated for each site, and the results of these five tests are listed in Table 3.

Table 3. Results of five spatial limits tests at North American sector. Average (over particular site) $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ are shown. Data were averaged over time period highlighted on Figure 18. Two left columns are improvement factors.

| Control site | $\bar{\varepsilon}_{IRTAM}^{control}$, MHz | $\bar{\varepsilon}_{IRTAM}^{assm}$, MHz | $\bar{\varepsilon}_{IRI}$, MHz | Improv., if control sites | Improv., if assimilated |
|---|---|---|---|---|---|
| Internal test | | | | | |
| BC840 | 0.44 | 0.35 | 0.77 | 1.75 | 2.20 |
| WP937 | 0.54 | 0.37 | 0.79 | 1.46 | 2.14 |
| AU930 | 0.64 | 0.41 | 0.86 | 1.34 | 2.10 |
| EG931 | 0.64 | 0.41 | 0.82 | 1.28 | 2.00 |
| External test | | | | | |
| IF843 | 0.51 | 0.37 | 0.71 | 1.39 | 1.92 |
| MHJ45 | 0.63 | 0.37 | 0.78 | 1.24 | 2.11 |
| PA836 | 0.63 | 0.31 | 0.81 | 1.29 | 2.61 |
| PRJ18 | 1.17 | 0.43 | 1.23 | 1.05 | 2.86 |
| Advanced external test | | | | | |
| BC840 | 0.62 | 0.35 | 0.77 | 1.24 | 2.20 |
| WP937 | 0.72 | 0.37 | 0.79 | 1.10 | 2.14 |
| IF843 | 0.60 | 0.37 | 0.71 | 1.18 | 1.92 |
| MHJ45 | 0.75 | 0.37 | 0.78 | 1.04 | 2.11 |
| PA836 | 0.69 | 0.31 | 0.81 | 1.17 | 2.61 |
| PRJ18 | 1.17 | 0.43 | 1.23 | 1.05 | 2.86 |
| Higher latitude test | | | | | |
| BC840 | 0.49 | 0.35 | 0.77 | 1.57 | 2.20 |
| WP937 | 0.60 | 0.37 | 0.79 | 1.32 | 2.14 |
| IF843 | 0.51 | 0.37 | 0.71 | 1.39 | 1.92 |
| MHJ45 | 0.63 | 0.37 | 0.78 | 1.24 | 2.11 |
| Lower latitude test | | | | | |
| PA836 | 0.67 | 0.31 | 0.81 | 1.21 | 2.61 |
| AU930 | 0.73 | 0.41 | 0.86 | 1.17 | 2.10 |
| EG931 | 0.65 | 0.41 | 0.82 | 1.26 | 2.00 |
| PRJ18 | 1.19 | 0.43 | 1.23 | 1.03 | 2.86 |

As it is expected, the $\bar{\varepsilon}_{IRTAM}^{control}$ stays lower than $\bar{\varepsilon}_{IRTAM}^{assm}$ for each individual site. For the internal test improvement factors for the sites BC840 and WP937 are larger than for other two sites AU930 and EG931. This resulted from the fact that BC840 and WP937

are geographically closer to nearest non-control site. For the external test, the accuracy appears to be worse than for the internal test. This asymmetry in accuracy is significant especially for two pairs of sites: BC840-IF843 and MHJ45-WP937. The improvement factor for BC840, when IF843 was excluded, was 1.75 and the improvement factor for the IF843, when BC840 is excluded, is 1.39. A similar relationship occurs for sites MHJ45 and WP937. One possible explanation of the asymmetry is the positive influence of the adjoining sites, which are PA836 for BC840 and PRJ18 for WP937. Relative position of the sites also matters. For the external test, the geometry was such that the second close sites, which are AU930 for IF843 and EG931 for MHJ45, are on the one line with the first close sites BC840 and WP937 correspondingly. AU930 is behind BC840 and EG931 is behind WP937, so the effects due to AU930 (for IF843) and EG931 (for MHJ45) stations are suppressed by BC840 and WP937 correspondingly. For the internal test, the geometry is different, and both IF843 and PA836 contribute the value at BC840, which turned out to give better result. This explains the major contribution to the observed asymmetry. Finer effects are given by all other sites.

From the advanced external test we see that even two sites, AU930 and EG931, can give a positive effect on the spatially far sites, although this effect is small. The results of the higher latitude test and lower latitude test show almost no improvement over the climatology. Assimilation of the data from the lower latitudes gives a larger positive effect on the simulation result at higher latitudes than vice versa. The results of the two last tests show that ionospheric characteristics vary more significantly with

latitude than with longitude. In other words, the IRTAM maps can be trusted further in zonal direction (along latitude circle) than in meridional direction.

It is illustrative to compare the results of the modeling with site-to-site observations correlation. Since ionospheric characteristics are local time dependent, I used a local time reference frame. Using geographical longitudes of any two sites, I found local time offset between any these sites. In quiet geomagnetic conditions, it is reasonable to assume that data for the same local time on similar geomagnetic latitude should correlate at a high order. To remove the main trend (climate), we subtract IRI model values from the observations and consider $\varepsilon_{IRI}$ . As it was described above, the IRTAM assimilative technique uses $\varepsilon_{IRI}$ to build new coefficients, so if the $\varepsilon_{IRI}$ correlate for different sites the IRTAM maps should also correlate for the same sites. Effects due to interpolation are discussed later. The correlation factor was calculated using Pearson product-moment correlation coefficient:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad (10)$$

where $x_i$ and $y_i$ represent two datasets, $\bar{x}$ and $\bar{y}$ are the mean values of corresponding datasets. Correlation factor *r* takes values between -1 to 1. If datasets perfectly correlate, *r* is 1. If datasets perfectly anti correlate, *r* is -1. If there is no correlation between

datasets, $r$ is 0. Table 4 shows the correlation between any two sites of eight sites chosen above. Table 4 also shows great circle distances between sites in each pair.

Table 4. Table of the $\varepsilon_{IRI}$ correlation for the sites in North American sector. Each cell is divided on two: the top cell contains the correlation factor; the bottom cell contains distance between sites on the sphere (great circle distance) in kilometers.

| | AU930 | BC840 | EG931 | IF843 | MHJ45 | PA836 | PRJ18 | WP937 |
|---|---|---|---|---|---|---|---|---|
| AU930 | | 0.560 | 0.671 | 0.355 | 0.150 | 0.479 | -0.024 | 0.363 |
| | | 1270 | 1070 | 1990 | 2690 | 2190 | 3350 | 2200 |
| BC840 | 0.560 | | 0.480 | 0.641 | 0.217 | 0.534 | -0.062 | 0.347 |
| | 1270 | | 2000 | 740 | 2820 | 1460 | 4360 | 2575 |
| EG931 | 0.671 | 0.480 | | 0.242 | 0.376 | 0.325 | 0.064 | 0.500 |
| | 1070 | 2000 | | 2730 | 1890 | 3200 | 2370 | 1300 |
| IF843 | 0.355 | 0.641 | 0.242 | | 0.200 | 0.552 | 0.012 | 0.194 |
| | 1990 | 740 | 2730 | | 3305 | 1205 | 5080 | 3170 |
| MHJ45 | 0.150 | 0.217 | 0.376 | 0.200 | | 0.147 | 0.187 | 0.517 |
| | 2690 | 2820 | 1890 | 3305 | | 4280 | 2710 | 620 |
| PA836 | 0.479 | 0.534 | 0.325 | 0.552 | 0.147 | | -0.013 | 0.245 |
| | 2190 | 1460 | 3200 | 1205 | 4280 | | 5540 | 4005 |
| PRJ18 | -0.024 | -0.062 | 0.064 | 0.012 | 0.187 | -0.013 | | 0.072 |
| | 3350 | 4360 | 2370 | 5080 | 2710 | 5540 | | 2305 |
| WP937 | 0.363 | 0.347 | 0.500 | 0.194 | 0.517 | 0.245 | 0.072 | |
| | 2200 | 2575 | 1300 | 3170 | 620 | 4005 | 2305 | |

As it is expected, data from spatially close sites correlate at higher extent, for example, sites AU930 with EG931 and BC840 with IF843. It is interesting that PRJ18 station almost does not correlate with others. This site is located in the close to low latitude region, while the others are pure middle-latitude stations. In general, data in Table 3 are in agreement with data in Table 4. If there is high correlation between data at two sites, and if one of the sites take part in the assimilation and another does not, the IRTAM representation of a characteristics at control site will be better than the

climatology representation. However, this behavior is expected only if two sites isolated from others are considered. In the case of multiple sites mutual geometry should be consider to estimate a characteristic representation at any given location. Moreover, the interpolation method suppresses a site's influence far from the site's locations, which may result in small improvement factor even if correlation is high. We can summarize spatial limits test if we come up with expression to express how $\bar{\varepsilon}_{IRTAM}^{control}$ is related to $\bar{\varepsilon}_{IRTAM}^{assm}$. We construct control site quality factor τ:

$$\tau = \frac{\bar{\varepsilon}_{IRI} - \bar{\varepsilon}_{IRTAM}^{control}}{\bar{\varepsilon}_{IRI} - \bar{\varepsilon}_{IRTAM}^{assm}} \tag{11}$$

This factor is equal 1 for the case when IRTAM representation at control site is as accurate as if the data from this site would be assimilated. Factor τ is equal 0 if IRTAM representation is the same as climatology representation. For each row in Table 4 we plot τ versus average distance to two closest neighbors, and the result is in Figure 20 .

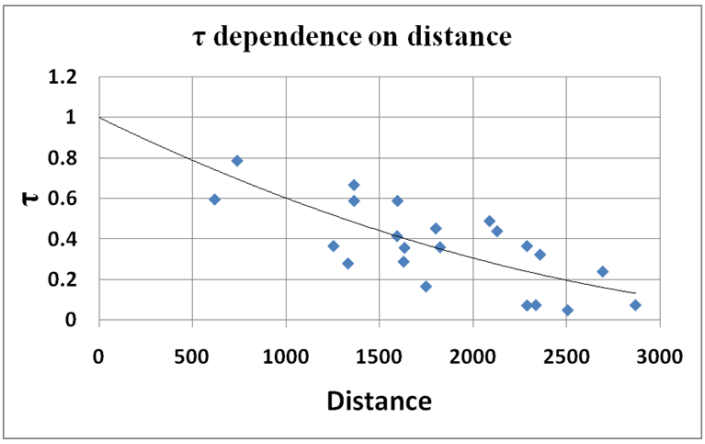



Figure 20. Control site quality factor's dependence on distance.

We see that τ drops to 0.5 at about 1300 km distance. Thus, within 1300 km IRTAM maps accurately represent ionospheric weather. Beyond 1300 km average IRTAM representation accuracy converges to the accuracy of the climatology representation.

Above, I explained how and why we chose eight sites that distributed relatively spatial compact and belong to middle latitude region. Now, I choose sites distributed roughly along meridian in the American sector. Data availability for a site in the sector from 55°S to 45°N by geographical latitude and from 57°W to 77°W by geographical longitude was found same way as before and is shown in Figure 21.

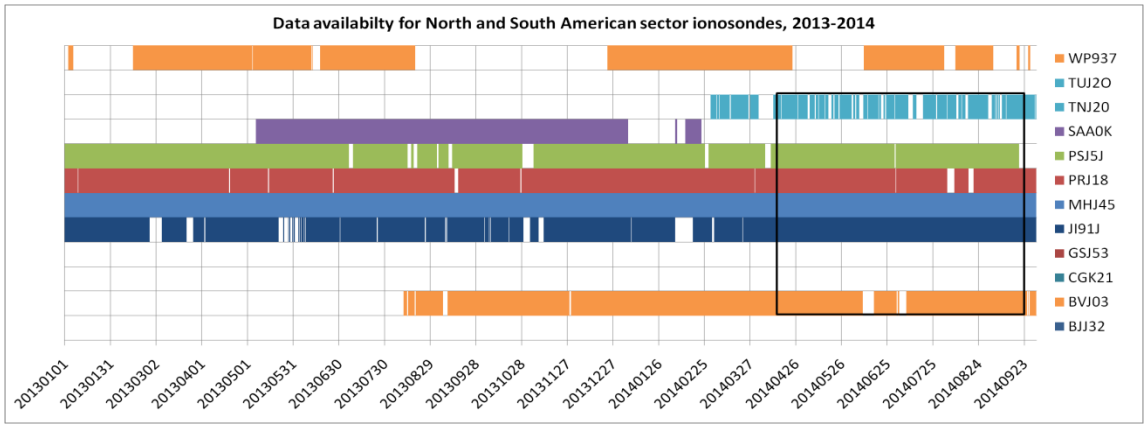



Figure 21. Same as Figure 18, but for sector from 55°S to 45°N and from 57°W to 77°W.

All other neighboring sites were excluded from the assimilation to avoid ambiguity in data interpretation. A black rectangle denotes time interval to perform further analysis. Although sites are distributed not along single meridian, meridian dimension as much greater than zonal dimension. The sites chosen for the analysis are MHJ45 (Millstone Hill 42.6°N, 71.6°W), PRJ18 (Ramey 18.5°N, 67.1°W), BVJ03 (Boa Vista 2.8°N, 60.7°W),

JI91J (Jicamarca 12.0°S, 76.8°W), TNJ20 (Tucuman AIS-INGV 26.9°S, 65.4°W), PSJ5J (Port Stanley 51.6°S, 57.9°W). The JI91J site is located at the magnetic equator and BVJ03 and TNJ20 are located in the equatorial ionization anomaly region. The other two sites, MHJ45 and PSJ5J, belong to middle latitude region, and PRJ18 location is discussed early in this subsection. Figure 22 shows geographical locations of these six sites. The figure is vertically shrunk, so apparent distance from site to site along meridian is smaller than in reality.

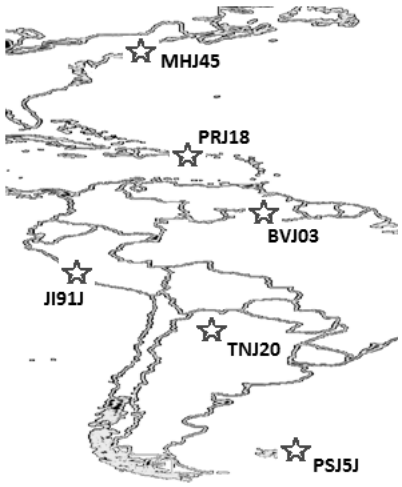



Figure 22. Geographical location of the six sites chosen by the rectangle on Figure 21.

Test configurations for this choice of the sites are:

1. Magnetic equator test: only JI91J site was taken as the control site and the others took part in the assimilation
2. Equatorial ionization anomaly test: BVJ03 and TNJ20 sites were the control sites and the others took part in the assimilation

3. Equatorial ionization anomaly + magnetic equator test: BVJ03, JI91J and TNJ20 sites were the control sites and the others took part in the assimilation
4. Middle latitudes test: MHJ45, PRJ18 and PSJ5J sites were the control sites and the others took part in the assimilation

The procedure is the same as described above, and the results of these four tests are listed in Table 5.

Table 5. Same as Table 3, but for four spatial limits tests at American sector.

| Site | $\bar{\varepsilon}_{IRTAM}\big\vert_{site}$, MHz if control site | $\bar{\varepsilon}_{IRTAM}\big\vert_{site}$, MHz if assimilated, | $\bar{\varepsilon}_{IRI}\big\vert_{site}$, MHz | Improv., if control sites | Improv., if assimilated |
|---|---|---|---|---|---|
| Magnetic equator test | | | | | |
| JI91J | 0.92 | 0.41 | 0.78 | 0.85 | 1.90 |
| Equatorial ionization anomaly test | | | | | |
| BVJ03 | 1.35 | 0.52 | 1.35 | 1.00 | 2.60 |
| TNJ20 | 1.62 | 0.56 | 1.66 | 1.02 | 2.96 |
| Equatorial ionization anomaly + magnetic equator test | | | | | |
| JI91J | 0.73 | 0.41 | 0.78 | 1.07 | 1.90 |
| BVJ03 | 1.35 | 0.52 | 1.35 | 1.00 | 2.60 |
| TNJ20 | 1.60 | 0.56 | 1.66 | 1.04 | 2.96 |
| Middle latitudes test | | | | | |
| MHJ45 | 0.91 | 0.50 | 0.96 | 1.05 | 1.92 |
| PRJ18 | 1.00 | 0.62 | 1.00 | 1.00 | 1.62 |
| PSJ5J | 1.23 | 0.50 | 1.25 | 1.02 | 2.50 |

First, I would like point out the result for JI91J site for magnetic equator test. As it clearly seen from Table 5, the IRTAM accuracy is even worse than IRI accuracy. This test was designed such that the observations from BVJ03 and TNJ20 sites, which are located in equatorial ionization anomaly region, were assimilated, and JI91J observations were not. The assimilative technique extended characteristic behavior from the equatorial

ionization anomaly to the magnetic equator. I concluded that there is a negative correlation between *foF2* behavior at the magnetic equator and equatorial ionization anomaly, which leaded to the improvement factor of less than one. It was expected that the improvement factor would also be less than one for BVJ03 and TNJ20 when the data from JI91J were assimilated (equatorial ionization anomaly test). The IRTAM accuracy in this test for BVJ03 and TNJ20 is close to the IRI accuracy, but is not worse than IRI accuracy. Observed asymmetry in improvement factors can be again explained by different geometries of the tests and was discussed in detail above. Figure 22 shows the other sites those contribute to the assimilative maps: MHJ45, PRJ18 and PSJ5J. The possible reason why there is no negative effect observed in *foF2* representation for BVJ03 and TNJ20, is in positive correlation of the data from BVJ03 and TNJ20 with data from MHJ45, PRJ18 and PSJ5J. As expected, if all three equatorial stations are excluded from assimilation, the IRTAM representation is similar to the IRI representation, which is shown by the equatorial ionization anomaly + magnetic equator test. The same result is observed for the middle latitude sites (see middle latitudes test). Thus, it is reasonable to state that, on average, there is no correlation between characteristics behavior at low and middle latitudes.

The product-moment correlation coefficient between any two sites in this case was calculated same way as for Table 4. The results are listed in Table 6.

Table 6. Same as Table 4, but for sites MHJ45, PRJ18, BVJ03, JI91J, TNJ20 and PSJ5J.

| | MHJ45 | PRJ18 | BVJ03 | JI91J | TNJ20 | PSJ5J |
|---|---|---|---|---|---|---|
| MHJ45 | | 0.045 | -0.033 | -0.048 | 0.073 | -0.069 |
| | | 2710 | 4550 | 6095 | 7750 | 10555 |
| PRJ18 | 0.045 | | -0.013 | -0.003 | -0.028 | -0.068 |
| | 2710 | | 1880 | 3555 | 5050 | 7845 |
| BVJ03 | -0.033 | -0.013 | | -0.035 | 0.218 | -0.081 |
| | 4550 | 1880 | | 2425 | 3340 | 6055 |
| JI91J | -0.048 | -0.003 | -0.035 | | -0.197 | 0.046 |
| | 6095 | 3555 | 2425 | | 2040 | 4720 |
| TNJ20 | 0.073 | -0.028 | 0.218 | -0.197 | | 0.046 |
| | 7750 | 5050 | 3340 | 2040 | | 2815 |
| PSJ5J | -0.069 | -0.068 | -0.081 | 0.046 | 0.046 | |
| | 10555 | 7845 | 6055 | 4720 | 2815 | |

As expected, there is negative correlation between data from TNJ20 and JI91J, also there is small positive correlation between BVJ03 and TNJ20. This supports the results shown in Table 4. On average, distances between sites in the meridional test are greater than distances between sites in North American sector. Thus, it is hard to draw general conclusion about the range at which IRTAM maps can still be trusted. Other meridional chains should be considered in order to conduct robust study on meridional variability.

These results reveal the necessity to redesign the interpolation technique used in the IRTAM. Currently, the interpolation technique treats all direction equally, so data extended equally far from sites locations in both latitude and longitude directions. It is reasonable to introduce a different set of rules for data interpolation/extrapolation in zonal and meridional directions.

## 3.6 Equatorial Post-Sunset Uplift of the Ionosphere

I have considered the particular cases when the IRTAM helps to improve climatology representation. It is well-known problem that the IRI underestimates the evening uplift of the ionosphere at magnetic equator in post-sunset time [*Adeniyi et al.*, 2003]. To understand this phenomenon, consider solar ionizing radiation and plasma recombination to neutrals. These two governing processes drive the ionospheric plasma production rate. When sunset solar terminator passes, there is no more ionizing radiation from the Sun, and the ionization is cut off. The recombination rate depends on density and stays the same. Sunset occurs at smaller altitudes first and only then at larger altitudes. The density drops with altitude, so the recombination rate is larger at smaller altitudes. Thus, around sunset, plasma production rate is higher at larger altitudes than at smaller altitudes. Plasma motion at magnetic equator that gives rise to the equatorial ionization anomaly also contributes significantly to the F layer peak uplift. As a reminder, plasma moves upward at magnetic equator and then sinks along magnetic field lines to the low latitudes. All this causes ionospheric peak density moves to higher altitudes, which is equivalent to the increase of *hmF2*. To avoid confusion, *foF2* is continuously decreasing, but moves to larger altitudes. This phenomenon is called post sunset uplift of the ionosphere at magnetic equator. The time scale of this process is about one hour, and the formalism includes only as high as $6^{th}$ order time harmonic, which allows to describe as fast as 4 hour period variations. *Adeniyi et al.* [2003] assumed that

higher order harmonics are needed to be able to track post-sunset uplift. Figure 23 shows post-sunset uplift as it is drawn by observations as well as its representations by IRTAM and IRI.

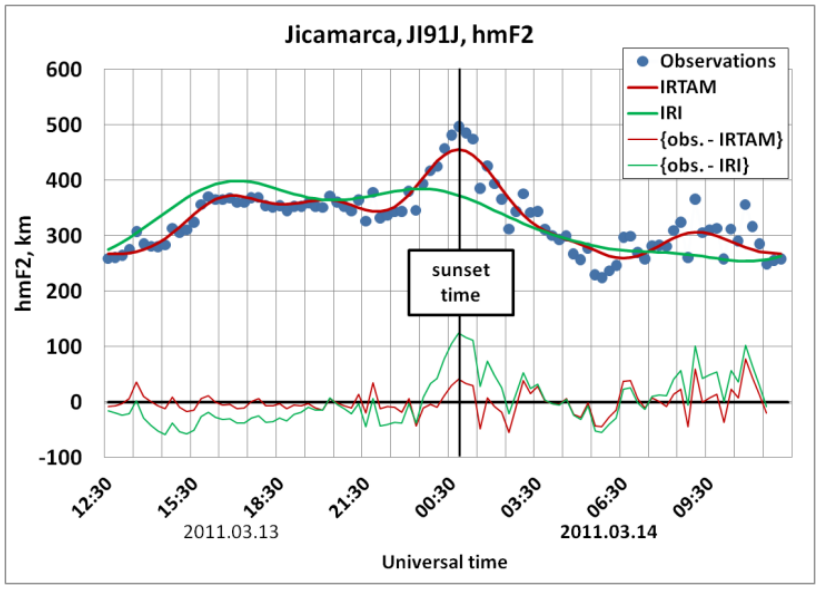



Figure 23. Post-sunset uplift of the equatorial ionosphere. Data for site Jicamarca, JI91J, (12.0°S, 76.8°W) are shown. Dots represent observations, red and green bold curves are IRTAM and IRI representations respectively. Red and green thin curves are differences between observations and model's representations for IRTAM and IRI respectively. Data are plotted versus UT. The vertical line show sunset at 300 km altitude.

IRTAM is able to capture the post-sunset uplift, although not perfectly. To make sure that the IRTAM routinely represent the peak, I collected all data available for 2003-2014 years at equatorial site Jicamarca, which has magnetic latitude of 1.7°S. For each particular day the sunset time at 300 km altitude was calculated. Each $\varepsilon_{IRTAM}$ or $\varepsilon_{IRI}$ was attributed to the particular local time relative to the sunset time. $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ data

were collected and counted separately whether they were positive or negative. Then the data were averaged. Thus, we obtain four values, $\bar{\varepsilon}_{IRTAM}^{pos}$, $\bar{\varepsilon}_{IRTAM}^{neg}$, $\bar{\varepsilon}_{IRI}^{pos}$ and $\bar{\varepsilon}_{IRI}^{neg}$, for each local time. To understand what we expect to see, consider the construction of $\varepsilon$, which is (observation - model). When a model underestimates observation, the difference is positive, and when a model overestimates observation, the difference is negative. Thus, it is expected to see positive maximum in the $\bar{\varepsilon}_{IRI}$ just after sunset, while no such maximum should be drawn by $\bar{\varepsilon}_{IRTAM}$. The result of this analysis is shown by Figure 24.

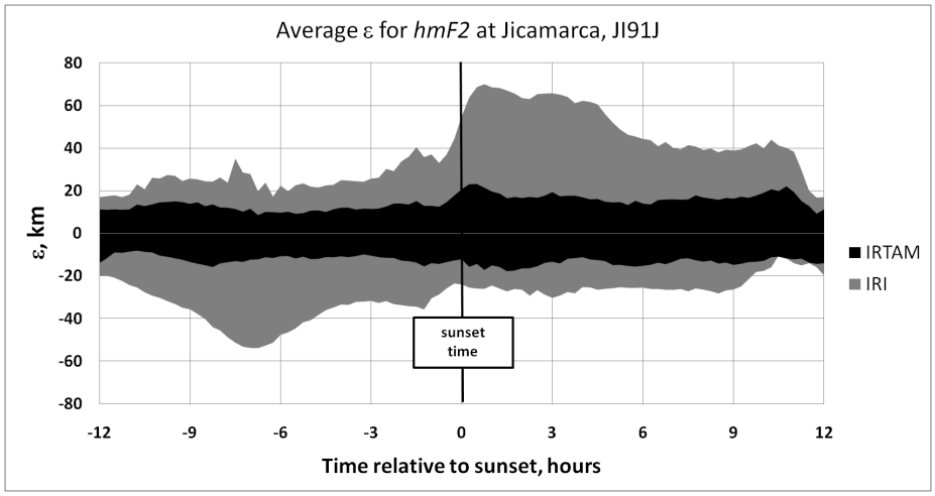



Figure 24. Average $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ dependence on local time relative to sunset time. The time 1:00 means that data correspond to 1 hour after sunset, and -2:00 means that data correspond to 2 hours before sunset. Zero corresponds to the time of local sunset. Only the JI91J site's measurements are used. $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ are averaged over 2003-2014 time interval. $\varepsilon_{IRTAM}$ and $\varepsilon_{IRI}$ are averaged separately for individual local time.

Indeed, there is positive maximum shown by $\bar{\varepsilon}_{IRI}$ data. Moreover, although it is not show in Figure 24, the count of data points for which $\varepsilon_{IRI}$ are positive is twice as

high as count of data points for which those are negative. This means that the IRI tends to underestimate observations more than two times frequently than overestimate observations. There is no strong positive peak in $\bar{\varepsilon}_{IRTAM}$ data at near sunset time. Thus, IRTAM successfully reproduces post-sunset ionospheric uplift on the equator. The only concern is that the ratio between occurrences of positive and negative $\varepsilon_{IRTAM}$ stays the same as for IRI, and this ratio is about two. At this point, the assumption of *Adeniyi et al.* [2003] about the inability of the IRI formalism to reproduce short time scale uplift is only partially true. As we see in Figure 23, the uplift can be resolved, but the amplitude is smaller.

## 3.7 Case Study 1: Interplanetary Shock 7 November 2004

Previously, only data averaged over long time period and/or over different geographical locations were presented. Although, this analysis is important, it is more interesting to see IRTAM in action during specific events. The ionospheric state during such events is far from its quiet time state, and this is a case when IRTAM can demonstrate way more reliable results than climatology IRI.

The first event I have considered is interplanetary shock that hits the Earth's magnetosphere at ~1830 UT on 7 November 2004. Description of the interplanetary shock is given in the previous section. Figure 25 shows solar wind parameters, including IMF, for time interval from 1600 UT on 7 November to 0000 UT on 8 November 2004.

These data are derived from National Aeronautics and Space Administration (NASA) Operating Missions as a Node on the Internet webpage (OMNIWeb). The shock appeared during extreme storm that was lasting from 7 to 11 November 2004, which will be discussed in the next subsection.

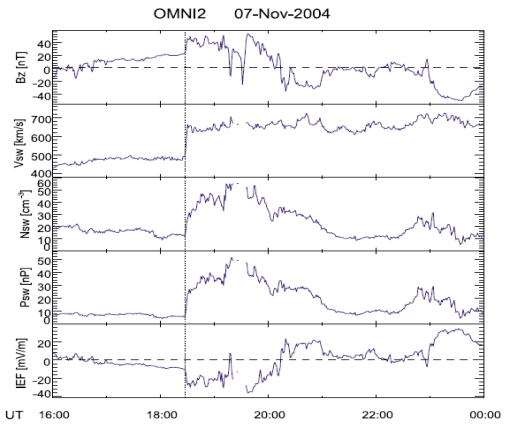



Figure 25. IMF and solar wind parameter taken from NASA OMNIWeb. From top are the interplanetary magnetic field *Bz*, solar wind flow speed, proton density, solar wind ion dynamic pressure and the interplanetary electric field. The front of the interplanetary shock is marked by the vertical dashed line. Adopted from *Zong et al.* [2010].

The shock was associated with northward vertical component of the IMF, north *Bz*, which persisted for long time, 107 min. Thus ionospheric effects could be attributed purely to the north *Bz* and not to the change in the IMF polarization [*Zong et al.*, 2010]. The comprehensive study of the ionospheric characteristics variations was conducted by *Zong et al.* [2010]. Longitudinal/latitudinal effects were studied using the observations in regions those are narrow in longitude/latitude and broad in latitude/longitude. To investigate longitudinal variations observation from North American and European

stations were used. To study latitudinal effects, observations from North American and South American stations were used. The two regions intersect at North American sector. The American sector was on the day side, and European sector was on the night side close to the sunset terminator, when the shock arrived. The major effects reported by *Zong et al.* [2010] are:

1. Almost instantaneous response on the day side close to local noon (subsolar point); observations: Boulder (40.0°N, 105.3°W), Millstone Hill (42.6°N, 71.5°W).
2. No significant immediate response on night side; observations: San Vito (40.6°N, 17.8°E), Athens (38.0°N, 23.6°E).
3. Compression of F region at magnetic equator when *Bz* was northward followed by uplift when *Bz* changed to southward.

The IRTAM provides global time dependent picture of the ionospheric characteristics variations, so there is no need in spatial region limitation. The IRTAM shows all major effects reported by *Zong et al.* [2010], although there are some discrepancies. For example, the IRTAM shows plasma density enhancements at European sector just after shock arrival which opposite to what is stated by *Zong et al.* [2010]. This discrepancy results from the fact that I used different baseline compared to *Zong et al.* [2010]. *Zong et al.* [2010] used observations before the shock, and I used climatology IRI representation. This is the main reason for observed discrepancy, detailed discussion of which is provided further is this subsection. Figure 26 shows global geographical variations followed by shock arrival at 1830 UT.

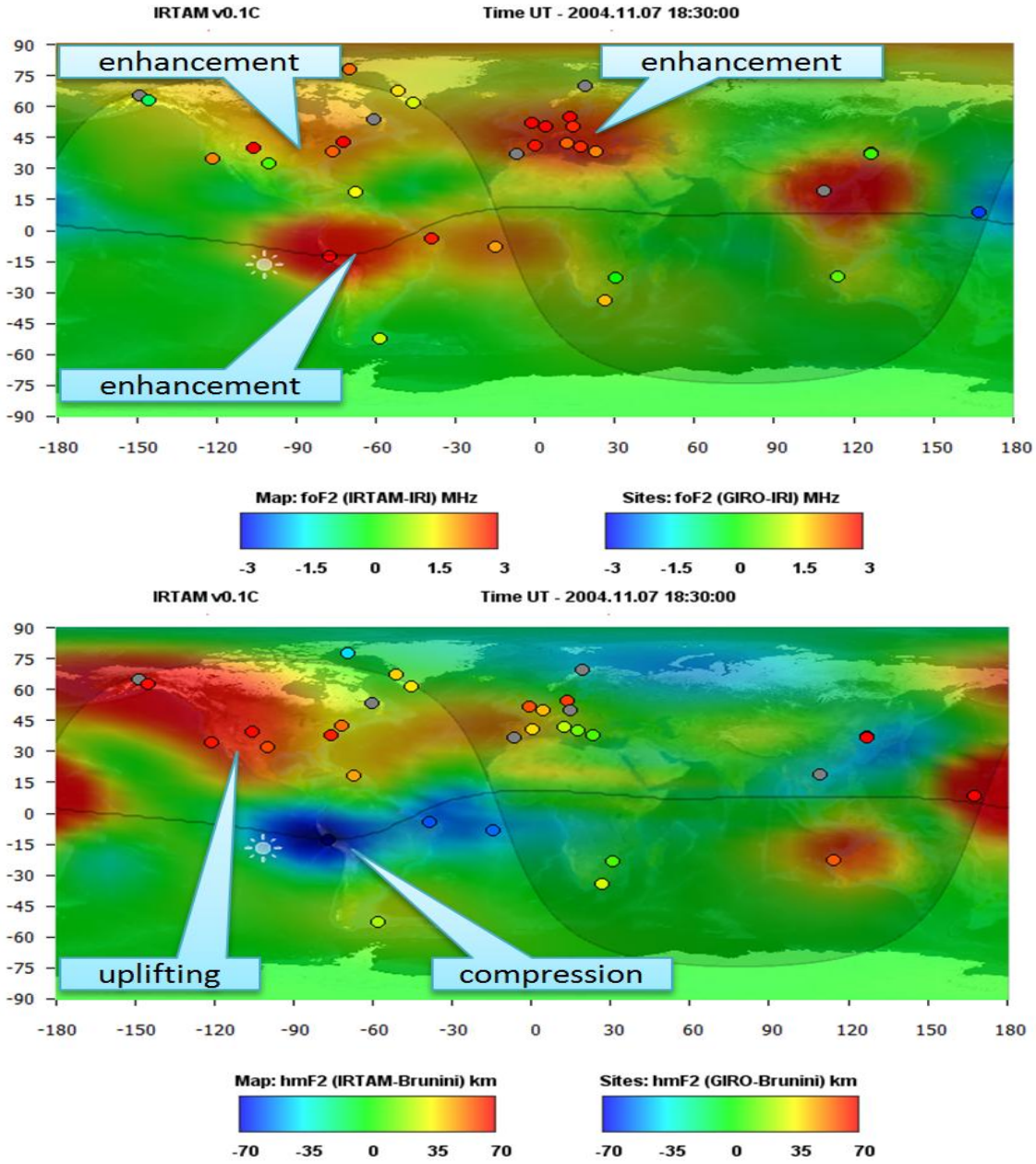



Figure 26. Ionospheric effects of the interplanetary shock passage at 1830 UT on 7 November 2004. *ΔfoF2* (top) and *ΔhmF2* (bottom) maps are shown. Maps are calculated for time just after shock hitted the Earth's magetosphere. Dots show $\varepsilon_{IRI}$ at sites locations. Shadow depicts night side, American sector is on the day side.

The red color on Figure 26 corresponds to enhancement of plasma density or uplift of the peak height, the blue color indicates to depletion of plasma density or compression of peak height. The green color means that there is no difference between climate and weather. Using the IRTAM maps, I found, that same signatures shown by Figure 26 appear not only after the shock arrival but actually start to appear several hours before the shock hit the Earth's magnetosphere. Figure 27 illustrates *ΔfoF2* map 2 hours before the shock arrival on 7 November and *ΔfoF2* map for same time, but on previous quiet day 6 November 2004.

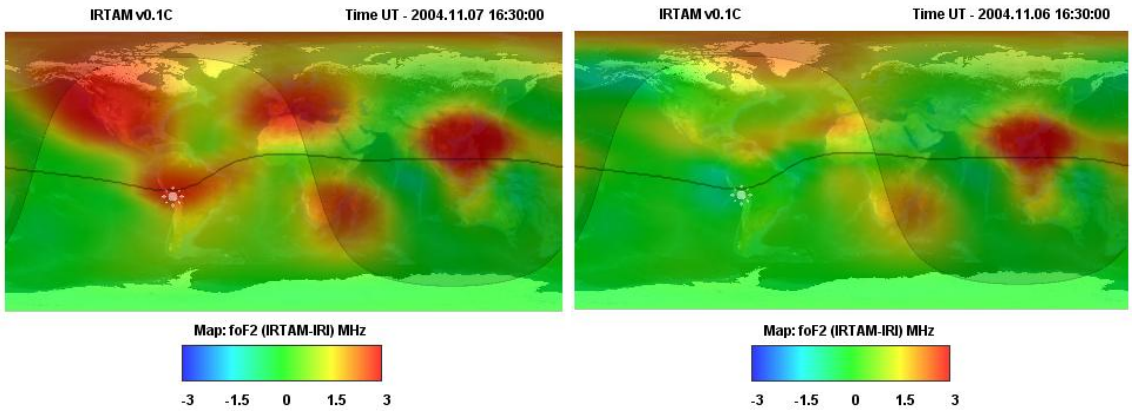



Figure 27. *ΔfoF2* maps for 1630 UT on 7 November 2004 (left) and 1630 UT on 6 November 2004 (right). 6 November 2004 is considered as quiet day and is provided for comparison. The red spot over the Asian sector, repeats from day to day. It is considered as constant deviation from the climate and should be disregarded on all maps.

By comparison of top panel of Figure 26 and left panel of Figure 27, it is clear why the IRTAM shows the enhancement of *foF2* that was not reported by *Zong et al.* [2010]. The enhancement appears before shock arrival and is seen if climatology representation is taken as the background. *Zong et al.* [2010] considered observations before shock arrival

as a baseline. These observations already contained disturbed time signatures. Thus, we were considering different things. I considered cumulative effect of the geomagnetic storm on the ionosphere, and *Zong et al.* [2010] considered ionospheric effects caused by shock passage only. To obtain the picture which is consistent with the results of *Zong et al.* [2010], I should take IRTAM *foF2* maps before the shock arrival and subtract them from the maps from the IRTAM *foF2* maps just after the shock arrival. The results of this comparison are not reported in this thesis in order to not confuse the reader by introducing new kind of maps. Here and in next two sections, *ΔfoF2* were calculated using only the IRI as baseline.

## 3.8 Case Study 2: Plasmasphere Refilling Effects on the Ionosphere

As it was mentioned above the shock that hits the Earth's magnetosphere at 1830 on 7 November 2004 was just one event in the sequence of the events during extreme geomagnetic storm lasted from 7 to 11 November 2004. The effects of the storm were studied by different researches, using numerous kinds of measurements [*Maruyama,* 2006; *Sahai et al.*, 2009; *Heelis and Coley,* 2007]. These measurements include, but not limited by in situ plasma density measurements, radio occultation measurements of the ionospheric TEC, GPS TEC measurements with ground network of receivers, ionospheric vertical drifts measurements. Where it was possible, I compared the IRTAM maps with the data from aforementioned papers. The IRTAM shows qualitative and quantitative

agreement with these data. I do not provide a detailed discussion of this comparison and rather concentrate on the phenomenon, which for my knowledge was not reported previously. This phenomenon appears in recovery phase of the storm, when DST index passes minimum and starts increasing. I plotted solar wind parameters in the form similar to one shown on Figure 25, but for the extended time interval from 0000 UT on 7 November 2004 to 2359 UT on 12 November 2004. The solar wind data are presented on Figure 28. Solar wind dynamic pressure and electric field are not shown since they can be derived from IMF, solar wind velocity and density. DST index is plotted on the IMF panel.

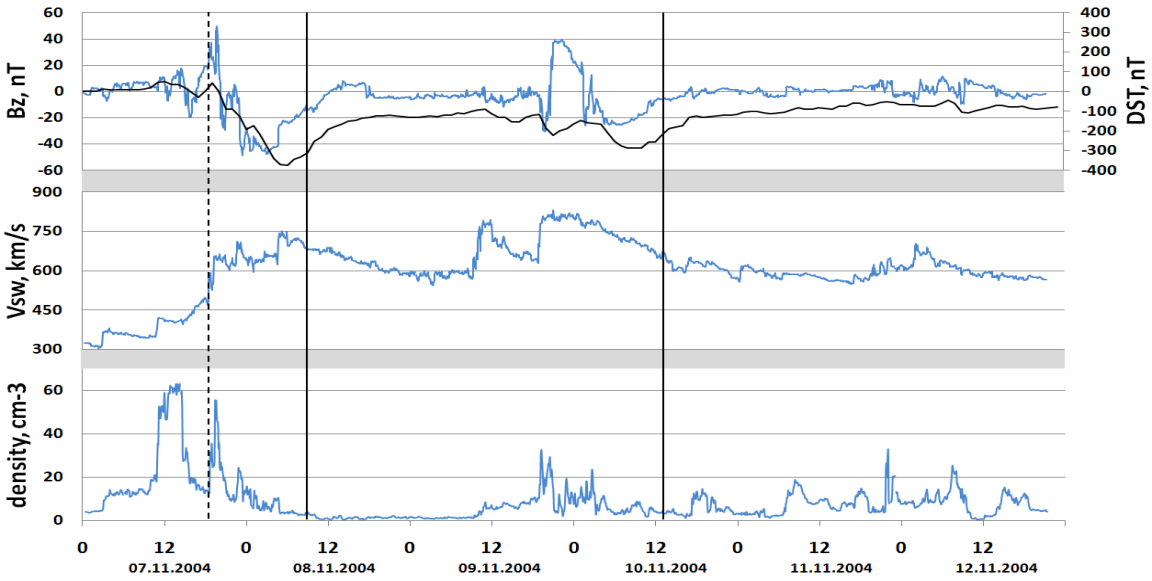



Figure 28. Data derived from NASA OMNIWeb for time interval from 0000 UT on 7 November 2004 to 2359 UT on 12 November 2004. The top panel shows *Bz* component of IMF together with DST index, middle and bottom panels show solar wind velocity and density correspondingly. Dashed vertical line shows the shock discussed above.

Figure 28 shows that disturbed time lasted for several days and two strong storms occurred. The phase of a storm is connected with DST index. Storm development corresponds to decreasing DST index, and recovery phase corresponds to increasing DST index. Two deep minimums in DST timeline occurred at 0600 UT on 8 November 2004 and 0900 UT on 10 November 2004. The phenomenon under investigation appears as strong plasma density depletion off magnetic equator up to middle latitudes and plasma density enhancement on the magnetic equator. The depletion appears to be stronger and larger scale than the enhancement, and further discussion is devoted primarily the depletion. This picture persists for more than 12 hours and is mapped on Figure 29. The time for which this map was calculated is marked as left solid line on Figure 28.

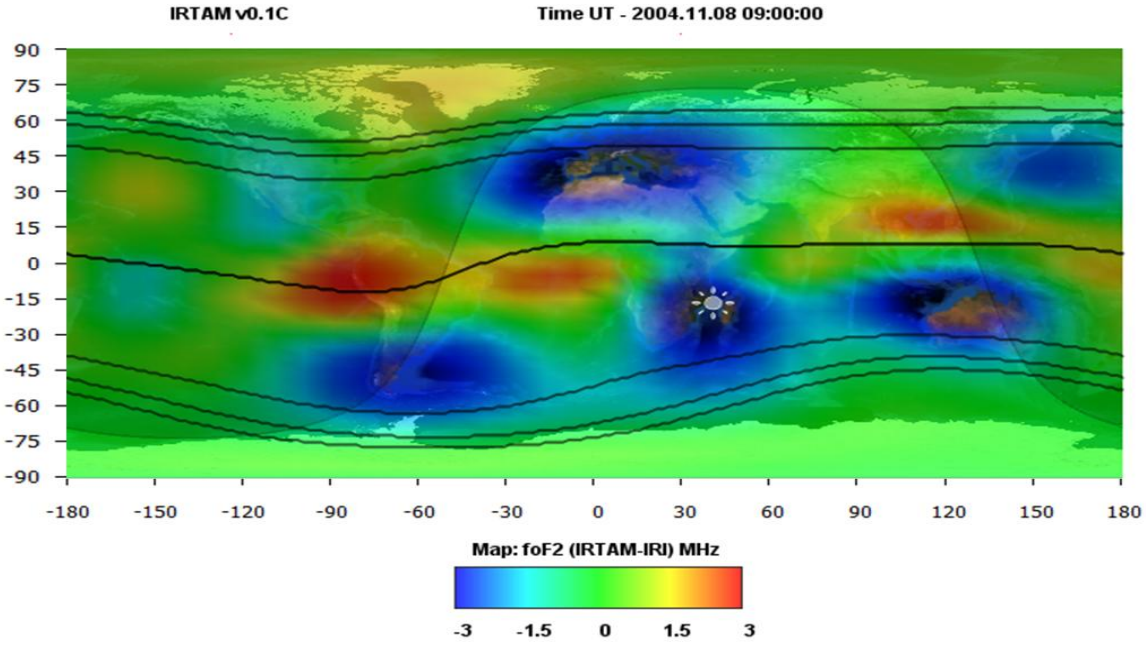



Figure 29. *ΔfoF2* map calculated for 0900 UT on 8 November 2004. Black lines represent magnetic equator (center line) and mapping of 2, 3, 4 L-shells on the Earth's surface.

The black lines are mappings of the L-shells on the Earth's surface. If the footprints of all field lines corresponding to a particular L-shell are plotted on the Earth's surface, we get two lines. One of these two lines is in Southern hemisphere and another is in Northern hemisphere. L-shells are shown in Figure 29 in order to depict which latitudes of the ionosphere plasmasphere is mapped to. As mentioned above, one possible mechanism responsible for features observed in Figure 29 can be attributed to the refilling of the plasmasphere. The plasmasphere is inner most part of the Earth's magnetosphere and extends up to 4 Earth's radii (L = 4). The plasma in the plasmasphere is of ionospheric origin. To make this conclusion about plasma density in plasmasphere, I have analyzed RPI plasmasphere plasma density profiles from 0000 UT on 1 November 2004 to 0000 UT on 9 November 2004. The time interval between 0000 UT on 1 November 2004 and 0000 UT on 7 November 2004 was considered as a quiet time, and 0000 UT on 7 November 2004 – 0000 UT on 9 November 2004 was considered as disturbed time. The goal was to find plasma density profiles measured for a similar L-shell and MLT, which belong to quiet and storm times. Relative position of the RPI orbit and the plasmasphere are not always the same, which make it difficult to find measurements made under the same conditions. Moreover, the RPI measurements do not always show detectible echo traces, and profile inversion is impossible in these cases. This long time interval is required in order to find comparable measurements. Figure 30 shows plasma density profiles measured roughly at 1030 MLT and 2230 MLT during quiet time and storm

time. 1030 MLT corresponds to the day side plasmasphere and 2230 MLT corresponds to the night side.

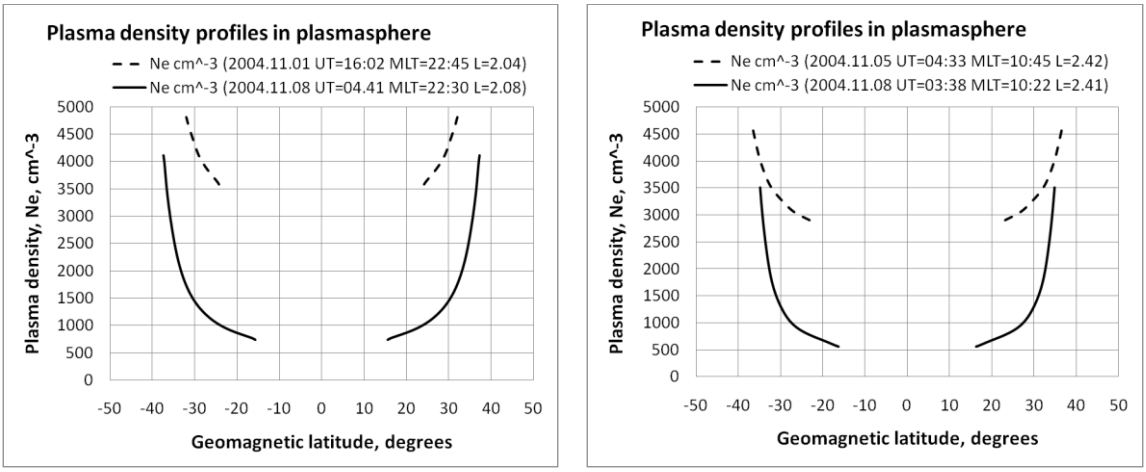



Figure 30. Plasma density profile along magnetic field lines derived from RPI plasmagrams. Plasma density is plotted versus magnetic latitude of the particular position at the magnetic field line. Dashed lines are profiles derived at quiet times, which serve as a baseline. Solid lines are profiles derived at storm time when plasma density depletion in the ionosphere was observed. Left panel shows profile measured for ~ 2230 MLT and ~2.06 L, and the right panel shows the profile measured at ~ 1030 MLT and ~2.42 L.

As seen from the plasma density profiles, the plasmasphere during the storm time is filled to less than 30% of its quiet time state on ±30 degrees of geomagnetic latitude. The plasma depletion in the plasmasphere is even stronger for smaller geomagnetic latitudes. The left panel shows that plasma density in the plasmasphere is depleted at least for the L-shell of 2. I could not find reliable measurements for the L-shell smaller than 2, which makes it difficult to draw conclusion of plasmapause position during the storm. Nevertheless, even depletion of the plasma in the plasmasphere on L-shell is enough to

describe *foF2* behavior shown in Figure 29. Largest plasma depletion in the ionosphere occurs on the day side. This could be explained by the ionosphere-plasmasphere plasma flow on the day and night sides. On the day side, even at quiet time, plasma flows from the ionosphere to the plasmasphere, on the night side vice versa. Thus, during the recovery phase of the storm, relatively empty flux tubes of the plasmasphere causes enhanced plasma outflow from the ionosphere on the day side. The night side is a little bit more complicated, since during quiet time, plasma flows from the plasmasphere to the ionosphere. During the recovery phase of the storm empty flux tubes do not contain a significant amount of plasma to flow into the ionosphere. The plasma flow direction depends on the ratio of the densities in plasmasphere and ionosphere. Quantitative calculations should be made to find the plasma flow direction, which are not provided here. Anyway, plasma depletion in the ionosphere is observed on the night side, however it is much weaker than on the day side. The plasma density on the day side is several times higher than on the night side, so the absolute *ΔfoF2* is higher on the day side. The relative depletion up to 50% of plasma density quite time value is almost the same on the day and night sides.

As it is seen in Figure 28, the second strong geomagnetic storm occurred a day later on 9-10 November 2004. The second storm was weaker than the first, which is depicted by the DST index. The right solid line in Figure 28 marks another time that was analyzed for plasma depletion. The conditions for the two analyzed time periods were similar, although were not exactly the same. The 7-8 November storm was followed by

about 24 hours period of almost zero solar wind impact. Solar wind velocity was continuously decreasing; *Bz* and solar wind density was decreasing until both of them reached approximately zero. These conditions make this storm favorable to study. Studied phenomenon could be linked to the recovery of the plasmasphere after the storm, and not linked to the extra impact from the solar wind. The 9-10 November storm was also followed by a period when variations in solar wind parameters were weak, but compared to the first storm, the density was higher. Figure 31 shows the *foF2* map observed during recovery phase of the second storm.

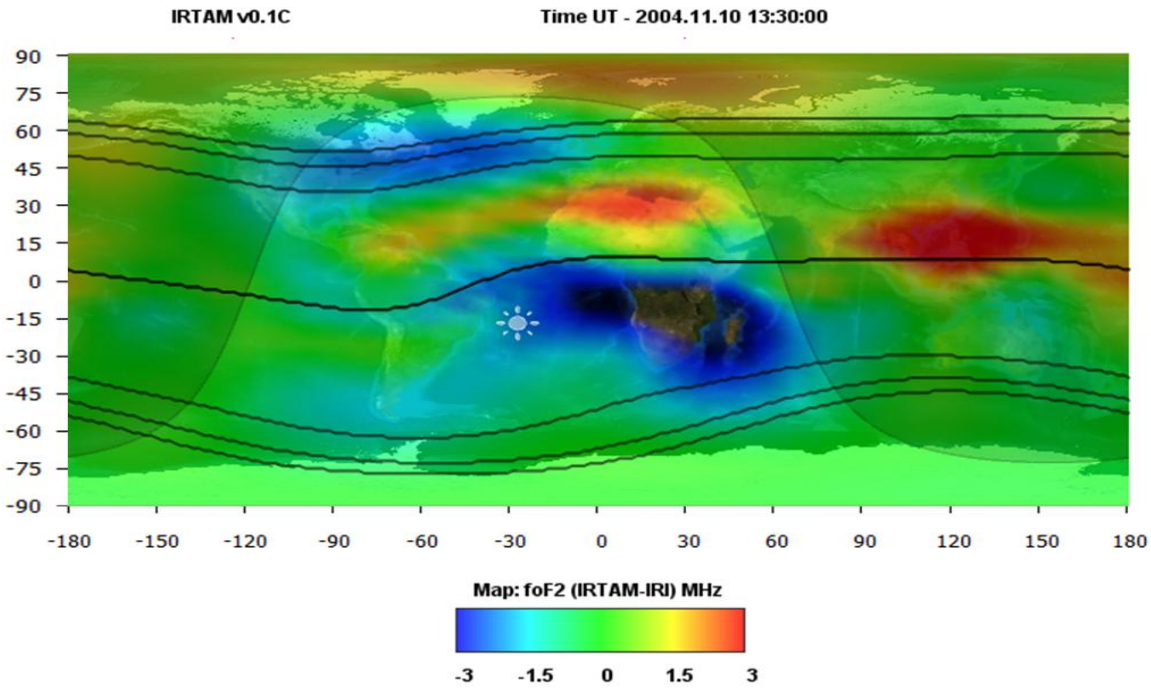



Figure 31. Same as Figure 29, but for 1330 UT on 10 November 2004.

The plasma deletion in Figure 31 is less pronounced than in the Figure 29 since the 9-10 November storm was weaker than the 7-8 November storm. Moreover, the picture

observed in Figure 31 seems to be shifted to the North compared to what was observed in Figure 29. Currently, I do not have a valid explanation for this asymmetry, and this will be studied separately. Also the RPI data could help to explain this asymmetry. Currently, I did not analyze these data and it is left for future research. The main features seem to be preserved, i.e. depletion appears mostly off magnetic equator and is stronger on the day side.

## 3.9 Case Study 3: Sudden Stratospheric Warming 2013

In this subsection we consider ionospheric variations driven from the neutral atmosphere. This is opposite to what has been considered in the previous subsection, where the solar wind was acting as a driver for the variations in the ionosphere. The event which I am discussed in this subsection is the January 2013 SSW [*Goncharenko et al.*, 2013]. To see low- and mid-latitude ionospheric effects due to this event, it is the most illustrative to consider GPS TEC measurements. Due to a great number of GPS receivers, TEC maps have good coverage, and can resolve small scale spatial variations of TEC. The TEC distributions for 1500 UT on 15 December 2012 are taken as a baseline and 1500 UT on 15 January 2013, the SSW peak, are shown in Figure 32.

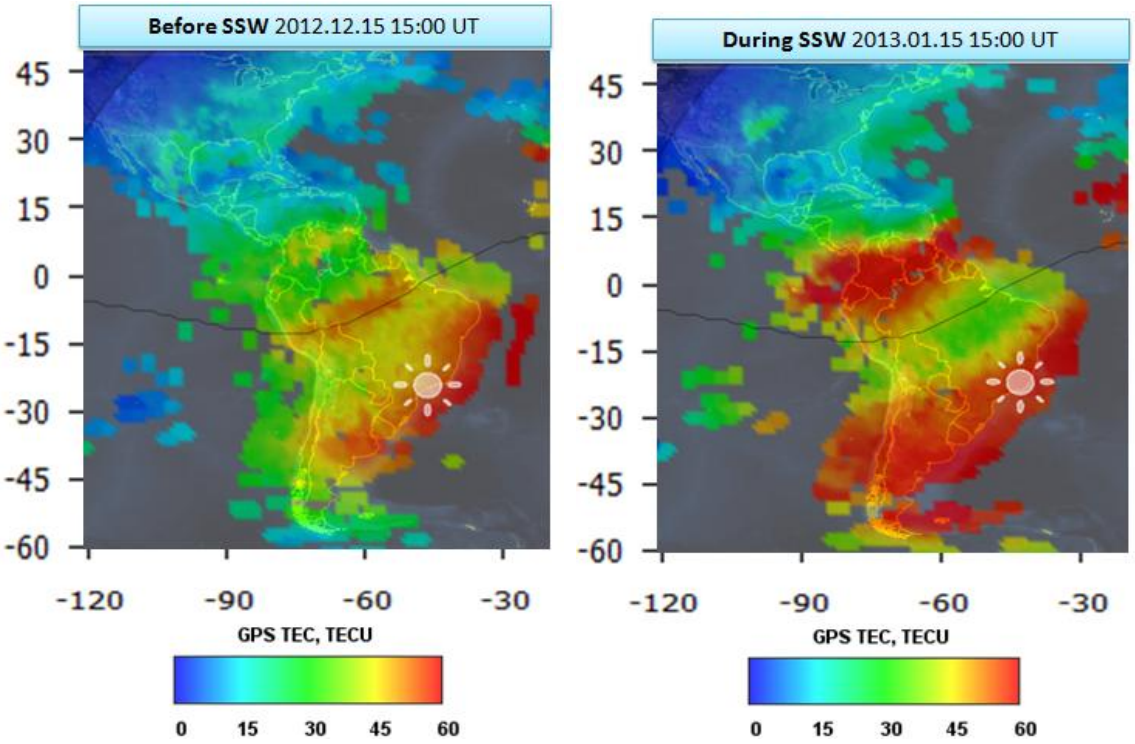



Figure 32. GPS TEC maps before (left panel) and during (right panel) the 2013 SSW. The magnetic equator is shown by the black line.

The left panel shows the usual distribution of TEC during winter time. TEC is larger at the magnetic equator, and smaller at the equatorial ionization anomaly. Since this map corresponds to December, the Southern hemisphere is subjected to more solar radiation than the Northern hemisphere. Thus, TEC is larger in Southern hemisphere than in the Northern hemisphere. This picture changed during the SSW. The right panel in Figure 32 corresponds to UT when the largest disturbances in the low latitude ionosphere appear. The right panel shows TEC depletion at the magnetic equator and TEC enhancement in northern hemisphere. Disturbances were observed during a long time period from December 2012 to January 2013 and are reported by *Goncharenko et al.* [2013]. In this

subsection, we discuss only ionospheric disturbances that appeared in the geographical sector shown in Figure 32. All other effects are not discussed. Since ionospheric disturbances are driven from the underlying neutral atmosphere, the behavior observed in Figure 32 is not necessary the same along the equator for the same local time and different longitudes. Unfortunately, GIRO site coverage does not allow a making robust conclusion of the *foF2* behavior at other longitudes.

To perform this case study, we compare the *foF2* IRTAM maps and GPS TEC maps. Since TEC and the IRTAM *foF2* maps at SSW time are under investigation, we need quiet time values of both TEC and *foF2* in order to make conclusions about the ionospheric variations. The measurements obtained at 1500 UT on 15 December 2012 were taken as a baseline for ΔTEC map. We consider the corresponding IRI climatology map as a baseline for the IRTAM. In order to be complete, the *ΔfoF2* map for 1500 UT on 15 December 2012, as well as for local time in early January 2013 (before SSW), did not show significant variations. Thus, ΔTEC can be obtained by subtracting the left panel of Figure 32 from the right panel, and *ΔfoF2* is obtained as usual using the IRI and IRTAM. Figure 33 shows *ΔfoF2* map together with ΔTEC map.

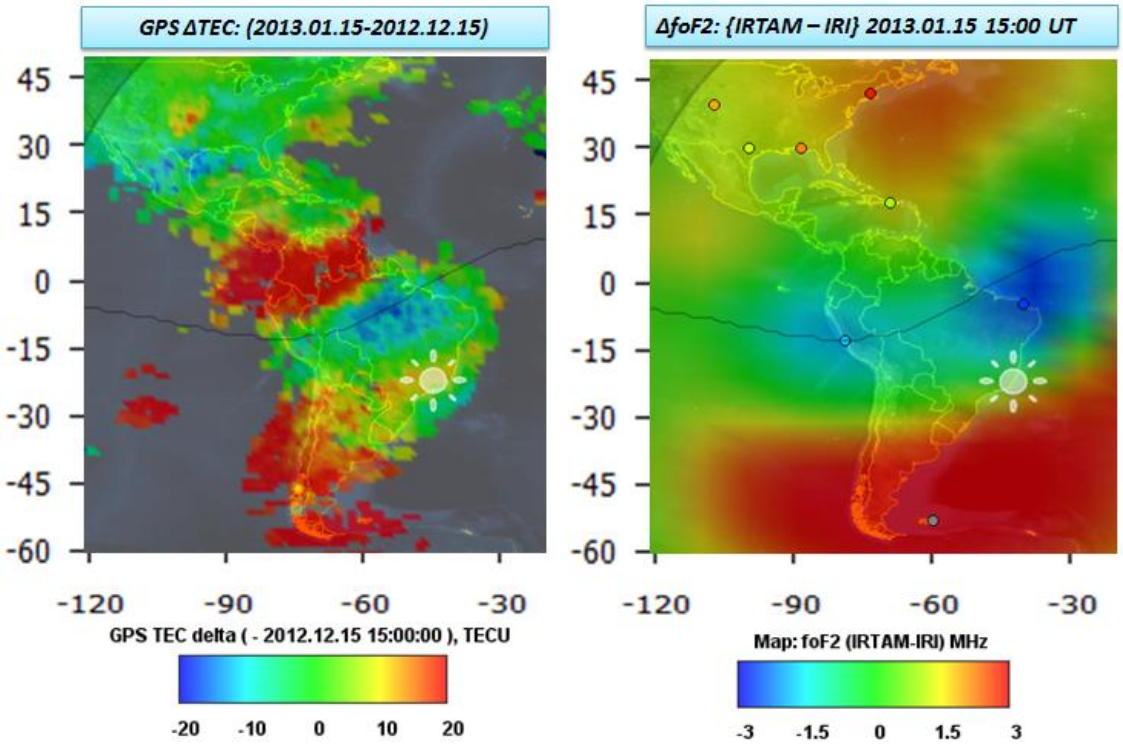



Figure 33. GPS ΔTEC map (on the left) and *ΔfoF2* (on the right). The same geographical sector as on Figure 32 is shown. Circles on the left panel correspond to the particular GIRO sites contributing data to the IRTAM. Color code is the same for map and sites.

TEC and *foF2* maps depict same features, which are enhanced TEC corresponding to increased values of *foF2*. Although the IRTAM does not reflect the increase in *foF2* to the North from the magnetic equator, the IRTAM is able to capture all other major effects discussed above. Increase of *foF2* to in northern hemisphere is absent on the IRTAM map due to missing measurements at this particular region during the SSW time. It is worth mentioning that the IRTAM uses only measurements from 8 sites to model *foF2* for this geographical sector. On the other hand, the number of GPS receivers is in the several hundred. Peak height disturbances, which could be obtained

only from the IRTAM, are much less in amplitude than *foF2* disturbances. This can be understood by the percentage *ΔfoF2* and *ΔhmF2* maps given in Figure 34.

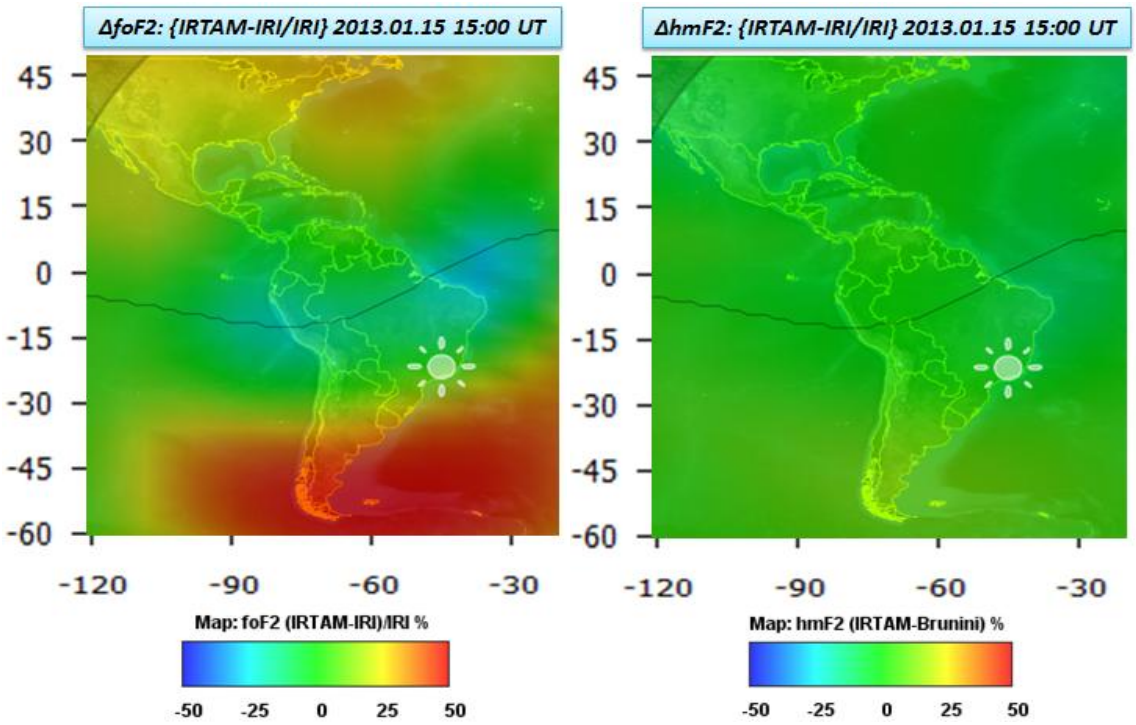



Figure 34. Percentage *ΔfoF2* (on the left) and *ΔhmF2* (on the right) maps during SSW.

This case study showed that if IRTAM maps and GNSS TEC could form a fruitful partnership, which will be useful for users.

# IV. DISCUSSION

## 4.1 Bias in hmF2 Histogram

First, we discuss the bias in the *hmF2* error histogram drawn by Figure 10. One possible reason for this bias is related to the uncertainties left after *foF2* and *M3000F2* were transformed to the *hmF2* [*Brunini et al.*, 2013]. The *hmF2* coefficients [*Brunini et al.*, 2013] are much "younger" than *foF2* coefficients [*Rush et al.*, 1989]. The *foF2* coefficients were subjected to different comparisons and reviews, and are trusted. The *hmF2* coefficients do not have this long history of validation. Moreover, *Brunini et al.* [2013] stated that *hmF2* fitting was not perfect and some residuals in the range from -10 to 10 km are left as errors of the transition between the IRI method of *hmF2* mapping and the direct *hmF2* mapping. These residuals add up to the errors due to *foF2* and *M3000F2* transformation into *hmF2,* which has been done according to *Bilitza et al.* [1979]. Nevertheless, if the global picture is considered, these errors should be distributed around zero such that mean value of $\varepsilon_{IRI}$ dataset is zero.

Another reason for the *foF2* and *hmF2* histogram discrepancy can arise from the measurements themselves. To understand this argument, we should consider how *foF2*

and *hmF2* observations are obtained. Both are deduced from measured ionograms, however the procedure is different. A typical ionogram is shown in Figure 35.

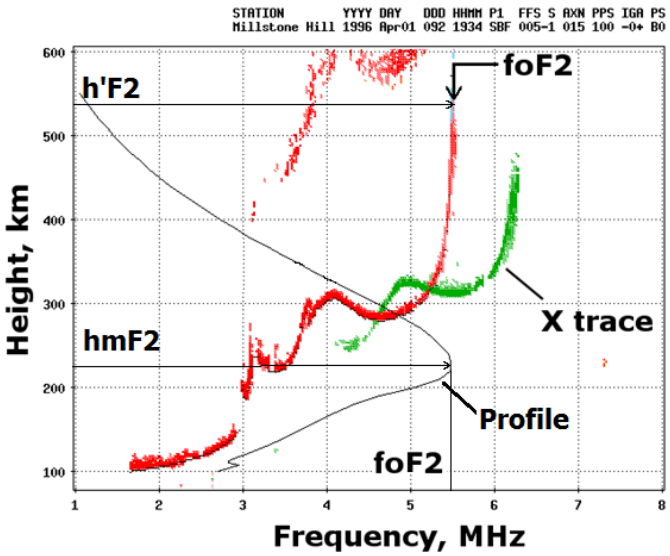



Figure 35. Example of an ionogram measured by the Millstone Hill Digisonde with autoscaled electron density profile. Retrieved from http://umlcar.uml.edu/DIDBase/. The red color refers to ordinary component of the signal (O trace), and green to the extraordinary component (X trace). The black curve is the electron density profile.

An ionogram represents the result of the sampling of the ionosphere with different frequencies. The transmitted wave reflects from the ionosphere at a certain height, given that its frequency is less than *foF2*, and returns to the ionosonde site with a certain delay. The reflection occurs at the height where the plasma density is such that plasma frequency coincides with the wave frequency. The delay between transmission and receiving can be recalculated to *virtual height* $h'$:

$$h' = \frac{1}{2} c \cdot t_{travel} \tag{12}$$

where $t_{travel}$ is the time transmission between detection of a radio wave, and $c$ is the speed of light in free space. When a radio wave travels in a plasma, its speed is not equal to speed of light in free space. The wave speed is reduced, and is slower as plasma density increases. The virtual height is always greater or equal to the real height, since $t_{travel}$ in plasma is greater than in free space. The virtual height of sounding signals plotted against frequency range is what constitutes an ionogram.

Critical frequency is directly deduced from an ionogram, which is just the maximal frequency of the O echo trace of F layer. Deduction of *hmF2* is more complicated and includes certain assumptions. Real peak height of the maximum density, *hmF2,* is related not only to measured virtual height, *h'F2*, but also depends on the underlying plasma. It is clear that radio waves which have a frequency equal to *foF2* have to travel all the way through the bottom side ionosphere (the ionosphere below *hmF2*). The same is also true for any point on the electron density profile. Thus, height of any particular electron density (plasma frequency) cannot be measured by using a single frequency. Samplings on all frequencies in range from 0 MHz to the frequency of the interest are required. After the required measurements are obtained, an inversion technique is applied in order to derive the electron density profile. There are several complications that can result in inaccurate values of *hmF2*, even if the ionogram showing a detectable trace and inversion is successful.

The first complication is the E valley, where the ionospheric E layer blacks out plasma above from being measured by an ionosonde. In Figure 35, the E valley corresponds to the local minimum above the E layer peak at around 110 km. The inversion technique should assume some model of the valley in order to obtain full electron density profile. This complication seems to be irrelevant to the bias shown by the bottom panel of Figure 10, since the valley shape model used in the Digisonde inversion technique is consistent with the valley model assumed by the IRI.

Another complication arises from the interpretation of the F layer cusp. The cusp is part of echo trace near *foF2* that has near vertical echoes. Quite often, the last echo is missing from, and the cusp is not depicted in, an ionogram. Missing echoes can happen due to various reasons not discussed here. If the cusp is not shown it is actually hard to say what the value for *h'F2* is correct. For *foF2* measurements, it is not that critical, since the uncertainty is less than half of the frequency step, which is usually in range 0.025-0.04 MHz. For *hmF2*, an undetermined cusp can result in a big uncertainty for *hmF2*. If the echo trace on ionogram is interpolated, *h'F2* is usually calculated according to a hyperbola fit. In this case, the *hmF2* can be 25 km larger compared to the case when no interpolation is performed. All of these factors introduce some ambiguity in *hmF2* deduction. Once the rules for ionogram interpretation are defined, *hmF2* values given by the same technique are consistent with each other. However, measurements resulting from other techniques can be different, even if the ionograms are the same in both cases [*Sauli et al.*, 2007]. Measurements that IRTAM assimilates are from the GIRO, and are

mostly provided by the Digisondes [*Reinisch*, 2009]. The profile inversion technique used for Digisondes is called NHPC [*Reinisch and Huang*, 1983; *Huang and Reinisch*, 1996]. This inversion technique implies no extrapolation beyond the last visible point of echo trace. According to *Chen et al.* [1994] the uncertainty of the *hmF2* measurements by the Digisondes is between -10 km to +10 km. So, there is possibility that *hmF2* is continuously underestimated. To make a valid and correct conclusion, extra *hmF2* validation efforts should conducted, using, for example, incoherent scatter radars which are able to measure *hmF2* directly.

## 4.2 IRTAM Error Sources.

A large portion of the previous section is devoted to the ability of the IRI formalism to represent same data which were assimilated. Although the IRTAM representation yields a two-fold improvement over the IRI representation, an average uncertainty of 0.31 MHz for *foF2* mapping remains. It is expected that observations fed into IRTAM would be represented perfectly. There are several reasons why this is not happening, and the residues are left. Generally, this is due to the fact that the expansion functions of the used formalism have limited order. The finite resolution of the formalism results in "smoothing" of original observations. Smoothed model representation can describe only limited time and spatial gradients in observations. Extra complications arise since temporal (13 functions) and spatial (76 functions) representations are coupled into

the total 988 coefficients. First, I considered smoothing due to limited order of time harmonics and then discuss spatial smoothing.

The temporal smoothing is an easy concept, which involves a lack of higher order harmonics to track fast temporal variations of a certain characteristic. The time variations in the IRTAM formalism are represented by 6$^{th}$ order Fourier series. This allows the model to resolve 4 hour period variations. Time variations with a period smaller than 4 hours cannot be represented, which is shown in Figure 36.

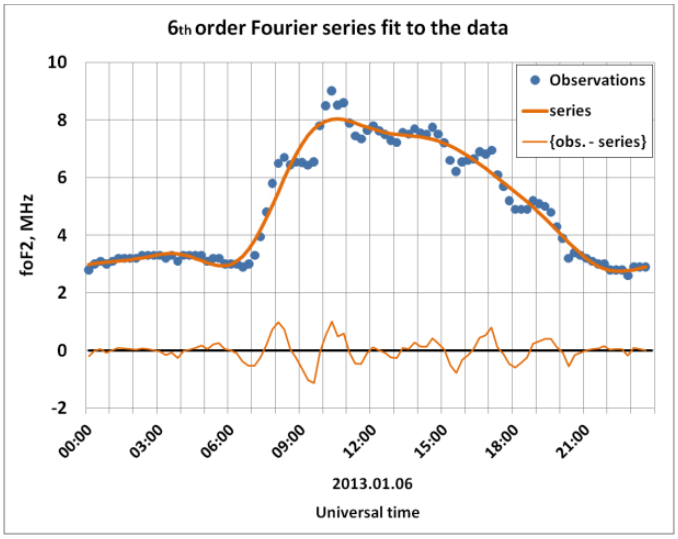



Figure 36. Illustration of the inability of 6$^{th}$ order harmonics to resove fast variations. The dots are the measurements at El Arenosillo (37.1°N, 6.7°W) site for the time interval indicated on the axis. The bold line is 6$^{th}$ order Fourier series fitted to the data, the thin line is residues left after fitting. Residues are originated from a limitation of the Fourier series order and show periodic varitions.

It should be pointed out that the bold line does not represent any particular model. This line is the best fit possible with the 6$^{th}$ order Fourier series. The IRTAM on top of

temporal variations will have restrictions due to spatial variations. The residues show a periodicity of about two hours, while the $6^{th}$ order Fourier series allows resolving as high as 4 hour period variations. One possibility to avoid this problem is to include higher order time harmonics. For a local model of the ionosphere, i.e. only at particular bounded geographical location, this can be done easily. For the IRTAM, it is not a straight forward task. Introduction of one extra order to existing time expansion will bring 152 (76*2=152) more coefficients. Remember, 76 is the number of spatial expansion functions, and the factor of 2 appears, because we have to include both *sin()* and *cos()* terms. The introduction of an increased number of coefficients is not the matter of computational time, and the expansion map into coefficients or synthesizing map from the coefficients can still be done very fast. The first real issue is that the IRTAM assimilative technique requires climatology coefficients. These coefficients represent the background state of a characteristic, and only corrections are calculated in order to match observations. Obviously, there are no climatology coefficients for potentially introduced higher order temporal harmonics in the original IRI coefficients set. This complication could be avoided by assumption that most of the time a $6^{th}$ order Fourier series is enough to accurately represent observations. Higher order harmonics are assumed to represent only variations connected with the ionospheric weather and not the climate. Figure 37 shows a typical situation of how good real observations could be represented by a $6^{th}$ order Fourier series. There some residues left, but they can be generally neglected.

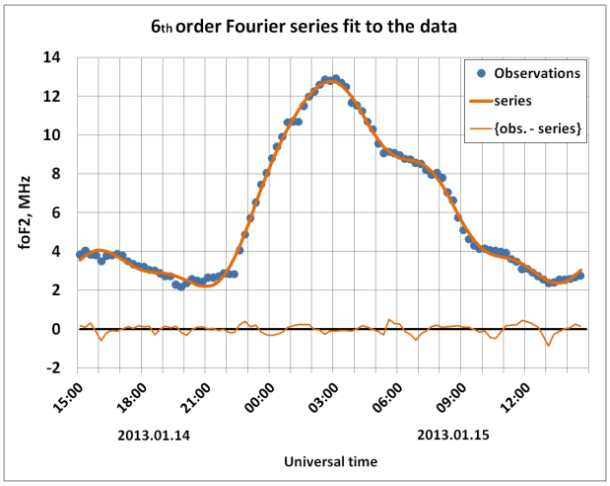



Figure 37. Illustration of accurate data representation by a 6$^{th}$ order Fourier series. Notation is the same as in Figure 15. The dots are the measurements at Jeju (33.4°N, 126.3°E) site for the time interval indicated on the figure.

The situation shown in Figure 37 is typical for the quiet undisturbed ionosphere. All variations faster than what is typically observed could be attributed to weather events. This allows us to assume that climatology background coefficients for these newly introduced harmonics are all zero. With this assumption, it is possible to introduce higher order time harmonics into the IRTAM. For time harmonics with order higher than 6, the coefficients are purely obtained as the result of assimilation. This will not break IRTAM-IRI compatibility, since the 988 coefficients formalism remains unchanged.

The issue of the climatology coefficients is resolved, so does it mean that now we can simply introduce higher order harmonics and increase the quality of the IRTAM? Unfortunately, there is another issue which is hard to avoid from a mathematical perspective. The temporal and spatial expansions are coupled in 988 coefficients. This

restricts variations of 13 temporal expansion coefficients from one spatial location to another. These coefficients cannot vary arbitrary, and the variations have to be allowed by spatial expansion, which is governed by 76 geographical functions. Since fast variations are linked to the weather events, they could be localized in some spatial region. The sites inside this region will draw these variations in the observations. Another site in neighborhood of the localized spatial region could not see this. Due to the limited number of geographic functions, only finite spatial gradients are allowed. Thus, the coupling between temporal and spatial coefficients may result in the situation, when observations are fitted well at the site, which draws fast variations. However, the same variations might be introduced at the other site, which does not show such variation. Hence, the characteristic representation appears to be good at the first site, but the representation is worse at the second site. Moreover, unphysical features are introduced at the location of the second site. The inverse could also happen, and the second site can suppress the real physical variations at location of the first site. Some further studies should be made in order to be able to say whether higher order time harmonics improve the quality of the model or not.

This last paragraph discussed the coupling between the spatial and temporal expansion and issues that can arise from it. This led us to the discussion of the spatial smoothing. This is similar to temporal smoothing only the limitation of the order and degree of geographical functions is concerned. The finite numbers of geographical functions result in finite resolution by latitude and longitude. In other words, sharp spatial

gradients may not be resolved. This is a major issue for equatorial ionosphere where sharp spatial gradients in *foF2* and *hmF2* maps occur frequently. For example, consider two spatially close sites, JJ433 (Jeju 33.4°N, 126.3°E) and IC437 (I-Cheon 37.1°N, 127.5°E), which are separated by 500 km. The JJ433 site is located inside the equatorial ionization anomaly and exhibits higher values of *foF2*. Even though IC437 is very close to JJ433 geographically, it is far enough for IC437 to appear slightly outside equatorial ionization anomaly. Because of that, IC437 shows smaller *foF2* values than JJ433. Figure 38 illustrates the inability of both models to reproduce big *foF2* variations with latitude.

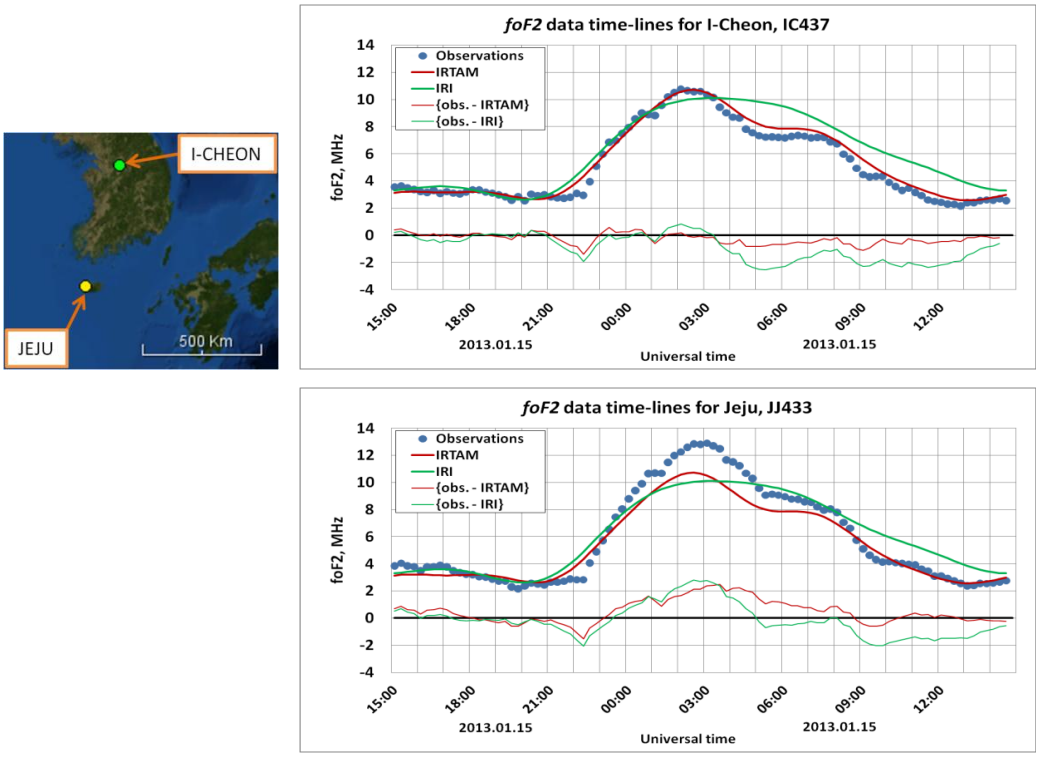



Figure 38. Illustration of sharp spatial gradients of *foF2* variations. Left panel shows two spatially close stations in South Korea: IC437 and JJ433. The panels on the right show observations (blue dots), IRTAM (red curve) and IRI (green curve) models.

The problem of sharp spatial gradients reproduction becomes more severe during space weather events, when sharp gradients appear frequently all over the globe. Sharp spatial gradients appear in data as shown in Figure 38. The problem of spatial resolution, by my reasoning, is more critical than the problem of temporal resolution. For time expansion, I can make the assumption that time variations faster than those already included in the IRI can be attributed purely to the ionospheric weather. In case of spatial expansion, the same could be also assumed, but this assumption is not as robust. Again, by the introduction of extra geographical functions the number of total coefficients will increase by 13. The climatology set of these extra coefficients cannot be assumed to be zero. Moreover, the form of the geographical functions is not as simple as for time harmonics. The construction of such new geographical functions should be subjected to very careful inspection and consideration.

Existing formalism describes ionospheric weather during quiet or even moderate disturbed time acceptably well. However, during severe space weather events the formalism should be enhanced in order to keep track of fast time variation and sharp spatial gradients of ionospheric characteristic behavior. This work is left for future versions of the IRTAM.

# V. CONCLUSION

This investigation was directed at verification and validation of the IRI-based real-time assimilative mapping (IRTAM) system used at UML GIRO to now-cast F2 layer critical frequency *foF2* and peak height *hmF2*. The *hmF2* direct mapping by the IRTAM is introduced in this thesis for the first time ever. Mapping of *foF2* was introduced previously by *Galkin et al.* [2012]. Mapping of *hmF2* is implemented with the same expansion formalism as *foF2* mapping and assimilation technique.

Overall, the IRTAM shows approximately two-fold improvement over the climatology IRI representations. Similar improvement is observed during storm time conditions or using only error-free validated observations. The variability of the *foF2* is higher for low-latitude ionosphere and *hmF2* varies more significantly at high latitudes. This is clearly seen from the IRI model comparison with the observational data, and is consistent with current knowledge of ionospheric dynamics. The IRTAM does not show significant dependence on geographical coordinates. Comparison of error histograms for *foF2* and *hmF2* reveals bias of *hmF2* data provided by climatology model of *Brunini et al.* [2013], which may be attributed to its construction artifacts; *foF2* climatology model [*Rush et al.*, 1989] shows no bias for the same conditions. It is therefore advisable to use *hmF2* values provided by IRTAM outside observatory locations with caution.

Testing of the spatial prediction capability of IRTAM outside observatory locations suggests that interpolation can be trusted further out in zonal direction than in the meridional direction. This makes it necessary to change the current implementation of the interpolation algorithm to treat longitude and latitude directions differently: such anisotropy will increase IRTAM accuracy of spatial prediction.

The IRTAM accuracy does not show significant dependence on the local time on different latitudes. The uniformity of the IRTAM deviation from observations implies that formalism used for the characteristics mapping performs equally well at different latitudes. Even though the IRTAM shows more than two times improvement over the IRI, significant residues between observations and the model are left unrepresented. On average, these residues are 0.31 MHz for *foF2* and 15 km for *hmF2*. When $\bar{\varepsilon}_M$ is calculated using only manually scaled measurements (error-free), while assimilation still uses all available data, $\bar{\varepsilon}_M$ decreases by 25%, which points to a noticeable effect of the measurement noise on the error statistics. In many cases, large model deviation from the data are actually explained by data errors of both ionospheric and non ionospheric origins. Another source of deviations was found to be the limited order of the classic spatial and temporal expansion implemented in IRI. Attempts to increase the order of time harmonics and geographical functions were made but insufficient spatial observational data coverage often leads to unstable performance and unreasonable prediction.

Case studies reveal an additional IRTAM utility in investigations of various space weather events. Timelines of the global IRTAM-IRI deviations appear a convenient tool for quick identification of all important effects that are evident in the observational data. Both statistical and case studies reveal the artifacts of the IRTAM interpolation over the areas of limited sensor coverage. The multi-cell structures often seen in the maps during strong disturbances (such as the example shown in Figure 29) are likely to result from the gradual suppression of the adjustments during the interpolation. While similar multi-cell *foF2* structure along the magnetic equator has been previously reported due to the neutral atmosphere weather impact [*Immel et al.*, 2006], the multi-cell morphology observed in Figure 29 is believed to be an artifact of the interpolation technique. Although the IRTAM is capable of certain spatial prediction due to its analysis of the diurnal harmonics of the ionospheric timeline, no-data areas well outside the sensor sites are still no different from the background climatology. Such modeling cannot be regarded as completely global. IRTAM should be considered as a multi-regional model in its current state. Growing number of the GIRO locations will increase the capability of the IRTAM to sense ionosphere globally. Another way to improve IRTAM is to employ data from other sources. Independent measurements can be used to validate results of the assimilation if they are made at the same location, or independent measurements can be added up to the ionosondes' measurements to increase spatial coverage. New measurements of characteristics, which are not directly measured by ionosondes, such as GPS TEC measurements or in situ measurements of the plasma density, will also be

useful. These measurements can be used to restrict the value of any ionospheric characteristic mapped by IRTAM. Either way, the first step is to implement full electron density profile real-time mapping in the IRTAM. The full electron density profile mapping is critical, since TEC and plasma density at a particular height can be directly derived from the electron density profile. The derived quantity can be directly compared to the independent measurement of the same quantity under the same conditions. Furthermore, full electron density profile might be corrected according to the results of this comparison. So far, only 2D maps of *foF2* and *hmF2* characteristics were presented. One step further will be to implement the mapping if the third main ionospheric characteristic, the *slab thickness B0* (see Figure 1). The electron density profile is mainly defined by *foF2*, *hmF2* and *B0*, so with the *B0* maps the IRTAM will be capable to yield a global 3D ionosphere.

During the development of the IRTAM and as a result of this thesis several useful products have been introduced. The database, called GAMBIT, was created and is currently routinely populated with the most recent coefficient for real-time *foF2* and *hmF2* mapping. The new coefficients are completely compatible with the climatology IRI coefficients, and can be used by IRI users without any modification. Most recent IRTAM real-time 24 hour IRTAM (weather) maps of two characteristics as well as (IRTAM – IRI) maps are posted since December 2012, and are routinely updated every 15 minutes on the web page http://giro.uml.edu/IRTAM/. The stand-alone desktop cross-platform application, called GAMBIT Explorer, is ready for release in December 2014. All

characteristics maps presented in this thesis are produced by GAMBIT Explorer. The IRTAM was presented at numerous scientific meetings and conferences, including 40$^{th}$ COSPAR assembly held in August 2014, where the community expressed interest in the IRTAM. The IRTAM will be further developed and hopefully become the general tool for scientific research in the field of Space Physics.

# VI. LITERATURE CITED

Adeniyi ,J.O., D. Bilitza, S.M. Radicella, A.A. Willoughby (2003), Equatorial F2-peak parameters in the IRI model, Advances in Space Research, Volume 31, Issue 3, 2003, Pages 507-512, ISSN 0273-1177

AIAA American National Standard (2011), Guide to Reference and Standard Ionospheric Models, *AIAA G 034-2011*, American Institute of Aeronautics and Astronautics, Reston, VA, 2011

Bilitza, D., M. Sheikh, R. Eyfrig (1979), A global model for the height of the F2-peak using M3000 values from the CCIR numerical map. Telecommun. J. 46 (9), 549–553, 1979.

Bilitza, D. (1990), International Reference Ionosphere 1990, 155 pages, National Space Science Data Center, NSSDC/WDC-A-R&S 90-22, Greenbelt, Maryland, November 1990.

Bilitza, D. (2006), The International Reference Ionosphere-Climatological Standard for the Ionosphere, *National aeronautics and space administration greenbelt md goddard space flight center*, 2006.

Bilitza, D., B. W. Reinisch, and L. A. McKinnell (2010), "International Reference Ionosphere 2010." In AGU Fall Meeting Abstracts, vol. 1, p. 1580. 2010.

Brunini, C., J. F. Conte, F. Azpilicueta, D. Bilitza (2013), A different method to update monthly median *hmF2* values, Advances in Space Research, Volume 51, Issue 12, 15 June 2013, Pages 2322-2332, ISSN 0273-1177. doi:10.1016/j.asr.2013.01.027.

Chen, C.F., B.W. Reinisch, J.L. Scali, X. Huang, R.R. Gamache, M.J. Buonsanto, B.D. Ward (1994), The accuracy of ionogram-derived N(h) profiles, Advances in Space Research, Volume 14, Issue 12, December 1994, Pages 43-46, ISSN 0273-1177, http://dx.doi.org/10.1016/0273-1177(94)90236-4.

Cherniak, Iu.V., I.E. Zakharenkova, D.A. Dzyubanov (2013), Accuracy of IRI profiles of ionospheric density and temperatures derived from comparisons to Kharkov incoherent scatter radar measurements, Advances in Space Research, Volume 51, Issue 4, 15 February 2013, Pages 639-646, ISSN 0273-1177, http://dx.doi.org/10.1016/j.asr.2011.12.022.

Damboldt, T., and P. Suessmann (2011), Information document on the analysis and validity of present ITU *foF2* and m(3000)f2 maps, *International Telecommunication Union*, question ITU-R 212-1/3. Retrieve from http://www.itu.int/md/R07-WP3L-C-0086/en

Dudeney, J.R (1975), A simple empirical method for estimating the height and semi-thickness of the F2-layer at the Argentine Islands Graham Land. British Antarctic Survey, Science Report 88, London, 1975.

Finlay, C. C., et al. (2010), International geomagnetic reference field: The eleventh generation, Geophys. J. Int., 183(3), 1216–1230. doi:10.1111/j.1365-246X.2010.04804.x.

Galkin, I. A., B. W. Reinisch, X. Huang, and D. Bilitza (2012), "Assimilation of GIRO data into a real-time IRI." *Radio Science* 47, no. null (2012): RS0L07. doi:10.1029/2011RS004952

Goldstein, J., M. Spasojevic, P. H. Reiff, B. R. Sandel,W. T. Forrester, D. L.Gallagher, and B. W. Reinisch (2003), Identifying the plasmapause in IMAGE EUV data using IMAGE RPI in situ steep density gradients, J. Geophys. Res., 108(A4), 1147, doi:10.1029/2002JA009475.

Goncharenko, L., J. L. Chau, P. Condor, A. Coster, and L. Benkevitch (2013), Ionospheric effects of sudden stratospheric warming during moderate-to-high solar activity: Case study of January 2013, Geophys. Res. Lett., 40, 4982–4986, doi:10.1002/grl.50980.

Heelis, R. A., and W. R. Coley (2007), Variations in the low- and middle-latitude topside ion concentration observed by DMSP during superstorm events, J. Geophys. Res., 112, A08310, doi:10.1029/2007JA012326.

Huang, X. and B.W. Reinisch (1996), Vertical electron density profiles from the Digisonde network, Advances in Space Research, Volume 18, Issue 6, 1996, Pages 121-129, ISSN 0273-1177.

Huang, X., B. W. Reinisch, I. Galkin, and D. Bilitza (2012), A Real-Time Assimilative Model for IRI, *EGU General Assembly 2012 Paper #EGU2012-12828*, 2012

Immel, T. J., E. Sagawa, S. L. England, S. B. Henderson, M. E. Hagan, S. B. Mende, H. U. Frey, C. M. Swenson, and L. J. Paxton (2006), Control of equatorial ionospheric morphology by atmospheric tides, Geophys. Res. Lett., 33, L15108, doi:10.1029/2006GL026161.

International Telecommunications Union (2009), *ITU-R reference ionospheric characteristics*, Recommendation P.1239-2 (10/2009). Retrieved from http://www.itu.int/rec/R-REC-P.1239/en on November 29, 2011.

Jones, W. B., R. M. Gallet (1962), Representation of diurnal and geographical variations of ionospheric data by numerical methods. ITU Telecommun. J. 29 (5), 129–149, 1962.

Jones, W. B., R. P. Graham, M. Leftin(1969), Advances in ionospheric mapping by numerical methods, ESSA Technical Report ERL 107-ITS 70, US Government Printing Office, Washington DC.

Loewe, C. A., and G. W. Prölss (1997), Classification and mean behavior of magnetic storms, J. Geophys. Res., 102(A7), 14209–14213, doi:10.1029/96JA04020.

Maruyama, T. (2006), Extreme enhancement in total electron content after sunset on 8 November 2004 and its connection with storm enhanced density, Geophys. Res. Lett., 33, L20111, doi:10.1029/2006GL027367.

Stanislawska, I., Belehaki, A., Jakowski, N., Zolesi, B., Gulyaeva, T., Cander, L., Reinisch, B., Pezzopane, M., Tsagouri, I., Tomasik, L., & Galkin, I., (2010) COST 296 scientific results designed for operational use, Annals of Geophysics, 52(3-4), 423-435. doi:10.4401/ag-4598, 2010.

NSWP, National Space Weather Program (2012), *Agency Roles and Missions*, http://www.spaceweathercenter.org/swop/nswp/1.html, retrieved on Dec 11, 2012.

Rawer, K. (1963), Propagation of decameter waves (HF band), in *Meteorological and Astronomical Influences on Radio Wave Propagation*, (Ed. B. Landmark) (New York Academic Press), pp. 221-250.

Reinisch, B.W., and X. Huang (1983), Automatic calculation of electron density profiles from digital ionograms 3: Processing of bottomside ionograms, Radio Sci., 18, 477-492, 1983.

Reinisch, B. W., et al. (2000), The radio plasma imager investigation on the IMAGE spacecraft, Space Sci. Rev., 91, 319.

Reinisch, B. W., et al. (2001a), First results from the radio plasma imager in IMAGE, Geophys. Res. Lett., 28, 1167.

Reinisch, B. W., X. Huang, P. Song, J. L. Green, S. F. Fung, V. M. Vasyliunas, D. L. Gallagher, and B. R. Sandel (2004), Plasmaspheric mass loss and refilling as a result of a magnetic storm, J. Geophys. Res., 109, A01202.

Reinisch, B. W., et al. (2009), New Digisonde for research and monitoring applications, Radio Sci., 44, RS0A24, doi:10.1029/2008RS004115.

Reinisch, B. W., and I. A. Galkin (2011), Global ionospheric radio observatory. *Earth, planets and space 63*, no. 4 (2011): 377-381.

Rush, C., M. Fox, D. Bilitza, K. Davies, L. McNamara, F. Stewart, M. PoKempner (1989), Ionospheric mapping – an update of fF2 coefficients, Telecomm. J., 56, 179-182, 1989.

Sauli, P., Z. Mosna, J. Boska, D. Kouba, J. Lastovicka, D. Altadill (2007), Comparison of true-height electron density profiles derived by POLAN and NHPC methods. Stud. Geophys. Geod. 51 (3), 2007, pp. 449–459.

Sahai, Y., et al. (2009), Effects observed in the Latin American sector ionospheric F region during the intense geomagnetic disturbances in the early part of November 2004, J. Geophys. Res., 114, A00A19, doi:10.1029/2007JA013007.

Scherliess, L., Schunk, R. W., Sojka, J. J., Thompson, D. C., & Zhu, L. (2006). Utah State University global assimilation of ionospheric measurements Gauss-Markov Kalman filter model of the ionosphere: model description and validation, *Journal of geophysical research, 111(A11)*, A11315. doi:10.1029/2006JA011712.

Wanliss, J. A., and K. M. Showalter (2006), High-resolution global storm index: Dst versus SYM-H, J. Geophys. Res., 111, A02202, doi:10.1029/2005JA011034.

Zong, Q.-G., B. W. Reinisch, P. Song, Y. Wei, and I. A. Galkin (2010), Dayside ionospheric response to the intense interplanetary shocks–solar wind discontinuities: Observations from the digisonde global ionospheric radio observatory, J. Geophys. Res., 115, A06304, doi:10.1029/2009JA014796.