Chapter 4

Ionospheric Tomography

The use of tomography to determine the ionospheric electron contents started with a 2D approach using polar satellites and placing the receivers along the ground track of these satellites ([53], [54], [55]). This approach has the limitations in space and time dictated by the pass of the satellites and the possible location of the receivers. The most common technique for solving the inverse problem in 2D ionospheric tomography uses Empirical Orthogonal Functions as a first set of basis functions to obtain the large scale evolution of the ionosphere. Later, a voxel approach is taken to finally characterize the residues.

The full constellation of GPS satellites together with a world-wide network of GPS receivers (the International GPS Service network, for instance) makes possible a global four dimensional tomographic solution of the ionospheric free electron density field. We will here describe the procedure, simulations and results as well as a calibration application to illustrate the potential of GPS ionospheric tomography.

4.1 Ionospheric tomography equations

It was shown in Chapter 2 that the *ionospheric combination* of observables was already a slant measurement of the ionospheric contents. The carrier-phase smoothing of the pseudorange reduced the level of noise. However, a remaining alignment error may exist and, in addition, other frequency dependent terms, such as instrumental delays, are also present. The slant ionospheric observable is then:

$$y(t_i) = \hat{P}_I(t_i) = \alpha_I I(t_i) + \hat{c}_{rcv}(t_i) + \hat{c}_{trx}(t_i) + c_{align} = \alpha_I I(t_i) + c_{rcv}(t_i) + c_{trx}(t_i) \quad (4.1)$$

This equation will be the base to construct the equations for ionospheric tomography: it relates the observable at a given time $y(t_i)$ (in m) with the integrated measurement of the ionospheric electron contents ($I(t_i)$, in el/m^2) at that time and along the same line, and

explicitly considers the instrumental delays. The assumed noise value of the observations is $\sigma_y = 0.1$ m.

If we consider the voxel-based tomography, the observable in Equation 4.1 can be related to the electronic density at each voxel (x_j) by linearizing the system and considering no bending of the ray across the atmosphere:

$$y_i = \alpha_I \left(a_1 x_1 + ... + a_{N_v} x_{N_v}\right) + x_{N_v+k} + x_{N_v+N_r+p}, \tag{4.2}$$

where it is considered that we have N_v voxels to solve for and that ray i involves ground station k and satellite p, N_r is the number of receivers, a_j is the length of the ray across voxel j and α_I is the constant relating the delay in meters to the volume density of electrons (see Section 2.2.5.1).

The instrumental delays can be considered to be constant during a period of 24 hours. Longer periods of data may require to consider them as stochastic parameters (see [56] for further details).

4.1.1 Differencing vs. direct approach in tomography

The presence of instrumental constants increases the number of unknowns: we have to solve for the electron density at each voxel and for $N_r + N_t - 1$ constants (note that the equations only relate the sum of pairs of constants and we hence need to set one to zero), where N_r is the number of receivers and N_t is the number of transmitters. This damps the numerical implementation of the processing.

This can be surmounted by differencing. As it is described in [57], the data are divided into phase connected arcs. Within each arc, each observation is substracted from a reference ray. This cancels the satellite and the receiver constants. Nevertheless, the noise power in the new observable is the addition of the noise power of the observable considered and the one from the reference ray (assumed to be uncorrelated gaussian noise). Therefore, the latter should be chosen to have the smallest noise. This corresponds to the observation at the highest elevation within each arc since the SNR is better and the multipath is lower. We rewrite our equations as:

$$y(t_i) - y(t_k) = \left(\mathbf{A_i} - \mathbf{A_k}\right) \alpha_I \mathbf{x} \tag{4.3}$$

where $y(t_i)$ is the ray considered, $y(t_k)$ is a the reference point of the arc and $\mathbf{A_i}$ and $\mathbf{A_k}$ are the corresponding rows in the jacobian matrix $\mathbf{A}$.

However, while differencing provides many practical advantages and the solution appears to be cleaner, it has two major drawbacks: first, the increase of the noise in the observables;

second, the constants in the estimation process act as nuisance parameters that are noise sinks, partially absorbing the effect of bad data or insufficient modeling.

Furthermore, the estimation of constants and the solution field in the same least-squares process provides the system with the absolute calibration capability ([2]). As we will show in Section 4.5, the tomographic technique can ingest data from sources other than GPS data (e.g. Radar Altimeter) and solve for the calibration constants of all the instruments.

4.2 Two-layers global solution

The first approach is to obtain a global solution for the ionospheric electron contents using data from the world-wide distributed IGS stations. The grid to be used is partly determined by the computer capability and by the expected resolution. We showed in Section 3.1.2.3 that the eigenvalue cut-off determines the noise in the solution. Considering that we write $\mathbf{S}$ (from Equation 3.17) in units of 10^5 meters (hkm), $\mathbf{x}$ is in Tera electrons/m^3, and $\mathbf{y}$ is in meters, then $\alpha_I = 1.05$ (m^3/Tera el)(m/hkm). If we want a resolution of 0.1 Tera el/m^3 and we expect a noise in the measurement, after phase-smoothing, of 0.1 m, then the minimum singular value ν should be 1 hkm since $\sigma_x \sim \sigma_y/(\alpha_I \nu)$. That is, we have to apply constraints to the solution so that the minimum eigenvalue of $\mathbf{S}^T \left(\mathbf{C}_{\hat{y}}\right)^{-1} \mathbf{S}$ is $\omega = \nu^2 = 1$ hkm^2. As we stated in Chapter 3, if the minimum singular value is ν =100 km, then a ray has to cross 100 km to provide information at a level greater than the noise. This distance is, in fact, an effective distance, since N close-by rays will carry information of the same vicinity and when combining all of them the noise will be downscaled by a factor of $\sqrt{N}$. Regarding the horizontal resolution, we need to consider the particular number of stations available and their distribution.

We select the date October 18th, 1995, for its high geomagnetic and solar activity indices (as distributed by the US National Geophysical Data Center (NGDC) and the National Space Science Data Center (NSSC)), as well as for the sudden decrease in the electron content measurements provided by the Geosynchronous Operational Environmental Satellites (GOES) (as supplied by the NGDC) For this day we have chosen about 60 IGS stations, mainly distributed over land. If we consider that the stations spacing is about the horizontal resolution then, if the stations were uniformly distributed over 1/3 of the Earth circumference (the equivalent in terms of land-sea proportion), then we would have some 7 stations to cover 12,000 km, yielding a 1,800 km horizontal resolution. Therefore, the grid has 20 divisions in longitude and 10 in latitude. The Kalman filtering will consider 120 minutes batches as a time span of the ionosphere large-scale dynamics. Defining this grid in horizontal and two layers in height of 100 km thick, we compute the average number of rays crossing each voxel during

120 minutes. It turns out that $N \approx 1$. We increase the thickness of the layers to 150 km in order to be well above the 100 km critical value.

Therefore we set the conditions for:

- 60 IGS stations
- grid of 20x10x2 with a total of 400 voxels plus 85 constants
- 120 minutes-batch Kalman filtering
- horizontal smoothing applied to increase the eigenvalues to be above 1 hkm^2

The results are shown in Figure 4.1 for UTC 12-14, as time series in Figure 4.2, and the histogram of the post-fit residuals is shown in Figure 4.3. These results were published in [44], making use of the correlation functional approach.

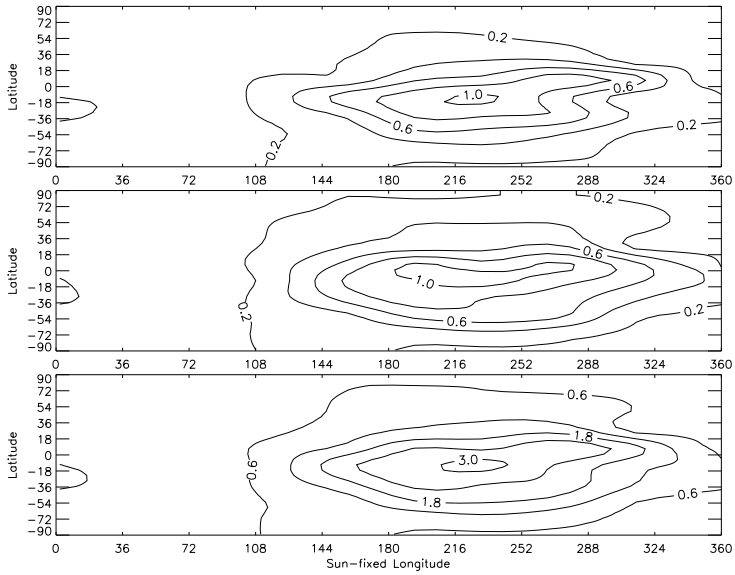

Figure 4.1: The result for UTC 12-14, showing from bottom up, the TEC in the two layers in electrons per square meter divided by 10^{17}, the lower layer and the higher one, with electron content in Tera-electrons per cubic meter.

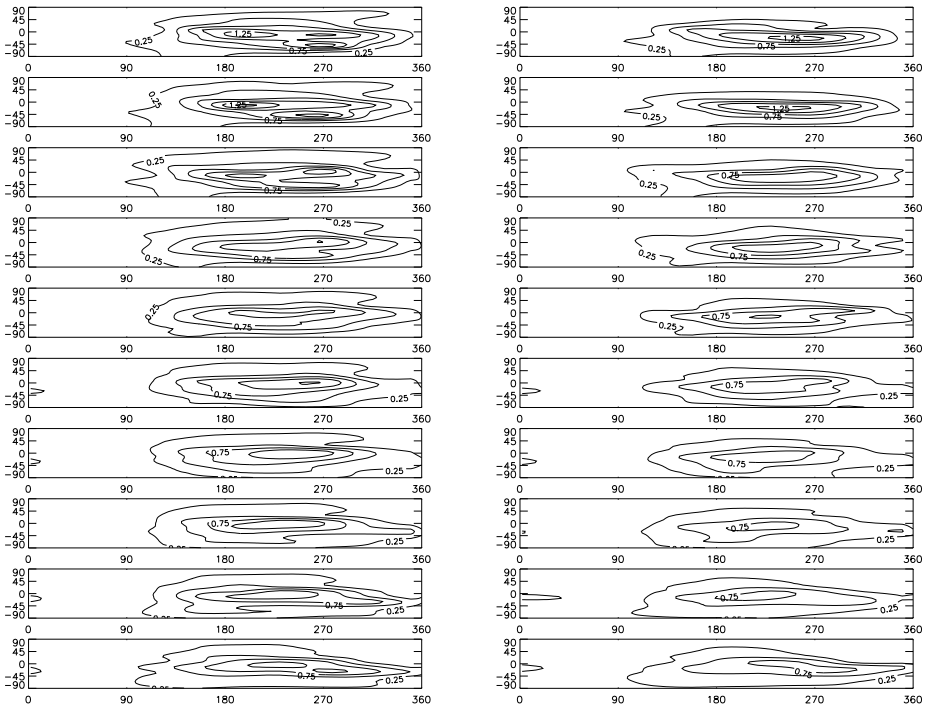

Figure 4.2: Time series solutions: on the left column, from bottom up, is the lower layer (200-350 km) starting at UTC 2-4 and up to UTC 20-22. On the right is the upper layer. Electron content is given in Tera electrons per cubic meter.

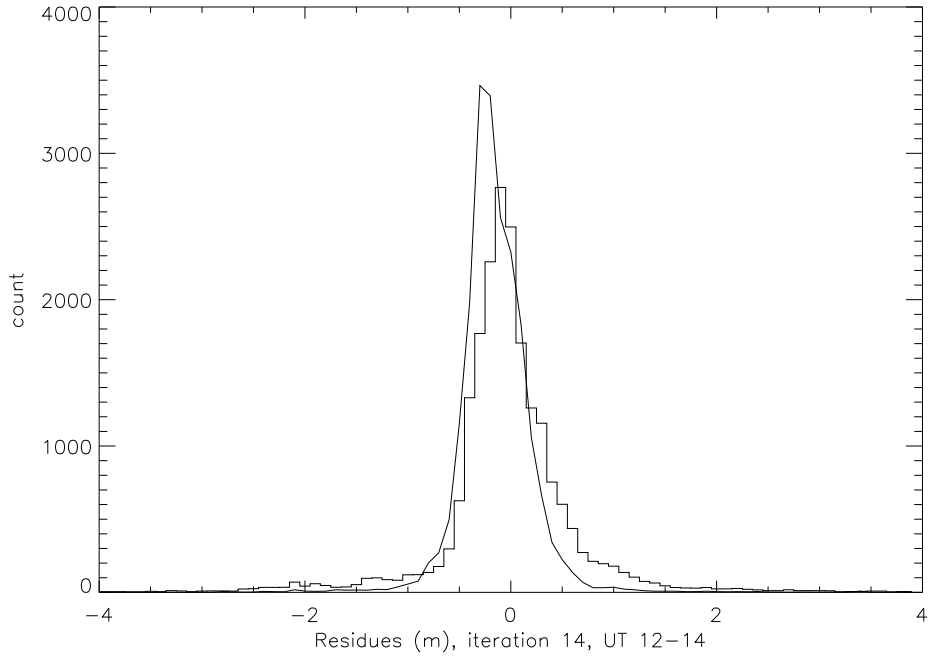

Figure 4.3: Histogram of the residues of the 2 layer ionospheric solution for different bin-size.

4.3 Tomography with Occultation data

Vertical resolution of ionospheric tomography can be greatly improved if there are occultation data available ([6]). The occultation technique was already used in the 70s to study the planetary atmospheres (for instance, the signals sent to the Earth by Mariner 5 spacecraft when occulting behind Venus were used to infer characteristics of its atmosphere, see [58]). GPS signals received by a Low Earth Orbiter (LEO) in the velocity or anti-velocity directions of its movement would lead to the same geometry (see Figure 4.4). The signal thus received would vertically scan the atmosphere. A first application is to retrieve the extra delay produced by the ionosphere to enhance the vertical resolution of ionospheric tomography.

4.3.1 The GPS/MET mission

The proof-of-concept experiment was the successful GPS/MET mission on board the Microlab-1 satellite (see [59] for details). The GPS/MET experiment consisted of a Turbo Rogue GPS receiver, modified to operate in a space environment and to track GPS signals in occultation conditions, and a hemispherical antenna placed facing the Anti-Velocity direction of the Microlab-1. The satellite orbit was circular, with a height of 775 km and 70^o inclination.

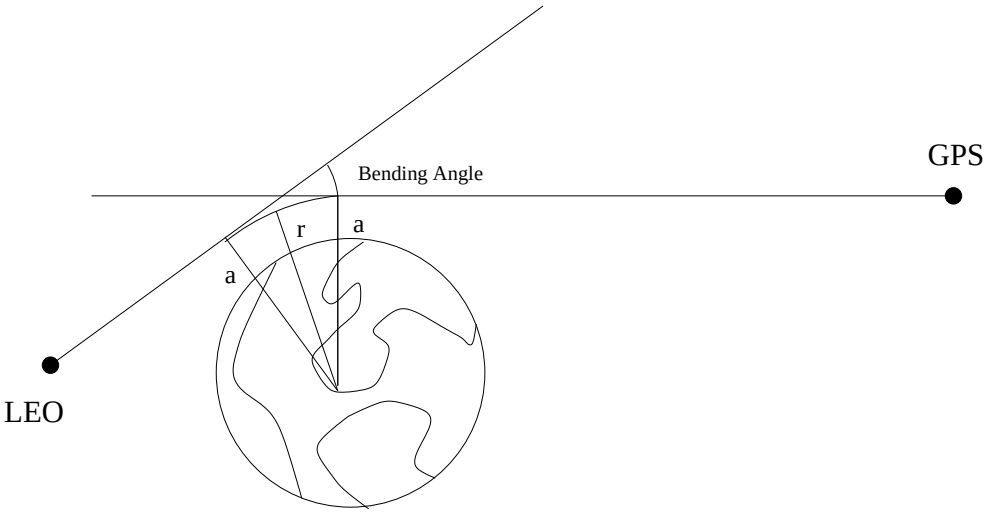

Figure 4.4: Geometry of an occultation.

The receiver tracked the GPS signals at three different rates, depending on the height of the impact parameter (a in Figure 4.4): 0.1 Hz, 1 Hz and 50 Hz for successively lower heights. The high rate covers about the last minute of the occultation. The first rate is used for navigation purposes and for ionospheric tomography studies; the second is used for ionospheric profiling (see [60], [61], [62]); and the high rate data are used to obtain refractivity profiles in the neutral atmosphere (see [20], and [63] for further details on the procedure). The mission was meant to freely provide the data to the scientific community. These data have been used as a reference for the design of further and more ambitious occultation experiments which are planned within the next years (see Table 4.1 for a the scheduled occultation missions).

Mission	**Agency**	**Scheduled**
Oersted	Denmark/JPL	1999
Sunsat	SouthAfrica/JPL	1999
CHAMP	GFZ/JPL	Jan 2000
SAC-C	Argentina/JPL	2000
COSMIC	UCAR/Taiwan	2002
MetOp	EUMETSAT/ESA	2002
NPOESS	IPO	2005

Table 4.1: Scheduled Missions incorporating a GPS occultation experiment. Dates are only tentative.

4.3.2 Localization of the occultations: the ring solution

While including occultation data in the tomographic solution enhances the vertical resolution, the location of the receiver at about 700 km high limits the existence of such data to a region

around the orbit of the LEO. In [48] it was shown that given a *uniform* field, with no horizontal features, and a 0-1-0 structure in vertical, the use of occultation data meant a big improvement in the tomographic reconstruction by reducing the number of zero eigenvalues. Because such field was constant in each layer, the information extracted in the region close to the LEO orbit (the GPS/MET mission orbit) was extrapolated to the rest by means of the horizontal smoothing. However, if we consider the locations of the intersections of occulting rays with the nominal height of the ionospheric electron contents peak, we can see that they form a "belt" around the Earth (in inertial coordinates), where the vertical resolution is actually enhanced (see Figure 4.5). Now, a realistic global solution with good vertical resolution using ground and occultation data may only be expected when the orbiting receiver is scanning the region with higher electron concentration and, hence, the poorly determined regions can be considered to be smoother, and basically dominated by the well determined region: that is, when the LEO orbit intersects the equatorial plane at the Sun's right ascension. This approximately happens, e.g., in the solution for the 21st February 1997, but it is the complete opposite (the orbit plane is normal to the direction of the Sun) on the 6th of January 1997. These two days correspond to prime times, when the steering of the Microlab-1 was such that the GPS antenna was facing the Earth limb in the anti-velocity direction and AS was turned off. We show in Figures 4.6 and 4.7 the results for these two days in a 24 hours solution. In the former, the location of the peak electron contents is slightly shifted in horizontal to the orbital plane right ascension; in height, the peak is at 250 km above the Earth's surface, and higher layers show a decay in electron density, as it is expected in ionospheric models. In addition the values coincide with those given by ionosondes. On the latter the peak value is more spread in horizontal and the resolution in the vertical dimension is rather poor, being the solutions for the layers at 400 km and at 550 km very similar, hence revealing the low vertical discrimination; because of this spread, the quantitative values at the peak are much lower than the ones given by ionosondes. We are therefore forced to focus our vertical resolution in the region of the LEO satellite orbit.

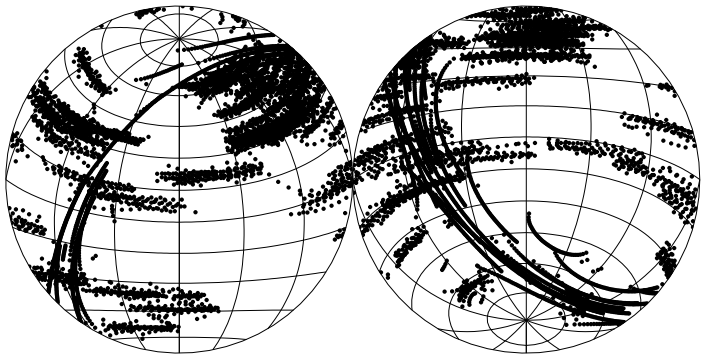

Figure 4.5: Intersection points of the rays with a 250 km layer height during three hours of data, in inertial coordinates for February 21, 1997. It can be appreciated that most of the occultation data are confined to a region around orbital plane (see arcs).

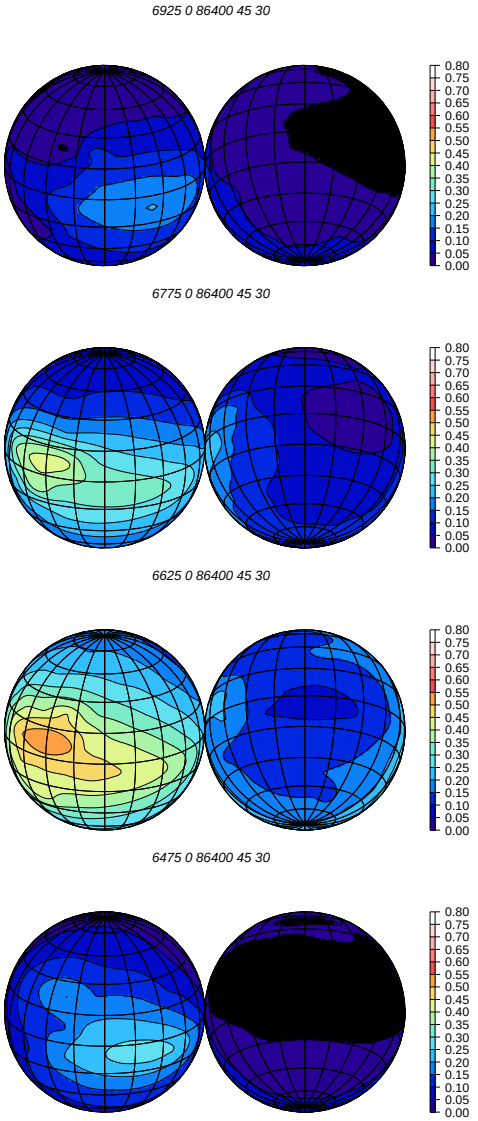

Figure 4.6: 24 hours solution for day 21st Feb 1997. The four layers are displayed from top to bottom corresponding to heights 6925, 6775, 6625, and 6475 km from the center of the Earth. Solution is given in inertial coordinates. The right ascension of the position of the maximum is correctly located in horizontal (except for the slight shift in 6925 km and 6475 km) and in height.

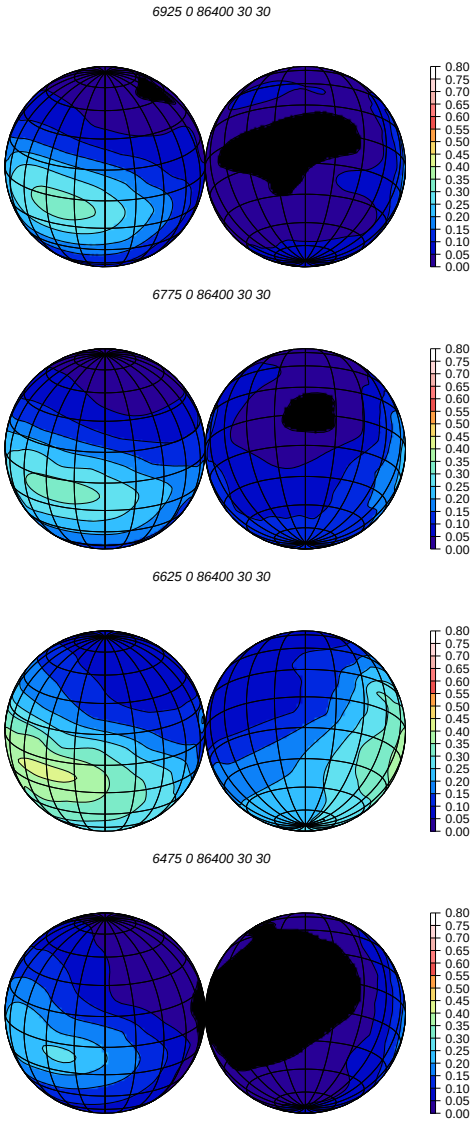

Figure 4.7: 24 hours solution for day 6th January 1997. The four layers are displayed from top to bottom corresponding to heights 6925, 6775, 6625, and 6475 km from the center of the Earth. Solution is given in inertial coordinates. In this case the orbit plane was normal to the direction of the Sun and the spread of the location of the peak in the horizontal and in the vertical is apparent.

Limiting our tomographic solution to the plane of the receiving satellite orbit may seem to revert our solution to the 2D tomography mentioned earlier. Nevertheless, since occultations do not take place in the exact anti-velocity direction but with some azimuth deviation angle (see Figure 4.8), the resulting scanned region still forms a 3D structure. The *ring* solution therefore, considers the first 10 degrees at each side of the orbital plane and up to six layers in height. In order to form the grid, the coordinate system is rotated so that the orbital plane is defined in the new reference system as having colatitude equal to $\pi/2$. Longitude is defined along the circumference and latitude being north or south of the orbital plane (see Figure 4.8). The new set of data to be analyzed consists of

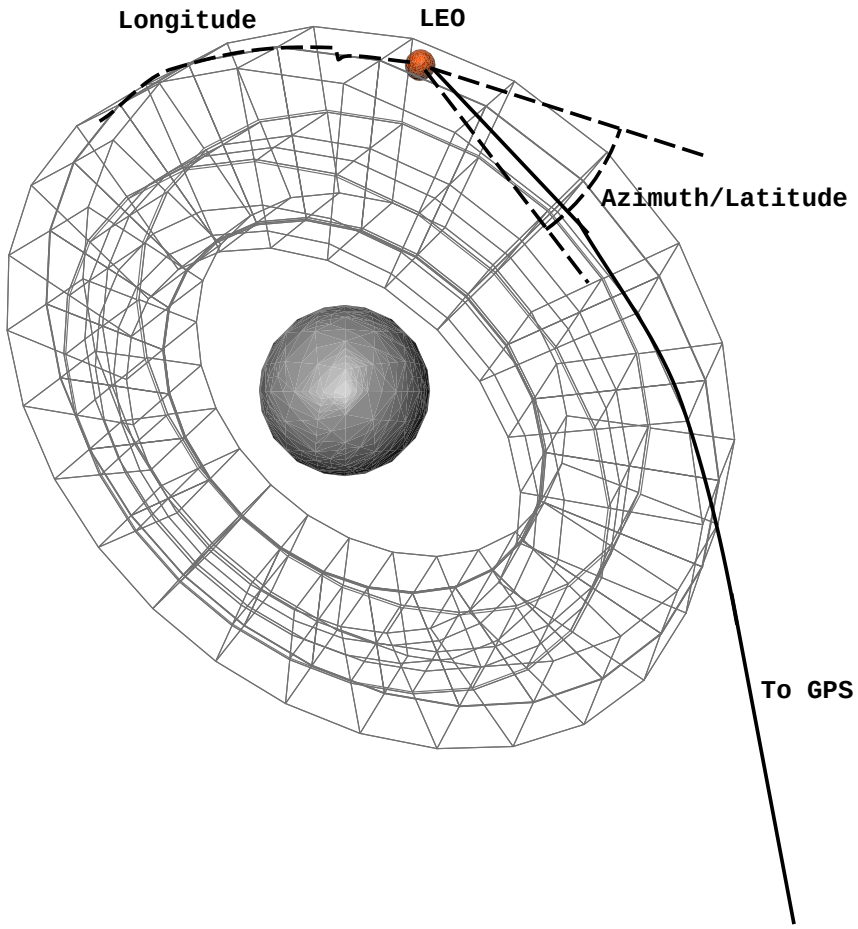

Figure 4.8: Grid used for the ring solution.

- Day selected: 21st February 1997, prime time for the GPS/MET mission and AS off.
- 100 IGS Ground Stations
- Low Rate (0.1 Hz) data from GPS/MET

- 4 divisions (5^o) in colatitude, 30 divisions in longitude (on the ring) and 6 layers in height
- 3 hours batch Kalman filtering

Results for UTC 12h-15h are shown in Figure 4.9. Note that the equatorial anomaly can be seen as having two peaks of electron density. From this study it becomes clear that, in order

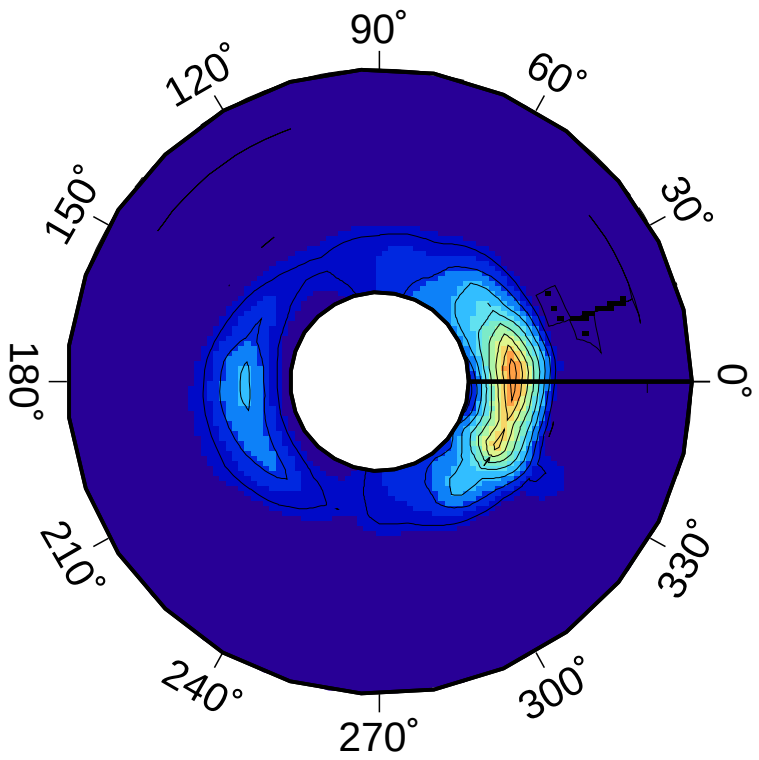

Figure 4.9: Ring solution for the UTC 12h - 15h: the radius is the height of the ionosphere: from the 100 km to 1300 km and the azimuth angles are the longitudes on the orbit plane, starting at the Sun's right ascension.

to have a good 4D map of the ionosphere, a constellation of sounding satellites is needed (such as the future mission COSMIC, http://cosmic.cosmic.ucar.edu for details).

4.4 Impact of the protonosphere: simulations using PIM

What is the impact of not modeling the protonospheric electron distribution? We can estimate this source of error using simulated data. The Parametrized Ionospheric Model

version 1.6 ([64], [65]) is used together with the actual geometric transmitter-receiver ray information for the same day. A full grid extending to the GPS orbit height is used to calculate the simulated delays, and the bias constants are set to zero. The question is then how accurate would the tomographic reconstruction be using a grid up to 700 km height with and without one layer above. As can be seen in Figure 4.10, the reconstruction up to 700 km height is quite accurate when including the layer above, since it absorbs the non uniform part of the protonosphere. As expected from previous work ([66]), the uniform section is absorbed by the bias constants. We therefore conclude that, in order to obtain accurate multi-layer representations of the ionosphere, the effect of the protonosphere has to be considered using an extra layer.

4.5 Ionospheric Calibration of Radar Altimeters

There are two main applications of GPS-based ionospheric tomography: a) to provide ionospheric maps on a near-real time basis, b) to provide a means for absolute ionospheric calibration of on-board instrumentation. The first application is readily seen and by introducing some modeling information into the system, data from a smaller number of ground stations may be used. Nowadays, Total Electron Content (TEC) is obtained from GPS measurements in real-time basis at different centers (some using techniques other than voxel-based tomography), such as the Jet Propulsion Laboratory where 15 minute updates of the state of the ionosphere are continuously produced [67]. The spatial resolution is poor because of the sparse set of data being used.

The second application reveals to be more powerful. In the voxel-based 4D tomographic direct approach, uncalibrated GPS data are ingested and the values of the instrumental biases of both the transmitters and the receivers are obtained as part of the solution together with the solution field; therefore we are considering the correlation between parameters and we obtain an absolute calibration of the system. Space-based instruments using two frequencies to cancel out the ionospheric extra delay may suffer from a bias difficult to calibrate for. A robust system such as tomography using voxels as the basis functions can efficiently combine data from different sources and consider the calibration constants as unknowns. Dual frequency radar altimeter data can be used to enhance the tomographic solution and to test this calibrating capability.

The NASA-CNES* TOPEX/POSEIDON experiment has been a very successful mission of radar altimetry (RA) (see [68]). The radar altimeter data comes from pulses operated at 13.6

*NASA, National Aeronautics and Space Administration, http://www.nasa.gov; CNES, Centre National d'Etudes Spatiales, http://www.cnes.fr

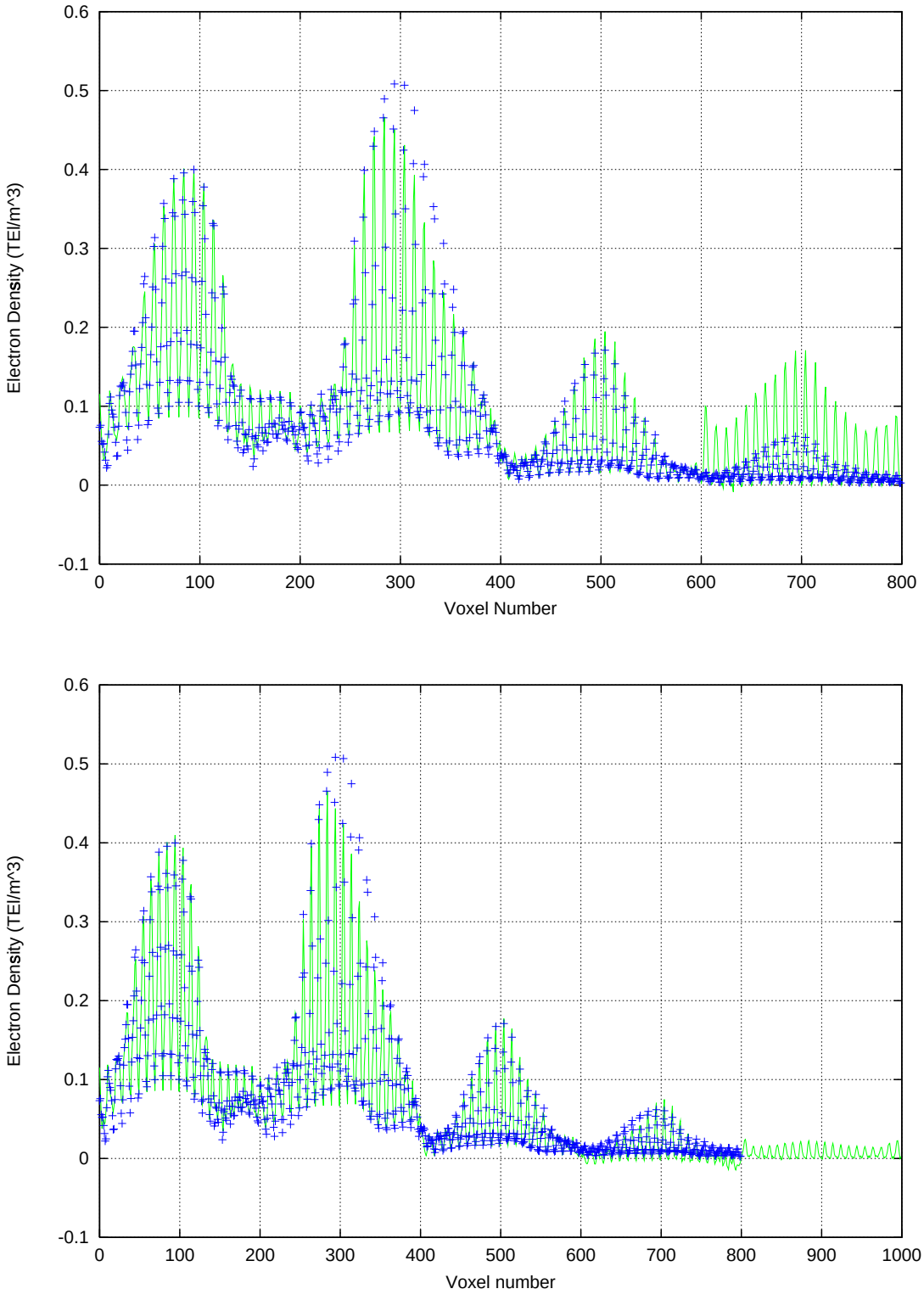

Figure 4.10: Simulation reconstruction (lines) and original field (crosses). The bias (up to voxel 800) and RMS in the reconstruction are 0.26 and 0.70 TECU, respectively. The bias and RMS for the layer above are 0.30 and 0.6 TECU.

GHz and they are corrected for the extra ionospheric delay with data at 5.3 GHz, again relying on the dispersive nature of the ionosphere. This extra delay, however, has to be periodically calibrated since the signals at both frequencies go through different instrumentation in the Radar Altimeter. The calibration can only be done with absolute maps. A natural way to do so is to ingest the ionospheric electron contents sensed by the RA into the tomographic solution and including a calibration constant in the equation. Radar altimeters, however, only sense the atmosphere vertically with no angle dependency and, thus, this calibration constant cannot be solved for using solely the RA ionospheric data, as in the case of the GPS data: when angle dependency is present, it is possible to discriminate the electron density field from constant terms affecting all the observations (instrumental biases, see [2]). In addition, the inclusion of RA data has to be balanced with the GPS data: the TOPEX/POSEIDON data are given at a rate of 1 s and have a much better spatial resolution (it scans a smaller area at a higher rate). If such high-rate very localized data are directly ingested to the linear system, the solution will be degraded and no possible calibration process will be effective. We must therefore downweight the RA data so that it does not dominate the ionospheric solution and the constant absorbs the inconsistency of the RA data with the GPS data. Although the calibration process could have been done a posteriori, once the ionospheric maps have been obtained from GPS data, ingesting the RA data as an input to the process has the advantage of considering the correlation between parameters and helps to obtain better solutions. Furthermore, the T/P satellite includes, for navigation purposes, a zenith looking GPS antenna which can be used as an extra GPS receiver to extract the protonosphere structure. The calibration of the T/P instrument using tomography was first presented in [69], where an a posteriori calibration process was followed. It was there claimed that GPS tomography provides ionospheric maps in an absolute manner. Later, the RA data was directly input to the system (see [70]) and the calibration constant thus obtained, matched previous results: in Table 4.2 we show the calibration constant results with the appropriate sigma value. We have listed the bias value and the rms of the data when comparing the radar altimeter data (NRA) with JPL 15 minutes map, the GPS tomography in the ring solution for 24 hours (RTGPS/24h), DORIS data using a Bent model (DORIS/Bent) (as provided by CNES with the T/P data, see [71]), the GPS tomography at global scale for 24 hours (TGPS/24h) and using a 3-hour Kalman filter (TGPS/3h), ionospheric maps from CODE (CODE/24h) and the GPS tomography solution considering one thin layer (TGPS/1layer). Finally, NRA data was averaged over 50^o along the orbit plane to resemble the voxel size in tomography and compared with the 24 hours TGPS. This last comparison yielded the best result of the a posteriori calibration. The ingestion of NRA data and zenith looking GPS data from T/P gave 2.98 TECU bias with a formal

error of 0.003 TECU, again reflecting that the a priori incorporation of the data is an optimal calibration procedure.

Fit	**Bias**	**RMS**
NRA-JPL/15min	0.7	2.5
NRA-RTGPS/24h	3.1	2.9
NRA-DORIS/Bent	4.1	3.0
NRA-TGPS/3h	3.4	3.2
NRA-TGPS/24h	3.2	3.8
NRA-CODE/24h	2.4	4.0
NRA-TGPS/1layer	1.0	4.5
NRA50^o-TGPS/24h	3.0	0.9
T(NRA+GPS)	2.98	$3 \cdot 10^{-3}$

Table 4.2: Fit results (figures are in TECU). The rms value in T(NRA+GPS), tomographic solution including NRA and GPS, is a formal error.

The ingestion of other sources of data (DORIS data, or any dual communications systems) shows a bright future for ionospheric tomography, since more and more data can be included in the solution while calibrating the systems. It is, in fact, an intermediate step to what it would be the ingestion of such radio data into a Numerical Space Weather Prediction Model, where the physics and dynamics of the system would be taken care of with finer detail, indeed leading to a multiple-source, absolute ionospheric monitoring system.