Document reference: IRM/GALOCAD/OUT310-1	Deliverable : OUT310-1	Date: 29 April 2008 Version: 2
Contract ref : GJU/06/2423/CTR/GALOCAD	Written by : I. Kutiev, P. Marinov, S. Fidanova	Verified by : S. Stankov

GALOCAD

Development of a Galileo Local Component for nowcasting and forecasting of atmospheric disturbances affecting the integrity of high precision Galileo applications.

WP 310 Technical Report :

"Correlation of ionospheric activity with other geophysical parameters"

Geophysical Institute, BAS, Acad. G. Bonchev str., bl. 3, Sofia 1113, Bulgaria

LIST OF ABBREVIATIONS

BDN	Belgian Dense Network
GI/BAS	The Geophysical Institute of the Bulgarian Academy of Sciences
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
LAT	Latitude
LONG	Longitude
LLT	Latitude Longitude Time (model)
LSQ	Least Square Fit
MSTID	Medium-Scale TID
RMI	Royal Meteorological Institute
ROB	Royal Observatory of Belgium
RMS	Root Mean Square
ROT	Rate of TEC
RTK	Real Time Kinematics
SSTID	Small-Scale Travelling Ionospheric Disturbance
TEC	Total Electron Content
TECU	TEC Unit (10^{16} electrons/m^2)
TAD	Travelling Atmospheric Disturbance
TID	Travelling Ionospheric Disturbance
WP	Work Package

TABLE OF CONTENT

1. Introduction

The ionospheric irregularity structures are known to propagate away from their areas of origin, driven by atmospheric density waves (Rawer, 1993). The moving ionospheric structures are known also as Travelling Ionospheric Disturbances (TIDs). TIDs are divided, according to their spectral parameters, into three main groups: large-scale TIDs with a wavelength more than 500 km and period of 0.5-3.0 hours and middle-scale TIDs (MSTIDs) with a wavelength of 50-500 km and periods 0.2-1.0 hours. The other group represents the smallest scale size TIDs (SSTIDs) with a wavelength less than 50 km and period of several minutes. MSTIDs have horizontal phase speed of 100-300 m/s and occur more frequently than LSTIDs. Their generation is not well understood, although many possible mechanisms have been proposed. All these mechanisms do not include geomagnetic activity as a primary driver, although some non-linear interactions with LSTIDs may also be assumed.

The great variety of sizes and periods makes identification and studies of irregularities effective in degrading GPS positioning accuracy difficult. The aim of WP 310 is to define the main characteristics of irregular TEC structures by approximating them with analytical functions, reveal their size and motion and then correlate these quantities with solar and geophysical parameters. The mathematical modeling is made possible through the use of the Belgian Dense Network (BDN) which consists of 67 permanent GPS stations. TEC data, obtained from all GPS satellites in view during a certain time window is approximated by 3-dimensional polynomial along latitude, longitude and time axes. Depending on the order, polynomial can capture disturbances with different size: low-order polynomials smooth out ionospheric disturbances, while higher order polinomials can capture localized TEC structures. The present approach subtracts low-order from higher-order approximations to localize disturbance structures and follow their movement across the area.

The present report summarizes activities and results obtained in completeing the WP tasks. Section 2 describes the way the database is compiled. Sections 3 provides the basic formulation of the so-called Latitude-Longitude-Time model (further denoted as the LLT model). Section 4 describes the main charactristics of the LLT model and its performance. Both, the long-term and short-term models are described in detail together with the evaluation of the model error, i.e. the standard deviation between the model results and the data. Section 5 provides the methodology of localizing ionospheric disturbances by using residuals from pairs of models with different degrees of polynomials and discusses the realibility of the results. Section 6 introduces a new method for spectral analysis of detected irregularities and shows how their spectral characteristics can be obtained. This question is further elaborated in Section 7 where the correlation of modeled disturbances and geomagnetic and solar activity indices is investigated. The last section 8 summarizes the results..

It has to be noted that the LLT model, localization technique and spectral analysis are original results, not yet published.

2. Compiling the database

The visual inspection of the lat/long plots is an essential part of the analysis. Figure 1 shows TEC data derived from GPS satellites in view between 11:00 and 12:00 UT on December 2, 2004 (Day 359). The magnitude of each TEC value is color coded and placed at coordinates where the respective signal ray path pierces the height of 400 km. The map area is framed in [-10°, 20°]E and [42°, 57°]N geographic scale. Further in the text, the "grad symbol" (°) will be omitted. The color scale is shown on the right, with 256 levels of spectra from blue to red. At least eight spots of data are seen in Figure 1, each representing a projection of BDN on ionospheric surface, seen from each of the eight satellites in view, and traced continuously during that hour. Although the data coverage is fragmented, the irregular structure is clearly seen. It is obvious that at any instant, a substantial part of the area does not contain data. To make a model over the data, we have first to divide the whole area into proper number of subareas (bins) and define the time period, during which enough data are accumulated in each bin. Then form time series of data in each bin. From the physical point of view, we need to detect irregularities as smaller in size, as possible, while the reliable modeling needs sufficient amount of data. We have to find a proper compromise between the two requirements: smaller bin size and statistical sufficiency of the data.

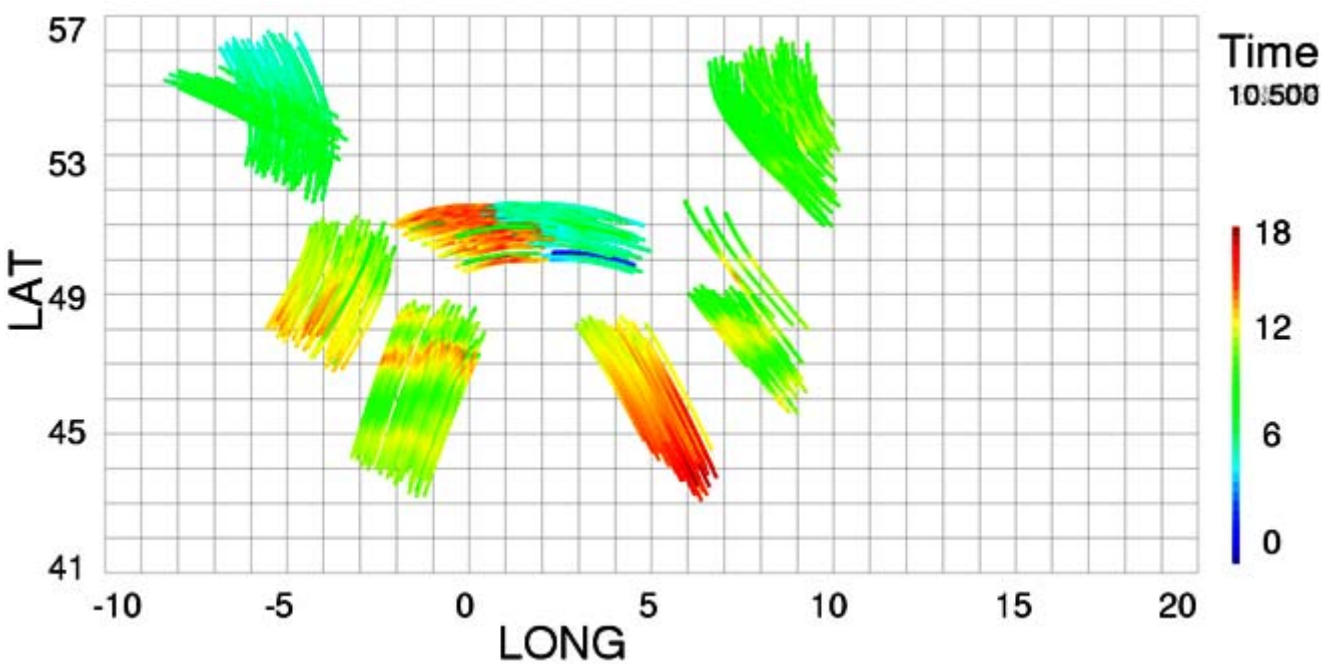

Figure 1: Data inferred from GPS satellites in view between 10:00 and 11:00 UT on December 24, 2004 (Day 359). TEC magnitudes (color-coded, units of $10^{16}/cm^2$) given at the ionospheric piercing points .

Figure 2 shows the number of TEC values (vertical bars) sorted in bins with size 2x2 and 1x1 accumulated within the same hour as in Figure 1. While in the larger (2x2) bins the number of accumulated data is roughly twice larger that in the smaller-size (1x1) bins, the area with non-empty bins remains almost the same. The non-empty bins are clustered in the central part of the mapped area.

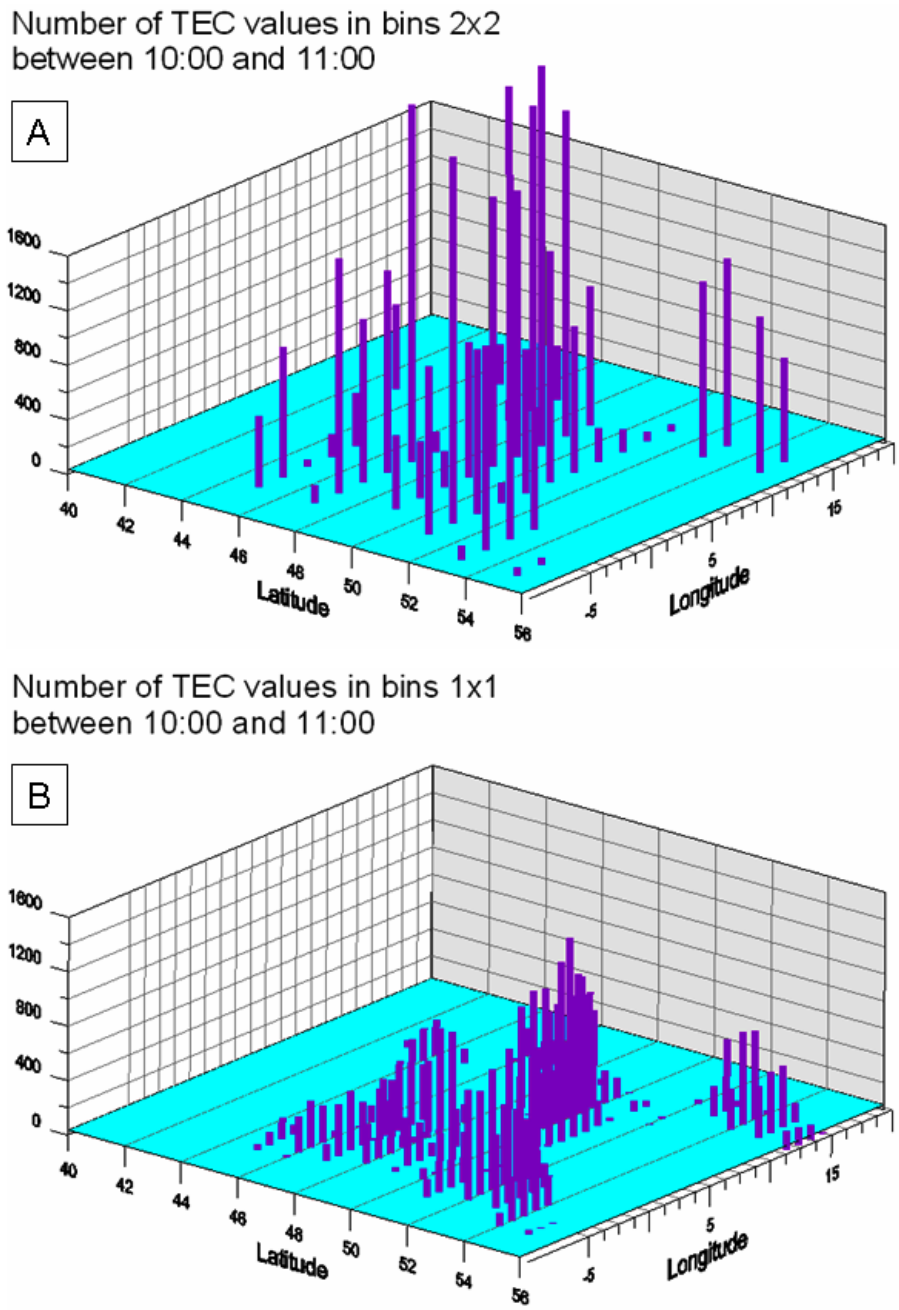

Figure 2: Number of TEC values (vertical bars) accumulated between 10:00 and 11:00 UT in each bin with size 2x2 (panel A) and size 1x1 (Panel B).

The following Figure 3 summarizes better the dynamics of filling the bins. The abscissa shows the percentage of non-empty bins (having more than 5 values) relative to the all 480 bins in the area considered. The upper plot shows the average filling rate from 10:00 UT, as a starting moment for accumulation, until 11:00 UT. At first 30 sec filled are 16% of larger bins and only 6% of the smaller bins. After an hour, filled are 45% and 32% respectively. The lower plot gives the same dynamics augmented in first 6 min, from 10.0 to 10.1. It is seen that the filling rate changes at a time lag of 6 min (0.1 hour). These estimates are roughly valid for the whole day 359 of 2004.

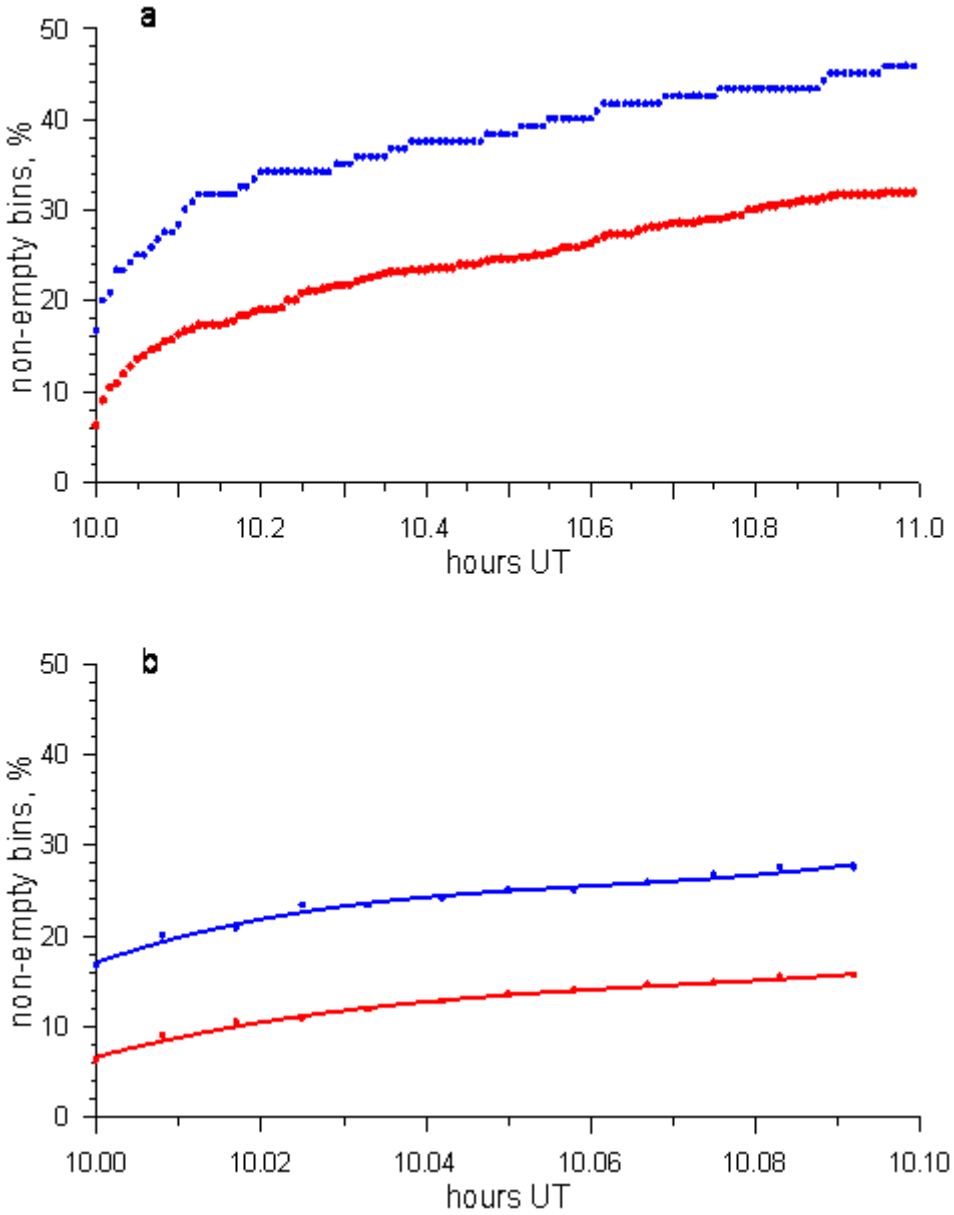

Figure 3: Percentage of non-empty bins relative to the all bins with size 2x2 (blue line) and 1x1 (red line). Accumulation starts accounting from 10.00 UT. Above: filling dynamics during the whole hour; below: filling in the first 6 min.

Based on the results presented in Fig. and Fig.3, we can optimize the database collection by significantly reducing the map area down to the frame [46, 52]N and [-1, 10]E, redefining the time accumulation (time step) to 6 min (0.1 hour), and selecting the smaller bin size 1x1. Thus we obtained that for the reduced area the rate of filling curves increases around 3 times, in proportion to those of Fig.3. For the smaller bin size, at 10:00 UT the filling starts from 20%, after 6 min reaches 40% and at the end of the hour approaches 67%. The reduced area significantly improves the availability of data, but still, a large part of the latitude/longitude/time space remains uncovered. The lack of sufficient data in the bins makes the task of modeling MSTID dynamics rather difficult. However, we avoid this disadvantage by introducing a special procedure of artificially filling the empty bins. In our modeling approach the irregular ionospheric structures are presented by the residuals from a pair of model approximations using polynomials of different degrees. Subtraction of the model pairs eliminates the effect of artificially filled bins; irregular structures are revealed by the measured data only.

3. LLT model

3.1 Basic formulation

As mentioned in Introduction, the present approach of studying MSTIDs requires approximation of acquired TEC data from BDN by analytical functions along the latitude, longitude and time. For convenience, we denote the approximation algorithm as the Latitude-Longitude-Time (LLT) model, although it is not yet a full-featured model. We use an analytic representation of the data as a function of 3 variables (LAT, LONG, TIME) = (x1, x2, t); where x1 is in the range [46, 52], x2 is in the range [-1, 11] and t is in the range [0, 24]. We denote the number of coefficients related to the variables x1, x2, and t as N1, N2 and N3 respectively. The analytic representation of the data is a polynomial of the following type:

$$F(x1,x2,x3;C)=\sum_{i1=1}^{N1}\sum_{i2=1}^{N2}\sum_{i3=1}^{N3}c(i1,i2,i3)B1(i1,x1)B2(i2,x2)B3(i3,t)$$

In order to express all functions by one variable u, we make following substitutions:

$$u1=-1+\frac{x1-46}{52-46}2=-1+\frac{x1-46}{3}$$
$$u2=-1+\frac{x2+1}{12}2=-1+\frac{x2+1}{6}$$
$$u3=\frac{(t-T_0)}{T}2\pi \quad for\ T=24h\ and$$
$$u3=\frac{(t-T_0)}{T}\pi \quad for\ T<24h$$

The time variations are approximated on a trigonometric basis:
$\{B3(i3,t)\}=\{1,\sin(u3),\cos(u3),\sin(2.u3),\cos(2.u3),\ldots\}$ for the standard-period (T=24h) model and
$\{B3(i3,t)\}=\{1,\cos(u3),\cos(2.u3),\ldots\}$ for the short-period (T<24h) model, T<24h.

The variables x1 and x2 are approximated by Tchebishev base functions:
$\{B1(i1,x1)\}=\{T_0(u1)=1,T_1(u1)=u1,\ldots,T_k(u1)=2.u1T_{k-1}(u1)-T_{k-2}(u1),\ldots\}$, i.e.
$T_k(u1)=\cos(k.\arccos(u1))$

B2(i2,x2) is obtained similarly to B1 by replacing i1, x1 and u1 with i2, x2, u2 respectively.

$C=\{c(i1,i2,i3)\mid i1=1,\ldots N1;i2=1,\ldots N2;i3=1,\ldots N3\}$.

The solution of the LSQ approximation is obtained by minimizing

$$\sum_{k=1}^{N}(f(k)-F(x1(k),x2(k),t(k);C))^2 ,$$

where $\{x1(k), x2(k), t(k), f(k)\}_{k=1}^{N}$ are data which we approximate. In domains with large gradients of the data, the Gibbs effect takes place (i.e. close to the area of the jump of the data the approximation achieves values greater than the maximal measured data and less than minimal one). To avoid this unacceptable effect, instead of approximating the function f(k), we approximate the function g(k)=lg(f(k)). In this case the approximated function is G(k) and $F(k) = 10^{G(k)}$.

3.2 Filling the empty bins

To constrain the approximation in acceptable limits, we assign values to the empty bins by using the following procedure. We average the measured TEC values in each non-empty bin. Every bin, except those at the boundaries, has 26 neighbor bins. In the first round, we choose the empty bins which have at least 20 non-empty neighbors and assign to each of them values, being average of all neighbor's averages. Considering the newly filled bins non-empty, procedure repeats the filling until no empty bins exist with 20 non-empty neighbors. In the second round, we consider those empty bins having 19 non-empty bins and each next round procedure reduces the number of required non-empty bins by one. Practically, 5-6 rounds are enough to fill all gaps in the 24 hours time range. We found that it is better to use both individual data and bin averages in the fitting procedure, thus giving some preference to the average values in the bins. This preference depends on the time range T and the number of coefficients N1 and N2. We found that for T=24 and the number of spatial coefficients 3 to 5, the best result was obtained (lowest standard deviation of model from data) when a weight of 66 is assigned to the average values. So, the best combination is to fit the model over the measured data and weighted average in the bins.

4. Model performance

We show LLT model performance for TEC data acquired on 24 December 2004 (day 359). We varied the values of N1, N2 and N3 and assessed the performance of the model by the root mean square (RMS) deviation of model from the data. The model approximation with a defined set of coefficients is denoted as "model (N1, N2, N3)", for example, "model (3, 3, 3)" denotes a model approximation with N1=3, N2=3, and N3=3 (time coefficients can be greater than 10). Approximating the time variations, we use two different time ranges: the whole day (T=24 hours) and short range (T is less than 2 hours). Both ranges have their own applications and will be described separately.

4.1 Approximations with time range T=24 hours

The top panel of Fig.4 is a model (5, 5, 21) representation of the data shown in Fig.1 in 1°x1° spatial grid size. The model presentation looks better when the model is shown in smaller grids, as in the bottom panel of Fig.4. Here the grid size is 0.2° x 0.3°. Two structures are well localized: a maximum at 49°N and 9.5°E and a minimum at 50.5°N and 5°E. It is seen that the minimum is visible in the data from Fig.1, while the maximum is found at a place where there are no data in this time interval. In this case, the model values are determined by the data from other time intervals.

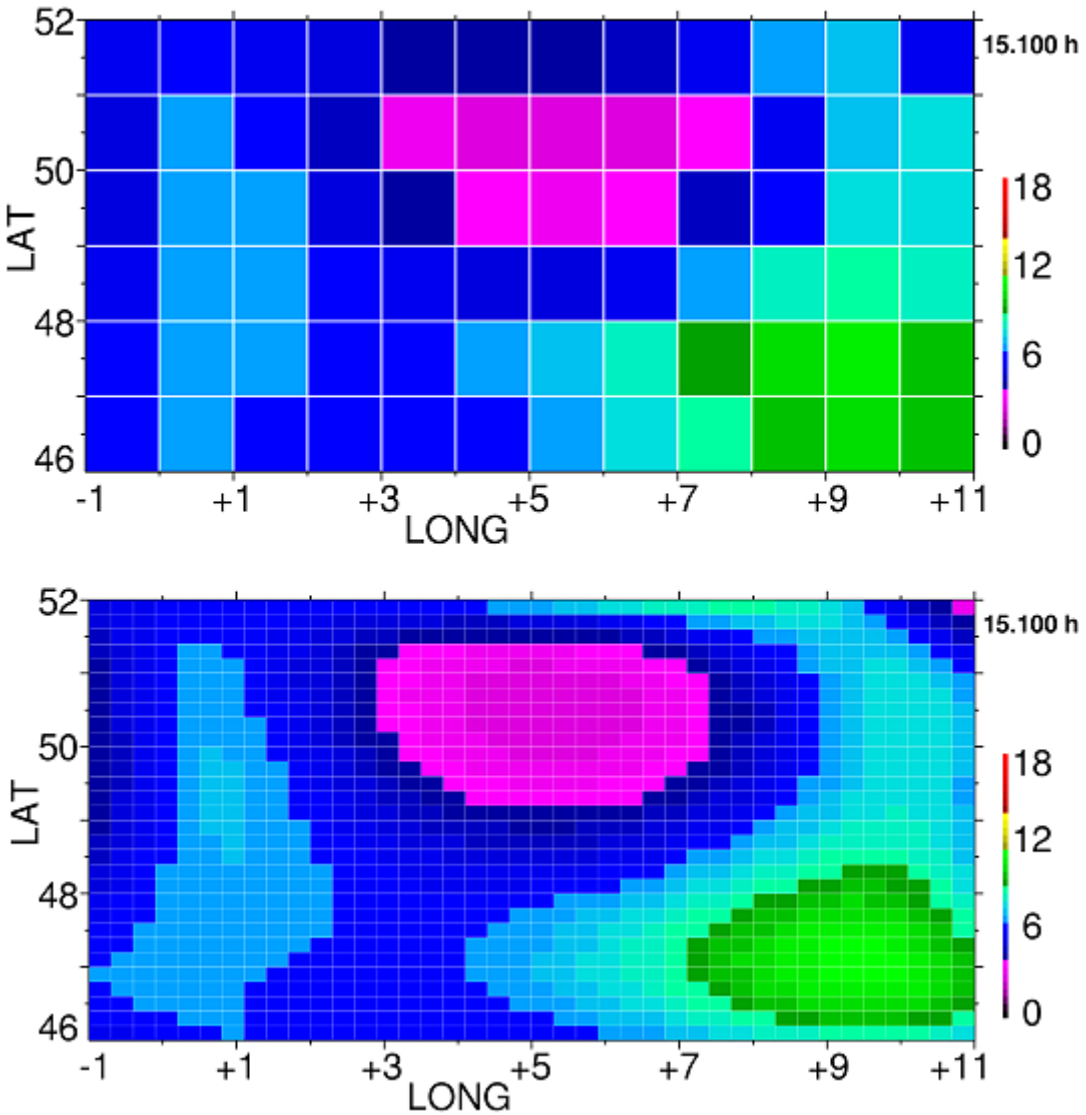

Figure 4: Model (5, 5 ,21) representation of the data shown in Fig. 1, using bin size 1ºx1º (top panel) and bin size 0.2º x 0.3º (bottom panel).

It is expected that the approximation of time variations in the range of T=24 is most important for the model performance. To give an idea how models with increasing N3 approaches the data, we show (Fig.5) the TEC data taken from the bin [49, 50]N and [3, 4]E as a function of time. Fig.5A represents the model (5, 5, 11) and Fig.5B represents the model (5, 5, 21). The upper panels of Fig.5A and Fig.5B show model approximation (red squares) and data (blue dots) during the whole day (T=24 hours). The lower panels show respective deviations of model values from the data (TEC error). Multiple curves represent the model every 0.2º inside the bin. Note, that the spread of model curves in not constant during the day, it depends on data configuration. This is an effect of small-scale structures in the data, not visible in the figure. The difference between the two approximations is well seen: the model (5,5,21) approximates better data variations.

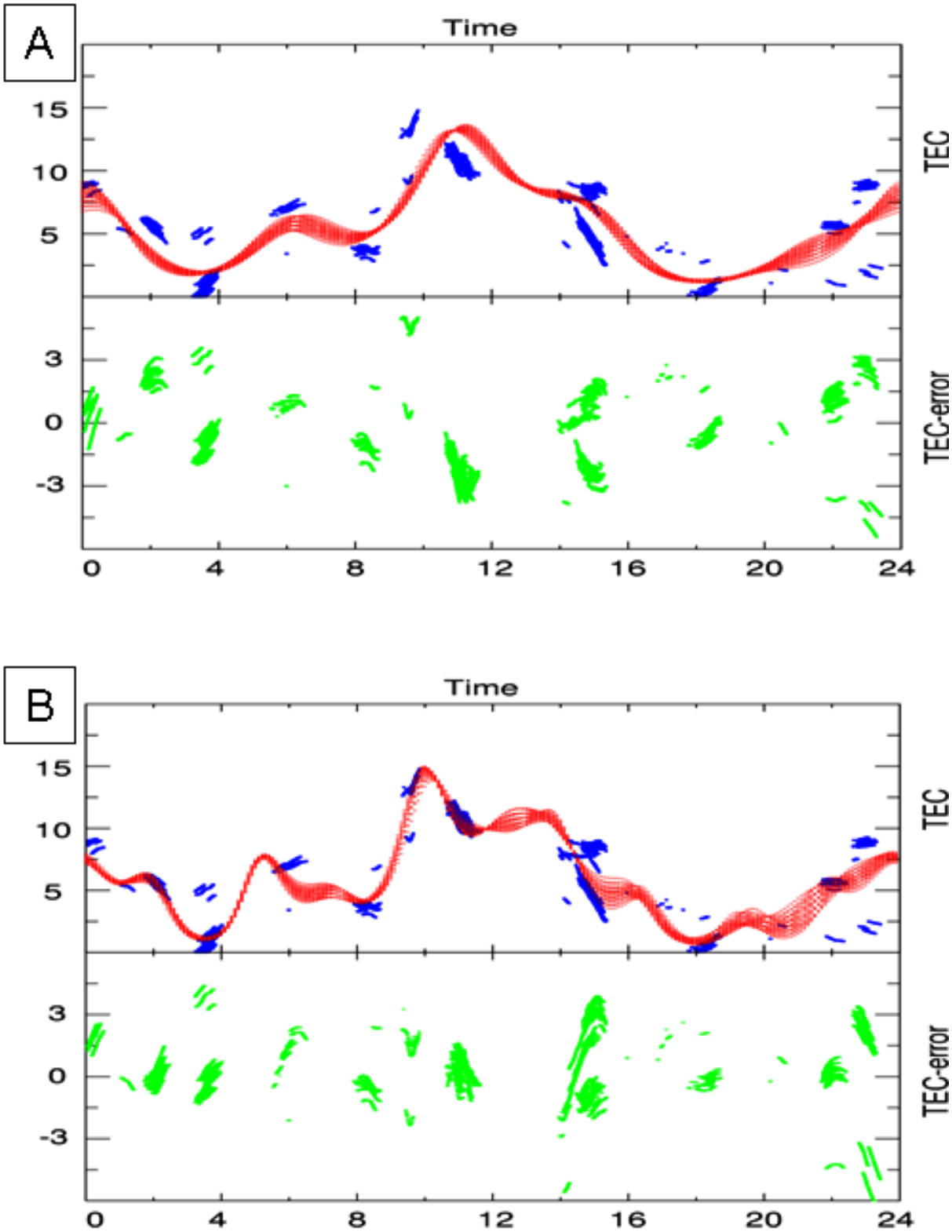

Figure 5: (Fig.5A): TEC data acquired the bin [49, 50]N and [3, 4]E during the whole day 359 of year 2004 as approximated by a model (5, 5, 11). Multiple curves are obtained by shifting Lat and Long by 0.2º inside the bin. Bottom panel: The model deviations from the data (TEC error) in TECU are given in green Color. (Fig.5B): the same as in Fig.5A but using the model (5,5,21).

4.2 Approximations with shorter time ranges T<24h

As was pointed out above, approximations over the whole time range of 24 hours require higher values of N3 and consequently, more computer resources. For MSTID studies in particular, the time range could be much smaller, but larger than 20 min (upper period limit). We run the model with two time ranges: T=48 min and T=96 min. To cover the whole day, the model was prepared to slide the time range starting from the beginning of the day, with a time shift of 6 min. The model value at any given moment is composed from the values of sequential short-range models, covering the given moment. For example, for models with T=96 min, the model value at 10:02 hours is composed by respective values from short-range models (08:30-10:02), (08:36-10:12),….(10:00-11:36), or totally from 16 models. The different values are weighted by a factor, reversely proportional to the time difference between their centers. Therefore, the model value at the moment 10:02 is an average of all weighted contributions. The same example, but for the time range T=48 min, the number of models, including the moment 10:02 are 8 (from 09:18 to 10:00).

Similarly to Figure 5, we represent in Figure 6a the model approximation (3, 3, 3) with T=96 min and time shift 12 min. The blue curve is the average of weighted individual model runs (red dots) and time shift 12 min. For better visualization, Figure 6b augments the model curves of Figure 6a in the time frame 10-12 UT. The individual short-range models, with several exceptions between 11:00 and 11:30, form a distinctive band around the main course of TEC variation. Note that the shorter T does not require high N3; here N3=3.

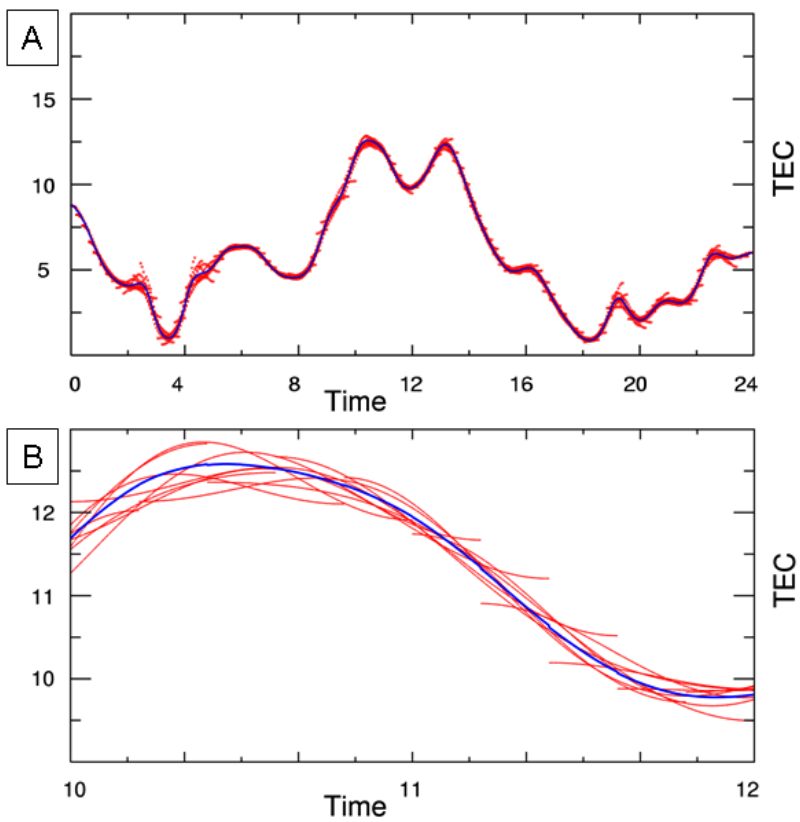

Figure 6: Panel A: Small dots represent individual model runs for approximation (3,3,03), with time range T= 96 min and time shift of 12 min. Blue curve is the weighted average from model runs in the range. Panel B: the same as in Panel A but the model runs in Fig. 8a are augmented in the time frame 10-12 UT

Increasing coefficients N1 and N2 also yields better approximations. Figure 7 represents model curves (red lines) composed by short-range (T=96 min) models (3,3, 05) in the upper panel and (5,5,05) in the bottom panel, for the same bin [49,50]N and [3,4]E, as above. For comparison, the green dots show the data and the step-like curve represents the average values (including the filled in) in the bins. Note, that the average models do not follow closely bin averages. It is clear that the model approximation (5,5,05) is closer to the data than approximation (3,3,05). We have found that the model performance is improved, when we increase N1, N2, along with N3, as they depend on each other.

We found that better approximation of the data is obtained with T=96 than T=48. The latter range is too short and the data are not enough for a reliable approximation. The larger time range, from one side, requires larger number of coefficient, but shorter ranges, from the other, do not contain enough data. Running the models with different time range (48 min, 96 min, 24 hours) we concluded that T=96 min is an optimum between these requirements.

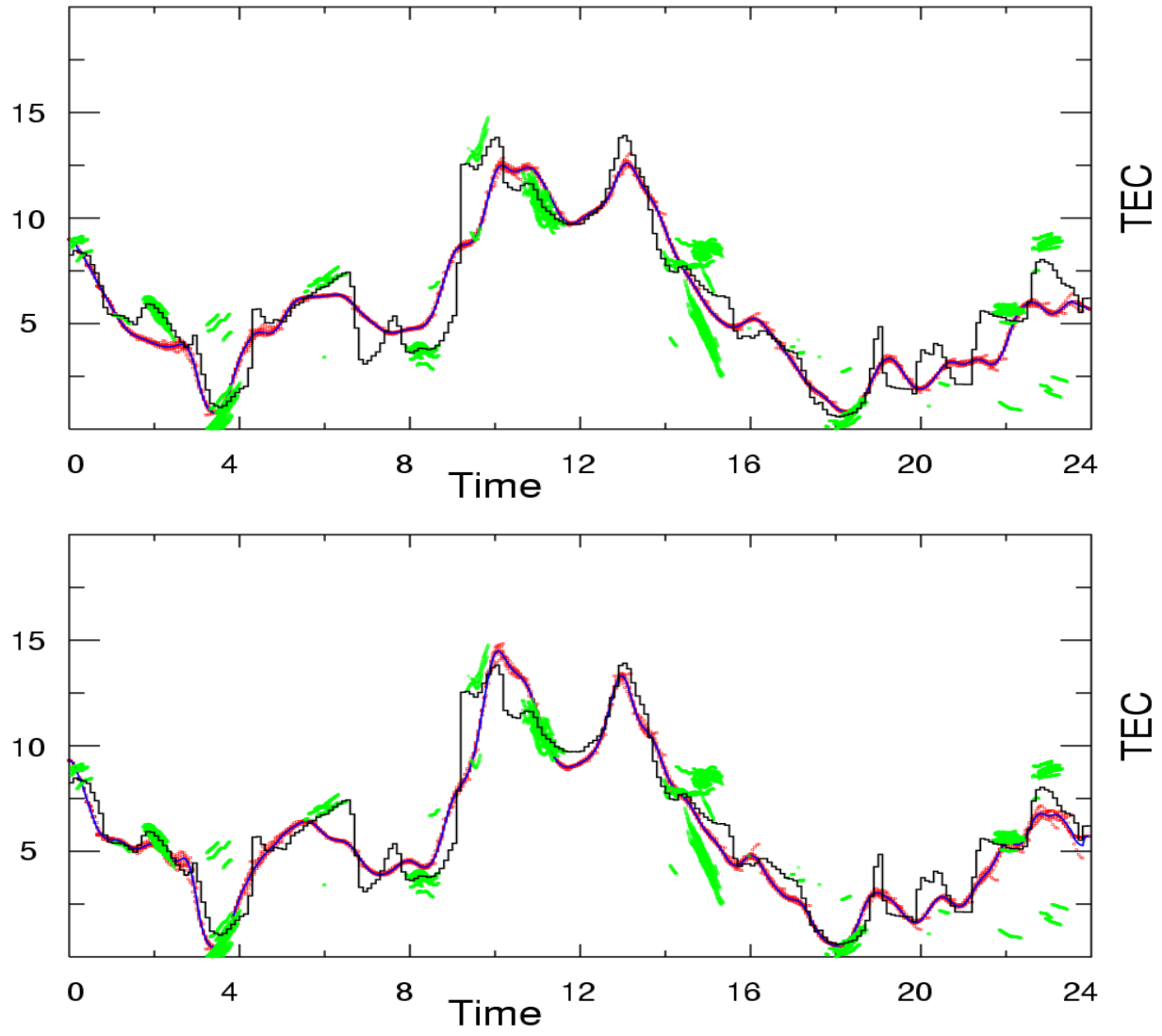

Figure 7: Data (green dots), bin averaged (step-like curve), short-range model (red dots) and weighted model curve (blue curve) fir the bin [49,50]N and [3,4]E. Top panel:. model approximation (3,3,05); Bottom panel: model approximation (5,5,05).

4.3 Error assessment

We used RMS (Root Mean Square) error as an estimate for assessing the model performance. RMS error, calculated over the whole day 359 of 2004, varies between 1 and 2 TEC units (TECU). In long models (T=24), RMS error varies between 1.4 for N3=25 and 1.8 for N3=13. Increasing or decreasing N1 and N2 yields a reverse change of RMS error. For short models, RMS error is more sensitive to N1 and N2, than to N3. For models with N1=N2=3, RMS error is around 1.5, while it decreases to 1.2 for N1=N2=5. It looks surprising that RMS error of the short models is less sensitive to N3 than to N1 and N2. Indeed, for a fixed number of spatial coefficients N1 and N2, RMS error slightly decreases with increasing N3 and this decrease is stronger for lower order spatial approximation. Figure 8 shows RMS error calculated for the bin [49, 50]N and [3, 4]E in the whole day 359 of 2004. In the upper plot, the RMS error of model (5, 5, 3) with T=48 min (red curve) and T=96 min (blue curve) is shown. Green curve represents the error of the long-range model (5, 5, 13) with T=24. The number of data points is given for reference by the black curve; the number of points is divided by 1000 to use the same scale in TECU. Both short-range models have practically the same error; the spike increase before 04:00 UT is due to the fact that TEC, obtained from two GPS satellites at same place and time, have quite different values (see Fig.7). The long-range model (5, 5, 13) has much higher error, especially daytime, although the time approximation uses 13 coefficients. This is another example of the advantages that short-term models inherit. The lower plot compares RMS error of three long-term models. The green curve represents the same model as above, the red and green curves represent models (5, 5, 25) and (5, 5, 31) respectively. It is clear, than a long-term model needs 31 coefficients of time approximation to reduce the error in the level that a short-range model with T=96 achieves. We can conclude that RMS error of LLT approximations is comparable (or even less) that the error of obtaining TEC from the GPS signals.

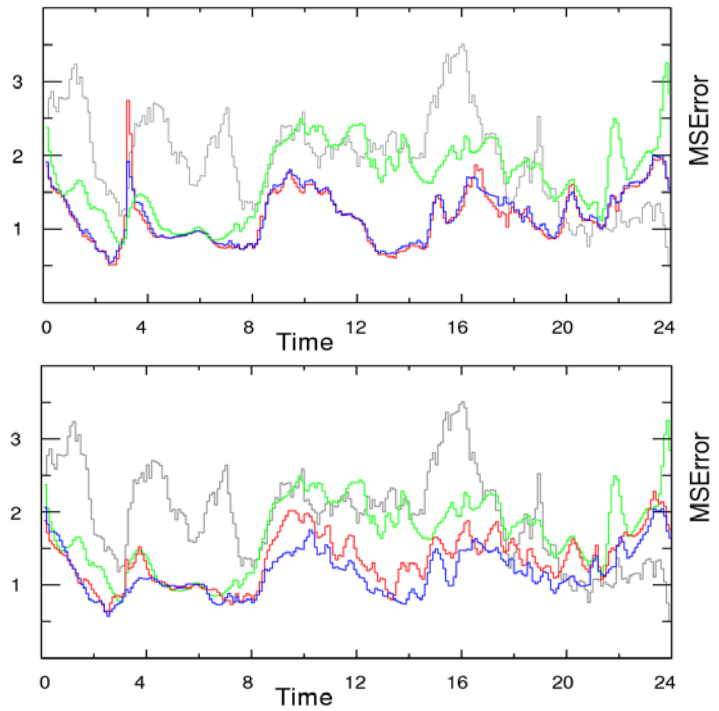

Figure 8: RMS error (color curves) in TECU calculated for the bin [49, 50]N and [3, 4]E on day 359 of 2004. The black curve refers to the same scale and show the number of data points divided by 1000. The upper plot shows RMS error of model (5, 5, 3) with T=48 min (red) and T=96 min (blue) as well as the model (5, 5, 13) with T=24 hours (green). The lower plot represents RMS error of models with T=24 hours: (5, 5, 25) in red, (5, 8, 13) in green, and (5, 8, 31) in blue.

5. Localizing the TEC disturbances

The key element of the proposed approach is the use of the LLT model for localization of TEC disturbances. As was shown above, higher order polynomials (larger number of coefficients) capture more details of the irregular structure of TEC, than the lower order polynomials. We subtract a pair of short time range models with slightly different number of coefficients and assume that the residuals represent local disturbances.

We consider the following criteria for specifying N1, N2, and N3 in order to capture smaller TEC structures. If X1 is the latitude range (46 to 52 degree), X2 is the range of longitudes (-1 to 11), and T is time range, the smallest size structures (threshold) which can be detected along each axis is defined by X1/(N1-1), X2/(N2-1), and T/(N3-1). These criteria are based on the fact, that a polynomial of k-degree can capture k-1 extremes (maximums and minimums). Subtracting models with different number of coefficients, we actually separate structures with size equal to the difference between both thresholds. For example, considering models (3, 3, 5) and (3, 3, 3), we see that the first model can localize structures with size S larger than 3°x 6° and duration D more than 24 min (as X1=6, X2=12 and T=96 min), while the second model localizes structures with the same size S, but with duration D longer than 48 min. When subtracting both models, we localize a structure with a size, exceeding 3°x 6° (S>3°x 6°) and duration between 24 min and 48 min (24<D<48). By varying N3, we can capture structures with apriori specified duration range. By analogy, we can subtract models with different N1 and N2 and localize structures with specified size range. Residuals, when subtracting the models (5, 5, 3) and (3, 3, 3), localize structures with size of 1.5°x 3°< S <3°x 6°. Duration, however, is not so strictly confined to particular time limits: all structures lasting longer than 48 min will be captured (D>48).

Figure 9 visualizes the difference between four pairs of short time range models (T=96 min) in the latitude/longitude frame from 17.5 to 18.1 UT. Each pair is shown in separate columns, as the time is increasing downwards. Comparing plots formed by residuals of the four pairs at any moment (given line), we see that the right (forth) column contains most detailed structure, which fade to the left. The time sequence clearly shows movement and development of well-defined TEC structures.

The first and second columns show the difference between models with the same spatial approximation. Any structure appearing in these columns are due to the difference in time approximation. The spatial approximation modulates only the magnitude of the structures, defined by the temporal approximation. Because the temporal approximation has trigonometric basis, N3=3 means approximation with two waves: the basic, with period of 96 min and its first harmonic, with period of 48 min. Similarly, N3=5 means four waves: the basic, with period of 96 min and its three harmonics, with period of 48, 32, and 24 min, respectively. Third and forth columns represent the difference between models, having the same temporal approximation, but the spatial surface is described by polynomials of 2nd and 4th degree. Figure 9 definitely shows that the difference of spatial approximation (columns 3 and 4) provides more detailed structure than the difference in temporal approximation (columns 1 and 2). The model error can easily explain the fact that the RMS difference between spatial approximations is larger than the RMS difference between temporal approximations with the same N1 and N2 coefficients. It is reasonable to assume that both, the spatial and temporal structures are coupled and characterize the size and duration of real disturbances.

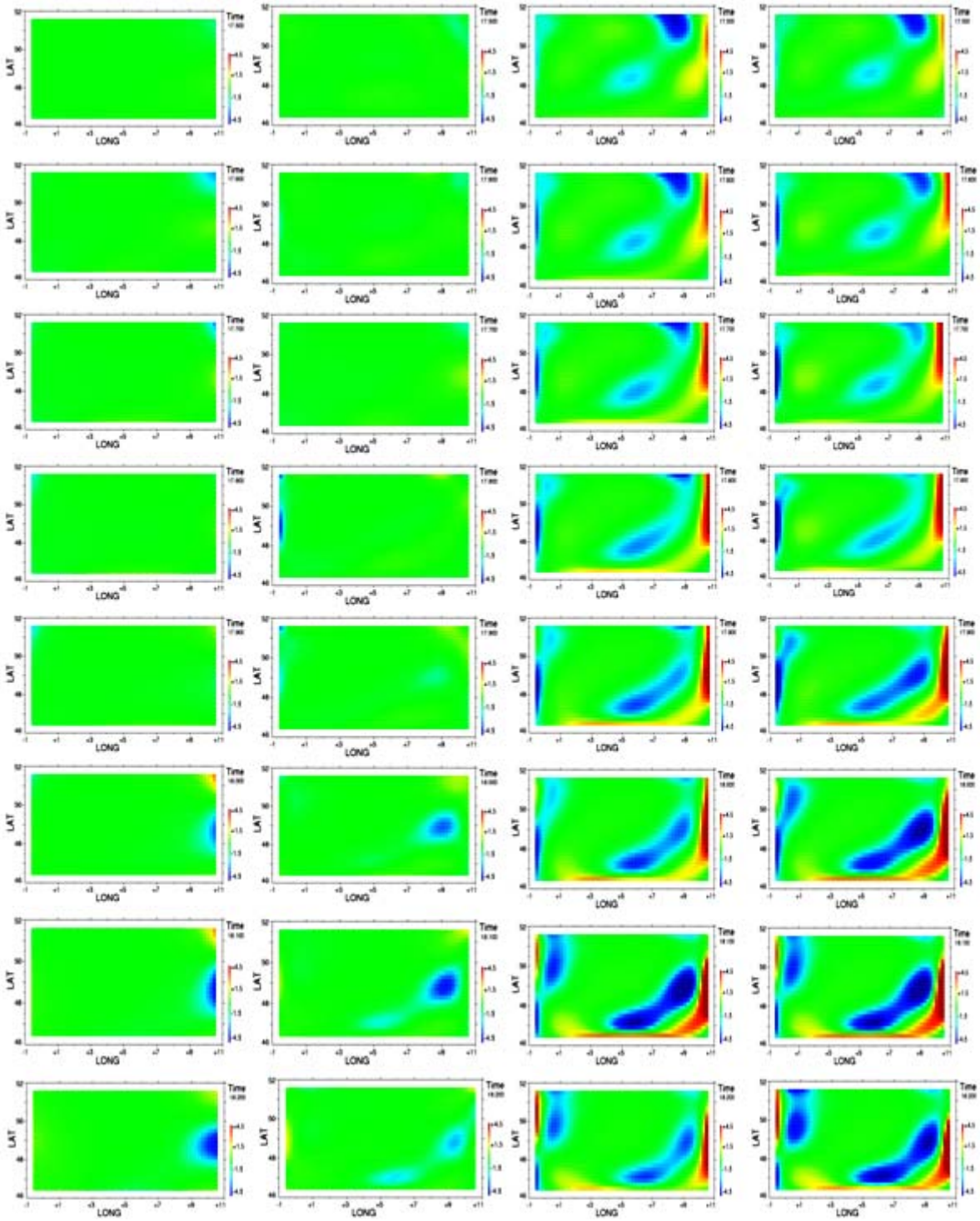

Figure 9: Differences between four pairs of short time range models (T=96 min) in the latitude/longitude frame from 17.5 to 18.1 UT. Model pair, size S, and duration D are given in the following Table 1.

Table 1.

column	*model pair*	*size S*	*duration D*
1	(3,3,3)-(3,3,5)	S> 3°x 6°	24<D<48
2	(5,5,3)-(5,5,5)	S> 1.5°x 3°	24<D<48
3	(3,3,3)-(5,5,3)	1.5°x 3°<S<3°x 6°	D>48
4	(3,3,5)-(5,5,5)	1.5°x 3°<S<3°x 6°	D>24

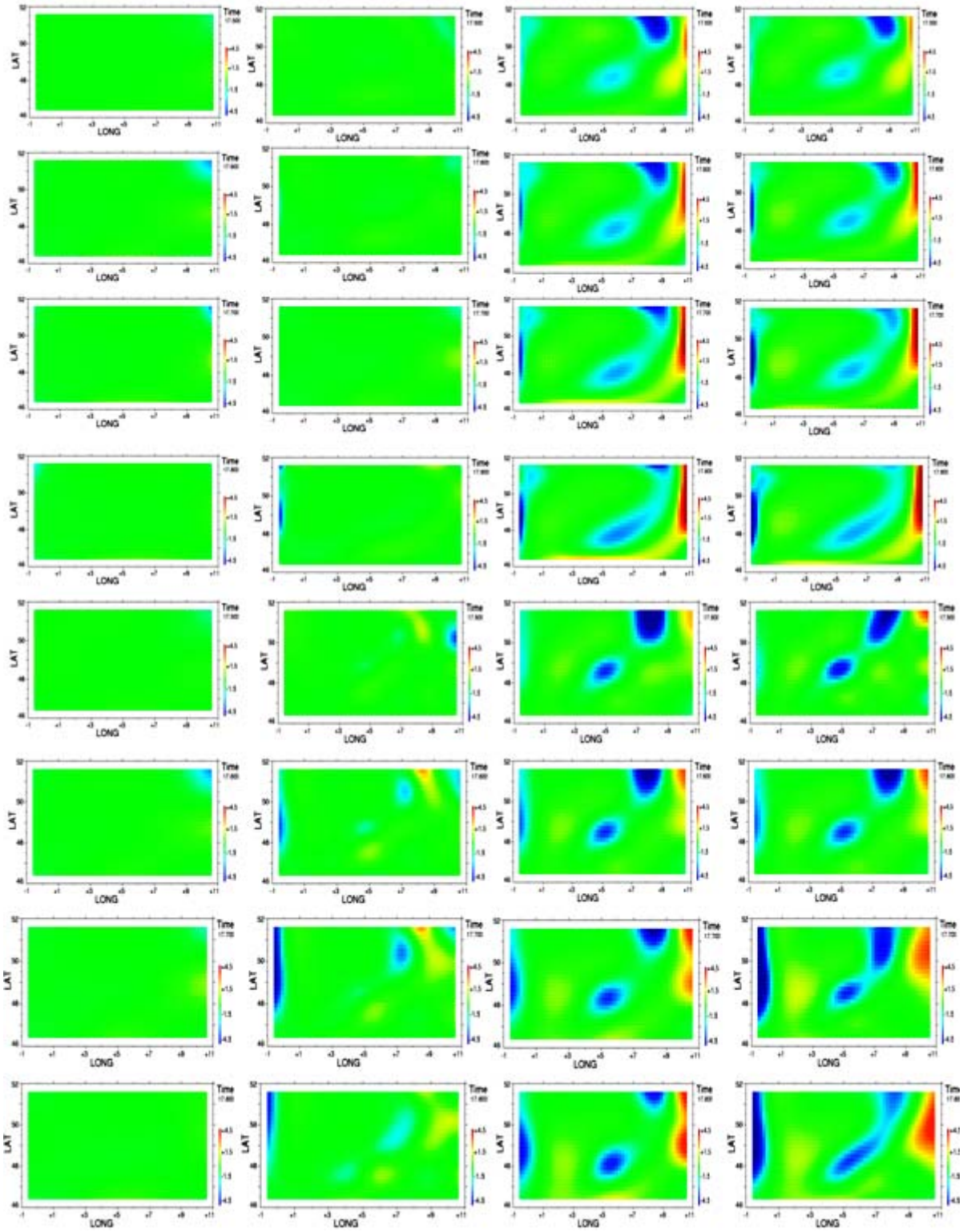

Figure 10: see Table 2 and description in the text

In the Figure 10, we compare first four lines of Figure 9 (17.5 to 17.8 UT) with another four pairs of models with larger difference of coefficients, taken for the same time interval. Model pair, size and duration for the lower part of Figure 10 are given in Table 2. By comparing the respective plots from the upper and lower parts of the figure, we see that the structures remain the same; in the lower part their magnitude is enlarged in both (positive and negative) directions. The larger differences in coefficients yield larger magnitude of disturbed structures. The graphic software marks with blank spots the magnitudes exceeding the specified range on the right of each plot. It has to be noted that the larger magnitude range decreases the color resolution, so we preferred to keep the resolution the same for all plots in the figures. The main conclusion that can be drawn is that localization is sustainable independently of the different approximations. The time development of the structures is physically reasonable.

The plots in the first two columns on the left-hand side of Fig.9 and Fig.10 show fewer details in comparison with the last two columns on the right-hand side. We can explain this fact with the constrained duration (12<D<48) imposed on the localized structures. On the contrary, the limited size in the other two columns does not filter out most of the structures. It is reasonable to assume that the characteristic duration is longer that 48 min while the specified size limits allow certain class of disturbances to be revealed. More work is needed to define the proper time and size ranges of main disturbances, and the results obtained so far are good basis for further development.

Table 2: legend for the lower 4 rows of Figure 10

column	*model pair*	*size S*	*duration D*
1	(3,3,3)-(3,3,9)	S> 3°x 6°	12<D<48
2	(5,7,3)-(5,7,9)	S> 1.5°x 2°	12<D<48
3	(3,3,3)-(5,7,9)	1.5°x 2°<S<3°x 6°	D>48
4	(3,3,9)-(5,7,9)	1.5°x 2°<S<3°x 6°	D>12

One important question is whether the residuals represent the real disturbed structure of TEC or not. It might be expected that higher order polynomials can form false extremes or large gradients at the map borders. Our experience, based on numerous runs of the model suggests that this is not very probable. From general considerations we conclude that the model can express extremes over the data availability areas only. In the areas without data, the bin averages form smoothed distribution and both, lower and higher order models would produce the same fit. That means that their residuals would be of negligible magnitude. Because the averages weight 66 times more than individual data, the model can form an extreme if the number of data points in the bin considerably exceeds 66. Another indirect evidence of capturing real disturbances comes from the time development of localized structures. Fig.8 and Fig.9 show that localized structures are sustainable and develop with time. Their behavior looks physically meaningful.

6. Spectral analysis

The LLT model offers a unique possibility to study the spectral characteristics of the observed TEC structures. Once the measured TEC are approximated by a model with a given set of coefficients, we can use DFT (Discrete Fourier Transform) method to obtain the power spectra as a function of time in each individual LAT/LONG bin. In the present analysis, the DFT algorithm and the power spectra is presented by Fourier coefficients obtained within a short time window. We slide the time window with a step *dt* and obtain coefficients along the time axis. In this way, by sliding the time window from 0 to 24 hours, we obtain diurnal variation of amplitude of each period considered. DFT method uses *sin* and *cos* functions to obtain amplitudes and phases of the basic wave and its harmonics. The DFT coefficients C_i are obtained as:

$$C_1 = \frac{1}{2} c_f \left(\sum_{n=1}^{N} z_n \right),$$

$$C_{2k} = c_f \left(\sum_{n=1}^{N} z_n \cos\left[\frac{2\pi}{N} k(n-1) \right] \right),$$

$$C_{2k+1} = c_f \left(\sum_{n=1}^{N} z_n \sin\left[\frac{2\pi}{N} k(n-1) \right] \right).$$

where N=2M+1is their number in the time window, C_k (k, n = 1,…M) are the coefficients of the DCT development, Z_n are the model TEC values; c_f is the normalization factor: $c_f = \sqrt{\frac{2}{N}}$. Wave amplitudes A_k and phases P_k are then obtained as:

$$A_k = \sqrt{C_{2k}^2 + C_{2k+1}^2}\text{, and } P_k = a\tan\left(\frac{C_{2k+1}}{C_{2k}} \right)$$

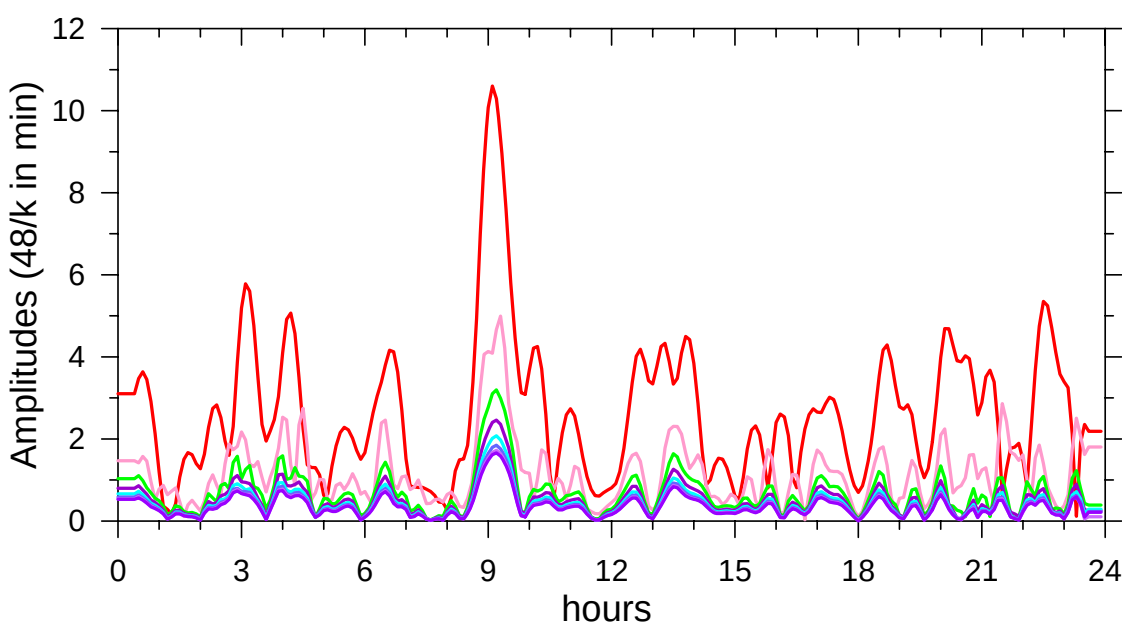

Figure 11: Amplitudes obtained by applying DFT to the model (5, 7, 9) with T=98 min. The basic wave (with period 48 min) and of its harmonics obtained in the bin [49, 50]N and [3, 4]E on day 359 of 2004. Red curve represents A_1, the other colored curves (from pink to violet) represent amplitude of 7 harmonics with increasing *k* number.

Fig.11 shows 8 amplitudes obtained by fitting DFT formulas to TEC values from the model (5, 7, 9) with time range T=96 min, in the bin [49, 50]N and [3, 4]E on day 359 of 2004. The time window for fitting DFT is 48 min, sliding with step of 12 min. Amplitudes and phases are assigned to the center of the time window. DFT consists of 8 waves, with a basic period of 48 min and its harmonics having periods of 48/k min. Figure 11 shows that the amplitude of the basic period is the largest and monotonically decay with the wave number k. It means that the main disturbances have equivalent periods larger than 48 min. Amplitudes vary with local time, expressing a marked maximum around 09 UT.

The lower panel of Fig.12 shows the power spectra as a wavelet. The ordinate shows the wave number, which represents periods in minutes as 48/k. The color scale on the right shows the magnitude of amplitudes in TECU. Vertical cross-section at any moment represents power spectra of the composing waves. The wavelet visualizes the dynamics of disturbances of TEC passing through the bin during the whole day. The upper panel reproduces the lower plot of Fig.7, in which the model (5, 5, 7) is shown by the red line, along with measured TEC (green dots) and the number of data points in units of 10^{16} cm^{-2} (black). There is no visible correlation between the location of measure TEC and power spectra. The latter is evidence of the reliability of present approach for spectral analysis.

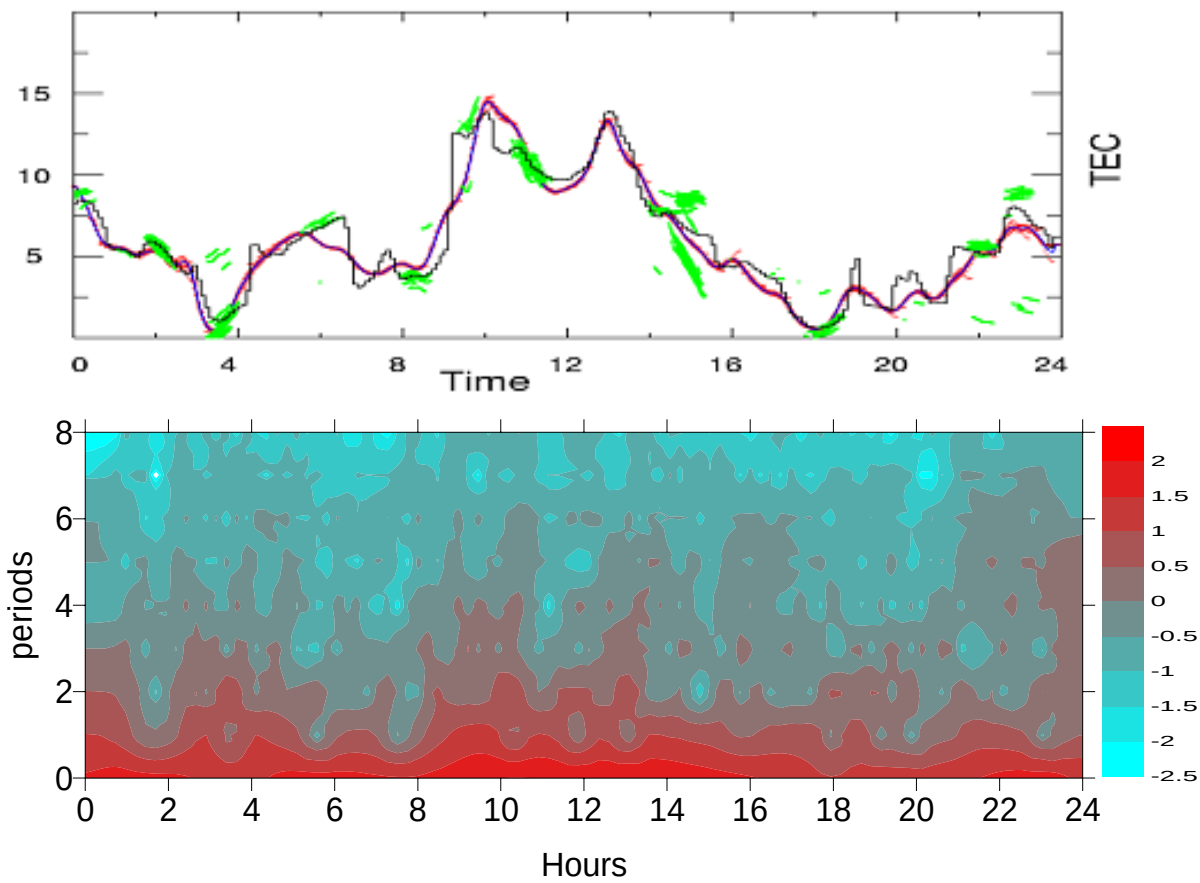

Figure 12: Upper panel: model (5, 5, 7) in red, measured TEC with green dots, and number of data points in units of 10^{16} cm^{-2} in black. Lower panel: the power spectra wavelet of amplitudes of Fig. 11. The magnitude of amplitudes is color coded in the right.

The physical interpretation of the observed wave structures is outside the scope of this work. In the next Fig.13 we demonstrate how the method of spectral analysis can be used to reveal some physical features of the observed phenomena. The upper row of the figure shows LAT/LONG maps of the residuals between models (5,5,3)-(5,5,5) obtained from 17.9 to 18.2 UT. It is seen the rise and decay of a small negative disturbance, stretched from northeast to southwest in the right part of the map. Because disturbed structure is revealed only by the residuals along the time axis (both spatial approximations are the same), according to Table 1 the size is larger than 1.5×3.0 and period is constrained between 24 and 48 min. Thus, we can consider that the observed structure reveals a part of a larger spectrum, limited within the above size and time ranges. The second and third rows from above show amplitudes of the basic period of 48 min and the first harmonic with period of 24 min (k=2). The last two rows show respective phases. It has to be noted that the phases are defined in the range [+π, -π] and the maps show discontinuity (sharp transition between red and blue) at places where phases go through π. Later we have to find another more attractive way to represent the phase development. It is evident that the disturbed structure in the first row is associated with increase of amplitudes (better seen for the basic wave) and the phase with a values close to π. The wave spatial structures closely resemble that of the density disturbance.

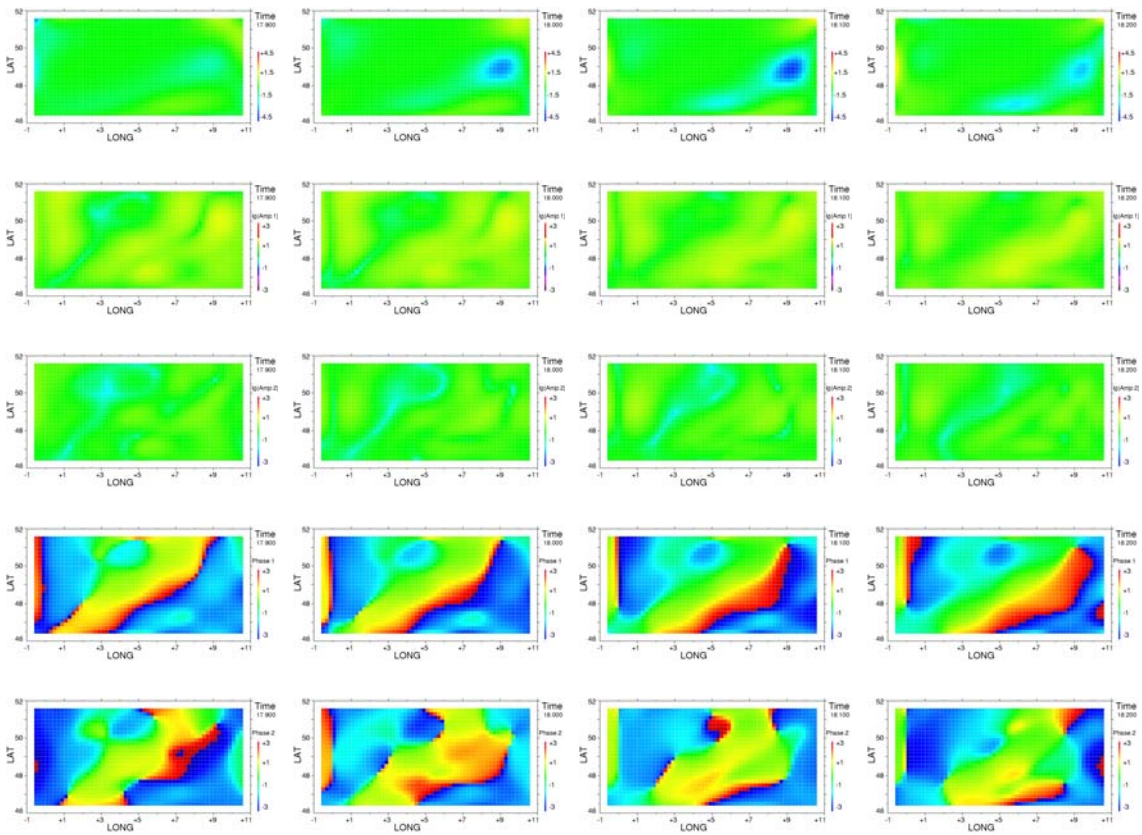

Figure 13: Residuals between models (5,5,3)-(5,5,5) (top row), wave amplitude with periods 48 min (2nd row) and 24 min (3rd row) and their respective phases (4th and 5th rows) for the 4 moments (columns) between 17.9 and 18.2 UT.

7. Correlation between modeled disturbances and geophysical indices

The main task of WP310 is to reveal the correlation of modeled disturbances and geomagnetic and solar activity indices. To study such a correlation, we need a long time series of data comprising at least several months (to account for seasonal effects) or several years (for solar cycle variations). It is obvious that for the present project we cannot form such large databases. To find a solution in the frame of the project we apply a comparative analysis by using an intermediate parameter, known to correlate closer with the modeled disturbance dynamics. As an intermediate parameter we use the RTK ionospheric intensity (or RTK events), defined and analyzed by Warnant et al, 2007a and 2007b. This parameter is defined as the standard deviation of de-trended TEC data from a single receiving station, averaged in 15 min interval. RTK ionospheric intensity is quantified in nine grades, depending on its values. We first will compare the RTK events obtained in the same period when model disturbances are available. Having obtained the consistency between them, we then will use a database, complied of RTK events during years 2000-2002 for correlation analysis. We will calculate correlation functions of RTK evens and geomagnetic indices K, Dst, and Hpi (Hemispheric Power index), and solar activity index F10.7 for a range of time lags. We assume that the correlation thus obtained will be representative for the ionospheric activity defined by model-extracted disturbances. These results will be presented in the final report of WP310.

8. Conclusions

We utilized a new approach for studying ionospheric disturbances. TEC, derived by the Belgian Dense Network of GPS stations within geographic area (46°x52°)N and (-1°x11°)E is approximated by 3-D polynomials of different degree. The corresponding model is named the Latitude-Longitude- Time (LLT) model. The residuals, when subtracting lower degree from higher degree polynomials, are considered representative of local density disturbances. We demonstrate that the present approach can localize TEC structures and follow their movement across the area. By using LLT model, we can determine the size, direction and speed of MSTIDs. We demonstrate also capability of the method of obtaining spectral characteristics of ionospheric disturbances by using a Fourier transform method. The knowledge of dynamics and spectral characteristics is a key factor of identifying and predicting of those ionospheric disturbances, most effectively degrading GPS position accuracy.

References

Rawer, K., 1993. Wave propagation in the ionosphere, Kluwer Academic Publishers, Dordrecht Boston London, pp.486.

Warnant, R., Kutiev, I., Marinov, P., Bavier, M., Lejeune, S., 2007a. Ionospheric and geomagnetic conditions during periods of degraded GPS position accuracy: 1. Monitoring variability in TEC which degrades the accuracy of Real Time Kinematics GPS applications, Adv. Space Res., 39, 875-880.

Warnant, R., Kutiev, I., Marinov, P., Bavier, M., Lejeune, S., 2007b, Ionospheric and geomagnetic conditions during periods of degraded GPS position accuracy: 2. RTK events during disturbed and quiet geomagnetic conditions, Adv. Space Res., 39, 881-888..