A New Technique to TEC Regional Modeling using a Neural Network.

Rodrigo F. Leandro

Geodetic Research Laboratory, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, Canada

BIOGRAPHY

Rodrigo Leandro is Ph.D. candidate at the Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada under supervision of Dr. Marcelo C. Santos. He holds a M. Sc. Eng. Degree in Civil Engineering from the University of São Paulo, in São Paulo, Brazil. He has been involved in research in the fields of Geodesy and satellite navigation.

ABSTRACT

In this paper we present a new technique of regional modeling of TEC (Total Electron Content), using a Neural Network model. This new model has the capability to predict TEC values derived from a GPS tracking network. Preliminary tests and respective results are shown. One of the main sources of errors of GPS measurements is the ionosphere refraction. As a dispersive medium, the ionosphere allows its influence to be computed by using dual frequency receivers. The use of two frequencies allow estimating the influence of ionosphere on GPS signal by the computation of TEC values, which have a direct relationship with the magnitude of the delay caused by the ionosphere. In the case of single frequency receivers it is necessary to use models that tell us how large the ionospheric refraction is. Such is the case of which the GPS broadcast message carries parameters of the Klobuchar model. One other alternative to single frequency users is to create a regional model based on a network of dual frequency receivers. In this case, the regional behaviour of ionosphere is modeled in a way that it is possible to estimate the TEC values inside or near this region. This regional model can be based on polynomials, for example. We have investigated a Neural Network-based model to the computation of regional TEC. The advantage from the use of this Neural Network model is that with the same model we can predict values for a station either within or outside the network, due to the adaptation capability of neural networks training process, that is an iterative adjust of the synaptic weights in function of residuals, using the training parameters. We have used data from the permanent GPS tracking network in Brazil (RBMC). We have tested the accuracy of the new model at all stations. To perform the tests TEC values were computed for each station of the network, except for a test station. After that the training parameters data set for the test station was formed, based on the TEC values of all other stations of the GPS network. The Neural Network was trained with these parameters, and tested by computing the TEC for the test station. This assessment was carried out several times, one for each station of the network. Preliminary assessment of results using our new technique shows a capability of retrieving around 85 % of TEC values for all stations. This means that we can correct the ionospheric delay at the same amount, due the direct relationship between both TEC and ionospheric delay.

INTRODUCTION

Ionospheric refraction is one of the most damaging effects on GPS signal. This effect is proportional to the total electron content (TEC), which is the number of free electrons contained in the ionospheric layer. Electrons of atmosphere are generated due to several factors, including solar activity. Figure 1 shows how solar radiation can create electrons in the atmosphere, forming the ionospheric layer.

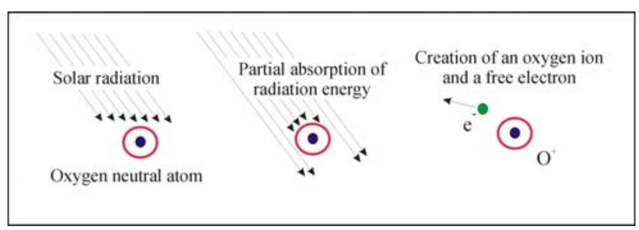

Figure 1. Creation of an oxygen ion and a free electron.

Once the TEC is known, it is possible to determine the delay caused by the ionosphere on GPS signal. Due to the dispersive characteristic of the ionosphere, the delay is a function of the frequency. It is possible to know the value of TEC using a dual frequency GPS receiver. Using the observations at both frequencies it is possible to compute the TEC value for the local where the station is. This computation will be explained in more detail in the next sections.

One alternative for single frequency receiver users is to use a regional model of TEC, generated by using data from a tracking network of dual frequency receivers. There are several ways to create such model. A network of receivers can generate a spatially distributed grid of TEC values. Using this grid it can be created a model from which is possible to estimate a TEC value to any position inside or near the region covered by the tracking network. Once the local TEC value is estimated, it is possible to correct the single frequency receiver observations. In this paper we present a new technique to regional TEC modeling, using a Neural Network approach. This new technique has the capability to predict TEC values derived from a GPS tracking network. Preliminary tests using the new technique indicate an average accuracy in the TEC values estimation of 97.5 %. In other words we can correct the ionospheric delay by the same amount, due to its direct relationship with TEC. These preliminary tests and respective results will be shown later in the paper.

TEC COMPUTATION USING A DUAL FREQUENCY RECEIVER

This section deals with the first step of our technique, that is the computation of the Vertical TEC (VTEC), using dual frequency observations. This computation allows the determination of VTEC values for each station of the tracking network. The model for VTEC computation presented here is a simple model, because our final objective is not to get a great precision in the VTEC determination for the tracking station itself, but a good estimation using our regional model for void areas, which is the main subject of this work. These same values can be computed using different techniques, probably providing a better quality input data to the regional model. However it will be shown that the final results obtained using our approach are satisfactory.

Let the equation of the carrier phase measurements on two frequencies (L1 and L2) be:

$$\lambda_1 \cdot \varphi^s_{r1}(t) = \rho^s_r(t) - I^s_{r1}(t) + T^s_{r1}(t) + \delta^s_{r1}(t)\cdot c + \lambda_1 \cdot N^s_{r1}, \quad (1)$$

$$\lambda_2 \cdot \varphi^s_{r2}(t) = \rho^s_r(t) - I^s_{r2}(t) + T^s_{r2}(t) + \delta^s_{r2}(t)\cdot c + \lambda_2 \cdot N^s_{r2}, \quad (2)$$

where λ_1 and λ_2 are the carrier phase wavelengths, in meters, $\varphi^s_{r1}(t)$ and $\varphi^s_{r2}(t)$ are the carrier phase measurements for a receiver r and a satellite s, in cycles, $\rho^s_r(t)$ is the geometric distance between the receiver and satellite antennas, in meters, $I^s_{r1}(t)$ and $I^s_{r2}(t)$ are the effects caused by the ionospheric refraction, in meters, $T^s_{r1}(t)$ and $T^s_{r2}(t)$ are the effects caused by the tropospheric refraction, in meters, $\delta^s_{r1}(t)$ and $\delta^s_{r2}(t)$ are the combinations of the satellite and receiver clock errors, in seconds, c is the speed of light, in meters per second, and N^s_{r1} and N^s_{r2} are the carrier phase ambiguities, in cycles.

Taking the difference between (1) and (2) we get:

$$\lambda_2 \cdot \varphi^s_{r2}(t) - \lambda_1 \cdot \varphi^s_{r1}(t) = I^s_{r1}(t) - I^s_{r2}(t) + \lambda_2 \cdot N^s_{r2} - \lambda_1 \cdot N^s_{r1}, \quad (3)$$

where the geometric distance, tropospheric delay and clock errors terms were cancelled out due their same behaviour in both frequencies. If there is not a cycle slip the ambiguity terms are constant. Let us combine the ambiguity terms of both frequencies into a constant, as follows:

$$C^s_r = \lambda_2 \cdot N^s_{r2} - \lambda_1 \cdot N^s_{r1}, \quad (4)$$

where C^s_r is the combination of the ambiguity terms of the two frequencies. Substituting (4) into (3) we get:

$$\lambda_2 \cdot \varphi^s_{r2}(t) - \lambda_1 \cdot \varphi^s_{r1}(t) = I^s_{r1}(t) - I^s_{r2}(t) + C^s_r(t). \quad (5)$$

The influence of ionosphere on GPS signals (I^s_{r1} and I^s_{r2}) can be computed according to (Hoffmann-Wellenhof, 2001):

$$I^s_{r1} = \frac{40.3 \cdot TEC}{(f_1)^2}, \quad (6)$$

and

$$I^s_{r2} = \frac{40.3 \cdot TEC}{(f_2)^2}, \quad (7)$$

where f_1 and f_2 are the frequencies of the L1 and L2 carrier signals, in units of Hz, and TEC is the total electron content in $\text{electrons} \cdot 10^{16} \cdot m^{-2}$. Substituting (6) and (7) into (5) we will obtain the following expression:

$$\lambda_2 \cdot \varphi^s_{r2}(t) - \lambda_1 \cdot \varphi^s_{r1}(t) = TEC \cdot \left(\frac{40.3}{(f_1)^2} - \frac{40.3}{(f_2)^2} \right) + C^s_r. \quad (8)$$

For simplification we assumed the TEC as a constant value during the period used for the computation. The choice of the size of such period is arbitrary, but it needs to be large enough to provide a good number of degrees of freedom in the adjustment and small enough to satisfy the assumption that TEC is constant over that period. In this work, we used periods of one hour for each determination of TEC. We can evaluate (8) as follows:

$$\lambda_2 \cdot \varphi_{r2}^{s}(t) - \lambda_1 \cdot \varphi_{r1}^{s}(t) = 0.1050 \cdot \mathrm{TEC} + C_r^s , \tag{9}$$

$$\mathrm{TEC} + 9.52 \cdot C_r^s = 9.52 \cdot \left(\lambda_2 \cdot \varphi_{r2}^{s}(t) - \lambda_1 \cdot \varphi_{r1}^{s}(t)\right). \tag{10}$$

TEC is defined as being the number of free electrons contained in a column with one meter squared of transversal section, along the path of the signal through the ionospheric layer. It is a number associated to an inclined trajectory with respect to the local zenith, as a function of the elevation angle of the satellite. In addition to that, the signal goes through the ionosphere at coordinates different from those of the station, at the ionospheric piercing point. To correct for the inclination and the position of the piercing point we can use mapping functions, as follows:

$$\mathrm{TEC} = M \cdot \mathrm{VTEC} , \tag{11}$$

where

$$M = \frac{1}{\sin(el)} . \tag{12}$$

where el is the elevation angle of the satellite at the observing station. Equations (11) and (12) allow us to determine VTEC instead of TEC. The mapping function used in this work is a simple bilinear model, as follows:

$$\mathrm{TEC} = M \cdot \left(a_0 + a_1 \cdot \Delta\varphi + a_2 \cdot \Delta\lambda\right), \tag{13}$$

where $\Delta\varphi$ is the latitude difference between the observation point and the ionospheric piercing point, $\Delta\lambda$ is the longitude difference between the observation point and the ionospheric piercing point, and a_0, a_1 and a_2 are the coefficients of the bilinear model to be adjusted. Substituting (13) into (10) we will get the final expression for TEC computation used in this work:

$$\begin{aligned} M \cdot \left(a_0 + a_1 \cdot \Delta\varphi + a_2 \cdot \Delta\lambda\right) + 9.52 \cdot C_r^s = \\ = 9.52 \cdot \left(\lambda_2 \cdot \varphi_{r2}^{s}(t) - \lambda_1 \cdot \varphi_{r1}^{s}(t)\right) \end{aligned} . \tag{14}$$

With this expression is possible to compute the parameters a_0, a_1 and a_2 and to determine VTEC for the tracking stations using measurements of several satellites during a certain period of time. The parameters of the model are the same to any satellite, but for each satellite included to the adjustment we will have an additional term C_r^s. Therefore, in this adjustment we will have (3 + ns) unknowns, where ns is the number of satellites used in the computation. The number of observations can be determined according to:

$$no_{total} = \sum_{s=1}^{ns} no(s) , \tag{15}$$

where no_{total} is the total number of observations used in the adjustment, and $no(s)$ is number of tracked epochs of the satellite s during the period of time used to the computation. If a cycle slip occur during this period, it is necessary to add another term to determine the combined ambiguities, loosing one degree of freedom. Depending on the size of the period it may be better to ignore such satellite to avoid the increasing of the ambiguity terms quantity.

The linear system formed by the several observations according to the equation (14) can be solved using the Parametric Least Squares Method. Performing this computation for each station of the GPS tracking network we will have a VTEC value associated to a coordinate wherever we have a station of the network. These values will be the input parameters of our Neural Network Model, which will perform the estimation of VTEC for any other point in or near the region covered by the network. The Neural Network Model will be discussed in the following section.

THE NEURAL NETWORK MODEL

A Neural Network is an information processing system formed by a big number of simple processing elements, called artificial neurons, or simply neurons.

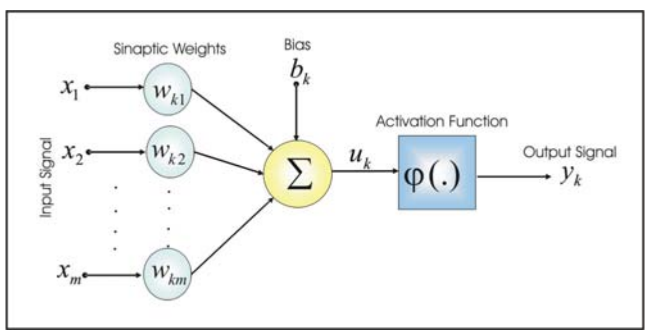

Figure 2. Nonlinear artificial neuron model (Adapted from Haykin, 1999).

A neuron computes its input as a linear combination of its input signal by using the synaptic weights. The synaptic weights play the role of parameters, which are adjusted at the training process (this procedure will be discussed later in this section). After that an activation function is applied to the neuron input to generate the neuron output (in the case of a single neuron it is already the output signal). One neuron may have one or more outputs, with the same value. In the case of a linear activation function, the neuron plays the role of a regression linear model. The processing of a neuron k can be represented by:

$$y_k = \varphi\left(\sum_{i=1}^{m}(x_i \cdot w_{ki}) + b_k\right), \tag{16}$$

where y_k is the neuron output, φ is the activation function, m is the number of input parameters, x_i is the i-th input parameter, w_{ki} is the i-th synaptic weight and b_k is the bias.

Typically the order of normalized amplitude of a neuron output is within the range [0,1], or alternatively [-1,1]. This range depends on the type of activation function used. The neural model also includes a term that is applied externally, called bias and represented by b_k. The bias has the function of increase or decrease the neuron input.

It is possible to introduce a functional link into the network as an additional layer of neurons, called a hidden layer. This layer can be composed of one or more neurons The input signal of the hidden layer neurons is generated by the output signal of the input layer. The output signal of the hidden layer is used to generate the input signal to the output layer. It is also possible to introduce not just one, but several hidden layers into the model.

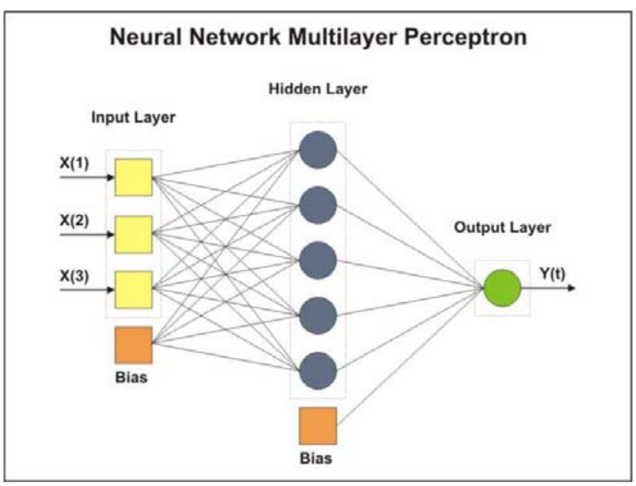

Figure 3. Neural Network Multilayer Perceptron

Figure 3 shows a scheme of a neural network with one hidden layer. In this example x(1), x(2) and x(3) are the input parameters and y(t) is the output parameter. Each element, excepting the biases, is a neuron. Each of these neurons is a processing element that works according to equation (16). The synaptic links (the lines in the draw) connect the different layers, carrying the output signal of a previous one to generate the input signal of the next one. Each synaptic link of the network has a corresponding synaptic weight that is applied to the flowing signal that is going through it.

Another issue of a neural network model is the number of neurons of each layer. This number is fixed to the input and output layers, in function of the input and output parameters. For the hidden layers this number is arbitrary. The model resulting from adding hidden layers between the input and output layers is called Multilayer Perceptron (MLP). The MLP is not the only type of neural network model, but is one of the most popular ones. In this work we have used a MLP.

It is necessary not just to know which model will be used (in our case the MLP), but also all its characteristics, such as the number of hidden layers, the number of neurons in each hidden layer, the activation function of each layer, etc. There are others more specific characteristics that will not be discussed here. The characteristics of the new model will be presented later in this section.

Once we have a model defined, it is necessary to train the neural network with data. Such data is composed by a set of know input and output parameters. The training process is not more than an adjustment of the synaptic weights to the data set. This adjustment attempts to decrease the residuals of the output of the network. The residuals are the difference between the computed output and the known output. Based on these residuals is performed an actualisation of the synaptic weights. Due to the complexity of neural networks the adjustment cannot be done with a direct computation. Therefore the so called training algorithms, which are a type of iterative adjustment of the synaptic weights, are used. One of these algorithms is the Backpropagation Training Algorithm, which is composed by two steps. The first one is the feed-forward, when the signal is propagated through the network, from the input layer to the output layer. After that the output value is compared with the known output and the residuals are computed. The second step is the feed-backward. In this step the errors are propagated through the network from the output layer to the input layer. During the feed-backward step the synaptic weights are adjusted. It is made several times to each parameter up to the residuals converge to a desired threshold value. After the training process we have a Neural Network

Model with adjusted synaptic weights according to the training parameters.

The presented model was created to estimate the VTEC for a certain position. The input parameters of the neural network model are Latitude and Longitude, while the output parameter is the VTEC. In this way, once the network is trained, it is possible to get a VTEC value for any location. The training parameters are the known coordinates and VTEC values of each station of the GPS network at a given time. Once the model is adjusted we can estimate a VTEC to any position inside or near the region covered by the GPS network to the given time.

Two hidden layers were used, each one with five neurons. The activation function of all layers (except the input one) is the hyperbolic tangent sigmoid function, represented in equation (17).

$$\varphi(x) = \frac{2}{1 + e^{-2 \cdot x}} - 1, \tag{17}$$

where x is the input signal of the neuron.

Figure 4 shows a scheme of the neural network model used. The techniques to apply the model with the GPS tracking network data are discussed in the following section.

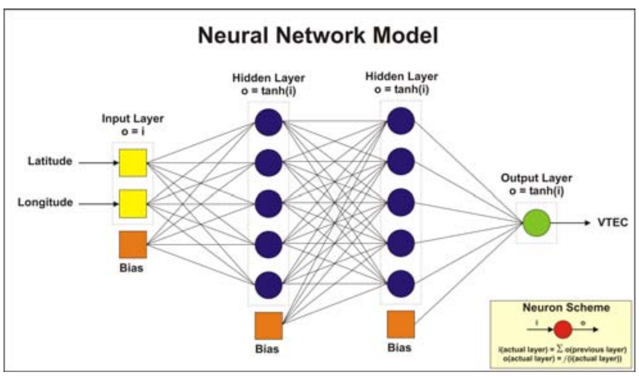

Figure 4. The Neural Network Model.

ANALYSIS STRATEGY

The data used in this work was obtained from the RBMC (Brazilian Continuous Monitoring Network), which is a GPS tracking network in Brazil. Figure 5 shows the configuration of such network.

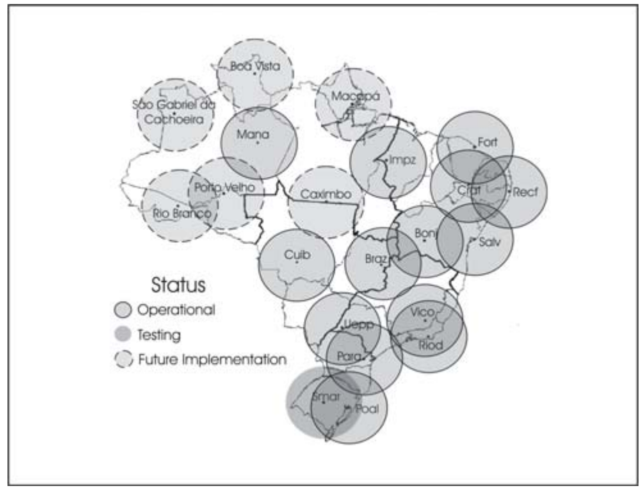

Figure 5. Stations of the RBMC

The advantage of using that network is due to the continental dimensions of Brazil, what can be considered one additional factor to test the capability of the model to estimate the TEC to long distances. Due to operational restrictions (not all stations are always operational) we did not use the whole network, but the stations BOMJ, BRAZ, CRAT, CUIB, IMPZ, PARA, POAL, RIOD, SALV, SMAR and UEPP, in a total of 11 stations. All stations were used either to calibrate the model or as a test station.

For each determination the test station data was not used during the training process of the neural network. After the training process the model was used to estimate the VTEC value for the test station position. This value is then compared with the known VTEC value obtained with the techniques expalined in previous sections. The difference between them shows the error of the prediction of the Neural Network Model. Performing this procedure to each of the 11 stations we could access the eficiency of the model everywhere. Using this technique we could analise the performance of the model for predictions inside and at the edges of the area covered by the network. Figure 6 shows a flowchart of the data processing for each given time.

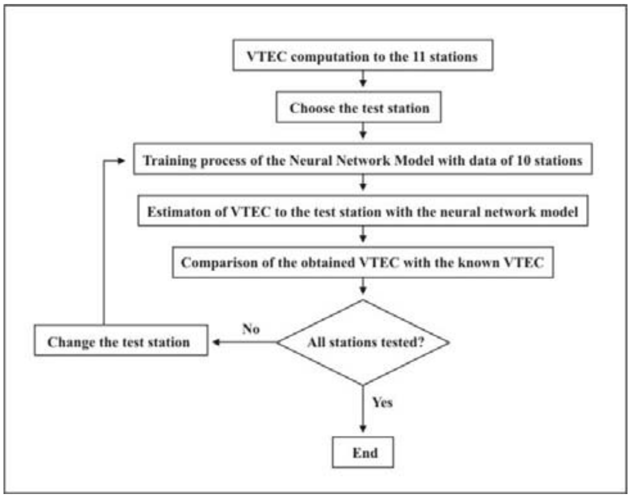

Figure 6. Flowchart of the data processing.

We performed these tests for two different periods of five days: One with low solar activity and the other with high solar activity. The low solar activity period covered days form February 1st 2004 to February 5th 2004. The high solar activity period covered days from October 26th 2003 to October 30th 2003. The index used to the analysis of solar activity was the solar radio flux. Figure 7 shows the behaviour of this index, within the chosen periods.

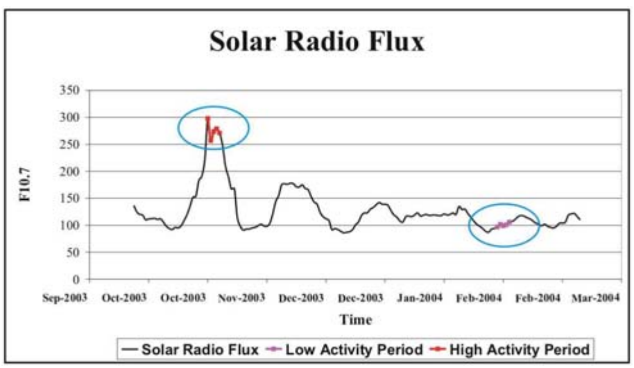

Figure 7. Solar radio flux.

For each day tests were performed to compute VTEC at 12, 14 and 16 hours (local time), corresponding to three predictions per day per station, resulting in a total of 318 predictions. The period used for the TEC computation was 1 hour. Prediction results are shown in the following section.

ANALISYS OF RESULTS

Results were analysed by assuming both absolute and elative errors. The absolute errors can be computed according to:

$$\text{Absolute Error} = \left|\text{VTEC}_e - \text{VTEC}\right|, \tag{18}$$

where VTEC is the computed value of VTEC, in TECU, and VTEC_e is the estimated value of VTEC, in TECU. The relative errors can be computed according to:

$$\text{Relative Error} = \frac{\text{Absolute Error}}{\text{VTEC}} \cdot 100 . \tag{19}$$

The less the absolute and relative errors are (as given by equations (18) and (19)), the closer are the predicted VTEC_e (given by our Neural Network Model) and the computed VTEC (determined from dual frequency receivers) used as reference.

Due to the direct relationship between TEC and ionospheric delay, according to equation (6) and (7), we can correct the ionospheric delay with a similar accuracy of the estimation of TEC. The results of VTEC estimations can be regarded as an estimated accuracy for correcting the ionospheric delay to single frequency receivers.

In this investigation 318 estimations were made with our new model, involving different stations, days and time of the day. The average absolute error of all estimations is equal to 3.7 TECU with standard deviation of 2.7 TECU (1 sigma). The average relative error was 14.9 %, with standard deviation of 10.9 % (1 sigma).

Figures 8 and 9 show the minimal absolute and relative errors for all stations, respectively.

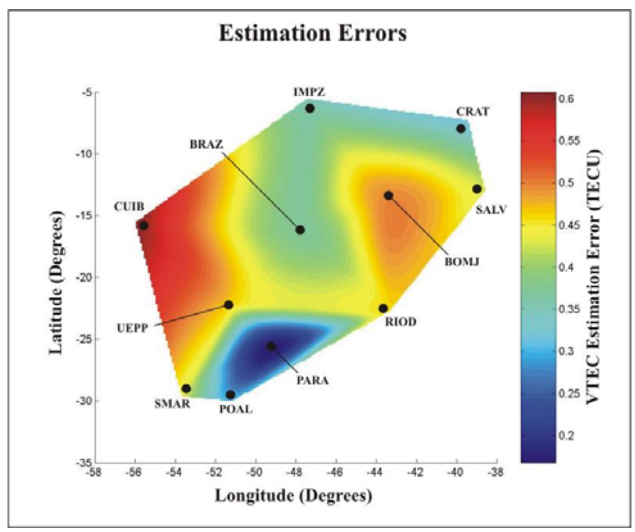

Figure 8. Minimal absolute errors for all stations.

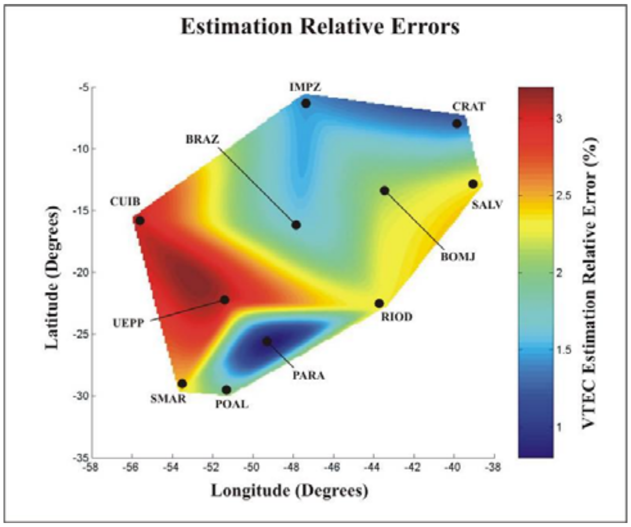

Figure 9. Minimal relative errors for all stations.

The worse average results were obtained for station IMPZ. It could be expected, since this station is the farthest from the others. The case of station IMPZ is an extrapolation case. The average results obtained for this station are 5.5 TECU and 18 % for absolute and relative errors, respectively

Figures 10 and 11 show the absolute and relative minimal errors, respectively, for all stations during the period of low activity.

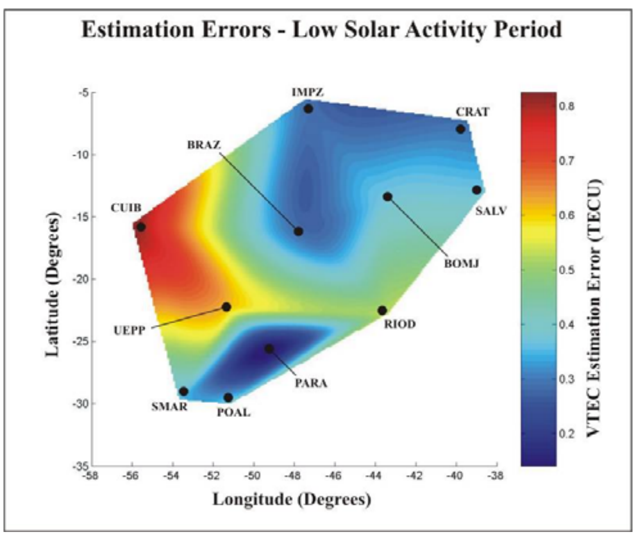

Figure 10. Minimal average errors during the low solar activity period.

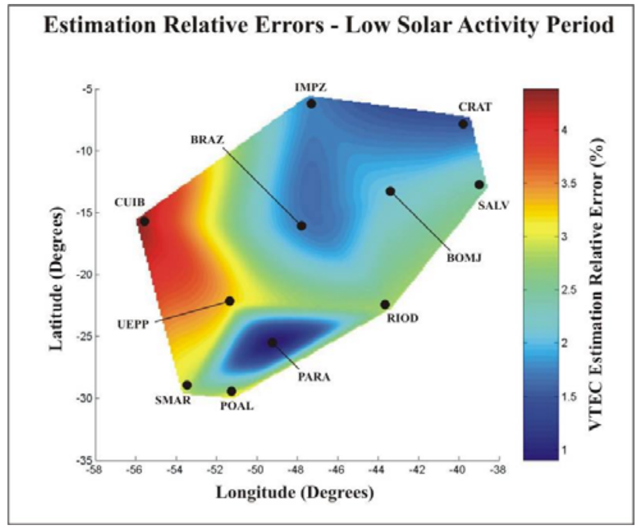

Figure 11. Relative minimal errors during the low solar activity period.

Figures 12 and 13 show the absolute and relative minimal errors, respectively, for all stations during the period of high activity.

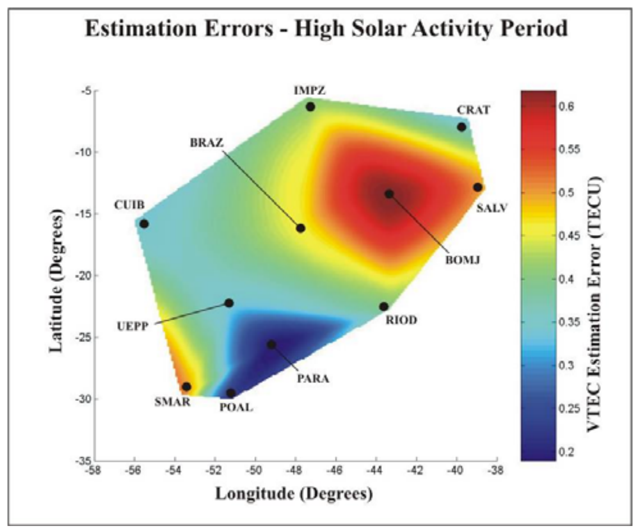

Figure 12. Absolute minimal errors during the high solar activity period.

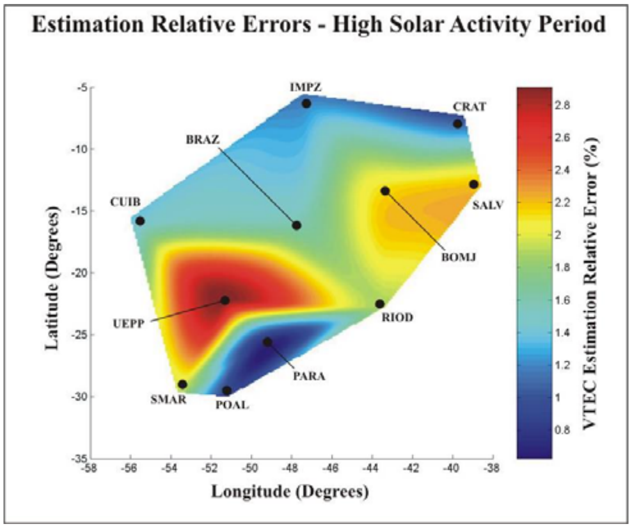

Figure 13. Relative minimal errors during the high solar activity period.

The absolute errors in both cases (low and high solar activity) were obtained in the same order of magnitude. Because the larger the TEC values during high activity periods, the smaller the relative errors in this situation than those obtained during the low activity period are.

CONCLUSIONS AND FUTURE RESEARCH

The model performed estimations with an average error of 3.7 TECU with standard deviation of 2.7 TECU (1 sigma). The average relative error was 14.9 %, with standard deviation of 10.9 % (1 sigma). This means that according to these preliminary results the new model allows to correct approximately 85 % of the ionospheric refraction to a single frequency receiver inside or outside the region. It can be concluded that the new model is adequate to predict VTEC values. The value of the standard deviations allow us to conclude that there was not great differences when comparing different stations, days, times or even solar activity levels.

The worse absolute and relative results were obtained in periods of high and low solar activity, respectively. As said before, this is due to a lower denominator value in equation (19), during low solar activity period. It can be also concluded that there is a stability of the estimations in terms of absolute values, during different situation with respect to solar activity. Therefore the average relative results for the high solar activity are better than those for low activity. However, the absolute errors during high solar activity period are not so good as those for calm periods.

Eventhought the spacing of tracking stations of the network used in this research is sparse, the model produced good estimations. With a larger number of stations it would be expected an even greater stability and confiability of the estimations of the model.

The worse average results were obtained for station IMPZ. It could be expected, since this station is the farthest from the others. The case of station IMPZ is an extrapolation case. Since the results are not bad (average of 5.5 TECU and 18 % for absolute and relative errors, respectively) the new model is a good model to estimate TEC values not just inside the region covered by the tracking network, but also outside it. An important consideration is the distance between IMPZ and the nearest station, of the order of eight hundred kilometres. It shows the capability of our model to extrapolate.

Future research is required to a complete validation of the model, assessing the efficiency of the new technique to different conditions of geomagnectic and solar activity. Comparison of the estimations of this new model with current models is another way to validate of technique.

Since it is concluded that the technique is a good way to modeling regional TEC, it can be possible go ahead and investigate similar techniques to global TEC modelling. Probably it would be necessary changes in the neural network configuration, based on the complexity of the problem.

ACKOWLEDGEMENTS

I would like to acknowledge the Brazilian Institute of Geography and Statistics (IBGE) for the use of the GPS data. Particular thanks to Mrs. Katia Pereira. Thanks to my supervisor Dr. Marcelo Santos. Funds for this research provided by NSERC.

REFERENCES

Haykin, S. (1999). Neural Networks – A Comprehensive Foundation. Prentice Hall – Upper Saddle River, New Jersey.

Hoffmann-Wellenhof, B. et. al. (2001). Global Positioning System: Theory and Practice. Springer-Verlag Wien New York.

Klobuchar, J.A. (1987). Ionospheric Time-Delay Corrections for Advanced Satellite Ranging System. AGARD Conference Proceedings – Propagation Limitations of Navigation and Positioning Systems.

Komjathy, A. and Langley, R. B., (1996). An Assessment of Predicted and Measured Ionospheric Total Electron Content Using a Regional GPS Network. Proceedings of the National Technical Meeting of the Institute of Navigation, pp. 615-624.

Langley, R. B. and Komjathy, A., (1996). High Precision Ionospheric Total Electron Count Mapping Using the Navstar Global Positioning System. American Geophysical Union, Western Pacific Geophysics Meeting, Brisbane, Australia.

McKinnell, L. (2002). A Neural Network based Ionospheric Model for the Bottomside electron density profile over Grahamstown, South Africa. Ph.D. Thesis, Rhodes University.

NorthWest Research Associates (2004). Space Weather: F10.7 Indices, http://www.nwra-az.com/spawx/f10.html, accessed March 1st.

Space Environment Center (2004). Old_indices, http://www.sec.noaa.gov/ftpmenu/indices/old_indices.html, accessed March 1st.

WELLS, D. et al. (1987). Guide to GPS Positioning. Canadian GPS Associates, Fredericton.