Soft Computing Technique for Recognition of Earthquake Precursor from Low Latitude Total Electron Content (TEC) Profiles

S. Kalita
Department of Instrumentation and USIC

Gauhati University

Assam, India

M. Devi
Department of Physics
Gauhati University

Assam, India

A. K. Barbara
Department of Physics
Gauhati University

Assam, India

P. H. Talukdar
Department of Instrumentation and USIC

Gauhati University

Assam, India

ABSTRACT
Total Electron Content (TEC) data from GPS are now used as tools for identifying an impending earthquake. The extraction of earthquake induced features from this parameter needs elaborate processing because it involves filtration of data with respect to disturbed day variations, contribution from multi path effects and also normal day-to-day fluctuations even during quiet days. In this paper we introduce, a peak detection algorithm, an automatic pattern matching approach and artificial neural network for processing GPS generated TEC record in where a template is framed from the time series of quiet day data for extraction of pre – seismic parameter. This matching algorithm is adopted to find deviation in TEC time series prior to an impending earthquake from the template and would be used as an index of earthquake induced signature. A few case studies using these techniques are presented in this paper. Work is based on TEC data collected from GPS at Gauhati (26º 10' N, 91º 45' E).

Keywords
Total Electron Content (TEC), GPS, pattern matching algorithm, earthquake.

1. INTRODUCTION
Effectiveness on the use of ionospheric parameters as tools for proposition of theories on prediction of earthquake has now become evident from considerable number of reported works [1,2,4,10,12]. However earthquake induced effects on the ionosphere are inherently complex in nature because even at normal times the ionosphere is influenced by solar geomagnetic ambience. In addition to this, in latitudes like that of Guwahati (26º 10' N, 91º 45' E) the effect of a special phenomenon known as Equatorial Anomaly makes such prediction process very complex [5,10,14]. Any attempt on identification of earthquake induced features in Total Electron Content (TEC) or F-layer electron density (foF2) of the ionosphere at the Appleton anomaly zone is therefore a difficult exercise. But on the other hand, difficult though, such endeavour offers opportunities for understanding the earthquake time dynamical and physical processes which interact with ionospheric and terrestrial seismomagnetic electrical systems [2,3,13]. In such an attempt GPS data on TEC taken at Guwahati, a seismically active Anomaly Crest station in sub Himalayan region, are examined by adopting a peak detection algorithm for detecting peak values of TEC profile, a pattern matching algorithm for measuring similarity between the sequences which varies with time and artificial neural network for processing of adopted methods.

2. ANALYSIS AND TECHNIQUE
2.1 The Peak Detection Method
Algorithm peak // one peak detection algorithms that uses peak function S_1

input $T = x_1, x_2, …, x_N, N$ // input time-series of N points

input k and h   // typically $1 \leq h \leq 3$

output O // set of peaks detected in T

begin

      $O = \varnothing$ // initially empty

for (i = 1; i < n; i++) **do**

a[i] = $S_1(k,i,x_i,T)$; // compute peak function value for each of the N points in T

end for

Compute the mean m' and standard deviation s' of all positive values in array a;

 for (i = 1; i < n; i++) **do** // remove local peaks which are "small" in global context

if (a[i] > 0 && (a[i] – m') >($h * s'$)) **then** $O = O \cup \{x_i\}$;

end if

end for

Order peaks in O in terms of increasing index in T

// retain only one peak out of any set of peaks within distance k of each other

for every adjacent pair of peaks x_i and x_j in O **do**

if $|j – i| \leq k$ **then** remove the smaller value of $\{x_i, x_j\}$ from O **end if end for**

end

2.2 The Dynamic Time Warping (DTW) Algorithm
In this approach a template is framed from five quietest day TEC profiles of a month. This group will be then used as a reference time series and defined as 'Template' for extracting a significant deviation that may appear on any day and its magnitude is marked by a parameter we call "certainty". The certainty magnitude will be maximum if any profile matches with the template. This deviation therefore addresses TEC peak as well as any change in shape between the two time series.

The technique uses a dynamic programming approach [6,7,8,11] to align a time series with a specified word template so that distance between the two is minimum. Specifically the pattern detection task involves searching a time series S for instances of a template T where,

$$S= s_1, s_2, s_3,.................,s_n$$
$$T= t_1, t_2, t_3,.............,t_m$$

The sequences S and T can be arranged to form a n-by-m grid where each grid point (i, j) corresponds to an alignment between elements s_i and t_j. A warping path W aligns the elements of S and T such that the distance between them is minimized. W can be written as

$$W= w_1, w_2, w_3,........., w_p$$

Where each w_k corresponds to a point $(i, j)_k$ in the grid. When there is no difference in timing, the warping path coincides with the diagonal line i=j. There are a number of techniques available for measuring 'distance' between two elements and this parameter can be expressed either in "magnitude of the difference" or "square of the difference" as given by the following equation.

$$\delta(i,j)= |s_i - t_j|$$
$$\delta(i,j)= (s_i - t_j)^2$$

To find the best match or alignment between these two sequences one need to find a path through the grid which minimizes the total distance between them. The procedures for computing this path involves finding all possible routes through the grid and for each one compute the overall distance. Now the dynamic time warping problem can be formally defined as a minimization over potential warping paths based on the cumulative distance for each path.

$$\text{DTW } (S, T)= min_w\left\{ \sum_{k=1}^{p} \delta(w_k) \right\} \quad (1)$$

Searching through all possible warping paths to obtain minimum distance is very tedious. To reduce the search space several types of restrictions can be used, some of these are

(i) Monotonicity: the path will not turn back on itself, i.e. for consecutive pairs w_{k-1} and w_k in W, $i_{k-1} \leq i_k$ and $j_{k-1} \leq j_k$.

(ii) Continuity: The path advances one step at a time i.e. $i_k - i_{k-1} \leq 1$ and $j_k - j_{k-1} \leq 1$.

(iii) Boundary condition: the path starts at the bottom left and ends at the top right.

(iv) Warping window condition: a good path is unlikely to wander very far from the diagonal. The distance that the path is allowed to wander is the window width.

(v) Slope constraint condition: The path should not be too steep or too shallow thereby avoiding excessively large movements in a single direction.

Now to formulate a dynamic programming problem we need a recurrence relation. Following is a recurrence relation which defines the cumulative distance

$$\gamma(i, j)= \delta(i, j) + min[\gamma(i-1, j), \gamma(i-1, j-1), \gamma(i, j-1)]$$

Other similar recurrence functions can also be used but the above relation is symmetric and it has been seen that symmetric formulation gives better result in speech recognition field etc. We will utilize this equation (2) for our application in TEC for identifying earthquake precursor.

The certainty factor (D) is then defined by the length normalized distance function

$$D(A,B)=\gamma(n,m)/(n+m) \quad (3)$$

where A and B are curves respectively of the template and the other time series.

The dynamic programming algorithm fills in a table of cumulative distances as the computation proceeds. Upon completion the optimal warping path can be found by tracing backward in the table.

2.3 Artificial Neural Network

An artificial neural network (ANN), usually called neural network (NN), is a mathematical model or computational model that is inspired by the structure and/or functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation.
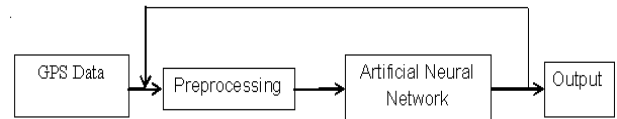

Figure 1:Block diagram of the model

A block diagram of the system used is shown in Figure 1. The raw time series values are preprocessed and then it is used as an input to the hidden layer and a multilayer perceptron is used to process the data. The back propagation technique is used for construction of the prediction model and following steps are utilized:

1. Artificial Neural Network (ANN) computes the certainty value of TEC time series obtained from DTW method for predicting earthquake.

2. The ANN computes peak vales from peak detection algorithm for developing rule for prediction of earthquake.

3. The output is the predicted earthquake event i.e. the result generated by the entire system.

3. RESULT
3.1 TEC PEAK Analysis Result

Peak values of TEC are calculated using the peak detection algorithm with the following features:

(i) There are normal fluctuations of TEC peak within Q-day Sd limits but in some days magnitude increases so much that it exceeds the maximum excursion bar of Sd limit of Q-day TEC (thick lines).

(ii) On attainment of the maximum, the peak TEC decreases and

(iii) On days near or at the trough of this decline, occurrence of a low latitude earthquake is noted.

Thus the features (i) and (ii) are flagged as identification marks for a possible earthquake. Based on these features we have plotted Figure 2 for predicting earthquake days with observed events
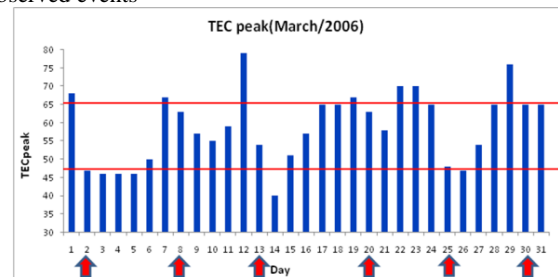

Figure.2: TEC day to day variation peak for March 2006.
indicates predicted earthquakes (Magnitude ≥ 5) occurrence days in the latitude region ($26 \pm 5°$). Threshold points are calculated and shown by thick lines in the figure.

3.2 Pattern Matching Technique Results

Case 1: Bhutan Earthquake of 21 September 2009

The 2009 Bhutan earthquake with M=6.1 magnitude occurred at 14:43 (82.5 EMT), on September 21, 2009 in the eastern region of Bhutan. The epicentre was at 180 kilometers (radius =445.85 km) east of the capital Thimphu and thus the receiving station at G.U. is within the earthquake preparatory zone. As DTW   technique demands framing of the template from five Q-days of this month for finding out "certainty" parameter,   we present in figure 3a the Q-day TEC feature along with one of the pre- earthquake day to highlight the differences between the two. The modification in shape of TEC before the earthquake day (see figure 3b) is significant compared to that for Q-day one. And it is expected that the certainty factor should be low before this event. Thus from a number of quiet day   TEC profiles a template for this particular month is created to measure distance of deviation between template and time series of other days of this month and the certainty factor is calculated [15].
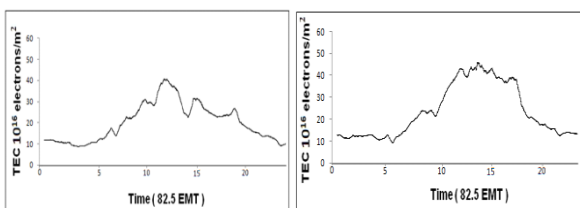
Figure. 3(a)   Figure. 3(b)

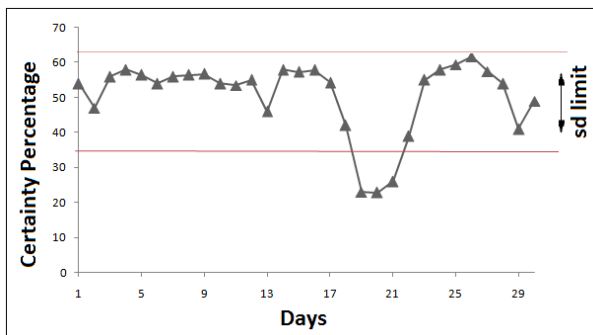**Figure 2. TEC profile for 3(a) Q-day and 3(b) pre earthquake day of 20.09.2009**

Figure 4: Certainty factor for the month of September 2009. Note the depletion of this factor during earthquake time; earthquake occurred on 21.9.2009.

The certainty percentage is plotted (Figure 4) against day for the entire month and it is clearly observed that the magnitude of this parameter has gone down below the sd limits (shown by red lines) from 19th Sep 2009 i.e two days prior to the earthquake and the parameter revives to its original value from 22nd Sep 2009. This indicates that it is possible to identify an earthquake day by calculating the certainty factor using the method.

Case 2: Strong China Earthquake of 12 May 2008

We now take one stronger recent earthquake to test the reliability of this approach.

The 2008 Sichuan earthquake or the Great Sichuan Earthquake was a deadly earthquake that measured at 8.0 M_s and 7.9 M_w occurred at 01.08 (82.5 EMT) on May 12, 2008. The quiet day template for this month has been made using the same approaches as for Bhutan earthquake case and certainty factor between the template and other days of this

month that is for May, 2008 is calculated and is presented in Figure 5 to the month of May 2008. Here too a clear depletion of this value one day prior to the earthquake is seen.
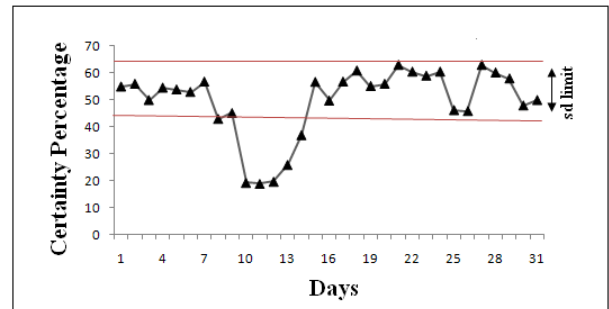
Figure 5: Certainty (D) percentage calculated by DTW method for the month of May, 2008. Note the decrease in the certainty factor prior to the earthquake day, marked by arrow head in the figure.

3.3 Artificial Neural Network result

The prediction method is implemented using MATLAB ver 7.0 of the mathwork inc. It is widely accepted mathematical tools provides a variety of toolboxes with built in algorithms and functions for optimization, neural network and genetic algorithm implementation. The proposed prediction model is consisting of 80 input nodes. The hidden layers in our architecture uses the hyperbolic tangent sigmoid transfer function which complies with artificial neural network's input data.Then the system is examined with both applied algorithm peak detection and DTW algorithm with different prediction events and then the accuracy of the prediction model is calculated.

Method	No. of predicted events	Accuracy
Peak detection	1	92%
DTW	1	94%
Peak detection	3	81%
DTW	3	87%
Peak detection	5	71%
DTW	5	86%

From the above table it is cleared that the DTW algorithm produces best result for predicting earthquake than the peak detection algorithm form GPS TEC time series data.

4. CONCLUSION AND FUTURE SCOPE

We have here adopted soft computing technique for identifying earthquake day profile from Q-day using Peak Detection technique as well as the Dynamic Time Warping(DTW) algorithm for pattern reorganization. This is the first step for utilizing this approach for identification of an impending earthquake. We have identified signature of strong earthquakes of M>6 when our observing station lies in the earthquake preparatory zone. However the limitation is that it cannot resolve fine structure (like changes in hours) because the enormous increase in size of the necessary matrix for computation such changes.

It is planned to adapt, in future, the fuzzy transformation approach for predicting TEC time series data. The resultant time series will be compared using DTW and peak detection technique and be processed by the Artificial Neural Network which will be utilized for identifying an impending earthquake in geomagnetic disturbance ambiance.

4. REFERENCES

[1] Assent, I., Kremer, H., Günnemann, S., Seidl, T., 2010. Pattern Detector: Fast Detection of Suspicious Stream Patterns for Immediate Reaction. In: Proceedings EDBT/ICDT 2010 Joint Conference, Electronic Conference.

[2] Depueva A. H., and Ruzhin Yu. Ya., 1995. Seismoionospheric fountain effect as analogue of active space experiment. Advances of. Space Research , 15, 12-15.

[3] Depueva A., Mikhailov A., Devi M., and Barbara A.K., 2007 . Spatial and time variations in critical frequencies of the ionospheric F region above the zone of equatorial earthquake preparation. Geomagnetism and Aeronomy, 47, 129-133.

[4] Devi M., Barman M.K., Barbara A.K., and Depueva A.H., 2001. Total electron content near anomaly crest as precursor of earthquake. Indian Journal of Radio Space Phys, 30, 209-213.

[5] Devi, M., Barbara, A.K., Barman, M.,1996. Ionospheric irregularities and storm-induced equatorial and high-latitude effects at the anomaly crest region. Radio Physics and Quantum Electronics, 39, 179-187.

[6] Efrat, A., Fan, Q., Venkatasubramanian, S., 2007. Curve matching, time warping, and light fields: New algorithms for computing similarity between curves. Journal of Mathematical Imaging 27(3), 203-216.

[7] Furtuna T. F.,2008. Dynamic Programming Algorithms in Speech Recognition, Revista Informatica Economică, 2(46).

[8] Han, J., Dong, G., Yin, Y., 1999. Efficient mining of partial periodic patterns in time series database. International journal of data engineering, 106-115.

[9] Hayakawa, M., 2004. Electromagnetic phenomena associated with earthquakes: A frontier in terrestrial electromagnetic noise environment. Recent Research Development in Geophysics, 6, 81-17.

[10] Huang, N .,1988. Geomagnetic and Ionosphere studies in Taiwan. Journal of Electrical Engineering 31, 281-294.

[11] Kuzmanic A., Zanchi, V., 2007. Hand shape classification using DTW and LCSS as similarity measures for vision-based gesture recognition system. In: EUROCON, The International Conference on "Computer as a Tool", pp. 264-269.

[12] Liu J.Y., Chen Y.I., Chou Y.J., and Tsai H.F., 2001. Variations of ionospheric total electron content during the Chi-Chi earthquake. Geophysical Research Letter, 28, 1383-1386.

[13] Parrot M., Berthelier J. J. , Lebreton J. P. , Sauvaud J. A .,Santolik O., and Blecki J.,2006. Examples of unusual ionospheric observations made by the DEMETER satellite over seismic region. Physics and Chemistry of the earth, 31, 486.

[14] Walker, G. O., and Chen H. F., 1989. Computer simulation of the seasonal variations of ionospheric equatorial anomaly in East Asia under solar minimum conditions. Journal of Atmospheric and Terrestrial Physics,51, 953-974.

[15] M. Devi, A. J. D. Sarma, S. Kalita A.K. Barbara and A. Depueva (2011) ,Adoptive techniques on extraction of pre-seismic parameters on Total Electron Content (TEC) at anomaly crest stations using GPS data Geomatics, Natural Hazards and Risk DOI:10.1080/19475705.2011.595831