LEAST SQUARE SUPPORT VECTOR MACHINE FOR DETECTION OF TEC-SEISMO-IONOSPHERIC ANOMALIES ASSOCIATED WITH THE POWERFUL NEPAL EARTHQUAKE (Mw=7.5) OF 25 APRIL 2015

M. Akhoondzadeh

Remote Sensing Department, School of Surveying and Geospatial Engineeringl, University college of Engineering, University of Tehran, North Amirabad Ave., Tehran, Iran - makhonz@ut.ac.ir

Commission VIII, WG VIII/1

KEY WORDS: Ionosphere, TEC, earthquake, anomaly, LSSVM.

ABSTRACT:

Due to the irrepalable devastations of strong earthquakes, accurate anomaly detection in time series of different precursors for creating a trustworthy early warning system has brought new challenges. In this paper the predictability of Least Square Support Vector Machine (LSSVM) has been investigated by forecasting the GPS-TEC (Total Electron Content) variations around the time and location of Nepal earthquake. In 77 km NW of Kathmandu in Nepal (28.147° N, 84.708° E, depth=15.0 km) a powerful earthquake of M_w=7.8 took place at 06:11:26 UTC on April 25, 2015. For comparing purpose, other two methods including Median and ANN (Artificial Neural Network) have been implemented. All implemented algorithms indicate on striking TEC anomalies 2 days prior to the main shock. Results reveal that LSSVM method is promising for TEC sesimo- ionospheric anomalies detection.

1.   INTRODUCTION

The pre-seismic disturbances occurring in lithosphere, atmosphere and ionosphere in absence of significant solar and geomagnetic perturbations are usually considered as earthquake precursors.

The ionospheric anomalies usually take place in the D, E and F layers, and may be observed 1 to 10 days prior to the earthquake and continue a few days after it. Many papers and reports have been published on satellite observations of the ionospheric plasma, the flux of charged particles, the DC electric field, electromagnetic waves and geomagnetic field associated with seismic activity (Parrot, 1995; Liu et al., 2004; Hayakawa and Molchanov, 2002; Pulinets and Boyarchuk, 2004; Akhoondzadeh, 2011). The pre-earthquake disturbances usually affect the TEC (Total Electron Content). In this study, GPS-TEC data have been downloaded via NASA Jet Propulsion Laboratory (JPL) website. Global Ionospheric Map (GIM) data constructed from a 5° × 2.5° (Longitude, Latitude) grid with a time resolution of 2 hours.

The accurate anomaly detection in time series of earthquake precursors is regarded as one of the most challenging tasks since the behaviors of precursors is complicated, dynamic and nonlinear. In addition, the affected factors in ionospheric precursors such as solar and geomagnetic indices intensify the complexity of the ionospheric precursors.

As a modified version of Support Vector Machine (SVM), Least Square SVM (LSSVM) was proposed by Suykens and Vandewalle in 1999. It retains principle of the Structural Risk Minimization (SRM) and has important improvement of calculating speed with traditional SVMs (Wang and Shang, 2014). It is because of changing inequality constraints into equations and takes a squared loss function. Therefore, LSSVM solves a system of equations instead of a quadratic programming problem. Due to the mentioned advantages of LSSVM, it has been implemented as a classification model (Wang and Shang, 2014). In this paper the predictability of LSSVM has been investigated by predicting the GPS-TEC variations around the time and location of Nepal earthquake.

2.   LSSVM METHOD

SVM training is a time consuming process specially when analyzing huge dataset. For this purpose, LSSVM is

proposed by Suykens and Vandewalle in 1999 to overcome these shortcomings.

Unlike the classical neural networks approach, SVM formulates the statistical learning problem as a quadratic programming with linear constraints, by the use of nonlinear kernels, high generalization ability, and sparseness of solution. However, for large-scale problems, the optimization process of SVM has high computational complexity, due to the high-dimensional matrix involved in the quadratic programming whose size is directly proportional to the training sample size (Zhou et al., 2011).

LS-SVM needs significantly less training effort than the standard SVM as a result of the model simplification.

The basic principle of SVM regression is to estimate the output variable y from $\varphi(x)$, a high dimensional feature space of the input vector $x = (x^1, x^2, ..., x^K)^T$, where K is the order of SVM. Thus, the general SVM regression model is

$$y = w^T \varphi(x) + b \qquad (1)$$

Where w and b are the weight vector and bias term respectively (Zhou et al., 2011).

Given a sample of training data $\{(x_i, y_i)\}_{i=1}^N$, LS-SVM determines the optimal weight vector and bias term by minimizing the following cost function, R,

$$\min_{w,e} R(w,e) = \frac{1}{2}\|w\|^2 + \frac{1}{2}\gamma\|e\|^2 \qquad (2)$$

subject to the equality constraints

$$y_i = w^T \varphi(x_i) + b + e_i \qquad i = 1,2,..,N \quad (3)$$

where $e = (e_1, e_2, ..., e_N)^T$. The first part of cost function regularizes weight sizes and penalizes large weights. Therefore, the weights tend to converge to similar values in that large weights cause excessive variance and hence deteriorate the generalization ability of LSSVM (Suykens and Vandewalle, 1999; Zhou et al., 2011).

The second part of Eq. (2) considers the regression error of all training data. The regularization parameter c controls the trade-off between the bias and variance of LS-SVM model. Note that the LSSVM model has equality constraints as shown in Eq. (3), rather than the inequality constraints with slack variables used in the standard SVM model. Moreover, as shown in Eq. (2), a squared loss function is

considered in the objective function of LS-SVM model, while the standard SVM model has a linear combination of slack variables in its objective function. These two modifications simplify the quadratic optimization problem for the standard SVM to be linear for LSSVM (Suykens and Vandewalle, 1999; Zhou et al., 2011).

The Lagrangian of Eq. (2) is

$$L(w,b,e,\lambda) = R(w,e) - \sum_{i=1}^N \lambda_i (w^T \varphi(x_i) + b + e_i - y_i)$$

$$(4)$$

where λ_i are the Lagrange multipliers. By the Karush–Kuhn–Tucker Theorem, the conditions of optimality are (Suykens and Vandewalle, 1999; Zhou et al., 2011)

$$\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^N \lambda_i \varphi(x_i)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^N \lambda_i = 0$$

$$\frac{\partial L}{\partial e_i} = 0 \rightarrow \lambda_i = \gamma e_i$$

$$\frac{\partial L}{\partial w} = 0 \rightarrow w^T \varphi(x_i) + b + e_i - y_i = 0$$

$$(5)$$

Thus, b and λ can be solved from the following set of linear equations after eliminating w and e,

$$\begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & K + \gamma^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (6)$$

where $\vec{1} = (1,1,...,1)^T$, I is the identity matrix, $y = (y_1, y_2, ..., y_N)^T$, $K = (k(x_i, x_j))_{i,j=1}^N$ is the kernel matrix, and $k(x_i, x_j) = \phi^T(x_i)\phi(x_j)$ is the kernel function (Suykens and Vandewalle, 1999; Zhou et al., 2011).

As a result, given vectors x and x_i, the LS-SVM regression model for estimating y in Eq. (1) becomes

$$y = \sum_{i=1}^{N} \lambda_i k(x_i, x_j) + b \quad (7)$$

In this study Gaussian kernel function has been used.

$$k_G(x, z) = \exp(-\|x - z\|^2 / \sigma^2). \quad (8)$$

where $\|.\|$ denotes the 2-norm, and σ is a constant determining the width of Gaussian kernel.

To implement the LSSVM method, training and testing data were initially set respectively to 60% and 40% of all TEC data. The input patterns in the LSSVM method are,

$x_4 = f(x_1, x_2, x_3)$

$x_5 = f(x_2, x_3, x_4)$

.

(9)

.

$x_N = f(x_{N-3}, x_{N-2}, x_{N-1})$

At each step, using the training data, the LSSVM method is implemented and the prediction error (*PE*) can be written as:

$$PE = \sum_{i=4}^{N} (x_i - \hat{x}_i) \quad (10)$$

Where, x_i and $\hat{x}_i$ are the actual value and the output from the LSSVM method, respectively.

Finally, the TEC value is predicted and then is compared to the true value in testing set. In the case of testing process, if the value of DX_i (i.e. the difference between the actual value X_i and the predicted value $\hat{X}_i$) is outside the pre-defined bounds $\mu \pm k \times \sigma$, (μ and σ are the mean and the standard deviation of DX_i values) the anomaly is detected.

3. OBSERVATIONS

In 77 km NW of Kathmandu in Nepal (28.147° N, 84.708° E, depth=15.0 km) a powerful earthquake of M_w=7.8 took place at 06:11:26 UTC on April 25, 2015 (Table 1).

In this study, geomagnetic and solar indices (i.e. K_p, A_p, D_{st} and F10.7) were used to distinguish seismic anomalies from the other anomalies related to the geomagnetic and solar activities. Figure 1 illustrates the variations of K_p, A_p, D_{st} and F10.7 indices, during the period of 01 March to 30 April 2015. An asterisk indicates the earthquake time. The X-axis represents the days relative to the earthquake day. The Y-axis represents the universal time coordinate.

Figure 2(a) shows TEC variations during the period of 01 March to 30 April 2015. To implement the Median method, *Dx* which will be called *DTEC* here, is calculated using Eq. (11).

$$x_{high} = M + k \times IQR$$

$$x_{low} = M - k \times IQR$$

$$x_{low} < x < x_{high} \quad \Rightarrow \quad -k < \frac{x-M}{IQR} < k \quad ; \quad Dx = \frac{x-M}{IQR}$$

$$(11)$$

Where x, x_{high}, x_{low}, M, IQR and Dx are the parameter value, higher bound, lower bound, median value, Interquartile range and differential of *x*, respectively. According to this, if the absolute value of *Dx* would be greater than *k*, ($|Dx| > k$), the behavior of the relevant parameter (*x*) is regarded as anomalous. Figure 2(b) shows variations of DTEC. Figure 2(c) shows detected TEC anomalies using the Median method based on: $|DTEC| > 1.75$ and without considering the solar and geomagnetic indices. Then to distinguish pre-earthquake anomalies from the other anomalies related to the geomagnetic activities, the five conditions of $|DTEC| > 1.75$, $Kp < 2.5$, $Dst > -20\ nt$, $Ap < 10$ and F10.7 <150, respectively, are jointly used using AND operator to construct the anomaly map (Figure 2(d)). The TEC value exceeds the higher bound ($M + 1.75 \times IQR$), 2 days before the earthquake time at 14:00, 16:00 and 18:00 UTC with the values of 18.68 %, 15.97% and 0.32% of the higher bound, respectively. It is seen that the other detected anomalies in Figure 2(c) have been masked by high geomagnetic activities (Figure 2(d)). In figure 2(d) increases (08.40% and 04.13%) in TEC are clearly observed at 16:00 and 18:00 UTC on earthquake date (Table 2).
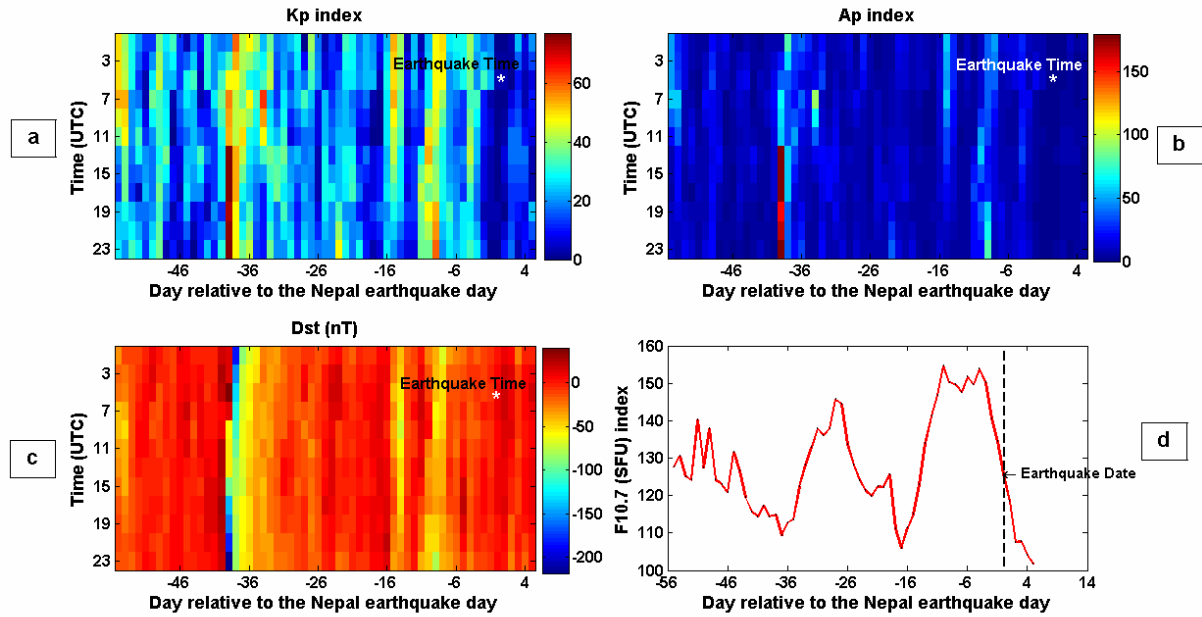
Fig. 1. a), b), c) and d) show respectively, the variations of K_p, A_p, D_{st} and solar radio flux (F10.7) indices during the period of 01 March to 30 April 2015. An asterisk indicates the earthquake time. The X-axis represents the days relative to the Nepal earthquake day.
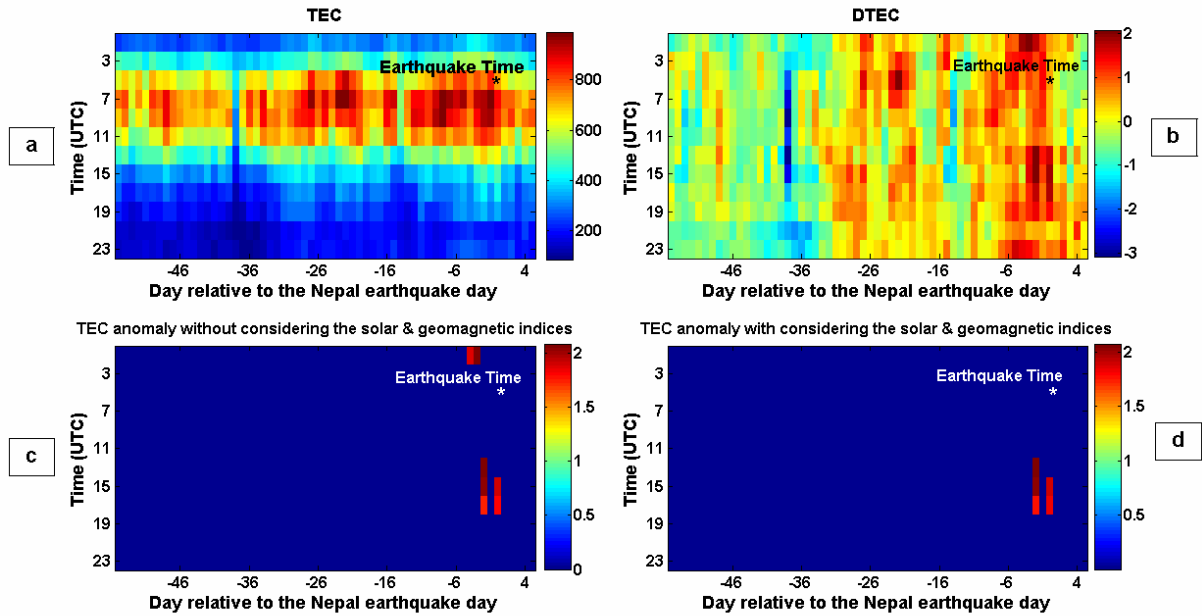

Fig. 2. a) TEC variations, b) DTEC variations after implementing the Median method, c) detected TEC anomalies using Median method without considering the geomagnetic conditions and d) detected TEC anomalies using Median method during quiet solar-geomagnetic conditions.

To implement the Artificial Neural Network (ANN) method (Akhoondzadeh, 2013), training data were set to 60% of all data. Using the training data, the pattern vectors in feature space are constructed. In the case of testing process, if the difference value PE_i between the actual value X_i and the predicted value $\hat{X}_i$, is outside the pre-defined bounds $\mu \pm 2.0 \times \sigma$, ($\mu$ and σ are the mean and the standard deviation of PE_i values) the anomaly is detected.

Red and green curves in figures 3(a) through (l) represent the observed and the predicted TEC values using the ANN method, respectively during the days selected as training and testing set. These figures indicate that the ANN method is a good estimator for non linear time series such as TEC variations. It can be seen that the ANN method efficiently predicts values during the time of testing.

Figures 4(a) through (l) represent the differences between the observed and the predicted TEC values during the testing data using the ANN method. Figure 5(a) is a representation of the differences values during the testing set. Figure 5(b) shows the DTEC values obtained from

$$Dx = \frac{x - \mu}{\sigma}$$

. In Figure 5(c), anomalous TEC values are only depicted at times when $|DTEC| > 2.0$. Then to distinguish pre-earthquake anomalies from the other anomalies related to the geomagnetic activities, the five conditions of $|DTEC| > 2.0$, $Kp < 2.5$, $Ap < 10$, $Dst > -20 \, nt$ and $F10.7 < 150$ respectively, are jointly used using AND operator to construct the anomaly map (Figure 5(d)). The DTEC value exceeds the upper bound ($\mu + 2.0 \times \sigma$) with the values of 8.14%, 2 days before the earthquake at 14:00 UTC, and also 4 days after the earthquake time at 22:00 UTC with the value of 38.49% (Table 2).

To implement the LSSVM method, training data were set to 60% of all data. Using the training data, the pattern vectors in feature space are constructed. In the case of testing process, if the difference value PE_i between the actual value X_i and the predicted value $\hat{X}_i$, is outside the pre-defined bounds $\mu \pm 2.0 \times \sigma$, ($\mu$ and σ are the mean and the standard deviation of PE_i values) the anomaly is detected.

Red and green curves in figures 6(a) through (l) represent the observed and the predicted TEC values using the LSSVM method, respectively during the days selected as training and testing set.

Figures 7(a) through (l) represent the differences between the observed and the predicted TEC values during the testing data using the LSSVM method. Figure 8(a) is a representation of the differences values between the observed and predicted values during the testing set. Figure 8(b) shows the DTEC values obtained from $Dx = \frac{x - \mu}{\sigma}$. In Figure 8(c), anomalous TEC values are only depicted at times when $|DTEC| > 2.0$. Then to distinguish pre-earthquake anomalies from the other anomalies related to the geomagnetic activities, the five conditions of $|DTEC| > 2.0$, $Kp < 2.5$, $Ap < 10$, $Dst > -20 \, nt$ and $F10.7 < 150$ respectively, are jointly used using AND operator to construct the anomaly map (Figure 8(d)). The DTEC value exceeds the lower bound ($\mu - 2.0 \times \sigma$) with the values of -26.18% and -30.89%, 2 days before the earthquake time at 02:00 and 24:00 UTC, and also with the value of -0.5%, on earthquake date at

Table 1. Characteristics of the Nepal earthquake and its main after shocks (reported by http://earthquake.usgs.gov/).

Date	Time (UTC)	Geographic Latitude, longitude	Magnitude (M$_W$)	Focal depth (km)
2015-04-25	06:11:26	28.147°N 84.708°E	7.8	15.0
2015-04-25	06:45:21	28.193°N 84.865°E	6.6	14.6
2015-04-26	07:09:10	27.782°N 85.997°E	6.7	17.3

Table 2. The main detected anomalies for Nepal earthquake using the implemented methods. *Value* is calculated by $p = \pm100\times(|Dx|-k)/k$. Day is relative to the earthquake day.

Method	Day	Time (UTC)	Value %
Median	0	16:00; 18:00	8.40; 4.13
	-2	14:00; 16:00; 18:00	18.68; 15.97; 0.32
ANN	-2	14:00	8.14
	+4	22:00	38.49
LSSVM	-2	2:00; 24:00	-26.18; 30.89
	0	20:00	0.5
	+1	16:00	0.71
	+3	12:00; 14:00; 16:00; 18:00; 20:00	67.20; 9.10; 15.45; 26.61; 7.89
	+4	22:00	-1.8
	+5	22:00	30.89
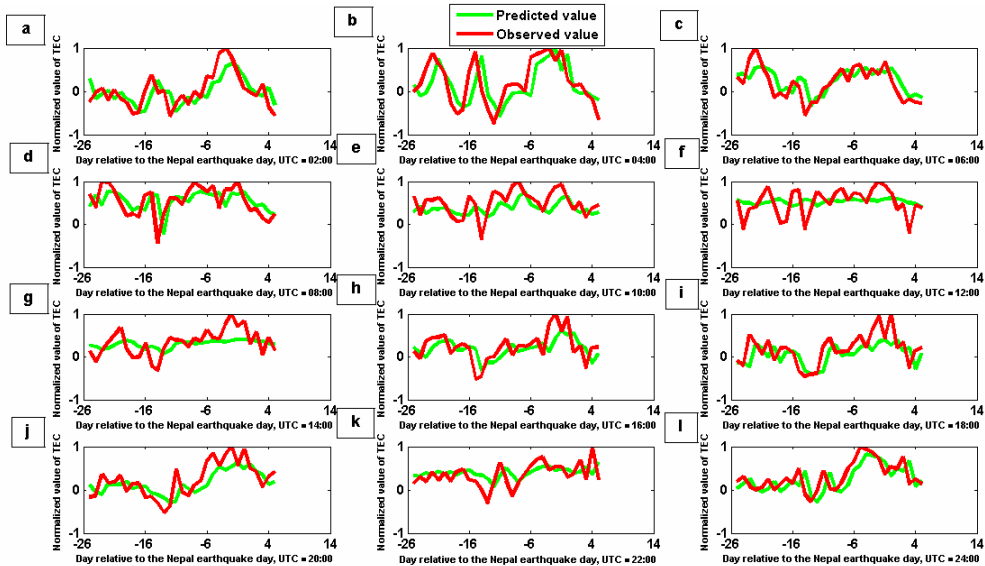
Fig. 3. a) through l) Variations of the observed (red curve) and predicted (green curve) TEC values obtained from ANN method during at different universal times. The X-axis represents the day relative to the Nepal earthquake day.
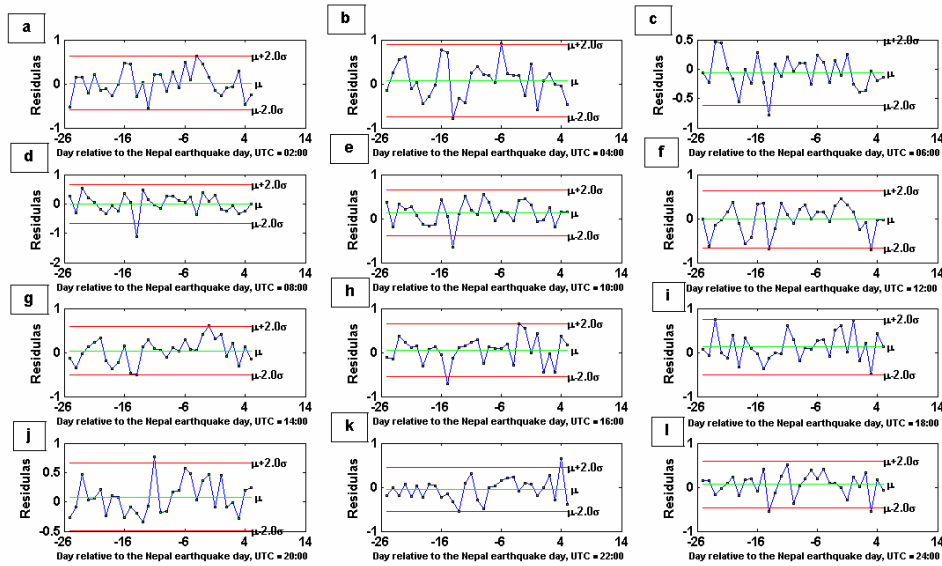
Fig. 4. a) through l) Variations of the differences between the observed and the predicted values of TEC obtained from ANN method on days selected as testing set at different universal times. The red horizontal lines indicate the upper and lower bounds ($\mu \pm 2 \times \sigma$). The green horizontal line indicates the mean value (μ). The X-axis represents the day relative to the Nepal earthquake day.
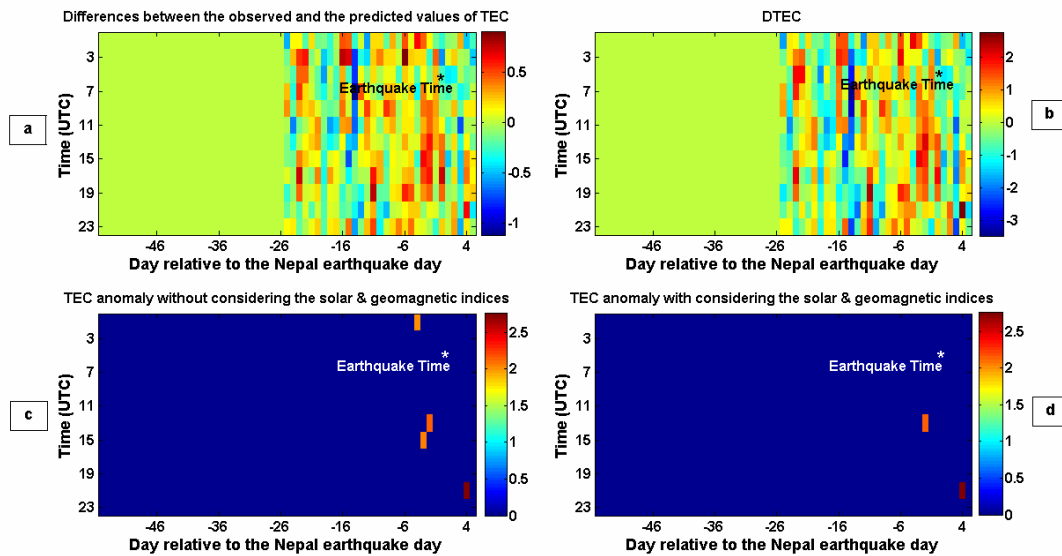

Fig. 5. a) Differences between the observed and the predicted values of TEC obtained from ANN method. b) DTEC variations. c) Detected anomalies using ANN method without considering the geomagnetic indices. d) Detected anomalies using ANN method with considering the geomagnetic indices. The X-axis represents the days relative to the Nepal earthquake day. The Y-axis represents the universal time coordinate.
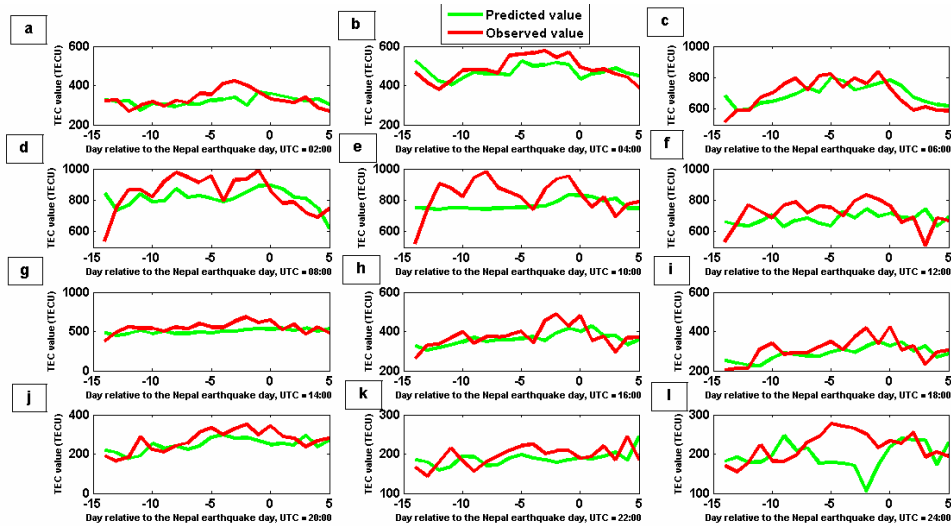
Fig. 6. a) through l) Variations of the observed (red curve) and predicted (green curve) TEC values obtained from LSSVM method during at different universal times. The X-axis represents the day relative to the Nepal earthquake day.
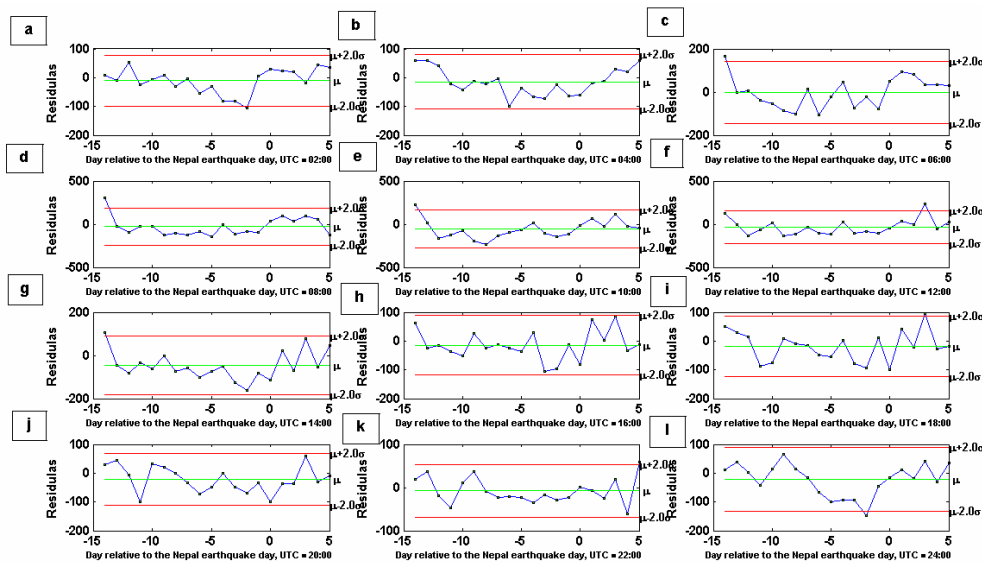

Fig. 7. a) through l) Variations of the differences between the observed and the predicted values of TEC obtained from LSSVM method on days selected as testing set at different universal times. The red horizontal lines indicate the upper and lower bounds ($\mu \pm 2 \times \sigma$). The green horizontal line indicates the mean value (μ). The X-axis represents the day relative to the Nepal earthquake day.
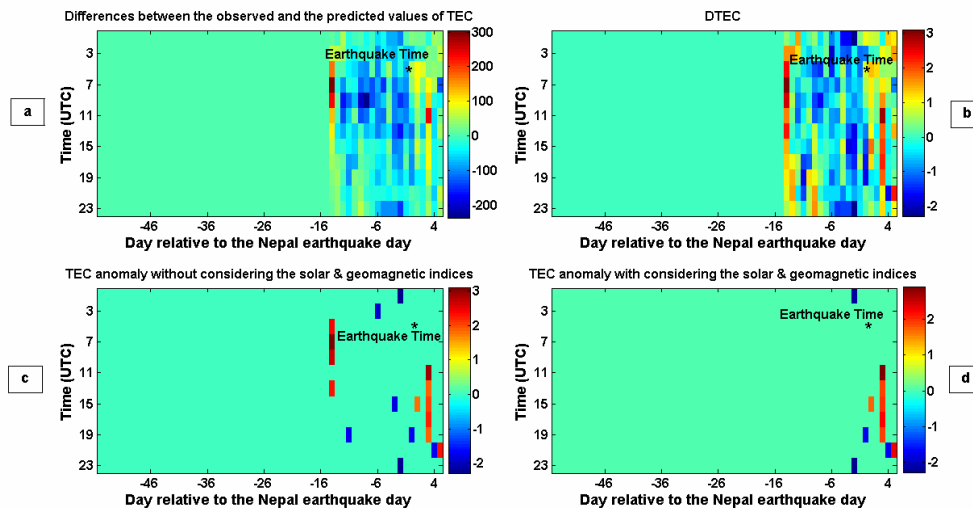
Fig. 8. a) Differences between the observed and the predicted values of TEC obtained from LSSVM method. b) DTEC variations. c) Detected anomalies using LSSVM method without considering the geomagnetic indices. d) Detected anomalies using LSSVM method with considering the geomagnetic indices. The X-axis represents the days relative to the Nepal earthquake day. The Y-axis represents the universal time coordinate.

20:00 UTC. Other post-seismic anomalies could be associated with the strong aftershocks (Table 1).

4. CONCLUSIONS

This paper attempts to acknowledge the ability of LSSVM as a good predictor to forecast the GPS-TEC variations around the time and location of Nepal earthquake. Also two classical and intelligent methods including Median and ANN have been implemented. Using all applied methods prominent TEC anomalies are observed 2 days prior to the main shock. Results reveal that LSSVM method is promising for TEC sesimo- ionospheric anomalies detection.

REFERENCES

Akhoondzadeh, M., 2011. Comparative study of the earthquake precursors obtained from satellite data. *PhD thesis*, University of Tehran, Surveying and Geomatics Engineering Department, Remote Sensing Division.

Akhoondzadeh, M., 2013. A MLP neural network as an investigator of TEC time series to detect seismo-ionospheric anomalies. *Advances in Space Research*, 51, pp. 2048-2057.

Hayakawa, M., and Molchanov, O.A., 2002. Seismo-Electromagnetics: Lithosphere-Atmosphere-Ionosphere Coupling. *Terra Scientific Publishing Co.* Tokyo, 477.

Liu, J.Y., Chuo, Y.J., Shan, S.J., Tsai, Y.B., Pulinets, S.A., and Yu, S.B., 2004. Pre-earthquake-ionospheric anomalies registered by continuous GPS TEC. *Ann. Geophys.*, 22, pp. 1585-1593.

Parrot, M., 1995. Use of satellites to detect seismo-electromagnetic effects, Main phenomenological features of ionospheric precursors of strong earthquakes. *Advances in Space Research,* 15 (11), pp. 1337-1347.

Pulinets, S. A., and Boyarchuk, K.A., 2004. Ionospheric Precursors of Earthquakes. Springer, Berlin.

Suykens, J. A. K., Vandewalle, J., 1999. Least squares support vector machine classifiers. *Neural processing letters*, 9 (3), pp. 293-300.

Wang, S., and Shang, W., 2014. Forecasting Direction of China Security Index 300 Movement with Least Squares Support Vector Machine. *Procedia Computer Science*, 31, pp. 869-874.

Zhou, J., Shi, J., Gong, L., 2011. Fine tuning support vector machines for short-term wind speed forecasting. *Energy Conversion and Management*, 52, pp. 990–1998.