



Search for earthquake-related perturbations in GPS-derived-TEC at Agra using support vector machine (SVM) algorithm

Swati*⁽¹⁾, Devbrat Pundhir⁽¹⁾, & Birbal Singh⁽²⁾

(1) Seismo-electromagnetics & Space Research Laboratory (SESRL), Department of Physics, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India-283105

email: swatirbs2709@gmail.com; devbratpundhir@gmail.com

(2) Seismo-electromagnetics & Space Research Laboratory (SESRL), Department of Electronics & Communication Engineering., Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India-283105, email: bbsagra@gmail.com

In the recent era, lots of work has been done on the identification and characterization of earthquake-related variations by applying well-established statistical techniques to ionospheric measurements but still, there are many imperfections in the identification of the true precursor of earthquake because the ionosphere is a highly complex and it can also be perturbed by others i.e. magnetic storms, solar flare, solar activity dust storms, etc. So it is very important to separate the impact of these parameters for the significant results which is not possible by using the conventional techniques.

With the advancement in technology and modern numerical tools the recompilation of more data, an opportunity that earthquake forecasting to improve its success probability exists. One data-driven solution to accurate forecasting of earthquake events is machine learning, which brings patterns in large datasets to the surface by finding complex mathematical relationships within the data. Algorithms that are commonly used to tackle classification and regression problems include Support Vector Machine (SVM) as one of the methods. SVM is a supervised machine learning approach based on the kernel method, often used for classification and regression missions. As we have employed the SVM to classify the ionospheric TEC enhancements before the earthquake events.

In this study, we have examined the GPS-TEC observed at Agra Station (27.2°N, 78°E) from the years 2010 to 2013. For this study, we have considered a total of 117 cases of earthquakes of magnitude $M \geq 4.5$ within an outer radius of 1000km around the Agra station. After checking the availability of data at our station, a total of 87 earthquakes ($M \geq 4.5$) are shortlisted for this study. The GPS-TEC data have been classified into two groups corresponding to these earthquakes. In the first group, GPS-TEC data of sixteen days before the occurrence of the earthquakes is to be considered anomalous data including the day of the earthquake also. In the second group, a few quiet days have been considered from the same dataset which is not influenced by any earthquake as well as due to other activity (i.e. magnetic storms, solar activity, dust storms, etc.). Further, GPS-TEC data have been analyzed by SVM using the Classification Learner app in MATLAB. Precisely, Linear SVM and Medium Gaussian SVM are used to analyze the dataset, as well as statistical techniques, are also used. The influence of other activities like magnetic storms, solar activity, etc. on GPS-TEC data is also considered.

The confusion matrixes are obtained by using the Linear and Medium Gaussian SVM and then, precision, accuracy, and recall are calculated mathematically. A total of 793 observations for the anomalous period and 515 for quiet days are obtained for the analysis and 90% of the data was used for training and 10% for the test set with the help of the 5-fold cross-validation method. For the training set, our model achieved an accuracy of 96.5% in the case of linear SVM, the precision of 97.9%, recall of 98.4%, and accuracy of 96.5%. For the test set our model achieved an accuracy of 61.3%, precision of 94.12%, recall of 63.16%, and accuracy of 61.25%. Further, in the case of medium Gaussian SVM, for the training set, our model achieved an accuracy of 96.2%, precision of 98.6%, recall of 97.36%, and accuracy of 96.21%. For the test set our model achieved an accuracy of 61.3%, precision of 98.03%, recall of 64.1%, and accuracy of 63.75%. The results interpret that GPS-TEC data observed at Agra station is significantly classified for the earthquake prediction purpose. Our model provides better results during training in comparison to the testing so it needs some modification for better results during real-time prediction.