

A Short-term foF2 Forecasting Method Using A Support Vector Machine Technique

Chun Chen⁺, Shuji Sun and Panpan Ban

National Key Laboratory of Electromagnetic Environment, China Research Institute of Radiowave Propagation, Qingdao, China

Abstract. Using data from seven ionosonde stations, Haikou and Chongqing, Guangzhou, Haikou, Chongqing, Beijing, Lanzhou, Changchun and Manzhouli stations, we have developed an empirical model by using Support Vector Machine, to predict the ionospheric foF2 parameter in storm-time. In this paper, we investigate whether foF2 during disturbed geomagnetic conditions at a single station can be well predicted one hour ahead by using some inputs to an SVM network, such as the latest foF2 observations, hourly quiet-time foF2 ($foF2_{QT}$), time, and the hourly time-weighted accumulation series derived from the geomagnetic index $ap(\tau)$. It indicated that the model described here can capture the low latitude storm time F2 layer variability at most times.

Keywords: Ionospheric storm-time forecasting, ionospheric critical frequency, support vector machine

1. Introduction

The ionosphere influences short-wave communication, space-terrestrial and inter-satellite links, but it is highly variable due to the influence of solar, geomagnetic and other sources. The F2 layer critical frequency, f_0F_2 , is one of the most important ionospheric parameters, controlled by local time, geographical latitude, solar and magnetic activity, the background atmospheric wind and other factors. When intense disturbances occur, the variation of f_0F_2 can reach or exceed 30% of f_0F_2 , and there is consequently a requirement to be able to forecast f_0F_2 during these events. Compared with the quiet ionospheric conditions, the situation is completely different under disturbed conditions related to geomagnetic storm events. However, knowledge of the ionospheric response during geomagnetic storms and its related process remain incomplete. To predict the ionospheric response during storms is a priority task.

It has been long known that ionospheric disturbances are usually related to geomagnetic activities that depend on the local time and season. During a geomagnetic storm, great changes of the electron density are different from the normally day-to-day variability of F region. The ionization density can either increase or decrease. These changes are denoted as negative or positive ionospheric storms, according to the conditions whether foF2 is below or above its “quiet value”, respectively. From several experiments and theoretical studies a phenomenological scenario of the ionospheric response to geomagnetic storms has emerged [e.g. *Buonsanto*, 1999; *Wintoft and Cander*, 2000], which made it possible to predict ionospheric response in F region positive and negative storms over mid-latitude stations several hours in advance using the auroral electrojet index (AE).

Currently available ionospheric forecasting models have shown a high degree of reliability during quiet conditions but they have proven inadequate during storm events [Cander, 2003]. The problem of forecasting the ionospheric disturbances associated with geomagnetic storms has already been studied since 1970s. Many geomagnetic indices were studied to find which of them could best forecast the ionospheric response to

⁺ Chun Chen
chenchun_qaz@yahoo.com.cn

geomagnetic storms [Mendillo, 1973]. Wrenn et al. [1987] used relative changes in foF2 data with respect to estimated quiet-time values as an ionospheric disturbance index (IDI), which depends on the geomagnetic activity to define a predictive scheme for foF2. Araujo-Pradere et al. [2002] developed an empirical storm-time correction model (STORM), which was designed to scale the quiet-time foF2 to account for storm-time changes in the ionosphere. This model was incorporated in the International Reference Ionosphere [Bilitza, 2001]. It provides an estimate of the expected change in the ionosphere during a period of increased geomagnetic activity. The STORM provides an output to “adjust” the quiet-time foF2 by calculating the correction factors.

The support vector machine (SVM) combines the training efficiency and simplicity of linear algorithms with the non-linear algorithms [Vapnik, 1995]. Owing to the toleration of high-dimensional and incomplete data, the SVM has been successfully applied in various fields, including the phonetic and literal recognition and the time-series prediction. In particular, Gavrishchaka and Ganguli [2001] used the SVM to model the solar wind-driven geomagnetic substorm activity characterized by the AE index.

2. The Support Vector Machine

SVMs have recently gained significant interests due to its excellent results in various applications. A brief description of the SVM [Vapnik, 1995, 1999; Cherkassky and Ma, 2004] is given as follows.

In SVM regression, the input x is first mapped onto an m -dimensional feature space using some fixed (non-linear) mapping, and then a linear model is constructed in this feature space. The linear model $f(x, w)$ is given by

$$f(x, w) = \sum_{i=1}^N w_i g_i(x) + b \quad (1)$$

Where $g_i(x)$, $i=1, \dots, N$, denotes a set of non-linear transformation, w_i , $i=1, \dots, N$, is a set of weight, b is the bias, and N the dimension of the feature space.

The regression estimates can be obtained by minimizing the empirical risk on the training data. Typical loss functions used for minimizing the empirical risk include the squared error and absolute error. The SVM regression uses a new type of loss function called ε -insensitive loss proposed by Vapnik [1999]:

$$|y - f(x)|_{\varepsilon} = \begin{cases} 0, & |y - f(x)| < \varepsilon \\ |y - f(x)| - \varepsilon, & \text{otherwise} \end{cases} \quad (2)$$

SVM regression performs linear regression in the high dimensional feature space using ε -insensitive loss and, at the same time, tries to reduce model complexity by minimizing $\|w\|^2$: This can be described by introducing (non-negative) slack variables ξ_i, ξ_i^* , $i=1, \dots, N$, to measure the deviation of training samples outside ε -insensitive zone. Thus, SVM regression is formulated as minimization of the following functional:

$$\text{minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*), \quad (3)$$

where C is a pre-specified value, under the constrains

$$\begin{aligned} y_i - f(x_i, w) &\leq \varepsilon + \xi_i, & i = 1, 2, \dots, N \\ f(x_i, w) - y_i &\geq \varepsilon + \xi_i^*, & i = 1, 2, \dots, N \\ \xi_i &\geq 0, & i = 1, 2, \dots, N \\ \xi_i^* &\geq 0, & i = 1, 2, \dots, N \end{aligned} \quad (4)$$

ξ_i, ξ_i^* represent upper and lower constraints on the outputs. The first step involving equation (3) is to minimize the Vapnik-Chervonenkis (VC) dimension, a parameter representing the complexity of the model. The second step involves minimizing the errors between the regression and target values, where the target values are the training samples.

To solve this optimization problem with constrains of inequality type one has to find a saddle point of the Lagrange functional :

$$\begin{aligned}
L(w, \xi_i, \xi_i^*, \alpha_i, \alpha_i^*, \beta_i, \beta_i^*) = & \\
\frac{1}{2} w^T w + C \sum_{i=1}^N (\xi_i + \xi_i^*) & \\
- \sum_{i=1}^N \alpha_i^* [y_i - w_i^T \varphi(x_i) - b + \varepsilon + \xi_i] & \quad (5) \\
- \sum_{i=1}^N \alpha_i [w_i^T \varphi(x_i) + b - y_i + \varepsilon + \xi_i] & \\
- \sum_{i=1}^N (\beta_i \xi_i + \beta_i^* \xi_i^*) &
\end{aligned}$$

The solution of this minimization problem can be solved to find coefficients $\alpha_i, \alpha_i^*, \beta_i, \beta_i^*, i = 1, \dots, N$, that maximize the quadratic form L_d . L_d is modified to

$$\begin{aligned}
L_d(\alpha, \alpha^*) = -\varepsilon \sum_{i=1}^N (\alpha_i^* - \alpha_i) & \\
- \frac{1}{2} \sum_{i,j=1}^N (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) \varphi(x_i)^T \varphi(x_j) & \quad (6)
\end{aligned}$$

subject to constraints

$$\begin{aligned}
\sum_{i=1}^N \alpha_i^* &= \sum_{i=1}^N \alpha_i \\
0 \leq \alpha_i^* &\leq C, \quad i = 1, 2, \dots, N \\
0 \leq \alpha_i &\leq C, \quad i = 1, 2, \dots, N
\end{aligned} \quad (7)$$

The optimization problem can be transformed into a dual problem, whose solution is given by

$$f(x) = \sum_{i=1}^N (\alpha_i^* - \alpha_i) x_i^T x + b, \quad (8)$$

where the dual variables are subject to constraints $0 \leq a, a^* \leq C$, b is the bias.

For non-linear regression problems, the training data can be mapped to a high-dimensional feature space H by a process, where the linear regression is performed. The kernel function $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$ is used here instead of $x_i^T x_j$ in linear regression problems. The Curb kernel function is selected as the kernel function in this study. This regression function is expressed as

$$f(x) = \sum_{i=1}^N (\alpha_i^* - \alpha_i) K(x_i, x) + b. \quad (9)$$

The major advantage of the kernel-based machine is that it decouples the number of free parameters (related to the machine capacity) from the size of the input space N , which can be very large or even infinite. Once the kernel function is chosen, kernel representation allows one to learn effectively by choosing the underlying map and dimension N [Gavrishchaka and Ganguli, 2001].

An SVM is a combination of a kernel-based architecture and a structural risk minimization (SRM) principle. The SVM minimizes an upper bound on the expected risk, which provides solid theoretical grounds for optimizing the generalization ability of SVM. The expected errors (risk) of the trained machine when applied to test data are bounded by the sum of two terms. The first term is the empirical error (risk) given by the mean error on the training set. The second is a function of the VC dimension, which is the measure of the machine capacity (i.e., the ability to learn a relation of a certain complexity). The VC dimension is related to the number of free parameters [Valerity and Ganguli, 2001]. The SVM performs a mapping process from the input to a high-dimensional feature space by using a kernel function, and finds out the relations between the input and the targets.

3. Datasets and model construction

Recent investigations have provided some insight and understandings in the ionospheric response to geomagnetic activity, which provides a useful step in characterizing ionospheric response to storms in a

relatively simple way. For example, taking into account a non-linear dependence of the integral of the ap and the ionospheric response, STORM model provides the correction for perturbed conditions by using data from 75 ionosonde stations and 43 storm events [Araujo-Pradere *et al.*, 2002]. Under the assumption that the ionospheric disturbance index is correlated to the integrated geomagnetic index $ap(\tau)$, a local model over Rome in mid-latitude region has been developed to forecast foF2 during geomagnetic storms and disturbed ionospheric conditions [Pietrella and Perrone, 2008].

For this purpose the hourly foF2 observed at Lanzhou station are used in this paper. Depending on the availability, the training data are from January 1958 and December 2000. The data from the years 2001 to 2006 are used to test the SVM capability and calculate forecast precision quantitatively.

4. Results and discussion

The optimal networks are tested to examine the expected performance using the test set during geomagnetic storms occurred from 2001 to 2006. The accuracy of the forecast is discussed in terms of the RMS error. The forecasted foF2 of the ELIFMOL are compared with the monthly median values and those of STORM_{MEDIAN} model, the persistence model and the STORM_{foF2QT} model, where the STORM_{MEDIAN} model and the STORM_{foF2QT} model are global models and the persistence model (Persistence) is a single station model. The comparisons show that under disturbed ionospheric conditions, the performance of ELIFMOL is always better than that of STORM_{MEDIAN}, STORM_{foF2QT} and Persistence.

Figure 1 shows the comparison between the observed foF2 (thin line) and forecasted foF2 (ELIFMOL, dashed line; STORM_{MEDIAN}, star marker; the monthly median values, dash-dotted line; STORM_{foF2QT}, dot marker) under disturbed ionospheric conditions for 2-4 October 2001 (moderately disturbed day). For this case, the day of 3 October 2001, holding a main phase of a storm, is characterized by moderately disturbed day with $ap(\tau=0.9)=43.22$. During the second half of the day the ionosphere shows a positive storm (compared the measurements with medians) with $ap(\tau=0.9)$ values increasing gradually. It can be found that STORM underestimates the foF2 values and bring a pseudo negative storm. However, ELIFMOL does give an acceptable positive storm prediction. It is also worthy known that during the first half of the day just before the initial phase of the current storm, i.e. part of the recovery phase of the former storm, the ionosphere almost keeps in quiet. This evolution is captured by ELIFMOL, but the foF2 values are also been underestimated by STORM_{MEDIAN} and STORM_{foF2QT}. The performance of ELIFMOL during this day is better than that of STORM_{MEDIAN} and STORM_{foF2QT}.

The foF2 is highly variable at time scales ranging from decades to seconds with the occurrence of ionospheric disturbances associated with geomagnetic storms. Unfortunately, the input time series does not contain sufficient information about the dynamics of these events to model their rapid changes accurately. For example, the information of the solar wind and the IMF, causing the geomagnetic activity, is not included in the forecasting procedure. The knowledge of the ionospheric response to the variation of the solar wind and IMF during geomagnetic storms is required to the development of a successful storm forecasting algorithm.

5. Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant Nos. 40974092, 61032009). The sunspot number and ap data are provided by the World Data Center for Geomagnetism at Kyoto University, Japan, and the ionospheric data by the Data Center of China Research Institute of Radiowave Propagation.

6. References

- [1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. (*references*)
- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

- [3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [4] K. Elissa, "Title of paper if known," unpublished.
- [5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.
- [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetism Japan, p. 301, 1982].
- [7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
Electronic Publication: Digital Object Identifiers (DOIs):
- [8] Article in a journal:
D. Kornack and P. Rakic, "Cell Proliferation without Neurogenesis in Adult Primate Neocortex," Science, vol. 294, Dec. 2001, pp. 2127-2130, doi:10.1126/science.1065467.
- [9] Article in a conference proceedings:
H. Goto, Y. Hasegawa, and M. Tanaka, "Efficient Scheduling Focusing on the Duality of MPL Representatives," Proc. IEEE Symp. Computational Intelligence in Scheduling (SCIS 07), IEEE Press, Dec. 2007, pp. 57-64, doi:10.1109/SCIS.2007.357670.

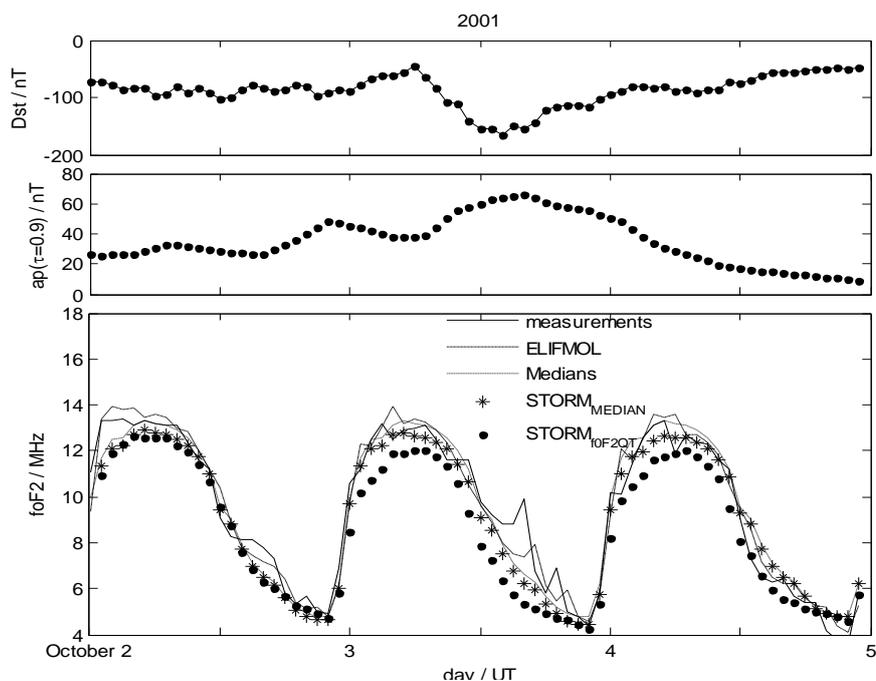


Figure 1. Comparison of the observed foF2 (thin line) and forecasted foF2 (ELIFMOL, dashed line; STORMMEDIAN, asterisks; the monthly median values, dotted line; STORMfoF2QT, dots) for disturbed ionospheric conditions on 2- 4 October 2001.