Communications and Network, **2013, 5, 11-14**
http://dx.doi.org/10.4236/cn.2013.53B2003 Published Online September 2013 (http://www.scirp.org/journal/cn)

Application of Grey Theory to Ionospheric Short-term Forecasting*

Yao Xiao, Zhonghui Gan, Yunjiang Liu, Man Li
Information and Navigation College, Air Force Engineering University, Xi'an 710077, China
Email: xiaozzi2014@sina.con

Received March, 2013

ABSTRACT

By analysis of historical data of the ionosphere, it is suggested to apply grey theory to ionospheric short-term forecasting, grey range information entropy is defined to determine the optimum grey length of the sample sequence, the prediction model based on residual error is constructed, and the observation data of multiple ionospheric observation stations in China are adopted for test. The prediction result indicates that the average grey range information entropy calculation results reflect the cyclical effects of solar rotation, precision of the forecasting method in high latitudes is higher than low latitudes, and its error is large relatively in more intense solar activity season, the effect of forecasting 1 day in advance of average relative residuals are less than 1 MHz, the average precision is more than 90%. It provides a new way of thinking for the ionospheric foF2 short-term forecast in the future.

Keywords: Ionosphere; Short-term Forecasting; Grey Theory

1. Introduction

Ionosphere is an important constituent part of solar terrestrial space. Owing to the influence from solar wind and geomagnetic field at the top and the impact from middle atmosphere at the bottom, ionosphere always changes along with time and space, and there also exist daily variation and time variation except the general diurnal variation, seasonal variation and 11-year solar cycle variation.

The critical frequency foF2 of the ionosphere F2 layer is one of the most important parameters of the ionosphere, measurement and prediction of this parameter are quite significant, and are also the hotspot of relevant domestic and overseas professional researches. For a long time, scholars from home and abroad have done a large number of studies on ionospheric short-term prediction, and raised multiple methods [1]: The autocorrelation method [2] using linear filter to deal with observation data forecast. Multiple linear regression method [3] using a large number of observation data training correlation coefficient forecast. The artificial neural network method [4, 5] simulates ionospheric nonlinear change process, and forecast scale agile. The ionosphere correction model during disturbance [6] takes full advantages of influence factors like geomagnetic latitude and season at the observation point to correct the prediction result. The integrated model [7] reasonably determines the weights of different forecast methods, and gives full play to the characteristics of each method forecast. Based on analysis of a large amount of historical data about foF2, this paper attempts to achieve the short-term forecasting based on grey theory

*Project supported by the Natural Science Foundation of Shaanxi Province (SJ08-2T06)

2. The Basic Principle of Grey Theory

Grey theory arose aimed at uncertainty issues like a small quantity of data and inexperience, and GM (1,1) model is the core of the grey theory prediction model[8], its working principle as shown in **Figure 1** shows.

In **Figure 1**, the parameter a to be estimated is the development coefficient, and b is the grey action. $z^{(1)}(k)$ is the white background value, and the value is the generation sequence of mean value near $x^{(1)}(k)$. Since $x^{(0)}(k)$ is the measured datum, it is the "whitening effect", and therefore, the action mechanism of GM(1,1) model is in line with grey cause whitening effect law. The specific expression such as Equation (1) shows.

$$x^{(0)}(k)+az^{(1)}(k)=b \tag{1}$$

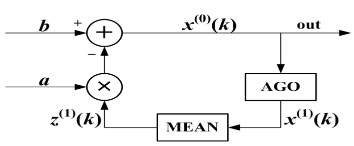

Figure 1. GM(1,1) model working principle diagram.

3. Grey Theory Modeling Analysis

Variation of the ionosphere is a complex and highly nonlinear process, and its variation rule cannot be described precisely by analytical methods, but its change within a short period is relatively stable. In a short time, the variation of foF2 has some certain correlation which can be used for prediction [9]. The inter-record gap of historical data about foF2 cannot be infinitely small, and it cannot be forecast by establishing white system, so the historical data used for prediction in a short time can be regarded as grey. Therefore, foF2 short-term forecasting research based on grey theory can be deployed.

3.1. Grey Range Information Entropy

The size of the selected sample data is related to the efficiency and precision of the prediction model. The recorded foF2 daily average is arranged as an equally spaced discrete time sample sequence *S*, such as Equation (2) shows:

$$S = \{ s_{-N}, s_{-N+1}, \ldots, s_{-2}, s_{-1} \} \quad (2)$$

In Equation (2), *N* is grey forecast length.

The mean value $\bar{S}$ of the sequence *S* is taken as the reference value, and the grey range measure can be adopted for data analysis according to the definition of grey relational coefficient and norms, to determine the optimum value of *N*.

The grey range measure of s_{-i} and $\bar{S}$ in the discrete sample is defined, as is shown in Equation (3):

$$Gd(s_{-i}, \bar{S}) = \frac{\xi \cdot \| d(S, \bar{S}) \|_\infty}{| s_{-i} - \bar{S} | + \xi \cdot \| d(S, \bar{S}) \|_\infty} \quad (3)$$

In Equation (3), ξ is the resolution coefficient, $\xi \in (0,1]$.

Based on information theory, the grey range information quantity is defined, as is shown in Equation (4):

$$GI(s_{-i}) = -\ln Gd(s_{-i}, \bar{S}) \quad (4)$$

The grey range information entropy of the sample sequence *S* is presented in Equation (5):

$$GH(S) = -\sum_{i=1}^{N} \frac{Gd(s_{-i}, \bar{S})}{\sum_{j=1}^{N} Gd(s_{-j}, \bar{S})} \cdot GI(s_{-i}) \quad (5)$$

For the sample sequence S, the smaller the GI(s-i) of the point datum s-i is, which is the smaller the grey range information quantity is, the smaller the uncertainty among the sample point data is [8]. The value magnitude of GH(S) reflects the average uncertainty of data in the sample sequence, and t the optimum value of N is taken from the sample sequence to minimize the value of GH(S).

3.2. Based on the Residual Error Correction GM(1,1) Model

S^1 is set as AGO sequence of the sample sequence *S*, then, according to the Equation (1) the sample sequence *S* prediction model of mathematical expression such as Equation (6) shows:

$$s(k) + a \cdot p^1(k) = b \quad (6)$$

In Equation (6), $s(k)$ is s_{-k} in the sample sequence *S*. $p^1(k)$ is the white background value, and the value is the generation sequence of mean value near S^1.

In addition, suppose $\hat{a} = (a, b)^T$ as the parameter sequence, and set:

$$X = \begin{bmatrix} s(2) \\ s(3) \\ \cdots \\ s(N) \end{bmatrix}, \quad C = \begin{bmatrix} -p^1(2) & 1 \\ -p^1(3) & 1 \\ \cdots & \cdots \\ -p^1(N) & 1 \end{bmatrix} \quad (7)$$

Therefore, the least square estimation parameter sequence of the Equation (6) is shown in Equation (8):

$$\hat{a} = (C^T C)^{-1} C^T X \quad (8)$$

The GM(1,1) model residual error is defined as is shown in Equation (9):

$$\varepsilon(k) = s(k) - \hat{s}(k) \quad (9)$$

In Equation (9), $\hat{s}(k)$ is the predicted value of the GM(1,1) model built by the sample sequence *S*.

In order to increase the prediction precision of GM(1,1) model, the GM(1,1) model of residual error value is established in this paper, and then the predicted value $\hat{\varepsilon}(N+m)$ of the residual error GM(1,1) is added into the original predicted value $\hat{s}(N+m)$, to correct the GM(1,1) model built by the sample sequence *S*.

The residual error sequence $\{\varepsilon(k)\}$ is truncated with a length of *N*-*l*, and the residual error end-piece sequence $\varepsilon_{\text{tail}}$ that can be used for modeling is obtained, as is indicated in Equation (10).

$$\varepsilon_{\text{tail}} = \{ |\varepsilon(l+1)|, |\varepsilon(l+2)|, \ldots, |\varepsilon(N)| \} \quad (10)$$

In a similar way, GM(1,1) model is established for the residual error truncation sequence $\varepsilon_{\text{tail}}$, and its time response function is shown in Equation (11):

$$\varepsilon_{\text{tail}}(k+1) = a_\varepsilon \cdot [b_\varepsilon / a_\varepsilon - \varepsilon(l)] \cdot e^{-a_\varepsilon (k-l)} \quad (11)$$

The residual error truncation prediction sequence $\hat{\varepsilon}_{\text{tail}}(k)$ is adopted to correct the prediction sequence $\hat{s}(k)$ of $s(k)$, and the corresponding corrected prediction value can be gained by deduction, as is shown in Equation (12):

$$\hat{s}_\varepsilon(k+1) = \begin{cases} (1 - e^a) \cdot (s(1) - \frac{b}{a}) \cdot e^{-ak} & k < l \\ (1 - e^a) \cdot (s(1) - \frac{b}{a}) \cdot e^{-ak} + \hat{\varepsilon}_{\text{tail}}(k+1) & k \ge l \end{cases} \quad (12)$$

In Equation (12), $\hat{\varepsilon}_{\text{tail}}(k+1)$ has the same symbol as the residual error end piece $\varepsilon(k+1)$.

The predicted value of appointed time zone can be obtained via GM(1,1) model, and the predicted value of the point m after the point N is

$$\hat{s}_{\varepsilon}(k) \to \hat{s}_{\varepsilon}(N+m),\ m = 1, 2, \ldots.$$

3.3. Error processing

According to the residual error value $s(k) - \hat{s}_{\varepsilon}(k)$ of the residual error GM(1,1) model, the relative residual error $\Delta(k)$ and average relative residual error $\Delta(avg)$ of this model can be defined, as is presented in Equation (13).

$$\begin{cases} \Delta(k) = \dfrac{s(k) - \hat{s}_{\varepsilon}(k)}{s(k)} \cdot 100\% \\ \Delta(avg) = \dfrac{1}{n-1}\sum_{k=2}^{N} |\Delta(k)| \end{cases} \tag{13}$$

In order to estimate the mean square error of the predicted value, set:

$$Q = (C^T C)^{-1}, \quad \delta = \sqrt{\frac{E \cdot E^T}{N-1}} \tag{14}$$

In which, $E = \left(\varepsilon'(1), \varepsilon'(2), \ldots, \varepsilon'(N)\right)$, and by deduction, the mean square error estimation Equation of the predicted value is shown in Equation (15):

$$\delta_{\hat{s}}(k+1) = \delta e^{-ak}\begin{bmatrix} (akp^1(1) - p^1(1) - kb)^2 \cdot Q_{11} + \\ Q_{22} + 2(akp^1(1) - p^1(1) - kb) \cdot Q_{12} \end{bmatrix}^{\frac{1}{2}} \tag{15}$$

The precision ρ of the predicted value can be expressed as:

$$\rho = (1 - \frac{\delta_{\hat{s}}(k+1)}{\hat{s}(k+1)}) \cdot 100\% \tag{16}$$

4. Forecasting Results and Analysis

The foF2 data adopted for analysis come from literature [10], ionospheric observation stations are listed in **Table 1**.

As is shown in Equation (5), the foF2 observed data recorded by these 4 observation stations in 2009 are calculated and tested, and the average values of the annual grey range information entropy under different grey forecast length values of N are presented in **Figure 2**.

In **Figure 2**, two valley values appear for the grey range information entropy, and they lie in two value intervals of N which are [6,9] and [26,28] respectively. Therefore, when the grey forecast length value of N is taken from [6,9] and [26,28], uncertainty among the sample sequence data can be regarded as the minimum. Since in the interval of [6,9], *GH* value vibrates and is instable, the grey forecast length value of N is set as 27 in this paper, which also reflects the 27-day periodic influence of solar rotation on the earth.

Table 1. Locations of the ionospheric stations.

Station Name	Northern Latitude/(°)	Eastern Longitude/(°)
Mohe(MH)	53.48	122.37
Beijing(BJ)	39.92	116.46
Wuhan(WH)	30.52	114.31
Sanya(SY)	18.15	109.30

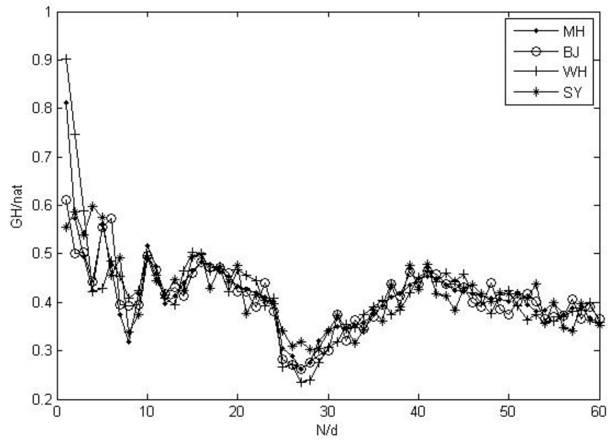

Figure 2. Average grey range information entropy of data from observation stations in 2009.

The residual error GM(1,1) model is established according to the measured data from stations in 2009 respectively, the prediction error is analyzed according to spring, summer, autumn and winter, and the average relative residual error $\Delta(avg)$ and the predicted value precision ρ of the prediction model 1 day and 2 days in advance in different seasons from different ionospheric observation stations are listed in **Table 2** and **Table 3**.

It can be gained from **Table 2** and **Table 3** that error of ionospheric short-term prediction is similar to variation characteristics of the ionosphere, and both of them are related to geographical positions, seasons and solar activities. Under the calculation of grey theory, the residual error and precision of the predicted value changes with variation of geographical positions, seasons and solar activity levels. The residual error of the prediction in autumn is large and the precision is low, which is related to the solar activity level. The residual error in low-latitude areas (Sanya Station) is larger than that in middle-latitude areas, which may be caused by the fact that China's low-latitude ionospheric observation stations are near the ionospheric equatorial anomaly hump where the ionospheric variation is severe. In general, the effect

Table 2. The $\Delta(avg)$ of the predicted value in different seasons from stations / MHz.

Station Name	One day in advance forecast				Two days in advance forecast			
	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter
MH	0.39	0.24	0.43	0.41	1.52	1.47	2.05	1.93
BJ	0.13	0.20	0.41	0.33	1.68	1.51	2.15	1.83
WH	0.62	0.53	0.66	0.61	1.89	1.93	2.70	2.22
SY	0.76	0.87	0.94	0.82	2.06	2.28	2.82	2.47

Table 3. The predicted value precision ρ in different seasons from stations / %.

Station Name	One day in advance forecast				Two days in advance forecast			
	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter
MH	96.2	94.5	91.2	95.3	83.1	75.3	72.1	77.4
BJ	97.4	97.9	94.3	96.9	84.4	78.9	74.7	80.2
WH	92.3	93.7	91.5	91.2	75.3	81.1	73.9	74.6
SY	91.6	90.9	90.6	93.7	72.1	73.6	70.7	71.2

of prediction 1 day in advance is better than the effect of prediction 2 days in advance.

5. Conclusions

By analysis of foF2 historical data from multiple ionospheric observation stations, grey theory is applied to short-term prediction, grey range information entropy is adopted to determine the optimum grey value of *N*, GM(1,1) prediction model is constructed, and the actually observed data in 2009 are used for test. It can be known by analyzing the predicted result that the prediction precision in middle-latitude areas is higher than that in low-latitude areas, and when the solar activity is relatively fierce, the prediction precision decreases. This method is simple, practical, feasible, and equipped with prediction precision, so it has some certain value of theoretical direction and engineering application for later studies on ionospheric prediction.

REFERENCES

[1] Y. K. Kong, J. Feng and W. Liu, *et al*., "Development on Short-term Prediction of Iononspheric Parameters," *Progress in Geophys*, Vol. 25, No. 6, 2010, pp. 1968-1976.

[2] R. Y. Liu, J. P. Wang and Y. W. Wu, *et al.*, "Method for Short-term Forecasting of Ionospheric Tatal Electron Content in Chinese Region," *Chinese Journal of Radio Science*, Vol. 26, No. 1, 2011, pp. 18- 24.

[3] R. Y. Liu, J. Wu and B. C. Zhang, "Development on Ionospheric Weather Prediction," *Chinese Journal of Radio Science*, Vol. 19, 2004, pp. 35- 40.

[4] C. Chen, Z. S. Wu, S. J. Sun, *et al*., "Forecasting of Ionospheric foF2 in China Using Neural Network. Chin. J. Space Sci., Vol. 31, No. 3, 2011, pp. 304-310.

[5] L. B. Weng, H. X. Fang, Z. Q. Miao, *et al*. Forecasting of Ionospheric TEC one Hour in Advance by Artificial Neural Network. Chin. J. Space Sci., Vol. 32, No. 2, 2012, pp. 204-208.

[6] D. J. Chen, J. Wu and X. Y. Wang, "A Technology Study of foF2 Forecasting during the Ionospheric Disturbance. Chinese J. Geophys. (in Chinese), Vol. 50, No. 1, 2007, pp. 18- 23.

[7] W. Liu, J. Feng, Q. Y. Kong, *et al*., "Integrated Model of Ionospheric Short-term Forecasting," *Chinese Journal of Radio Science*, Vol. 25, No. 3, 2010, pp. 492- 498.

[8] J. L. Deng, "Grey Prediction and Grey Decision," Wuhan: Huazhong University of Science and Technology Press, 2002.

[9] D. Y. Huang, "Quasi-Real-Time Forecast Method of the Shortwave Circuit MUF," *Journal of China Institute of Communications*, Vol. 12, No. 6, 1991, pp. 103- 105.

[10] W. X. Wan, B. Q. Ning, L. B. Liu, *et al.*, "Nowcasting the Ionospheric Total Electron Content over China," *Progress in Geophys*, Vol. 22, No. 4, 2007, pp. 1040- 1045.