Study on Gas Emission Rate Prediction based on Chaos Analysis

Liu Yang\textsuperscript{a}, Shi Qingjun\textsuperscript{a}, Li Jing\textsuperscript{a}, Ma Huibin\textsuperscript{a}; Liu Desheng\textsuperscript{a} a*

\textsuperscript{a} School of Information and Electronics Technology, Jiamusi University, Jiamusi, Heilongjiang, 154007 China

Abstract

In order to realize the dynamic prediction on gas emission rate and avoid constructing a model, a study is carried out through chaos theory on the gas emission rate in this paper. On the basis of testing and verifying the gas emission rate to have chaotic characteristics, the Cao method is adopted to recognize embedding dimension and the mutual information method is used to recognize time delay, to reconstruct the phase spaces equivalent to the original system. In phase space, the prediction model base on both local region method and global method to realize the short-term prediction on the gas emission rate. The global method based on the BP neural network shows a good performance. Thus, the application of the chaos theory to the prediction on the gas emission rate is feasible.

© 2011 Published by Elsevier Ltd. Open access under CC BY-NC-ND license. Selection and/or peer-review under responsibility of ICAE2011.

Keywords: Prediction on Gas Emission Rate, Chaotic Time Series, Reconstruction of Phase Space, BP Neural Network

1. Introduction

At present, the gas monitoring system installed in the mines of China can only implement a real-time monitoring on the gas emission rate and can't realize the prediction and warning on gas emission rate. The traditional prediction methods such as mine statistics method \cite{1}, different-source method \cite{2} and velocity method \cite{3} all are static. However, the neural network method realizing the dynamic prediction has the shortcomings, such as selecting the influencing factors of gas emission rate subjectively, have difficulties to acquire some accurate parameters as well as the poor ability in the generalization of small samples. The prediction method, based on the chaotic analysis, can recover the chaotic attractor characteristics by analyzing the internal laws of the gas emission rate time series, and hence reconstructs the phase space equivalent to the original dynamical system. Constructing the prediction model in the phase space to realize the dynamic prediction of the gas emission rate is of quite good prediction accuracy and reliability, and therefore is a very important development direction of the prediction on gas emission rate.

* Corresponding author. Tel.; fax. E-mail address: liuyang@guigu.org

Open access under CC BY-NC-ND license.
2. Chaotic identification of gas emission rate time series

Lyapunov index is used to describe the adjacent-track dispersion degree, and identifies the system's chaotic characteristics based on whether the tracks have the characteristics of divergent movement. When the biggest Lyapunov is greater than 0, the time series can be recognized to have the chaotic characteristics. In this paper, through the gas monitoring system in the Hegang Nanshan mine of China, 1500 samples data are collected, and the sampling cycle is 5 minutes.

Use the small-data method [4] to solve the largest Lyapunov index, and the steps are shown below: (1) use Cao method and the mutual information method to respectively solve the embedding dimension (m) and the time delay (τ), and reconstruct the phase space; (2) find the point \( X(i) \) which each point \( X(0) \) is the most adjacent to in the phase space, and calculate the ith discrete time distance \( d_i \) between \( X(t) \) and \( X(i) \); (3) Solve the mean \( y(i) \) of the \( \ln d_i \) of all of t for each \( i \); in the figure of \( i \) and \( y(i) \), the method of the minimum squares is used to make a regression straight line, and the slope of the line is the largest Lyapunov index.

The largest Lyapunov index of the gas emission rate time series is solved to be 0.0053. This verifies the gas emission rate time series have chaotic characteristics.

3. Reconstruction of phase space

Chaos is a phenomenon, which seems to have no rules and to be pseudorandom in a deterministic system. Packard thought that the evolution of any component in a chaotic system is decided by other components which are interacted with it. According to Takens theorem [4], if \( x(t), t = 1, 2, \ldots, N \) is the time series, the embedding dimension (m) and time delay (τ) are selected, to get the reconstructed phase space \( X(t) = [x(t), x(t + \tau), \ldots, x(t + (m-1)\tau)] \) \( (t = 1, \ldots, M; M = N - (m-1)\tau) \). Thus, the phase space reconstruction is selecting appropriate embedding dimension and time delay.

3.1 Embedding dimension

Embedding dimension is an important parameter for the phase space reconstruction. If the embedding dimension (m) is too large, it makes the space too complex and hence increases fit difficulty, and also excessive dimensions can cause greater errors and reduce the prediction accuracy. However, if the m is too small, it can't recover the form of chaotic attractors and can't reflect the original system dynamics characteristics, making the constructed model deviated greatly [5]. Cao method is used to solve embedding dimensions in this paper, as shown below:  

\[
\begin{align*}
E(m) &= \frac{1}{N - m \tau} \sum_{i=1}^{N} a(i,m) \\
a(i,m) &= \left\| \frac{X(m+1) - X_{\text{mix}}(m+1)}{X(m) - X_{\text{mix}}(m)} \right\| \\
X_{\text{mix}}(m+1) &\text{ is the vector that has the most adjacent to } (1 \leq n(i,m) \leq N - m \tau) \\
E(m) &= \frac{1}{N - m \tau} \sum_{i=1}^{N} a(i,m) \\
E(m) &= E(m+1) / E(m), m = m + 1 (m \leq 14)
\end{align*}
\]

If \( E(m) \) increase as embedding dimensions increases and reaches saturation, \( m+1 \) is the best embedding dimensions. As shown in figure, the number of embedding dimensions is 8.
3.2 Time delay

Time delay is another main parameter in the phase space reconstruction. In this paper, the time delay is recognized based on the first minimum of mutual information function, namely, the mutual information method [6]:

1. Construct the 2D phase graphic of gas emission rate time series \( \{x(t)\} \) and order \((x,y)=[x(t),x(t+\tau)]\), \(\tau=1\);
2. divide attractors in the 2D phase graphic into equidistant grids;
3. if \(x_i \leq x(i) \leq x_i + \Delta x, y_j \leq y(j) \leq y_j + \Delta y, i,j=1,2,...,N\), the point \([x(i),y(j)]\) falls in rectangle frame and also meets \((k-1)\Delta x \leq x(i) - x_j < k\Delta x, (l-1)\Delta y \leq y(i) - y_j < l\Delta y, k=1,2,...,M, l=1,2,...,M\), and the point \([x(i),y(j)]\) falls in the small grid \(\Delta_{k,l}\) and the times to fall in the small grid \((k,l)\) is \(N_{xy}\), and also the number of the points to fall in the small grid \(k-1\) to \(K\) is \(N_{x}\) and the times to fall in the \(1-1\) to \(1\) is \(N_{y}\).

\[
p[x(i)] = N_x / N, \quad p[y(i)] = N_y / N, \quad p[x(i), y(i)] = N_{xy} / N, \quad N \text{ is the number of all sampling points }
\]

and can be substituted into the equation below.

\[
H(X) = -\sum_{i=1}^{N} P(x_i) \log P(x_i)
\]

\[
H(Y) = -\sum_{j=1}^{N} P(y_j) \log P(y_j)
\]

\[
l(X,Y) = H(X,Y) - H(X) - H(Y)
\]

The mutual information function value \(I(\tau)\) of the delay time \(\tau\) is solved.

4. Order \(\tau = \tau + 1(\tau \leq 20)\), and return to 2.

As the first minimum of the mutual information function is the best time delay as shown in figure 2, the time delay is solved to be six.

![Figure 2: Mutual Information Function](image)

4. Gas emission rate time series prediction

Per the embedding dimension \(m\) and delay time \(\tau\), a phase space equivalent to the original system is reconstructed. Thus, there is a mapping \(F: R^n \rightarrow R\) to make \(X(t+\eta) = F\{x(t), x(t-\tau), ..., x[(t-(m-1)\tau)]\}\).

In the above, \(\eta\) is the prediction step length. Based on modeling situations, time series prediction is divided into local method and global method.

4.1 The first-order weighted local prediction method

The first-order weighted local prediction method [7] is a local prediction method, which makes the prediction by using multiple neighbour points around the central point to fit \(X(t+\eta) = a + bX(t)\). Each
prediction step length \( \eta \), respectively uses the minimum squares method to calculate related \( a, b \) values and construct the prediction model; the number of neighbour points is \( N > m + 1 \) (\( N=9 \) in this experiment), and the prediction value is separated from the prediction vector \( X(t+1) \). As the first-order weighted local prediction method uses the fit relation of multiple neighbour points to construct the prediction model, it has a good predicting performance and accuracy when compared with using a single neighbor point to fit the largest Lyapunov index prediction model.

### 4.2 Prediction method based on the BP neural network

The prediction method based on the BP neural network is global. In the phase space, the BP neural network is used to fit the mapping relation between \( X(t) \) and \( x(t+h) \) to construct multi-step prediction model \( x(t+h) = f_s(X(t)) \) for the \( h \)th step prediction. The parameters of the BP neural network are below: BP neural network has three layers; the input layer neural number is the number (8) of embedding dimensions; the output layer neural number is 1; the hidden layer neural number trains different neural network; the RMS error and the largest error of the training results are used to evaluate the constructed neural network; ultimately the layer neural number is decided to be 15. As the sample data of the gas emission rate time series is distributed between \( 0 \sim 1 \), the transfer function of hidden layer and output layer selects the logarithmic S-shaped function logsig (Output interval: \( 0 \sim 1 \)). The training function uses the numerically optimized Levenberg-Marquardt to train the function trainlm; the learning function uses the gradient-decreasing dynamics learning function learngdm; the network can evaluate the function to use the mean square error function mse. The prediction result of the gas emission rate by the first-order weighted local method and the BP neural network model is in figure 3.

![Figure 3: Prediction results of first-order weighted local method & BP neural network model](image)

The first-order weighted local prediction method can be implemented easily. If the local prediction model could be constructed in real time and accurately, it can get a good performance, but shows instability in experiment. However, the BP neural network prediction method has a good performance and shows good stability and accuracy within 50 steps, while the model's performance is reduced quickly with the increase of the prediction steps, indicating the prediction method based on the chaos analysis unable to finish a long-term prediction.

### 4.3 Analysis of the prediction model performance

To make an objective evaluation on the prediction model, the root mean square error \( E_{rmse} \) is introduced to evaluate the prediction model.

\[
E_{rmse} = \left\{ \frac{1}{n-1} \sum_{t=1}^{n} [y(t) - x(t)]^2 \right\}^{1/2}
\]

In the equation (5), \( y(t) \) is the expected output series; \( x(t) \) is the actual output series of the prediction model; \( n \) is the number of samples. The \( E_{rmse} \) reflects the degree of error between predicted value and
expected value and its value is larger than or equal to 0, but is 0 if the prediction has no errors. The comparison of $E_{\text{rmse}}$ between two prediction models is as shown in Table 1.

Table 1: Comparison of the Error Function $E_{\text{rmse}}$ between two Prediction Models

<table>
<thead>
<tr>
<th>Predicted Steps</th>
<th>First-order weighted local method</th>
<th>Global Method Based on BP Neural Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Steps</td>
<td>0.0017</td>
<td>0.0008</td>
</tr>
<tr>
<td>20 Steps</td>
<td>0.0059</td>
<td>0.0020</td>
</tr>
<tr>
<td>35 Steps</td>
<td>0.0093</td>
<td>0.0023</td>
</tr>
<tr>
<td>50 Steps</td>
<td>0.0086</td>
<td>0.0023</td>
</tr>
</tbody>
</table>

From Table 1, it can be seen that the global prediction model based on the BP neural network has a good performance; if the predicted steps are less than 50, the results always keep a smaller deviation, and have good, stable overall performance.

5. Conclusion

To realize the dynamic prediction on gas emission rate and prevent the subjectivity in the prediction model construction, chaotic time series analysis methods are used to study the gas emission rate prediction in this paper, and the main results are as shown below. First, gas emission rate time series are with chaos characteristics, so chaos theory can be used to analyze and predict gas emission rate. Second, when the prediction model is being constructed in the reconstructed phase space, the global prediction method based on BP neural network is better than the first-order weighted local prediction method. The global prediction model shows a good performance in step, accuracy and stability predictions. This indicates that the application of chaos theory to the prediction on the gas emission rate time series is feasible. Third, under the influence of the chaos system initial value sensitivity, the accurate short-term prediction can only be made on the gas emission rate.

Fund Project: This paper is supported by Jiamusi University Key Project (No. LZ2010-013)

References


