Dynamic Analysis of Blast Furnace Sensor Data using Cross-recurrence Quantification Strategies

X C Sun, B Wei, J h Gao, J C Fu and Z G Li*

College of Artificial Intelligence, North China University of Science and Technology, Tangshan, Hebei, 063210, China

*Corresponding author’s e-mail: lizhigang@ncst.edu.cn;

Abstract. This paper investigates impact degree of blast furnace related elements towards blast furnace gas (BFG) production. BFG is a by-product in the steel industry, which is one of the enterprise's most essential energy resources. While because multiple factors affect BFG production it has characteristics of large fluctuations. Most works focus on finding a satisfactory method or improving the accuracy of existing methods to predict BFG production. There are no special studies on the factors that affect the production of BFG. Finding the elements that affect BFG production is benefit to production of BFG, which has a significance in economy. We propose a novel framework, combining cross recurrence plot (CRP) and cross recurrence quantification analysis (CRQA). Moreover, it supplies a general method to convert time series of BFG related data into high-dimensional space. This is the first analytical framework that attempts to reveal the inherent dynamic similarities of blast furnace gas-related elements. The experimental results demonstrate that this framework can realize the visualization of the time series. In addition, the results also identify the factor that has the greatest impact on blast furnace gas production by quantitative analysis.

1. Introduction
Last few years a new industrial revolution has emerged, which is called Industry 4.0. The development and technological advancement of Industry 4.0 provides a series of feasible solutions to the growing demand for information technology in the manufacturing industries [1]. Sensors and cognitive techniques are crucial components of Industry 4.0 [2]. Sensors are able to accumulate data and transmit data to database. Cognitive techniques, including automation, machine learning and data analytics, generally are used to evaluate the dynamic complex factors that contribute to production.

BFG is an essential byproduct of steel industry, which is used as auxiliary fuel for power plants, and heat treatment, coking, reheating furnaces [3]. It has characteristics of large output and large fluctuations. Consequently, BFG resource related analysis and predictions have become a noticeable research field. Matinoet al.forecasted blast furnace gas production applying echo state neural networks [4]. Shiet al.implement recognition of blast furnace gas flow center distribution using infrared image processing [5]. From the papers mentioned above, it can be found that most people’s works focus on finding a satisfactory prediction method by the BFG data itself. However, there are no special studies on the factors that affect the production of BFG. Finding out the elements that are mostly related to blast furnace gas production is very meaningful for predicting blast furnace gas production.

In the recent past a new approach called CRQA which concentrate on the analysis of nonlinear data has been popular. It is used to analyze short, noisy and non-stationary data [6]. The recurrence of state is a crucial property of complex dynamical systems. The Recurrence analysis was originally a graphical
approach that is used to demonstrate periodic ways of activity in the system over time. It uses data to generate CRP. CRP is able to visualize inherent relations between two nonlinear time series [7]. CRQA is a qualitative analysis of CRP, quantifying how often two systems show similar changes or movement patterns [8]. Because of this fabulous aspect, CRQA has been used in various fields widely, including medicine, social analytics. For example, Timothy et al., proposed an EEG classification method for mild cognitive impairment, which combines synchronization features and complexity applying recurrence analysis. [9]. Considering the possible nonlinear dynamics, Liu utilized the CRP method on the current and voltage fluctuation time series. Inspired by these broad prospects, we use CRP and CRQA to study the correlation between BFG production and blast furnace related factors.

In this paper, a new and practical framework is proposed for analyzing factors affecting BFG production. This framework helps to reveal impact degree of factors through phase space reconstruction, the establishment of CRP and corresponding quantitative analysis. As far as we know, CRP and CRQA have never been applied in seeking the potential influencing factors of BFG production. Experimental results illustrate that the framework provides clear visualization and effective quantitative analysis of each factor. At the same time, the degree of impact of different factors on BFG production can be determined. The main contributions of this paper can be listed in the following.

- In this paper, high dimensional phase space reconstruction of blast furnace related data is realized. Reconstruction parameters are determined by the mutual information method and false nearest neighbor method. The experimental results prove that choosing the appropriate parameters is the key to constructing the phase space.
- We provide a method to visualize the dynamic characteristics of blast furnace gas data. Results show that this method can intuitively reflect the correlation between various factors and BFG production. These correlations are represented by graphical features in different CRPs.
- Quantitative analysis of blast furnace related data is finished through CRQA. It is a qualitative measure of the CRPs, aiming to obtain the similarity between the blast furnace related data and the BFG production. Compared to CRPs, CRQA can describe graphical features in more detail.

The architecture of the paper is divided into four sections. Section II firstly introduces recursive theory. Then it describes the choice of parameters, including delay time and embedding dimensions. CRP and CRQA are described at the last. In Section III, the impact of some blast furnace related data towards BFG production is analyzed in four steps. Finally, Section IV offers concluding comments and prospective research orientations.

![Schematic diagram of cross-recurrence quantification analysis of blast furnace gas](image)

**Figure 1.** Schematic diagram of cross-recurrence quantification analysis of blast furnace gas

**2. Methodology**

**2.1. RP Analysis framework**

The analysis framework is shown in Figure 1. The most important part is the phase space reconstruction. The reconstructed phase space is equivalent to the original BFG one-dimensional time series in the geometric topological sense. Then use an appropriate threshold to construct a recursive matrix, and use the recursive matrix to draw the corresponding CRP.
2.2. Reconstruction

For a given time series $S = \{s_i | i = 1, 2, \cdots, N\}$, After reconstruction the higher dimensional space is as follows:

$$S(h) = \{s_i, s_{i+\tau}, \cdots, s_{i+(h-1)\tau}\}$$ (1)

Where $\tau$ means the delay time of BFG, $h$ refers to the embedding dimension of BFG series, $N$ is the number of this sequence, $i$ is a positive integer ranging from 1 to $N-(h-1)\tau$.

It’s the most vital part to select proper delay time and embedding dimension in phase space reconstruction. The methods for choosing them are described in the following.

(1) Delay time theorem: In this paper, we use the mutual information method to calculate the delay time. Entropy is a measure of uncertainty in the information theory. The average amount of information required to describe [10] is measured utilizing the entropy. The entropy of BFG series $E = \{e_1, e_2, \cdots, e_N\}$ is defined by $H(E)$, where $e_i$ refers to the possible values that $E$ can take. $H(E)$ is defined as follows

$$H(E) = -\sum_{i=1}^{N} p(e_i) \log_2 \left( p(e_i) \right)$$ (2)

Where $p(e_i)$ is the possibility function of $e_i$.

The joint entropy of two series $E$ and $T = \{t_1, t_2, \cdots, t_M\}$ is expressed as

$$H(E, T) = -\sum_{i=1}^{N} \sum_{j=1}^{M} p(e_i, t_j) \log_2 \left( p(e_i, t_j) \right)$$ (3)

Where $p(e_i, t_j)$ is the joint possibility function of $E$ and $T$.

Mutual Information is the amount of information that both series share. It computed as

$$F(E, T) = H(E) + H(T) - H(E, T)$$ (4)

The first minimum value of $F$ is the suitable delay time of the series to reconstruct phase space.

(2) Embedding dimension theorem: In this section, we apply a false nearest neighbor (FNN) algorithm to calculate the dimensionality of BFG series. The FNN method takes advantage of the local property of the reconstructed phase space, which is a breakthrough in determining the embedding dimension. Its principle is simple and easy to understand. Besides, the calculation of it is convenient. This method is used widely. For most systems, it gives the appropriate result.

Suppose that phase space formed by the BFG sequence is $w$-dimensional. Define $S(h)$ as the nearest neighbor of $S'(h)$. The distance between $S(h)$ and $S'(h)$ is $D_w(h)$. It can be calculated by the following

$$D_w(h) = D_w - S(h) - S'(h)$$ (5)

As the dimension increases to $w+1$, the distance changes. We use $D_{w+1}(h)$ to represent the new distance. It satisfies the following equation:

$$D_{w+1}^2(h) = D_w^2(h) + \| S(h + \tau w) - S'(h + \tau w) \|$$ (6)

When $D_w(h)$ is much greater than $D_{w+1}(h)$, the $S'(h)$ is false. $Z(h, w)$ is defined to judge the false nearest neighbor. It is given by

$$Z(h, w) = \frac{S(h + \tau w) - S'(h + \tau w)}{D_{w+1}(h)}$$ (7)

$D_\tau$ is a positive integer, which is more than 10 and less than 50. If $Z(h, w) > D_\tau$, $S'(h)$ is the false nearest neighbor. According to the embedding dimension theory, the ratio of the false nearest neighbors is calculated from the minimum value, then gradually increases the embedding dimension. When the ratio is near to 0 and stable, the dimension at this time is the minimum value of the required embedding dimension.

2.3. Cross recurrence plot

After the phase space reconstruction is completed, the CRP can be established. The CRP is able to analyze the similarities of two different sequences [11]. After the BFG time series $X$ and one blast furnace element time series $Y$ are reconstructed separately, the corresponding embedding dimensions
and delay time are obtained. Same reconstruction parameters ought to be selected to reconstruct two different BFG series into a same phase space. In the process of constructing cross recurrence plot, if the embedding dimensions and delay time of the two sequences are inconsistent, the principle of selecting parameters is smaller time delay and higher embedding dimension.

Recurrence matrix is able to determine the CRP. The calculation method of the recurrence matrix is given in following equation

\[
R_{m,n}(\varepsilon) = \Theta(\varepsilon - S_m - S_n), m, n = 1, 2, \ldots, N
\]

\(N\) has the same meaning as equation 1. \(\Theta(\cdot)\) is a Heaviside function, which is computed as follows

\[
\Theta(x) = \begin{cases} 
1, & x \geq 0 \\
0, & x < 0 
\end{cases}
\]

\(\|X\|\) is the absolute value of \(X\). \(\varepsilon\) represents the recurrence threshold.

The recurrence threshold is a vital parameter in the process of establishing CRP. Selection of the Recurrence threshold is a trade-off of to have a threshold as small as possible. Meanwhile a sufficient number of recurrence structures are needed. Considering the above conditions, the fixed percentile recurrence rate method is used to determine recurrence threshold. Obviously, the calculated recurrence matrix consists of 0 and 1. CRP is the 2−D visualization of this matrix. In this plot, the black point refers to 1 while the white point refers to 0.

2.4. Cross recurrence quantitative analysis

Zbilut and Webber et al. proposed an advanced approach to capture the characteristics of the recurrence plot [12]. This method is called recurrence quantification analysis (RQA). We use an improved system named CRQA. CRQA supplies a series of parameters to quantify the structure of a CRP. In this work, for finding factors that affect BFG output we extracted a quantitative measurement parameter using cross recurrence toolbox [13]. It is briefly introduced below.

Maximal diagonal line (LMax) is a numerical analysis parameter. A research has shown that by far LMax feature generated the highest discriminative power between CRPs [14]. So in this paper, it is regarded as the measure of similarity, defined by

\[
LMax = \max \left\{ \left\| \sum_{i=1}^{N_i} 1 \right\| \right\}
\]

Where \(N_i\) represents number of diagonal lines in the recurrence plot.

3. Simulation experiment

3.1. Blast furnace elements data acquisition

All blast furnace relevant data used in this paper is from a steel industry. Data for one week is collected. In this paper blast furnace gas total amount (BFGTA) is used to represent the production of blast furnace gas. Factors we select come from the two components of the blast furnace, which include hot blast stove (HBS) and blower.

The hot blast stove is an equipment used in the steelmaking process, which uses fuel to generate flue gas and has heat exchange with air to generate hot air for fire heating.

| Element | BFGTA | HBSF | HBST | HBSP | BOBP | BOBOA | BOBTA |
|---------|-------|------|------|------|------|-------|-------|
| Delay Time(s) | 2 | 2 | 10 | 2 | 10 | 3 | 3 |
| Embedding Dimension | 5 | 13 | 4 | 9 | 5 | 5 | 6 |

The blower can gather a part of the air and increase the air pressure by pressurizing to form a blast furnace blast with a certain pressure and flow rate. Because of the important role of blower and HBS in the production of BFG, we choose to study the data collected from these two devices.
Data acquisition process is shown as figure 1. Flow, temperature and pressure of HBS are factors we firstly research. They are abbreviated as HBSF, HBST and HBSP. Lastly, blower on blast pressure (BOBP), blower on blast oxygen amount (BOBOA) and blower on blast total amount (BOBTA) are picked. Due to space limitations, only parameter calculation, CRP of first day is detailed given. All CRQA results are given in table 3.

Figure 2. Time delay determination of blast furnace relevant factors.

Figure 3. Embedding dimension determination of blast furnace relevant factors.

Table 2. New reconstruction parameters using parameters principle.

| Element  | HBSF | HBST | HBSP | BOBP | BOBOA | BOBTA |
|----------|------|------|------|------|-------|-------|
| Delay Time(s) | 2    | 2    | 2    | 2    | 2     | 2     |
| Embedding Dimension | 13   | 5    | 9    | 5    | 5     | 6     |
3.2. Calculation of phase space reconstruction parameters

Figure 2 shows the process of calculating the time delay using the mutual information method. As HBSF, shown in figure 2 (a), when the delay time is 2 second, mutual information of HBSF is lower than mutual information in 1 second and 3 seconds. In other words, 2 seconds is the proper delay time. So we take 2 seconds as delay time. The determination of embedding dimensions based on FNN is in figure 3. Take BOBP , shown in figure 3 (d), as an example. Because FNN percentage is stably and near to 0 when dimension is 5, we determine 3 as embedding dimension. All parameters are given in table 1. According to the selecting reconstruction parameters principle, new reconstruction parameters are computed, as shown in table 2.

| Element   | HBSF | HBST | HBSP | BOBP | BOBOA | BOBTA |
|-----------|------|------|------|------|-------|-------|
| First Day | 47   | 17   | 26   | 17   | 4     | 28    |
| Second Day| 49   | 29   | 31   | 25   | 21    | 25    |
| Third Day | 24   | 25   | 30   | 21   | 25    | 23    |
| Fourth Day| 37   | 34   | 31   | 34   | 33    | 28    |
| Fifth day | 27   | 18   | 25   | 15   | 18    | 21    |
| Sixth Day | 23   | 19   | 24   | 21   | 19    | 22    |
| Seventh day| 31  | 18   | 29   | 21   | 28    | 23    |

3.3. Blast furnace elements cross recurrence plot visualization

The cross recurrence plots formed by BFG and other blast furnace related elements are demonstrated in figure 4. The structures that appear on these plots can be roughly divided into three types of graphics. The first one is a line segment parallelling to the diagonal direction.

It means a strong correlation between BFG and element. As we can see in figure 4 (a) and 4 (c), there are some lines parallelling to the diagonal, indicating HBSF and HBSP are greatly relevant to the production of BFG.
3.4. Quantitative numerical analysis of blast furnace elements
The measure parameter LMax reveals which elements are the most relevant factor with BFG. LMax computed by data of a week is listed in table 3. From table 3 it is obvious that in most cases HBSF and HBSP rank highest. These results are consistent with our observation in figure 4, where HBSF and HBSP have the most obvious diagonal structure. CRP and CRQA both prove HBSF and HBSP are the most relevant factors with BFG production.

4. Conclusion
We propose a novel framework analyzing BFG production in this paper. The framework implements functions such as phase space reconstruction, data visualization and quantitative analysis. The experimental results show that the framework has a good ability judging correlation between two series. In future work, the prediction of BFG production combining most relevant blast furnace elements will be studied.

Acknowledgments
This work was supported in part by the Science and Technology Project of Tangshan, Grant 19150230E, in part by the S&T Major Project of the Science and Technology Ministry of China, Grant 2017YFE0135700, and in part by the Science and Technology Project of Hebei Education Department, Grant ZD2021088.

References
[1] Rossit, D A., Tohmé, F., Frutos, M. (2019) Industry 4.0: smart scheduling. International Journal of Production Research, 12: 3802-3813.
[2] Xu, L D., Xu, E L., Li, L. (2018) Industry 4.0: state of the art and future trends. International Journal of Production Research, 8: 2941-2962.
[3] Han, Z., Zhao, J., Wang, W., et al. (2016) A two-stage method for predicting and scheduling energy in an oxygen/nitrogen system of the steel industry. Control Engineering Practice, 52: 35-45.
[4] Matino, I., Dettori, S., Colla, V., et al. (2019) Application of echo state neural networks to forecast blast furnace gas production: pave the way to off-gas optimized management. Energy Procedia, 158: 4037-4042.
[5] Shi, L., Wen, Y., Zhao, G., et al. (2016) Recognition of blast furnace gas flow center distribution based on infrared image processing. Journal of Iron and Steel Research International, 3: 203-209.
[6] Wallot, S., Leonardi, G. (2018) Analyzing multivariate dynamics using cross-recurrence quantification analysis (crqa), diagonal-cross-recurrence profiles (dcrp), and multidimensional recurrence quantification analysis (mdrqa)--a tutorial in r. Frontiers in psychology, 9: 2232.
[7] Ziaei-Halimejani, H., Zarghami, R., Mostoufi, N. (2017) Investigation of hydrodynamics of gas-solid fluidized beds using cross recurrence quantification analysis. Advanced Powder Technology, 4: 1237-1248.
[8] Fusaroli, R., Konvalinka, I., Wallot, S. (2014) Analyzing social interactions: the promises and challenges of using cross recurrence quantification analysis. Springer, Cham.
[9] Timothy, L T., Krishna, B M., Nair, U. (2017) Classification of mild cognitive impairment EEG using combined recurrence and cross recurrence quantification analysis. International Journal of Psychophysiology, 120: 86-95.
[10] Tang, J., Liu, F., Zhang, W., et al. (2016) Exploring dynamic property of traffic flow time series in multi-states based on complex networks: Phase space reconstruction versus visibility graph. Physica A: Statistical Mechanics and its Applications, 450: 635-648.
[11] Marwan, N., Thiel, M., Nowaczyk, N R. (2002) Cross recurrence plot based synchronization of time series. Nonlinear processes in Geophysics, 3/4: 325-331.
[12] Desai, U., Martis, R J., Acharya, U R., et al. (2016) Diagnosis of multiclass tachycardia beats
using recurrence quantification analysis and ensemble classifiers. Journal of Mechanics in Medicine and Biology, 01: 1640005.

[13] Marwan, N., Romano, M C., Thiel, M., et al. (2007) Recurrence plots for the analysis of complex systems. Physics reports, 5-6: 237-329.

[14] Serra, J., Serra, X., Andrzejak, R G. (2009) Cross recurrence quantification for cover song identification. New Journal of Physics, 9: 093017.