Wiener Model Identification of Blast Furnace Ironmaking Process

Jiu-sun ZENG,¹ Xiang-guan LIU,¹ Chuan-hou GAO,³ Shi-hua LUO² and Ling JIAN³

1) The Institute of System Optimum Technique, Zhejiang University, Hangzhou, 310027, China.
E-mail: gachou@zju.edu.cn (C. H. Gao) 2) School of Information Management, Jiangxi University of Finance and Economics, Nanchang, 331305, China. 3) College of Mathematics and Computer Science, China University of Petroleum, Dongying, 257061, China.

(Received on May 28, 2008; accepted on September 8, 2008)

To account for the nonlinearity of blast furnace ironmaking process, a nonlinear Wiener model identification algorithm is presented. The system consists of a linear time invariant (LTI) subsystem followed by a static nonlinearity. The inverse of the nonlinearity is assumed to be a linear combination of known nonlinear basis functions and the linear subspace algorithm is used to identify the model. The inputs to the model are parameters regarded to be most responsible for the fluctuation of thermal state in blast furnace while the output to the model is silicon content in hot metal. The identified Wiener model is then tested on datasets obtained from No. 6 Blast Furnace from Baotou Steel. It is found that the blast furnace of concern is a short memory system, so that for each prediction the Wiener method is retrained. It is shown that the retrained model well improves the predictive accuracy.

KEY WORDS: blast furnace iron-making; silicon content; Wiener model; subspace algorithm.

1. Introduction

Blast furnace ironmaking is one of the most energy intensive industrial processes in the world. Over the past several decades, significant progress has been made in reducing energy intensity. One solution is to keep silicon content in hot metal at a low level. And to achieve this, future information of silicon content is needed before blast furnace operators take corrective actions. This is one of the reasons that prediction of silicon content has attracted considerable research interests; among them are models like statistical model¹ and fuzzy models,²³ neural networks and nonlinear methods.⁴⁻⁶ Recently, several researchers have paid attention to the chaotic and fractal characteristics of silicon content in ironmaking process,⁷⁻⁹ which clearly indicates a strong nonlinearity. It is also shown in Ref. 8 that both linear and nonlinear dynamics are present in the blast furnace ironmaking process and to explore the nonlinearity (fractal characteristics) of the system, a linear filtering method is applied to get rid of the linear dynamics. Thus a natural thinking is to deal with the linear dynamics and nonlinear dynamics in a compound model, and the nonlinear Wiener model is our choice.

A Wiener model consists of a linear time invariant (LTI) system followed by a static nonlinearity. Wiener models have the capability of approximating with arbitrary accuracy, any fading memory nonlinear time-invariant system.⁹ There are many successful applications of Wiener model, for example, pH control systems¹⁰ and distillation columns.¹¹ Several methods have been proposed in the literature for the identification of Wiener models.¹²⁻¹⁴ More recently, the linear subspace identification methods¹⁵⁻¹⁷ is extended to the identification of Wiener models,¹⁸⁻¹⁹ which use the subspace methods to identify the linear block and basis functions to approximate the nonlinear block.

The subspace methods adopt the state space model form, which is very convenient for estimation, filtering, prediction and control and has been widely used to describe industrial processes. The subspace methods can be elegantly implemented using well-known numerical linear algebra algorithms like singular value decomposition (SVD) and QR factorization. It offers an attractive alternative to input–output methods due to simple and general parameterization of multiple–input multiple-output (MIMO) system.

The present work deals with application of subspace based Wiener model to the prediction of silicon content in hot metal. It accounts for the nonlinearity characteristic of silicon content and is numerical robust compared to other nonlinear methods like neural networks. The proposed model is then tested in datasets from No. 6 blast furnace of Baotou Steel in China. It is found that the identified model can give good prediction for short-term but the deviation for the long-term is great. To solve this problem, the model is retrained for each prediction and the performance is significantly improved.

2. Process Description

A blast furnace is a type of metallurgical furnace used for smelting to produce molten iron. In a blast furnace, fuel and ore are continuously dumped through the top of furnace, while air (or pure oxygen) is blown into the bottom of
the chamber, so that chemical reactions take place throughout the furnace as the material moves downward. The end products are usually molten iron and slag phases tapped from the bottom, and flue gases exiting from the top of the furnace.

The critical operating parameters are the temperature and silicon content in hot metal. It has been shown that silicon content has a linear approximation relation with temperature. Furthermore, the silicon content is a good measure of the heat content and the course of the blast furnace process. Therefore, only silicon content is chosen as an output variable for the process model. In the meantime, a large number of variables are monitored and they become the basis for the blast furnace operators to judge the status of blast furnace. Since not all monitored process variables have an influence on the fluctuation of silicon content in hot metal, selections must be made. According to engineering knowledge and controllability of variables, we select out 7 process variables as the model inputs and silicon content in hot metal as the model output to construct the predictive model. These process variables include: the quantity of blast, the temperature of blast, the pressure of blast, the quantity of coal powder, the pressure of top gas, the coke ratio and the quantity of oxygen blasted. Delay time for each input variable for this specific blast furnace is given in Table 1.

### Table 1. Delay times for input variables.

| Variable                          | Delay time, h |
|----------------------------------|---------------|
| the quantity of blast            | 2.5           |
| the temperature of blast         | 4             |
| the pressure of blast            | 2.5           |
| the quantity of coal powder      | 2             |
| the pressure of top gas          | 5             |
| the coke ratio                   | 6             |
| the quantity of oxygen blasted   | 2             |

3. Subspace Methods

Subspace methods are a relatively recent development in the field of system identification which is most suitable for multi-variable identification. There are 3 main subspace methods, i.e., N4SID developed by van Overschee and De Moor,15) CVA developed by Larimore16) and MOESP developed by Verhaegen and Dewilde.17) For the 3 algorithms, N4SID and CVA have been seen much more applications. The CVA algorithm was based on mathematical statistics and time series analysis methods while the N4SID algorithm is more closely related to linear system theory. For the present work, we use the N4SID algorithm to identify the linear block of the system. The N4SID algorithm identifies a state-space model,

\[ x_{k+1} = Ax_k + Bu_k + Ke_k, \quad y_k = Cx_k + Du_k + e_k \tag{1} \]

where \( A, B, C \) and \( D \) are the system matrices of appropriate dimensions, and \( x_k \in \mathbb{R}^n \) is the state vector of variables, \( u_k \in \mathbb{R}^m \) is the vector of measured inputs, \( y_k \in \mathbb{R}^m \) is the vector of measured outputs, \( e_k \in \mathbb{R}^m \) is the vector of innovations which is assumed to be a white noise series and the matrix \( K \) is the Kalman gain. The state vector has a very particular meaning for subspace algorithms. The state-space model structure allows for any noise models, i.e., any linear model structure (ARX, ARMAX, OE18)) can be represented by a state-space model.

For subspace algorithms, the state vector, \( x_k \) is defined by a linear combination of past inputs and outputs,

\[ q_k = [y_{k-N}^T \ldots y_{k-1}^T u_{k-N}^T \ldots u_{k-1}^T]^T \tag{2} \]

\[ x_k = Jq_k \tag{3} \]

where \( q_k \) is referred to as the “past” at sample \( k \). The dimension of the past is specified by the number of lags \( N \). The state vector \( x_k \) is computed from data, and is not specified a priori. Once \( J \) has been determined, as will be described below, the state vector can be estimated by Eq. (3).

The state-space model matrices can then be estimated via linear least squares regression,

\[ \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \text{Cov}(x_k, u_k) \times \text{Cov}^{-1}(y_k, u_k) \tag{4} \]

The parameter estimation step of subspace algorithms can vary, but all algorithms proceed in the same general fashion: estimate the state vector from the “past”, and then estimate the state-space matrices using “current” values for state, input and output vectors.

The calculation of \( J \) distinguishes the various subspace algorithms from one another. In the N4SID approach, \( J \) results from a series of geometric arguments based on the set of matrix equations for a linear system. For CVA, N4SID and all other subspace methods, the key step for calculating \( J \) can be written as weighted singular value decomposition,

\[ \text{svd}(W_1 \text{Cov}(\tilde{f}, q)W_2) = [U_1 \quad U_2] \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} \tilde{V}_1^T \\ \tilde{V}_2^T \end{bmatrix} \tag{5} \]

where \( W_1 \) and \( W_2 \) are weighting matrices, \( \text{Cov}(\tilde{f}, q) \) is the covariance matrix of the conditional future \( \tilde{f} \) and the past \( q \). \( S_1 \) is the sub-matrix representing most variance of \( W_1 \text{Cov}(\tilde{f}, q)W \) and \( S_2 \) is the sub-matrix representing the remaining variance, they are both diagonal matrices. \( U_1, U_2 \), \( V_1 \) and \( V_2 \) are matrices whose dimensions correspond to \( S_1 \) and \( S_2 \). \( J \) is calculated from the lower dimensional subspace defined by \( U_1 \). Each subspace algorithm defines the weighting matrices \( W_1 \) and \( W_2 \) differently.

In principle, different subspace algorithms would identify the same state-space model if weighted appropriately. However, other important differences exist between the subspace algorithms: the exact numerical and statistical estimation procedures, determining the dimension of the past, i.e. \( N \) in Eq. (2), and selecting the model order. Ideally, the
model order corresponds to the number of singular values in Eq. (5) greater than zero or some very small value \( \varepsilon \). For N4SID, the order is selected ad hoc by the user, usually looking for a “knee” in the plot of singular values, or by noting where the singular values fall below a specified critical value.

4. Subspace Based Wiener Model Identification

A Wiener model is depicted in Fig. 1. It consists of a linear dynamic subsystem \( G \) followed by a static nonlinearity \( \varphi \). The linear dynamic subsystem \( G \) has a state space representation of the form which is much similar to the model structure in Eq. (1):

\[
x_{k+1} = Ax_k + Bu_k + Ke_k, \quad s_k =Cx_k + Du_k + e_k \tag{6}
\]

where \( s_k \in \mathbb{R}^m \) is the intermediate output.

It is assumed that the static nonlinear function \( \varphi(\cdot) \): \( \mathbb{R}^m \rightarrow \mathbb{R}^r \) is invertible, and that its inverse function \( \varphi^{-1}(\cdot) \) can be described as a linear combination of basis functions in the form:

\[
s_k = \varphi^{-1}(y_k) = \sum_{i=1}^{r} \alpha_i g_i(y_k) \tag{7}
\]

where \( g_i(\cdot): \mathbb{R}^m \rightarrow \mathbb{R}^r \), \( i = 1, \ldots, r \), are the assumed known nonlinear basis functions, and \( \alpha_i \in \mathbb{R}^{r \times m}, (i = 1, \ldots, r) \) are unknown matrix parameters. Typically, the basis functions are polynomials (any smooth function in an interval that can be represented with arbitrary accuracy by a polynomial of sufficiently high order) but they can also be other basis functions like wavelets or radial basis functions.

With this representation for the static nonlinearity, the second equation of Eq. (6) can be written as:

\[
\alpha Y_k = \sum_{i=1}^{r} \alpha_i g_i(y_k) = Cx_k + Du_k + e_k \tag{8}
\]

where \( \alpha = [\alpha_1, \ldots, \alpha_r], \quad Y_k = [g_1(y_k), \ldots, g_r(y_k)]^T \). The Wiener model can then be described as:

\[
x_{k+1} = Ax_k + Bu_k + Ke_k, \quad Y_k = \tilde{C}x_k + \tilde{D}u_k + \tilde{e}_k \tag{9}
\]

with \( \tilde{C} = \alpha^T C, \tilde{D} = \alpha^T D, \tilde{e}_k = \alpha^T e_k \), and where \( \alpha^T \) stands for the left pseudo-inverse of \( \alpha \). It can be seen from Eq. (9) that the parameterization Eqs. (6) and (7) is not unique, since any parameter matrices \( \beta C, \beta D \) and \( \alpha^T \beta^{-1} \), for some nonsingular matrix \( \beta \in \mathbb{R}^{r \times r} \), provide the same description as Eq. (9). A technique that can be used to obtain uniqueness is to normalize the parameter matrices \( \alpha^T \), that is to assume \( \|\alpha^T\|=1 \), where \( \|\cdot\| \) is the Euclidean norm. Under this assumption the parameterization Eqs. (6) and (7) is unique. Thus Eq. (9) can be seen as a state-space realization of an LTI system whose output \( Y_k \) is transformed by the nonlinear and known basis function \( g_i(\cdot) \) version of the original output \( y_k \).

Based on this representation, the N4SID algorithm (and any other subspace algorithm) can then be used to obtain estimates of the system matrices \( A, B, \tilde{C} \) and \( \tilde{D} \) respectively. The problem is how to compute estimates of the matrices \( C, D \) and \( \alpha \). Matrices \( \tilde{C} \) and \( \tilde{D} \) can be expressed in a combined form as:

\[
[\tilde{C} \; \tilde{D}] = \alpha^T[C \; D] \tag{10}
\]

The best (in the mean-squares sense) estimates of matrices \( C, D \) and \( \alpha \) are such that:

\[
(\hat{C}, \hat{D}, \hat{\alpha}) = \arg \min_{C,D,\alpha} \mathbb{E} \left\{ \| \hat{Y} - \alpha^T[C \; D] \|_2 \right\} \tag{11}
\]

where \( \| \cdot \|_2 \) is the power of Euclidean norm. The solution to this minimization problem is provided by a singular value decomposition of the matrix \( [\tilde{C} \; \tilde{D}] \). The subspace identification algorithm for the Wiener model can be summarized as follows:

Step 1: Compute estimates (\( \hat{A}, \hat{B}, \hat{C}, \hat{D} \)) of the system matrices (\( A, B, C, D \)) in Eq. (9) using N4SID algorithm for LTI system.

Step 2: Compute the economy size SVD of \( [\hat{C} \; \hat{D}] \), and the partition of this decomposition as:

\[
[\hat{C} \; \hat{D}] = U_s \Sigma V_s^T = [U_1 \; U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} V_1^T V_2^T \tag{12}
\]

where \( \Sigma_1 \) is a diagonal matrix containing the \( s \) nonzero singular values (\( \sigma_i, i = 1, \ldots, s \)) of \([\hat{C} \; \hat{D}]\) in non-increasing order, and where the unitary matrices \( U_i = [u_1, u_2, \ldots, u_s] \in \mathbb{R}^{m \times s} \) and \( V_i = [v_1, v_2, \ldots, v_s] \in \mathbb{R}^{p \times s} \) contain the corresponding left and right singular vectors, respectively. The partition of SVD in Eq. (12) is such that the following dimensions for the matrices \( \Sigma_1, U_1 \) and \( V_1 \) hold, \( \Sigma_1 \in \mathbb{R}^{s \times s}, U_1 \in \mathbb{R}^{s \times s} \) and \( V_1 \in \mathbb{R}^{p \times s} \).

Step 3: Compute the estimates of the parameter matrices \( C, D \) and \( \alpha \) as \([\hat{C} \; \hat{D}] = \Sigma_1 V_1^T \) and \( \alpha = U_1^T \).

5. Simulation of Algorithm

The proposed Wiener model is then applied to consecutive data (with the size of \( 1 \, 000 \)) collected from No. 6 blast furnace of Baotou Steel in China, and the volume of the blast furnace is 2,500 m\(^3\). The iron ore used for this blast furnace is from the famous Baiyunebo mineral in Baotou, Inner Mongolia of China, which contains materials like rare earth elements, niobium and other metals. The complex ingredients of the iron ore make the ironmaking process extremely unstable, and the prediction of silicon content becomes more difficult.

In order to determine the model order of both the Wiener and linear models, identification experiments were performed for the data. The N4SID was then used in the first step of the algorithm and a sixth-order model was estimated for the linear block in the Wiener model. On the other side, the following third-order polynomial:

\[
\varphi^{-1}(y_k) = -0.4445 y_{k-1} - 0.6094 y_{k-2} - 0.6506 y_{k-3} \tag{13}
\]

was estimated for the inverse of the nonlinear block in the Wiener model. The estimated nonlinear characteristic is represented in Fig. 2. Generally, the range of \( y_k \) (here \( y_k \) is the silicon content) is between 0 and 1,6, but to show the
nonlinear characteristic more clearly, we extend it to $-6$ and $6$.

To evaluate the model, the estimated model is then used for estimation and prediction. The first 700 datasets are used for estimation and the remaining 300 datasets are for validation. What should be noted is that the model output is no longer the system output $y_k$ but the transformed version of $Y_k$, but since the basis functions are used with the first element being linear, the system output $y_k$ can be directly obtained from $Y_k$ (in this case, $Y_k = [y_k, y_k^2, y_k^3]$). The result of estimation and prediction are displayed in Fig. 3. It can be seen from Fig. 3 that for prediction, the identified model function well at the beginning but deviates much after a certain time. This clearly shows that the ironmaking process is a short memory system. The reason may lie on the fact that the blast furnace ironmaking is a process influenced by many factors and under constant change. As the production goes on, the influence gradually accumulates and production conditions change significantly. To solve this problem, for each prediction, the model is retrained with a new sample from validation set and the oldest training set will be discarded. The new predictions are shown in Fig. 4. It can be seen from Fig. 4 that the retrained models function better than the model without retraining.

To further evaluate the performance of the Wiener model, three error measurements, MSE (mean square error) and the MAPE (mean absolute percentage error) and CC (correlation coefficients) are used to evaluate the accuracy of the model. The MAPE is as follows:

$$
MAPE = \frac{1}{n} \sum_{k=1}^{n} \frac{|y_k - \hat{y}_k|}{y_k} \quad \text{(14)}
$$

where $y_k$ is the actual output at instance $k$ and $\hat{y}_k$ is the predicted output at instance $k$. $n$ is the length of the output sequence. The smaller the value of MAPE, the better the predictive accuracy will be. Generally, the signification of MAPE is presented in Table 2.

And the MSE is as follows:

$$
MSE = \frac{1}{n} \sum_{k=1}^{n} (y_k - \hat{y}_k)^2 \quad \text{(15)}
$$

The results are compared with results obtained from linear subspace method and Wiener model without retraining as in Table 3.

It can be see from Table 3 that the retrained Wiener models function much better than the Wiener model without retraining. Both the MAPE and MSE value is smaller when the retrained Wiener models are adopted. For a process under fierce fluctuation, We can say that good predictions are achieved.
6. Conclusions

The subspace based Wiener model has been successfully used for prediction of silicon content in hot metal. The proposed model, with a cascade structure of linear and nonlinear block, is suitable for identification of the blast furnace ironmaking process, which is also a mixture of linear and nonlinear dynamics. By exploring the inverse of the nonlinear block, the Wiener model is then transformed into a multiple-input and multiple-output (MIMO) linear model which can be easily estimated from the linear subspace methods. The identified model is then used for prediction and it is found that the ironmaking process is a short memory system. A retraining method is then employed to make more accurate predictions. The Wiener model well accounts the nonlinearity of the ironmaking process and there are many successful predictive control applications for this kind of model structure. To better control the ironmaking process, further application of the Wiener model structure should be extended to the obtaining of a predictive controller for the blast furnace ironmaking process.

Acknowledgements

The third author wishes to thank sponsor and financial support from Zhejiang Provincial Natural Science Foundation of China under Grant No. Y107110, Research Fund for the Doctoral Program of Higher Education of China (for new teachers) under Grant No. 20070335161, and the fourth author would like to thank Scientific Research Fund of Jiangxi Provincial Education Department under Grant No. GJJ08358 for financial support.

REFERENCES

1) T. Bhattacharya: ISIJ Int., 45 (2005), 1943.
2) S. H Luo and J. S. Zeng: Proc. of Fuzzy Infomation and Engineering, Springer, Guangzhou, (2007), 411.
3) Q. H. Li and X. G. Liu: J. Iron Steel Res. Int., 12 (2005), 1.
4) J. Jimenez, J. Mochon, J. S. Ayala and F. Obeso: ISIJ Int., 44 (2004), 573.
5) H. Saxén and F. Pettersson: ISIJ Int., 47 (2007), 1732.
6) F. Pettersson, N. Chakraborti and H. Saxen: Appl. Soft Comput., 7 (2007), 387.
7) C. H. Gao, Z. M. Zhou and J. X. Qian: J. Iron Steel Res. Int., 12 (2005), 3.
8) S. H. Luo, X. G. Liu and J. S. Zeng: ISIJ Int., 47 (2007), 1102.
9) S. Boyd and L. Chua: IEEE Trans. Circ. Syst., 32 (1985), 1150.
10) J. C. Gomez, A. Jutan and E. Baeyens: IEEE Proc. Control Theory Appl., 151 (2004), 329.
11) H. Bloemen, C. Chou, T. van den Boom, M. Verhaegen and T. Backx: J. Process Control, 11 (2001), 601.
12) S. L. Lacy, R. S. Erwin and D. Bernstein: J. Dyn. Sys., Meas. Control, 123 (2001), 566.
13) R. Raich, G. T. Zhou and M. Viberg: IEEE Trans. Automat. Control, 50 (2005), 1629.
14) J. Gomez and E. Baeyens: Proc. 15th Int. Federation of Automatic Control World Congress, Barcelona, (2002), 2849.
15) S. L. Lacy and D. S. Bernstein: Int. J. Control, 78 (2005), 906.
16) W. Larimore: Proc. of the 29th IEEE Conf. on Decision and Control, IEEE, Hawaii, (1990), 596.
17) M. Verhaegenen: Automatica, 30 (1994), 61.
18) J. Gomez and E. Baeyens: Proc. 15th Int. Federation of Automatic Control World Congress, Barcelona, (2002), 2849.
19) S. L. Lacy and D. S. Bernstein: Int. J. Control, 78 (2005), 906.
20) R. D. Martin, F. Obeso, J. Mochón, R. Barea and J. Jiménez: Ironmaking Steelmaking, 34 (2007), 241.
21) L. Ljung: System Identification, 2nd ed., Prentice-Hall, Upper Saddle River, NJ, (1999), 158.