AIChE

A Chaos-Based Iterated Multistep Predictor for Blast Furnace Ironmaking Process

Chuanhou Gao
Institute of Systems Optimization Techniques, Dept. of Mathematics, Zhejiang University, Hangzhou 310027, China

Jiming Chen
Institute of Industrial Process Control, Dept. of Control, Zhejiang University, Hangzhou 310027, China

Jiusun Zeng
Institute of Systems Optimization Techniques, Dept. of Mathematics, Zhejiang University, Hangzhou 310027, China

Xueyi Liu
Dept. of Mathematics, China Jiliang University, Hangzhou 310018, China

Youxian Sun
Institute of Industrial Process Control, Dept. of Control, Zhejiang University, Hangzhou 310027, China

DOI 10.1002/aic.11724
Published online March 11, 2009 in Wiley InterScience (www.interscience.wiley.com).

The prediction and control of the inner thermal state of a blast furnace, represented as silicon content in blast furnace hot metal, pose a great challenge because of complex chemical reactions and transfer phenomena taking place in blast furnace ironmaking process. In this article, a chaos-based iterated multistep predictor is designed for predicting the silicon content in blast furnace hot metal collected from a pint-sized blast furnace. The reasonable agreement between the predicted values and the observed values indicates that the established high dimensional chaotic predictor can predict the evolvement of silicon series well, which conversely render the strong indication of existing deterministic mechanism ruling the dynamics of complex blast furnace ironmaking process, i.e., a high-dimensional chaotic system is suitable for representing the blast furnace system. The results may serve as guidelines for characterizing blast furnace ironmaking process, an extremely complex but fascinating field, with chaos in the future investigation. © 2009 American Institute of Chemical Engineers *AIChE J,* 55: 947–962, 2009
Keywords: blast furnace ironmaking process, silicon content in hot metal, chaos, prediction, iterated multistep predictor

Introduction

The main purpose of blast furnace (BF) is to chemically reduce and physically convert iron oxides into liquid iron called ''hot metal'' or ''pig iron.'' In the manufacturing of iron and steel, blast furnace ironmaking process (BFIP) is a crucial unit operation and consumes about 70% of the whole

Correspondence concerning this article should be addressed to J. Chen at jmchen@ieee.org.

energy input to the integrated route of steel production. Recently, due to the industrial and increasing social need for iron and steel plus increased prices for raw materials and reducing agents, there has been a growing awareness of modeling and controlling complex BFIP for increasing productivity and cutting potential cost. However, the inner phenomena of ironmaking BF are extremely challenging for the human mind to understand. The phenomena that take place inside the ironmaking blast furnace are usually regarded as one of the most complex phenomena in industrial process because of the high temperature, high pressure, multiphase simultaneously coexisting, interacting with each other and exchanging momentum, mass and energy, and so on. Besides, the sealed structure of BF and sensible probing cost make it almost impractical to detect adequately in-furnace state and behavior. Thus, it poses a great challenge to build an accurate model to simulate BFIP.

In the past decades, extensive thermodynamic and kinetic research and operation analysis had been carried out on BFIP.[1,2] Based on these studies, a fundamental understanding of BFIP has been progressively established and much progress has been made in constructing mathematical models to simulate BFIP. Candidate examples can be categorized by the number of dimensions treated in the models as one-dimensional,[3] two-dimensional,[4] and three-dimensional models[5] (See more references therein). Objectively speaking, these models can, to some extent, exhibit good performance in revealing the inner behavior of ironmaking BF, but they have some unavoidable shortcomings when they are used in practice, for example, weak real time performance, excessive assumptions and simplification for establishing models, difficulty of finding the analytical solutions for equations, etc. It can be said that a fully reliable mechanism model is yet to emerge, and BFIP remains a partial mystery. Today, the BF operation still depends on the experience and intuition of skilled operators. At the core of the current dilemma lies the difficulty in finding an accurate predictive model for the temperature of blast furnace hot metal. Overcoming this hurdle may open an exciting possibility for simulating BFIP.

As is well known, the objectives of the ironmaking BF are to obtain a good and homogenous quality of hot metal, reduce the energy consumption, and increase productivity while maintaining long BF life.[6] Among these objectives a good and homogenous quality of hot metal is the first important one. In the BFIP the quality of hot metal is usually determined by its temperature and the content of crucial components, mostly C, Si, S, which thus need to be controlled stringently within acceptable upper and lower bounds. Here attentions are mainly focused on the investigation of silicon content in BF hot metal since it has long been recognized as one of the chief indicators to represent the inner thermal state of the BF. A high silicon content in hot metal often implies a high furnace temperature, which corresponds to a high cost of coke and excessive generation of heat, while a decreasing silicon content typically expresses the possibility of a cooling of BF, which may lead to a chilled hearth in the worst case.[7] Therefore, to achieve a high quality of hot metal and a stable performance of BF, the silicon content in hot metal has to be maintained within proper bounds. Meanwhile, from the view point of saving energy, the silicon content in hot metal should be as low as possible on the premise of obtaining hot metal with high quality.

The above reasons make the prediction of silicon content in hot metal become a task of great practical significance for melters. In view of the shortcomings of analytical models, data-driven models have recently been extensively explored in an attempt to shed light on the complex interrelations between variables in the BFIP. A significant amount of tools and algorithms have been applied in this area, including neural net,[8–12] nonlinear time series analysis,[13,14] partial least square,[15,16] fuzzy theory,[17] regression model,[7] etc. The main basis of these applications is the fact that many tools have universal nonlinear approximation capacities. Roughly speaking, these models can serve as potential candidates to predict silicon content in hot metal, but it is always difficult to choose an appropriate tool from all kinds of tools, and to explain why this tool is better than that one for a given BF or why one model can work while others cannot work.

A sticking point causing the above difficulties, in our view, is that the inherent mechanism inside the BF is neglected in the process of constructing data-driven models. In fact, the optimal tool to simulate and model a phenomenon should be based on the understanding of the inherent mechanism behind such a phenomenon. Thus, it is a primary task to nail down which inherent mechanism of the evolvement of silicon content in hot metal is, regular, chaotic or random. Our early works indicate that there exists a deterministic mechanism, i.e., chaos, in the evolvement of silicon content in hot metal by calculating fractal dimension,[18] power spectrum,[19] Lyapunov exponents,[20] and forecast entropy.[21] The same results are observed from directly investigating time series data of BF temperature by Miyano et al.[13] Based on the chaotic features exhibited by the evolvement of silicon content in hot metal and the fact that chaos admits short-term prediction while refuses long-term prediction,[22] this article is devoted to the construction of a chaotic predictor, i.e., a chaos-based iterated multistep predictive model, for silicon prediction. The results may serve as guidelines for constructing a valid data-driven model for blast furnace system.

The remainder of this article is organized as follows: In the next section, a brief introduction of the BFIP is made. This will be followed by a detailed statement of the current problem, including silicon transfer mechanism inside BFIP, preliminary knowledge on phase space reconstruction, and design of chaotic multistep predictor. Next, the designed chaotic predictor is used to predict the silicon content in BF hot metal for the given examples, and its advantages and potential pitfalls are discussed. Finally, conclusions and points of possible future research are summarized.

Blast Furnace Ironmaking Process

Blast furnace ironmaking process is a highly complex industrial semibatch process. Its complexity originates from not only the in-furnace complex process chemistry but also the impossibility of direct measurements of many in-furnace variables due to high temperature, pressurized operation, and nearly sealed structure. When a blast furnace runs, it acts as a huge countercurrent heat and mass exchanger and metallurgical reactor inside which highly complex transport

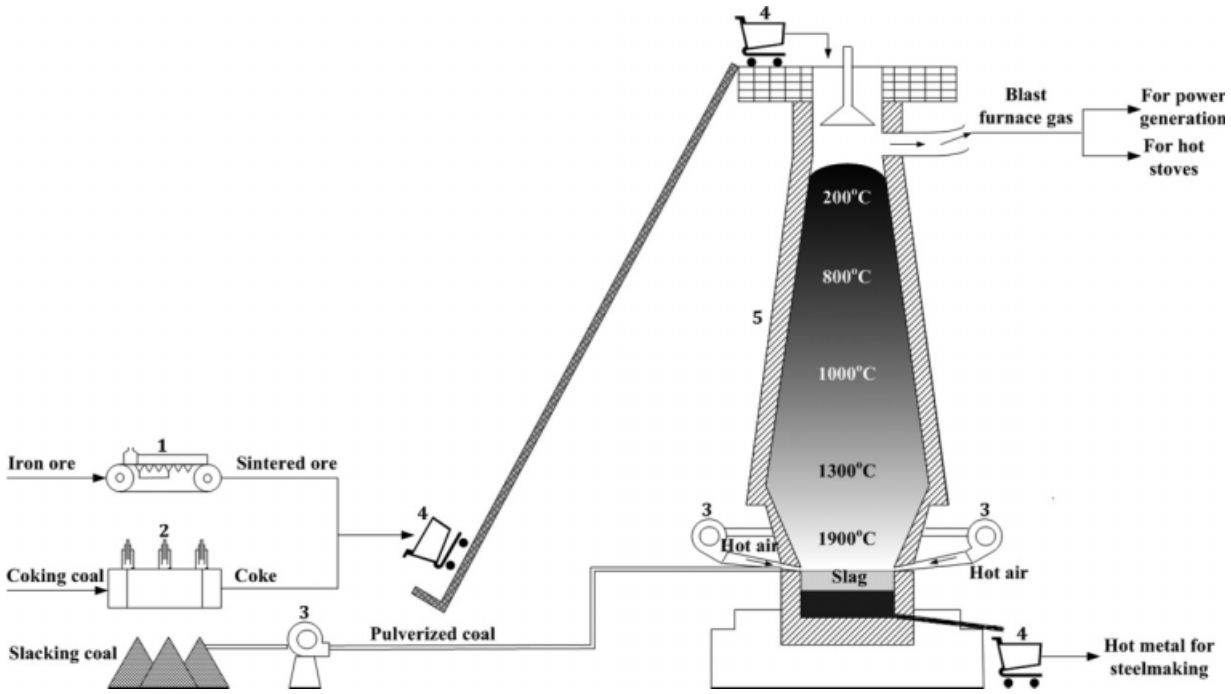

Figure 1. Flow chart of blast furnace ironmaking process: (1) Sintering machine; (2) coke oven; (3) blower; (4) skip car; (5) blast furnace.

phenomena and chemical reactions occur. Generally, according to the state of burdens moving along the BF, the ironmaking BF is roughly divided into five different zones from the top to the bottom, called lumpy zone, cohesive zone, active coke zone, Hearth-Deadman (The dead man is the hearth coke bed), and tuyere raceways, respectively.[23]

Figure 1 shows the flow chart of BFIP. At the beginning of ironmaking, the solid raw materials, including iron ore and coking coal, pretreated to be sintered ore and coke, respectively, with definite quantities of fluxes, mostly limestone, are charged layer by layer into the furnace from the top. Preheated air is injected together with pulverized coal or oil at the bottom through the tuyeres located just above the hearth. As the ironmaking proceeds, the solid burden descends gradually, partly because of the combustion of coke and partly because of the melting and softening of iron oxides,[24] while the hot air moves upward continuously through the furnace and reacts with coke to produce CO. The generated CO then reduces the ore and forms the final product, hot metal accumulating in the hearth. In the meantime, some impurities called slag, mostly CaO and SiO_2, are yielded as a by-product floating on the top of hot metal. The residence time of ore and hot gas within BF is about 8 h and a few seconds, respectively.[23] The main chemical reactions during this stage are:

$$2C + O_2 = 2CO \ (\text{under the condition of high temperature}) \quad (1)$$

$$Fe_2O_3 + 3CO = 2Fe + 3CO_2 \quad (2)$$

$$CaCO_3 = CaO + CO_2 \quad (3)$$

$$CaO + SiO_2 = CaSiO_3 \quad (4)$$

Along with reaction (1), a great deal of heat energy is produced, which makes the highest temperature within BF be close to 1900°C. At the end of the ironmaking, the generated blast furnace gas, mostly CO and CO_2, associated with part of unused original nitrogen, are released from the top of the furnace, part of which enter into the stoves and are employed to heat the fresh blown air, while the liquid hot metal and slag are periodically removed through several tapholes near the bottom of the furnace for further handling. The whole ironmaking time will last 6 to 8 h. This process is repeated for a very prolonged period, about 6 to 8 years for a BF.

Chaotic Prediction for Blast Furnace Ironmaking Process

Silicon transfer mechanism inside blast furnace ironmaking process

Hot metal is the immediate product of an ironmaking BF. In the hot metal, there are several alloying elements dissolved, key components including C, Si, S. Among them, the silicon content in BF hot metal is an important index reflecting the hot metal quality and energy utilization ratio of BF. It often acts as a chief indicator to represent the in-furnace thermal state due to the exceeding difficulty in directly measuring the in-furnace temperature. Thus, the silicon content in BF hot metal should be controlled within acceptable upper bound and lower bound for a smooth operation of BF. To achieve this, understanding of silicon transfer mechanism inside BF is necessary.

The silicon inside the BF is originally in the form of silica, mostly SiO_2, existing in the BF burdens when it is

introduced into BF. As the BFIP proceeds, SiO_2 is reduced by CO to generate volatile $SiO_{(g)}$, and then $SiO_{(g)}$ continuous to react with carbon saturated iron droplets producing Si, which results in silicon transfer to the iron.[1] The related chemical reactions are

$$(SiO_2) + CO = SiO_{(g)} + CO_2 \tag{5}$$

$$SiO_{(g)} + C = Si + CO \tag{6}$$

$$CO_2 + C = 2CO \tag{7}$$

Combining these three reactions yields the whole silicon transfer reaction as

$$(SiO_2) + 2C = Si + 2CO \tag{8}$$

Extensive kinetic research indicates that the silicon content in BF hot metal is mainly determined by reaction (6) but not the direct reaction between SiO_2 and carbon saturated iron droplets.[1] The former has fast rate of silicon transfer while the latter's reaction rate is too slow.

Preliminary knowledge on phase space reconstruction

Chaotic predictor plays an import role on forecasting the future values of time series collected from many complex systems like atmosphere,[25] stock market,[26] fluidized bed,[27] hydrology,[28] etc. Of these types of systems, a universal feature is the extreme difficulty in building valid mathematical models, often being high-order differential equations which always exist in theory, for process description through mechanism analysis. As any high-order differential equation can be mathematically converted to a set of first-order autonomous equations, the system represented by the latter is equivalent to that described by the former. Naturally, attention is turned to the dynamical system, also called phase space, constituted by the dynamical variables emerging in all the first-order equations. According to Takens' theorem,[29] a topological equivalence of such phase space may be reconstructed from time series produced by system, and will keep some statistical quantities invariant, such as dimension, Lyapunov exponents, etc. Thus, nonlinear time series analysis becomes a powerful tool for shedding light on the complex dynamics of such systems. The main basis is that an experimentally measured time series of every physical variable of system contains enough information on revealing the dynamic behavior of the whole system.

Consider an m-dimensional dynamical system, where m means the number of dynamical variables included in all the first-order autonomous equations and is often quite large and even infinite in practice. For this dynamical system, a time series $\{s_i \equiv s(t_i) \equiv s(i\Delta t)\}_{i=1}^{N_{\text{dat}}}$ can be obtained by measuring a single scalar function $s = g(\mathbf{u})$ at the frequency of $1/\Delta t$, in which $\mathbf{u} \in \mathbf{R}^m$ is the state variable evolving in a finite-dimensional attractor M. From this, one can reconstruct a d-dimensional phase space by using time-delayed coordinates of the form[30]

$$\mathbf{S}_i = \left(s_i, s_{i+\tau}, s_{i+2\tau}, \ldots, s_{i+(d-1)\tau}\right)^{\mathrm{T}}, \quad i = 1, 2, \ldots, N_d, \tag{9}$$

Here τ is a positive integer and the product of τ and Δt is usually called delay time, d is the embedding dimension related to irreducible degrees of freedom of the system, and $N_d = N_{\text{dat}} - (d-1)\tau$ is the size of vectors $\{\mathbf{S}_i\}$. Compared with the original phase space, this one constituted by $\mathbf{S}$ is called reconstructed phase space. Takens' theorem[29] asserts that the pattern $\mathbf{S}_i$ has a one-to-one correspondence to the variable $\mathbf{u}_i$ of the original phase space if $\tau, d \gg N_d$ and $d > 2d_M$, where d_M is the box-counting dimension of attractor M. Namely, there exists a diffeomorphism between the reconstructed dynamics $\mathbf{F}$: $\mathbf{R}^d \rightarrow \mathbf{R}^d$ and the dynamics of original phase space, provided that τ and d are appropriately chosen. Thus, prediction of future values of time series $\{s_i\}_{i=1}^{N_{\text{dat}}}$ may be realized through $\mathbf{F}$ while lying on the intrinsic unknown dynamics of the real system.

Chaos-based iterated multistep prediction

In the reconstructed phase space, the dynamics $\mathbf{F}$: $\mathbf{R}^d \rightarrow \mathbf{R}^d$ with $\mathbf{F} = (f_1, f_2, \ldots, f_d)^{\mathrm{T}}$ can define a one-step map describing the evolvement of the reconstructed vectors as[31]

$$\begin{cases} \mathbf{S}_2 = \mathbf{F}^1(\mathbf{S}_1) \\ \mathbf{S}_3 = \mathbf{F}^1(\mathbf{S}_2) \\ \ldots \\ \mathbf{S}_{N_d} = \mathbf{F}^1(\mathbf{S}_{N_d-1}) \end{cases} \tag{10}$$

Assume $\hat{\mathbf{F}}$ to be the good representations of $\mathbf{F}$ and h to be the measurement function of the reconstructed vectors through which the dth entry of vectors may be available, then the future value of the time series $\{s_i\}_{i=1}^{N_{\text{dat}}}, s_{N_{\text{dat}}+1}$, may be approximately obtained by

$$\begin{aligned} \hat{\mathbf{S}}_{N_d+1} \approx \mathbf{S}_{N_d+1} &= \begin{pmatrix} s_{N_d+1} \\ s_{N_d+1+\tau} \\ \cdots \\ s_{N_d+1+(d-2)\tau} = s_{N_{\text{dat}}+1-\tau} \\ s_{N_{\text{dat}}+1} \end{pmatrix} \\ &= \mathbf{F}^1(\mathbf{S}_{N_d}) \approx \hat{\mathbf{F}}^1 \begin{pmatrix} s_{N_d} \\ s_{N_d+\tau} \\ \cdots \\ s_{N_d+(d-2)\tau} = s_{N_{\text{dat}}-\tau} \\ s_{N_{\text{dat}}} \end{pmatrix}, \\ \hat{s}_{N_{\text{dat}}+1} \approx h(\hat{\mathbf{S}}_{N_d+1}) &\approx (h \bullet \hat{\mathbf{F}}^1)(\mathbf{S}_{N_d}) \approx \widehat{f_d^1}(\mathbf{S}_{N_d}) \end{aligned} \tag{11}$$

Mathematically, this one-step map can induce a multistep map through $\mathbf{F}^1$ operating on itself continuously, e.g., a p-step map $\mathbf{F}^p$ can be constructed according to

$$\mathbf{S}_{i+p} = \underbrace{\mathbf{F}^1(\cdots(\mathbf{F}^1}_{p}(\mathbf{F}^1(\mathbf{S}_i)))) = \mathbf{F}^p(\mathbf{S}_i) \tag{12}$$

where $\mathbf{S}_{i+p} = \left(s_{i+p}, s_{i+p+\tau}, s_{i+p+2\tau}, \ldots, s_{i+p+(d-1)\tau}\right)^{\mathrm{T}}$, p is the ahead step of prediction and may be 1,2,3,…. This implies that

it is possible to make a multistep prediction for the reconstructed vectors $\{\mathbf{S}_i\}_{i=1}^{Nd}$ in theory.

As we talk about multistep prediction, there are two possible methods to realize it. One is that the prediction of generating $\mathbf{S}_{i+p}$ ($p \geq 2$) is directly from $\mathbf{S}_i$, which is called direct multistep prediction, the other is that $\mathbf{S}_{i+p}$ ($p \geq 2$) is predicted based on $\hat{\mathbf{S}}_{i+p-1}, \hat{\mathbf{S}}_{i+p-2}, \ldots, \hat{\mathbf{S}}_{i+1}, \mathbf{S}_i$, which is called iterated multistep prediction.[32] Here, let $\hat{\mathbf{S}}_{i+j}$ be the predicted value of $\mathbf{S}_{i+j}(j = 1, 2, \ldots, p-1)$. The iterated multistep prediction may be implemented by

$$\begin{cases} \hat{\mathbf{S}}_{i+1} \approx \hat{\mathbf{F}}^1(\mathbf{S}_i) = \hat{\mathbf{F}}^1((s_i, s_{i+\tau}, \ldots, s_{i+(d-1)\tau})^{\mathrm{T}}), \\ \hat{s}_{i+1+(d-1)\tau} \approx h(\hat{\mathbf{S}}_{i+1}) \approx \widehat{f_d^1}(\mathbf{S}_i) = \widehat{f_d^1}((s_i, s_{i+\tau}, \ldots, s_{i+(d-1)\tau})^{\mathrm{T}}), \\ \hat{\mathbf{S}}_{i+2} \approx \hat{\mathbf{F}}^1(\hat{\mathbf{S}}_{i+1}) = \hat{\mathbf{F}}^1((s_{i+1}, s_{i+1+\tau}, \ldots, \hat{s}_{i+1+(d-1)\tau})^{\mathrm{T}}), \\ \hat{s}_{i+2+(d-1)\tau} \approx h(\hat{\mathbf{S}}_{i+2}) \approx \widehat{f_d^1}(\hat{\mathbf{S}}_{i+1}) = \widehat{f_d^1}((s_{i+1}, s_{i+1+\tau}, \ldots, \hat{s}_{i+1+(d-1)\tau})^{\mathrm{T}}), \\ \cdots \\ \hat{\mathbf{S}}_{i+p} \approx \hat{\mathbf{F}}^1(\hat{\mathbf{S}}_{i+p-1}) = \hat{\mathbf{F}}^1((s_{i+p-1}, \ldots, \hat{s}_{i+p-1+(d-2)\tau}, \hat{s}_{i+p-1+(d-1)\tau})^{\mathrm{T}}), \\ \hat{s}_{i+p+(d-1)\tau} \approx h(\hat{\mathbf{S}}_{i+p}) \approx \widehat{f_d^1}(\hat{\mathbf{S}}_{i+p-1}) = \widehat{f_d^1}((s_{i+p-1}, \ldots, \hat{s}_{i+p-1+(d-2)\tau}, \hat{s}_{i+p-1+(d-1)\tau})^{\mathrm{T}}) \end{cases} \tag{13}$$

Only from Eqs. 12 and 13, it is hard to say which method is better when making multistep predictions. Theoretically, direct multistep prediction has more uncertainty of future and will suffer more difficulty than the one-step prediction, while iterated multistep prediction does not consider the accumulated errors in the inner single step prediction.[32] Although McNames[32] and Casdagli[33] advocate that iterated multistep prediction is more accurate than direct multistep prediction, both of these methods are used in the subsequent research.

In the process of seeking a proper approximator of $\mathbf{F}$, there are two possible methods to express $\hat{\mathbf{F}}$, i.e., global approximation method and local approximation method. The global approximation method allows that $\hat{\mathbf{F}}$ is valid globally in all phase space whenever the prediction starts. This usually needs a large number of parameters in the map $\hat{\mathbf{F}}$, which makes the model too over-parameterized to be used in practice. Local approximation method, on the other hand, demands that $\hat{\mathbf{F}}$ be only valid locally in phase space, such as in a close neighborhood in phase space.[34] In the process of modeling with local approximation method, the domain of $\mathbf{F}$ is frequently divided into many subsets, and $\hat{\mathbf{F}}$ is sought in every subset. Compared with global approximation method, local approximation one has little complexity in approximating $\mathbf{F}$ without degrading the model accuracy. Thus, in the subsequent investigation, local approximation method is adopted for predicting silicon content in blast furnace hot metal.

The main idea of local approximation method is to predict the future by analogs of past event, i.e., using nearby states to predict the future.[34] According to this idea, expanding $\mathbf{F}^1(\mathbf{S}_{N_d})$ of Eq. 11 about a fiducial orbit $\mathbf{S}_{nN_d}$ in a Taylor series then gives

$$\begin{aligned} \mathbf{S}_{N_d+1} &= \mathbf{F}^1(\mathbf{S}_{N_d}) = \mathbf{F}^1(\mathbf{S}_{nN_d} + \mathbf{S}_{N_d} - \mathbf{S}_{nN_d}) \\ &= \mathbf{F}^1(\mathbf{S}_{nN_d}) + \left.\frac{\partial \mathbf{F}^1}{\partial \mathbf{S}_{N_d}}\right|_{\mathbf{S}_{nN_d}} \cdot (\mathbf{S}_{N_d} - \mathbf{S}_{nN_d}) + \ldots \end{aligned} \tag{14a}$$

where $\mathbf{S}_{nN_d}$ is the nearest neighbor point to $\mathbf{S}_{N_d}$ with respect to the Euclidean norm

$$\|\mathbf{S}_{nN_d} - \mathbf{S}_{N_d}\| = \left\{ \sum_{\alpha=0}^{d-1} [s_{nN_d+\alpha} - s_{N_d+\alpha}]^2 \right\}^{1/2} \tag{14b}$$

and the suffix index, nN_d, satisfies $nN_d < N_d$, $\left.\frac{\partial \mathbf{F}^1}{\partial \mathbf{S}_{N_d}}\right|_{\mathbf{S}_{nN_d}}$ representing the $d \times d$ Jacobian matrix of $\mathbf{F}^1$ evaluated along the fiducial orbit $\mathbf{S}_{nN_d}$. Thus, the dth entry of $\mathbf{S}_{N_d+1}$, $s_{N_{\mathrm{dal}}+1}$, is written as the following expression

$$\begin{aligned} s_{N_{\mathrm{dat}}+1} = {} & f_d^1(\mathbf{S}_{nN_d}) + \sum_{\beta=1}^{d} Df_{d\,\beta}^1(\mathbf{S}_{nN_d})(s_\beta(N_d) - s_\beta(nN_d)) \\ & + \sum_{\beta=1}^{d}\sum_{\gamma=1}^{d} D^2 f_{d\,\beta\gamma}^1(\mathbf{S}_{nN_d})(s_\beta(N_d) - s_\beta(nN_d))(s_\gamma(N_d) \\ & - s_\gamma(nN_d)) + \cdots \end{aligned} \tag{15a}$$

where

$$Df_{d\,\beta}^1(\mathbf{S}_{nN_d}) = \left.\frac{\partial f_d^1}{\partial s_\beta}\right|_{\mathbf{S}_{nN_d}}, \quad D^2 f_{d\,\beta\gamma}^1(\mathbf{S}_{nN_d}) = \left.\frac{1}{2!}\frac{\partial^2 f_d^1}{\partial s_\beta \partial s_\gamma}\right|_{\mathbf{S}_{nN_d}}, \quad \ldots(\beta, \gamma = 1, 2, \ldots, d) \tag{15b}$$

and

$$s_\beta(N_d) = s_{N_d+(\beta-1)\tau}, \quad f_d^1(\mathbf{S}_{nN_d}) = s_{nN_d+(d-1)\tau+1} \tag{15c}$$

To realize the prediction of $s_{N_{\mathrm{dat}}+1}$, these unknown parameters $Df_{d\,\beta}^1(\mathbf{S}_{nN_d}), D^2 f_{d\,\beta\gamma}^1(\mathbf{S}_{nN_d}), \ldots$ should be calculated.

Based on the work of Sano and Sawada[35] and Brown and Bryant,[36] the calculation of $Df_{d\,\beta}^1(\mathbf{S}_{nN_d})$, $D^2 f_{d\,\beta\gamma}^1(\mathbf{S}_{nN_d}), \ldots$ may be realized from the real measured data. Let $\mathbf{S}_{i^r}(i^r < i)$ represent the rth nearest neighbor point to $\mathbf{S}_i$, $r = 1,2,\ldots,n$, $i = 1,2,\ldots,N_d$, and $\mathbf{Z}_i^r$ represent the displace vector between them, then

$$\mathbf{Z}_i^r = \mathbf{S}_{i^r} - \mathbf{S}_i \tag{16}$$

After the evolvement of a time interval Δt, the trajectory point $\mathbf{S}_i$ will proceed to $\mathbf{S}_{i+1}$ and the neighboring point $\mathbf{S}_{i^r}$ to $\mathbf{S}_{i^r+1}$. Thus, the vector distance $\mathbf{Z}_i^r$ becomes

$$\mathbf{Z}_{i+1}^r = \mathbf{S}_{i^r+1} - \mathbf{S}_{i+1} \tag{17}$$

Submitting $\mathbf{S}_{i^r+1} = \mathbf{F}^1(\mathbf{S}_{i^r})$ and $\mathbf{S}_{i+1} = \mathbf{F}^1(\mathbf{S}_i)$ into Eq. 17 then gives

$$\mathbf{Z}_{i+1}^r = \mathbf{F}^1(\mathbf{S}_{i^r}) - \mathbf{F}^1(\mathbf{S}_i) = \mathbf{F}^1(\mathbf{Z}_i^r + \mathbf{S}_i) - \mathbf{F}^1(\mathbf{S}_i) \tag{18}$$

Denote $z^r(j)$ as the jth component of $\mathbf{Z}_i^r$, $j = 1,2,\ldots,d$, then $z_{i+1}^r(d)$ can be written as

$$z_{i+1}^r(d) = f_d^1(\mathbf{Z}_i^r + \mathbf{S}_i) - f_d^1(\mathbf{S}_i) \tag{19}$$

Also, the function $f_d^1(\mathbf{Z}_i^r + \mathbf{S}_i)$ is expanded in a Taylor series about the fiducial orbit point $\mathbf{S}_i$, then Eq. 19 becomes the following

$$\begin{aligned} z_{i+1}^r(d) &= \sum_{\beta=1}^{d} z_i^r(\beta) \frac{\partial f_d^1}{\partial s_\beta}\bigg|_{\mathbf{S}_i} \\ &\quad + \sum_{\beta=1}^{d}\sum_{\gamma=1}^{d} z_i^r(\beta) z_i^r(\gamma) \frac{1}{2!} \frac{\partial^2 f_d^1}{\partial s_\beta \partial s_\gamma}\bigg|_{\mathbf{S}_i} + \ldots \\ &= \sum_{\beta=1}^{d} z_i^r(\beta) Df_{d\beta}^1(\mathbf{S}_i) + \sum_{\beta=1}^{d}\sum_{\gamma=1}^{d} z_i^r(\beta) z_i^r(\gamma) D^2 f_{d\beta\gamma}^1(\mathbf{S}_i) + \ldots \end{aligned} \tag{20}$$

Since all $\mathbf{Z}_i^r$, $i = 1,2,\ldots,N_d - 1$ are known, $Df_{d\,\beta}^1(\mathbf{S}_i)$, $D^2 f_{d\,\beta\gamma}^1(\mathbf{S}_i), \ldots$ can be estimated by a least-squares fitting if the number of neighbors to $\mathbf{S}_i$, n, is large enough. The n value depends on the number of unknown parameters included in Eq. 20. Seen from the Eq. 20, there are $d + d^2 + \ldots$ parameters $Df_{d1}^1(\mathbf{S}_i), \ldots, Df_{dd}^1(\mathbf{S}_i)$, $D^2 f_{d11}^1(\mathbf{S}_i), \ldots, D^2 f_{d\,1d}^1(\mathbf{S}_i), \ldots D^2 f_{d\,d1}^1(\mathbf{S}_i), \ldots, D^2 f_{d\,dd}^1(\mathbf{S}_i), \ldots$ needed to be estimated when β, γ change from 1 to d increasingly. In the case of the mixed partial derivatives of function $f_d^1(\cdot)$ existing and being continuous in domain, $D^2 f_{d\,\beta\gamma}^1(\mathbf{S}_i)$ equals to $D^2 f_{d\,\gamma\beta}^1(\mathbf{S}_i)$. Thus, the number of unknown parameters will reduce greatly. In this case, for a given Taylor series of order n_{Tay} and embedding dimension d, the number of unknown parameters, denoted as N_p, in Eq. 20 satisfies[36]

$$N_p = \left(\prod_{k=1}^{n_{\mathrm{Tay}}} \frac{d+k}{k}\right) - 1 \tag{21}$$

Therefore, n is at least N_p for estimating values of parameters via least-squares fitting. Minimizing the average of squared residual error of Eq. 20 would obtain the optimal estimation of $D^2 f_{d\,\beta}^1(\mathbf{S}_i), D^2 f_{d\,\beta\gamma}^1(\mathbf{S}_i), \ldots$. Define $\mathbf{Y}$, $\mathbf{A}$, and $\mathbf{X}$ as

$$\begin{aligned} \mathbf{Y}_{n\times 1} &= (z_{i+1}^1(d), z_{i+1}^2(d), \ldots, z_{i+1}^n(d))^{\mathrm{T}}, \\ \mathbf{A}_{N_p\times 1} &= (Df_{d1}^1(\mathbf{S}_i), \ldots, Df_{dd}^1(\mathbf{S}_i), D^2 f_{d11}^1(\mathbf{S}_i), \ldots, D^2 f_{d1d}^1(\mathbf{S}_i), D^2 f_{d22}^1(\mathbf{S}_i), \ldots, D^2 f_{ddd}^1(\mathbf{S}_i), \ldots)^{\mathrm{T}}, \\ \mathbf{X}_{n\times N_p} &= \begin{pmatrix} z_i^1(1) & z_i^1(2) & \ldots & z_i^1(d) & z_i^1(1)z_i^1(1) & \ldots & z_i^1(1)z_i^1(d) & z_i^1(2)z_i^1(2) & \ldots & z_i^1(d)z_i^1(d) & \ldots \\ z_i^2(1) & z_i^2(2) & \ldots & z_i^2(d) & z_i^2(1)z_i^2(1) & \ldots & z_i^2(1)z_i^2(d) & z_i^2(2)z_i^2(2) & \ldots & z_i^2(d)z_i^2(d) & \ldots \\ \vdots & \vdots & \ldots & \vdots & \vdots & & \vdots\ \ldots & \vdots & \ldots & & \\ z_i^n(1) & z_i^n(2) & \ldots & z_i^n(d) & z_i^n(1)z_i^n(1) & \ldots & z_i^n(1)z_i^n(d) & z_i^n(2)z_i^n(2) & \ldots & z_i^n(d)z_i^n(d) & \ldots \end{pmatrix} \end{aligned} \tag{22}$$

then Eq. 20 may be rewritten as

$$\mathbf{Y} = \mathbf{X}\mathbf{A} \tag{23}$$

Thus, the least-squares fitting is to solve

$$\min_{\mathbf{A}} \Phi = \frac{1}{n}||\mathbf{Y} - \mathbf{X}\mathbf{A}||^2 \tag{24}$$

Let a_k denote the $(k,1)$ component of matrix $\mathbf{A}$, $k = 1,2,\ldots,N_p$, then one can easily get the following expression for $\mathbf{A}$ by setting $\partial\Phi/\partial a_k = 0$

$$\mathbf{A} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y} \tag{25}$$

In the Eq. 24, there is no difference for different neighbor point of $\mathbf{S}_i$ to contribute to fitting parameters. Strictly speaking, the obtained results are coarse in this case. Because different neighbor point has different effect on the model output, weighted least-squares fitting should be adopted to achieve more accurate estimation of parameters. To this task, attention is turned to solve the following problem

$$\min_{\mathbf{A}} \Phi = (\mathbf{Y} - \mathbf{X}\mathbf{A})^{\mathrm{T}}\Lambda(\mathbf{Y} - \mathbf{X}\mathbf{A}) \tag{26a}$$

where Λ is the weighting function matrix and expressed as

$$\Lambda = \begin{bmatrix} \omega_1 & 0 & \ldots & 0 \\ 0 & \omega_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \ldots & \omega_n \end{bmatrix} \tag{26b}$$

Here, ω_k ($k = 1,2,\ldots,n$) is the weighting function of the kth neighbor point to $\mathbf{S}_i$. Typically, ω_k is defined as a function of the distance between $\mathbf{S}_i$ and $\mathbf{S}_{i^k}$. It is generally deemed that the shape of the weighting function has no strong effect on the accuracy of model as long as the weighting function is a non-negative, monotonically decreasing and smooth function of the distance.[32] Based on the fact that the nearest neighbor point has the strongest effect on the model output, the following shape of weighting function is selected for the subsequent investigation

$$\omega_k = \exp(-l_k) \Big/ \sum_{k=1}^{n} \exp(-l_k) \tag{27a}$$

where

$$l_k = \|\mathbf{S}_{i^k} - \mathbf{S}_i\|, \quad k = 1,2,\ldots,n \tag{27b}$$

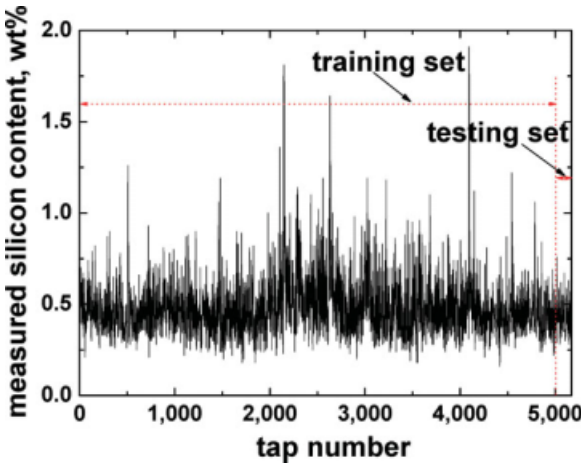

Figure 2. Time series of silicon content in hot metal collected from the selected BF.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Similarly, applying condition, Eq. 26a, and further setting $\partial\Phi/\partial a_k = 0$ will give the expression for $\mathbf{A}$ as follows

$$\mathbf{A} = (\mathbf{X}^{\mathrm{T}}\Lambda\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\Lambda\mathbf{Y} \tag{28}$$

Noted that Eq. 28 is equivalent to Eq. 25 when Λ is an identity matrix $\mathbf{I}$.

Up to now all coefficients $Df^1_{d\,\beta}(\mathbf{S}_i), D^2f^1_{d\,\beta\gamma}(\mathbf{S}_i), \ldots, \beta,\gamma,\ldots = 1,2,\ldots,d$ evaluated along $\mathbf{S}_i$ have been estimated using least-squares fitting. Repeating the above procedure will get $Df^1_{d\,\beta}, D^2f^1_{d\,\beta\gamma}, \ldots$ evaluated along any other fiducial orbit. Thus, the one-step prediction would be realized using Eq. 15, so would be the iterated multistep prediction by combining Eqs. 15 and 13.

The above constructed chaotic predictor is very different from the one reported in our early work[20] where the predictor is in the form of

$$\mathbf{S}_{N_d} = b_1\mathbf{S}_{N_d-1} + b_2\mathbf{1}, \quad b_1, b_2 \in \mathbf{R}, \mathbf{1} = \underbrace{(1,\ldots,1)}_{d}^{\mathrm{T}} \tag{29}$$

i.e., suppose a linear relationship between $\mathbf{S}_{Nd}$ and $\mathbf{S}_{Nd-1}$. The unknown parameters b_1 and b_2 are fitted by following the evolvement law of the neighboring vectors of $\mathbf{S}_{N_d-1}$ in a small ball with radius of ε. Unlike the early model, the current model, Eq. 15a, is based on the Taylor expansion of the project f^1_d but not a simple linear expression. This makes the current model more reasonable and accurate in principle.

Case Study: Chaotic Prediction for Silicon Content in Blast Furnace Hot Metal

Experimental data

In the present work, time series data set of silicon content in BF hot metal (weight, %) collected from a pint-sized BF with a volume of $\sim$750 m^3 is used for understanding and predicting the dynamics of complex BFIP. The actual silicon amounts are measured through analyzing the chemical components of hot metal sample. When the ironmaking ends, the liquid hot metal flows into a ladle from which the hot metal sample is taken, and then sent to laboratory for compositions measurement. To make the analysis results more accurate, two samples are taken per tap at the time of the volume of liquid hot metal attaining one third of the ladle's and two-thirds of the ladle's, respectively. The arithmetical mean of these two analysis results is the desired. The above procedure is often called ladle-wise analysis. Figure 2 exhibits the silicon sequences measured from the selected BF. The silicon time series consists of the values of silicon content observed from continuous 5167 batches of hot metal and the sampling frequency is about 0.5 h^{-1}.

A glance at the Figure 2 may suggest that the oscillating behavior of silicon time series is very complex, and it seems impossible to discover the inherent law and further to predict the evolvement of silicon sequences. A close look, however, reveals a very subtle structure, namely, a short-term cycle is visible and the self-similarity will be shown if the series are overlapped each other. Also, it should be pointed out that there are some large upsets in the silicon sequences in Figure 2. Unlike usual data pretreatment, these obvious large observations are not removed designedly from the data set for the subsequent investigation. The main reason is that the chaos-based predictor depends strongly on the following the historical patterns of time series evolvement, also seen from Eq. 15a, which thus needs the training sample keep continuous or no missing values at the tap number. Therefore, it is unadvisable to remove these large upsets from the data set. To obtain deeper understanding on this data set, some of the important statistical properties, such as mean, standard deviation (SD), etc., should be calculated. Shown in Table 1 are the results, where Pseudo SD tests the normality of the data with PSD $>$ SD representing light tailed distribution, PSD $=$ SD representing normal distribution, and PSD $<$ SD representing heavy tailed distribution; Skew tests the skewness of the data with Skew $>$ 0 representing right or positive skewness, Skew $=$ 0 representing symmetry and Skew $<$ 0 representing left or negative skewness; Kurtosis tests the relative peakness or flatness of the data with Kurtosis $>$ 0 representing peaked distribution, Kurtosis $=$ 0 representing normal distribution and Kurtosis $<$ 0 representing flat distribution.[9] Based on the statistics of the series, it can be concluded that the studied data set has the characteristics of fluctuating tempestuously, heavy tailed distribution, right skewness distribution and peaked distribution. These indicate that the studied data is not normally distributed and cannot act as a good sample for model training,[9] which implies it an arduous

Table 1. Statistical Properties of Silicon Sequences of the Selected BF

Statistical Properties	Value
Minimum value	0.16
Maximum value	1.91
Mean	0.475
Standard deviation	0.154
Pseudo SD	−0.126
Coefficient of variation	0.323
Skew	2.369
Kurtosis	12.643

task to predict hot metal silicon using silicon historical information.

For the valid use of chaos-based iterated multistep predictor, the parameters appearing in the Eq. 22 need to be estimated. Thus, the data set depicted in Figure 2 is segmented into two groups, i.e., the training set and the testing set. The former contains 5000 data points used to estimate the model parameters while the latter contains 167 data points used to verify the validity of chaotic predictor. The basis for partitioning the studied data set as above is the fact that, on the one hand, the chaos-based predictive model is implemented by finding analogs of past events,[34] which implies that the model will fail to predict new pattern, and thereby the size of training set should be large enough to include past events as many as possible, on the other hand, the size of testing set should be large enough to make the prediction results more convinced. Of course, a possible disadvantage of choosing long training set is that there may be inconsistency in finding analogous past events, e.g., analogous past events may give contrary evolvement of future. As a result, the prediction accuracy may be reduced if a contrary historical pattern is chosen for prediction. Investigation of the effect of a particular training set and a test horizon on the prediction accuracy is a plagued but worthwhile task. In this work, the possible disadvantage of long training set is ignored to obtain more historical patterns in the record. Therefore, the training set is chosen as long as possible from the available actual data. In all the process of making prediction, the size of training set remains unchanged while the training sample is updated continuously when a new prediction is implemented. Namely, the first point in the last training set will be removed and the new observed silicon content will be included to constitute the new training set for next prediction.

Determination of the reconstructed parameters

In the chaotic predictor, Eq. 15, there are many parameters required to be estimated before making prediction. First of all, the reconstructed parameters of delay time τ and embedding dimension d, reflecting some important features of real system, should be determined. With respect to τ, it may be arbitrary for an infinite and noise free time series in principle.[29] However, for a real finite time series it should be chosen carefully. An accurate delay time can not only make the reconstructed vectors maximally independent but also preserve the dynamical information of these vectors. There exist lots of methods for determining the delay time including quantitative analysis such as autocorrelation function,[37] mutual information[38] and forecast entropy,[39] and qualitative analysis such as comparison of phase space picture with pseudo-phase space picture[30] which is obtained by plotting $s(t_i)$ vs. its τ shifted version, $s(t_i + \tau\Delta t)$. Among the quantitative analysis methods, mutual information technique[38] is being broadly adopted for its superior performance in measuring a more general relationship between two variables. Based on this technique,[38] the best choice of τ corresponds to the first local minimum of mutual information. However, even being a popular method, it is not always applied successfully in dealing with practical time series. For instance, some investigators[27,40,41] reported that there is not the minimum mutual information within any reasonable area in their studied cases. Analogous results are encountered in our experimental data. Therefore, it is necessary to use other techniques to determine the τ. Our earlier work[21] indicated that $\tau = 1$ is suited for reconstructing phase space when a shorter silicon record collected from the same blast furnace is analyzed through forecast entropy technique. In this regard, $\tau = 1$ seems available for the present study. Here, as an assistant means, the qualitative analysis method,[30] i.e., comparison of phase plots to pseudo-phase plots, is employed to make further verification of the result of delay time. Compared with the quantitative analysis methods, the qualitative analysis seems to be rough, but it is often designated to determine delay time for reconstructing phase space by finding similar pseudo-phase plots to phase plots,[31] and thus yields broad application in practice,[27,40] especially in the case that the quantitative analysis methods fail. Hence, it is still of practical significance to make delay time analysis using the qualitative method. Based on this method, an appropriate τ is dependent on the geometrical similarity of the phase portrait of experimental data dynamics to its pseudo-phase portrait.[30] Figures 3 and 4 exhibit the phase portrait and pseudo-phase portrait of observed silicon signal dynamics, respectively. Noticeably, only delays of up to $\tau = 6$ are included in Figure 4 since low delay time parts are usually interesting for reconstructing phase space. Comparison of Figure 3 with Figure 4 indicates that the geometric structure of Figure 4a is more like that of Figure 3 than others. Thus, $\tau = 1$ is believed to be the appropriate delay time for phase space reconstruction.

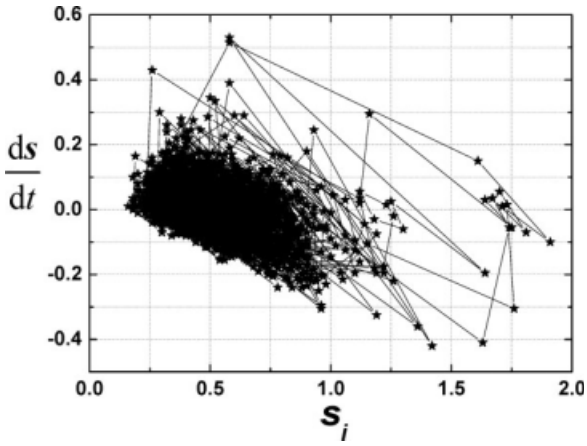

Figure 3. Phase portrait of silicon sequences collected from the selected BF.

Here, the derivative ds/dt is approximated by $ds/dt \approx \Delta s/\Delta t = (s_i - s_{i-1})/\Delta t$.

Like τ, embedding dimension d is also an important parameter for reconstructing phase space. The optimal d should satisfy that the reconstructed orbits do not intersect or overlap with each other. G-P algorithm,[42] false neighbors,[43] singular value decomposition,[44] C-C method,[45] forecast entropy,[39] etc. are the frequently adopted methods to estimate d. Among these methods, the way of false neighbors is to look for false neighbors in phase space at a given embedding dimension. The criterion for determining optimal

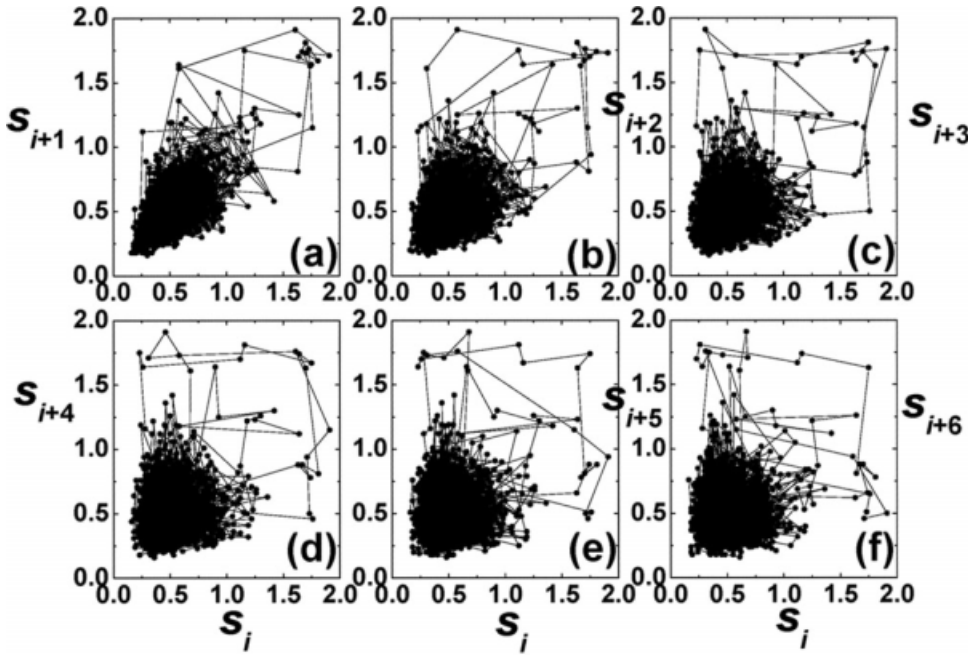

Figure 4. Pseudo phase portrait of silicon sequences collected from the selected BF at low delay time: (a) $\tau = 1$; (b) $\tau = 2$; (c) $\tau = 3$; (d) $\tau = 4$; (e) $\tau = 5$; (f) $\tau = 6$.

d depends on the information at which embedding dimension the percentage of false neighbors will drop to zero. Compared with other methods, false neighbor method has advantages of computing simply and guaranteeing to find the minimum embedding dimension. The result of optimal d, however, may be very different if different threshold values are chosen in saying that a neighbor is false. In order to avoid this subjectivity, a modified version of false neighbor was presented by Cao.[46] In his work, a variable only dependent on d and τ is defined as

$$E(d) = \frac{1}{N_{\text{dat}} - \tau} \sum_{i=1}^{N_{\text{dat}}-\tau} a(i,d) \tag{30a}$$

where

$$a(i,d) = \frac{\|\mathbf{S}_i(d+1) - \mathbf{S}_{i^n}(d+1)\|_{l_\infty}}{\|\mathbf{S}_i - \mathbf{S}_{i^n}\|_{l_\infty}}, \quad i = 1,2,\ldots,N_{\text{dat}} - d\tau \tag{30b}$$

Here $\mathbf{S}_{i^n}$ is the nearest neighbor point of $\mathbf{S}_i$ in terms of l_∞ norm and the suffix i^n satisfies $1 \le i^n \le N_{\text{dat}} - d\tau$, $i^n \ne i$; $\mathbf{S}_i(d+1) = (s_i, s_{i+\tau}, \ldots, s_{i+(d-1)\tau}, s_{i+d\tau})$ and $\mathbf{S}_{i^n}(d+1) = (s_{i^n}, s_{i^n+\tau}, \ldots, s_{i^n+(d-1)\tau}, s_{i^n+d\tau})$. If $\|\mathbf{S}_i - \mathbf{S}_{i^n}\|_{l_\infty} = 0$, the second nearest neighbor is selected, and so forth. The criterion of determining d is to observe the starting point of plateau of the following expression

$$E1(d) = E(d+1)/E(d) \tag{31}$$

with d increasing. The optimal embedding dimension is considered as the one at which $E1(d)$ stops changing.

Figure 5 shows the dependence of $E1(d)$ and its relative variation on embedding dimension. Clearly, at low embedding dimension, $E1(d)$ drops rapidly and a big dip emerges at $d = 3$ after which $E1(d)$ increases slowly and nearly stops changing at $d = 18$. Therefore, 18 is the optimal candidate of embedding dimension for the selected BF. This indicates that the evolvement of silicon content in BF hot metal could be viewed from a high-dimensional chaotic dynamic perspective, and 18 independent macroscopic variables are at least required to capture the important features of underlying dynamics. Unfortunately, there is not any information on the candidate variables at this stage. On the basis of the work of Bhattacharya[15] and Saxén and Pettersson,[12] these variables

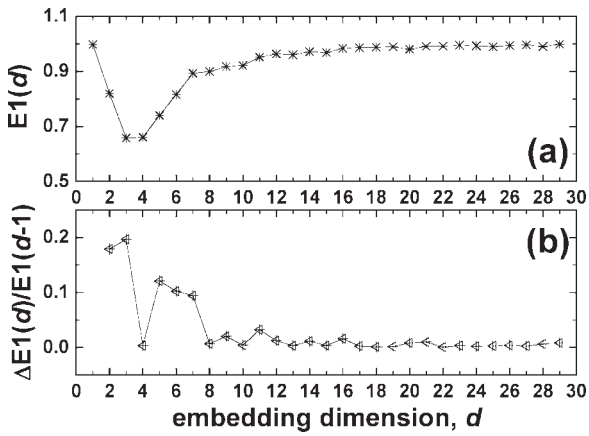

Figure 5. Dependence of $E1(d)$ and its relative variety on embedding dimension: (a) $E1(d)$; (b) $\Delta E1(d)/E1(d-1) = [E1(d) - E1(d-1)]/E1(d-1)$.

may be blast volume, pressure and temperature, gas permeability, oxygen enrichment, pulverized coal injection rate, ore-to-coke ratio, flame temperature, energy at tuyeres, solution loss rate, tuyere heat loss, gas CO utilization, the sum of CO and CO_2 content, temperature of top gas and time elapsed from last taping, etc. In addition, the historical values of some variables over the last 8 h[9] and the specific blast volume, pressure and temperature,[12] expressed per ton of hot metal, are also considered to have important effect on the silicon content in hot metal. More metallurgical knowledge and statistical analysis (such as principle component analysis, etc.) are needed to single out the candidate 18 independent variables from those involved in BFIP.

Predictor implementation and performance evaluation

In the process of implementing chaotic predictor, except for the reconstructed parameters d and τ, there are $Df^1_{d\,\beta}, D^2f^1_{d\,\beta\gamma}, \ldots$ required to be evaluated along a certain specific fiducial orbit. These parameters values will have a great effect on the prediction results, and thus should be treated with caution. Seen from Eq. 25 or 28, in the case of given reconstructed parameters, the values of $Df^1_{d\,\beta}, D^2f^1_{d\,\beta\gamma}, \ldots$ rely mainly on the properties of neighbors of the given fiducial orbit, such as the number of neighbors, n. In principle, n is an integer lying in $[N_p,N_d)$, but its exact value is not known a priori. Therefore, to make the chaotic predictor optimal, a trial and error procedure has to be utilized to find the optimal n, i.e., by increasing n and choosing the one that gives the best prediction results. Prediction success is evaluated by measuring the agreement between the actual values and predicted values in terms of the following criteria:

(i) the percentage of hitting the target

$$e_1 = \frac{1}{N}\left(\sum_{k=1}^{N} H_k\right) \times 100\% \tag{32a}$$

where N is the size of the testing set,

$$H_k = \begin{cases} 1 & |s_k - \hat{s}_{k|k-p}| t < 0.1 \\ 0 & \text{else} \end{cases} \tag{32b}$$

$\hat{s}_{k|k-p}$ represents the p-step predicted value ahead at time $(k-p)$, which means that the prediction is successful in the case of $\left|s_k - \hat{s}_{k|k-p}\right| < 0.1$;

(ii) the correlation coefficient between the actual values and the predicted values

$$e_2 = \frac{\frac{1}{N}\sum_{k=1}^{N}\left(s_k - \langle s\rangle\right)\left(\hat{s}_{k|k-p} - \langle \hat{s}\rangle\right)}{\sigma(s)\sigma(\hat{s})} \tag{33}$$

where $\langle s\rangle$, $\langle \hat{s}\rangle$ are the average values of the actual values and the predicted values, respectively, and $\sigma(s)$, $\sigma(\hat{s})$ representing their standard deviation;

(iii) the root mean square error

$$e_3 = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(s_k - \hat{s}_{k|k-p}\right)^2} \tag{34}$$

and (iv) the coefficient of efficiency defined as the ratio of the mean square error to the variance of testing set, then subtracted from 1, i.e.,

$$e_4 = 1 - \frac{\frac{1}{N}\sum_{k=1}^{N}\left(s_k - \hat{s}_{k|k-p}\right)^2}{\sigma^2(s)} \tag{35}$$

Among these criteria, e_1 is the most important one in metallurgical field and can provide important guide for taking correct actions on operating a BF; e_2 and e_3 are the universal criteria for evaluating the performance of a predictor; e_4 lies in $(-\infty,1]$ and is used to decide whether the mean square error is large or not, with the value closer to 1 indicating better prediction.

Based on the above decided reconstructed parameters values and evaluating criteria, the testing set presented in Figure 2 is predicted with one-step chaotic predictor for optimizing the number of neighbors. At this stage, Eq. 25, i.e., in the case of $\Lambda = \mathbf{I}$, is used to fitting parameters $Df^1_{d\,\beta}, D^2f^1_{d\,\beta\gamma}, \ldots$. Since each evaluating criterion is not a continuous function of n, it is impossible to achieve the optimal n through rigorous mathematical analysis. Thus, the procedure of optimizing n is carried out in manner of relative coarseness, i.e., by increasing n from 500 to 4500 with an increment of 500 every time, firstly; then finding the local optimal n; afterwards continuing to optimize n around the local optimal n with an increment of 25 every time; finally obtaining approximate optimal n. Figure 6 exhibits the optimization plot of number of neighbors for the selected BF. Here, the first-order and second-order Taylor expansion series (also called the order of chaotic predictor in the following text) of $s_{N_{\text{dat}}+1}$ about the nearest neighbors point $\mathbf{S}_{nN_d}$ are considered, respectively.

As can be seen from the left part of Figure 6, the local optimal number of neighbors emerges at $n = 1000$ and $n = 2500$ in the case of $n_{\text{Tay}} = 1$ and $n_{\text{Tay}} = 2$, respectively, around which more careful optimization process is performed. Shown in the right part of Figure 6 are the enlargement plots of the selected areas with ellipse, from which it is easy to conclude that the approximate optimal n is 1225 at $n_{\text{Tay}} = 1$ because all the evaluating criteria are consistent, while it is difficult to find the optimal n at $n_{\text{Tay}} = 2$ because of the inconsistency from different evaluating criterion. Here, given the importance of e_1 in the metallurgical field, the approximate optimal n at $n_{\text{Tay}} = 2$ is taken as $n = 2600$ in terms of the maximum e_1. The detailed results may be found in Table 2. Therefore, as the selected sample is concerned, the approximate optimal number of neighbors is 1225 at $n_{\text{Tay}} = 1$ and 2600 at $n_{\text{Tay}} = 2$ for the designed one-step chaotic predictor, respectively. These optimal numbers of neighbors at corresponding order of chaotic predictor will be adopted to make one-step prediction and multistep prediction of silicon content in the subsequent study.

Having all parameters values, the local chaotic predictor, Eq. 15, is now employed for predicting silicon content in hot metal collected from the selected BF. Remarkably, here weighed least-square method, i.e., Eq. 28 is used to fit parameters $Df^1_{d\,\beta}, D^2f^1_{d\,\beta\gamma}, \ldots$ for the purpose of improving prediction accuracy. The one-step, direct two-step and iterated two-step prediction of silicon content are made at different

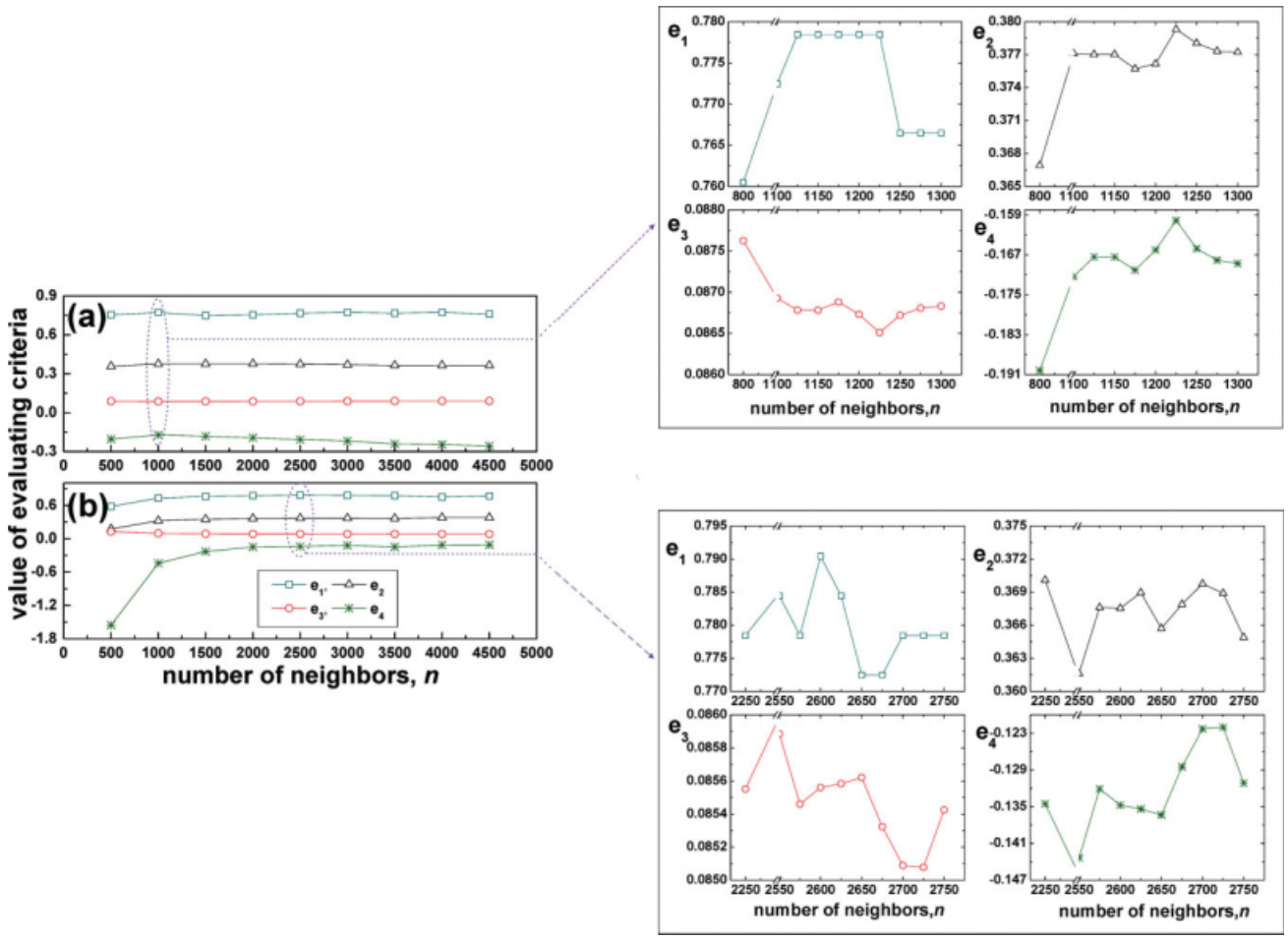

Figure 6. Prediction accuracy vs. number of neighbors at different order of chaotic predictor for the purpose of optimization: (a) $n_{Tay} = 1$; (b) $n_{Tay} = 2$.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

order of chaotic predictor. During the process of making iterated two-step prediction, the one-step predicted values at $n_{Tay} = 2$ are taken to form the reconstructed vectors according to Eq. 13. Displayed in Figures 7–9 are the results of one-step, direct two-step and iterated two-step prediction, respectively. In these figures, the two dotted horizontal lines indicate ± 0.1, the acceptable boundary of prediction error for saying that the prediction is successful in terms of e_1; $n_{Tay} = 0$ means that a zeroth-order Taylor expansion series of $s_{N_{dat}+1}$ is taken as a predictor, i.e., directly taking the evolvement of the dth entry of $\mathbf{S}_{nN_d}$ as the predicted value, which is the simplest local prediction method. A summary of these prediction results is reported in Table 3. From the results displayed in Figures 7–9 and Table 3, the following information is clear: (i) in the case of making prediction with the same type, such as one-step prediction, the prediction accuracy is basically increasing as the chaotic predictor changes from zeroth-order to second-order; (ii) in the case of employing chaotic predictor with the same order, one-step prediction has better results than two-step prediction; (iii) as the earlier investigators reported,[32,33] iterated two-step chaotic predictor has superior performance to direct two-step predictor except at $n_{Tay} = 0$ in terms of e_2 and at $n_{Tay} = 2$ in terms of e_1. These phenomena can be easily explained: (i) because higher order Taylor expansion series can approach the desired function with higher precision in theory, second-

Table 2. The Optimal Number of Neighbors at Different Order of Chaotic Predictor for the Selected BF

Order of Chaotic Predictor (n_{Tay})	Optimal Number of Neighbors (n)	Evaluating Criteria			
		e_1	e_2	e_3	e_4
1	1225	77.84%	0.3793	0.08651	−0.1601
2	2600	79.04%	0.3675	0.08556	−0.1348

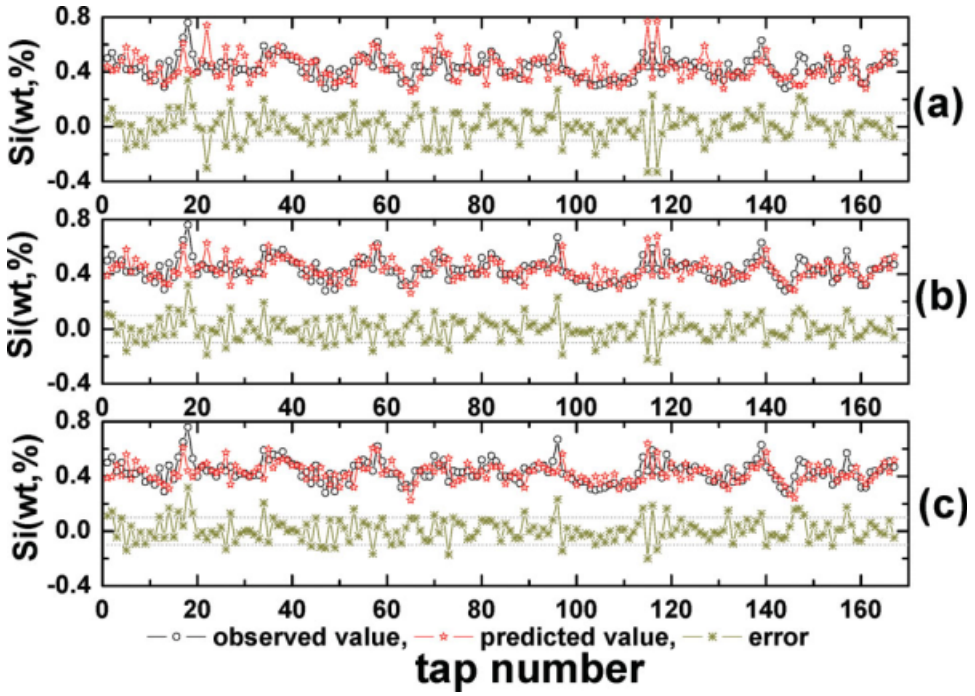

Figure 7. One-step prediction for the selected BF at different order of chaotic predictor: (a) $n_{Tay} = 0$; (b) $n_{Tay} = 1$; (c) $n_{Tay} = 2$.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

order chaotic predictor is more accurate than first-order and zeroth-order one; (ii) since multistep predictor is to forecast the future value at more time step ahead than one-step predictor, the historical information is little for the former; (iii) iterated two-step predictor takes one-step predicted value as new future information, then based on this, continues to

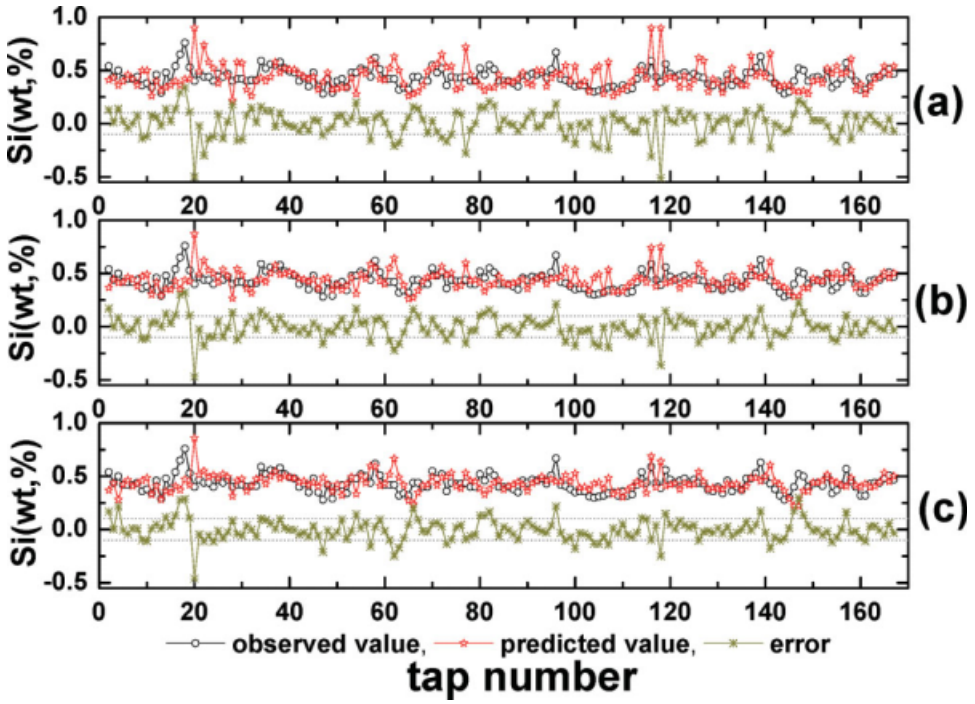

Figure 8. Direct two-step prediction for the selected BF at different order of chaotic predictor: (a) $n_{Tay} = 0$; (b) $n_{Tay} = 1$; (c) $n_{Tay} = 2$.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

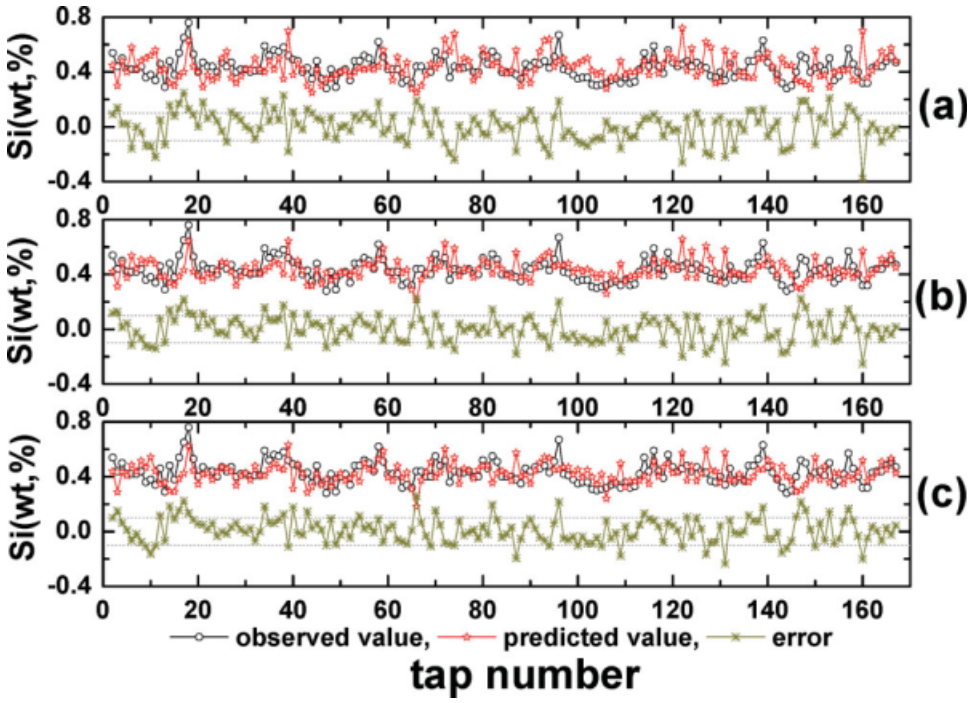

Figure 9. Iterated two-step prediction for the selected BF at different order of chaotic predictor: (a) $n_{Tay} = 0$; (b) $n_{Tay} = 1$; (c) $n_{Tay} = 2$.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

make one-step prediction, which is deemed to utilizing more information to predict future value, therefore it has better results than direct two-step prediction.

A further look at these figures and Table 3 may suggest that the designed local chaotic predictor can represent the evolvement of silicon content in hot metal measured from the studied BF well from the view point of practicality. Not only the minor fluctuations but also some major trends in the silicon sequences are captured well. As far as one-step prediction is concerned, the percentage of hitting the target, 79.04%, and the root mean square error, 0.08538, are very encouraging. Even for difficult two-step prediction, these two criteria attain 73.49% and 0.09288, respectively, with iterated multistep prediction method. These results reveal that the predicted values of silicon content are in reasonable agreement with the experimental data, which further indicate that an 18-dimensional local chaotic predictor is suitable for understanding the dynamics of silicon sequences for the selected case. Although some literature report better result, such as $e_1 = 88.7\%$[10] and $e_3 = 0.074$,[12] the current predictive model has still enough competitive power compared with other existing models. The main dominance is that only silicon content historical data is utilized for building predictive model in this study while many other variables involved in BFIP, such as blast volume, blast temperature, coke rate, etc. are considered for modeling in the literature,[10,12] even including expert knowledge.[10] The current chaotic predictive model can work well even if other variables involved in BFIP are scarce. Furthermore, multistep prediction of silicon content is easily realized using the current predictive model. Actually, it is not objective to say a model better than others just on the basis of the value of evaluating criteria. This is

Table 3. Prediction Results at Different Order of Chaotic Predictor for the Selected BF

Type of Prediction	Evaluating Criteria	$n_{Tay} = 0$	$n_{Tay} = 1$	$n_{Tay} = 2$
One-step prediction	e_1	71.26%	77.84%	79.04%
	e_2	0.3214	0.3793	0.3688
	e_3	0.1022	0.08648	0.08538
	e_4	−0.6199	−0.1593	−0.1299
Direct two-step prediction	e_1	62.65%	71.08%	75.30%
	e_2	0.2209	0.2256	0.2139
	e_3	0.1264	0.1048	0.1010
	e_4	−1.4705	−0.7001	−0.5790
Iterated two-step prediction	e_1	64.46%	71.69%	73.49%
	e_2	0.1992	0.2703	0.2778
	e_3	0.1100	0.09333	0.09288
	e_4	−0.8712	−0.3474	−0.3346

especially true for evaluating the performance of silicon content predictor. A scientific evaluation should compare the performance criteria achieved from the same BF, including the same training set and same testing set, with different predictors. Since for a single BF, the performance criteria, such as the percentage of hitting the target, may be quite different using different training set. Even the size of testing set, N, has a great effect on evaluating criteria. To put the findings of the study, listed in Table 3, in a more relevant perspective, the silicon values of testing set shown in Figure 2 are also made one-step prediction by a naive prediction model, $\hat{s_{i+1}} = s_i$. In this way the hit rate of predicting silicon, e_1, is about 74.85%, which is lower than the results listed in Table 3 when $n_{\mathrm{Tay}} = 1$ and $n_{\mathrm{Tay}} = 2$. It seems that such result is basically satisfied. However, compared with the established chaotic predictor, this simple prediction model is just an empirical model and lack of potential to be further improved.

Based on the above prediction results, it is obvious that a high dimensional chaotic system can capture the important features of silicon series evolvement dynamics well for the studied BF. This in turn renders strong indication of the presence of a deterministic mechanism ruling the dynamics of BFIP. In the meantime, the good but not perfect prediction results imply that the underlying dynamics has a deterministic component but is not completely deterministic. A possible application of the current results is that a deeper understanding on the complex BFIP is obtained which can provide a novel tool, chaos, for characterizing BFIP, including prediction, control, and optimization. In addition, the predicted silicon values may also act as a suggestion for the operators to make appropriate control actions in advance, which further serves as a means for optimizing the quality of hot metal and decreasing the production costs, or directly act as output error feedback for correcting the next decision during the process optimization. At the same time, it should be pointed out that the current chaotic predictive model is a kind of time series model in nature. Some intrinsic limitation of time series model, such as strong inertia, being only able to predict the existing pattern in the training set but failing to predict the new pattern, are difficult to be overcome. Such shortcomings can be observed from Figure 7 to 9, namely, some predicted values have a short time lag with the corresponding experimental data. A possible way to weaken the effect of strong inertia limitation is to replace the autoregressive terms of the present model by other key variables involved in BFIP through multivariable time series phase space reconstruction. Generally speaking, the autoregressive terms in the silicon predictive model are considered as the source of producing high inertia.[12] The second limitation may, in theory, be overcome by increasing the size of the training set so that more historical law may be learnt for predicting future. Investigations on improving the accuracy of chaotic predictor in these directions are underway.

Conclusions and Points of Possible Future Research

The main contribution of this article is to establish a chaos-based iterated multistep predictor for predicting silicon content in hot metal produced from complex blast furnace ironmaking process. Through a typical example, a time series data of silicon content in hot metal measured from a pint-sized BF, the designed predictor exhibits good performance. The acceptable percentage of hitting the target and low root mean square error, 79.04%, 0.08538 and 73.49%, 0.09288 for one-step prediction and iterated two-step prediction, respectively, indicate that the reconstructed 18 dimensional chaotic system can explain the studied silicon sequence dynamics well, which further indicate that there is a deterministic mechanism ruling the dynamics of BFIP. Such results may serve as guidelines for characterizing complex BFIP using chaos in the future work.

However, it should also be noted that, despite knowing an 18 dimensional chaotic system suitable for representing the dynamics of the studied BFIP, these 18 independent variables are unknown at this stage. Multidimensional time series phase space reconstruction or principle component analysis may provide more information about this. As a kind of time series model, the established chaotic predictive model has inevitably intrinsic limitation of time series model, such as high inertia, failing to predict new pattern, etc. Taking other important variables involved in BFIP, such as blast volume, blast temperature, coke rate, etc. into account for constructing new chaotic predictor is an emphasis of future investigation to reduce high inertia of model. The incorporation of these exogenous inputs, i.e., constructing a multivariate chaotic predictor, is deemed to be able to further improve the prediction accuracy and also throw more light on how to control complex ironmaking process. In the meantime, increasing the size of the training set, judiciously selecting the training set and testing set, extending the order of chaotic predictor, and applying chaotic predictive method to other blast furnace system are all worth investigating. Additionally, the established chaos-based predictor is a data-driven model in nature, which does not consider the effect of chemical reactions related to silicon, i.e., silicon transfer mechanism, on the model predictions. More future investigation may focus on constructing a silicon prediction model taking process mechanism, production data, and even expert experience, into consideration.

In conclusion, notwithstanding some limitations, the chaos-based iterated multistep predictor is an alternative tool to predict the evolvement of silicon content in blast furnace hot metal, especially in the absence of other variables information involved in BFIP. It is hoped that this method can promote the predictive technology of silicon content in blast furnace hot metal progress, and provide a possibility for simulating complex BFIP.

Acknowledgements

The authors are grateful to 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, Joint Funds of NSFC-Guangdong under Grant No. U0735003, and Natural Science Foundation of China under Grant No. 10826100, 60604029 for financial support of this work.

Notation

$\mathbf{1}$ = a d-dimensional column vector with all entries 1
BFIP = blast furnace ironmaking process
BF = blast furnace
b_1,b_2 = unknown parameters in Eq. 29
Df = coefficients of first-order Taylor expansion series

D^2f = coefficients of second-order Taylor expansion series
d = embedding dimension
d_M = box-counting dimension of attractor M
$E1$ = criterion to determining embedding dimension
e_1 = percentage of hitting the target
e_2 = correlation coefficient
e_3 = root mean square error
e_4 = coefficient of efficiency
$\mathbf{F}$ = a map vector
$\hat{\mathbf{F}}$ = approximation of $\mathbf{F}$
$\mathbf{F}^1$ = one-step map
$\hat{\mathbf{F}}^1$ = approximation of one-step map
$\mathbf{F}^p$ = p-step map
f = the entry of map vector $\mathbf{F}$
h = measurement function of $\mathbf{S}$
$\mathbf{I}$ = identity matrix
k = circle index
l = Euclidean distance
l_∞ = a kind of norm
M = a finite-dimensional attractor
m = dimension of dynamical system
N = size of the testing set
N_d = number of reconstructed phase space
N_{dat} = size of silicon time series
Np = number of unknown parameters emerging in Eq. 20
n = number of nearest neighbor points
n_{Tay} = order of Taylor expansion series
PSD = pseudo standard deviation
p = ahead step of prediction
r = parameter index
SD = standard deviation
$\mathbf{S}$ = reconstructed vector
$\hat{\mathbf{S}}$ = predicted value of $\mathbf{S}$
s = silicon content in blast furnace hot metal, weight, %
$\hat{s}$ = predicted value of s
$\langle s \rangle$ = mean value of s
$\langle \hat{s} \rangle$ = mean value of $\hat{s}$
t = sampling time, h
Δt = sampling interval, h^{-1}
$\mathbf{u}$ = state variable of m-dimensional dynamical system
$\mathbf{Y}$, $\mathbf{A}$, $\mathbf{X}$ = defined matrix in Eq. 22
$\mathbf{Z}$ = displace vector
z = entry of $\mathbf{Z}$

Greek letters

α, β, γ = parameter index
Φ = performance index
Λ = weighting function matrix
σ = standard deviation
τ = delay time
ω = weighting function

Superscript

T = transpose of a matrix

Subscripts

i, i^r, j = parameter index
$k|k-p$ = p-step predicted value ahead at time $(k - p)$

Literature Cited

1. Ozturk B, Fruehan RJ. Kinetics of the reaction of SiO(g) with carbon saturated iron. *Metall Trans B*. 1985;16:121–127.
2. Sugawara K, Morimoto K, Sugawara T, Dranoff JS. Dynamic behavior of iron forms in rapid reduction of carbon-coated iron ore. *AIChE J*. 1999;45:574–580.
3. Bi XG, Torrssel K, Wijk O. Prediction of the blast furnace process by a mathematical model. *ISIJ Int*. 1992;32:481–488.
4. Nogami H, Chu MS, Yagi J. Multi-dimensional transient mathematical simulator of blast furnace process based on multi-fluid and kinetic theories. *Comput Chem Eng*. 2005;29:2438–2448.
5. Nishioka K, Maeda T, Shimizu M. A three-dimensional mathematical modeling of drainage behavior in blast furnace hearth. *ISIJ Int*. 2005;45:669–676.
6. Waller M, Saxén H. On the development of predictive models with applications to a metallurgical process. *Ind Eng Chem Res*. 2000;39:982–988.
7. Waller M, Saxén H. Time-varying event-internal trends in predictive modeling-methods with applications to ladlewise analyses of hot metal silicon content. *Ind Eng Chem Res*. 2003;42:85–90.
8. Singh H, Sridhar NV, Deo B. Artificial neural nets for prediction of silicon content of blast furnace hot metal. *Steel Res*. 1996;67:521–527.
9. Radhakrishnan VR, Mohamed AR. Neural networks for the identification and control of blast furnace hot metal quality. *J Process Control*. 2000;10:509–524.
10. Chen J. A predictive system for blast furnaces by integrating a neural network with qualitative analysis. *Eng Appl Artif Intell*. 2001;14:77–85.
11. Jiménez J, Mochón J, de Ayalai JS, Obeso F. Blast furnace hot metal temperature prediction through neural networks-based models. *ISIJ Int*. 2004;44:573–580.
12. Saxén H, Pettersson F. Nonlinear prediction of the hot metal silicon content in the blast furnace. *ISIJ Int*. 2007;47:1732–1737.
13. Miyano T, Kimoto S, Shibuta H, Nakashima K, Ikenaga Y, Aihara K. Time series analysis and prediction on complex dynamical behavior observed in a blast furnace. *Physica D*. 2000;135:305–330.
14. Waller M, Saxén H. Application of nonlinear time series analysis to the prediction of silicon content of pig iron. *ISIJ Int*. 2002;42:316–318.
15. Bhattacharya T. Prediction of silicon content in blast furnace hot metal using partial least squares (PLS). *ISIJ Int*. 2005;45:1943–1945.
16. Hao XJ, Shen FM, Du G, Shen YS, Xie Z. A blast furnace prediction model combining neural network with partial least square regression. *Steel Res Int*. 2005;76:694–699.
17. Martín RD, Obeso F, Mochón J,Barea1 R Jiménez J. Hot metal temperature prediction in blast furnace using advanced model based on fuzzy logic tools. *Ironmak Steelmak*. 2007;34:241–247.
18. Zhou ZM. Measurement of time-dependent fractal dimension for time series of silicon content in pig iron. *Physica A*. 2007;376:133–138.
19. Gao CH, Qian JX. Time-dependent fractal characteristics on time series of silicon content in hot metal of blast furnace. *ISIJ Int*. 2005;45:1269–1271.
20. Gao CH, Zhou ZM, Qian JX. Chaotic identification and prediction of silicon content in hot metal. *J Iron Steel Res Int*. 2005;12:3.
21. Gao CH, Zhou ZM, Chen JM. Assessing the predictability for blast furnace system through nonlinear time series analysis. *Ind Eng Chem Res*. 2008;47:3037–3045.
22. Ingraham RL. *A Survey of Nonlinear Dynamics ("Chaos Theory")*. Singapore: World Scientific, 1992.
23. Nightingale RJ, Dippenaar RJ, Lu WK. Developments in blast furnace process control at Port Kembla based on process fundamentals. *Metall Mater Trans B*. 2000;31:993–1003.
24. Saxén H. Simplified simulation of the transient behavior of temperatures in the upper shaft of the blast furnace. *Metall Mater Trans B*. 1998;29:691–697.
25. Lorenz EN. Dimension of weather and climate attractors. *Nature*. 1991;353:241–244.
26. Brock WA, Hommes CH. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *J Econ Dyn Control*. 2000;24:799–831.
27. Zhao GB, Chen JZ, Yang YR. Predictive model and deterministic mechanism in a bubbling fluidized bed. *AIChE J*. 2001;47:1524–1532.
28. Ng WW, Panu US, Lennox WC. Chaos based analytical techniques for daily extreme hydrological observations. *J Hydrol*. 2007;342:17–41.
29. Takens F. *Detecting strange attractor in fluid turbulence*. In: Rand D, Young LS, editors. *Lecture Notes in Mathematics*. Berlin: Springer, 1981:366–381.
30. Packard NH, Crutchfield JP, Farmer JD, Shaw RS. Geometry from a time series. *Phys Rev Lett*. 1980;45:712–716.

31. Moon FC. *Chaotic and Fractal Dynamics: An Introduction for Applied Scientists and Engineers*. New York: Wiley, 1992.
32. McNames J. Local averaging optimization for chaotic time series prediction. *Neurocomputing*. 2002;48:279–297.
33. Casdagli M. Nonlinear prediction of chaotic time series. *Physica D*. 1989;35:335–356.
34. Farmer JD, Sidorowich JJ. Predicting chaotic time series. *Phys Rev Lett*. 1987;59:845–848.
35. Sano M, Sawada Y. Measurement of Lyapunov spectrum from a chaotic time series. *Phys Rev Lett*. 1985;55:1082–1085.
36. Brown R, Bryant P. Computing the Lyapunov spectrum of a dynamical system from an observed time series. *Phys Rev A*. 1991;43:2787–2806.
37. Albano AM, Muench J, Schwartz C. Singular-value decomposition and the Grassberger-Procaccia algorithm. *Phys Rev A*. 1998;38: 3017–3026.
38. Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. *Phys Rev A*. 1996;33:1134–1140.
39. Yao WG, Essex C, Yu P, Davison M. Measure of predictability. *Phys Rev E*. 2004;69:066121.
40. Daw CS, Halow JS. Evaluation and control of fluidization quality through chaotic time series analysis of pressure drop measurements. *AIChE Symp Ser*. 1993;89:103–122.
41. Pence DV, Beasley DE, Riester JB. Deterministic chaotic behavior of heat transfer in gas fluidized beds. *J Heat Transfer*. 1995;117:465–472.
42. Grassberger P, Procaccia I. Characterization of strange attractors. *Phys Rev Lett*. 1983;50:346–349.
43. Kennel MB, Brown R, Abarbanel HDI. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys Rev A*. 1992;45:3403–3411.
44. Broomhead DS, King GP. Extracting qualitative dynamics from experimental data. *Physica D*. 1986;20:217–236.
45. Kim HS, Eykholt R, Salas JD. Nonlinear dynamics, delay time, and embedding windows. *Physica D*. 1999;127:48–60.
46. Cao LY. Practical method for determining the minimum embedding dimension of a scale time series. *Physica D*. 1997;110:43–50.

Manuscript received Jun. 4, 2008, and revision received Sept. 21, 2008.