

高炉铁水含硅量序列的动力学结构分析*

赵 敏 刘祥官 郜传厚†

(浙江大学数学系, 系统优化技术研究所, 杭州 310027)

(2007 年 6 月 4 日收到, 2007 年 9 月 27 日收到修改稿)

以邯钢钢铁公司 7 号高炉在线采集的 2000 炉铁水含硅量 $[Si]$ 数据为样本, 对 $[Si]$ 时间序列作了基于逆序数的平稳性检验. 然后, 在关联积分的基础上, 定义了衡量不同时间序列间动力学相似性的“距离”, 通过等分采集得到的 $[Si]$ 序列, 计算子序列间的“距离”, 发现了高炉冶炼过程中存在显著的动力学结构突变性, 最后应用 DVV 算法分析动力学性质变动下, 高炉铁水含硅量 $[Si]$ 的可预测性.

关键词: 高炉冶炼, 平稳性, 动力学相似性, 可预测性

PACC: 0545

1. 引 言

高炉冶炼系统的建模和控制是当前冶金科技发展的前沿课题, 其中高炉炉温的预测和控制是难点所在^[1]. 在高炉炼铁工艺中, 通常以铁水硅含量 $[Si]$ 反映高炉炉内的热状况(炉温) $[Si]$ 的控制与高炉冶炼过程的稳定性、生产效率(利用系数)、能耗和铁水质量之间有密切关系, 对 $[Si]$ 的预报和控制往往是衡量高炉工长操作水平的重要依据^[2], 因而本文选取了 $[Si]$ 序列作为分析的对象.

对时间序列进行分析时, 大多数的研究都是建立在时间序列是平稳的这一条件成立基础上的; 因此对实际工业系统观测到的时间序列(例如 $[Si]$ 序列)做平稳性检验显得十分必要. 当序列不满足平稳性时, 数学上至关重要的遍历性定理不复成立^[3]. 影响系统的某些参数或条件随时间发生的调整造成了整个系统内在演化方程(动力学结构)的显著改变, 从而在观测序列中体现了出来^[4].

目前关于时间序列动力学结构突变性研究的理论和方法还比较少, 并没有一套使人们普遍接受的方法认定多长的, 哪些空间点属于相近的动力系统. 本文以邯钢钢铁公司 7 号高炉在线采集的 2000 炉铁水含硅量 $[Si]$ 数据为样本(如图 1), 首先利用基于逆序数的检验方法对 $[Si]$ 序列做平稳性分析. 然后基于关联积分的概念, 定义了衡量不同时间序列间

动力学相似性的“距离”, 将其应用于人工构造的含噪声的分段时间序列, 验证了该“距离”辨识动力学突变的有效性; 通过计算 $[Si]$ 数据等分子序列间的“距离”, 给出了高炉动力学演变分析结果. 分析结果表明高炉冶炼过程中, 其冶炼规律在某些时间段存在突变性. 那么在这种动力学突变性的影响下, 高炉铁水含硅量 $[Si]$ 是可以预测的吗? 为此, 本文应用基于相空间重构的 DVV 算法, 通过分析 $[Si]$ 序列, 并与理想的自回归白噪声序列、确定性序列、混沌序列相比较, 进一步检验高炉冶炼动力学演变的确定性.

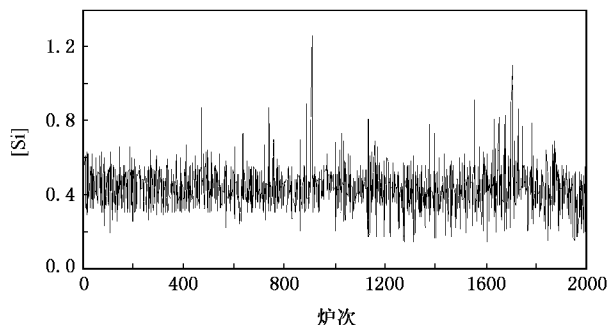


图 1 邯钢 7 号高炉 $[Si]$ 原始时间序列

2. $[Si]$ 时间序列的平稳性非参数检验

时间序列分析的基本用途是根据过去预报未来, 因此, 通常要求平稳性假定. 称一个随机过程

* 国家科技成果重点推广项目(批准号: 2005EC000166)和浙江省自然科学基金(批准号: Y107110)资助的课题.

† 通讯联系人, E-mail: gaouchou@zju.edu.cn

$X(t)$ 为平稳过程(严格地说是弱平稳过程),若对所有的 t, h, k 有 $E(Y_t) = E(Y_{t+k}); \text{var}(Y_t) = \text{var}(Y_{t+k}); \text{cov}(Y_t, Y_{t+k}) = \text{cov}(Y_{t+k}, Y_{t+h+k})$. 检验平稳性运用较为广泛的方法是单位根检验^[5],但这类检验方法要求对原序列做线性回归,而实际复杂系统产生的时间序列很难保证线性性质,因此本文采用基于逆序数的平稳检验方法对 [Si] 时间序列进行检验^[5].

逆序平稳性检验方法的基本原理是:如果数据平稳,则分段子序列的均值与方差应无显著差异,其具体步骤是:

1) 将原时间序列 $X(t) (1 \leq t \leq N)$ 等分成 k 个长度为 n 的子序列,记为

$$\{x_{i1}, x_{i2}, \dots, x_{in}\}, i = 1, \dots, k;$$

2) 计算各子序列的均值 \bar{x}_i 与方差 σ_i^2 ;

3) 统计均值与方差序列的逆序数

$$A = \sum_{i < j} a_{ij}, B = \sum_{i < j} b_{ij},$$

其中

$$a_{ij} = \begin{cases} 1, & \text{当 } i < j \text{ 时 } \bar{x}_i > \bar{x}_j, \\ 0, & \text{其他,} \end{cases}$$

$$b_{ij} = \begin{cases} 1, & \text{当 } i < j \text{ 时 } \sigma_i^2 > \sigma_j^2, \\ 0, & \text{其他;} \end{cases}$$

4) 计算统计量

$$u = \frac{A + \frac{1}{2} - \frac{k(k-1)}{4}}{\sqrt{k(2k^2 + 3k - 5)/72}},$$

$$v = \frac{B + \frac{1}{2} - \frac{k(k-1)}{4}}{\sqrt{k(2k^2 + 3k - 5)/72}},$$

假若原序列平稳,则当 $k > 10$ 时, u, v 均渐进服从标准正态分布.在 0.05 显著性水平下,当 $|u|$ 或 $|v|$ 大于 1.96 时,拒绝序列均值或方差平稳的假设.

表 1 给出了 [Si] 样本序列的平稳性检验结果,从表中可以看出对应于不同的分段数,高炉铁水 [Si] 序列均拒绝方差平稳的假设,对于均值则没有明显的证据拒绝其平稳性,这主要是由于高炉冶炼过程中,为了保持炼铁过程平稳顺行,高炉操作工长力图将铁水含硅量钉在一个理想的目标值附近,因而硅序列呈现均值回复的平稳性特征;另一方面高炉冶炼过程包括配料、上料、布料、鼓风、富氧、喷煤等众多子工序,对高炉铁水含硅量的影响因素错综复杂,高炉工长虽然可以大体上控制炉内温度向凉或向热的趋势,却很难把握控制的手段和强度,更主要的是受到内部惯性和外部控制的持续影响,其内

在的演化方程可能随之发生突变,造成实际生产中,先前行之有效的控制方法在某些时刻失去作用甚至产生相反的效果,本文认为正是这种系统内部动力学方程演变的复杂性造成了硅序列的非平稳性,使得诸多预报和控制模型纷纷失效.下面就从动力学突变性的角度对 [Si] 序列非平稳性成因做进一步地分析.

表 1 [Si] 时间序列的平稳性检验结果

分段数 k	分段子序列长度	统计量 u	统计量 v
25	80	1.5648	-2.2654
20	100	1.7493	-2.8227
16	125	1.3856	-2.5733

3. [Si] 序列非平稳性成因的探测

对于时间序列 $X(t) (1 \leq t \leq n)$ 做相空间重构,重构向量记为 $\mathbf{x}_k^m = [x(k), \dots, x(k + (m-1)\tau)]$,其中 τ 为延迟时间, m 为嵌入空间维数,则 $X(t) (1 \leq t \leq n)$ 关联积分 $C_m(x, x)$ 的计算公式为

$$C_m(x, x) = P(\|\mathbf{x}_i^m - \mathbf{x}_j^m\| < r) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \Theta(r - \|\mathbf{x}_i^m - \mathbf{x}_j^m\|),$$

其中 $\|\cdot\|$ 取 ∞ -范数, $\Theta(\cdot)$ 为 Heaviside 函数.

关联积分是个累积分布函数,表示相空间中任意两点之间距离小于 r 的概率.对于两个等长的时间序列 $X(t), Y(t)$,可以定义所谓的交叉关联积分

$$C_m(x, y) = P(\|\mathbf{x}_i^m - \mathbf{y}_j^m\| < r),$$

关联积分具有一定区分潜在动力学结构的能力,但它对嵌入维数的选择依赖较大,不利于分析较短的时间序列,因此,需要引入更稳健的参数,物理量 S_m

$$= \frac{C_{m+1}^{[6]}}{C_m^{[6]}}$$

反映了前 $m\tau$ 项的序列信息对于序列演化的影响能力,比率 $R_m = \frac{S_{m+1}^{[6]}}{S_m^{[6]}}$ 反映的是序列的第 $(m+1)\tau$ 项延迟量对序列的贡献程度.将 [Si] 序列分割成长度为 80 的 25 个等长子序列,当 $\tau = 1, r = 0.5\sigma_{[Si]}$ ($\sigma_{[Si]}$ 为 [Si] 序列的标准差)时,分别计算每段的 C_2, S_1, R_1 ,而后取平均值,标准差和相对方差,可得表 2.

表 2 [Si] 序列各物理特征量的统计性质比较

	C_2	S_1	R_1
均值	0.01291	0.232	4.184
标准差	0.00082	0.012	0.053
相对方差	0.064	0.051	0.013

从表 2 可以看出 R_m 具有较小的相对方差,因而可以更可靠地识别两个时间序列(子序列)的动力学差异.基于 R_m ,定义度量时间序列间动力学“距离”^[6]的物理量 $d_m(x, y) = \max(d_m(x \leftarrow y), d_m(y \leftarrow x))$,其中 $d_m(x \leftarrow y) = \sqrt{|R_m(x, x) - R_m(x, y)|}$, $d_m(y \leftarrow x) = \sqrt{|R_m(y, y) - R_m(y, x)|}$,这样定义的 $d_m(x, y)$ 满足对称性的要求.如果 $X(t), Y(t)$ 具有相近的内部机理,则 $d_m(x, y)$ 具有相应较小的数值.因此可以通过计算时间序列的分段子序列间的 d_m 来研究时间序列动力学性质的突变性.

对

$$x(t) = \begin{cases} 2\sin(0.5t) + 1.5\cos(0.2t) + 0.1 + \sigma\varepsilon_1, & t \leq 1000, \\ \tan(\pi t) + 2\sin(0.2t) + 2 + \sigma\varepsilon_2, & 1000 < t \leq 2000 \end{cases}$$

产生的理想时间序列根据 d_m 做动力学演变分析,验证其有效性.当 $t \leq 1000$ 时,子序列由正弦和余弦函数叠加得到, ε_1 为白噪声服从标准正态分布, σ 为无噪声时序列的标准差;当 $1000 < t \leq 2000$ 时,子序列由正切和正弦函数叠加而成, ε_2 为白噪声, $t = 1000$ 为动力学结构突变点.取宽度为 n 的滑动窗口,分别计算 $x(t)$ 中 n 至 $2000 - n$ 各点左右两子序列的距离 d_m (其中 $\tau = 1, r = 0.5\sigma_x, n = 100, m = 2$),结果如图 2 所示.

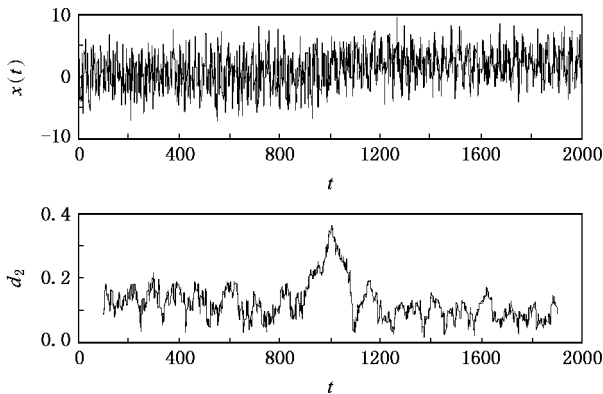


图 2 理想时间序列动力学突变性检验结果

从图 2 可以看出 d_m 曲线在 900—1100 存在明显的大值区,且在 $t = 1004$ 时, d_m 取到最大值 0.361.构造统计量 $\tilde{e}(i) = e(i)/\tilde{\sigma}$,其中 $e(i) = d_m(i) - \tilde{d}_m(i)$, $\tilde{d}_m(i)$ 表示除去 $d_m(i)$ 后 d_m 序列的平均值, $\tilde{\sigma}$ 为序列 $e(i)$ 的标准差;可以证明 $\tilde{e}(i)$ 渐进服从标准正态分布, $P(\tilde{e}(i) > 1.96) = 0.05$.如果 $|\tilde{e}(i)| > 1.96$,我们就认为在第 i 点原序列确实发

生了动力学性质的改变. $\tilde{e}(1004) = 4.183$,检验结果表明该点为动力学突变点,这与实际突变点非常接近,检验偏差 $\eta = \frac{|t - t_0|}{N} \times 100\% = 0.2\%$ ($t_0 = 1000, N = 2000$).

运用上述方法对实际的高炉含硅量 [Si] 时间序列做动力学演变分析,根据平稳性检验结果选取窗口宽度为 80,分别计算各点左右子序列之间的距离 d_m (其中 $\tau = 1, r = 0.5\sigma_{[Si]}, m = 2$),所得结果如图 3 所示.

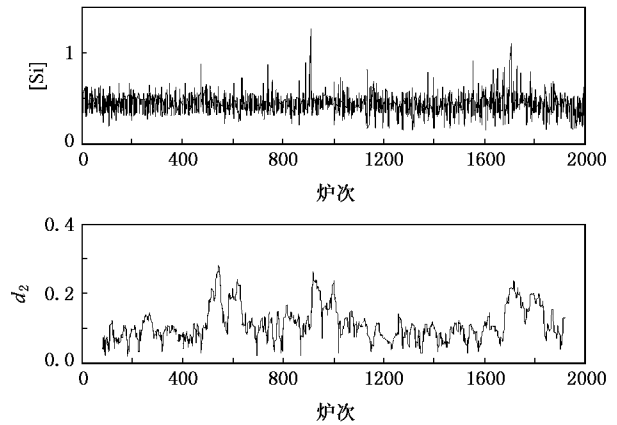


图 3 [Si] 时间序列动力学突变性检验结果

可以看到 d_m 曲线存在三个明显的大值区域,对这些区域的极值点做检验, d_m 分别在 $t = 543, 921, 1720$ 时,取到局部最大值 0.278, 0.261, 0.238, 对应的检验统计量 $\tilde{e}(543) = 3.507, \tilde{e}(921) = 3.1, \tilde{e}(1720) = 2.617$,因此可以认为在这些冶炼炉次附近,高炉冶炼机理发生了较大的改变.除去这三个区域, d_m 在 $t = 816$ 时,取最大值 0.166,检验统计量 $\tilde{e}(816) = 1.132$,动力学突变不具备显著性.与 [Si] 时间序列对照,容易发现动力学突变点 $t = 921, 1720$ 之前 [Si] 出现了较大的波动,炉长根据观测到的 [Si],在操作制度上进行调整(如改变喷煤量,调节料速等),这些外部因素可能是造成冶炼机理发生变化的主要原因;而 $t = 543$ 附近 [Si] 波动相对比较平缓,突变点的形成更可能是由于冶炼过程热量的积累,炉体内矿石分布结构的变化等内部原因引起的.值得注意的是 [Si] 序列的 d_m 曲线大值区域跨度比较长,且存在反复,这可能是由于高炉冶炼的大滞后性^[11],各种因素的影响效果不容易立竿见影,炉长的操作往往需要多次校正所造成的.

4. [Si] 序列的可预测性分析

高炉冶炼过程的内部动力学结构变动如此频繁,那么高炉冶炼过程是可以预测然后加以控制的吗?由于 [Si] 序列是反映高炉内部热状况的关键指标,接下来本文就运用 DVV 算法分析 [Si] 序列蕴含的确定性因素,定性判断 [Si] 序列的可预测性。

DVV 算法^[7]是基于相空间重构的一种比较新颖的算法,通过计算所谓的目标方差较好地判断了时间序列确定性的强弱,具体步骤如下:

1) 对于时间序列 { x(t) | t = 1, ..., N } 运用 C-C 算法^[8]确定最优延迟时间和嵌入窗,根据所确定的参数做相空间重构,延迟向量记为

$$\bar{x}_k = [x(k), \dots, x(k + (m - 1)\tau)],$$
$$k = 1, \dots, N - (m - 1)\tau ;$$

记录延迟向量 x_k 扩充的目标值 $x(k + m\tau)$ 。

2) 计算延迟向量两两间的距离

$$\text{dis}(x_{k'} - x_{k''}) = \max_{i=0}^{m-1} | x(k' + i\tau) - x(k'' + i\tau) |,$$

然后计算这些距离的均值 μ 和标准差 σ 。

3) 在区间 $[\mu - n_d\sigma, \mu + n_d\sigma]$ 上等间隔选取 r , 对于给定的 r , 给出集合

$$\Omega_k(m, r) = \{ x_i \mid \text{dis}(x_i - x_k) \leq r \},$$
$$k = 1, \dots, N - (m - 1)\tau ;$$

参数 n_d 用来调节距离 r 的选择区域。

4) 计算集合 $\Omega_k(m, r)$ 中所有延迟向量扩充目标值 $x_{k+m\tau}$ 的方差 σ_k^2 , 目标方差 $\sigma^{*2}(m, r) = \text{avg}(\sigma_k^2) \sigma^2$, σ^2 为原序列 { x(t) | t = 1, ..., N } 的方差, $\text{avg}(\sigma_k^2)$ 为所有诸如 $\Omega_k(m, r)$ 的集合相应方差的平均值。

DVV 算法的物理意义为:如果数据的确定性较强,那么临近相似的延迟向量演化的目标值也应该相似,因此,计算所得的目标方差应较小,随距离 r 的增大目标方差的增大(图像斜率)不应过大。图 4 给出了 DVV 算法对 [Si] 序列的仿真结果,为了更好地理解 [Si] 序列计算结果,文中给出了三种典型的时间序列,用 DVV 算法做运算,以便比较。

1) 自回归模型

$$x_k = 1.79x_{k-1} - 1.85x_{k-2} + 1.27x_{k-3} - 0.41x_{k-4} + \epsilon_k,$$

其中 { ϵ_k } 是独立同分布的白噪声过程,取标准正态分布生成的四个随机数作为初始值,迭代模型产生

2000 个数据点,算法计算结果如图 5 所示。

2) 确定性系统

$$\frac{dx}{dt} = y, \frac{dy}{dt} = -0.25y - x - 0.3\cos(t),$$

取 (10, 17) 为初始点,应用变步长四阶五级 Runge-Kutta-Felberg 算法,产生分量 x 的 2000 个数据点,数据分析结果如图 6。

3) Lorenz 系统

$$\frac{dx}{dt} = 10y - 10x, \frac{dy}{dt} = 28x - y - xz, \frac{dz}{dt} = xy - \frac{8}{3}z,$$

取 (15.34, 13.68, 37.91) 为初始点,应用 Runge-Kutta-Felberg 算法求系统数值解,产生 2000 个分量 x 的数据点,对序列做确定性分析,结果如图 7。

为了便于比较各种情况计算得到的目标方差,各图中距离 r 一致规范在 [0, 1] 区间内。

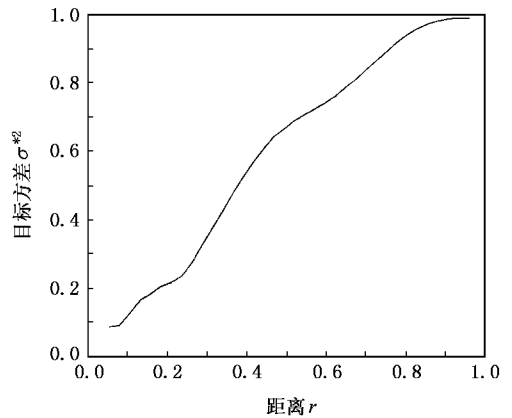


图 4 [Si] 序列的 DVV 分析结果

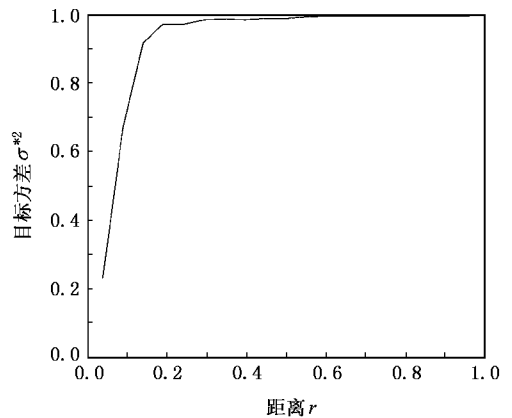


图 5 自回归模型的 DVV 分析结果

自回归模型

$$x_k = 1.79x_{k-1} - 1.85x_{k-2} + 1.27x_{k-3} - 0.41x_{k-4} + \epsilon_k$$

具有白噪声的强干扰,由于这种随机性的影响,要对

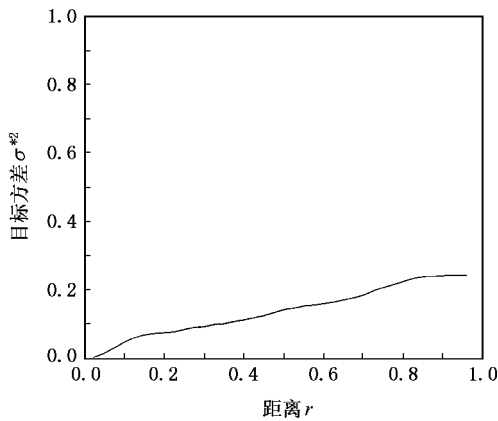


图 6 确定性系统的 DVV 分析结果

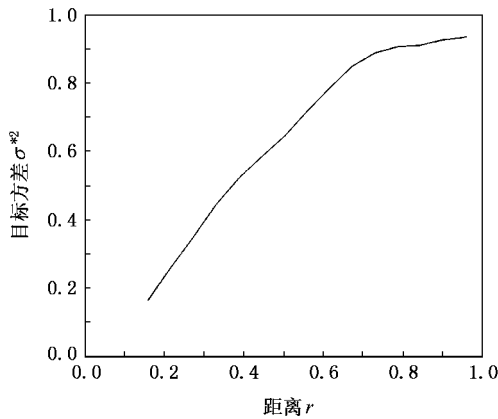


图 7 Lorenz 系统的 DVV 分析结果

数据做准确的预报几乎是不可能的,因此目标方差 $\sigma^{*2}(m, r)$ 数值较大,且随距离 r 增加而迅速增大;对于系统 $\frac{dx}{dt} = y, \frac{dy}{dt} = -0.25y - x - 0.3\cos(t)$,可算出分量 x 的解析通解为 $x = C_1 \exp(-t/8) \sin(3\sqrt{7}t/8) + C_2 \exp(-t/8) \cos(3\sqrt{7}t/8) - 6\sin t/5$,只要知道系统的初始值,就可以准确地预知系统的发展情况,因此,计算得到的目标方差很小,随距离增长也很缓慢.混沌系统虽然具有随机不确定性,但这是非线性

确定系统内在的属性,具有丰富的内部层次结构,混沌信号可以通过重构体现在嵌入空间的一个低维流形上,其重构的轨迹可以短期预测,因而图 7 体现出介于确定性和随机性之间的图像特征.

与上述三种典型的时间序列图像比较 [Si] 序列图像(图 4)中目标方差 $\sigma^{*2}(m, r)$ 相对于自回归模型数值较小,随 r 增加的幅度也比较平缓;从表 3 给出的各个模型相空间重构和目标方差结果也可以看出 [Si] 序列的最小目标方差介于确定性系统和 Lorenz 系统之间,因而可以认为高炉系统动力学突变并不是盲目地、随机地发生的,而是蕴含着较强确定性的行为,这使得对 [Si] 序列做满足工业要求的短期预报成为可能.

表 3 C-C 算法和 DVV 算法的数值结果

模型	随机自回归模型	确定性系统	Lorenz 系统	邯钢铁序列
最优延迟时间	4	14	11	6
嵌入窗	11	157	72	32
最小目标方差	0.2292	0.0027	0.1654	0.086

5. 结论及意义

利用逆序平稳性检验方法分析了邯钢 7 号高炉在线采集的 2000 炉 [Si] 时间序列,分析结果表明 [Si] 序列的方差不具备平稳性,然后,从高炉冶炼过程内部动力学突变性的角度解释 [Si] 序列的不平稳性,通过关联积分定义了度量时间序列间动力学“距离”的物理量 d_m ,从而有效地检测了高炉冶炼规律发生突变的具体炉次;最后运用 DVV 算法对 [Si] 序列做定性的确定性分析;虽然高炉冶炼过程的内部动力学结构变动频繁造成诸多预报和控制模型应用效果不理想^[9],但是这种演化过程是由众多有序行为(喷煤、调节风量、高炉热惯性等)组合而成的, DVV 检验结果体现出较强的确定性,对 [Si] 序列做满足工业要求的短期预测控制^[10]是可行的.

[1] Liu X G, Liu F 2003 *Optimization and Intelligent Control System of BF Ironmaking Process* (Beijing: Metallurgical Industry Press) p31 (in Chinese) [刘祥官、刘芳 2003 高炉炼铁过程优化与智能控制系统(北京:冶金工业出版社)第 31 页]

[2] Biswas A K 1984 *Principles of Blast Furnace Ironmaking* (SBA Publication, Calcutta)

[3] Yang P C, Zhou X J 2005 *Acta Meteorologica Sinica* 63 5 (in Chinese) [杨培才、周秀骥 2005 气象学报 63 5]

- [4] Gribkov D , Gribkova V 2000 *Phys. Rev. E* **61** 6
1(in Chinese) 陆振波、蔡志明、姜可宇 2007 系统仿真学报 **19** 11]
- [5] Harris R , Sollis R 2003 *Applied time series modeling and forecasting* (John Wiley & Sons Ltd).
- [6] Manuca R , Savit R 1996 *Physica D* **99** 134
- [7] Gautama T , Mandic D P , Van Hulle M M 2004 *IEEE Trans. Biomed. Eng.* **51** 728
- [8] Lu Z B , Cai Z M , Jiang K Y 2007 *Journal of System Simulation* **19**
- [9] Luo S H , Liu X G 2006 *Acta Phys. Sin.* **55** 3343(in Chinese) [罗世华 刘祥官 2006 物理学报 **55** 3343]
- [10] Gao C H , Zhou Z M 2004 *Acta Phys. Sin.* **53** 4092(in Chinese) [郜传厚、周志敏 2004 物理学报 **53** 4092]

The analysis of dynamic structure of silicon content in blast furnace hot metal^{*}

Zhao Min Liu Xiang-Guan Gao Chuan-Hou[†]

(*The Institute of System Optimum Technique , Zhejiang University , Hangzhou 310027 , China*)

(Received 4 June 2007 ; revised manuscript received 27 September 2007)

Abstract

With Si time series of No.7 blast furnace(BF) at Han Iron and Steel Group Co. as sample space , the stationarity of Si time series is tested based on the number of reversed order. A statistical measure of dynamic closeness between two orbits is defined using correlation integral. Divide the Si series into equal-length windows , according to the calculation of the characterization between these windows , the dynamic abrupt changes in BF ironmaking process are detected. Finally the novel DVV(delay vector variance) method is used to examine the predictability of Si time series during the dynamic evolution.

Keywords : BF ironmaking , stationarity , dynamical closeness , predictability

PACC : 0545

^{*} Project supported by the National Ministry of Science and Technology (Grant No. 2005EC000166) and the Natural Science Foundation of Zhejiang Province , China (Grant No. Y107110).

[†] Corresponding author. E-mail : gaochou@zju.edu.cn