

改进的数据驱动子空间算法求解钢铁企业能源预测问题

张颜颜^{1,3}, 唐立新^{1,2}

(1. 东北大学 物流优化与控制研究所, 辽宁 沈阳 110819;

2. 辽宁省制造系统与物流优化重点实验室, 辽宁 沈阳 110819; 3. 流程工业综合自动化国家重点实验室, 辽宁 沈阳 110819)

摘要: 本文以钢铁企业生产与能源系统作为研究背景, 提出一种数据驱动的子空间方法(data-driven subspace, DDS)预测各生产工序的能源消耗. 为了准确表达该问题, 本文全面分析了能源消耗和回收的特点, 以提取建模因素; 为了设计有效的求解方法, 对实际生产和数据的特征进行了分析. 为了提高预测准确率, 文中引入了反馈因子和遗忘因子来改进子空间方法, 因子的取值采用粒子群算法(particle swarm optimization, PSO)来优化. 对实际生产数据的测试验证了本文所提出的方法的有效性, 该结果能够为钢铁企业的能源预测和管理提供有效的决策支持.

关键词: 数据驱动子空间; 粒子群优化; 能源预测

中图分类号: TP29 **文献标识码:** A

Improved data-driven subspace algorithm for energy prediction in iron and steel industry

ZHANG Yan-yan^{1,3}, TANG Li-xin^{1,2}

(1. The Logistics Institute, Northeastern University, Shenyang Liaoning 110819, China;

2. Liaoning Key Laboratory of Manufacturing System and Logistics, Shenyang Liaoning 110819, China;

3. State Key Laboratory of Synthetical Automation for Process Industries, Shenyang Liaoning 110819, China)

Abstract: Using the production and energy system in Iron and Steel industry as the research background, we propose a data-driven subspace (DDS) method for predicting the energy consumption of production operations. The characteristics of energy consumption and regeneration are fully investigated to find the crucial factors for building the model. The features of practical conditions and data are analyzed in designing the efficient solving method. The subspace method is improved by introducing the feedback factor and the forgetting factor, values of which are optimized by particle swarm optimization (PSO) algorithm in order to improve the prediction accuracy. The performance of the improved method is demonstrated by experimental tests using the practical data, which provide beneficial results in energy prediction and management.

Key words: data-driven subspace; particle swarm optimization; energy prediction

1 引言(Introduction)

能源是现代工业的基础. 在经济持续强劲增长和城市化进程的带动下, 一次能源和二次能源都面临巨大的需求增长. 作为典型的高能耗工业之一, 国内钢铁企业的能源消耗量占全国能耗总量的10%以上; 而且国内钢铁行业的综合能耗远高于世界先进水平, 能源利用率仅达到30%. 钢铁企业能源消耗的特点是能源介质种类多(包括煤、焦炭、煤气、电力、蒸汽等20几种能源介质)、消耗量大. 随着能源价格的上涨, 能源成本已在总生产成本中占有很大比例. 在研究钢铁能源问题时, 由于要考虑能源消耗过程中伴随生成的二次能源, 并且与生产紧密相关, 整个系统变得复杂, 增加了问题的难度. 通过实施先进

的技术手段, 对于能源需求的趋势提供准确的估计, 合理控制能源投入, 是实现节能降耗并进一步降低生产成本的重要手段之一, 也是钢铁企业增强市场竞争力、提高企业的经济效益和信息管理水平的重要途径.

2 问题描述及文献综述(Problem description and literature review)

2.1 问题描述(Problem description)

能源消耗过程机理复杂, 且没有确定的模型, 因此能耗预测包括两个步骤: 首先根据能源消耗样本数据对问题进行建模, 然后设计有效的求解方法, 对未来计划期内的能源消耗和回收情况进行预测. 为

了建立能充分表达问题特征的数学模型,要对建模对象的能源消耗规律进行全面分析以提炼出正确的输入输出变量。比如,要预测加热炉的能源消耗,必须要考虑加热方式、冷坯和热坯的比例等显著影响能源燃料消耗的因素。本文的研究对象是炼铁工序高炉煤气(blast furnace gas, BFG)发生量预测以及热轧产线的能源消耗预测。

2.2 文献综述(Literature review)

由于能源系统具有多介质、多工序、能源消耗复杂和非线性的特点,仅靠历史数据无法建立准确的机理模型,因此高精度的能耗预测一直是能源管理的难点。文献[1]综述了目前用于能源预测的主要方法,分析了各自的优缺点,归纳了建模过程中需要考虑的问题。Liu等^[2]采用时间序列法对焦炉煤气消耗量和发生量进行预测,并通过将高斯过程用于回波网络来改进算法。文献[3]提出一种改进的回声状态网络方法,采用时间序列方法根据某钢铁企业高炉煤气发生量的历史数据预测其未来的发生量。文献[4]采用MATLAB方法研究了某钢铁企业的能源预测问题,但未有关于该问题的研究背景的详细描述。文献[5]设计了遗传算法(genetic algorithm, GA)和后向传播(back propagation, BP)网络的混合方法,针对某炼铁厂的氧气消耗量预测验证了混合算法的有效性。

目前,关于钢铁企业能源预测的研究相对较少,并且已有的研究成果多关注某一种能源介质,缺乏系统性。实际上,钢铁生产与能源是一个涉及多种设备、多种能源介质的复杂系统。从整个工序或全厂角度对能源消耗进行预测并且对预测结果进行分析和评价的系统性研究,对于企业实现能源介质优化配置具有重要的实际意义。基于上述考虑,本文首先从实际生产中抽象出能源预测问题,然后根据问题特点设计了有效的求解算法。

常见的预测方法有人工神经网络、线性回归等。人工神经网络的优点是具有良好的非线性映射能力、并行计算能力以及自适应学习能力,但是某些人工神经网络算法,如基于梯度下降的BP算法收敛速度慢,易陷入局部极小点,且对网络初始权值、阈值和自身的学习率等参数极为敏感。线性回归方法采用数理统计中回归分析方法来确定两种或两种以上变量间相互依赖的定量关系,不具备对数据的学习能力。本文通过分析能源消耗及历史数据的特点,设计了数据驱动的子空间方法。

子空间方法是20世纪90年代早期提出的一种新型的系統识别方法,适于求解先验知识较少的多变量系统辨识问题。由于其在数值计算上的优势,已受到了预测控制和辨识领域的广泛关注^[6]。

随着自动化技术的发展和钢铁企业管理水平的提高,实际生产过程中能耗数据的产生量和存储量

都随之增加,这类数据中包括大量的生产信息,比如能源的消耗规律、设备或生产过程的变化规律和趋势等。这些因素促成了子空间等数据驱动类算法的出现,这些算法的核心思想是充分利用历史数据获得其蕴含的规律。子空间方法基于状态空间理论,不必建立系统模型,而是通过计算识别出状态空间模型;即,该方法直接由历史数据(输入&输出)及未来的输入数据推导出空间矩阵来完成对未来输出变量的预测,因此,该方法克服了传统方法对模型的依赖性。

为了改进子空间算法的性能,Goethals等^[7]提出了一种对多输入多输出的H系统(Hammerstein systems)的辨识方法,该方法采用最小二乘支持向量机(least squares support vector machines, LS-SVM)回归来代替辨识过程中使用常规的最小二乘求中间矩阵的步骤,结果表明即使有输出噪声存在,该方法也能从输入输出数据中辨识出线性系统或非线性系统。Favoreel等^[8]将子空间算法用于辨识受白色输入限制的多输入多输出双线性系统。结果表明,若系统的输入是白色的且相互间独立,可以通过对数据矩阵的行子空间的正交投影来获得该未知双线性系统的状态空间估计。Van Overschee等^[9]研究了多输入多输出线性离散时间系统的辨识问题,给出了通过对给定的数据进行逆傅里叶变换得到状态空间系统矩阵的方法。Huang and Kadali^[10]详细阐述了数据驱动的思想,即利用子空间方法直接从输入输出数据设计预测器的方法。Huang等^[11]使用数据驱动的子空间方法建立了一个数据驱动预测控制规则,并应用于燃料电池系统。所提出的方法能够处理输出变量在线测量数据不完备的情况。上述改进算法基本思路是由输入输出Hankel矩阵投影的行子空间和列子空间来获取模型参数。

本文通过引入反馈因子和遗忘因子对标准DDS方法进行改进,因子的取值采用PSO算法来寻优,该方法是1995年由Kenned and Eberhart^[12-13]提出的基于群体智能的方法,采用微粒种群对探索最优空间问题建模,粒子的轨迹同时影响其自身和其他粒子的搜索模式。搜索朝着有希望的方向进行,每个粒子的速度方向由其自身的最好位置和整个种群的最好位置综合决定。粒子群算法已成功求解许多优化问题,但将该算法与子空间相结合的研究还不多见。

3 DDS的改进策略(The improving strategy of DDS)

3.1 参数及符号说明(Notations)

文中采用的参数及变量说明如下:

A, B, C, D : 状态空间系统矩阵;

$x_t \in \mathbb{R}^n$: 采样时刻 t 的系统状态;

$u_t \in \mathbb{R}^l$: 输入变量在时刻 t 的测量值;

$y_t \in \mathbb{R}^m$: 输出变量在时刻 t 的测量值;

$e_t \in \mathbb{R}^m$: 白噪声序列;

K : Kalman滤波增益矩阵;

U_p : u_t 的历史数据的Hankel矩阵;

U_f : u_t 的未来数据的Hankel矩阵;

Γ_N : 扩展可观测矩阵;

H_N^d : 下三角确定性Toeplitz矩阵;

H_N^s : 下三角随机Toeplitz矩阵;

Δ_N^d : A, B 的逆扩展可控矩阵;

Δ_N^s : A, K 的逆扩展可控矩阵;

L_w : 历史输入输出数据的子空间矩阵;

L_u : 包含过程Markov参数的子空间矩阵;

H : 反馈因子;

F : 反馈矩阵;

E_p : e_t 的历史数据的Hankel矩阵;

W_p : $\begin{pmatrix} Y_p \\ U_p \end{pmatrix}$.

3.2 DDS算法简介(DDS algorithm introduction)

子空间的理论基础之一是状态空间理论, 通过使用脉冲响应系数组成的Hankel矩阵求出系统的状态空间模型^[14]. 由于可靠的脉冲响应较难获取, 人们考虑直接用输入输出数据求取状态空间模型. Overschee and De Moor^[15]首先提出了用于求解确定性系统和随机系统的子空间辨识算法, 基本思路是首先构造输入输出数据的Hankel矩阵, 然后对Hankel矩阵进行投影运算, 最后利用线性代数工具, 如矩阵奇异值分解(singular value decomposition, SVD), 正交三角分解(QR)来求解状态空间模型.

状态空间模型的标准形如下式:

$$x_{t+1} = Ax_t + Bu_t + Ke_t, \quad (1)$$

$$y_t = Cx_t + Du_t + e_t. \quad (2)$$

由式(1)–(2)推导出子空间方程如式(3)–(5)所示:

$$Y_f = \Gamma_N X_f + H_N^d U_f + H_N^s E_f, \quad (3)$$

$$Y_p = \Gamma_N X_p + H_N^d U_p + H_N^s E_p, \quad (4)$$

$$X_f = A_N X_p + \Delta_N^d U_p + \Delta_N^s E_p. \quad (5)$$

这里: p 表示过去(past), f 表示未来(future), 各自对应的Hankel矩阵定义如下:

$$U_p \triangleq U_{0|N-1} = \begin{bmatrix} u_0 u_1 & \cdots & u_{j-1} \\ u_1 u_2 & \cdots & u_j \\ \vdots & & \vdots \\ u_{N-1} u_N & \cdots & u_{N+j-2} \end{bmatrix}, \quad (6)$$

$$U_f \triangleq U_{N|2N-1} = \begin{bmatrix} u_N u_{N+1} & \cdots & u_{N+j-1} \\ u_{N+1} u_{N+2} & \cdots & u_{N+j} \\ \vdots & & \vdots \\ u_{2N-1} u_{2N} & \cdots & u_{2N+j-2} \end{bmatrix}. \quad (7)$$

下标 $t_1|t_2$ 表示子空间矩阵的第1列从时间 t_1 开始到时间 t_2 结束. “future”类Hankel矩阵的第1列与“past”类Hankel矩阵的第1列在时间上具有连续性, 两类矩阵相应的列也具有此规律.

基于子空间算法的预测方法可描述为: 以过去的输入输出 W_p 和未来的输入 U_f 作为输入, 通过计算预测出最优的未来输出 Y_f 的值.

$\hat{Y}_f$ 是最小二乘预测值, 表达为

$$\min_{L_w L_u} \|Y_f - (L_w L_u) \begin{bmatrix} W_p \\ U_f \end{bmatrix}\|_N^2. \quad (8)$$

式(8)的解, 是从列空间到和的列空间的正交映射, 由式(9)给出:

$$Y_f = \hat{Y}_f / \left\langle \begin{bmatrix} W_p \\ U_f \end{bmatrix} \right\rangle = \underbrace{Y_f / U_f W_p}_{L_w W_p} + \underbrace{Y_f / W_p U_f}_{L_u U_f}. \quad (9)$$

子空间方程 $\hat{Y}_f = L_w W_p + L_u W_f$ 用来预测. 可见, 采用子空间算法预测, 是通过挖掘输入输出数据中包含的信息和知识来认知系统, 消除了对机理模型的依赖, 同时避免了建模过程中可能引入的各种误差及不确定性.

本文研究的两个问题为: 高炉煤气发生量预测问题和热轧能源消耗预测问题. 高炉煤气是炼铁过程中最重要的副产品, 也是钢铁企业炼焦、加热炉和电厂的主要燃料之一. 在该问题中, 高炉煤气发生量设计为输出变量 Y , 其大小受到送风流量和热风炉流量的影响, 后两者设计为输入变量 U . 在热轧工序中, 影响能源消耗量的因素包括产量和气温, 因此将这两个因素设计为二维输入变量 U , 输出变量 Y 是未来的能源消耗量, 本文重点研究热轧产线的电力消耗预测, 它是热轧产线最重要的能源介质.

上述两个问题用来测试数据驱动的子空间算法在时间维预测问题方面的性能. 在测试过程中, 首先将数据分为训练数据和测试数据, 前者作为样本数据来训练算法得到Hankel矩阵, 后者用来检验算法的预测精度.

3.3 DDS的改进策略(The improving strategy of DDS)

标准子空间方法遵循其基本步骤求解问题时, 面对不同的问题表现出不同的性能. 通过实验测试, 笔者发现标准的子空间方法对本文研究的问题在性能

上有很大的改进空间(如表3和5所示). 通过分析问题的特征和子空间算法的求解步骤, 本文提出了改进策略, 考虑在预测模型中增加两类因子.

1) 对子空间方程增加反馈因子.

由输入输出数据导出的子空间矩阵包含着输入输出数据间的关系, 这一特点使算法对数据的选取较为敏感; 即, 当未来的输入 U_f 与过去的输入 U_p 在数值上相差较大时, 预测值与实际值间的偏差比较大. 为了克服这一缺点, 考虑在预测值的基础上, 加入一个反馈调节矩阵.

反馈矩阵的计算式为 $F = H * E_p$, 则基于数据的子空间预测模型表达为

$$\hat{Y}_f = L_w W_p + L_u U_f + F. \quad (10)$$

本文通过使用粒子群优化算法对反馈调节因子 H 进行寻优, 力图找到使预测值与实际值间的差值 F 达到最小的反馈调节因子 H . 算法在具体实现时将输入输出数据详细分为 U_{train} 和 Y_{train} 用于训练、 U_{test} 和 Y_{test} 用于测试.

2) 对 $U_p, Y_p, U_f, U_{\text{train}}$ 加入遗忘因子.

当子空间算法用于时间维度上的预测时, 历史数据中越接近当前时间的数据能更好地反映系统当前的状态, 所以在预测时尽量缩小远离当前时刻的历史数据的权重^[16]. 故考虑加入遗忘因子, 使远离当前时刻的历史数据的权重小于近期数据的权重, 遗忘因子采取 $\alpha(0 < \alpha < 1)$ 的指数形式.

$$U_p \triangleq U_{0|N-1|N+j-2} = \begin{bmatrix} \alpha^{N+j-2}u_0 & \alpha^{N+j-3}u_1 & \cdots & \alpha^{N-1}u_{j-1} \\ \alpha^{N+j-3}u_1 & \alpha^{N+j-4}u_2 & \cdots & \alpha^{N-2}u_j \\ \vdots & \vdots & & \vdots \\ \alpha^{j-1}u_{N-1} & \alpha^{j-2}u_N & \cdots & \alpha u_{N+j-2} \end{bmatrix}. \quad (11)$$

遗忘因子 α 的指数是通过 $N+j-2$ 减去对应元素的时间下标得到. 由于 U_p 和 U_f 的Hankel矩阵对应列的时间下标具有连续性, 将两者组合, 得到的带有遗忘因子的 U_{pf} 的新矩阵表示为

$$U_{\text{pf}} \triangleq U_{0|2N-1|2N+j-2} = \begin{bmatrix} \alpha^{2N+j-2}u_0 & \alpha^{2N+j-3}u_1 & \cdots & \alpha^{2N-1}u_{j-1} \\ \alpha^{2N+j-3}u_1 & \alpha^{2N+j-4}u_2 & \cdots & \alpha^{2N-2}u_j \\ \vdots & \vdots & & \vdots \\ \alpha^{N+j-1}u_{N+1} & \alpha^{N+j-2}u_N & \cdots & \alpha^N u_{N+j-2} \\ \alpha^{N+j-2}u_N & \alpha^{N+j-3}u_{N+1} & \cdots & \alpha^{N-1}u_{N+j-1} \\ \alpha^{N+j-3}u_{N+1} & \alpha^{N+j-4}u_{N+2} & \cdots & \alpha^{N-2}u_{N+j} \\ \vdots & \vdots & & \vdots \\ \alpha^{j-1}u_{2N-1} & \alpha^{j-2}u_{2N} & \cdots & \alpha u_{2N+j-2} \end{bmatrix}. \quad (12)$$

新的 Y_p Hankel矩阵为

$$Y_p \triangleq Y_{0|N-1|N+j-2} = \begin{bmatrix} \alpha^{N+j-2}y_0 & \alpha^{N+j-3}y_1 & \cdots & \alpha^{N-1}y_{j-1} \\ \alpha^{N+j-3}y_1 & \alpha^{N+j-4}y_2 & \cdots & \alpha^{N-2}y_j \\ \vdots & \vdots & & \vdots \\ \alpha^{j-1}y_{N-1} & \alpha^{j-2}y_N & \cdots & \alpha y_{N+j-2} \end{bmatrix}. \quad (13)$$

改进的子空间预测模型如下:

$$\hat{Y}_f = F + (L_w L_u) \cdot \begin{bmatrix} \alpha^{2N+j-2}u_0 & \alpha^{2N+j-3}u_1 & \cdots & \alpha^{2N-1}u_{j-1} \\ \alpha^{2N+j-3}u_1 & \alpha^{2N+j-4}u_2 & \cdots & \alpha^{2N-2}u_j \\ \cdots & \cdots & \cdots & \cdots \\ \alpha^{N+j-1}u_{N+1} & \alpha^{N+j-2}u_N & \cdots & \alpha^N u_{N+j-2} \\ \alpha^{2N+j-2}y_0 & \alpha^{2N+j-3}y_1 & \cdots & \alpha^{2N-1}y_{j-1} \\ \vdots & \vdots & & \vdots \\ \alpha^{N+j-1}y_{N-1} & \alpha^{N+j-2}y_N & \cdots & \alpha^N y_{N+j-2} \\ \alpha^{N+j-2}u_N & \alpha^{N+j-3}u_{N+1} & \cdots & \alpha^{N-1}u_{N+j-1} \\ \vdots & \vdots & & \vdots \\ \alpha^{j-1}u_{2N-1} & \alpha^{j-2}u_{2N} & \cdots & \alpha u_{2N+j-2} \end{bmatrix}. \quad (14)$$

3.4 改进的DDS的基本步骤(The procedures of improved DDS)

对反馈因子和遗忘因子的取值寻优的PSO算法设计如下:

参数符号:

I : 种群中的粒子总数;

D_I : PSO搜索空间的维数;

Q : 最大迭代次数;

d : 维数标志, $d = 1, 2, \cdots, D_I$;

i : 粒子标志, $i = 1, 2, \cdots, I$;

q : 迭代次数标志, $q = 1, 2, \cdots, Q$;

V_i : 粒子 i 的速度, $V_i = v_{i1}, v_{i2}, \cdots, v_{iD_I}$;

P_i : 粒子 i 的位置, $P_i = p_{i1}, p_{i2}, \cdots, p_{iD_I}$;

O_i : 粒子 i 的当前最好解, $O_i = o_{i1}, o_{i2}, \cdots, o_{iD_I}$;

G_b : 种群的当前最好解, $G_b = g_{b1}, g_{b2}, \cdots, g_{bD_I}$;

c_1, c_2 : 学习因子(飞向个体最好粒子和全局最好粒子时的步长);

r_1, r_2 : $[0, 1]$ 之间的随机数;

w : 惯性权重.

粒子的速度和位置的更新方式如下:

$$v_{id}^{q+1} = wv_{id}^q + c_1r_1(o_{id}^q - p_{id}^q) + c_2r_2(g_{bd}^q - p_{id}^q), \quad (15)$$

$$p_{id}^{q+1} = p_{id}^q + v_{id}^{q+1}, \quad (16)$$

其中: 个体学习系数 c_1 和社会学习系数 c_2 都按常规方法设为2. 惯性权重 w 起到扰动的作用, 能够防止算法过早收敛, 根据文献报道和实验测试, 本文的取值为0.98. PSO的适值函数定义为预测偏差.

引入PSO算法后, 改进的子空间预测方法的详细求解步骤归纳如下:

步骤 1 构造 $U_p, U_f, U_{\text{train}}, U_{\text{test}}$ 和 $Y_p, Y_f, Y_{\text{train}}$ 的Hankel 矩阵;

步骤 2 用子空间预测模型计算子空间矩阵 L_w 和 L_u ;

步骤 3 用 U_{train} 的Hankel矩阵计算出训练时预测值;

步骤 4 计算步骤3中得到预测值与实际输出值 Y_{train} 之间的差值, 将其乘以PSO优化后的反馈因子 H 的值, 得到 F , 重复进行此步骤直到得到 F 的最小值;

步骤 5 用 U_{test} 的Hankel矩阵预测式(14)的输出.

4 实验结果(Experimental results)

本文提出的算法在配置为Inter(R) Core(TM) 2 Duo CPU (2.33GHz) CPU的电脑上用C++语言编程实现, 操作系统为Windows XP. 实验数据取自某国内大型钢铁企业的实际数据.

为了验证本章提出的算法的性能, 设计了两个实验: 炼铁工序的高炉煤气发生量预测和热轧工序的电力消耗预测. 测试之前, 先采用统计工具或曲线分析炼铁和热轧工序的实际数据, 以获得稳定的、全面的样本数据. 稳定指的是样本数据要对应于正常的稳定生产, 全面指的是样本数据要包含尽可能丰富的数据类型. 在实验中考考虑两种改进策略:

- 1) 只引入反馈因子 H ;
- 2) 引入反馈因子和遗忘因子.

为了测试改进策略的性能, 实验中考考虑了3种方案:

SDDS(standard version of data-driven based subspace algorithm): 标准形式的数据驱动子空间算法;

DDSF(data-driven based subspace algorithm with feedback factor): 带反馈因子的数据驱动子空间算法;

DDSFF(data-driven based subspace algorithm with feedback factor and forgetting factor): 带反馈因子和遗忘因子的数据驱动子空间算法.

反馈因子和遗忘因子的取值由PSO算法来寻优,

PSO算法的参数如表1所示.

表 1 PSO参数取值列表
Table 1 The parameter list of PSO

参数	取值
种群规模	50
最大迭代次数	100
惯性权重	0.98
个体认知系数	2
社会系数	2
精英解学习系数	2
反馈调节因子 H 的下界值	0
反馈调节因子 H 的上界值	1
遗忘因子的下界	0
遗忘因子的上界	1
遗忘因子的最大速度	5
反馈因子的最大速度	2

4.1 BFG发生量预测(BFG generation prediction)

表2和表3给出了高炉煤气发生量预测问题所采用的测试数据及实验结果. DDSF的反馈因子的取值为0.009; DDSFF的反馈因子和遗忘因子的取值分别为0.62和0.893. 表3中的实验结果表明:

1) 与SDDS的结果相比, DDSF对预测偏差的降低不明显, 说明反馈因子对标准子空间算法的性能改进量大;

2) DDSFF显著提高了预测精度, 虽然第4组数据未得到改进, 但其精度也在可接受的范围内.

表 2 BFG发生量的输入输出数据
Table 2 Input and output data of BFG generation

输入变量 U		输出变量 Y
送风流量/ m^3	热风炉流量/ m^3	BFC发生量/ m^3
6807.08	227.53	528.93
6809.12	234.53	528.07
6805.05	235.11	513.98
6819.8	236.2	531.38
6812.17	235.58	533.42
6808.61	231.22	479.33
6814.2	232.74	424.94
6799.96	236.61	444.26
6808.61	237.52	497.81
6814.2	236.34	517.16
6817.26	235	510.71
6825.4	236.81	517.44
6813.19	235.97	521.41
6801.99	235.39	514.83
6801.49	235.61	513.08

表 3 BFG 发生量实验结果
Table 3 The experimental results of BFG generation

实际值/ m^3	预测值/ m^3			偏差/%		
	SDDS	DDSF	DDSFF	SDDS	DDSF	DDSFF
513.05	495.31	495.45	508.65	3.4577	3.4307	0.8568
499.8	455.63	456.19	509.36	8.8368	8.7255	1.9127
496.19	469.08	469.55	516.57	5.4637	5.3688	4.1064
501.35	515.57	515.56	525.55	2.8361	2.8348	4.8265
507.59	541.67	541.41	535.75	6.713	6.6624	5.5479

4.2 热轧工序能源消耗量预测(Energy consumption prediction in hot rolling operation)

表4和表5分别给出了热轧工序能源预测问题所采用的测试数据集及实验结果. 在DDSF中, 反馈因子的取值为0.024; 在DDSFF中, 反馈因子和遗忘因子的取值分别为0.15和0.948.

表5的结果表明:

- 1) 与SDDS相比, DDSF对预测偏差的降低量不甚明显, 再次说明反馈因子的引入对标准子空间算法性能的改善更大;
- 2) DDSFF策略对预测精度的改进量最明显, 表明了该策略的有效性;
- 3) 与高炉煤气发生量的预测结果相比, 热轧产线的预测精度偏低, 未来需要考虑更多更全面的因素影响热轧产线能源介质消耗量的因素.

表 4 热轧工序电力消耗输入输出数据
Table 4 Input and output data of hot rolling operation

输入变量 U		输出变量 Y
气温/ $^{\circ}\text{C}$	产量/ t	消耗量/ $(\text{kWh} \cdot \text{t}^{-1})$
3.6	31.01	93.6
6	27.36	92.8
8.6	29.58	96.6
12.8	26.12	94.9
21.1	30.52	96.3
24.8	23.26	98.5
28.3	32.66	94.9
29.7	32.58	92.8
24.2	29.23	95.8
17.1	30	97
11.9	27	98
6.5	30	97
2.2	22.82	107
7.3	22.54	102.3
9.8	18.56	115.2

表 5 热轧工序实验结果
Table 5 The experimental results of hot rolling operation

实际值/ $(\text{kWh} \cdot \text{t}^{-1})$	预测值/ $(\text{kWh} \cdot \text{t}^{-1})$			偏差/%		
	SDDS	DDSF	DDSFF	SDDS	DDSF	DDSFF
114.2	95.81	95.88	98.32	16.101	16.042	13.92
105.8	94.34	94.42	94.92	10.827	10.749	10.26
102.8	90.87	91.37	94.08	11.609	11.122	8.465
97.3	93.71	93.97	94.26	3.692	3.419	3.126
98.6	96.19	96.78	97.1	2.442	1.844	1.526

5 结论(Conclusion)

本文研究了钢铁企业炼铁工序的高炉煤气发生量预测问题和热轧工序的能源消耗量预测问题. 通过对能源消耗和影响因素的全面分析, 建立了基于数据的预测模型. 设计了数据驱动的子空间求解方法, 由过去的输入输出数据和未来的已知信息直接推导出子空间矩阵. 提出了反馈因子用以减弱算法对数据的依赖性以及对所选数据的敏

感性. 由于越接近当前时刻的数据越能准确反映系统状态, 对于不同时间的数据引入了遗忘因子使其对系统有不同的影响程度. 反馈因子和遗忘因子的取值采用PSO算法进行优化. 上述两种子空间方法的改进策略的性能通过测试钢铁企业的实际数据来验证. 结果表明, 改进的子空间方法在预测精度方面优于标准形式的子空间方法, 预测结果能够为企业制定准确的能源预算提供有效的

参考和决策支持.

参考文献(References):

- [1] 张加云, 张德江, 陈蕾. 钢铁企业能耗预测方法综述 [J]. 冶金能源, 2009, 28(2): 11 – 18.
(ZHANG Jiayun, ZHANG Dejiang, CHEN Lei. Summary of the methods of energy consumption forecast in iron and steel enterprises [J]. *Energy for Metallurgical Industry*, 2009, 28(2): 11 – 18.)
- [2] LIU Y, ZHAO J, WANG W. A time series based prediction method for a coke oven gas system in steel industry [J]. *ICIC Express Letters*, 2010, 4(4): 1373 – 1378.
- [3] 刘颖, 赵珺, 王伟, 等. 基于数据的改进回声状态网络在高炉煤气发生量预测中的应用 [J]. 自动化学报, 2009, 35(6): 731 – 738.
(LIU Ying, ZHAO Jun, WANG Wei, et al. Improved echo state network based on data-driven and its application to prediction of blast furnace gas output [J]. *Acta Automatica Sinica*, 2009, 35(6): 731 – 738.)
- [4] 张全, 刘渺, 凌振华, 等. 钢铁企业能耗预测系统的设计 [J]. 冶金动力, 2006, 2: 67 – 70.
(ZHANG Quan, LIU Miao, LING Zhenhua, et al. Design of energy consumption prediction system of iron and steel enterprises [J]. *Metallurgical Power*, 2006, 2: 67 – 70.)
- [5] 杨宏韬, 张德江, 李秀兰, 等. 遗传神经网络能耗预测模型在钢铁企业中的应用 [J]. 长春工业大学学报(自然科学版), 2007, 28(增刊): 186 – 189.
(YANG Hongtao, ZHANG Dejiang, LI Xiulan, et al. Application of energy consumption estimation model based on generic neural in steel manufacture [J]. *Journal of Changchun University of Technology (Natural Science Edition)*, 2007, 28(Suppl): 186 – 189.)
- [6] 杨华, 李少远. 一种完全数据驱动的子空间辨识与鲁棒预测控制器设计 [J]. 控制理论与应用, 2007, 24(5): 732 – 742.
(YANG Hua, LI Shaoyuan. A novel robust predictive controller design based on data-driven subspace identification [J]. *Control Theory & Applications*, 2007, 24(5): 732 – 742.)
- [7] GOETHALS I, PELCKMANS K, SUYKENS J, et al. Subspace identification of hamerstein systems using least squares support vector machines [J]. *IEEE Transactions on Automatica Control*, 2005, 50(10): 1509 – 1519.
- [8] FAVOREEL W, DE MOOR B, VAN OVERSCHEE P. Subspace identification of bilinear systems subject to white inputs [J]. *IEEE Transactions on Automatica Control*, 1999, 44(6): 1157 – 1165.
- [9] VAN OVERSCHEE P, DE MOOR B, DEHANDSCHUTTER W. A subspace algorithm for the identification of discrete time frequency domain power spectra [J]. *Automatica*, 1997, 33(12): 2147 – 2157.
- [10] HUANG B, KADALI R. *Dynamic Modeling, Predictive Control and Performancemonitoring* [M]. London: Springer-Verlag, 2008.
- [11] HUANG X, HUANG R B, CHEN T W. Data-driven predictive control for solid oxide fuel cells [J]. *Journal of Process Control*, 2007, 17(2): 103 – 114.
- [12] KENNEDY J, EBERHART R. Particle swarm optimization [C] // *Proceedings of IEEE International Conference on Neural Networks*. Piscataway, NJ: IEEE Service Center, 1995, IV: 1942 – 1948.
- [13] EBERHART R, KENNEDY J. A new optimizer using particle swarm theory [C] // *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*. Piscataway, NJ: IEEE Service Center, 1995: 39 – 43.
- [14] 杨华. 基于子空间方法的系统辨识及预测控制设计 [D]. 上海: 上海交通大学, 2007.
(YANG Hua. *Subspace methods for system identification and predictive control design* [D]. Shanghai: Shanghai Jiao Tong University, 2007.)
- [15] VAN OVERSCHEE P, DE MOOR B. *Subspace Identification for Linear Systems: Theory, Implementation, Applications* [M]. Norwell: Kluwer Academic Publishers, 1996.
- [16] 杨华, 李少远. 一种新的基于遗忘因子的递推子空间辨识算法. 控制理论与应用, 2009, 26(1): 69 – 72.
(YANG Hua, LI Shaoyuan. A novel recursive MOESP subspace identification algorithm based on forgetting factor [J]. *Control Theory & Applications*, 2009, 26(1): 69 – 72.)

作者简介:

张颜颜 (1975–), 女, 讲师, 博士, 主要从事能源优化、生产调度等方面的研究工作, E-mail: zhangyanyan@ise.neu.edu.cn;

唐立新 (1966–), 男, 教授, 博士生导师, 主要从事最优控制理论及应用、生产物流计划调度与智能优化等方面的研究工作, E-mail: qhjytlx@mail.neu.edu.cn.