

文章编号: 1001-0920(2010)08-1178-06

基于最小二乘支持向量机的焦炉煤气柜位预测模型及应用

张晓平¹, 赵 珺¹, 王 伟¹, 丛力群², 冯为民², 陈伟昌³

(1. 大连理工大学 信息与控制研究中心, 大连 116023; 2. 上海宝信软件股份有限公司
自动化部, 上海 201203; 3. 上海宝山钢铁股份有限公司 能源部, 上海 200431)

摘要: 针对焦炉煤气柜位难以机理预测问题, 通过分析煤气的产消及柜位变化特点, 建立了基于最小二乘支持向量机的柜位预测模型. 构造梯度网格搜索算法优选模型参数和大样本筛选方法选取训练样本, 从而提高了预测精度. 上海宝钢实际煤气数据的仿真结果表明, 所建模型参数选取耗时少, 预测效果良好, 可为煤气的平衡调度提供科学指导.

关键词: 焦炉煤气系统; 柜位预测; 最小二乘支持向量机; 快速留一法; 梯度网格搜索

中图分类号: TP181

文献标识码: A

COG holder level prediction model based on least square support vector machine and its application

ZHANG Xiao-ping¹, ZHAO Jun¹, WANG Wei¹, CONG Li-qun², FENG Wei-min², CHEN Wei-chang³

(1. Research Center of Information and Control, Dalian University of Technology, Dalian 116023, China; 2. Automation Department, Shanghai Baosight Software Company Limited, Shanghai 201203, China; 3. Energy Department, Shanghai Baosteel Company Limited, Shanghai 200431, China. Correspondent: ZHANG Xiao-ping, E-mail: zhang_xiaoping@163.com)

Abstract: Aiming at the prediction for coke oven gas holder level in steel enterprises, which is very difficult to be modeled using the mechanism modeling, a gasholder level prediction model combined with the analysis of the gas production-consumption and level variation is established based on the least square support vector machine. A gradient grid search algorithm for selecting the model's parameters and an effective big samples selection approach to build the training samples are proposed to improve the prediction accuracy. The simulation results using the practical gas data in Shanghai Baosteel show that, the proposed method has shorter parameter optimization time and better performance, and can provide scientific guidance for the gas balance scheduling process.

Key words: Coke oven gas; Holder level prediction; Least square support vector machine; Fast leave one out method; Gradient grid search

1 引言

钢铁炼焦过程中产生的焦炉煤气(COG)是各生产环节所需的重要二次能源. 由于其比热高, 产量大, 有效合理地利用将直接影响着企业的能耗标准和产出成本. 但由于煤气柜的容量限制和各用户的实时煤气产消复杂多变, 煤气供需不平衡经常发生. 分析焦炉煤气的产消特点, 实现煤气柜位的有效预测, 对煤气的合理调度具有重要的指导作用.

目前, 国内外关于煤气柜位预测的研究相对较少. 文献[1]通过分析煤气柜位与煤气产消差的变化

特点, 建立了人工机理柜位预测模型, 利用前一刻的柜位和产消差来预测当前时刻的柜位, 但未考虑各用户煤气产消变化对柜位的影响. [2]建立了基于时间序列的柜位预测模型, 仅根据前一段时间柜位情况来预测未来柜位, 同样忽略了实际煤气产消变化的影响. [3]将煤气柜进出流量作为因变量建立了柜位变化量模型, 但该模型应用时, 与实际生产情况偏差较大. 鉴于焦炉煤气柜位预测问题的复杂机理和高度非线性等特点, 采用基于数据驱动建模的方法来构建预测模型是一种有效的尝试. 此类方法包括人工神经网络

收稿日期: 2009-07-21; 修回日期: 2009-11-18.

基金项目: 国家863计划项目(2007AA04Z1A9).

作者简介: 张晓平(1980—), 女, 山东淄博人, 博士生, 从事钢铁生产及能源计划的研究; 王伟(1955—), 男(满族), 辽宁鞍山人, 教授, 博士生导师, 从事预测控制、复杂工业过程建模与控制等研究.

络(ANN)和支持向量机(SVM)等. 其中基于结构风险最小化原则的最小二乘支持向量机^[4](LSSVM)的模型结构比ANN更容易确定, 算法较SVM简单快速, 已成功应用于较多实际复杂过程的建模^[5-7]. 所建模型的精度取决于模型参数的优化选取及训练样本的合理构造, 但常用的参数优化方法中, 基于 l 折交叉验证的普通网格搜索参数优化^[8]需要重复对矩阵求逆, 耗时较长, 且搜索过程存在盲目性, 不能快速获得较优的参数; 基于贝叶斯框架的优选法^[9]需要求解非线性方程组, 计算过程复杂, 仅适合优化机理简单的模型. 而构造训练样本时, 若连续选取的样本太少, 则不能完全包括建模对象的各状态变化, 导致模型不完整, 推广性极差; 反之, 若样本太多, 则存在冗余, 模型不仅花费大量训练时间, 精度也得不到保证.

本文针对焦炉煤气柜位预测问题, 通过分析煤气的产消变化, 建立了基于LSSVM的柜位预测模型. 鉴于上述LSSVM预测建模时存在的问题, 提出了梯度网格参数优化算法来优选模型参数. 优化过程中, 采用快速留一法验证, 缩短了优化时间; 以全样本留一预测误差平方和的梯度来指导网格搜索, 获得了更好的优化结果. 构造了大样本筛选方法来选取模型的训练样本, 避免了连续样本下模型不准确的问题. 利用上海宝钢实际焦炉煤气数据对所建柜位预测模型进行有效性和可行性验证, 结果表明所建模型参数优化耗时少, 且预测效果好于现场人工机理推理模型和基于普通网格搜索及连续样本的LSSVM预测模型.

2 问题描述

典型COG系统由煤气发生源、消耗用户、调节用户、煤气输送管网和煤气柜组成. 焦炉在炼焦过程中产生的COG一部分通过煤气管网输送给各个消耗用户供其生产, 如焦炉自身、高炉热风炉、连铸、冷轧工序的加热炉等; 一部分分配给调节用户(CCP电厂、发电厂和低压锅炉)用于发电和产生蒸汽, 剩余的储存在煤气柜中. 但由于管网中各用户的煤气产消变化不平稳(如图1(a)~图1(g)), COG的供需经常出现不平衡, 导致柜位的频繁上升和下降(如图1(h)). 考虑煤气柜的容量限制和安全运行要求, 煤气调度人员需要时时根据柜位的变化来合理增减各调节用户的煤气用量, 确保柜位的稳定运行, 从而最大化利用过量煤气, 实现零放散.

3 COG柜位预测模型

3.1 LSSVM预测模型

文献[1]将前时刻柜位与用户煤气流量产消差作为柜位影响因素, 建立如下人工机理柜位预测模型:

$$H_t = H_{t-1} + \Delta\text{flow}_{t-1}. \quad (1)$$

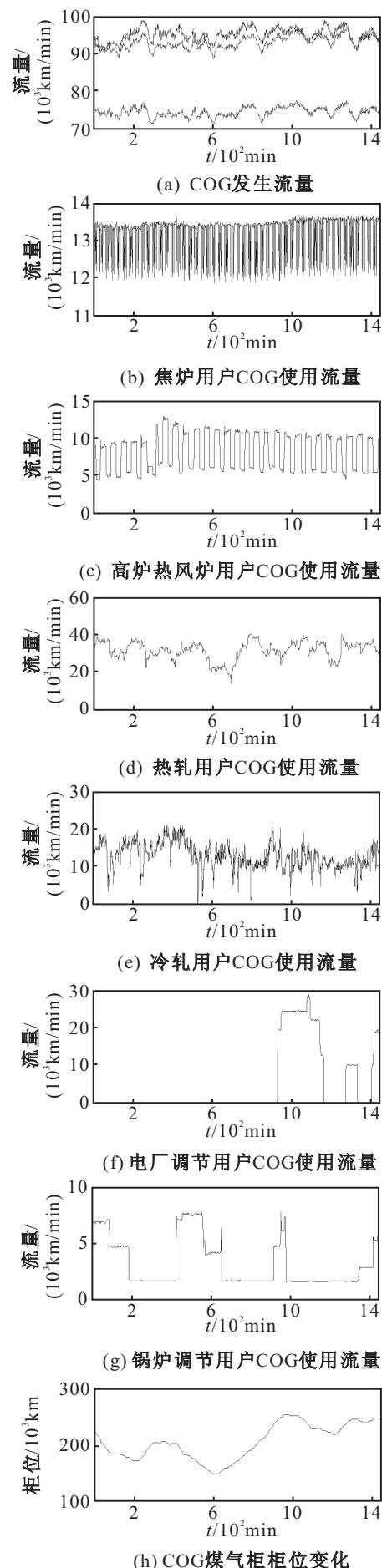


图1 各用户COG发生和使用流量变化及柜位变化

然而在实际生产中,产消差 Δflow 不能完全反映各用户煤气流量对柜位 H 的影响. 此外, 由于不同的流量变化和管网距离, 各用户煤气流量对柜位的影响也各不相同. 可见, COG 柜位与各用户煤气产消流量及自身柜位之间呈现复杂的非线性关系, 并非上述的简单线性关系. 为此, 本文建立了基于 LSSVM 的柜位预测模型. 训练样本集为

$$S = \{(U_{t_i-1}, H_{t_i-1}), H_{t_i}\}, i = 0, 1, \dots, N\}.$$

其中: U_{t_i-1} 和 H_{t_i-1} 组成模型输入样本 X_{t_i-1} , H_{t_i-1} 表示前一时刻的柜位; H_{t_i} 为模型输出样本, 表示当前时刻的柜位; $U_{t_i-1} \in R^m$ 为影响柜位的 m 个用户前一时刻煤气发生和使用流量. 综合考虑经验风险和模型复杂度, LSSVM 柜位预测模型即为求解如下约束优化问题^[4]:

$$\begin{aligned} \min J_1\{w, b, e\} &= \frac{1}{2}w^T w + C \sum_{i=1}^N e_{t_i}^2, \\ \text{s.t. } H_{t_i} &= w^T \varphi(X_{t_i-1}) + b + e_{t_i}. \end{aligned} \quad (2)$$

其中: $w \in R^p$, $b \in R$ 分别为高维空间中预测模型的系数和常数项; C 为惩罚因子; $e_{t_i} \in R$ 为第 i 柜位样本的拟合误差; $\varphi(\cdot): R^{m+1} \rightarrow R^p$ 为输入样本空间至核空间的映射函数. 引入 Lagrange 乘子 $\alpha > 0$, LSSVM 柜位预测模型可转换为求解如下无约束最优化问题:

$$\begin{aligned} L\{w, b, e, \alpha\} &= \\ &\frac{1}{2}w^T w + C \sum_{i=1}^N e_{t_i}^2 - \\ &\sum_{i=1}^N \alpha_i (w^T \varphi(X_{t_i-1}) + b + e_{t_i} - H_{t_i}). \end{aligned} \quad (3)$$

由最优性 KKT(Karush-Kuhn-Tucker) 条件可得到关于预测模型系数的线性方程组为

$$\begin{bmatrix} K + I/C & 1 \\ 1^T & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} H \\ 0 \end{bmatrix}. \quad (4)$$

其中: K 为 $N \times N$ 的核矩阵, 其元素为

$$k_{t_i-1, t_j-1} = k(X_{t_i-1}, X_{t_j-1}) = \varphi(X_{t_i-1})^T \varphi(X_{t_j-1}),$$

$k(\cdot, \cdot)$ 为核函数, 一般常取径向基函数

$$k(X_{t_i}, X_{t_j}) = e^{-\|X_{t_i} - X_{t_j}\|^2 / 2\sigma^2};$$

I 为单位矩阵; $1 = [1, 1, \dots, 1]^T$; H 为柜位输出样本组成的向量. 求解式 (4) 可得到柜位 LSSVM 预测模型的表达式为

$$H_t(X_{t-1}) = \sum_{i=1}^N \alpha_i k(X_{t-1}, X_{t_i-1}) + b. \quad (5)$$

若预测时段为 l , 本文采用在线学习算法^[10]构建在线递推预测模型, 即每预测一个时刻柜位, 便将该时刻的柜位预测值增加到下一时刻的训练样本中, 重新训练模型进行下一时刻的预测, 而递增矩阵的逆由

初始矩阵的逆得到, 从而大大缩短了预测时间. 从所建预测模型也可看出, 核函数及其参数 σ 决定了模型的表达形式, 惩罚因子 C 折衷模型的复杂度和经验风险. 如果 σ 和 C 选取不当, 则会造成预测模型过拟合和欠拟合, 致使预测能力下降. 此外, 训练样本 S 的构成也影响模型的准确性: 样本集过大, 求解耗时; 而样本集过小, 模型不完整. 本文分别就这两方面进行改进来优化柜位预测模型.

3.2 梯度网格参数优化算法

由于 l 折交叉验证普通网格搜索过程中的反复求逆和缺乏指导的网格遍历跳跃搜索会消耗大量时间, 无法满足现场实时性要求, 且优化结果也不理想, 本文提出基于网格搜索和梯度法相结合的参数优化方法. 采用快速留一验证过程^[11], 缩短了参数优化时间, 并基于快速留一法导出网格梯度搜索方向来避免普通网格搜索方向的盲目性, 使搜索始终向小的留一检验误差方向进行, 从而获得较优的模型参数.

3.2.1 快速留一法交叉验证及网格梯度

记

$$A = \begin{bmatrix} K + I/C & 1 \\ 1^T & 0 \end{bmatrix}, B = \begin{bmatrix} H \\ 0 \end{bmatrix}, s = \begin{bmatrix} a \\ b \end{bmatrix},$$

则对模型进行第 i 次留一交叉验证时, 柜位 LSSVM 预测模型的系数为

$$s_i = s_{i-} - \frac{s(i)}{A^{-1}(i, i)} A^{-1}(i^-, i). \quad (6)$$

其中: s_{i-} 为全训练样本下模型系数 s 中去掉第 i 个元素的系数列向量; $s(i)$ 为 s 的第 i 个元素; $A^{-1}(i^-, i)$ 为全训练样本下 A^{-1} 的第 i 列去掉第 i 个元素的列向量; $A^{-1}(i, i)$ 为 A^{-1} 第 i 对角元素. 令 $\tilde{K} = [K \ 1]$, 则全样本留一法交叉检验误差 PRSEE (Predicted residual sum of squares) 为

$$\text{PRSEE} = \sum_{i=1}^N (\tilde{K}(i, i^-) s_i - H_i)^2. \quad (7)$$

由式 (7) 可知, N 次参数验证, 只需对 A 求逆一次, 相比 l 重交叉验证的多次求逆, 计算量显著减小, 缩短了验证过程. 即基于式 (7) 对参数 (C, σ) 进行优化, 使全体检验误差 PRSEE 达到最小. 鉴于网格搜索时其方向的盲目遍历性导致优化结果不理想, 本文导出 PRSEE 对参数 (C, σ) 的偏导数作为指导网格搜索的梯度方向, 推导过程如下:

Step 1: 求 \tilde{K} 和 $k(X_i, X_j)$ 对参数 σ 的偏导数, 即

$$\begin{aligned} \frac{\partial \tilde{K}}{\partial \sigma} &= \begin{bmatrix} \frac{\partial K}{\partial \sigma} & 0 \end{bmatrix}, \\ \frac{\partial k(X_i, X_j)}{\partial \sigma} &= -\sigma^{-3} k(X_i, X_j) \|X_i - X_j\|^2. \end{aligned} \quad (8)$$

Step 2: 求 M^{-1} 对参数 (C, σ) 的偏导数, 即

$$\frac{\partial A^{-1}}{\partial C} = -A^{-1} \begin{bmatrix} -C^{-2}I & 0 \\ 0^T & 0 \end{bmatrix} A^{-1},$$

$$\frac{\partial A^{-1}}{\partial \sigma} = -A^{-1} \begin{bmatrix} \partial K/\partial \sigma & 0 \\ 0^T & 0 \end{bmatrix} A^{-1}. \quad (9)$$

Step 3: 求系数 s 对参数 (C, σ) 的偏导数, 即

$$\frac{\partial s}{\partial C} = \frac{\partial A^{-1}}{\partial C} B = -A^{-1} \begin{bmatrix} -C^{-2}I & 0 \\ 0^T & 0 \end{bmatrix} A^{-1} B,$$

$$\frac{\partial A^{-1}}{\partial \sigma} = \frac{\partial A^{-1}}{\partial C} B = -A^{-1} \begin{bmatrix} \partial K/\partial \sigma & 0 \\ 0^T & 0 \end{bmatrix} A^{-1} B. \quad (10)$$

Step 4: 求系数 s_i 对参数 (C, σ) 的偏导数, 即

$$\frac{\partial s_i}{\partial C} = \frac{\partial s}{\partial C}(i^-) - \frac{s(i) \frac{\partial A^{-1}}{\partial C}(i^-, i) + \frac{\partial s}{\partial C}(i) A^{-1}(i^-, i)}{A^{-1}(i, i)} + \frac{s(i) A^{-1}(i^-, i) \frac{\partial A^{-1}}{\partial C}(i, i)}{[A^{-1}(i, i)]^2} + \frac{\partial s}{\partial \sigma}(i^-) - \frac{s(i) \frac{\partial A^{-1}}{\partial \sigma}(i^-, i) \frac{\partial s}{\partial \sigma}(i) A^{-1}(i, i)}{A^{-1}(i, i)} + \frac{s(i) A^{-1}(i^-, i) \frac{\partial A^{-1}}{\partial \sigma}(i, i)}{[A^{-1}(i, i)]^2} \frac{\partial s_i}{\partial \sigma}. \quad (11)$$

Step 5: 求 PRSEE 对参数 (C, σ) 的偏导数, 即

$$\frac{\partial \text{PRSEE}}{\partial C} = \sum_{i=1}^N 2(\tilde{K}(i, i^-) s_i - H_i) \frac{\partial s_i}{\partial C},$$

$$\frac{\partial \text{PRSEE}}{\partial \sigma} = \sum_{i=1}^N 2(\tilde{K}(i, i^-) s_i - H_i) \times \left(\frac{\partial \tilde{K}}{\partial \sigma}(i, i^-) s_i + \tilde{K}(i, i^-) \frac{\partial s_i}{\partial \sigma} \right). \quad (12)$$

Step 6: 将式 (9)~(11) 代入 (12), 得到网格搜索的梯度方向。

3.2.2 梯度网格搜索参数优化步骤

结合全样本留一法交叉检验误差和网格梯度公式, 给出本文构造的梯度网格参数优化步骤如下:

Step 1: 采用不等间距的方法划分网格, 对于参数 (C, σ) 按一定的间隔比例各划分成 100 份长度不等的网格;

Step 2: 在 (C, σ) 网格点组成的二维参数平面上, 找出中心网格参数点及其周围网格步长为 10 的 8 个网格参数点 (图 2 中以①标注);

Step 3: 通过式 (7) 计算这 9 个参数点的快速留一法交叉验证误差 PRSEE, 选择最小的 PRSEE 对应的网格参数点 $(C, \sigma)_{\min \text{PRSEE}}$;

Step 4: 根据式 (12) 计算参数 $(C, \sigma)_{\min \text{PRSEE}}$ 对应的梯度值作为网格梯度方向, 采用最速下降法更

新 (C, σ) , 并将离更新参数点最近的网格点 (如图 2 中以②标注) 作为新中心点, 同时将网格步长减 1, 选取其周围的 8 个点;

Step 5: 重复 Step 3 和 Step 4, 直至网格步长减为零, 若此时的 PRSEE 不满足要求, 则重置网格步长, 重复以上步骤, 直至搜索到满足要求的最优参数点 (如图 2 中的黑点) 为止。

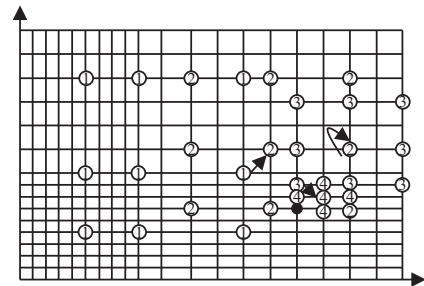


图 2 梯度网格搜索过程

如图 2 所示, 对于每次优化, 若初始步长为 10, 则采用普通网格搜索需要进行 100×100 次计算, 而采用梯度网格法最多需要进行 $10 \times 9 = 90$ 次运算即可得到模型参数, 节省了大量运算时间。另外, 由最速下降梯度来指导网格搜索方向, 可避免普通网格搜索的盲目搜索。实际数据实验表明, 采用本文参数优化算法能将参数搜索时间控制在 2 min 之内, 并且 PRSEE 满足要求。

3.3 大样本训练集筛选方法

考虑实际煤气调度时, 主要通过增减调节用户消耗来缓冲柜位的升降, 对柜位的影响最为直接。因此, 大样本训练集筛选方法建立样本集时主要考虑调节用户的各种运行情况。设调节用户个数为 k , 则选取步骤如下:

Step 1: 初始化 $j = 1$;

Step 2: 遍历 COG 历史数据, 记录第 j 个调整用户煤气流量的调整时刻 t_j 和被调整前后 30 min 所对应的时刻集 $\{t_{j-30}, \dots, t_{j-1}, t_j, t_{j+1}, \dots, t_{j+30}\}$;

Step 3: 令 $j = j + 1$, 重复 Step 3, 直至 $j = k$;

Step 4: 取 k 个调整用户时刻集的交集, 得到样本筛选时刻集 t_s , 采样每个时刻的各用户煤气流量和柜位作为输入样本, 下一时刻柜位作为输出样本。

采用上述方法建立的样本集不仅大大减少了连续所选样本中的冗余样本, 而且能够全面包含各种运行状态下的煤气流量数据, 从而提高了模型的准确性。在实际应用中, 若预测精度不满足要求, 则为了使模型反映系统最新的运行状况, 可以定时地采样新的训练样本集, 重新训练模型。

4 实例验证及应用

选取上海宝钢能源中心焦炉煤气系统 2009 年 2

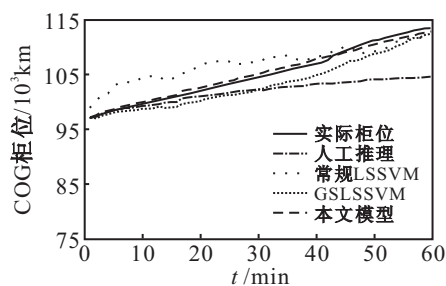
月7日至4月15日的数据对模型进行检验,采样间隔为1 min. 该系统每小时大约产生 300 km^3 COG, 联网的煤气产消用户多达80个, 且有两个煤气柜, 容积分别为 150 km^3 和 300 km^3 . 通常将容量较大的2号柜作为联网柜, 1号柜作为辅助柜. 当2号柜位值过低或过高时, 若1号柜能进行补偿, 则关闭2号柜, 将1号柜进行联网. 因此, 本文分别依据各个柜子联网运行数据建立了两个柜位预测模型, 并在实际应用中根据联网柜的变化自动切换预测模型. 在建模过程中, 为了减少冗余输入对柜位模型精度的影响, 借助现场煤气调度人员工作经验和灰色关联分析方法选取其中26个对柜位变化影响较大的用户作为模型的输入, 即 $m = 26$.

采用本文大样本数据选取方法, 在2个多月的生产数据中, 对于两个煤气柜各找到4028个和3578个数据点, 致使样本减少到约3天的数据量. 均选其中2000个组成模型的训练样本集, 设模型参数 C 和 σ 的搜索范围分别为 $[1, 10000]$ 和 $[0.1, 1000]$, 按第3.2节所述参数优化方法选取参数 (C, σ) , 其中: C 在 $[0.1, 100]$, $[100, 1000]$ 和 $[1000, 10000]$ 上各平分

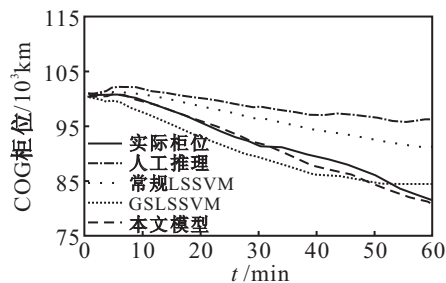
60个网格点; σ 在 $[0.1, 100]$, $[100, 500]$ 和 $[500, 1000]$ 上各平分成30, 30和40个网格点. 最终优选的1号煤气柜预测模型的参数为 $C = 60.972$, $\sigma = 8.316$; 2号煤气柜预测模型的参数为 $C = 100.441$ 和 $\sigma = 74.526$.

为了说明本文所建模型的预测性能, 采用人工推理^[1]、基于连续样本的常规LSSVM、基于普通网格搜索的LSSVM(GSLSSVM)分别建立相应的柜位预测模型与之比较, 其中核函数、训练样本个数与本文模型相同. 选取柜位不同变化趋势下60个样本点用于测试, 采用归一化均方根误差NRMSE、平均绝对误差MAE和模型建立耗时CT作为预测性能指标. 图3和图4分别为4种模型预测1号焦炉煤气柜和2号焦炉煤气柜上升趋势、下降趋势和平稳变化趋势的效果对比图. 表1给出了这4种预测模型的预测性能指标比较结果.

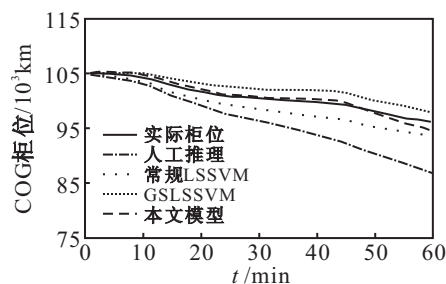
从图3和图4可以看出, 本文所建预测模型的预测效果最好, 能很好地跟踪柜位的升降和平稳变化; 而GSLSSVM次之, 常规LSSVM较差, 但结果依赖于参数和训练样本的选取; 人工推理预测效果最差, 当预测长度超过10 min后, 柜位预测值就开始偏



(a) 上升趋势

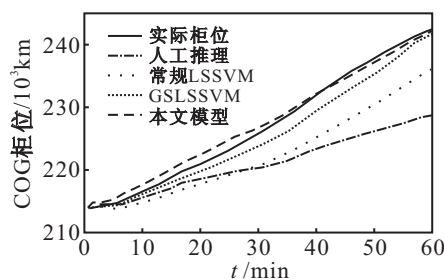


(b) 下降趋势

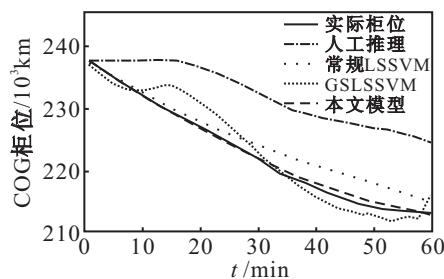


(c) 平稳趋势

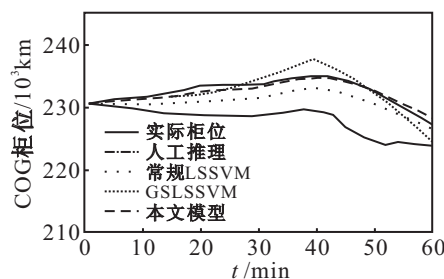
图3 1号COG柜位变化趋势预测效果



(a) 上升趋势



(b) 下降趋势



(c) 平稳趋势

图4 2号COG柜位变化趋势预测效果

离实际柜位,并随着预测长度的增加偏离越来越大。从表1所列各模型的预测性能指标中也可看出,本文模型产生的NRMSE最小,与实际柜位的平均偏离误差MAE保持在 1 km^3 以内。模型建立耗时较少,仅需 1 min 左右,满足了实际预测时间和精度要求。而GSLSSVM耗时最多,需要 5 min 才能获得预测结果;常规LSSVM耗时较少,但产生的误差均大于本文模型;人工推理模型耗时最少,但误差却最大,MAE最大达 9 km^3 。其原因是本文模型考虑了影响柜位的众多变化因素,较人工机理模型更合理;对模型参数优化进行了改进,采用快速留一法交叉验证并设计梯度网格搜索,较GSLSSVM中基于 l 折交叉验证的普通网格搜索更加快速有效;采用大样本训练集筛选方法选取样本集,避免了常规LSSVM的连续选取样本带来的样本集构造不全面的缺点,使模型能在较完整的样本集下构建,从而获得较高的预测精度。

表1 本文模型与人工推理^[1]、常规LSSVM和GSLSSVM模型的预测性能比较

预测对象	柜位变化	预测模型	NRMSE	MAE	CT/s
1号COG柜	上升趋势	人工推理	0.9575	3.3091	9.7125
		常规LSSVM	0.4757	2.7330	20.3516
		GSLSSVM	0.1805	1.7500	310.7115
		本文模型	0.0150	0.5038	59.7468
	下降趋势	人工推理	1.5177	6.5516	9.1597
		常规LSSVM	0.6267	4.1880	20.0181
		GSLSSVM	0.1385	2.1746	302.5411
		本文模型	0.0200	0.6781	58.7989
	平稳趋势	人工推理	3.8990	4.1485	9.7614
		常规LSSVM	0.6006	1.1443	21.1826
		GSLSSVM	0.3464	1.4247	311.9547
		本文模型	0.0272	0.4352	60.5463
2号COG柜	上升趋势	人工推理	0.6369	5.9918	10.5652
		常规LSSVM	0.3036	4.5855	20.3358
		GSLSSVM	0.0863	1.5362	300.698
		本文模型	0.0096	0.7888	59.6531
	下降趋势	人工推理	1.667	4.7277	10.1028
		常规LSSVM	0.5319	1.5854	21.0179
		GSLSSVM	0.0734	1.7192	311.374
		本文模型	0.0042	0.3859	59.4453
	平稳趋势	人工推理	7.2099	9.3539	9.3144
		常规LSSVM	1.2932	1.9846	20.3661
		GSLSSVM	0.4459	0.9666	307.2765
		本文模型	0.0717	0.4231	60.1124

5 结 论

准确预测焦炉煤气柜的柜位变化是焦炉煤气系统进行有效平衡调整的主要依据,本文以宝钢焦炉煤气系统为背景,提出了基于网格梯度参数优化和大样本选取的LSSVM柜位预测模型,解决了常规网格搜索参数优化耗时较长的问题,克服了连续样本选取对模型构建精度的影响。通过对现场柜位不同变化趋势

预测,结果表明所建模型的预测精度高于目前现场的人工推理模型,可以为焦炉煤气调度提供更有力的支持,同时模型建立耗时比较少,满足现场实时性要求。基于本文所建模型的宝钢焦炉煤气柜位预测系统已试运行于宝钢能源中心,预测效果良好。然而,实际煤气系统中,有时两个煤气柜同时运行在煤气管网中,今后工作应该对这种情况加以分析并建立相应的多柜位预测模型。

参考文献(References)

- [1] Chu Y H, Kim J H, Moon S J, et al. Control of gasholder level by trend prediction based on time-series analysis and process heuristics[C]. The 7th Int Symposium on Advanced Control of Chemical Processes. Hong Kong, 2004.
- [2] Iwao Higashi. Energy balance of steel mills and utilization of byproduct gases[J]. Trans of the Iron and Steel Institute of Japan, 1982, 22(1): 57-65.
- [3] 明德廷, 李娟, 尹怡欣. 钢铁企业煤气优化调度模型研究[J]. 计算机工程与设计, 2008, 29(6): 1575-1578. (Ming D T, Li J, Yin Y X. Research of gas optimal scheduling model in iron and steel enterprises[J]. Computer Engineering and Design, 2008, 29(6): 1575-1578.)
- [4] Suykens J A K, Van Gestel T, De Brabanter J, et al. Least squares support vector machines[M]. Singapore: World Scientific, 2002.
- [5] Gao L J, Tay F E H. Support vector machine with adaptive parameters in financial time series forecasting[J]. IEEE Trans on Neural Networks, 2003, 14(6): 1506-1518.
- [6] 黄景涛, 马龙华, 茅建波, 等. 基于支持向量回归的300MW电站锅炉再热汽温建模[J]. 中国电机工程学报, 2006, 26(7): 19-24. (Huang J T, Ma L H, Mao J B, et al. Modeling research of the reheat steam temperature of 300MW boiler based on support vector regression[J]. Proc of the Chinese Society for Electrical Engineering, 2006, 26(7): 19-24.)
- [7] 吕志军, 杨建国, 项前, 等. 基于支持向量机的纺纱质量预测模型研究[J]. 控制与决策, 2007, 23(7): 693-696. (Lv Z J, Yang J G, Xiang Q, et al. Research on support vector machines based on predictive model for yarn quality[J]. Control and Decision, 2007, 23(7): 693-696.)
- [8] De Kruijff B J, De Vries T J A. Pruning error minimization in least squares support vector machines[J]. IEEE Trans on Neural Networks, 2003, 14(3): 696-702.
- [9] Van Gestel T, Suykens J A K. Financial time series prediction using least squares support vector machines within the evidence framework[J]. IEEE Trans on Neural Networks, 2001, 12(4): 809-821.