

灰色特征加权 LSSVM 在砷盐除钴中的应用^①

伍铁斌^{②*} 李勇刚^{③*} 阳春华* 孙 备* 桂卫华*

(* 中南大学信息科学与工程学院 长沙 410083)

(** 湖南人文科技学院通信与控制工程系 娄底 417000)

摘 要 针对砷盐净化除钴中影响钴离子浓度预测的因素影响程度不同和数据噪声造成的钴离子浓度预测精度较低的问题,提出一种新的钴离子浓度预测模型——灰色特征加权最小二乘支持向量机(LSSVM)模型。该模型采用灰色累加的方法削弱原始数据中的噪声,并通过对砷盐除钴反应影响较为明显的因素的分析,根据影响因素的重要程度分别赋予其不同的特征权重,以提高 LSSVM 的预测精度;利用动态分级微粒群算法的快速收敛性和全局优化能力,优化选取 LSSVM 模型的两个关键参数——惩罚因子 C 和核参数 σ ,以避免参数选择的盲目性。基于选取的砷盐净化除钴过程生产数据进行了仿真,结果表明,灰色特征加权 LSSVM 的预测精度高,预测值能很好地跟踪实际钴离子浓度的变化趋势,满足砷盐除钴过程钴离子浓度预测的要求。

关键词 特征加权,最小二乘支持向量机(LSSVM),钴离子,微粒群算法,灰色累加

0 引言

净化除钴是湿法炼锌生产中极其重要的过程,钴离子浓度过高会引起锌反溶(即“烧板”)^[1]。由于净化除钴过程流程长和钴离子浓度化验的滞后性,导致除钴过程操作参数无法及时调节,从而影响净化除钴过程的优化运行;因此研究钴离子浓度预测具有重要意义。本文针对钴离子浓度影响因素及数据噪声对钴离子浓度预测精度的影响进行了研究,提出了一种灰色特征加权最小二乘支持向量机(least squares support vector machine, LSSVM)预测模型,该模型能满足砷盐除钴过程对钴离子浓度预测的要求。

1 相关研究

砷盐净化除钴过程中,氧化还原电位(oxidation reduction potential, ORP)、反应器含固量、溶液温度、反滴定(back titration, BT)值、入口溶液主要离子浓度等因素对置换除钴影响较大,且各因素对除钴反应速率的影响程度差别较大,具有强非线性特点,采

用多元线性回归等简单的回归模型无法满足钴离子预测的要求。文献[2]提出了基于神经网络与灰色模型的砷盐净化除钴过程钴离子浓度预测模型,仿真证明了该模型的可行性,但神经网络存在收敛速度较慢、容易陷入局部最优等问题。由于最小二乘支持向量机(least squares support vector machine, LSSVM)在回归建模中能有效避免过学习、局部最小问题^[3],因而它在非线性系统预测建模中得到广泛应用^[4-7]。文献[8]提出了一种结合小波分析和 LSSVM 方法的净化除钴过程钴离子浓度预测方法,但利用小波重构信号时会丢失原有时域特征,且没有考虑数据噪声的处理。文献[9]采用模糊支持向量机预测钴离子浓度,对不同时间段的样本赋予不同的权值,取得了一定的效果,但是该模型对数据噪声的处理能力较差。

针对砷盐净化除钴机理复杂和强非线性等特点而采用 LSSVM 预测钴离子浓度的方法面临以下挑战:由于砷盐净化除钴过程现场噪音和振动大,溶液腐蚀性强,使得生产数据含有大量噪声,会严重恶化 LSSVM 的预测性能;此外,砷盐净化除钴过程影响

① 国家自然科学基金(61174133, 61273159),国家杰出青年科学基金(61025015)和国家“十二五”科技支撑计划(2012BAF03B05)资助项目。

② 男,1981 年生,博士生;研究方向:复杂系统建模,优化与控制;E-mail: wutiebings1@163.com

③ 通讯作者, E-mail: liyonggang@csu.edu.cn
(收稿日期:2011-12-13)

因素众多且对除钴的影响程度不一样,如果对各影响因素赋予相同权重,会降低 LSSVM 预测精度^[10]。为此,本研究采用灰色累加方法削弱原始数据的噪声,根据影响砷盐净化除钴因素的重要程度赋予不同特征加权值,并取得显著效果。针对 LSSVM 模型参数难以确定^[11],本研究采用动态分级微粒群算法 (particle swarm optimization, PSO) 优化选取模型的惩罚因子 C 和核函数参数 σ 。

选取硫酸锌溶液净化除钴过程生产数据进行仿真的结果表明,本研究给出的钴离子浓度预测模型的预测精度高于 BP 神经网络和 LSSVM,且更能反映实际钴离子浓度变化的趋势,证明了该方法的合理性和有效性,能满足砷盐净化除钴生产过程钴离子浓度在线预测的要求。

2 砷盐净化除钴过程分析及灰色特征加权 LSSVM 预测模型建立

2.1 砷盐净化除钴机理分析

某锌冶炼企业砷盐 (As_2O_3) 净化除钴过程如图 1 所示。砷盐除钴是在反应器中溶液的含固量 (返回底流调节, 固体主要为促进除钴的合金)、反滴定 (BT) 值 (1BT 定义为溶液添加硫酸 0.4g/L 使得甲基橙由橙黄色变成红色)、温度和 Cu^{2+} (净化分除铜、除钴和除镉三段, 除铜工序留有适量铜离子^[12]) 等合理的情况下, 向溶液中加入 As_2O_3 和锌粉, 从而使溶液中的 Co^{3+} 以 $CoAs$ 合金形式沉淀, 达到净化除钴的要求。

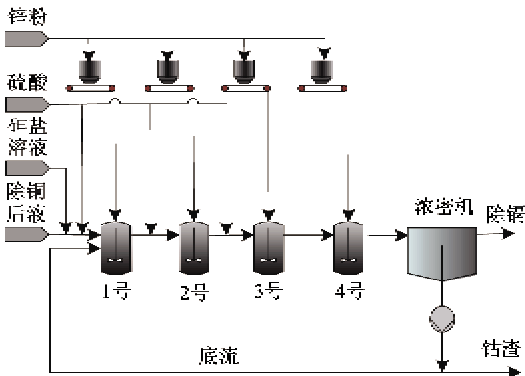


图 1 砷盐净化除钴工艺流程

控制可操作参数,使钴离子浓度在 4 个反应器中按某个锌粉总消耗量最少的梯度下降。1 号反应器出口的钴离子浓度的合理控制最重要,1 号出口钴离子浓度的准确预测影响整个流程优化控制的成败。

砷盐除钴反应为一阶动力学反应,动力学方程为^[13]

$$\frac{dC(t)}{dt} = -K(t) \times C(t) \tag{1}$$

其中 C 为钴离子浓度, K 为反应速率 (在同一个反应器内, 反应器入口参数基本相同时, 将 K 视为常数)。

通过偏相关性分析可知^[14], 影响 1 号反应器除钴速率的主要因素有: 氧化还原电位 (即 ORP, 电位越负, 溶液还原性越强, 反之氧化性越强; 采用 ORP 控制锌粉的添加量, 保持溶液合适的氧化还原性)、反应器中溶液的含固量、溶液温度、溶液 BT 值、除铜后液的铜离子、钴离子浓度、添加的砷盐。

图 2 是 1 号反应器 ORP 与 K 的关系 (条件: 含固量为 55g/L、温度 75℃、BT 为 1.2、入口钴离子浓度 30mg/L、铜离子浓度 300mg/L 和砷浓度 120mg/L 左右), 图中的数据点都是由 2 个原始数据的平均值组成 (图 3 至图 8 类同), 从中可以看出, ORP 越负, 钴离子的沉淀速率也越大, 影响极为显著。

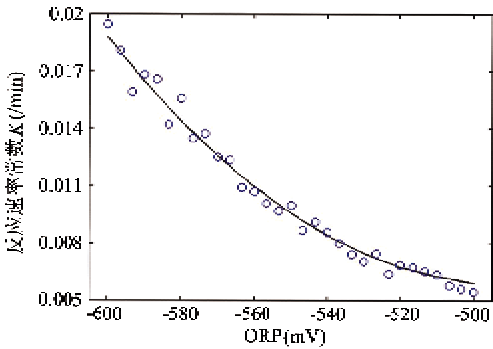


图 2 1 号反应器 ORP 与 K 关系

图 3 是 1 号反应器含固量与相对反应速率常数 K 的关系 (在其它条件相近, 仅含固量不同时得出的关系; K 并非实际值, 而是相对值, 即将某一条件下的 K 值定义为 1, 其它的与之相对比, 称为相对反应速率常数, 图 4 至图 8 类同)。

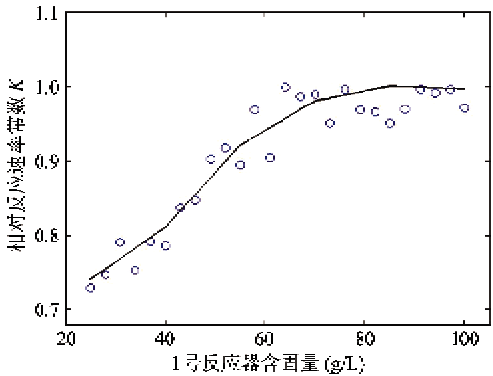


图 3 1 号反应器含固量与 K 关系

图 4 是 1 号反应器温度与相对反应速率常数 K 的关系。一般将温度控制在 $70 \sim 85^\circ\text{C}$ 。在该范围内,温度对除钴反应速率影响相对较小。

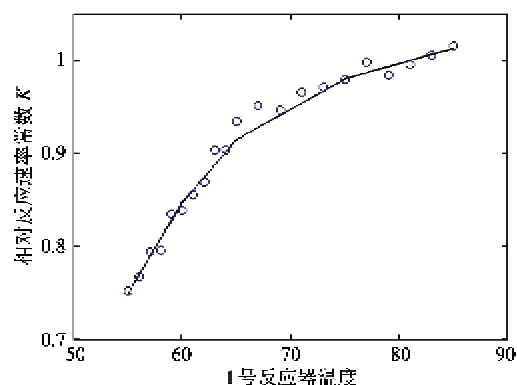


图 4 1 号反应器温度与 K 关系

图 5 是 1 号反应器 BT 值与相对反应速率常数 K 的关系。BT 值太大,意味着碱式硫酸盐的含量过高,会阻碍钴离子和锌粉的接触,降低除钴速率;BT 值太小时,利于氢离子和锌粉反应,故降低了钴离子还原的几率。

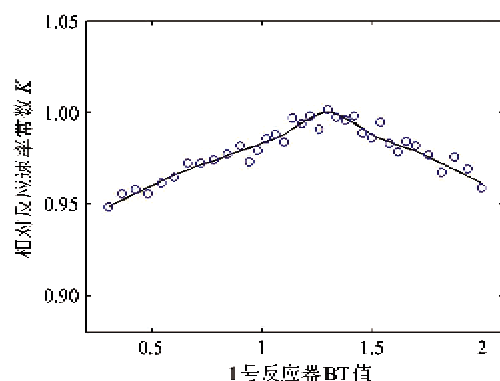


图 5 1 号反应器 BT 值与 K 的关系

图 6 是 1 号反应器入口钴离子浓度与相对反应速率常数 K 的关系。显然,入口钴离子浓度越高,

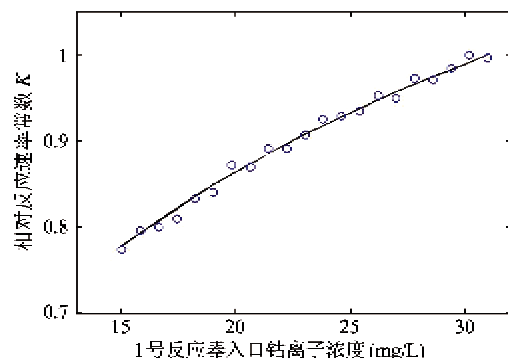


图 6 1 号反应器入口钴离子浓度与 K 的关系

与锌粉碰撞的几率越大,故反应速率越大。

图 7 是 1 号反应器入口铜离子浓度与相对反应速率常数 K 的关系。由于铜的标准电极电位 ($\text{Cu}/\text{Cu}^{2+}; +0.337\text{V}$) 比钴 ($\text{Co}/\text{Co}^{2+}; -0.277\text{V}$) 要正很多,所以铜离子比钴容易还原,铜离子浓度过高时会影响钴离子的还原。

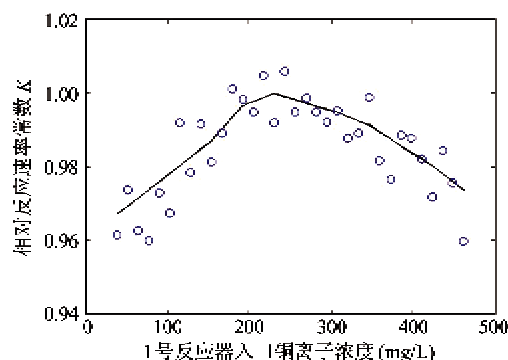


图 7 1 号反应器入口铜离子浓度与 K 的关系

图 8 是 1 号反应器入口砷盐浓度与相对反应速率常数 K 的关系(实际将砷盐溶解于浓的氢氧化钠,生成 NaAsO_2 ,酸性溶液中以 HAsO_2 形式存在)。还原后的砷与钴结合成砷化钴的合金,使得钴不易复溶,且合金能促进除钴反应的进行。但由于砷的标准电极电位 ($\text{HAsO}_2/\text{As}; 0.248\text{V}$) 远比钴的正,因此当砷的浓度高到一定程度时同样会影响除钴的进行。

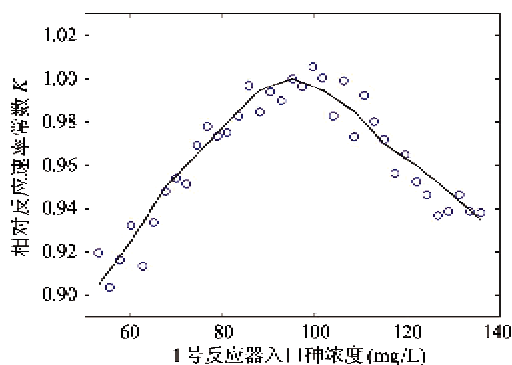


图 8 1 号反应器入口砷浓度与 K 的关系

从图 2 至图 8 的分析得出,各因素对除钴反应速率的影响程度差别比较大,且影响具有强非线性特点;采用多元线性回归等简单的回归建模方法显然难以满足预测钴离子浓度要求。

本文采用适合小样本、非线性系统建模的 LSS-VM 建立钴离子浓度预测模型。

2.2 特征加权最小二乘支持向量机

在支持向量机回归中,训练样本集 $Tg = (x_i, y_i), x_i \in R^M, y_i \in R, i = 1, 2, \dots, N$ 。利用映射函数 $\phi(x)$ 将样本映射到高维特征空间,转化成高维特征空间中线性函数估计问题:

$$y(x) = w^T \phi(x) + b \quad (2)$$

式中, w 为权向量, b 为偏移量。根据结构风险最小化原理转化成等式约束:

$$\min_{w, b, e} J(w, e) = \frac{1}{2} w^T w + \frac{C}{2} \sum_{i=1}^N e_i^2 \quad (3)$$

$$\text{s. t. } y_i = w^T \phi(x_i) + b + e_i \quad (i = 1, \dots, N) \quad (4)$$

式中, C 为惩罚系数, e_i 为松弛变量。采用 Lagrange 方法进一步转换为无约束优化问题:

$$L(w, b, e, \alpha) = J(w, e) - \sum_{i=1}^N \alpha_i (w^T \phi(x_i) + b + e_i - y_i) \quad (5)$$

式中 $\alpha_i (i = 1, \dots, N)$ 为 Lagrange 乘子。

对(5)式中变量求偏导,解得 α, b , 得

$$y(x) = \sum_{i=1}^N \alpha_i k(x, x_i) + b \quad (6)$$

$k(x_k, x_i)$ 称为核函数,本文采用 RBF 核函数:

$$k(x_k, x_i) = \exp(-\|x_k - x_i\|^2 / \sigma^2) \quad (7)$$

由于 LSSVM 假定每个自变量对因变量的贡献相同,会导致核函数的计算可能会被一些影响程度弱的自变量所支配,从而影响 LSSVM 的预测或分类性能。

根据自变量 $x_i (i = 1, 2, \dots, N)$ 的重要程度赋予不同的特征权重 $w_i (i = 1, 2, \dots, N)$, 令

$$W = \begin{bmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_N \end{bmatrix}$$

N 阶对角矩阵 W 定义为特征重要度矩阵,特征加权的 RBF 核函数变形为

$$k_w(x_k, x_i) = \exp(-\|x_k^T W - x_i^T W\|^2 / \sigma^2) \quad (8)$$

由机理分析可知,影响砷盐除钴效果的主要有:1 号反应器 ORP 值 (x_1)、除铜后液的流量 (x_2 , 流量越大,停留时间越短)、溶液中 Co^{2+} 浓度 (x_3)、底流的返回流量 (x_4)、底流中固体的含量 (x_5)、 As_2O_3 的添加量 (x_6)、溶液 BT 值 (x_7)、除铜后液温度 (x_8)、溶液中 Cu^{2+} 浓度 (x_9)。

显然,各影响因素对砷盐净化除钴的影响程度是不一样的,在利用 LSSVM 进行预测建模时,如果对各因素赋予相同权重,会导致 LSSVM 的核函数计算被除铜后液铜离子浓度等影响相对较弱的因素支配,会降低 LSSVM 的预测性能。根据 9 个影响因素 $x_1, \dots, x_9$ 对砷盐净化除钴影响程度,赋予不同的特征权重 $w_1, \dots, w_9$, 分别为 0.18, 0.15, 0.13, 0.12, 0.12, 0.10, 0.08, 0.07, 0.05。

由于 LSSVM 的惩罚参数 C 和核函数参数 σ 对预测能力影响较大,本文采用动态分级微粒群算法优化选择两个参数 C 和 σ , 减少主观经验选择参数的盲目性,提高模型的预测能力。

2.3 动态分级微粒群算法

标准微粒群算法如下公式所示^[15]:

$$v_{im}(k+1) = \omega(k)v_{im}(k) + c_1 r_1 (pbest_{im} - x_{im}(k)) + c_2 r_2 (gbest_m - x_{im}(k)) \quad (9)$$

$$x_{im}(k+1) = x_{im}(k) + v_{im}(k+1) \quad (10)$$

式中, c_1 和 c_2 分别为自我学习因子和社会学习因子,一般都取常数 2。 r_1 和 r_2 为 $[0, 1]$ 之间的随机数。 $pbest_{im}$ 和 $gbest_m$ 分别是第 i 个微粒和全部微粒第 m 维经过的最好位置分量。

标准微粒群算法收敛速度较慢,且在寻优后期随着种群多样性迅速降低,极易陷入局部最优解^[16],难以寻优得到 LSSVM 较好的一组参数。

动态分级微粒群算法基本思想是:将种群 ($N = N1 + N2$) 分为两组,一组用于全局搜索 ($N1$), 一组用于局部搜索 ($N2$), 并在寻优过程中进行信息交换,并根据进化代数动态调整 $N1$ 和 $N2$ 。

动态分级的微粒群算法具体步骤如下:

步骤 1: 设置算法参数,采用混沌方法产生初始种群。

步骤 2: 计算各粒子的适应度值,按适应度值从小到大分为两组,第一组 $N1$ 个粒子,第二组 $N2$ 个粒子。随着进化代数增加, $N1$ 减小(减小到 $N2$), $N2$ 增加(增加到 $N1$); 即算法进化的前期偏重全局寻优,后期偏重局部寻优。

步骤 3: 第一组进行局部搜索:先按式(9)和式(10)更新位置和速率,然后将该组各粒子与群体全局最优值 $gbest$ 以概率 p_c 进行交叉操作,使得粒子在最优值附近进行局部搜索;然后更新该组 $pbest1$ 和 $gbest1$ 。

第二组进行全局搜索:寻优后期,较优的粒子具有趋同性,即种群多样性降低,极易陷入局部最优。先按式(9)和式(10)更新位置和速率,然后将该组

粒子以概率 p_m 进行变异操作,使算法能有效地跳出局部最优解。更新该组 $pbest2$ 和 $gbest2$ 。

步骤 4:比较 $gbest1$ 和 $gbest2$, 将最好的做为群体的全局最优值 $gbest$ 输出。

步骤 5:判断算法是否结束,若结束则输出最优值,否则转步骤 2。

2.4 灰色特征加权 LSSVM 的建立

由于砷盐净化除钴过程具有溶液腐蚀性强、现场噪音大和振动大等特点,生产数据不可避免地含有大量噪声,因而削弱或消除对 LSSVM 的预测性能有影响的噪声非常重要。

灰色预测方法中对原始数据序列进行“累加生成”处理,能削弱原始数据序列中噪声的影响。本文将基于 GM(1,1) 灰色模型思想和特征加权 LSSVM 预测模型结合起来,如图 9 所示。

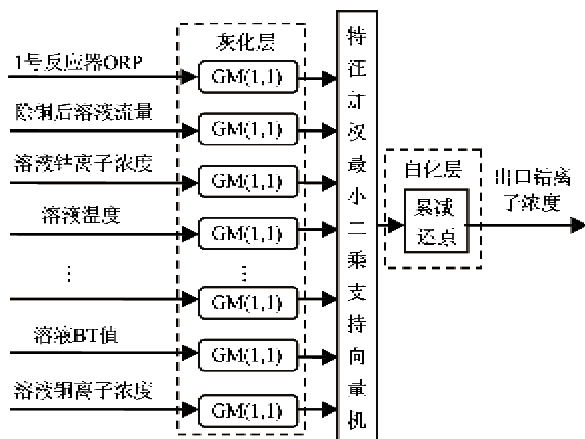


图 9 灰色特征加权 LSSVM 预测模型

灰色特征加权 LSSVM 组合预测模型建立步骤如下：

步骤 1: 生成原始序列 $X_m^{(0)} (m = 1, 2, \dots, 9)$;
 $X_m^{(0)} = (x_m^{(0)}(1), x_m^{(0)}(2), \dots, x_m^{(0)}(N)) (N$ 为样本的大小);
 $Y^{(0)}: Y^{(0)} = (y^{(0)}(1), y^{(0)}(2), y^{(0)}(N))$, 对数据进行归一化处理后再采用灰色累加, 得到累加生成序列 (1-AGO), 记为 $X_m^{(1)} = (x_m^{(1)}(1), x_m^{(1)}(2), \dots, x_m^{(1)}(N))$, 其中 $x_m^{(1)}(k) = \sum_{i=1}^k x_m^{(0)}(i), k = 1, 2, \dots, N; Y^{(1)} = (y^{(1)}(1), y^{(1)}(2), \dots, y^{(1)}(N)), y_m^{(1)}(k) = \sum_{i=1}^k y_m^{(0)}(i), k = 1, 2, \dots, N$ 。

步骤 2: 将灰色累加后的数据做为新样本集, 即得特征加权 LSSVM 的样本集为 $Tg = (x_m^{(1)}(k), y^{(1)}(k))$ (其中: $x_m^{(1)}(k)$ 为 9 维输入, $y^{(1)}(k)$ 为一

维输出, $k = 1, 2, \dots, N$)。

步骤 3: 通过累减还原即得原数据序列的预测模型。

本文采用动态分级粒子群算法来优化选择对 LSSVM 模型的学习和泛化能力影响很大的两个参数。

3 模型验证与分析

为验证灰色特征加权 LSSVM 模型的有效性, 采用 2011 年 6 月某锌冶炼企业砷盐除钴过程实际数据 156 组, 其中 106 组做训练样本集, 50 组做测试样本集。

本文采用动态分级 PSO 算法寻优 LSSVM 两个参数 σ 和 C , 为避免搜索的盲目性, 应该设定参数的寻优范围^[17]; 核参数 σ 的寻优范围设定为 $[0.001, 50]$, 惩罚参数 C 的寻优范围为 $[0.001, 500]$ 。

动态分级的 PSO 算法参数设置对算法的性能有较大的影响; 惯性权重 ω 取 0.9, 微粒数取 30 (对于待寻优参数较少的优化问题, 取 30 个微粒可以得到较好的结果)^[18]; 初始 $N1 = 20$ (寻优初期, 注重全局搜索, 故 $N1$ 初值一般为 $N2$ 的 2 倍), $N2 = 10$ 。为产生足够多的新个体, 加速算法的收敛, 交叉概率 p_c 取 0.7; 为使算法有效的跳出局部最优解, 又不至于退化到随机搜索, 变异概率 p_m 取 0.2; 最大迭代次数 $ITER_{max} = 50$ (对于两个变量的寻优, 进化到 50 代可以取得较好的效果)。寻优得灰色特征加权 LSSVM 两个参数 σ 和 C 分别为 5.7 和 181.8。

图 10、图 11 和图 12 为 3 种预测模型预测结果及相对误差绝对值 Ea 的仿真结果图, 表 1 为 3 种预

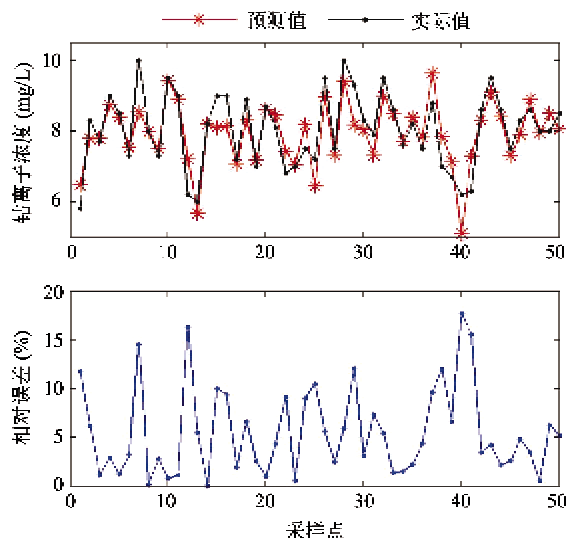


图 10 灰色加权 LSSVM 仿真结果

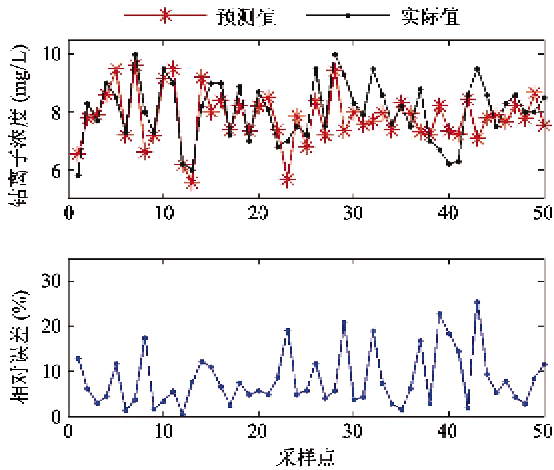


图 11 LSSVM 仿真结果

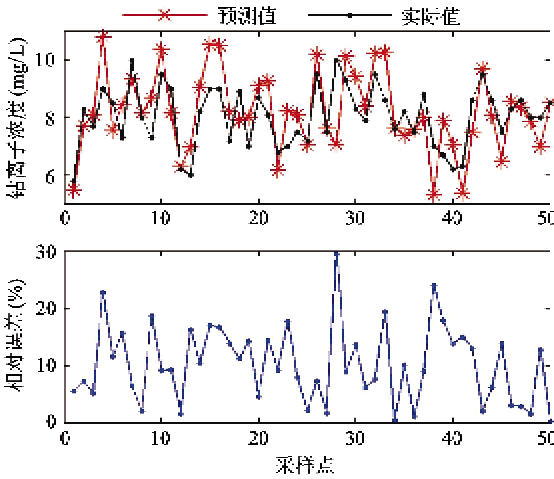


图 12 BP 神经网络仿真结果

表 1 三种预测模型结果比较

预测模型	Ea 平均值 (%)	Ea_{max} (%)
灰色特征加权 LSSVM	5.53	17.72
LSSVM	8.25	25.68
BP 神经网络	10.18	29.51

测模型误差的统计结果。

由图 10、图 11 和图 12 及表 1 可知,BP 神经网络的预测精度较低,且预测值不能跟踪实际值的变化趋势,实际应用可能会因为判断失误而造成操作不合理。LSSVM 的预测精度和跟踪性能得到了一定的提高,但由于砷盐净化除钴生产数据存在噪声,且各影响因素的重要程度不一样,LSSVM 的预测效果难以满足砷盐净化除钴过程钴离子浓度在线预测的要求。灰色特征加权 LSSVM 预测模型的预测值较好地反应了实际值的变化趋势,能满足钴离子浓度在线预测的要求。

4 结 论

砷盐除钴过程环境恶劣,生产数据存在大量噪声,利用含有大量噪声的数据建模会严重恶化 LSSVM 的预测性能,采用灰色累加的思想,可显著削弱原始数据中的噪声。影响砷盐净化除钴的因素很多,但这些因素的重要程度并不完全一样,如果将其影响看成是相同的,会导致核函数的计算可能会被一些重要程度低的自变量所支配,因而需引入特征重要度矩阵,并利用动态分级微粒群算法对特征加权 LSSVM 的参数进行寻优,得到使特征加权 LSSVM 预测精度最高的一组最优参数 (σ , C)。对砷盐净化除钴过程中钴离子浓度的预测结果表明,本文提出的灰色特征加权 LSSVM 预测模型能适应数据含有噪声、影响因素多且影响程度差别比较大的砷盐净化除钴过程的预测建模。该方法对其它类似生产过程具有一定借鉴意义。

参考文献

[1] Bockman O, Østvold T. Products formed during cobalt cementation on zinc in zinc sulfate electrolytes. *Hydrometallurgy*, 2000,54(2):65-78

[2] 晏密英,桂卫华,王凌云. 基于神经网络补偿灰色预测误差的钴离子浓度预测研究. *计算机与应用化学*, 2008,25(7): 805-808

[3] Suykens J, Vandewalle J. Least squares support vector machine Classifiers. *Neural Processing Letters*, 1999,19(3): 293-300

[4] Esen H, Ozgen F, Esen M. Modeling of a new solar air heater through least squares support vector machine. *Expert Systems with Applications*, 2009, 36(7): 10673-10682

[5] 周开军,阳春华,牟学民等. 基于泡沫特征与 LSSVM 的浮选回收率预测. *仪器仪表学报*, 2009,30(6): 1295-1300

[6] Zhao J, Zhang X P, Wang W. Gradient optimized LSSVM for prediction of gas consumption in steel industry. *ICIC Express Letters*, 2011,4(6):2069-2073

[7] Zhen C, Yi H. Modeling and simulation of the GA-LSSVM-based soft measurement for torque. *Applied Mechanics and Materials*, 2010,45(1):733-737

[8] 刘梅花,桂卫华,阳春华等. 基于 WA 和 LS-SVM 的净化过程钴离子浓度预测. *计算机测量与控制*, 2009,17(4):652-655

[9] 朱红求,阳春华,桂卫华. 基于模糊聚类和模糊支持向

- 量机的湿法炼锌净化除钴过程建模. 高技术通讯, 2010, 20(10):1068-1071
- [10] 汪建华, 田盛丰, 黄厚宽. 特征加权支持向量机. 电子与信息学报, 2009, 31(3):515-519
- [11] Chappelle O, Vapnik V, Bousquet O. Choosing multiple parameters for support vector machines. *Machine Learning*, 2002, 46(1):131-160
- [12] Jari N. Statistical analysis of cobalt removal from zinc electrolyte using the arsenic-activated process. *Hydrometallurgy*, 2004, 73(1):123-132
- [13] Sakiohi G. Purification of zinc leaching solution: Mechanism of removal of cobalt by zinc dust with arsenious oxide and copper. *Metallurgical review of MMJJ*, 1997, 14(1):37-52
- [14] Jacob C, Cohen P. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (Third Edition). Lawrence Erlbaum Associates, New Jersey, American, 2003. 69-71
- [15] Kannana S, Mary S, Slochanal R, et al. Application of particle swarm optimization technique and its variants to generation expansion planning problem. *Electric Power Systems Research*, 2004, 70(8): 203-210
- [16] 杨智, 陈志堂, 范正平等. 基于改进粒子群优化算法的 PID 控制器整定. 控制理论与应用, 2010, 27(10):1345-1352
- [17] 朱红求, 阳春华, 王觉等. 基于混合 QPSO 的 LS-SVM 参数优化及其应用. 中南大学学报(自然科学版), 2011, 42(4):1000-1003
- [18] 龚纯, 王正林著. 精通 MATLAB 最优化计算. 北京, 2009. 270-298

Application of grey feature weighted LSSVM to removal of cobalt with arsenic

Wu Tiebin^{***}, Li Yonggang^{*}, Yang Chunhua, Sun Bei^{*}, Gui Weihua^{*}

(^{*} School of Information Science & Engineering, Central South University, Changsha 410083)

(^{**} Department of Communications & Control Engineering,

Hunan Institute of Humanities Science and Technology, Loudi 417000)

Abstract

To solve the problem in removal of cobalt with arsenic that the cobalt ion concentration prediction accuracy is always low because of the data noise and multiple influence factors, a novel grey feature weighted least squares support vector machine (LSSVM) model for cobalt ion concentration prediction is proposed. The model uses the grey accumulation method to weaken the influences of the noise in primary data, and based on the investigation of the main influence factors to the prediction of the cobalt ion concentration, sets feature weights to every influence factors according to their importance to improve the predictive accuracy. The two parameters of LSSVM model, the penalty factor C and the core parameter σ , are optimized by the dynamic hierarchical particle swarm optimization algorithm which has the abilities of fast convergence and global optimization, so that the blindness in choice of model parameters can be avoided. The simulation results show that the grey feature weighted LSSVM model has the higher prediction accuracy, and its predicting values can well follow the changing trend of the real cobalt ion concentration. The proposed model satisfies the requirements of on-line prediction of the cobalt ion concentration in the cobalt removal process with arsenic.

Key words: feature-weighted, least squares support vector machine (LSSVM), cobalt ion, particle swarm optimization, grey accumulation