

# 基于粒子群支持向量回归优化的 循环流化床床温预测分析

黄纯颖, 曾庆敏, 陈玲红, 吴学成, 岑可法

(浙江大学 能源清洁利用国家重点实验室, 浙江 杭州 310027)

**摘要:** 为了优化循环流化床锅炉(CFB)的床温系统控制,使用支持向量回归(SVR)人工智能方法进行建模预测;为提高预测结果的精准性,引入互信息法则(MI)及主成分分析法(PCA)作为输入特征选择手段,同时使用粒子群算法(PSO)进行参数寻优。研究结果表明,MI法则进行数据预处理并利用PSO算法优化后的SVR模型能够精确地预测在不同运行工况下的床温变化,且预测误差较小,拥有较强的泛化能力。

**关键词:** 循环流化床; 床温; 支持向量回归; 粒子群优化; 特征选择

中图分类号: X701.3

文献标识码: A

文章编号: 1004-3950(2022)03-0011-07

DOI:10.16189/j.cnki.nygc.2022.03.003

## Prediction and analysis of bed temperature on circulating fluidized bed based on particle swarm support vector regression optimization

HUANG Chun-ying, ZENG Qing-min, CHEN Ling-hong, WU Xue-cheng, CEN Ke-fa

(State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, China)

**Abstract:** In order to optimize the bed temperature control of circulating fluidized boiler(CFB), support vector regression(SVR) artificial intelligence method is used for modeling and prediction. To improve the accuracy of the prediction results, while using mutual information law(MI) and principal component analysis(PCA) as input feature selection methods, particle swarm optimization(PSO) is used for parameter optimization. The research result shows that after data preprocessing by the MI law and optimized by the PSO algorithm, the SVR model can accurately predict the bed temperature changes under different operating conditions, and the prediction error is little with strong generalization ability.

**Key words:** circulating fluidized bed; bed temperature; support vector regression; particle swarm optimization; feature selection

## 0 引言

循环流化床(circulating fluidized bed, CFB)具有负荷调节快、脱硫效率高、污染排放低等优点,近年来一直是能源清洁利用的重要技术<sup>[1]</sup>。CFB的最大特点是整个燃烧过程中的流动状态,在燃烧期间气体作为流动介质,许多固体燃料和床料在炉中燃烧从而产生许多小颗粒的飞粉,循环灰分离器将飞出的灰粉通过管道送回炉中再次燃烧以此保证燃烧效率。此外炉中反复进行低温燃烧和等级燃烧,最终提高了脱硫效率<sup>[2]</sup>。因此,大

力发展CFB燃烧技术,研究相应的优化控制方式也是一大趋势。

CFB锅炉与常规煤粉炉另一大差别是其独特的控制系统,特别是锅炉床温的控制。CFB的床温水平对于锅炉燃烧安全性与经济性会产生极大影响。除此之外,考虑脱硫效果的最佳床温运行区间在850~900℃,温度过低,石灰石煅烧速度显著下降;温度过高,脱硫产物CaSO<sub>4</sub>会在还原性气氛中被大量分解<sup>[2-4]</sup>。CFB的床温运行系统具有非线性、参数时变和多变量耦合等特性,故极难通过简单机理建模来进行预测控制研究<sup>[5]</sup>。

收稿日期: 2021-12-13

作者简介: 黄纯颖(1997-),女,福建泉州人,硕士研究生,主要研究方向为循环流化床燃烧系统建模仿真。

通讯作者: 陈玲红,女,浙江绍兴人,教授, E-mail: chenlh@zju.edu.cn

对于 CFB 锅炉这样复杂的燃烧系统的控制,重要的是建立一个可以准确反映其燃烧系统对象特性的模型。目前床温特性建模主要有以下几种方法:(1)通过大量的现场试验收集数据,根据试验数据建立回归关系,这种方式需要花费一定的人力、物力成本,试验工况的种类也会有局限性<sup>[6]</sup>;(2)根据 CFB 锅炉系统物理或化学变化规律进行机理建模,这是最原始的床温建模法,可以从原理角度反映床温变化趋势,但机理模型往往存在需要合理的简化假设,无法保证在各个工况下的预测精度<sup>[7]</sup>;(3)利用计算流体力学、传热学等机理模拟炉内燃烧过程,求得炉内床温分布情况,这种方法从精确度来说较第二种已有了很大的提高,且被证实有良好的效果<sup>[8]</sup>,但由于其涉及到复杂的机理建模和计算步长,需要高端的计算机配置及较长的计算时间才能够完成建模的全过程,经济性不高<sup>[9]</sup>。此外,由于 CFB 锅炉运行时有入炉组分复杂、多边形强等特点,给 CFB 的传统建模方式带来了困难,这就需要 CFB 床温特性建模需要有较好的自适应能力,而这项能力是以上三种方式所欠缺的。

随着科学技术及人工智能的发展,电厂的分散控制系统也日益完善,基于神经网络(artificial neural networks,ANN)及支持向量回归(support vector regression,SVR)对电厂燃烧特性进行预测建模的方法得到广泛的关注及研究。基于电厂大数据进行建模时,SVR 应用了结构性风险小化(structure risk minimization,SRM)原则,在非线性和高阶模式识别的解决中展示了其独特的优点,并且与 ANN 相比泛化能力更强<sup>[10,11]</sup>。同时,SVR 引入了可以将输入空间映射到相应的特征空间的核函数,从而对样本数据在高维特征空间建立非线性模型,即可以使用核函数将影响床温变化的因素数据向高维空间映射,提高预测模型的精度。廖伟等指出,支持向量回归不再使用传统的经验风险最小化原则,同时最小化学习算法的失灵敏度损失系数  $\epsilon$  是 Min-Max 理论中均值绝对误差准则的推广,该理论具有较强的能力<sup>[12]</sup>。

本文研究对象为某热电厂 130 t/h CFB 锅炉,整体床温模型设计如图 1 所示。首先,从控制运行的角度出发,确定该锅炉的床温影响因素,基于该电厂分散控制系统整合而得的 CFB 锅炉数据样本,对样本数据进行清洗后,再进行数据预处理

及归一化操作,接着利用不同的特征选择方法对输入变量进行处理,以此来提高模型预测准确性,同时使用粒子群算法(particle swarm optimization,PSO)对模型参数进行寻优,进一步提高预测准确性及模型适用性,在此基础上利用 SVR 对该 CFB 锅炉床温进行回归预测,从而实现对该电厂床温的预测控制研究。

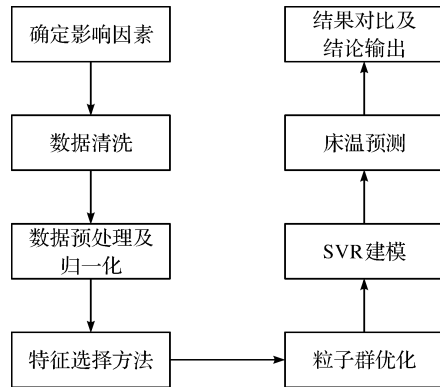


图 1 床温模型设计流程图

## 1 锅炉概况

本研究 130 t/h CFB 锅炉为单锅筒、自然循环、集中下降管、II 型布置、单级高温分离循环系统的循环流化床锅炉,采用半露天布置。锅炉由一个膜式水冷壁炉膛、两个汽冷式旋风分离器和一个尾部竖井烟道及其上的蒸汽包组成。运行时过热蒸汽采用锅炉给水喷水减温。

其炉膛结构与物料入炉位置如图 2 所示,两个给煤口与两个石灰石口均位于炉膛前墙。各床温测点如图 3 所示,前后墙分别有 4 个测温点,测点均在炉膛下方密相区。

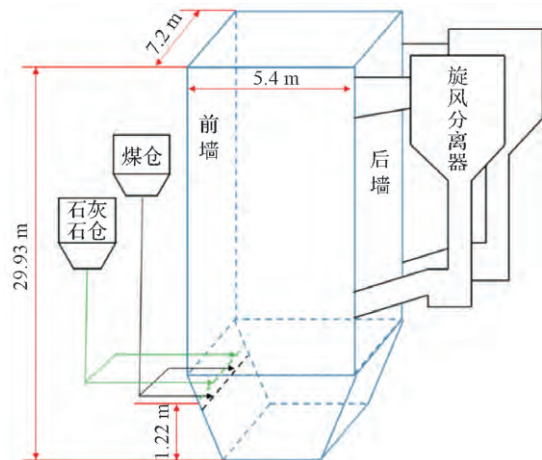


图 2 CFB 锅炉结构图

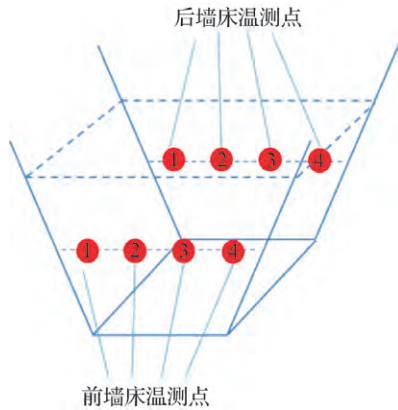


图3 CFB 锅炉密相区床温测点示意图

## 2 建模原理

### 2.1 支持向量回归原理

SVR 是在支持向量机 (support vector machine, SVM) 的基础上发展而来的, 最开始 SVM 主要被应用于分类问题, 后来 SVR 被逐渐应用在回归估计中。

SVR 模型建立的基本思想: 通过一个非线性映射将训练集和样本集的输入参数映射到高维特征空间, 在此特征空间中利用以下公式进行线性建模:

$$f(x) = [w \cdot \varphi(x)] + b \quad (1)$$

$\varphi: R^m \rightarrow F \quad w \in F$

其中  $b$  为阈值。在式 (1) 的基础上, SVM 通过对式 (2) 极小化来取得回归函数:

$$\min \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \right\} \quad (2)$$

约束条件为:

$$St: \begin{cases} y_i - w \cdot x_i - b \leq \varepsilon + \xi_i \\ w \cdot x_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \quad i = 1, \dots, l \end{cases} \quad (3)$$

其中, 常数  $C > 0$ , 是一种对于超出误差的样本的惩罚程度的表征手段;  $\xi_i$  和  $\xi_i^*$  为松弛因子;  $\varepsilon$  为不敏感损失函数。以上的优化问题可转化成 Lagrange 对偶问题:

$$\max \left[ -\frac{1}{2} \sum_{i,j=1}^l (a_i - a_i^*) (a_j - a_j^*) K(X_i, X_j) + \sum_{i=1}^l y_i (a_i - a_i^*) - \varepsilon \sum_{i=1}^l (a_i + a_i^*) \right] \quad (4)$$

约束条件为:

$$St: \begin{cases} \sum_{i=1}^l (a_i - a_i^*) = 0 \\ 0 \leq a_i, a_i^* \leq C \quad i = 0, \dots, l \end{cases} \quad (5)$$

通过对式 (4) (5) 对偶问题的求解, 得到回归问题的决策函数:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \quad (6)$$

式中:  $K(x_i, x_j)$  为核函数, 即一种从低维空间到高维空间的映射方法。不同的核函数会使得模型处理结果有所不同。根据经验, 本文选择高斯径向基核函数进行研究:

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

式中:  $\sigma$  为宽度参数, 对径向作用的范围进行控制。

### 2.2 粒子群优化算法原理

PSO 优化算法是一种进化计算技术, 它的特点是简单、快速及高效, 在人工智能及深度学习方面有极大的应用潜力。PSO 与遗传算法类似, 是一种以群体为基础的优化工具, 不同的是 PSO 不需要进行交叉和变异的操作, 而是粒子在解空间自行迭代搜寻最优解。

PSO 的算法流程如下。

#### (1) 初始化

首先, 设置最大迭代次数、粒子最大速度、目标函数自变量个数及搜索空间的位置信息, 之后随即初始化速度、位置及粒子群规模。

#### (2) 搜寻全局最优解

从每个粒子找到的最优解, 即个体极值中, 搜寻到一个全局最优解, 并与历史全局最优比较后进行更新。

#### (3) 更新速度和位置公式

$$V_{id} = \omega V_{id} + C_1 \text{random}(0, 1) (P_{id} - X_{id}) + C_2 \text{random}(0, 1) (P_{gd} - X_{id}) \quad (8)$$

$$X_{id} = X_{id} + V_{id} \quad (9)$$

式中:  $\omega$  为惯性因子, 其大小决定了 PSO 算法的寻优性能;  $C_1$  和  $C_2$  为加速常数, 一般取  $C_1 = C_2 \in [0, 4]$ ;  $X_{id}$  为第  $i$  个变量在第  $d$  维的当前位置;  $V_{id}$  是第  $i$  个变量在第  $d$  维的速度;  $P_{id}$  为第  $i$  个变量在第  $d$  维的个体极值点的位置;  $P_{gd}$  为整个种群在第  $d$  维的全局极值点的位置。

(4) 终止条件

当达到最大迭代数或相邻两代偏差位于指定范围内时 搜寻停止。

寻优过程可简化为如图 4 所示:

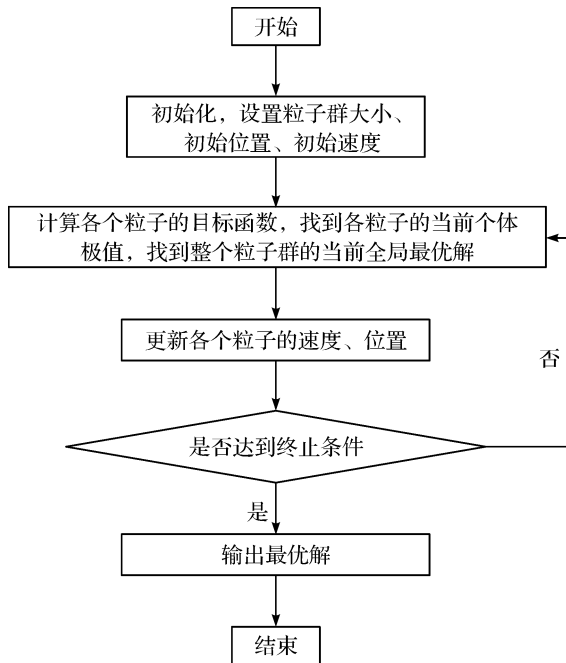


图 4 粒子群寻优路径

### 3 模型建立

#### 3.1 变量选择

CFB 锅炉床温是由多变量耦合作用影响的对象,有研究表明,CFB 床温的影响因素可总结为运行负荷、一次风量、二次风量、给煤等<sup>[13]</sup>,锅炉在正常工况下运行时,通过对以上因素的调整可以进行床温控制,本研究在以上影响因素的基础上,加入引风机电流、石灰石入炉频率等可调节参数及其他影响参数,拓宽变量选择范围,以某热电厂 130 t/h CFB 锅炉为研究对象,使用表 1 中所列参数作为因变量,利用 DCS 系统数据对于床温变化水平进行实时建模。

#### 3.2 数据预处理——归一化

选取某热电厂 DCS 系统中 2020 年 8 月 1 日 0:00 至 2020 年 10 月 12 日 24:00 共 2348 条数据,首先对异常值、空值等进行清洗处理,经过数据清洗操作后筛选出表 2 中的 2166 条数据进行进一步研究。

表 1 变量选择清单

参数	单位
瞬时给煤量	t/h
主蒸汽流量	
给水流量	
风室压力	kPa
一次风量	km <sup>3</sup> /h
二次风量	
石灰石给料频率	Hz
一次风机电流	A
二次风机电流	
引风机电流	
循环风机电流	
返料风机电流	
炉前石灰石风机电流	
烟气氧含量	%

表 2 热电厂 DCS 数据

序号	一次风机电流	一次风量	风室压力	……	给煤量
1	24.81	39541.28	5.45	……	13.57
2	24.33	38557.38	5.53	……	13.16
3	23.81	37510.06	6	……	13.28
4	23.53	36637.14	6.25	……	13.97
……	……	……	……	……	……
2163	27.08	37434.15	7.53	……	14.36
2164	27.21	36411.45	7.57	……	13.84
2165	26.82	37511.89	7.53	……	14.52
2166	27.18	38099.58	7.51	……	13.97

在进行特征选择及建模之前,为了避免参数间量纲以及数量级的不同造成模型性能的降低,接着对筛选清洗完毕的热电厂月运行 DCS 数据执行归一化操作。通过归一化操作可以把数据映射到 [0, 1] 之间,从而消除各大特征的量纲对最终模拟结果的影响。归一化变化的函数为:

$$f: x \rightarrow y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (10)$$

其中  $x, y \in R^n$ ,  $R^n$  为实向量空间。

#### 3.3 特征选择

数据和特征决定了机器学习的上限,而模型和算法只是逼近这个上限。好的特征选择能够提升模型的性能,更能帮助理解数据的特点、底层结构,这对进一步改善模型、算法都有着重要作用<sup>[14]</sup>。故在本文分析中,选择合适的特征来帮助建模是有必要的,这有助于提高模型预测的准确性,并为后续优化研究工作的进一步开展提供参考。



### 3.3.1 互信息法

互信息 (mutual information, MI) 数值结果表示两个变量 (如  $X$  与  $Y$ ) 是否具有变化关系, 及其关系的强弱, 其常用于非线性建模中的变量筛选。如果  $(X, Y) \sim p(x, y)$ ,  $X, Y$  之间的互信息  $I(X; Y)$  定义为:

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (11)$$

利用互信息法对特征进行处理后, 各特征排序如图 5 所示, 根据表 3 互信息值与相关程度的关系, 互信息值越接近 1, 两个变量间的关联性越强。故为了保证模型效果及预测准确性, 在以上可调节 CFB 锅炉参数中根据互信息值大小进行筛选, 将关联性较高, 即选择大于 0.2 的共 11 个特征进行进一步研究。

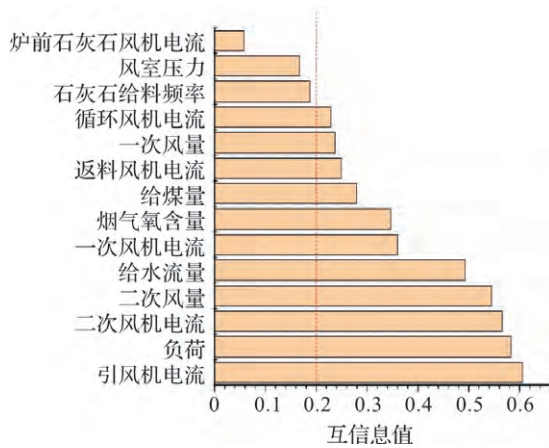


图5 各参数互信息值

表3 互信息值与相关程度关系<sup>[15]</sup>

互信息值的绝对值	相关程度
$0.00 \leq  r  < 0.10$	不相关
$0.10 \leq  r  < 0.20$	极低相关
$0.20 \leq  r  < 0.40$	低度相关
$0.40 \leq  r  < 0.70$	中度相关
$0.70 \leq  r  < 0.90$	高度相关
$0.90 \leq  r  \leq 1.00$	极高相关

### 3.3.2 主成分分析法

主成分分析 (principal component analysis, PCA) 是一种常用的数据分析方法, 可以识别影响过程监控参数变化的主要因素或大量过程变量中指标的下降, 通常可以用来降低高维数据空间的维数。另外, PCA 是基于原始数据空间的, 它通过构造一组新的变量来降低原始数据空间的维数, 然后

从新的映射空间中提取主要的变化信息, 提取统计特征, 从而构造出原始数据空间特征的新解<sup>[16]</sup>。

对 14 个参数变量进行 PCA 分析, 选取累计贡献率为 80% 以上的成分作为主成分, 最终得到四个主成分结果 (见图 6)。

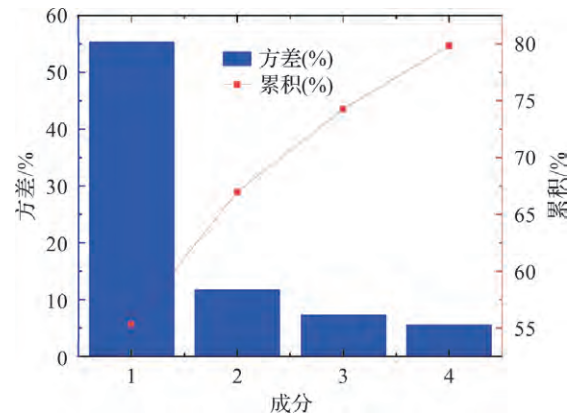


图6 主成分分析结果

### 3.4 参数优化

将源数据随机划分为训练集与测试集, 其中训练集占 75%, 测试集占 25%。训练最终模型前, 为了使模型达到最佳性能, 需要确定 SVR 算法中的惩罚系数  $C$  和径向核函数  $g$ , 针对以上两种特征选择方法, 本文使用 PSO 算法进行优化参数选择。

使用 matlab 软件, 在一定范围内基于 PSO 方法对  $C$  和  $g$  参数进行取值。将训练集作为原始数据进行关于  $C$  和  $g$  参数的计算, 选择准确率最高的一组作为最终的最佳参数。当存在多组准确率高的  $C$  和  $g$  参数的组合时, 则选择  $C$  值最小的参数组合, 若存在多组  $C$  值最小组, 则选取搜索到的第一组  $C$  和  $g$  参数, 通过这样的参数选取方式, 可以避免机器学习过程中产生过学习的状态, 从而降低模型的泛化能力<sup>[17]</sup>。

基于上述参数选取规则, 进行 PSO 算法寻优过程后, 分别得到不同特征选取方式下的适应度变化趋势结果如图 7 和图 8 所示, 对于 MI 法则得到的最优参数组合为:  $C = 4.7682$ ,  $g = 1.4134$ ; 对于 PCA 法则得到的最优参数组合为:  $C = 0.1000$ ,  $g = 6.0163$ 。故对于不同的特征选择方式来说, 寻优得到了不同的组合结果, 为避免机器学习机制的过学习状态, 在寻优过程中通常选取较小的惩罚参数  $C$  值<sup>[18]</sup>。

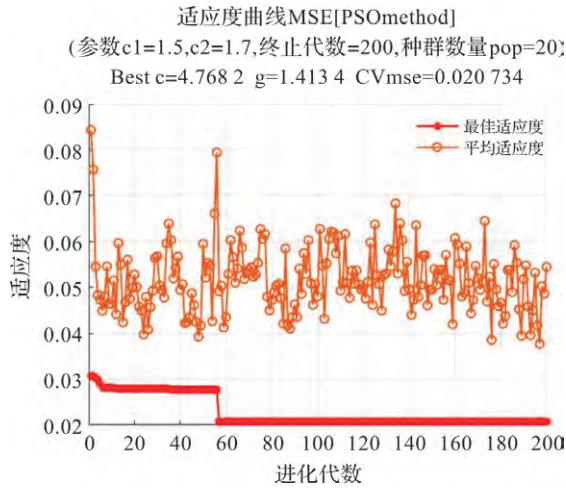


图7 互信息法寻优结果

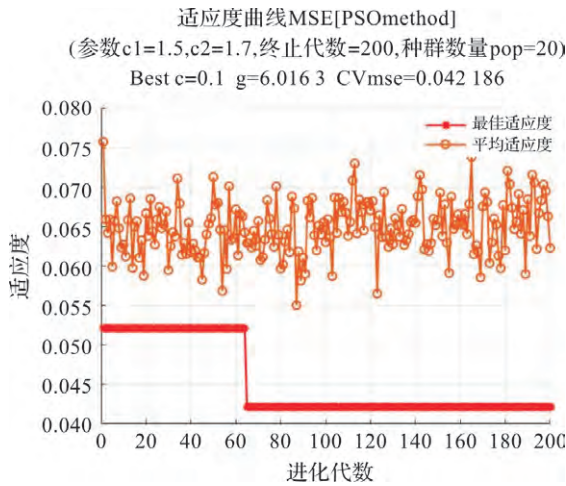


图8 主成分分析寻优结果

### 3.5 模型仿真结果

在对比仿真结果之前,引入指标评价体系评估模型效果:

$$\text{均方根误差 RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (12)$$

相关系数

$$R = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (13)$$

平均绝对百分比误差

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (14)$$

其中  $\hat{y}_i$  为模型预测值  $y_i$  为实际值。

在进行特征选择及粒子群参数寻优后,最终模型仿真结果如图9-图11所示。由仿真图形

可以得出,经过互信息化特征选择的 PSO-SVR 模型拟合结果与实际床温结果更为接近,即该 MI-PSO-SVR 仿真模型能够对该锅炉床温变化情况做出准确预测。由表5的模型评价指标可知,MI-PSO-SVR 的模型均方根误差仅 4.37,相关系数平方高达 0.9189,绝对百分比误差低至 0.38%。相反,经过 PCA 分析处理过后的模型呈现出较差的拟合水平,仿真结果的均方根误差高达 11.52,相关系数平方仅为 -0.2219,MAPE 高达 1.04%,可被认为是劣质模型。

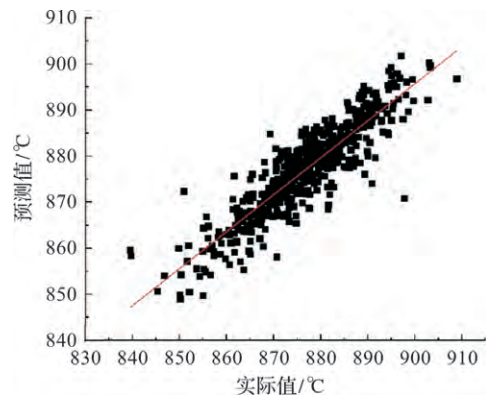


图9 SVR 仿真结果

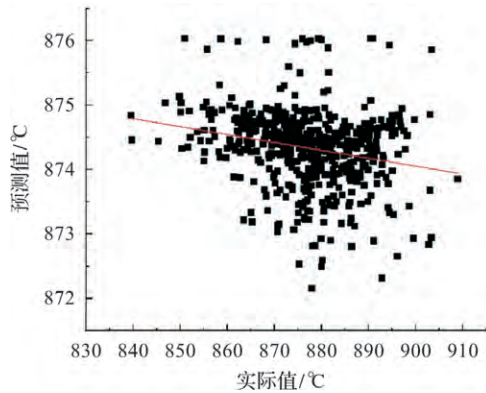


图10 PCA-PSO-SVR 仿真结果

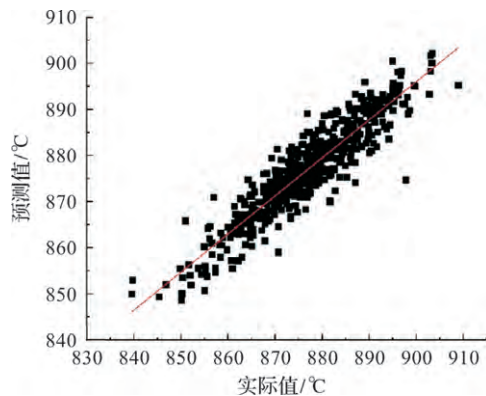


图11 MI-PSO-SVR 仿真结果

表5 模拟仿真结果汇总表

方法	$C$	$g$	$R$	RMSE	MAPE
SVR	默认		0.8979	4.87	0.41%
PCA-PSO-SVR	0.1000	6.0163	-0.2219	11.52	1.04%
MI-PSO-SVR	4.7682	1.4134	0.9189	4.37	0.38%

## 4 结 论

对于由相互耦合的多变量控制影响的 CFB 锅炉燃烧系统,为完成精准建模预测及控制,本文研究了某热电厂 130 t/h CFB 锅炉床温运行状况,结合其实际运行数据,在基于研究经验及机理原理所选择的床温影响因素的基础上,加入了更多可调节参数,最终完成了该 CFB 锅炉运行床温预测模型的建立,为锅炉床温优化控制奠定了基础。

本文对比了不同特征选择方法下的建模精准性,结果表明,使用 MI 法则处理可以有效提高模型预测结果准确性,这也表明,在利用优化算法提高预测效率同时,输入参数的选择也十分重要。参数选取的合理性有助于减少欠学习或过学习的现象,从而帮助预测精度获得提升。然而,被广泛使用的 PCA 法则不适用于本文的 SVR 建模过程,这是由于该特征选择方法在执行过程中令原 DCS 数据集结构产生丢失所致。

通过本文研究得到,MI 法则进行数据预处理并利用 PSO 算法优化后的 SVR 模型能够精确地预测在不同运行工况下的床温,且预测误差较小,拥有较强的泛化能力。

### 参考文献:

[1] 沈雅丽. 基于广义简单解耦的多变量系统自抗扰分散控制[D]. 北京: 华北电力大学, 2016.

[2] 岑可法. 循环流化床锅炉理论与运行[M]. 北京: 中国电力出版社, 1998.

[3] 朱国栋、徐 洋. 循环流化床锅炉设计与计算[M]. 北京: 清华大学出版社, 2004.

[4] 蒋敏华、肖 平. 大型循环流化床锅炉技术[M]. 北京: 中国电力出版社, 2009.

[5] HONG F, LONG D, CHEN J, *et al.* Modeling for the Bed Temperature 2D-Interval Prediction of CFB Boilers Based on Long-Short Term Memory Network [J]. En-

ergy, 2020, 194: 116733.

[6] 韩 静. 循环流化床垃圾焚烧炉燃烧和污染物生成的数值模拟[D]. 保定: 华北电力大学, 2010.

[7] 马素霞、杨献勇. 循环流化床锅炉燃烧系统的动态特性研究[J]. 中国电机工程学报, 2006, 26(9): 1-6.

[8] STEFANITSIS D, NESIADIS A, KOUTITA K, *et al.* Simulation of a CFB boiler integrated with a thermal energy storage system during transient operation [J]. Frontiers in Energy Research, 2020, 8: 169.

[9] ADAMCZYK W P, KOZOUB P, KLIMANEK A, *et al.* Numerical simulations of the industrial circulating fluidized bed boiler under air-and oxy-fuel combustion [J]. Applied Thermal Engineering, 2015, 87: 127-136.

[10] BORGES F, PINTO A, RIBEIRO D, *et al.* An unsupervised method based on support vector machines and higher-order statistics for mechanical faults detection [J]. IEEE Latin America Transactions, 2020, 18(6): 1093-1101.

[11] CHERKASSKY V. The nature of statistical learning theory [J]. IEEE Transactions on Neural Networks, 1997, 8(6): 1564.

[12] WEI L, CHEN J, YAN G. Bed temperature modeling and forecasting of circulating fluidized bed boiler based on SMO algorithm [J]. IEEE, 2010.

[13] ZHANG Y, ZHANG B, WU Z. Multi-model modeling of CFB boiler bed temperature system based on principal component analysis [J]. IEEE Access, 2019, 8: 389-399.

[14] LI J, CHENG K, WANG S, *et al.* Feature selection: A data perspective [J]. Acm Computing Surveys, 2016, 50(6): 1-45.

[15] 吴 田、曹志博、方玉群, 等. 基于最大信息系数的同塔双回输电线路感应电影响因素敏感性分析 [J]. 电测与仪表, 2021(7): 74-80.

[16] LEVER J, KRZYWINSKI M, ALTMAN N. Points of significance: principal component analysis [J]. Nature Methods, 2017, 14(7): 641-642.

[17] MATLAB 中文论坛. MATLAB 神经网络 30 个案例分析 [R]. 2010.

[18] 尹凌霄. 电站燃煤锅炉燃烧优化的智能技术研究 [D]. 南京: 东南大学, 2014.

(责任编辑 杨启岳)