

基于粒子群算法的煤气化过程 LS_SVM 预测模型^①

谢文君¹, 曹根牛¹, 李怀毅²

¹(西安科技大学 理学院, 西安 710054)

²(黄河上游水电开发有限责任公司, 西宁 810000)

摘 要: 针对最小二乘支持向量机参数选择对模型性能的重要影响, 并且以往的参数优选方法效果差且耗时长这一问题, 提出基于粒子群算法优化最小二乘支持向量机预测模型. 该模型用最小二乘支持向量机理论建立, 用粒子群算法优化模型参数. 论文将此模型用于预测评价固定床煤气化气化效果的三个主要性能指标(气体热值、气化效率、气体产率), 通过现场实际数据仿真结果表明, 该算法有效地提高了模型预测精度, 验证了此模型的可靠性和可用性.

关键词: 最小二乘支持向量机; 粒子群优化算法; 参数优化; 煤气化过程

Coal Gasification LS_SVM Forecasting Model Based on Particle Swarm Algorithm

XIE Wen-Jun¹, CAO Gen-Niu¹, LI Huai-Yi²

¹(School of College of Science, Xi'an University and Technology, Xi'an Shanxi 710054, China)

²(HuangHe Hydropower Development CO., LTD., Xi ning Qinghai 810000, China)

Abstract: According to the least squares support vector machine (LS_SVM) parameter selection has important influence on the model of performance, and conventional parameter optimization methods' effect is poor and time-consuming, this paper present a least squares support vector machine prediction model which based on particle swarm algorithm. The model based on least squares support vector machine theories, and with particle swarm algorithm to optimize the model parameters. In this paper we use the model to predict three main performance indexes of evaluating coal gasification effect of the fixed bed (gas heating value, gasification efficiency, gas production rate), through the practical data's simulation results show that the algorithm can effectively improve the prediction accuracy of the model, and the model's reliability and usability has been verified.

Key words: Least Squares-Support Vector Machine(LS_SVM); Particle Swarm Optimization(PSO); parameter optimization; coal gasification process

近年来, 我国煤炭资源十分紧缺, 人们为了达到有效利用煤炭资源并且尽可能多的得到高品质合成气的双赢目地, 对评价煤气化效果的主要性能指标的预测显得尤为重要. 但是, 由于煤气化过程涉及了许多物理化学反应, 是一个复杂的、强非线性的耦合性系统, 因此很难分析其内部机理建立准确的预测模型. 目前建立的煤气化过程预测模型有 BP 神经网络预测模型^[1]、运用拉格朗日方法建立的预测模型^[2]等.

最小二乘支持向量机算法具有可靠的全局最优性, 适用于小样本、非线性问题, 并且避免了维数灾难, 已

成功应用于实际复杂过程的建模, 文献[3-5]利用最小二乘支持向量机算法建立煤气化相关参数的预测模型并取得了一定的成果. 预测模型中参数的选取对模型的学习精度和泛化能力十分重要, 已有的文献[6-8]中通常采用梯度下降法、试凑法或遍历法来选取预测模型参数, 但这些方法不仅耗时长且有可能陷入局部最优解. 因此, 本文将具有全局优化性能的粒子群优化算法与最小二乘支持向量机模型结合, 采用粒子群算法优化最小二乘支持向量机参数, 从而提高模型预测精度.

① 收稿时间:2012-10-11;收到修改稿时间:2012-11-26

1 粒子群优化算法原理

粒子群算法(PSO)源于对鸟群捕食行为的研究,是一种进化算法. 粒子群算法的数学描述如下:

随机初始化粒子群 $pop = \{X_1, X_2, \dots, X_N\}$, 其中 $X_i = (X_{i1}, X_{i2}, \dots, X_{id})^T$, $V_i = (v_{i1}, v_{i2}, \dots, v_{id})$, ($i = 1, 2, \dots, N$) 分别为第 i 个粒子在 d 维空间上的位置和速度. 计算每个粒子的适应度值, 得到第 i 个粒子的自身最优解 p_i 和整个粒子群的最优解 p_g . 每次迭代时, 每个粒子根据公式(7)(8)更新自己的速度和位置. 如此反复迭代, 直达到最大迭代次数.

$$v_{i,j}(t+1) = wv_{i,j}(t) + c_1r_1[p_{i,j} - x_{i,j}(t)] + c_2r_2[p_{g,j} - x_{i,j}(t)] \quad (1)$$

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1) \quad i=1,2,\dots,n \quad (2)$$

$$j=1,2,\dots,d$$

其中, $i=1,2,\dots,n$ 表示粒子个数, $j=1,2,\dots,d$ 表示维数. w 为惯性权重, c_1, c_2 为正的学习因子, r_1, r_2 为 0 到 1 之间的均匀分布的随机数, t 为迭代次数.

2 预测实例与分析

2.1 基于 LS_SVM 的固定床煤气化过程预测模型建立

氧煤比是氧气的进料量与煤的进料量的比值, 气化过程中, 氧煤比的增大会提高合成气中有效气体成分($CO+H_2$)含量, 增加合成气热值和碳转化率. 但是, 随着氧气量的增加出口煤气中无效气体成分($CO+H_2$)随之增加, 从而降低气化效率. 水煤比是水蒸气的进料量与煤的进料量的比值, 气化过程中, 加入的水蒸气会提高合成气中有效气体含量, 但水煤比过高时, 会使出口气体中无效气体成分($CO+H_2$)增加, 直接影响气化效果^[9].

综上所述, 氧煤比和水煤比是煤气化过程中重要的工艺参数, 其值将直接影响出口煤气中的气体组成^[10]、气化效率、气体热值和产气率等.

本文以氧煤比和水煤比作为模型的输入量, 以气体热值、气化效率和产气率为模型输出量, 建立固定床煤气化过程的 LS_SVM 预测模型, 如图 1 所示.

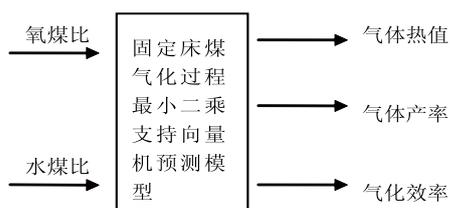


图 1 固定床煤气化过程支持向量机模型

对给定的训练样本 $D = \{(x_i, y_i) | i=1, 2, \dots, l\}$, $x_i \in R^n$ 是输入数据, $y_i \in R^n$ 是对应的输出数据, l 为训练样本个数, n 为数据向量空间维数. 由最小二乘支持向量机算法^[11]可得到函数估计:

$$[y_{LHV}, y_{\%}, y_{\eta}] = \sum_{i=1}^l a_i K(x, x_i) + b \quad (3)$$

其中, $K(x_i, x_j) = \varphi(x_i) \times \varphi(x_j)$ 为核函数. 参数 a 和 b 为:

$$\begin{bmatrix} b \\ a \end{bmatrix} = \Phi^{-1} \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (4)$$

上式中,

$$\Phi = \begin{bmatrix} 0 & 1 & & 1 \\ 1 & K(x_1, x_1) + \frac{1}{c} & \dots & K(x_1, x_l) \\ \vdots & \vdots & & \vdots \\ 1 & K(x_l, x_1) & \dots & K(x_l, x_l) + \frac{1}{c} \end{bmatrix}$$

其中, c 为惩罚参数.

2.2 基于 PSO 的 LS_SVM 参数优化算法

本文选用高斯径向基核函数(RBF), 则已建立的最小二乘支持向量机模型为:

$$[y_{LHV}, y_{\%}, y_{\eta}] = \sum_{i=1}^l \alpha_i \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) + b \quad (5)$$

其中包含两个未知参数: 惩罚因子 C , 核函数宽度 σ^2 . 通过对模型参数对 (C, σ^2) 的最优选择可使得最小二乘支持向量机模型达到最佳状态. 每一个粒子代表 LS_SVM 模型中的一对参数 (C, σ^2) , 选取均方差作为适应度函数:

$$E = \frac{1}{l} \sum_{i=1}^l (y_i - \hat{y}_i)^2 \quad (6)$$

其中, E 表示适应值, l 表示样本个数, y_i 是第 i 个样本的实际输出值, \hat{y}_i 是第 i 个样本的预测输出值. 粒子群优化 LS_SVM 参数步骤下:

1) PSO 的参数设置: 设置粒子种群数 N 、最大迭代次数 M 、粒子维数 d 、最大速度 V_{max} , 惯性权重 w , 学习因子 c_1 和 c_2 ;

2) 群定训练样本集, 根据经验和试算确定 (C, σ^2) 的大致范围, 初始化各粒子的位置向量 X_i 和速度向量 V_i , 设粒子 i 的当前最优位置为 p_i , $p_i = X_i (i=1, 2, \dots, N)$;

3) 将训练样本以及各粒子位置带入线性方程组(4)

中，计算出每个粒子相对应的模型参数 $(\alpha_1, \alpha_2, \dots, \alpha_i, b)$ ；再将每个粒子对应的模型参数和训练样本带入(5)式计算出每个粒子的预测值；

4) 将训练样本输出值和已计算出的预测值带入(6)式中，计算出每个粒子的适应度值；

5) 根据计算出的每个粒子的适应度值，将当前各粒子的位置存储在各粒子的最优位置 p_i 中，将粒子群中适应度值最优的个体的位置存储在全局最优 p_g 中；

6) 根据(1)(2)式更新粒子的速度和位置；

7) 对更新的每个粒子计算其适应度值，将其适应度值与其上一代最优值比较，若较好，则将更新的粒子作为当前最优的位置 p_i ；

8) 比较当前所有粒子的个体最优位置 p_i 和全局最优位置 p_g 的适应度值，若 p_i 优于 p_g ，则更新 p_g ；

9) 若迭代次数已达到最大迭代次数或用已搜索到的最优位置计算出的适应值已满足给定适应值的要求时，则停止迭代，否则转步骤 6)；

10) 将测试样本输入已训练好的 LS_SVM 模型预测中，检验此模型的可靠性。

3 仿真结果与验证

本文选取煤炭粒度为 5~50mm 的陕北神府煤为气化原料，在 3.0MP 压力下以固定床煤气化过程为例。选用 30 组样本数据，其中 20 组为 PSO_LSSVM 的训练样本，10 组为检验样本。通过试算法选取优化参数范围为 $c \in (1, 1000), \sigma \in (0.1, 10)$ 。粒子种群规模 $N=100$ ，最大迭代次数 $M=100$ ，学习因子 $c_1 = c_2 = 2$ 。基于 MATLAB 软件平台，经 PSO 优化，得到 LS_SVM 模型最优参数为 (998.8724, 0.1001)，参数优化过程如图 2 所示，从图中可看出，经过 20 次迭代后，最优适应度值达到最小值 2×10^{-4} 。为了检验 PSO 优化 LS_SVM 参数的实际效果，选取遗传算法优化 LS_SVM 模型参数作为对比。采用遗传算法优化 LS_SVM，模型参数优化结果为 (944.6351, 0.1607)。

通过两种算法对模型训练的基础上，采用 10 组数据对模型进行仿真验证。图 3、图 4、图 5 分别给出了采用两种优化算法建立的 LS_SVM 预测模型对 10 个测试数据的测试结果，测试结果分别为气体热值、气化效率和气体产率的检验样本的拟合效果图。表 1 为两种算法检验误差比较。为了将两种算法进行性能比较，表 2 给出了两种算法得出的最优适应度值、适应

度值达到最优时的最大迭代次数的比较结果。

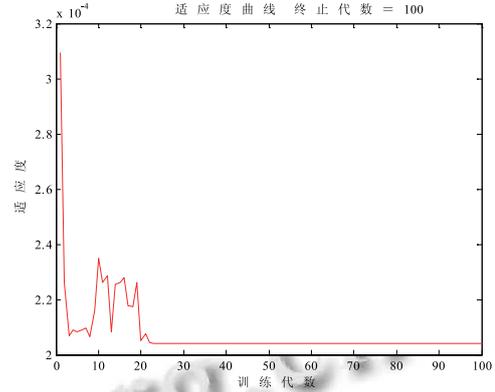


图 2 PSO 参数寻优过程

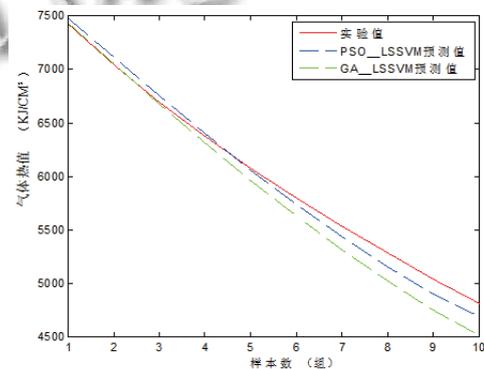


图 3 气体热值检验样本拟合结果

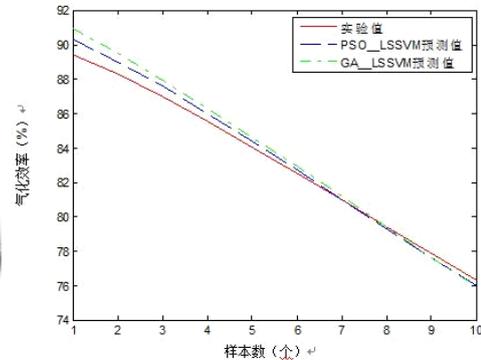


图 4 气化效率样本检验拟合结果

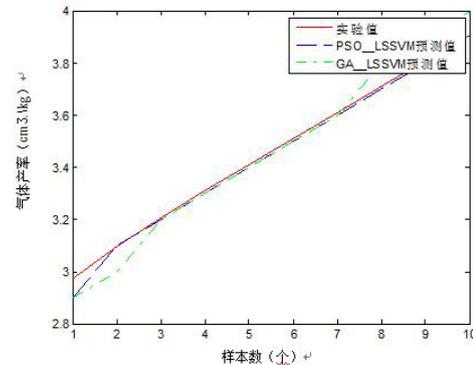


图 5 气体产率样本检验拟合结果

表1 两种算法的检验误差

训练方法	训练样本数	参数对 (C, σ) 的训练结果	检验样本数	气体热值平均相对误差 (%)	气化效率平均相对误差 (%)	气体产率平均相对误差 (%)
PSO_LSSVM	20	(998.8724, 0.1001)	10	1.42%	1.89%	2.59%
GA_LSSVM	20	(944.6351, 0.1607)	10	2.75%	3.48%	4.48%

由检验拟合图 3-5 和表 1 可得出, 测试阶段, 由 PSO_LSSVM 模型训练得出的气体热值的平均相对误差为 1.42%; 气化效率的平均相对误差为 1.89%; 气体产率的平均相对误差为 2.59%。而 GA_LSSVM 模型训练的气体热值、气化效率、气体产率的平均相对误差分别为 2.75%, 3.48%, 4.48%。从而明显得出: 在优化固定床煤气化过程 LS_SVM 预测模型时, 粒子群算法比遗传算法的测试误差小, 因此, 粒子群算法比遗传算法在优化预测模型方面更准确, 即 PSO_LSSVM 模型比 GA_LSSVM 模型具有更好的学习能力和泛化能力。

由表 2 可得出, 与遗传算法相比, 使用粒子群算法优化煤气化过程 LS_SVM 预测模型参数, 对煤气化过程指标的预测具有良好的效果, 同时 PSO 具有更快的收敛速度和更优的适应度值。因此, PSO_LSSVM 煤气化过程预测模型具有更好的实用性。

表2 PSO 与 GA 比较结果

LS_SVM	PSO	GA
C	998.8724	944.6351
σ^2	0.1001	0.1607
最优适应度值	2.0657×10^{-4}	2.1142×10^{-4}
最大迭代次数	20	26

4 结语

煤气化过程内部机理十分复杂, 因而很难建立准确的机理模型。本文在分析了煤气化过程的一些主要影响因素和性能指标的基础上, 运用最小二乘支持向量机理论, 以氧煤比和水煤比为模型输入变量, 以气体热值、气化效率和气体产率为模型输出变量, 建立了固定床煤气化过程的预测模型。为了使该预测模型有较好的学习精度和泛化能力, 利用 PSO 优化模型参数。通过实验数据检测, PSO 比 GA 在 LS_SVM 参数优化方面效果更好。

本文基于 LS_SVM 建立的固定床煤气化过程预测模型为多输入多输出模型, 而大多数文献中都是建立

的多输入单输出或单输入单输出模型, 这些模型在相关应用邻域具有一定的局限性。因此, 本文建立的多输入多输出模型在对其他领域的应用具有实际的指导意义和参考价值。

参考文献

- 1 Cui yang, Xu Long. Prediction model of coal catalytic gasification based on the improved BP neural network. Journal of Fuel Chemistry and Technology, 2011, 39(2): 90-93.
- 2 Jun Xie, Wenqi Zhong. Eulerian-Lagrangian method for three-dimensional simulation of fluidized bed coal gasification. Advanced Powder Technology, 2012, 66(6): 1285-1295.
- 3 王春林. 大型电站锅炉配煤及燃烧优化的支持向量机建模与实验研究. 浙江: 浙江大学, 2007.
- 4 陶少辉, 史书阳等. LS_SVM 模型在线校正的替代法及软测量应用. 化工自动化及仪表, 2010, 37(8): 15-18.
- 5 HC Yang, SB Zhang, KZ Deng, PJ Du. Research into a feature method of hyperspectral imagery using PSO and SVM. Journal of China University of Mining & Technology, 2007, 17(4): 473-478.
- 6 WRIGHT A H. Genetic Algorithms for Real Parameter Optimization. Foundations of Genetic Algorithms[M]. Rawlins G J E, EDSGA: Morgan Kaufmann, 1991: 205-218.
- 7 V Cherkassky, Y Ma. Practical selection of SVM parameters and noise estimation for SVM regression. Neural networks, 2004, 17(1): 113-126.
- 8 Wenjian Wang, Zongben Xua, Weizhen Lu, et al. Determination of the spread parameter in the Gaussian Kernel for classification and regression. Neurocomputing, 2003, 55(3): 271-290.
- 9 张大晶. 浅析相关参数对气流床粉煤加压气化的影响. 煤化工, 2008, 138(5): 44-46.
- 10 陈超, 姚洪, 岳良等. 氧煤比与蒸汽煤比对粉煤气化炉影响的数值模拟分析[C]. 华中科技大学煤燃烧国家重点实验室, 2010.
- 11 邓乃杨, 田英杰. 数据挖掘中的新方法——支持向量机[M]. 第2版. 北京: 科学出版, 2006, 12.