

基于贝叶斯优化-随机森林回归的燃煤锅炉 NO_x 预测模型

孙胡彬¹, 杨建国¹, 金宏伟², 屠海彪², 周晓亮³, 赵虹¹

(1. 浙江大学 能源清洁利用国家重点实验室, 杭州 310027;

2. 浙江浙能台州第二发电有限责任公司, 浙江台州 317109;

3. 杭州集益科技有限公司, 杭州 310012)

摘要: 根据某超超临界 1 050 MW 燃煤机组实际运行数据, 采用随机森林(RF)算法建立燃煤锅炉炉膛出口烟气中 NO_x 质量浓度预测模型, 并利用贝叶斯优化(BO)进行超参数寻优, 将 BO-RF 模型与网格搜索优化的 RF 模型(GSO-RF)进行对比。为了更好地评价预测模型, 以平均绝对百分比误差 δ_{MAPE} 和决定系数 R^2 作为评价指标, 将所建立的 BO-RF 模型与目前常见的基于贝叶斯优化的 BP 神经网络(BO-BPNN)模型、最小二乘支持向量机(BO-LSSVM)模型进行比较。结果表明: BO-RF 模型比 GSO-RF 模型的预测精度更高, 且 BO-RF 模型的 δ_{MAPE} 为 1.478%, R^2 为 0.916 2, 均优于 BO-BPNN 模型和 BO-LSSVM 模型的预测结果, 证明 BO-RF 模型具有更高的预测精度和更优的泛化性能。

关键词: NO_x; 预测模型; 随机森林; 贝叶斯优化

中图分类号: X773

文献标志码: A

学科分类号: 470.30

NO_x Prediction Model for Coal-fired Boiler Based on Bayesian Optimization-Random Forest Regression

SUN Hubin¹, YANG Jianguo¹, JIN Hongwei², TU Haibiao²,
ZHOU Xiaoliang³, ZHAO Hong¹

(1. State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, China;

2. Zhejiang Zheneng Taizhou Second Power Generation Co., Ltd., Taizhou 317109, Zhejiang Province, China; 3. Hangzhou Jiyi Technology Co., Ltd., Hangzhou 310012, China)

Abstract: Based on the actual operating data of a coal-fired boiler in a 1 050 MW ultra-supercritical unit, a random forest (RF) algorithm was used to establish a prediction model for NO_x concentration in the flue gas at the furnace outlet of the coal-fired boiler, and Bayesian optimization (BO) was used to optimize the hyperparameters. Then the BO-RF model was compared with the grid search optimized RF model (GSO-RF). In order to better evaluate the prediction model, the established BO-RF model was compared with the current common error Bayesian optimized back propagation neural network (BO-BPNN) model and Bayesian optimized least square support vector machine (BO-LSSVM) model, using the average absolute percentage error δ_{MAPE} and the coefficient of determination R^2 as evaluation indicators. Results show that

收稿日期: 2021-11-09 修订日期: 2022-03-04

作者简介: 孙胡彬(1996—), 男, 浙江慈溪人, 硕士研究生, 研究方向为电厂数据挖掘。

赵虹(通信作者), 男, 教授, 硕士, 电话(Tel.): 13805747305; E-mail: zhaohong@zju.edu.cn.

the prediction accuracy of the BO-RF model is higher than that of the GSO-RF model, and the δ_{MAPE} of the BO-RF model is 1.478%, the R^2 is 0.916 2, which are better than the prediction results of the BO-BPNN model and the BO-LSSVM model, indicating that the BO-RF model has higher prediction accuracy and better generalization performance.

Key words: NO_x; prediction model; random forest; Bayesian optimization

燃煤电厂是我国主要的电力来源,根据 bp 最新发布的数据,2020 年我国的燃煤发电量占总发电量的 63.2%^[1]。而燃煤电厂排放的 NO_x 是主要的大气污染物,严重危害环境和人体健康。此外,为了满足“超低排放”的要求,需要对 NO_x 进行烟气后处理,而锅炉燃烧过程中较高的 NO_x 生成量会增加额外的烟气脱硝费用,从而增加发电成本^[2]。目前,NO_x 排放控制方式主要有燃烧过程控制和烟气后处理 2 种^[3],而针对燃煤锅炉炉膛出口烟气中 NO_x 生成量的预测是其中的重要环节。因此,精确预测燃煤锅炉 NO_x 生成量,对配煤掺烧、燃烧优化以及污染物超低排放控制等具有重要意义。

锅炉系统是具有强耦合性的复杂系统,其 NO_x 生成特性受燃料种类、锅炉负荷、燃烧器类型和运行方式等多种因素影响,涉及湍流燃烧、传热传质等多种过程,难以用简单过程反映,建立基于反应动力学原理的 NO_x 生成预测模型难度较大^[4]。

近年来,随着人工智能的快速发展,大量机器学习或深度学习算法被应用到电站锅炉建模领域,实现了高度非线性问题的预测,使其中以数据挖掘、机器学习等为核心技术的电站锅炉 NO_x 浓度预测模型被大量提出。其中,以神经网络模型^[5-8]和支持向量机模型^[9-13]为代表,这 2 种算法被广泛验证并应用在锅炉 NO_x 预测领域。但是,神经网络存在易陷入局部最小值和过拟合的问题,影响模型的泛化能力,且需要大量样本训练模型;支持向量机是一种基于核函数的算法,泛化能力在很大程度上依赖于核函数的选择,且在处理大规模样本数据集时学习效率不高^[14]。换言之,支持向量机的参数选择人工影响较大,不利于预测精度和训练速度的进一步提高。

随机森林(RF)由 Breiman 于 2001 年提出,是一种基学习器为决策树的集成学习算法^[15-16],具有预测精度高、训练速度快、调节参数少和抗过拟合能力强等优点,且适用于各种数据集运算,被誉为“代表集成学习技术水平的方法”。

针对神经网络和支持向量机模型存在的问题,笔者基于某 1 050 MW 超超临界机组实际运行数据,采用贝叶斯优化(BO)算法进行超参数寻优,建

立基于随机森林回归的锅炉 NO_x 生成预测(BO-RF)模型。并将该模型与网格搜索法优化的随机森林模型(GSO-RF)、贝叶斯优化的 BP 神经网络模型(BO-BPNN)、贝叶斯优化的最小二乘支持向量机模型(BO-LSSVM)进行比较。

1 随机森林与贝叶斯优化

1.1 随机森林

随机森林在样本选择中基于 Bootstrap 自助采样法,对给定包含 m 个原始样本的数据集进行 m 次有放回的采样,最终得到含 m 个样本的采样集,根据概率公式:

$$\lim_{m \rightarrow \infty} \left(1 - \frac{1}{m}\right)^m = \frac{1}{e} \approx 0.368 \quad (1)$$

采样集包含初始数据集 63.2% 的样本。重复进行 T 轮抽取,训练 T 棵基决策树并进行组合形成随机森林。随机森林示意图如图 1 所示,其中 D 为原始样本, $D_1 \sim D_T$ 为 T 个采样集。

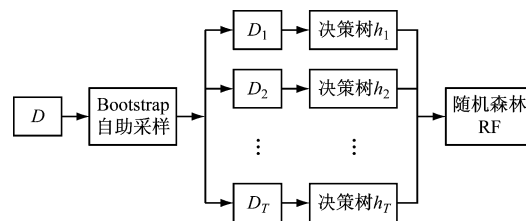


图 1 随机森林示意图

Fig. 1 Schematic diagram of RF

随机森林是 Bagging 算法的一个扩展变体,在决策树的训练过程中引入了随机属性选择。对于随机森林中基决策树的每个结点,先从结点的属性集合(假设 d 个属性)中随机选择一个包含 k 个属性的子集,再从这个子集中选择一个最优属性进行划分。因此,随机森林的基学习器的多样性不仅来自于对初始训练集采样的样本扰动,还来自于属性扰动,使得随机森林最终的泛化性能可通过基学习器之间差异度的增加而进一步提升。此外,由于在对决策树进行划分时仅需要考察一个属性子集,而 Bagging 则需要考虑所有属性,因此随机森林的训练效率也通常要优于 Bagging^[17]。

随机森林回归采用的基决策树为分类与回归树 (CART), 它基于平方误差最小化准则:

$$\min_{j,s} \left\{ \min_{c_1, c_2} \left[\sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right] \right\} \quad (2)$$

其中, j 为变量属性; s 为对应的变量值; 作为切分变量与切分点, 定义 2 个区域 R_1 和 R_2 , c_1, c_2 分别为 2 个区域包含的数据实例的输出 y_i 的均值。

$$R_1(j, s) = \{x | x^{(j)} \leq s\} \quad (3)$$

$$R_2(j, s) = \{x | x^{(j)} > s\} \quad (4)$$

$$c_1 = \text{ave}(y_i | x_i \in R_1(j, s)) \quad (5)$$

$$c_2 = \text{ave}(y_i | x_i \in R_2(j, s)) \quad (6)$$

在训练数据集所在的输入空间中, 递归地将每个区域划分为 2 个子区域并决定每个子区域上的输出值, 构建二叉决策树 $h(x)$ 。最终, 随机森林回归的输出结果为 T 棵 CART 回归树输出值的平均值 $H(x)$:

$$H(x) = \frac{1}{T} \sum_{i=1}^T h_i(x) \quad (7)$$

随机森林算法的步骤如下:

(1) 利用 Bootstrap 自助采样法, 生成训练子集;

(2) 采取随机属性选择的方法进行结点分裂, 构建单棵 CART 回归树;

(3) 重复步骤 (1)、步骤 (2), 构建 T 棵 CART 回归树, 每棵树自由生长, 不进行剪枝, 成为随机森林;

(4) 将 T 棵 CART 回归树的输出值取均值作为随机森林的输出结果 $H(x)$ 。

1.2 贝叶斯优化

贝叶斯优化^[18]是一种有效的全局优化算法, 本质上属于概率模型, 是在概率机器学习和人工智能领域的先进技术之一^[19]。它是一种基于模型的序贯优化方法, 根据对未知目标函数 f 获取的信息, 找到下一个评估位置, 从而最快地达到最优解^[20]。优化过程中利用贝叶斯定理:

$$p(f | D_{1:t}) = \frac{p(D_{1:t} | f) p(f)}{p(D_{1:t})} \quad (8)$$

$$D_{1:t} = \{(x_1, y_1), (x_2, y_2), \dots, (x_t, y_t)\} \quad (9)$$

$$y_t = f(x_t) + \epsilon_t \quad (10)$$

式中: $D_{1:t}$ 为已观测集合; x_t 为决策向量; y_t 为观测值; t 为观测集合个数 (迭代次数); ϵ_t 为观测误差; $p(D_{1:t} | f)$ 为 y 的似然分布; $p(f)$ 为未知目标函数 f 的先验概率分布; $p(D_{1:t})$ 为 f 的边际似然分布, 用于优化超参数; $p(f | D_{1:t})$ 为 f 的后验概率分布, 后

验概率分布表示修正先验分布后未知目标函数的置信度。

贝叶斯优化包括概率代理模型和采集函数 2 个核心部分。本文主要介绍的概率代理模型为非参数模型中的高斯过程, 采集函数采用置信边界策略。

1.2.1 概率代理模型

采用高斯过程 (Gaussian Processes, GPs) 作为概率代理模型, 高斯过程被广泛应用在回归、分类以及许多需要推断黑箱函数的领域中。高斯过程由一个均值函数 $m(x)$ 和一个半正定的协方差函数 $k(x, x')$ 构成:

$$f(x) \sim gp(m(x), k(x, x')) \quad (11)$$

$$m(x) = E[f(x)] \quad (12)$$

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))] \quad (13)$$

通常设置均值函数 $m(x) = 0$ 。

首先假设一个 0 均值的先验分布 $p(f | \mathbf{X}, \boldsymbol{\theta})$:

$$p(f | \mathbf{X}, \boldsymbol{\theta}) = N(0, \boldsymbol{\Sigma}) \quad (14)$$

$$\boldsymbol{\Sigma}_{i,j} = k(x_i, x_j) \quad (15)$$

式中: \mathbf{X} 为训练集, $\mathbf{X} = \{x_1, x_2, \dots, x_t\}$; f 为未知函数的集合, $f = \{f(x_1), f(x_2), \dots, f(x_t)\}$; $\boldsymbol{\Sigma}$ 为 $k(x, x')$ 构成的协方差矩阵; $\boldsymbol{\theta}$ 为超参数。

首先假设噪声 ϵ 存在且满足独立同分布的高斯分布, 进一步得到似然分布:

$$p(y | f) = N(f, \sigma^2 \mathbf{I}) \quad (16)$$

其中, $y = \{y_1, y_2, \dots, y_t\}$, σ 为标准差, \mathbf{I} 表示单位矩阵。

由先验分布和似然分布公式, 得到边际似然分布:

$$p(y | \mathbf{X}, \boldsymbol{\theta}) = \int p(y | f) p(f | \mathbf{X}, \boldsymbol{\theta}) df = N(0, \boldsymbol{\Sigma} + \sigma^2 \mathbf{I}) \quad (17)$$

通常通过最大化该边际似然分布来优化超参数 $\boldsymbol{\theta}$ 。根据高斯过程的性质, 存在如下联合分布:

$$\begin{bmatrix} y \\ f_* \end{bmatrix} \sim N \left[0, \begin{bmatrix} \boldsymbol{\Sigma} + \sigma^2 \mathbf{I} & K_* \\ K_*^\top & K_{**} \end{bmatrix} \right] \quad (18)$$

其中, f_* 表示预测函数值, \mathbf{X}_* 表示预测输入, $K_*^\top = \{k(x_1, \mathbf{X}_*), k(x_2, \mathbf{X}_*), \dots, k(x_t, \mathbf{X}_*)\}$, $K_{**} = k(\mathbf{X}_*, \mathbf{X}_*)$ 。

由联合分布公式得到以下预测分布:

$$p(f_* | \mathbf{X}, y, \mathbf{X}_*) = N(\langle f_* \rangle, \text{cov}(f_*)) \quad (19)$$

$$\langle f_* \rangle = K_*^\top [\boldsymbol{\Sigma} + \sigma^2 \mathbf{I}]^{-1} y \quad (20)$$

$$\text{cov}(f_*) = K_{**} - K_*^\top [\boldsymbol{\Sigma} + \sigma^2 \mathbf{I}]^{-1} K_* \quad (21)$$

式中: $\langle f_* \rangle$ 为预测均值; $\text{cov}(f_*)$ 为预测协方差。

通过明确指定先验均值函数 $m(x)$,可在增加模型解释性的同时方便先验信息的表达,此时预测协方差不变,预测均值变为:

$$\langle f_* \rangle = m(X_*) + K_*^T [\Sigma + \sigma^2 I]^{-1} (y - m(X_*)) \tag{22}$$

为了简便,在实际应用中通常假设先验均值函数恒为 0,后验均值并不限制为 0,因此该假设对后验准确性几乎没有影响。

1.2.2 采集函数

采集函数的目的主要是寻找下一个评估点,置信边界策略(GP-UCB)是一种针对高斯过程的评估策略,UCB 表示置信上界,在求解目标函数最大值时,UCB 策略的采集函数 α_t 为:

$$\alpha_t(x; D_{1:t}) = \mu_t(x) + \sqrt{\beta_t} \sigma_t(x) \tag{23}$$

式中: μ_t 为后验均值; σ_t 为后验标准差; β_t 为平衡期望和方差的参数。

贝叶斯优化的流程如下:

- (1) 最大化采集函数,得到下一个评估点 x_t 。
- (2) 评估目标函数 y_t 。
- (3) 整合数据 $D_{1:t}$,更新概率代理模型。
- (4) 迭代到指定次数,输出最优评估点。

2 随机森林回归的 NO_x 预测模型

2.1 数据预处理与特征变量选择

选取某超超临界 1 050 MW 燃煤机组的实际运行数据,共 1 440 组。首先对数据进行预处理,包括进行缺失值、零值的检验与处理。然后,对数据进行基本观察,发现锅炉负荷变化较大,所取的数据样本可能存在非稳态工况数据,对模型的影响较大,因此需要进行数据的异常值检测与剔除。通过利用数据可视化(折线图、散点图)、 3σ 法则和箱型图等方式,剔除了 27 组粗大误差数据与异常数据,最终选择 1 413 组数据。

结合不同的特征对 NO_x 生成的影响,选用 33 个输入参数(见表 1),输出参数为炉膛出口 NO_x 质量浓度(mg/m³)。

2.2 建立预测模型

基于随机森林回归的 NO_x 预测模型流程图如图 2 所示。首先将数据集按 8:2 的比例分为训练集和测试集,训练集采用 5 折交叉验证,利用贝叶斯优化算法对预测模型中 5 个超参数进行寻优(见表 2),以 5 折交叉验证的决定系数 R^2 均值作为目标函数,设置迭代次数,最终输出最优超参数组合并在训练集上训练 NO_x 预测模型,利用测试集进行模

型效果的评估。

表 1 输入参数

Tab. 1 Input parameters

参数	个数	数值范围
全水分质量分数 $w(M_1)/\%$	1	11.5~13.0
内在水分质量分数 $w(M_{ad})/\%$	1	3.5~5.5
空气干燥基灰分质量分数 $w(A_{ad})/\%$	1	14.0~22.5
空气干燥基挥发分质量分数 $w(V_{ad})/\%$	1	28.5~36.0
空气干燥基硫分质量分数 $w(S_{ad})/\%$	1	0.55~0.70
收到基低位发热量 $Q_{ar,net}/(kJ \cdot kg^{-1})$	1	20 500~23 500
总煤量/(t · h ⁻¹)	1	200~450
锅炉负荷/MW	1	540~1 050
总风量/(t · h ⁻¹)	1	1 800~3 250
二次风量/(t · h ⁻¹)	1	1 300~2 900
A~F 磨煤机入口一次风量/(t · h ⁻¹)	6	0~180
前后墙 A、B 侧燃尽风挡板开度/%	4	99~100
前后墙 A、B 侧各三层二次风门开度/%	12	20~100
炉膛氧量/%	1	1.5~5.5

为了更好地评价预测模型,选择平均绝对百分比误差(δ_{MAPE})和决定系数(R^2)作为评价指标。计算公式如下:

$$\delta_{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y'_i - y_i}{y_i} \right| \tag{24}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y'_i - y_{mean})^2} \tag{25}$$

式中: n 为数据个数; y'_i 为预测值; y_i 为实际值; y_{mean} 为预测值的平均值。

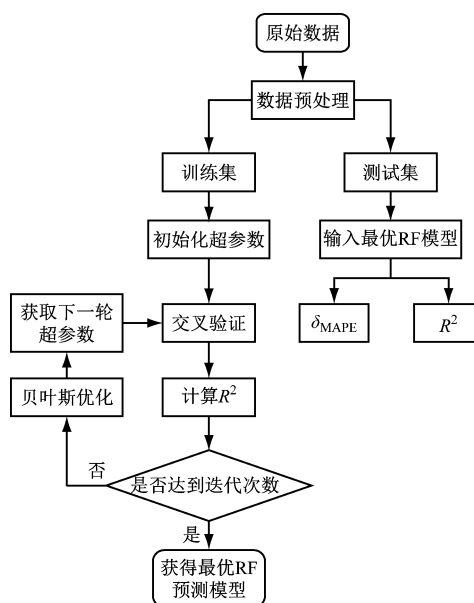


图 2 随机森林预测模型流程图

Fig. 2 Flow chart of the RF prediction model

根据经验,选择迭代次数为 200 次,经过 200 次迭代寻优后的模型参数结果见表 2。

表 2 最优 RF 模型参数结构
Tab. 2 The optimal RF parameter structure

参数	符号	数值
数目棵树	n_estimators	188
数目深度	max_depth	15
特征选择个数	max_features	0.9
叶子节点最小样本数	min_samples_leaf	1
内部节点划分最小样本数	min_samples_split	2

3 结果与分析

3.1 BO-RF 模型预测结果

最终通过寻优到的最佳超参数,在训练集和测试集上的预测结果如图 3 所示。从图 3 可以看出,经过 BO-RF 模型计算出的预测值与真实值差距较小,拟合效果好,表明预测值能较好反映实际的真实值情况,说明该 BO-RF 模型有着较好的预测能力。

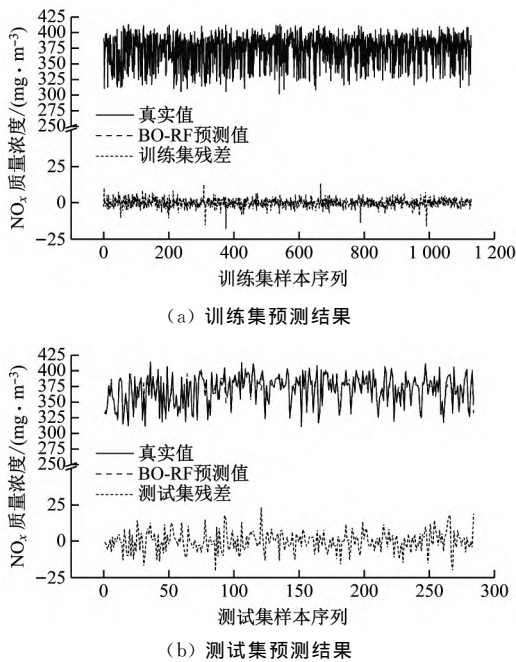


图 3 BO-RF 模型预测结果

Fig. 3 Predicted results of the BO-RF model

同时,为了验证贝叶斯优化的优势,将 BO-RF 模型与建立采用传统网格搜索法优化的 RF 模型进行比较。表 3 给出了 BO-RF 模型和 GSO-RF 模型的性能对比。从表 3 可以看出,BO-RF 模型在训练集和测试集上的 δ_{MAPE} 和 R^2 都比 GSO-RF 模型表现得更好,表明 BO-RF 模型训练效果更好,同时泛化能力更佳,且网格搜索法是在一个范围内不断进

表 3 2 种 RF 模型性能对比

Tab. 3 Performance comparison of two RF models

模型	训练集		测试集	
	$\delta_{MAPE}/\%$	R^2	$\delta_{MAPE}/\%$	R^2
GSO-RF	0.878	0.973 5	1.611	0.891 8
BO-RF	0.606	0.986 5	1.478	0.916 2

行超参数排列组合,在维度较多时并不适用,对后期的模型优化存在瓶颈。

3.2 多模型比较

为了验证 BO-RF 模型的可靠性和先进性,将基于同一训练集建立的 BO-RF 模型与 BO-BPNN 模型、BO-LSSVM 模型进行比较,结果如图 4 所示。表 4 给出了 3 种模型预测的评价指标情况。

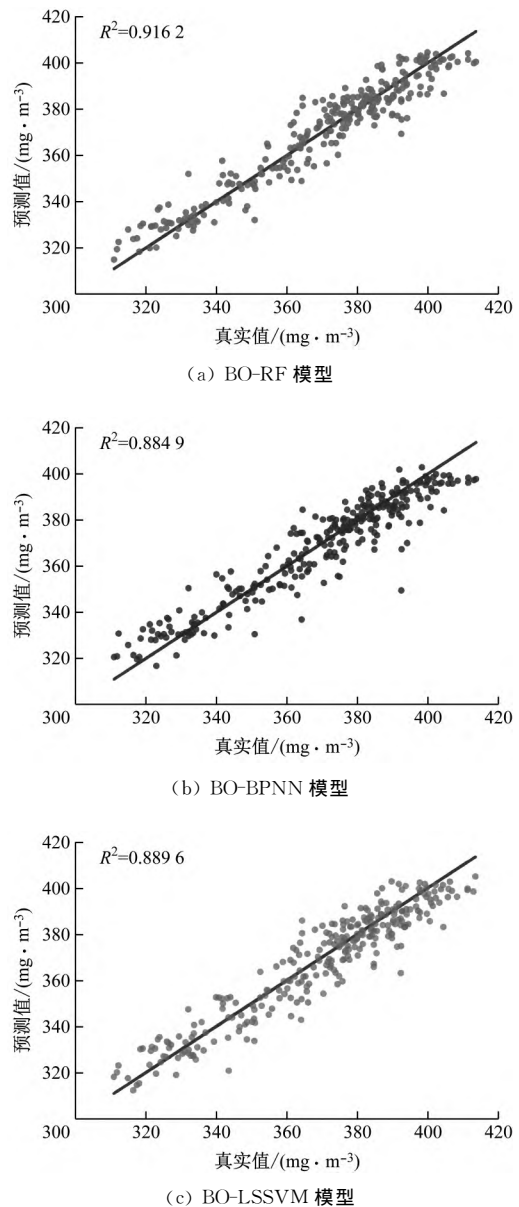


图 4 3 种模型预测结果对比

Fig. 4 Comparison of prediction results of three models

表4 3种模型评价指标

Tab. 4 Evaluation indexes of three models

模型	$\delta_{MAPE}/\%$	R^2
BO-RF	1.478	0.916 2
BO-BPNN	1.661	0.884 9
BO-LSSVM	1.766	0.889 6

从图4和表4可以看出,BO-RF模型的 δ_{MAPE} 在3种模型中最小, R^2 在3种模型中最大,说明与其他2种模型相比,BO-RF模型具有更高的预测精度和更优越的泛化性能,可以更好地完成高维多变量非线性拟合。

4 结 论

(1) 针对燃煤锅炉在运行过程中生成的NO_x浓度预测问题,建立了基于随机森林回归的NO_x预测模型。并采用贝叶斯优化算法对随机森林RF模型进行超参数寻优。结果表明,采用贝叶斯优化的随机森林RF模型的预测精度比传统网格搜索法优化模型的预测精度有明显提升。

(2) 将所建立的BO-RF模型与BO-BPNN模型、BO-LSSVM模型进行比较,BO-RF模型的 δ_{MAPE} 在3种模型中最小, R^2 在3种模型中最大,证明BO-RF模型具有更好的预测精度与泛化性能。

参考文献:

- [1] bp. bp statistical review of world energy 2021[EB/OL]. [2021-07-08]. <http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html>.
- [2] 杜振,钱徐悦,何胜,等. 燃煤电厂烟气SCR脱硝成本分析与优化[J]. 中国电力, 2013, 46(10): 124-128.
DU Zhen, QIAN Xuyue, HE Sheng, et al. Analysis and optimization for operating costs of gas SCR De-NO_x in coal-fired power plants[J]. **Electric Power**, 2013, 46(10): 124-128.
- [3] 梁志宏. 基于我国新大气污染排放标准下的燃煤锅炉高效低NO_x协调优化系统研究及工程应用[J]. 中国电机工程学报, 2014, 34(增刊1): 122-129.
LIANG Zhihong. Study and engineering application of high efficiency and low NO_x coordinated optimization control system for coal-fired boilers based on new air pollutant emission standard[J]. **Proceedings of the CSEE**, 2014, 34(Sup1): 122-129.
- [4] 卓建坤,焦伟红,宋少鹏,等. 锅炉燃烧优化中NO_x

预测模型研究进展[J]. 燃烧科学与技术, 2016, 22(6): 531-540.

ZHUO Jiankun, JIAO Weihong, SONG Shaopeng, et al. A review on nitrogen oxides prediction model in combustion optimization of boilers[J]. **Journal of Combustion Science and Technology**, 2016, 22(6): 531-540.

- [5] 谷丽景,李永华,李路. 电站锅炉燃烧优化混合模型预测[J]. 中国电机工程学报, 2015, 35(9): 2231-2237.
GU Lijing, LI Yonghua, LI Lu. Hybrid model prediction of utility boiler combustion optimization[J]. **Proceedings of the CSEE**, 2015, 35(9): 2231-2237.
- [6] 牛培峰,史春见,刘楠,等. 基于自适应量子灰狼算法的锅炉NO_x预测模型[J]. 动力工程学报, 2018, 38(4): 278-285, 315.
NIU Peifeng, SHI Chunjian, LIU Nan, et al. Prediction model for boiler NO_x emission based on adaptive quantum grey wolf optimization[J]. **Journal of Chinese Society of Power Engineering**, 2018, 38(4): 278-285, 315.
- [7] 杨国田,张涛,王英男,等. 基于长短期记忆神经网络的火电厂NO_x排放预测模型[J]. 热力发电, 2018, 47(10): 12-17.
YANG Guotian, ZHANG Tao, WANG Yingnan, et al. Prediction model for NO_x emissions from thermal power plants based on long-short-term memory neural network[J]. **Thermal Power Generation**, 2018, 47(10): 12-17.
- [8] 王文广,赵文杰. 基于GRU神经网络的燃煤电站NO_x排放预测模型[J]. 华北电力大学学报(自然科学版), 2020, 47(1): 96-103.
WANG Wenguang, ZHAO Wenjie. NO_x emission prediction model based on GRU neural network in coal-fired power station[J]. **Journal of North China Electric Power University (Natural Science Edition)**, 2020, 47(1): 96-103.
- [9] 顾燕萍,赵文杰,吴占松. 基于最小二乘支持向量机的电站锅炉燃烧优化[J]. 中国电机工程学报, 2010, 30(17): 91-97.
GU Yanping, ZHAO Wenjie, WU Zhansong. Combustion optimization for utility boiler based on least square-support vector machine[J]. **Proceedings of the CSEE**, 2010, 30(17): 91-97.
- [10] 刘吉臻,吕游,杨婷婷. 基于变量选择的锅炉NO_x排放的最小二乘支持向量机建模[J]. 中国电机工程学报, 2012, 32(20): 102-107.
LIU Jizhen, LYU You, YANG Tingting. Least

- squares support vector machine modeling on NO_x emission of boilers based on variable selection[J]. **Proceedings of the CSEE**, 2012, 32(20): 102-107.
- [11] 袁洪, 宋选锋, 赵征. 基于 LSSVM-ARMA 的电站锅炉 NO_x 排放量动态软测量的研究[J]. **电力科学与工程**, 2017, 33(12): 50-54.
- YUAN Hong, SONG Xuanfeng, ZHAO Zheng. Research on dynamic soft sensing of NO_x emission for utility boiler based on LSSVM-ARMA [J]. **Electric Power Science and Engineering**, 2017, 33(12): 50-54.
- [12] 杨建国, 叶凌云, 赵敏, 等. 基于煤种掺烧模式的锅炉燃烧 NO_x 预测模型[J]. **动力工程学报**, 2017, 37(11): 870-875.
- YANG Jianguo, YE Lingyun, ZHAO Min, et al. NO_x prediction model for multi-coal admixing combustion in boiler [J]. **Journal of Chinese Society of Power Engineering**, 2017, 37(11): 870-875.
- [13] 丁续达, 刘潇, 金秀章. 基于压缩感知最小二乘支持向量机的 NO_x 软测量模型[J]. **热力发电**, 2018, 47(3): 76-81.
- DING Xuda, LIU Xiao, JIN Xiuzhang. A soft sensor model for NO_x concentration based on CS-LSSVM [J]. **Thermal Power Generation**, 2018, 47(3): 76-81.
- [14] 丁世飞, 齐丙娟, 谭红艳. 支持向量机理论与算法研究综述[J]. **电子科技大学学报**, 2011, 40(1): 1-10.
- DING Shifei, QI Bingjuan, TAN Hongyan. An overview on theory and algorithm of support vector machines [J]. **Journal of University of Electronic Science and Technology of China**, 2011, 40(1): 1-10.
- [15] BREIMAN L. Random forests [J]. **Machine Learning**, 2001, 45(1): 5-32.
- [16] 王奕森, 夏树涛. 集成学习之随机森林算法综述[J]. **信息技术**, 2018, 12(1): 49-55.
- WANG Yisen, XIA Shutao. A survey of random forests algorithms [J]. **Information and Communications Technologies**, 2018, 12(1): 49-55.
- [17] 周志华. **机器学习** [M]. 北京: 清华大学出版社, 2016: 179-181.
- [18] 崔佳旭, 杨博. 贝叶斯优化方法和应用综述[J]. **软件学报**, 2018, 29(10): 3068-3090.
- CUI Jiayu, YANG Bo. Survey on Bayesian optimization methodology and applications [J]. **Journal of Software**, 2018, 29(10): 3068-3090.
- [19] GHABRAMANI Z. Probabilistic machine learning and artificial intelligence [J]. **Nature**, 2015, 521(7553): 452-459.
- [20] JONES D R, SCHONLAU M, WELCH W J. Efficient global optimization of expensive black-box functions [J]. **Journal of Global Optimization**, 1998, 13(4): 455-492.
- LIU Hao, ZHU Tong, ZHANG Tao. Evaluation and analysis of CCHP-ORC system for different buildings in Shanghai [J]. **Proceedings of the CSEE**, 2016, 36(12): 3198-3205.
- [19] REN Fukang, WANG Jiangjiang, ZHU Sitong, et al. Multi-objective optimization of combined cooling, heating and power system integrated with solar and geothermal energies [J]. **Energy Conversion and Management**, 2019, 199: 111866.

(上接第 909 页)

YANG Hengyue, LIU Qingrong, RUAN Yingjun. Selection of typical daily cooling and heating load of CCHP system based on *k*-means clustering algorithm [J]. **Thermal Power Generation**, 2021, 50(3): 84-90.

- [18] 刘豪, 朱彤, 张涛. 上海地区不同类型建筑的 CCHP-ORC 系统评价与分析 [J]. **中国电机工程学报**, 2016, 36(12): 3198-3205.