

# 数据挖掘技术在燃煤锅炉与垃圾焚烧锅炉中的应用 现状及展望

陆俊林<sup>1</sup>, 应雨轩<sup>2</sup>, 林晓青<sup>2</sup>, 李晓东<sup>2</sup>

(1. 浙江大学 工程师学院, 浙江 杭州 310000; 2. 浙江大学 热能工程研究所, 浙江 杭州 310012)

**摘要:**总结了数据挖掘技术在燃煤锅炉故障诊断、燃烧优化、污染物减排及机组优化运行等方面的应用现状,分析了关联规则、聚类分析、神经网络和支持向量机等数据挖掘算法在锅炉优化运行和污染物排放控制中的优缺点。分析表明:人工神经网络鲁棒性强、可自学习且适用面广,未来可基于焚烧机理并耦合其他算法进行工程应用;对于在高控制要求下智能化工况优化空间大的垃圾焚烧锅炉中的发展及应用,建议将数据挖掘技术与云计算平台结合,并考虑垃圾焚烧过程的实际工况和特性进一步开发数据预处理方法,扩大动态数据采集范围,提高模型的实际运行效率和泛化能力。

**关键词:**数据挖掘; 燃烧锅炉; 垃圾焚烧; 智慧能源

中图分类号:TK229.6, TK11<sup>+</sup>4

文献标识码:A

DOI:10.16146/j.cnki.rndlgc.2022.10.002

[引用本文格式]陆俊林,应雨轩,林晓青,等.数据挖掘技术在燃煤锅炉与垃圾焚烧锅炉中的应用现状及展望[J].热能动力工程,2022,37(10):16-26. LU Jun-lin, YING Yu-xuan, LIN Xiao-qing, et al. Status and outlook of data mining technology in coal-fired boiler and waste incineration boiler[J]. Journal of Engineering for Thermal Energy and Power, 2022, 37(10):16-26.

## Status and Outlook of Data Mining Technology in Coal-fired Boiler and Waste Incineration Boiler

LU Jun-lin<sup>1</sup>, YING Yu-xuan<sup>2</sup>, LIN Xiao-qing<sup>2</sup>, LI Xiao-dong<sup>2</sup>

(1. Polytechnic Institute, Zhejiang University, Hangzhou, China, Post Code:310000; 2. Institute for Thermal Power Engineering, Zhejiang University, Hangzhou, China, Post Code:310012)

**Abstract:** Based on the advantages and disadvantages of data mining algorithms that applied in coal-fired boiler optimization operation and pollutant emission control, including association rules, cluster analysis, neural networks as well as support vector machines, this paper reviews the status on fault diagnosis, boiler combustion optimization, pollutant emission reduction and unit optimization operation by data mining technology. The analysis shows that artificial neural network is relatively widely used due to its strong robustness and self-learning ability. In the future, based on the incineration mechanism and coupled with other algorithms, it can be further applied in engineering. Furthermore, on basis of existing application of data mining technology, the future development and application of the intelligent waste incineration boiler that has largely optimization space under high control requirements have been prospected. We suggest to combine the data mining technology with the cloud computing platform, consider the actual working conditions and characteristics of the waste incineration process, and further develop data preprocessing methods to expand the dynamic data sampling range and improve the actual operation efficiency and generalization a-

收稿日期:2021-11-01; 修订日期:2021-12-24

基金项目:国家重点研发计划项目(2020YFC1910100)

**Fund-supported Project:** National Key Research and Development Program of China(2020YFC1910100)

作者简介:林晓青(1988-),男,福建漳州人,浙江大学副教授.

通讯作者:陆俊林(1997-),男,广西南宁人,浙江大学硕士研究生.

bility of the model.

**Key words:** data mining, coal-fired boiler, waste incineration, smart energy

## 引 言

国内火力发电厂基于厂级监控系统(SIS)、集散控制系统(DCS)等平台,采集了大量运行数据并储存于数据库。但由于火电机组运行受多种复杂因素影响,数据往往存在噪声、缺失等不足,传统的算法难以建立能准确描述其运行特性的数学模型,导致大量数据利用困难<sup>[1]</sup>。而数据挖掘基于人工智能、机器学习、模式识别、统计学、数据库和可视化等技术,可以从数据库中挖掘出隐含的、未知的及有价值的信息。黄景涛等人<sup>[2]</sup>采用支持向量回归算法建模,并利用现场实测参数进行回归分析;刘宝玲等人<sup>[3]</sup>采用 k-mean 算法对电厂各种运行工况数据进行划分。目前,数据挖掘技术的研究尚存在部分模型泛化性能差、应用面窄及工程应用场景不足等问题,在复杂工况下的工业锅炉(如垃圾焚烧锅炉)中应用较少。

垃圾焚烧技术已成为我国生活垃圾处置的主要方式,截至 2020 年底,全国城市生活垃圾无害化处理场(厂)共 1 287 座,其中垃圾焚烧厂 463 座,焚烧处理能力 56.78 万 t/d,占比超过 58.93%<sup>[4]</sup>。与燃煤电厂相比,垃圾焚烧电厂入炉燃烧组分复杂、分布不均,焚烧工况波动更大,污染物排放限制更严格<sup>[5-6]</sup>。而垃圾焚烧电厂智能化运行处于起步阶段,数据挖掘技术在垃圾焚烧电厂中的应用尚未落地,进一步开展并完善针对垃圾焚烧运行数据库的数据挖掘技术,将推动垃圾焚烧的智慧化发展。

本文分析了数据挖掘在燃煤电厂中的应用,多角度分析与总结了数据挖掘技术在燃煤锅炉优化运行和污染物排放控制方面的技术关键和难点,并对垃圾焚烧行业数据挖掘技术的应用进行梳理,进一步探索数据挖掘技术在垃圾焚烧锅炉应用的前景。

## 1 数据挖掘的概念以及常用算法

数据挖掘过程如图 1 所示,主要可以分 3 步:数据准备、规律寻找、结果的分析 and 表达。数据准备阶

段从数据源中选取并整合用于挖掘的数据集;规律寻找阶段从数据集中找出规律;结果分析和表达是将规律以用户可理解的方式表示出来。

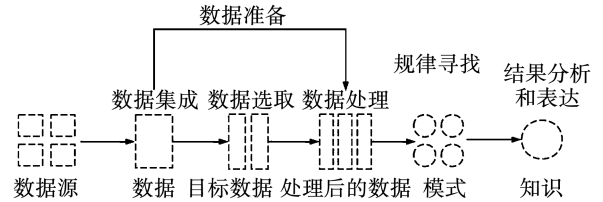


图 1 数据挖掘过程<sup>[7]</sup>

Fig. 1 Data mining process<sup>[7]</sup>

常用的数据挖掘算法包括聚类模式分析、关联规则、神经网络、遗传算法、支持向量机、模糊集理论和粗糙集理论等,如表 1 所示。聚类分析算法,主要将具有相似特征的事物归聚为一类,通过静态分类方法把相似对象归为一个子集或组别,不同子集或组别之间互不相关;关联规则算法的目标在于找出数据之间有意义的关系,主要采用支持度、置信度和频繁项集合 3 个参数描述;神经网络主要从信息处理角度模拟人脑神经网络,建立某种简单模型,按照特定规则进行连接构成的信息处理网络,主要包含输入层、隐藏层和输出层 3 个部分;遗传算法是基于生物进化理论而演变来的组合优化方式,将基因寻找适应度高的染色体的过程应用到求解问题的过程中;支持向量机(Support Vector Machines, SVM)是一种以统计学理论为基础的二分类模型,具有严格的理论和数学基础, SVM 的基本思想是将输入的样本空间利用核函数映射到高维的特征空间,然后在高维空间中求出最优超平面,该最优超平面可以分开两类样本,并使得样本间隔达到最大;模糊集理论是处理模糊性和不确定性等不完备问题的数学工具,采用隶属度描述,元素根据隶属度的值从属于某个集合,与模糊集理论类似;粗糙集理论也是处理模糊性和不确定性等不完备问题的数学工具,主要是在所研究问题保持不变的前提下通过知识约简把重复、冗余的属性去除,挖掘出潜在的新知识与数据内部的关联特征。

表 1 数据挖掘常用算法

Tab. 1 Common algorithms for data mining

分类	常用算法	特点	适用场景
聚类分析	$k$ -均值、模糊 $c$ -均值算法等	依赖 $k$ 值和初始聚类中心	常被用于分析复杂参数间的关系
关联规则	Apriori, FP-growth 算法等	适合事务数据库的挖掘,但计算耗时较长	常用频繁项集挖掘的方法来获取数据间隐含规则
神经网络	自组织神经网络、前向神经网络和后向神经网络等	适应性强、黑箱模型、解释性差,易陷入局部最优	常用于电厂的非线性过程中,如故障诊断、优化运行等
遗传算法	遗传算法、自适应遗传算法等	具备良好的全局搜索能力、擅长数据聚类	遗传算法收敛速度慢,局部搜索能力差常与神经网络结合使用
支持向量机	支持向量机、最小二乘支持向量机等	特征相关性不敏感、核函数选取困难,训练时间较长	常用于处理非线性问题
模糊理论	模糊聚类算法、模糊神经网络等	采用隶属度描述数据集	处理模糊和不确定性问题、常与聚类、关联规则、神经网络等结合
粗糙集理论	粗糙神经网络、粗糙遗传算法等	不需要提供任何先验知识或其他额外的信息	处理模糊性和不确定性等不完备问题的数学工具

## 2 数据挖掘技术在燃煤锅炉中的应用现状

### 2.1 故障诊断

发电企业生产过程具有工艺复杂、设备资产高度集中等特点,虽然建立了以集散控制系统(DCS)、厂级监控系统(SIS)和信息管理系统(MIS)为核心的信息化架构,但多数还处于机组的过程监视、报表与经济性能计算阶段,在设备的诊断、预警等功能方面相对较弱。结合数据挖掘等人工智能技术开发高效、便捷的故障诊断系统已经成为研究的热点问题。Cai 等人<sup>[8]</sup>基于 SVM 算法建立燃煤发电机组性能

在线监测模型,实现了锅炉飞灰含碳量和汽轮机排气焓的在线监测。图 2 为 300 MW 机组 SVM 模型监测数据和实测数据的对比。从图中可以看出,SVM 模型的计算结果与实测结果基本吻合。

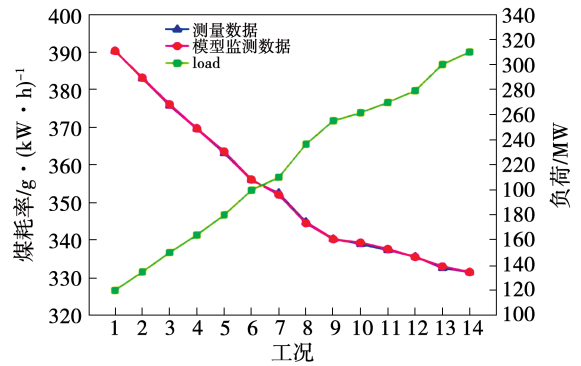


图 2 基于 300 MV 机组的 SVM 模型监测数据和测量数据对比<sup>[8]</sup>

Fig. 2 Comparison of monitoring data and measurement data of SVM model based on 300 MV unit<sup>[8]</sup>

系统间相互耦合,某个组件故障也会影响其他组件运行,所以设计故障诊断系统时需要考虑多种情况。Ma 等人<sup>[9]</sup>考虑到故障发生后参数的动态变化,将故障症状分为语义症状和趋势症状,并采用改进的模糊理论分别进行了计算。语义症状和趋势症状的模糊表达式为:

$$\mu(y) = \begin{cases} 1, & y \geq y_0 + A \\ \frac{y - y_0}{A}, & y_0 - A < y < y_0 + A \\ -1, & y \leq y_0 - A \end{cases} \quad (1)$$

$$\mu(y') = \begin{cases} 1, & y' \geq B \\ \frac{y'}{B}, & -B < y' < B \\ -1, & y' \leq -B \end{cases} \quad (2)$$

式中:  $y$ —特征参数;  $y_0$ — $y$  在不同工作点的参考值;  $A$ —不同故障下参数变化的最大值;  $y'$ — $y$  的最大变化率;  $B$ —常数。

采用模糊表达方法优化了数据训练集,结合 BP 神经网络对多工况下的机组运行数据进行了故障诊断。

设备故障是长期积累的结果,当设备仍在可控范围内运行时,提前预警可以避免进一步损失或事

故的发生。现有的故障预警方法主要有<sup>[10]</sup>故障特征方法、概率统计方法、状态分类方法和状态预测方法。其中,故障特征方法实验成本较高且建模机制比较困难;概率统计方法基于特征参数值和概率分布进行预测,但此过程需要大量故障数据;状态分类方法把故障警告看作分类问题,仅仅需要正常的历史数据,数据获取容易但是相应的分类方法复杂;状态预测方法基于历史数据建立设备正常状态模型,并采用异常函数进行异常状态捕获。与前 3 种方法相比,状态预测方法物理意义明确,适用性更强。但是现有的很多状态预测方法过于复杂,实际应用时响应较慢。数据挖掘算法的引入,可以有效提高预测效率。张维等人<sup>[10]</sup>以 DSC 大量运行数据作为支撑,基于改进模糊关联规则的模糊推理离线建立了正常状态的知识库并对其进行关联规则挖掘;Xiao 等人<sup>[11]</sup>提出了一种证据理论框架下基于证据  $k$  临近规则(EKNN)的状态监测预警方法,如图 3 所示,采用自适应的折扣因子使预警边界适应训练样本的局部分布特征,提高了模型的鲁棒性和效率。

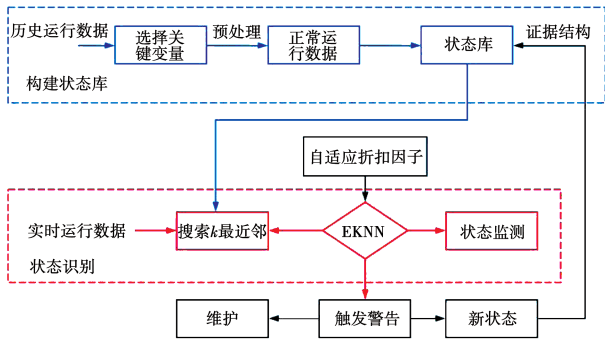


图 3 证据理论框架下基于证据  $k$  临近规则的状态监测预警方法<sup>[11]</sup>

Fig. 3 Condition monitoring and early warning method based on evidence  $k$  proximity rule under the framework of evidence theory<sup>[11]</sup>

综上,针对火电机组设备的故障诊断,神经网络、支持向量机等数据挖掘技术取得了良好的效果;考虑到设备之间的耦合,需要对多工况状态进行挖掘,研究人员引入了神经网络、模糊特征计算等对原始算法进行优化。

## 2.2 锅炉的燃烧优化

锅炉燃烧是一个多输入输出的非线性、强耦合

过程,通过建立锅炉燃烧系统模型,并结合非线性寻优技术,找出不同负荷、不同燃烧状况下锅炉燃烧参数的最优值,从而指导锅炉燃烧优化调整是当前研究的重点<sup>[12-13]</sup>。锅炉的燃烧优化技术可以分为 4 类<sup>[14-16]</sup>:(1) 通过设备层面的优化设计实现锅炉燃烧优化调整;(2) 通过燃烧调整实验的方式确定优化方案;(3) 通过在线监测锅炉燃烧参数指导优化调整;(4) 通过先进控制逻辑、控制算法或人工智能技术实现燃烧优化。这 4 类技术各有优缺点,其中第 4 类技术能够充分利用机组的运行数据,直接实现锅炉高效低排,具有成本低、风险小、效果显著等优点。韩义等人<sup>[17]</sup>建立 BP 神经网络燃烧仿真模型,并结合遗传算法实时计算和调整运行参数,使得锅炉效率与  $\text{NO}_x$  排放质量浓度达到最优。表 2 为燃烧优化后的试验数据对比。可以看出,优化后  $\text{NO}_x$  排放质量浓度降至  $289 \text{ mg/m}^3$ ,锅炉效率为 92.02%。

表 2 燃烧优化试验数据对比<sup>[17]</sup>

Tab. 2 Comparison of combustion optimization test data<sup>[17]</sup>

参数	优化前	优化后
功率/MW	330	330
煤量/ $\text{t} \cdot \text{h}^{-1}$	176.5	175.3
A 层二次风门开度/%	65.5	25.0
B 层二次风门开度/%	65	25.0
C 层二次风门开度/%	20	20.0
D 层二次风门开度/%	68.5	36.0
E 层二次风门开度/%	60	35.0
OFA1 二次风门开度/%	80	100
OFA2 二次风门开度/%	80	100
二次风压/kPa	1.35	1.11
一次风压/kPa	9.91	9.81
氧量/%	3.1	1.13
$\text{NO}_x$ 质量浓度/ $\text{mg} \cdot \text{m}^{-3}$	348	289
锅炉效率/%	92.37	92.02

锅炉燃烧优化是一个典型的多目标优化问题:

$$\begin{cases} [\min_{y_{\text{NO}_x}}, \max_{y_{\eta}}] \\ a(i) \leq x(i) \leq b(i), i = 1, 2, \dots, m \end{cases} \quad (3)$$

式中:  $[y_{\text{NO}_x}, y_{\eta}] = f(x(i))$ ;  $f$ —映射关系;  $x(i)$ —第  $i$  个优化变量;  $[a(i), b(i)]$ —第  $i$  个优化变量的

取值范围;  $m$ —优化变量数。

在该类问题的研究中一般通过加权因子将多目标优化问题转化为单目标优化问题,但是这样每次只能得到一个帕雷托解(有效解),无法充分利用数据,增大计算难度,并且加权系数的改变依赖决策者的经验,容易造成更优帕雷托解遗漏。李素芬等人<sup>[18]</sup>在电站锅炉  $\text{NO}_x$  排放与效率模型的基础上采用多目标粒子群算法优化锅炉运行参数,使获得的解为帕雷托最优解集,即为  $\text{NO}_x$  排放量最小和锅炉效率最大两个目标之间的折中曲线,如图4所示。

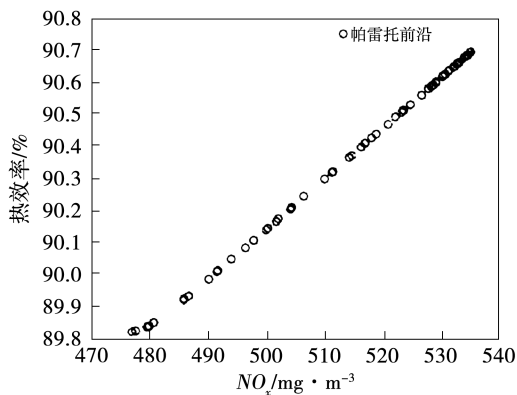


图4  $\text{NO}_x$  排放量和锅炉效率优化曲线<sup>[18]</sup>

Fig. 4  $\text{NO}_x$  emission and boiler efficiency optimization curve<sup>[18]</sup>

在上述针对锅炉高效低排的研究中,并没有考虑到机组在大范围变负荷情况下动态优化的需求。对此,Si 等人<sup>[19]</sup>考虑到传统自适应更新样本方法的局限性,引入了一种样本去除准则:

$$\begin{cases} D_1 = \{(x_i, y_i) : \|x_i - x_{\text{new}}\| > a, i = 1, \dots, N\} \\ D_2 = \bigcap_{j=1}^m \{(x_i, y_i) : |x_{i,j} - x_{\text{new},j}| < a_j\} \end{cases} \quad (4)$$

式中:  $a, a_j$ —阈值;  $(x_i, y_i)$ —训练集第  $i$  个样本点参数;  $(x_{\text{new}}, y_{\text{new}})$ —新给定的样本点。

通过改进的在线支持向量回归模型,逼近特征函数,模型参数根据运行条件而适应性变化,结合非支配排序的遗传算法(NSFA-II)进行多目标优化,获得了SCR系统和锅炉的最佳运行设置。综上,数据挖掘在锅炉燃烧优化中的应用主要是通过神经网络预测并结合遗传算法进行寻优,在实际应用中还

要考虑锅炉的运行状态,进一步对算法模型进行改进。

在锅炉燃烧优化的研究中,由于基于人工智能算法的数据挖掘技术依据锅炉系统历史数据,不需要建立精确的数学模型,具有低成本、低风险及效果显著等优点,在燃煤电厂的节能减排研究中应用广泛。针对神经网络与寻优过程中的多目标优化问题,研究人员对排放模型进行了改进,优化神经网络或者引入了支持向量机等算法提高了模型的精确度与计算效率;引入改进的遗传算法改进寻优过程,使得计算结果更加精确。例如,非支配排序的遗传算法进一步提高了预测和优化效果。但该类研究仍缺乏工程实际应用,未来可针对实际应用过程,选取或进一步改进算法模型。针对锅炉实际运行问题,通过在线优化、增量计算等实现了锅炉的动态燃烧优化。

### 2.3 污染物减排

电站中一般采用烟气在线监测系统(Continuous Emission Monitoring System, CEMS)测量烟气污染物的成分,但CEMS成本较高,且工作环境恶劣影响其使用周期<sup>[20]</sup>。软测量采用间接测量的思路,利用易测过程变量与难以直接测量变量间的数学关系,采用计算或估算等方法建立数学模型,结合计算机技术实现待测变量的测量<sup>[21]</sup>。白建云等人<sup>[22]</sup>针对循环流化床锅炉  $\text{NO}_x$  排放直接测量困难的问题,选择烟气中氧体积分数和机组负荷作为输入,采用BP神经网络建立软测量模型,间接测出  $\text{NO}_x$  的质量浓度;BP神经网络模型如图5所示。BP神经网络模型的第1隐含层输出为:

$$u_i = f\left(\sum_{m=1}^M \omega_{im} x_m + b^1\right) \quad (5)$$

第2隐含层输出为:

$$v_i = f\left(\sum_{i=1}^I \omega_{ji} u_i + b^2\right) \quad (6)$$

输出节点为:

$$y = f\left(\sum_{j=1}^J \omega_{pj} u_j + b^3\right) \quad (7)$$

其中,

$$f(x) = \frac{1}{1 + e^{-x}} \quad (8)$$

式中:  $b^1, b^2, b^3$  — 阈值;  $i, j$  — 任一神经元;  $\omega_{im}$  — 输入层和第 1 隐层权重;  $\omega_{ji}$  — 第 1 隐层和第 2 隐层权重;

$\omega_{pj}$  — 第 2 隐层和输出层权重;  $u_i, v_i, y$  — 各层输出。

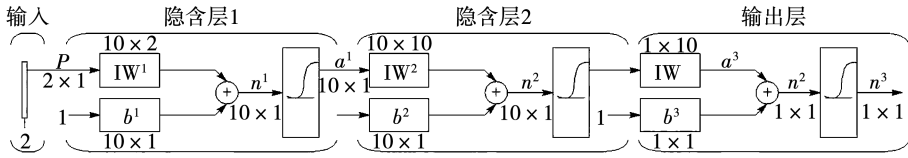


图 5 BP 神经网络模型<sup>[22]</sup>

Fig. 5 BP neural network model<sup>[22]</sup>

但 BP 神经网络模型会耗费大量的时间且容易陷入局部最优值,限制了 BP 神经网络的预测推广能力<sup>[23]</sup>。对此,刘锁清等人<sup>[24]</sup>在权值的变化量中引入动量项:

$$\Delta\omega_{ip}(n) = \alpha\Delta\omega_{ip}(n-1) + \eta\delta_p(n)v_i(n) \quad (9)$$

式中:  $n$  — 样本数;  $\delta$  — 局部梯度;  $\eta$  — 学习步长;  $\omega$  — 权重;  $\alpha$  — 动量项。

改进后的 BP 神经网络预测模型,对脱硫效率的预测精度更高;牛培峰等人<sup>[25]</sup>针对神经网络的局限性在  $\text{NO}_x$  排放特性的研究中引入 SVM,建立了循环流化床锅炉  $\text{NO}_x$  排放特性预测模型,考虑 SVM 过于依赖参数选择,引入果蝇优化算法对 SVM 参数进行优化。

综上,针对烟气污染物直接测量成本高的问题,引入软测量的理念,结合智能算法,可以实现烟气污染物较低成本的预测。软测量方法主要基于神经网络、SVM 等构建关联模型,具有成本低、响应快速的优点;但同样存在需要大量前馈数据、计算准确性等问题。具体来说,由于人工神经网络在非线性系统中具有良好的应用能力,因此在烟气污染物预测中被广泛应用。但由于其计算耗费时间长且容易陷入局部极小值,在应用过程中引入了支持向量机(SVM)模型。

### 2.4 机组优化运行

图 6 为电站运行优化结构示意图。电厂运行优化是在对设备和系统离线研究分析的基础上,在线对运行参数进行评价和决策,以经济性、安全性为导向,指导机组运行<sup>[26]</sup>。电厂运行优化的目标主要包含 4 个方面:在线性能计算、运行目标值确定、偏差

分析和运行优化操作指导<sup>[27]</sup>。

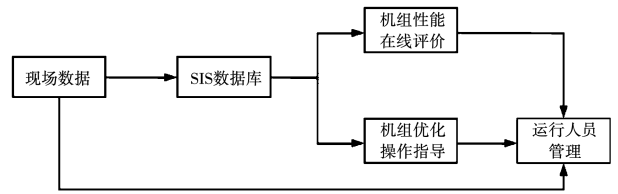


图 6 火电站运行优化示意图<sup>[28]</sup>

Fig. 6 Schematic diagram of thermal power plant operation optimization<sup>[28]</sup>

目标值是在满足安全生产的条件下,机组在当前运行工况与环境条件下达到最佳运行状态时,各个运行参数和性能指标的值,优化目标值的确定是机组经济性诊断和机组优化运行的基础<sup>[29]</sup>。优化目标值确定的方法主要有:(1)设计参数法;(2)变工况计算法;(3)优化试验法;(4)统计分析最优法;(5)人工智能算法寻优<sup>[30]</sup>。设计参数法操作简单规则清晰,但是随着机组运行得到的设计参数目标值与实际比误差较大;变工况计算法在理论上可行,但计算结果受模型精度影响较大,并且容易偏离实际值;优化试验法在机组优化实验完成初期比较适用,但试验耗时耗力,计算周期长,导致精度下降;统计分析最优法计算量大,过程复杂且方法单一,预测结果对数据的精确性依赖很大<sup>[30-31]</sup>;人工智能算法针对机组运行的历史数据,采用数据挖掘、机器学习方法进行寻优,该方法能综合各种影响因素,可以有效避免热力过程难以进行数值描述的难题,结果可信度高。王培红等人<sup>[1]</sup>基于 Apriori 算法编制了数据挖掘分析软件,得到了运行参数之间的关系:

$$(a_1 \sim a_2) - (b_1 \sim b_2) - (c_1 \sim c_2) \quad (10)$$

式中： $a$ —负荷； $b$ —运行监控参数； $c$ —锅炉效率，且  $a_1 < a_2, b_1 < b_2, c_1 < c_2$ 。

根据所运行参数之间的关联，可以为锅炉的优化运行提供参考，获得运行优化基准值。表 3 为排烟氧量基准值。

表 3 排烟氧量基准值<sup>[1]</sup>

Tab.3 Reference value of flue gas oxygen content<sup>[1]</sup>

工况	机组负荷/MW	基准氧量/%
1	67.5 ~ 92.9	8.749 ~ 9.555
2	92.9 ~ 100.5	7.789 ~ 8.749
3	100.5 ~ 107.5	7.101 ~ 7.798
4	107.5 ~ 122.7	5.359 ~ 6.433
5	122.7 ~ 128.2	3.887 ~ 5.359

电厂数据始终处于动态更新中，外界负荷、运行人员操作和燃料热值等也会对数据造成影响，所以火电机组运行数据是典型随时间变化、无规则、无限增长的动态数据<sup>[32]</sup>。而传统的数据挖掘是从静态的历史数据库中获取数据，数据动态更新，使得之前发现的知识不适用于现在，因此数据挖掘在机组优化运行中，需要考虑数据的大规模动态增长。牛成林<sup>[29]</sup>基于大量历史运行数据，利用模糊数值型关联规则算法确定了特定负荷工况下氧量最优值以及所有的频繁项集，并且考虑到由于机组运行状态的改变引起最优值更新，将增量挖掘引入关联规则挖掘过程，使得挖掘结果更具有合理性、准确性和可操作性。图 7 为增量挖掘的运行优化目标值确定系统结构图。

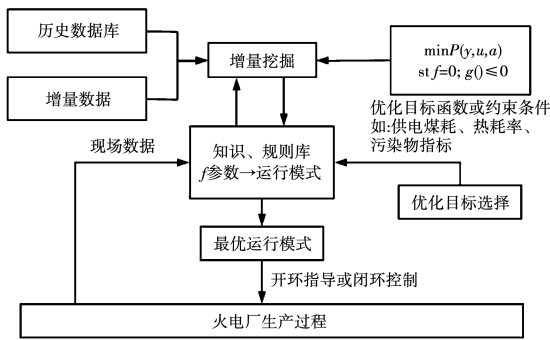


图 7 基于增量挖掘的运行优化目标值确定系统结构图<sup>[29]</sup>

Fig.7 System structure diagram determined by operational optimization target value based on incremental mining<sup>[29]</sup>

综上，在机组确定优化目标值的过程中，针对传统方法的缺陷，可以引入数据挖掘算法对电厂运行过程中产生的海量数据进行挖掘，寻找数据之间的关联性，从而得到一系列最优工况下的参数。但是考虑到火电机组的运行数据是无规则、动态增加的，所以需要传统的数据挖掘算法进行改进。对此可以引入增量挖掘思想进行动态的数据挖掘，从而能更加准确地反映机组的状况，对机组的优化运行、故障检测等有很好的指导作用。

### 3 数据挖掘技术在垃圾焚烧电厂的应用现状

在针对垃圾焚烧电厂优化的研究中，周志成<sup>[33]</sup>引入粗糙集理论选取垃圾火焰特征量并结合 BP 神经网络构建了垃圾焚烧炉燃烧状态诊断模型。图 8 为基于神经网络的垃圾焚烧诊断原理。

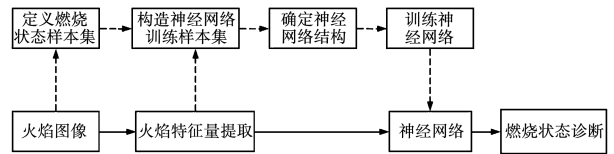


图 8 基于神经网络的垃圾焚烧诊断原理<sup>[33]</sup>

Fig.8 Principles of waste incineration diagnosis based on neural network<sup>[33]</sup>

姜明男等人<sup>[34]</sup>以大型垃圾焚烧炉为对象，基于支持向量机(SVM)建立了垃圾焚烧炉运行参数预测模型，将焚烧状态相关参数作为输入，预测了燃烧炉排速度、燃尽炉排速度、一次风变频控制柜输出频率和二次风变频控制柜输出频率，预测过程如图 9 所示。

垃圾热值变化会显著影响垃圾焚烧炉的稳定运行，大部分垃圾热值计算预测主要采用垃圾的物理成分作为模型输入<sup>[35-37]</sup>，然而现场定量进行垃圾成分分析比较困难。对此马晓茜等人<sup>[38]</sup>通过 Garson 方法和主成分分析法(PCA)对垃圾焚烧电厂在线运行数据进行分析，将处理后的数据结合 BP 神经网络和遗传算法，在线预测了入炉垃圾热值，其结构如图 10 所示。

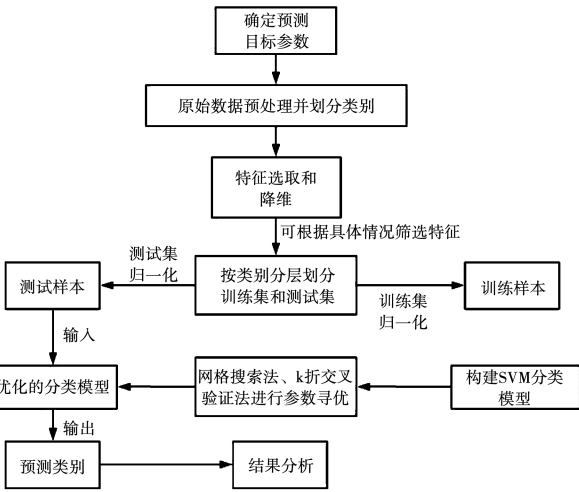


图 9 基于 SVM 的垃圾焚烧炉运行参数预测流程图<sup>[34]</sup>

Fig. 9 SVM-based flow chart for predicting operating parameters of waste incinerator<sup>[34]</sup>

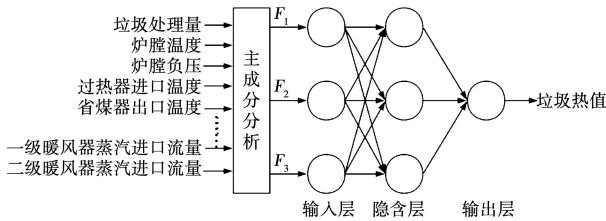


图 10 PCA-BP 网络结构<sup>[38]</sup>

Fig. 10 PCA-BP network structure<sup>[38]</sup>

数据挖掘在垃圾焚烧过程中的应用主要集中在对垃圾热值、燃烧状态、污染物排放的预测和控制方面。鉴于垃圾焚烧受到更复杂因素的影响,对模型的稳定性、鲁棒性等都有更高的要求,在应用过程可以引入其他优化算法如遗传算法、粒子群算法、模糊算法等,进一步提高模型的精度和鲁棒性。

垃圾焚烧过程会产生多种污染物,主要有粉尘、有毒重金属、酸性气体( $\text{NO}_x$ ,  $\text{SO}_2$ ,  $\text{HCl}$ ,  $\text{HF}$  等)和痕量有机污染物(如二噁英等),其中以二噁英毒性和检测难度最大,对二噁英检测与减排是研究热点<sup>[39]</sup>。常用的监测方法有离线检测法、指示物/关联物在线间接检测法和软测量方法。离线检测法需要专门的实验室和相应的仪器设备,耗时长,指示物/关联物在线间接检测法成本较高;软测量方法基于运行数据,成本低,耗时短,是较为理想的检测方

法<sup>[40]</sup>。肖晓东等人<sup>[41]</sup>通过分析二噁英类污染物的排放与工况参数和常规污染物的关系,以炉膛温度、锅炉出口烟温、烟气流量、 $\text{SO}_2$  质量浓度、 $\text{HCl}$  质量浓度及颗粒物浓度为输入,采用支持向量回归方法对二噁英类总浓度及毒性当量的小样本数据进行预测,模型性能整体表现良好;乔俊飞等人<sup>[42]</sup>在二噁英软测量的研究方面,针对标记样本难以获取以及小样本高维数据特征的选择问题,提出了基于多层特征选择的固废焚烧过程二噁英排放浓度软测量模型,如图 11 所示。

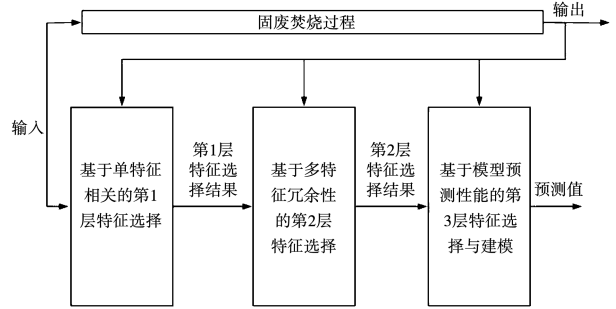


图 11 基于多层特征选择的固废焚烧过程二噁英排放浓度软测量模型<sup>[42]</sup>

Fig. 11 Soft-sensing model of dioxin emission concentration in solid waste incineration process based on multi-layer feature selection<sup>[42]</sup>

数据挖掘在垃圾焚烧电厂的应用较为广泛,如基于炉膛火焰的智能识别、垃圾焚烧炉的诊断、垃圾热值的预测等,常用的算法为人工神经网络、支持向量机、聚类算法等。由于二噁英生成过程机理复杂,测量困难,采用软测量的方法,结合神经网络、支持向量机等算法对二噁英进行测量和预测可以为排放控制以及电厂运行提供参考。

## 4 结论

结合故障诊断、锅炉燃烧优化、污染物减排以及机组运行优化等过程,梳理了数据挖掘技术在火电机组的应用现状,针对当前垃圾焚烧电厂智能化运行的需求,阐述了数据挖掘在垃圾焚烧行业的应用前景。

(1) 目前,设备安全运行的故障诊断主要基于过程监控报表及人工诊断。未来可结合数据挖掘思



想,运用神经网络、SVM、遗传算法及粒子群算法等进行设备的故障诊断,并在此基础上根据机组实际运行状况对设备进行故障预警,提高机组的经济安全运行能力。

(2) 烟气中与  $\text{NO}_x$  等污染物有关的运行参数与锅炉效率直接相关,降低排放的同时会使锅炉效率降低。锅炉燃烧优化主要以在线监测燃烧参数指导优化调整为主,未来可进一步结合数据挖掘的思想,在预测模型的基础上采用遗传算法、粒子群算法、人工蜂群算法等进行寻优,在满足锅炉高效运行的同时污染物排放达到最低。

(3) 污染物控制以在线连续监测为主,软测量作为一种间接在线测量方法,可结合数据挖掘技术构建常规烟气污染物预测模型,较好适应非线性系统。神经网络在非线性系统中具有良好的适应能力,结合人工神经网络与软测量,对烟气污染物进行在线预测,可以很好地解决烟气在线监测系统成本较高的问题;通过引入 SVM 等算法,可有效克服神经网络易陷入局部最优值、计算时间较长的缺点。

(4) 机组运行优化目标值的确定是机组经济性诊断和机组运行优化的基础。由于运行环境发生改变,常常使得机组偏离最优工况运行,所以需要确定实际运行状态下机组处于最优工况时操作参数的范围。优化目标值确定主要来自于厂家设计值、优化实验、统计分析或热力计算。未来可采用关联规则、模糊理论等数据挖掘算法进行挖掘;考虑到数据的动态变化、数据量过大等问题,可以引入增量挖掘思想将数据抓取范围从静态数据拓展至动态数据进行算法的对应优化,从而提高机组的运行效率。

(5) 相对传统燃煤电厂,数据挖掘算法在垃圾焚烧电厂中应用起步较晚,常用于燃烧诊断、垃圾热值预测等研究中。其中,神经网络、SVM 等算法应用相对比较广。数据挖掘和软测量结合,对二噁英进行在线测量,可以为二噁英的排放控制提供指导。

#### 参考文献:

[1] 王培红,陈强,董益华,等.数据挖掘及其在电厂 SIS 中的应用[J].电力系统自动化,2004(8):76-79.  
WANG Pei-hong, CHEN Qiang, DONG Yi-hua, et al. Data mining

and its application in power plant SIS[J]. Automation of Electric Power Systems, 2004(8):76-79.

- [2] 黄景涛,马龙华,茅建波,等.基于支持向量回归的 300 MW 电站锅炉再热汽温建模[J].中国电机工程学报,2006(7):19-24.  
HUANG Jing-tao, MA Long-hua, MAO Jian-bo, et al. Modeling of 300 MW power station boiler reheat steam temperature based on support vector regression[J]. Proceedings of the CSEE, 2006(7):19-24.
- [3] 刘宝玲,何钧.基于数据挖掘及 SIS 的工况划分方法研究[J].南昌工程学院学报,2009,28(6):36-39.  
LIU Bao-ling, HE Jun. Research on working condition division method based on data mining and SIS[J]. Journal of Nanchang Institute of Technology, 2009, 28(6):36-39.
- [4] 国家统计局.2019 中国统计年鉴[M].北京:中国统计出版社,2019.  
National bureau of statistics. 2019 China statistical yearbook[M]. Beijing: China Statistics Press, 2019.
- [5] GB 13223-2011, 火电厂大气污染物排放标准[S].  
GB 13223-2011, Emission standards of air pollutants for thermal power plants[S].
- [6] GB 18485-2014, 生活垃圾焚烧污染控制标准[S].  
GB 18485-2014, Pollution control standards for domestic waste incineration[S].
- [7] 潘召涛.基于数据挖掘的火电厂风机故障预警研究[J].机械管理开发,2021,36(2):119-120.  
PAN Zhao-tao. Research on early warning of fan failure in thermal power plant based on data mining[J]. Machinery Management Development, 2021, 36(2):119-120.
- [8] CAI J, MA X, LI Q. On-line monitoring the performance of coal-fired power unit: a method based on support vector machine[J]. Applied Thermal Engineering, 2009, 29(11/12):2308-2319.
- [9] MA L, LEE K Y. Fuzzy neural network approach for fault diagnosis of power plant thermal system under different operating points: 2008 IEEE power and energy society general meeting-conversion and delivery of electrical energy in the 21st century[C]. Pittsburgh: IEEE, 2008.
- [10] 张维,刘吉臻,高明明.基于数据挖掘的循环流化床锅炉辅机故障预警[J].动力工程学报,2019,39(10):826-833.  
ZHANG Wei, LIU Ji-zhen, GAO Ming-ming. Fault early warning of circulating fluidized bed boiler auxiliary equipment based on data mining[J]. Journal of Chinese Society of Power Engineering, 2019, 39(10):826-833.
- [11] CHEN X, WANG P, HAO Y, et al. Evidential KNN-based condition monitoring and early warning method with applications in

- power plant[J]. *Neurocomputing*,2018,315:18-32.
- [12] 曹庆才,高德欣,刘芳.基于神经网络与遗传算法的锅炉燃烧优化系统设计[J].*自动化技术与应用*,2016,35(6):10-14.  
CAO Qing-cai,GAO De-xin,LIU Fang. Design of boiler combustion optimization system based on neural network and genetic algorithm[J]. *Technology of Automation and Applications*,2016,35(6):10-14.
- [13] 张国斌,张光明,刘永江,等.基于数据挖掘案例推理的电站锅炉燃烧优化[J].*热能动力工程*,2021,36(3):114-121.  
ZHANG Guo-bin,ZHANG Guang-ming,LIU Yong-jiang, et al. Combustion optimization of power station boilers based on data mining case reasoning[J]. *Journal of Engineering for Thermal Energy and Power*,2021,36(3):114-121.
- [14] 周建新,樊征兵,司凤琪,等.电站锅炉燃烧优化技术研究发展综述[J].*锅炉技术*,2008(5):33-36.  
ZHOU Jian-xin,FAN Zheng-bing,SI Feng-qi, et al. Overview of research and development of combustion optimization technology for power station boilers [J]. *Boiler Technology*, 2008 (5): 33-36.
- [15] 孔亮,张毅,丁艳军,等.电站锅炉燃烧优化控制技术综述[J].*电力设备*,2006(2):19-22.  
KONG Liang,ZHANG Yi,DING Yan-jun, et al. Overview of combustion optimization control technology for power station boilers [J]. *Electric Power Equipment*,2006(2):19-22.
- [16] 徐军伟,宋兆龙,王磊.电站锅炉燃烧优化技术现状和发展动向[J].*江苏电机工程*,2005(3):6-7.  
XU Jun-wei,SONG Zhao-long,WANG Lei. Status quo and development trends of combustion optimization technology for power station boilers [J]. *Jiangsu Electrical Engineering*, 2005 (3): 6-7.
- [17] 韩义,孙英波,于英利,等.遗传算法与BP神经网络在低氮燃烧优化中的应用[J].*内蒙古电力技术*,2017,35(6):32-35.  
HAN Yi,SUN Ying-bo,YU Ying-li, et al. Application of genetic algorithm and BP neural network in low nitrogen combustion optimization[J]. *Inner Mongolia Electric Power Technology*,2017,35(6):32-35.
- [18] 李素芬,吴祯祥,葛玉林.电站锅炉NO<sub>x</sub>排放与效率的混合建模及优化[J].*热科学与技术*,2007(1):26-31.  
LI Su-fen,WU Zhen-xiang,GE Yu-lin. Hybrid modeling and optimization of NO<sub>x</sub> emissions and efficiency in power station boilers [J]. *Thermal Science and Technology*,2007(1):26-31.
- [19] SI F,ROMERO C E,YAO Z, et al. Optimization of coal-fired boiler SCRs based on modified support vector machine models and genetic algorithms[J]. *Fuel*,2009,88(5):806-816.
- [20] 吕游,刘吉臻,杨婷婷,等.基于PLS特征提取和LS-SVM结合的NO<sub>x</sub>排放特性建模[J].*仪器仪表学报*,2013,34(11):2418-2424.  
LYU You,LIU Ji-zhen,YANG Ting-ting, et al. Modeling of NO<sub>x</sub> emission characteristics based on the combination of PLS feature extraction and LS-SVM[J]. *Chinese Journal of Scientific Instrument*,2013,34(11):2418-2424.
- [21] 乔弘.火电厂热工参数软测量关键技术和方法研究[D].北京:华北电力大学,2009.  
QIAO Hong. Research on key technologies and methods of soft measurement of thermal parameters in thermal power plants[D]. Beijing:North China Electric Power University,2009.
- [22] 白建云,朱竹军,张培华.基于BP神经网络的循环流化床锅炉生成NO<sub>x</sub>质量浓度在线软测量[J].*热力发电*,2016,45(12):78-83.  
BAI Jian-yun,ZHU Zhu-jun,ZHANG Pei-hua. Online soft measurement of mass concentration of NO<sub>x</sub> generated by circulating fluidized bed boiler based on BP neural network [J]. *Thermal Power Generation*,2016,45(12):78-83.
- [23] 魏辉,陆方,罗永浩,等.燃煤锅炉高效、低NO<sub>x</sub>运行策略的研究[J].*动力工程*,2008(3):361-366.  
WEI Hui,LU Fang,LUO Yong-hao, et al. Research on high-efficiency and low NO<sub>x</sub> operation strategies of coal-fired boilers [J]. *Power Engineering*,2008(3):361-366.
- [24] 刘锁清,翟奕博,张耀华,等.动量改进BP神经网络在湿法脱硫系统中的应用[J].*自动化技术与应用*,2016,35(8):23-25.  
LIU Suo-qing,ZHAI Yi-bo,ZHANG Yao-hua, et al. Application of momentum-improved BP neural network in wet desulfurization system [J]. *Automation Technology and Application*, 2016, 35 (8): 23-25.
- [25] 牛培峰,麻红波,李国强,等.基于支持向量机和果蝇优化算法的循环流化床锅炉NO<sub>x</sub>排放特性研究[J].*动力工程学报*,2013,33(4):267-271.  
NIU Pei-feng,MA Hong-bo,LI Guo-qiang, et al. Research on NO<sub>x</sub> emission characteristics of circulating fluidized bed boiler based on support vector machine and Drosophila optimization algorithm [J]. *Journal of Chinese Society of Power Engineering*, 2013,33(4):267-271.
- [26] 陈丹丹.基于粗集的电站运行数据分析与运行优化[D].保定:华北电力大学,2013.  
CHEN Dan-dan. Power plant operation data analysis and operation optimization based on rough set [D]. Baoding:North China Electric Power University,2013.

- [27] 李建强. 基于数据挖掘的电站运行优化理论研究与应用[D]. 保定:华北电力大学,2006.  
LI Jian-qiang. Research and application of power plant operation optimization theory based on data mining[D]. Baoding: North China Electric Power University,2006.
- [28] 张仁杰. 粗糙集理论在电站运行优化中的应用研究[D]. 北京:华北电力大学,2011.  
ZHANG Ren-jie. Application research of rough set theory in power plant operation optimization[D]. Beijing: North China Electric Power University,2011.
- [29] 牛成林. 增量数据挖掘及其在电站运行优化中的理论研究及应用[D]. 北京:华北电力大学,2010.  
NIU Cheng-lin. Incremental data mining and its theoretical research and application in power plant operation optimization[D]. Beijing: North China Electric Power University,2010.
- [30] 万祥. 基于大数据挖掘技术的火电机组运行优化研究[D]. 武汉:武汉大学,2017.  
WAN Xiang. Research on optimization of thermal power unit operation based on big data mining technology[D]. Wuhan: Wuhan University,2017.
- [31] 钱瑾,王培红,李琳. 聚类算法在锅炉运行参数基准值分析中的应用[J]. 中国电机工程学报,2007(23):71-74.  
QIAN Jin, WANG Pei-hong, LI Lin. Application of clustering algorithm in analysis of boiler operating parameters reference value[J]. Proceedings of the CSEE,2007(23):71-74.
- [32] 冉鹏. 基于动态数据挖掘的电站热力系统运行优化方法研究[D]. 保定:华北电力大学,2012.  
RAN Peng. Research on optimization method of power plant thermal system operation based on dynamic data mining[D]. Baoding: North China Electric Power University,2012.
- [33] 周志成. 基于图像处理和人工智能的垃圾焚烧炉燃烧状态诊断研究[D]. 南京:东南大学,2015.  
ZHOU Zhi-cheng. Research on diagnosis of combustion state of waste incinerator based on image processing and artificial intelligence[D]. Nanjing: Southeast University,2015.
- [34] 姜明男,汪守康,何俊捷,等. 基于支持向量机的大型生活垃圾焚烧炉排炉运行参数预测[J]. 中国电机工程学报,2022,42(1):1-14.  
JIANG Ming-nan, WANG Shou-kang, HE Jun-jie, et al. Operation parameter prediction of large-scale domestic waste incineration grate furnace based on support vector machine[J]. Proceedings of the CSEE,2022,42(1):1-14.
- [35] CHANG N, CHANG Y, CHEN W C. Evaluation of heat value and its prediction for refuse-derived fuel[J]. Science of the Total Environment,1997,197(1-3):139-148.
- [36] TIAN W D, WEI X L, WU D Y, et al. Analysis of ingredient and heating value of municipal solid waste[J]. Journal of Environmental Sciences,2001,13:87-91.
- [37] KATHIRAVALE S, YUNUS M N M, SOPIAN K, et al. Modeling the heating value of municipal solid waste[J]. Fuel,2003,82(9):1119-1125.
- [38] 马晓茜,谢泽琼. 基于BP神经网络的垃圾热值预测模型[J]. 科技导报,2012,30(23):46-50.  
MA Xiao-qian, XIE Ze-qiong. A prediction model of garbage calorific value based on BP neural network[J]. Science & Technology Review,2012,30(23):46-50.
- [39] 应雨轩,林晓青,吴昂键,等. 生活垃圾智慧焚烧的研究现状及展望[J]. 化工学报,2021,72(2):886-900.  
YING Yu-xuan, LIN Xiao-qing, WU Ang-jian, et al. Research status and prospects of smart incineration of domestic waste[J]. CIESC Journal,2021,72(2):886-900.
- [40] 乔俊飞,郭子豪,汤健. 面向城市固废焚烧过程的二噁英排放浓度检测方法综述[J]. 自动化学报,2020,46(6):1063-1089.  
QIAO Jun-fei, GUO Zi-hao, TANG Jian. Summary of detection methods for dioxin emission concentration in the process of municipal solid waste incineration[J]. Acta Automatica Sinica,2020,46(6):1063-1089.
- [41] 肖晓东,卢加伟,海景,等. 垃圾焚烧烟气中二噁英类浓度的支持向量回归预测[J]. 可再生能源,2017,35(8):1107-1114.  
XIAO Xiao-dong, LU Jia-wei, HAI Jing, et al. Support vector regression prediction of dioxins concentration in waste incineration flue gas[J]. Renewable Energy,2017,35(8):1107-1114.
- [42] 乔俊飞,郭子豪,汤健. 基于多层特征选择的固废焚烧过程二噁英排放浓度软测量[J]. 信息与控制,2021,50(1):75-87.  
QIAO Jun-fei, GUO Zi-hao, TANG Jian. Soft measurement of dioxin emission concentration in solid waste incineration process based on multi-layer feature selection[J]. Information and Control,2021,50(1):75-87.