

基于潜在特征选择性集成建模的二噁英排放浓度软测量

汤健^{1,2} 乔俊飞^{1,2} 郭子豪^{1,2}

摘要 二噁英 (Dioxin, DXN) 是导致城市固废焚烧 (Municipal solid waste incineration, MSWI) 建厂存在“邻避现象”的主要原因之一。工业现场多采用离线化验手段检测 DXN 浓度, 难以满足污染物减排控制的需求。针对上述问题, 本文提出了基于潜在特征选择性集成 (Selective ensemble, SEN) 建模的 DXN 排放浓度软测量方法。首先, 采用主元分析 (Principal component analysis, PCA) 分别提取依据工艺阶段子系统及全流程系统过程变量的潜在特征, 并依据预设贡献率阈值进行特征初选; 接着, 采用互信息 (Mutual information, MI) 度量初选特征与 DXN 间的相关性, 并自适应确定再选的上下限及阈值; 最后, 采用具有超参数自适应选择机制的最小二乘-支持向量机 (Least squares — support vector machine, LS-SVM) 算法建立多源特征的候选子模型, 基于分支定界 (Branch and bound, BB) 优化和预测误差信息熵加权算法进行集成子模型的优化选择和加权组合, 进而得到软测量模型。基于某 MSWI 焚烧厂 DXN 检测数据仿真验证了所提方法的有效性。

关键词 城市固废焚烧, 二噁英, 多源潜在特征, 最小二乘-支持向量机, 选择性集成建模

引用格式 汤健, 乔俊飞, 郭子豪. 基于潜在特征选择性集成建模的二噁英排放浓度软测量. 自动化学报, 2022, 48(1): 223–238

DOI 10.16383/j.aas.c190254

Dioxin Emission Concentration Soft Measurement Based on Multi-source Latent Feature Selective Ensemble Modeling for Municipal Solid Waste Incineration Process

TANG Jian^{1,2} QIAO Jun-Fei^{1,2} GUO Zi-Hao^{1,2}

Abstract One of the main reasons leading to “not in my backyard (NIMBY)” of municipal solid waste incineration (MSWI) plant construction is dioxin (DXN) emission from such process, which is a highly toxic substance to the ecological environment. In practical industrial process, the DXN emission concentration is detected by off-line. It is difficult to meet the requirements of optimal control. Aim at the above problem, a new DXN emission concentration soft measurement approach based on multi-source latent feature selective ensemble (SEN) modeling is proposed. Firstly, MSWI process is divided into different subsystems according to industrial processes. Principal component analysis (PCA) was used to extract their latent features. Primary selection of these features is made based on empirical pre-set threshold of contribution rate. Then, mutual information (MI) is used to measure the correlation between these primary selected features and DXN. The upper and lower limits and thresholds for re-selected feature are adaptively determined. Finally, based on the re-selected feature, the least squares-support vector machine (LS-SVM) algorithm with hyper-parameter adaptive selection mechanism is used to construct sub-models. A strategy based on branch and bound (BB) and prediction error information entropy weighting algorithm is used to select sub-model and calculate the weight coefficient. Thus, an SEN soft sensing model is obtained. The proposed method is verified by using DXN detection data of MSWI process in Beijing.

Key words Municipal solid waste incineration (MSWI), dioxin (DXN), multi-source latent feature, least squares-support vector machine (LS-SVM), selective ensemble (SEN) modeling

Citation Tang Jian, Qiao Jun-Fei, Guo Zi-Hao. Dioxin emission concentration soft measurement based on multi-source latent feature selective ensemble modeling for municipal solid waste incineration process. *Acta Automatica Sinica*, 2022, 48(1): 223–238

收稿日期 2019-03-27 录用日期 2019-06-27

Manuscript received March 27, 2019; accepted June 27, 2019

国家自然科学基金 (62073006, 62021003), 北京市自然科学基金 (4212032, 4192009), 科学技术部国家重点研发计划 (2018YFC1900800-5), 矿冶过程自动控制技术国家 (北京市) 重点实验室 (BGRIMM-KZSKL-2020-02) 资助

Supported by National Natural Science Foundation of China (62073006, 62021003), Beijing Natural Science Foundation (4212032, 4192009), National Key Research and Development Program of the Ministry of Science and Technology (2018YFC1900800-5),

and Beijing Key Laboratory of Process Automation in Mining and Metallurgy (BGRIMM-KZSKL-2020-02)

本文责任编辑 刘艳军

Recommended by Associate Editor LIU Yan-Jun

1. 北京工业大学信息学部 北京 100124 2. 计算智能与智能系统北京市重点实验室 北京 100124

1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124 2. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124

如何基于运行优化控制策略降低复杂工业过程的能源消耗和污染排放, 是国内外流程工业企业所面临的急需解决难题^[1-3]. 焚烧是进行城市固废 (Municipal solid waste, MSW) 处理的主要技术手段^[4]. 对于发展中国家的 MSW 焚烧 (MSW incineration, MSWI) 企业, 最为紧迫的问题是如何降低焚烧造成的污染排放^[5-6], 其中急需在线监视和优化控制的是造成焚烧建厂“邻避现象 (Not in my backyard, NIMBY)” (指居民或单位因担心垃圾场、核电厂、殡仪馆类的建设项目对身体健康、环境质量和资产价值等带来负面影响而激发的嫌恶情结, 滋生“不要建在我家后院”的心理, 并采取强烈、坚决, 甚至高度情绪化的集体反对和抗争行为) 主要原因之一的剧毒物质二噁英 (Dioxin, DXN) 的排放浓度^[7-9]. MSWI 企业当前主要关注如何基于优化的运行参数实现 DXN 排放的最小化^[10]. 目前, 除配套先进的尾气处理装置外, 普遍采用“3T1E”准则间接控制 DXN 排放^[11], 即: 焚烧炉内高于 850 °C 的温度 (Temperature, T)、超过 2 秒 (Time, T) 的烟气停留时间、较大的湍流程度 (Turbulence, T) 和合适的过量空气系数 (Excess oxygen, E). 当前 MSWI 企业难以进行以降低 DXN 排放为直接目标的运行优化和反馈控制, 其主要原因是: 1) DXN 排放浓度的机理模型难以构建^[12]; 2) 以月或季为周期的离线直接检测焚烧尾气方式不能提供实时反馈的 DXN 排放浓度值^[13]. 近年来的研究热点是基于指示物/关联物对 DXN 排放进行在线间接检测^[14-15], 但 these 方法固有的设备复杂、造价昂贵、检测滞后等原因导致其难以用于 MSWI 过程的运行优化和反馈控制^[13].

数据驱动软测量技术可用于需要离线化验的难以检测参数 (如本文中的二噁英) 的在线估计^[16-17]. MSWI 过程包括固废焚烧、蒸汽发电、烟气处理及尾气排放等多个阶段, 其所包含的数百维过程变量间具有较大的冗余性与互补性; 显然, 这些不同阶段与 DXN 的产生、燃烧、吸收、再合成等过程相关, 有必要结合工艺过程划分为不同阶段子系统以便于能够量化表征对 DXN 排放浓度软测量的贡献度, 同时能够保留全部过程变量, 进而避免有用过程变量的缺失. 因此, 依据 MSWI 过程的特点, 可将 DXN 排放浓度软测量归结为一类面向小样本高维数据的建模问题. 文献 [18] 指出, 模型输入维数和低价值训练样本的增加使得获取完备训练样本的难度增大. 文献 [19] 定义了维数约简后的建模样本与约简特征之比, 指出该值应满足构建鲁棒学习模型的需

求. 因此, 针对 MSWI 过程具有小样本高维特性的 DXN 排放建模数据进行维数约简是必要的.

目前较为常用的方法是基于机理或经验通过特征选择实现维数约简. 以依据经验选择的部分过程变量为输入, 文献 [20-21] 通过采用多年前欧美研究机构所收集的少量样本, 基于线性回归、人工神经网络 (Artificial neural network, ANN) 等算法构建 DXN 排放浓度软测量模型. 近年来, 我国台湾地区针对实际焚烧过程, 首先初选部分过程变量, 再结合相关性分析和主元分析 (Principal component analysis, PCA) 进行特征选择, 最后基于 BP 神经网络 (Back propagation neural network, BPNN) 进行 DXN 排放浓度建模^[22]; 但 BPNN 具有易陷入局部最小、易过拟合和面向小样本数据建模泛化性能差等缺点. 理论上, 基于结构风险最小化准则的支持向量机 (Support vector machine, SVM) 算法能够有效建模小样本数据^[23-24], 但其需求解二次规划 (Quadratic programming, QP) 问题且超参数难以自适应选择. 最小二乘-支持向量机 (Least squares SVM, LS-SVM) 通过求解线性等式克服 QP 问题, 其超参数可通过优化算法得到^[25-27], 但耗时且只能得到次优解^[28]. 上述方法均以部分过程变量为输入构建传统单模型, 其泛化性有待于提高, 并且难以表征 MSWI 不同工艺阶段对 DXN 模型的贡献率. 此外, 上述研究也缺少对 LS-SVM 超参数的自适应选择机制.

针对工业过程机理复杂性难以有效地进行特征选择以及变量间具有的强共线性, 能够提取高维数据蕴含变化的主元分析 (PCA) 是工业过程难以检测参数软测量中较为常用的潜在特征提取方法^[29], 但贡献率低的主元建模会降低预测稳定性^[30]. 此外, 基于上述非监督方法所提取的蕴含原始过程变量主要变化的潜在特征与难测参数间的相关性却可能较弱. 因此, 有必要对贡献率满足要求的潜在变量进行再次选择.

此外, 针对 MSWI 过程的不同阶段子系统和全流程系统所提取的潜在特征可视为表征不同局部和全局特性的多源信息. 理论和经验分析表明, 面向多源信息采用选择性集成 (Selective ensemble, SEN) 机制构建的软测量模型具有更佳的稳定性和鲁棒性^[30]. 文献 [31] 综述了集成子模型多样性的构造策略, 指出训练样本重采样包括划分训练样本 (样本空间)、划分或变换特征变量 (特征空间) 等, 基于特征空间的集成构造策略在模型预测性能上较优. 针对小样本多源高维谱数据, Tang 等^[32] 提出基

于选择性融合多源特征和多工况样本的 SEN 潜结构映射模型. 文献 [32-33] 提出了基于随机采样样本空间的 SEN 神经网络模型和潜结构映射模型. 文献 [34] 提出基于子空间的集成学习通用框架. 文献 [35] 提出了在特征子空间内随机采样样本空间的面向多尺度机械信号的双层 SEN 潜结构映射模型. 文献 [36] 提出的 SEN 神经网络模型分别构建候选子模型和选择集成子模型及计算其权重. 但上述方法均未进行模型参数自适应机制的研究. 采用与文献 [20] 相同的建模数据, 文献 [37] 提出了基于候选超参数的 SEN 建模方法, 但该方法难以描述当前实际 MSWI 过程的真实特性, 并且难以有效表征 DXN 生成、燃烧、吸附和排放过程的多阶段特性.

综上所述, 依据 MSWI 过程的多阶段特性对不同阶段子系统进行非监督潜在特征的提取与度量, 构建具有自适应超参数选择和 SEN 机制的 DXN 排放浓度软测量模型的研究还未见报道. 因此, 本文所提的基于潜在特征 SEN 的 DXN 排放浓度软测量方法的创新点表现在: 1) 采用 PCA 提取依工艺流程划分阶段子系统和 MSWI 全流程系统的潜在特征, 并依据预设的主元贡献率阈值进行多源潜在特征初选, 保证预测稳定性和避免特征选择不当造成的信息损失; 2) 采用互信息 (Mutual information, MI) 度量初选潜在特征并进行选择以保证再选潜在特征与 DXN 间的相关性, 自适应确定多源

潜在特征再选的上下限及阈值; 3) 采用具有超参数自适应选择机制的 LS-SVM 算法和自适应确定集成子模型尺寸、集成子模型及其加权系数的 SEN 机制构建 DXN 排放浓度软测量模型, 确保具有互补特性的最佳子系统能够选择性融合.

1 面向 DXN 排放过程的 MSWI 描述

MSWI 的主要设备包括焚烧炉、移动炉排、废锅和尾气处理等设备, 其中, 焚烧炉将 MSW 转化为残渣、灰尘、烟气与热量, 位于焚烧炉底部的移动炉排促使 MSW 有效和完全燃烧, 废锅产生的蒸汽用于推动汽轮机产生电力, 烟气中的灰尘和污染物通过尾气处理设备净化后排入大气. 其过程如图 1 所示.

由图 1 可知, MSWI 过程包含 DXN 生成、燃烧、吸附和排放等阶段, G1、G2 和 G3 等不同阶段的烟气中所包含的 DXN 浓度具有差异性. 通常, 为保证焚烧炉内的有害物质能够有效和完全分解, 烟气温度的应该达到 $850\text{ }^{\circ}\text{C}$ 并保持 2 s 以上. 用于 MSW 焚烧的一次风从炉排底部喷入, 并通过引入湍流和保证过量氧的供应使得二次风能够辅助进行烟气的完全燃烧. 在烟气冷却过程中, 进行焚烧矿渣和废锅底灰的处置和收集后排出烟气 (G1). 活性炭和石灰被注入反应器, 用于移除酸性气体和吸收 DXN 及一些重金属, 烟气然后再进入袋式过滤器. 在反应器和袋式过滤器内产生的飞灰被注入混湿设备后排放烟气 (G2). 引风机将烟气 (G2) 吸入到烟囱, 烟囱

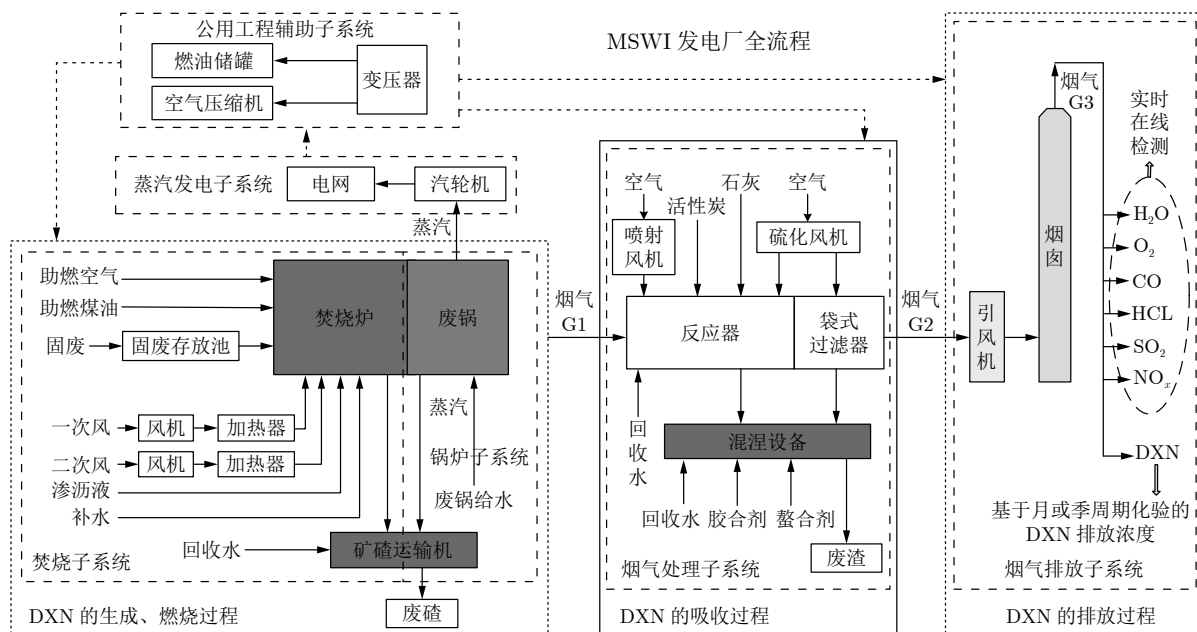


图 1 基于 DXN 视角的 MSWI 过程描述

Fig. 1 MSWI process description based on DXN perspective

产生排放至大气中的烟气 (G3), 其包含 HCL、SO₂、NO_x 和 HF 等多种能够实时在线检测的污染物浓度, 以及具有长周期、高成本等特点需离线才能化验的 DXN 浓度. 由上述描述可知, DXN 排放浓度与 MSWI 过程不同阶段的易检测过程变量均具有相关性. 由图 1 可知, MSWI 过程按工艺流程至少可分为焚烧、锅炉、烟气处理、蒸汽发电、烟气排放、公用工程辅助共 6 个阶段子系统. 对于 DXN 排放浓度建模, 本文将不同的阶段子系统及全流程系统均视为多源信息.

2 建模策略

依据上述分析, 本文提一种基于潜在特征 SEN 建模的 DXN 排放浓度软测量方法, 包括广义子系统划分模块、潜在特征提取与初选模块、潜在特征度量与再选模块、自适应选择性集成建模模块, 如图 2 所示.

在图 2 中, $X^i \in \mathbf{R}^{N \times M^i}$ 表示从第 i 个子系统所采集的全部过程变量; $Z_{\text{FeAll}}^i \in \mathbf{R}^{N \times M_{\text{FeAll}}^i}$ 表示针对第 i 个子系统的全部过程变量, 采用 PCA 提取的数量为 M_{FeAll}^i 的全部潜在特征; $Z_{\text{FeSe1st}}^i \in \mathbf{R}^{N \times M_{\text{FeSe1st}}^i}$ 表示针对第 i 个子系统的全部潜在特征, 依据设定阈值 θ_{Contri} 选择的数量为 M_{FeSe1st}^i 的初选潜在特征; $Z_{\text{FeSe2nd}}^i \in \mathbf{R}^{N \times M_{\text{FeSe2nd}}^i}$ 表示对第 i 个子系统的初选潜在特征与 DXN 间的相关性采用 MI 度量后, 基于

阈值选择的数量为 M_{FeSe2nd}^i 的再选潜在特征; K_{er}^i 和 R_{eg}^i 表示为第 i 个子模型所选择的核参数和正则化参数, 即超参数对, 本文将记为 $\{K_{\text{er}}^i, R_{\text{eg}}^i\}$; \hat{y}^i 表示第 i 个子模型的预测输出; y 和 \hat{y} 表示 DXN 排放浓度软测量模型的真值和预测输出.

上述模块的功能是: 1) 广义子系统划分模块. 基于工艺流程或经验知识将过程变量分组为蕴含不同局部信息的阶段子系统, 并将全流程系统作为包含全局信息的广义子系统. 2) 潜在特征提取与初选模块. 采用 PCA 提取不同阶段子系统和 MSWI 全流程系统所包含过程变量的全部潜在特征, 基于依据经验设定的潜在特征贡献率阈值获得多源初选潜在特征, 其目的是防止采用较小贡献率的潜在特征建模造成预测性能的不稳定. 3) 潜在特征度量与再选模块. 采用 MI 度量初选潜在特征与 DXN 排放浓度间的关系, 并结合软测量模型预测性能自适应确定再选潜在特征, 其目的是使得所选的多源潜在特征与 DXN 间具有较好的映射关系. 4) 自适应选择性集成建模模块. 采用超参数自适应选择策略构建基于不同广义子系统再选潜在特征的最佳预测性能子模型, 结合分支定界和预测误差信息熵加权算法自适应地选择子模型和计算其加权系数, 其目的是选择具有较好冗余与互补特性子模型进行融合, 以提高 SEN 软测量模型的预测性能. 本文中采用的公式符号及其说明如表 1 所示.

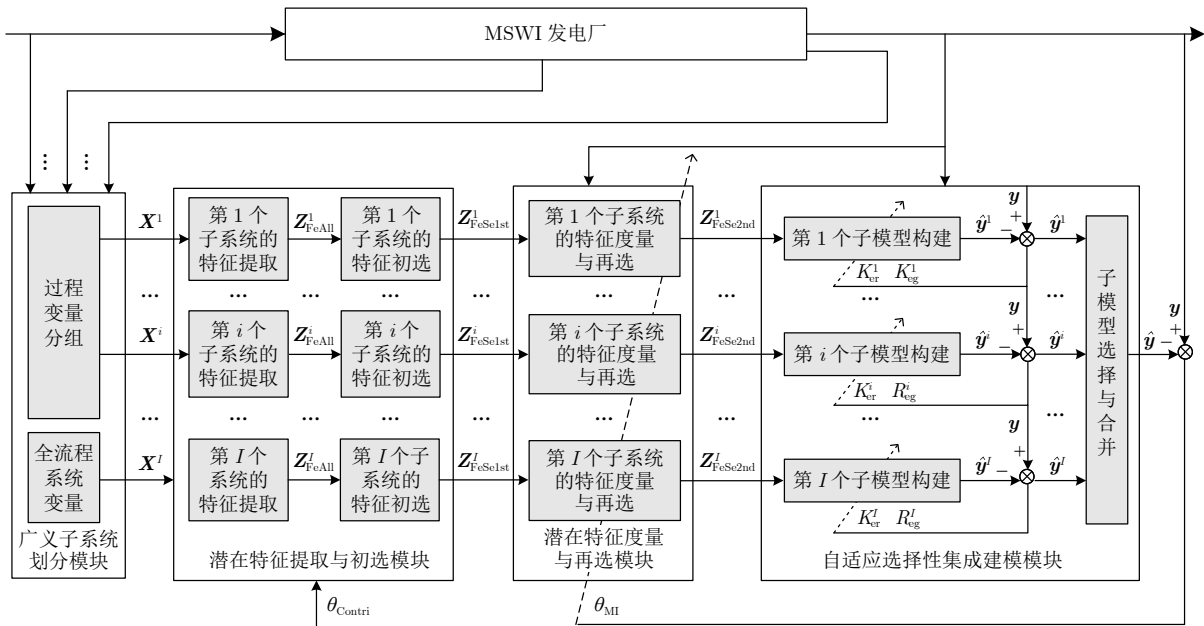


图 2 基于潜在特征 SEN 建模的 DXN 排放浓度软测量策略

Fig. 2 Soft sensing strategy of DXN emission concentration based on latent feature SEN modeling

表 1 本文中的公式符号及其说明汇总表
Table 1 Summary of formula symbols and their explanations in this paper

符号	含义	符号	含义
\mathbf{y}	DXN 排放浓度软测量模型的真值	$\hat{\mathbf{y}}$	DXN 排放浓度软测量模型的预测输出
N	建模样本数量	M	输入过程变量数量
\mathbf{X}	MSWI 全流程系统的输入数据	\mathbf{X}^i	第 i 个子系统的输入数据
$I - 1$	MSWI 全流程系统划分子系统个数	M^i	第 i 个子系统包含的过程变量个数
$\mathbf{Z}_{\text{FeAll}}^i$	第 i 个子系统的过程变量采用 PCA 提取的全部潜在特征	M_{FeAll}^i	第 i 个子系统的过程变量采用 PCA 提取的全部潜在特征的数量
$\mathbf{Z}_{\text{FeSe1st}}^i$	第 i 个子系统的初选潜在特征	θ_{Contri}	对全部潜在特征进行初选的设置阈值
M_{FeSe1st}^i	第 i 个子系统初选潜在特征的数量	M_{FeSe2nd}^i	第 i 个子系统再选潜在特征的数量
$\mathbf{Z}_{\text{FeSe2nd}}^i$	第 i 个子系统的再选潜在特征	θ_{MI}	再选潜在特征的选择阈值 θ_{MI}
$(K_{\text{er}}^i, R_{\text{eg}}^i)$	第 i 个子模型的核参数和正则化参数, 即超参数对	i	第 i 个子模型的预测输出
$\mathbf{t}_{m_{\text{FeAll}}^i}^i$	第 i 个子系统的第 m_{FeAll}^i 个主元的得分向量	$\mathbf{p}_{m_{\text{FeAll}}^i}^i$	第 i 个子系统的第 m_{FeAll}^i 个主元的载荷向量
\mathbf{T}^i	第 i 个子系统的得分矩阵	\mathbf{P}^i	第 i 个子系统的载荷矩阵
$\lambda_{m_{\text{FeAll}}^i}^i$	第 i 个子系统的第 m_{FeAll}^i 个载荷向量 $\mathbf{p}_{m_{\text{FeAll}}^i}^i$ 相对应的特征值	$\theta_{m_{\text{FeAll}}^i}^i$	第 i 个子系统的第 m_{FeAll}^i 个潜在特征的贡献率
$\xi_{m_{\text{FeAll}}^i}^i$	第 i 个子系统的第 m_{FeAll}^i 个潜在特征是否被选中的标记值	ξ_{MI}^i	第 i 个子系统的初选潜在特征 $\mathbf{z}_{m_{\text{FeSe1st}}^i}^i$ 与 DXN 排放浓度间的 MI 值
$\theta_{\text{Contri}}^{\text{Uplimit}}$	潜在特征再选阈值的上限值	$\theta_{\text{Contri}}^{\text{Downlimit}}$	潜在特征再选阈值的下限值
$\theta_{\text{Contri}}^{\text{Step}}$	潜在特征再选阈值的固定步长	$\beta_{m_{\text{FeSe1st}}^i}^i$	第 i 个子系统的第 m_{FeSe1st}^i 个初选潜在特征是否被选中的标记值
\mathbf{w}^i	第 i 个子模型的权重系数	b^i	第 i 个子模型的偏置系数
β^i	第 i 个子模型的拉格朗日算子向量	ζ^i	第 i 个子模型的预测误差向量
M_{para}	候选超参数矩阵	$\{K_{\text{er}}^i, R_{\text{eg}}^i\}$	第 i 个子模型在 M_{para} 中自适应选择的超参数对
K	候选核参数数量	R	候选惩罚参数数量
$J = K \times R$	超参数矩阵中的超参数对的数量	$\{(K_{\text{er}}^{\text{initial}})^i, (R_{\text{eg}}^{\text{initial}})^i\}$	第 i 个子模型在采用网格搜索策略在矩阵 M_{para} 中初选的超参数对
$(K_{\text{er}}^{\text{vector}})^i$	依据初选超参数对计算的新候选核参数向量	$(R_{\text{eg}}^{\text{vector}})^i$	依据初选超参数对计算的新候选惩罚参数向量
N_{ker}	新候选核参数的数量	N_{reg}	新候选惩罚参数的数量
$k_{\text{supara}}^{\text{down}}, k_{\text{supara}}^{\text{up}}$	确定超参数向量的收缩和扩张因子	$f^i(\cdot)$	第 i 个子模型
$f^i_{\text{sel}}(\cdot)$	第 i_{sel} 个集成子模型	$w_{i_{\text{sel}}}$	第 i_{sel} 个集成子模型的加权系数
$\hat{y}_{i_{\text{sel}}}$	第 i_{sel} 个集成子模型的预测值	$K_{\text{er}}^{i_{\text{sel}}}, R_{\text{eg}}^{i_{\text{sel}}}$	第 i_{sel} 个集成子模型的超参数
$(\hat{y}_{i_{\text{sel}}})_n$	第 n 个样本基于第 i_{sel} 个集成子模型的预测值	$(e_{i_{\text{sel}}})_n$	第 n 个样本基于第 i_{sel} 个集成子模型的相对预测误差
$E_{i_{\text{sel}}}$	第 i_{sel} 个集成子模型的预测误差信息熵		

3 建模实现

3.1 广义子系统划分模块

本文中, 模型输入数据 $\mathbf{X} \in \mathbf{R}^{N \times M}$ 包括 N 个样本 (行) 和 M 个变量 (列), 其源于 MSWI 过程数量为 $(I - 1)$ 的子系统. 为表征 MSWI 全流程系统所蕴含的全局信息, 将其视为广义上的第 I 个子系统.

若将第 i 个子系统的建模数据表示为 $\mathbf{X}^i \in \mathbf{R}^{N \times M^i}$, 即存在如下关系:

$$\mathbf{X} = [\mathbf{X}^1, \dots, \mathbf{X}^i, \dots, \mathbf{X}^I] = \{\mathbf{X}^i\}_{i=1}^{I-1} \quad (1)$$

$$M = M^1 + \dots + M^i + \dots + M^{I-1} = \sum_{i=1}^{I-1} M^i \quad (2)$$

其中, I 表示子系统个数, M^i 表示第 i 个子系统包含的变量个数.

相应地, 输出数据 $\mathbf{y} = \{y_n\}_{n=1}^N$ 包括 N 个样本 (行), 其来源于离线化验的 DXN 排放浓度检测数据. 显然, 输入/输出数据在时间尺度上具有较大的

差异性: 过程变量以秒为单位在 DCS 系统采集与存储, DXN 排放浓度以月/季为周期离线化验获得, 故存在 $N \ll M$.

3.2 潜在特征提取与初选模块

以第 i 个子系统为例, 首先采用 PCA 提取输入变量的潜在特征. 将输入数据 \mathbf{X}^i 进行零均值、1 方差的标准化处理后, 将其分解为

$$\mathbf{X}^i = \mathbf{t}_{1_{\text{FeAll}}}^i (\mathbf{p}_{1_{\text{FeAll}}}^i)^T + \cdots + \mathbf{t}_{m_{\text{FeAll}}}^i (\mathbf{p}_{m_{\text{FeAll}}}^i)^T + \cdots + \mathbf{t}_{M_{\text{FeAll}}}^i (\mathbf{p}_{M_{\text{FeAll}}}^i)^T \quad (3)$$

其中, $\mathbf{t}_{m_{\text{FeAll}}}^i$ 和 $\mathbf{p}_{m_{\text{FeAll}}}^i$ 表示第 m_{FeAll}^i 个主元 (Principal component, PC) 的得分和载荷向量, M_{FeAll}^i 表示对第 i 个子系统所提取潜在特征数量, 其计算式为

$$M_{\text{FeAll}}^i = \text{rank}(\mathbf{X}^i) \quad (4)$$

基于式 (3), 从数据 \mathbf{X}^i 所提取的全部潜在特征可表示为

$$\mathbf{T}^i = [\mathbf{t}_{1_{\text{FeAll}}}^i, \cdots, \mathbf{t}_{m_{\text{FeAll}}}^i, \cdots, \mathbf{t}_{M_{\text{FeAll}}}^i] \quad (5)$$

其中, $\mathbf{T}^i \in \mathbf{R}^{N \times M_{\text{FeAll}}^i}$ 表示得分矩阵, 是输入数据 \mathbf{X}^i 在载荷矩阵 \mathbf{P}^i 方向上的正交映射; \mathbf{P}^i 采用下式表示

$$\mathbf{P}^i = [\mathbf{p}_{1_{\text{FeAll}}}^i, \cdots, \mathbf{p}_{m_{\text{FeAll}}}^i, \cdots, \mathbf{p}_{M_{\text{FeAll}}}^i] \quad (6)$$

其中, $\mathbf{P}^i \in \mathbf{R}^{M \times M_{\text{FeAll}}^i}$.

进而, 从数据 \mathbf{X}^i 所提取的潜在特征可表示为

$$\mathbf{Z}_{\text{FeAll}}^i = [\mathbf{Z}_{\text{FeAll}}^1, \cdots, \mathbf{Z}_{\text{FeAll}}^i, \cdots, \mathbf{Z}_{\text{FeAll}}^I] = \{\mathbf{Z}_{\text{FeAll}}^i\}_{i=1}^I \quad (7)$$

其中, $\mathbf{Z}_{\text{FeAll}}^i \in \mathbf{R}^{N \times M_{\text{FeAll}}^i}$.

进一步, 全部潜在特征可表示为

$$\mathbf{Z}_{\text{FeAll}} = [\mathbf{Z}_{\text{FeAll}}^1, \cdots, \mathbf{Z}_{\text{FeAll}}^i, \cdots, \mathbf{Z}_{\text{FeAll}}^I] = \{\mathbf{Z}_{\text{FeAll}}^i\}_{i=1}^I \quad (8)$$

研究表明, 采用贡献率较小的潜在特征建模会导致模型预测性能的不稳定. 此处, 将与第 m_{FeAll}^i 个载荷向量 $\mathbf{p}_{m_{\text{FeAll}}}^i$ 对应的特征值记为 $\lambda_{m_{\text{FeAll}}}^i$, 相应的第 m_{FeAll}^i 个潜在特征 $\mathbf{z}_{m_{\text{FeAll}}}^i$ 的贡献率 $\theta_{m_{\text{FeAll}}}^i$, 可采用下式计算

$$\theta_{m_{\text{FeAll}}}^i = \frac{\lambda_{m_{\text{FeAll}}}^i}{\sum_{m_{\text{FeAll}}^i=1}^{M_{\text{FeAll}}^i} \lambda_{m_{\text{FeAll}}}^i} \times 100\% \quad (9)$$

将依据经验选择的阈值记为 θ_{Contri} (其默认取值为 1), 采用如下规则对全部潜在特征进行初次选择,

$$\xi_{m_{\text{FeAll}}}^i = \begin{cases} 1, & \theta_{m_{\text{FeAll}}}^i \geq \theta_{\text{Contri}} \\ 0, & \theta_{m_{\text{FeAll}}}^i < \theta_{\text{Contri}} \end{cases} \quad (10)$$

其中, $\xi_{m_{\text{FeAll}}}^i$ 表示第 m_{FeAll}^i 个潜在特征是否被选中的标记值, 其值为 1 表示该潜在特征被初次选中.

进而, 将针对第系统的初选潜在特征表示为

$$\mathbf{Z}_{\text{FeSelst}}^i = [\mathbf{z}_{1_{\text{FeSelst}}}^i, \cdots, \mathbf{z}_{m_{\text{FeSelst}}}^i, \cdots, \mathbf{z}_{M_{\text{FeSelst}}}^i] = \left[\left\{ \left(\mathbf{z}_{1_{\text{FeSelst}}}^i \right)_n \right\}_{n=1}^N, \cdots, \left\{ \left(\mathbf{z}_{m_{\text{FeSelst}}}^i \right)_n \right\}_{n=1}^N, \cdots, \left\{ \left(\mathbf{z}_{M_{\text{FeSelst}}}^i \right)_n \right\}_{n=1}^N \right] = \left\{ \left(\mathbf{z}_{\text{FeSelst}}^i \right)_n \right\}_{n=1}^N \quad (11)$$

进一步, 全部初选潜在特征 $\mathbf{Z}_{\text{FeSelst}}$ 可表示为

$$\mathbf{Z}_{\text{FeSelst}} = [\mathbf{Z}_{\text{FeSelst}}^1, \cdots, \mathbf{Z}_{\text{FeSelst}}^i, \cdots, \mathbf{Z}_{\text{FeSelst}}^I] = \left\{ \mathbf{Z}_{\text{FeSelst}}^i \right\}_{i=1}^I \quad (12)$$

3.3 潜在特征度量与再选模块

采用非监督方式提取的初选潜在特征相互独立, 但却未考虑这些特征与 DXN 间的相关性. 仍以第 i 个子系统为例, 将初选潜在特征 $\mathbf{z}_{m_{\text{FeSelst}}}^i$ 与 DXN 间的互信息 (MI) 值记为 $\xi_{\text{MI}}^{m_{\text{FeSelst}}^i}$, 计算式为

$$\xi_{\text{MI}}^{m_{\text{FeSelst}}^i} = p_{\text{prob}}(\mathbf{z}_{m_{\text{FeSelst}}}^i, \mathbf{y}) \lg \left(\frac{p_{\text{prob}}(\mathbf{z}_{m_{\text{FeSelst}}}^i, \mathbf{y})}{p_{\text{prob}}(\mathbf{z}_{m_{\text{FeSelst}}}^i) p_{\text{prob}}(\mathbf{y})} \right) = H(\mathbf{y}) - H(\mathbf{y} | \mathbf{z}_{m_{\text{FeSelst}}}^i) \quad (13)$$

其中, $p_{\text{prob}}(\mathbf{z}_{m_{\text{FeSelst}}}^i)$ 和 $p_{\text{prob}}(\mathbf{y})$ 表示 $\mathbf{z}_{m_{\text{FeSelst}}}^i$ 和 \mathbf{y} 的边缘概率密度, $p_{\text{prob}}(\mathbf{z}_{m_{\text{FeSelst}}}^i, \mathbf{y})$ 表示联合概率密度, $H(\mathbf{y} | \mathbf{z}_{m_{\text{FeSelst}}}^i)$ 表示条件熵, $H(\mathbf{y})$ 表示信息熵.

此处, 将依据软测量模型的预测性能确定阈值, 其上限值 $\theta_{\text{Contri}}^{\text{Uplimit}}$ 、下限值 $\theta_{\text{Contri}}^{\text{Downlimit}}$ 及固定步长 $\theta_{\text{Contri}}^{\text{Step}}$ 的计算式见式 (14)~(16):

$$\theta_{\text{Contri}}^{\text{Uplimit}} = \min \left(\max \left(\xi_{\text{MI}}^{m_{\text{FeSelst}}^1}, \cdots, \max \left(\xi_{\text{MI}}^{m_{\text{FeSelst}}^i}, \cdots, \max \left(\xi_{\text{MI}}^{m_{\text{FeSelst}}^I} \right) \right) \right) \right) \quad (14)$$

$$\theta_{\text{Contri}}^{\text{Downlimit}} = \max \left(\min \left(\xi_{\text{MI}}^{m_{\text{FeSelst}}^1}, \cdots, \min \left(\xi_{\text{MI}}^{m_{\text{FeSelst}}^i}, \cdots, \min \left(\xi_{\text{MI}}^{m_{\text{FeSelst}}^I} \right) \right) \right) \right) \quad (15)$$

$$\theta_{\text{Contri}}^{\text{Step}} = \frac{\theta_{\text{Contri}}^{\text{Uplimit}} - \theta_{\text{Contri}}^{\text{Downlimit}}}{N_{\text{Contri}}^{\text{Step}}} \quad (16)$$

式中, 函数 $\max(\cdot)$ 和 $\min(\cdot)$ 分别表示取最大值和最小值, $N_{\text{Contri}}^{\text{Step}}$ 表示依据经验确定的候选阈值数量 (其默认值为 10).

将选定阈值记为 θ_{Contri} , 其值在 $\theta_{\text{Contri}}^{\text{Uplimit}}$ 和 $\theta_{\text{Contri}}^{\text{Downlimit}}$ 之间依据预测性能进行自适应选择. 采用如下规则对初选潜在特征进行再选:

$$\beta_{m_{\text{FeSe1st}}^i} = \begin{cases} 1, & \xi_{\text{MI}}^{m_{\text{FeSe1st}}^i} \geq \theta_{\text{Contri}} \\ 0, & \xi_{\text{MI}}^{m_{\text{FeSe1st}}^i} < \theta_{\text{Contri}} \end{cases} \quad (17)$$

其中, $\beta_{m_{\text{FeSe1st}}^i}$ 表示第 m_{FeSe1st}^i 个初选潜在特征是否被选中的标记值, 其值为 1 表示该潜在特征被再次选中.

进一步, 第 i 个系统的再选潜在特征表示为

$$\mathbf{Z}_{\text{FeSe2nd}}^i = \left[\mathbf{z}_{1_{\text{FeSe2nd}}}^i, \dots, \mathbf{z}_{m_{\text{FeSe2nd}}^i}^i, \dots, \mathbf{z}_{M_{\text{FeSe2nd}}^i}^i \right] = \left[\left\{ \left(\mathbf{z}_{1_{\text{FeSe2nd}}}^i \right)_n \right\}_{n=1}^N, \dots, \left\{ \left(\mathbf{z}_{m_{\text{FeSe2nd}}^i}^i \right)_n \right\}_{n=1}^N, \dots, \left\{ \left(\mathbf{z}_{M_{\text{FeSe2nd}}^i}^i \right)_n \right\}_{n=1}^N \right] = \left\{ \left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_n \right\}_{n=1}^N \quad (18)$$

因此, 全部再选潜在特征 $\mathbf{Z}_{\text{FeSe2nd}}$ 可表示为

$$\mathbf{Z}_{\text{FeSe2nd}} = \left[\mathbf{Z}_{\text{FeSe2nd}}^1, \dots, \mathbf{Z}_{\text{FeSe2nd}}^i, \dots, \mathbf{Z}_{\text{FeSe2nd}}^I \right] = \left\{ \mathbf{Z}_{\text{FeSe2nd}}^I \right\}_{i=1}^I \quad (19)$$

3.4 自适应选择性集成建模模块

以第 i 个子系统为例, 描述基于再选潜在特征 $\mathbf{Z}_{\text{FeSe2nd}}^i$ 和超参数对 $\{K_{\text{er}}^i, R_{\text{eg}}^i\}$ 构建 DXN 子模型的过程. 首先, 将 $\left\{ \left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_n \right\}_{n=1}^N$ 通过映射 $\varphi(\cdot)$ 变换到高维特征空间, 求解如下优化问题:

$$\begin{cases} \min_{\mathbf{w}^i, b^i} O_{\text{LS-SVM}} = \frac{1}{2} (\mathbf{w}^i)^T \mathbf{w}^i + \frac{1}{2} R_{\text{eg}}^i \sum_{n=1}^N (\zeta_n^i)^2 \\ \text{s.t. } \hat{y}_n^i = (\mathbf{w}^i)^T \varphi \left(\left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_n \right) + b^i + \zeta_n^i \end{cases} \quad (20)$$

其中, \mathbf{w}^i 表示权重系数, b^i 表示偏置, ζ_n^i 是第 n 个样本的预测误差.

采用拉格朗日方法可得到

$$L^i(\mathbf{w}^i, b^i, \zeta^i, \beta^i) = \frac{1}{2} (\mathbf{w}^i)^T \mathbf{w}^i + \frac{1}{2} \sum_{n=1}^N (\zeta_n^i)^2 - \sum_{n=1}^N \beta_n^i \left[(\mathbf{w}^i)^T \varphi \left(\left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_n \right) + b^i + \zeta_n^i - \hat{y}_n^i \right] \quad (21)$$

其中, $\beta^i = [\beta_1^i, \dots, \beta_n^i, \dots, \beta_N^i]$ 表示拉格朗日算子向量, $\zeta^i = [\zeta_1^i, \dots, \zeta_n^i, \dots, \zeta_N^i]$ 表示预测误差向量.

对上述公式进行求解,

$$\begin{cases} \frac{\partial L^i}{\partial \mathbf{w}^i} = 0 \\ \frac{\partial L^i}{\partial b^i} = 0 \\ \frac{\partial L^i}{\partial \zeta^i} = 0 \\ \frac{\partial L^i}{\partial \beta^i} = 0 \end{cases} \quad (22)$$

将所采用的核函数表示为

$$\Omega_{\text{ker}}^i \left(\mathbf{z}_{\text{FeSe2nd}}^i, \left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_n \right) = \left\langle \varphi \left(\mathbf{z}_{\text{FeSe2nd}}^i \right) \times \varphi \left(\left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_n \right) \right\rangle \quad (23)$$

进一步, 将 LS-SVM 问题转换为求解以下线性系统 (24) (见本页下方).

通过求解式 (24), 得到 β^i 和 b^i .

进而, 基于 LS-SVM 构建的子模型可表示为

$$\hat{y}^i = \sum_{n=1}^N \beta_n^i \cdot \Omega_{\text{ker}}^i \left(\mathbf{z}_{\text{FeSe2nd}}^i, \left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_n \right) + b^i \quad (25)$$

上述排放浓度子模型的超参数自适应选择机制采用下述的两步法实现.

步骤 1. 采用网格搜索策略以子模型的预测性能为目标函数, 在候选超参数矩阵 M_{para} 中自适应选择初始超参数对 $\{(K_{\text{er}}^{\text{initial}})^i, (R_{\text{eg}}^{\text{initial}})^i\}$. 其中, 候选超参数矩阵 M_{para} 如式 (26) 所示.

$$\begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & \Omega_{\text{ker}}^i \left(\left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_1, \left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_1 \right) + \frac{1}{R_{\text{eg}}^i} & \dots & \Omega_{\text{ker}}^i \left(\left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_1, \left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_N \right) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \Omega_{\text{ker}}^i \left(\left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_N, \left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_1 \right) & \dots & \Omega_{\text{ker}}^i \left(\left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_N, \left(\mathbf{z}_{\text{FeSe2nd}}^i \right)_N \right) + \frac{1}{R_{\text{eg}}^i} \end{bmatrix} \times \begin{bmatrix} b^i \\ \beta_1^i \\ \vdots \\ \beta_N^i \end{bmatrix} = \begin{bmatrix} 1 \\ y_1^i \\ \vdots \\ y_N^i \end{bmatrix} \quad (24)$$

$$M_{\text{para}} = \begin{bmatrix} [K_{\text{er}}^1, R_{\text{eg}}^1] & \cdots & [K_{\text{er}}^1, R_{\text{eg}}^r] & \cdots & [K_{\text{er}}^1, R_{\text{eg}}^R] \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ [K_{\text{er}}^k, R_{\text{eg}}^1] & \cdots & [K_{\text{er}}^k, R_{\text{eg}}^r] & \cdots & [K_{\text{er}}^k, R_{\text{eg}}^R] \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ [K_{\text{er}}^K, R_{\text{eg}}^1] & \cdots & [K_{\text{er}}^K, R_{\text{eg}}^r] & \cdots & [K_{\text{er}}^K, R_{\text{eg}}^R] \end{bmatrix}_{K \times R} \quad (26)$$

其中, $k = 1, \dots, K$, K 表示候选核参数的数量 $r = 1, \dots, R$, R 表示候选惩罚参数的数量; $[K_{\text{er}}^k, R_{\text{eg}}^r]$ 表示由第 k 个核参数和第 r 个惩罚参数组成的超参数对, 为 M_{para} 中的第 j 个元素, 即存在的对应关系为 $M_{\text{para}}^j = [K_{\text{er}}^k, R_{\text{eg}}^r]$; $j = 1, \dots, J$, $J = K \times R$ 表示候选超参数对的数量. 可见, 初次采用网格搜索策略为第 i 个子模型所选择的超参数对 $\{(K_{\text{er}}^{\text{initial}})^i, (R_{\text{eg}}^{\text{initial}})^i\}$ 是 M_{para} 中一个元素, 即存在 $\{(K_{\text{er}}^{\text{initial}})^i, (R_{\text{eg}}^{\text{initial}})^i\} \in M_{\text{para}}$.

步骤 2. 基于 $\{(K_{\text{er}}^{\text{initial}})^i, (R_{\text{eg}}^{\text{initial}})^i\}$, 采用式 (27) 和式 (28) 获得新的候选核参数向量 $(\mathbf{K}_{\text{er}}^{\text{vector}})^i$ 和惩罚参数向量 $(\mathbf{R}_{\text{eg}}^{\text{vector}})^i$

$$\begin{aligned} (\mathbf{K}_{\text{er}}^{\text{vector}})^i &= (K_{\text{er}}^{\text{initial}})^i / k_{\text{supara}}^{\text{down}} : \\ & \frac{k_{\text{supara}}^{\text{up}} \times (K_{\text{er}}^{\text{initial}})^i - (K_{\text{er}}^{\text{initial}})^i / k_{\text{supara}}^{\text{down}}}{N_{\text{ker}}} : \\ & k_{\text{supara}}^{\text{up}} \times (K_{\text{er}}^{\text{initial}})^i \end{aligned} \quad (27)$$

$$\begin{aligned} (\mathbf{R}_{\text{eg}}^{\text{vector}})^i &= (R_{\text{eg}}^{\text{initial}})^i / k_{\text{supara}}^{\text{down}} : \\ & \frac{k_{\text{supara}}^{\text{up}} \times (R_{\text{eg}}^{\text{initial}})^i - (R_{\text{eg}}^{\text{initial}})^i / k_{\text{supara}}^{\text{down}}}{N_{\text{ker}}} : \\ & k_{\text{supara}}^{\text{up}} \times (R_{\text{eg}}^{\text{initial}})^i \end{aligned} \quad (28)$$

其中, N_{ker} 和 N_{reg} 表示设定的新的候选核参数和惩罚参数的数量; $k_{\text{supara}}^{\text{down}}$ 和 $k_{\text{supara}}^{\text{up}}$ 为设定的超参数收缩和扩张因子 (默认值均为 10).

再次采用网格搜索策略获得第 i 个子模型的超参数对 $\{K_{\text{er}}^i, R_{\text{eg}}^i\}$.

对全部子系统及 MSWI 全流程系统执行上述过程, 所构建的候选子模型预测输出的集合可表示为

$$\begin{aligned} \hat{Y} &= [\hat{y}^1, \dots, \hat{y}^i, \dots, \hat{y}^I] = \{\hat{y}^i\}_{i=1}^I = \\ & [f^1(K_{\text{er}}^1, R_{\text{eg}}^1, \mathbf{Z}_{\text{FeSe2nd}}^1), \dots, \\ & f^i(K_{\text{er}}^i, R_{\text{eg}}^i, \mathbf{Z}_{\text{FeSe2nd}}^i), \dots, \\ & f^I(K_{\text{er}}^I, R_{\text{eg}}^I, \mathbf{Z}_{\text{FeSe2nd}}^I)] = \\ & \{f^i(K_{\text{er}}^i, R_{\text{eg}}^i, \mathbf{Z}_{\text{FeSe2nd}}^i)\}_{i=1}^I \end{aligned} \quad (29)$$

其中, $f^i(\cdot)$ 表示第 i 个子模型.

结合基于分支定界的最优化选择算法和基于预测误差的信息熵加权算法对上述候选子模型进行自适应的选择和加权. 在给定候选子模型和加权算法后, 集成子模型的选择类似最优组合特征选择^[30]. 面向数量有限的候选子模型, 通常采用的策略为: 多次耦合运行优化与加权算法构建集成尺寸为 $2 \sim (I - 1)$ 的多个 SEN 模型, 基于预测性能排序得到优选的 SEN 模型. 假定本文最终软测量模型的集成尺寸为 I_{sel} , 其预测输出值 \hat{y} 为

$$\hat{y} = \sum_{i_{\text{sel}}=1}^{I_{\text{sel}}} w_{i_{\text{sel}}} \hat{y}_{i_{\text{sel}}} = \sum_{i_{\text{sel}}=1}^{I_{\text{sel}}} w_{i_{\text{sel}}} \times f^{i_{\text{sel}}}(K_{\text{er}}^{i_{\text{sel}}}, R_{\text{eg}}^{i_{\text{sel}}}, \mathbf{Z}_{\text{FeSe2nd}}^{i_{\text{sel}}}) \quad (30)$$

其中, $f^{i_{\text{sel}}}(\cdot)$ 表示第 i_{sel} 个集成子模型, $w_{i_{\text{sel}}}$ 和 $\hat{y}_{i_{\text{sel}}}$ 表示其相应的加权系数及预测值, $K_{\text{er}}^{i_{\text{sel}}}$, $R_{\text{eg}}^{i_{\text{sel}}}$ 和 $\mathbf{Z}_{\text{FeSe2nd}}^{i_{\text{sel}}}$ 表示其超参数和输入特征. 对比式 (29) 可知, 存在如下关系:

$$\begin{cases} \{\hat{y}_{i_{\text{sel}}}\}_{i_{\text{sel}}=1}^{I_{\text{sel}}} \in \{\hat{y}^i\}_{i=1}^I \\ \{f^{i_{\text{sel}}}(K_{\text{er}}^{i_{\text{sel}}}, R_{\text{eg}}^{i_{\text{sel}}}, \mathbf{Z}_{\text{FeSe2nd}}^{i_{\text{sel}}})\}_{i_{\text{sel}}=1}^{I_{\text{sel}}} \in \\ \{f^i(K_{\text{er}}^i, R_{\text{eg}}^i, \mathbf{Z}_{\text{FeSe2nd}}^i)\}_{i=1}^I \end{cases} \quad (31)$$

利用集成子模型的预测值和真值, 基于预测误差信息熵的加权算法得到, $w_{i_{\text{sel}}}$ 由下式得到:

$$w_{i_{\text{sel}}} = \frac{1}{I_{\text{sel}} - 1} \left(\frac{1 - (1 - E_{i_{\text{sel}}})}{\sum_{i_{\text{sel}}=1}^{I_{\text{sel}}} (1 - E_{i_{\text{sel}}})} \right) \quad (32)$$

$$E_{i_{\text{sel}}} = \frac{1}{\ln N} \sum_{n=1}^N \left(\frac{(e_{i_{\text{sel}}})_n}{\sum_{n=1}^N (e_{i_{\text{sel}}})_n} \right) \ln \left(\frac{(e_{i_{\text{sel}}})_n}{\sum_{n=1}^N (e_{i_{\text{sel}}})_n} \right) \quad (33)$$

$$(e_{i_{\text{sel}}})_n = \begin{cases} (\hat{y}_{i_{\text{sel}}})_n - \frac{y_n}{y_n}, & 0 \leq \left| (\hat{y}_{i_{\text{sel}}})_n - \frac{y_n}{y_n} \right| < 1 \\ 1, & \left| (\hat{y}_{i_{\text{sel}}})_n - \frac{y_n}{y_n} \right| \geq 1 \end{cases} \quad (34)$$

其中, $(\hat{y}_{i_{\text{sel}}})_n$ 和 $(e_{i_{\text{sel}}})_n$ 表示第 n 个样本基于第 i_{sel} 个集成子模型 $f^{i_{\text{sel}}}(\cdot)$ 的预测值和相对预测误差, $E_{i_{\text{sel}}}$ 表示第 i_{sel} 个集成的子模型的预测误差信息熵.

4 应用研究

4.1 建模数据描述

本文数据源于北京某 MSWI 焚烧厂的 1# 炉,

涵盖了2012年~2018年DXN排放浓度检测样本39个,输入变量的维数286维(删除部分数据残缺的变量)。可见,输入特征维数远远超过建模样本数量,进行维数约简非常必要。本文将建模数据等分训练和测试两部分

注1. 本文研究的目的包括对工业过程数据可用性验证和算法有效性验证2个方面。

4.2 建模方法

4.2.1 广义子系统划分、潜在特征提取与初选的结果

本文中,将依据工艺流程划分的阶段子系统焚烧、锅炉、烟气处理、蒸汽发电、烟气排放和公用工程分别标记为Incinerator、Boiler、Flue gas、Steam、Stack和Common,并认定包含差异化的局部信息的标准子系统。为表示焚烧过程变量的整体变化特性,将包含全部变量的MSWI全流程系统也作为表征全局信息的广义子系统。因此,广义上讲,本文共包含7个子系统。采用PCA提取的前6个潜在特征的累积贡献率如图3所示。

图3表明:不同子系统前6个PC的累积贡献率均达到了80。

考虑到采用贡献率较低的潜在变量建模会导致模型预测性能的不稳定,本文将 θ_{Contri} 取为1进行潜在特征的初选。初选潜在特征及其贡献率如表2所示。

由表2可知:1)不同阶段子系统的初选潜在特征数量依次为13、6、9、8、6和12,其中Incinerator(焚烧)子系统和Steam(蒸汽发电)子系统选择的数量为最多的13个和最少的5个,为各自原始过程变量数量的1/6和1/10,表明输入变量间较强的共线性和阶段子系统间的差异性;2)Stack(烟气排放)子系统包含的过程变量和提取的潜在变量数量均为6个,表明排放烟气包含的HCL、SO₂等组分之间具有差异性;3)MSWI全流程系统所初选的潜在特征数量为15个,高于其他阶段子系统,从其蕴含全局信息的视角这也是合理的。

4.2.2 潜在特征度量与再选结果

此处采用MI方法度量上述初选潜在特征与DXN间的相关性,如图4所示。

图4表明:1)从全部子系统及全流程系统初选潜在特征的第1个特征的MI值可知,虽然其能够表征原始过程变量中的最大部分变化,但其与DXN间的相关性却是最小的,这表明进行潜在特征的再选是必要的;2)第1个潜在特征的MI最小值(0.4429)源于MSWI全流程系统,最大值(0.7182)源于Stack(烟气排放)子系统,且后者是前者的1.6倍,表明按照工艺流程划分为不同子系统进行特征提取是非常合理的;3)除第1个潜在特征外,其他潜在

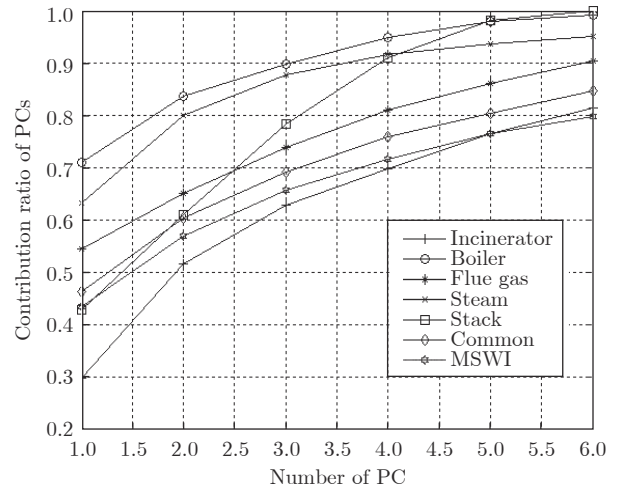


图3 不同功能子系统的前6个PC的累积贡献率

Fig.3 Cumulative contribution rate of the first six PCs of different functional subsystems

特征虽然在表征过程变量的贡献率上是逐渐下降的,但在表征MI值上却未呈现出任何规律性,表明以与DXN的相关性大小为准则进行潜在特征度量和再选择是必要的。

初选潜在特征MI值的极值统计如表3所示。

表3表明:潜在特征与DXN相关性的强弱与其被提取的顺序和贡献率不相关。如,针对Incinerator(焚烧)子系统,其潜在特征的最大MI值0.8559来自于贡献率仅为1.514。

结合第4.2节所述算法和表3度量结果,MI阈值的上限和下限自适应地确定为0.7882和0.7182,将 $N_{\text{Contri}}^{\text{Step}}$ 取为默认值10,可知步长为0.006999。结合软测量模型性能,最终确定阈值为0.7882。再选潜在特征的数量和MI值的统计结果如表4所示。

由表4可知,最终为全部子系统和全流程系统选择的再选潜在特征数量分为5、2、1、3、2、6和1,这表明与DXN具有较高相关性的潜在变量的分布与数量是不同的,构建不同的预测子模型更能体现其在建模贡献率上的差异性和互补性。

4.2.3 自适应选择集成建模

本文中,将候选正则化参数与核参数的集合分别预先选择为{0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000, 2000, 4000, 6000, 8000, 10000, 20000, 40000, 60000, 80000, 160000}和{0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000, 1600, 3200, 6400, 12800, 25600, 51200, 102400}。结合第3节所确定的再选潜在特征为输入,采用网格搜索方法进行LS-SVM超参数自适应寻优的第1次和第2次的曲线如图5所示。

基于上述结果,上述子模型自适应选择的超参数对分别为(109, 109)、(10000, 25.75)、(5.950,

表 2 初选潜在特征的数量及其贡献率
Table 2 Number of the primary selected latent feature and their contribution ratio

子系统代号		Incinerator	Boiler	Flue gas	Steam	Stack	Common	MSWI
特征编号	1	29.90	70.99	54.57	63.34	42.91	46.33	43.58
	2	21.75	12.66	10.42	16.56	18.06	14.10	13.40
	3	11.14	6.058	8.901	7.691	17.30	8.653	8.761
	4	6.952	5.014	7.146	3.906	12.65	6.798	5.921
	5	6.635	3.036	5.041	2.030	7.211	4.483	4.822
	6	5.075	1.356	4.269	1.533	1.854	4.221	3.246
	7	3.792	—	3.237	1.184	—	3.501	3.071
	8	3.208	—	2.584	1.007	—	2.842	2.919
	9	2.784	—	1.190	—	—	2.116	2.444
	10	1.846	—	—	—	—	1.494	2.138
	11	1.514	—	—	—	—	1.256	1.911
	12	1.283	—	—	—	—	1.164	1.731
	13	1.129	—	—	—	—	—	1.481
	14	—	—	—	—	—	—	1.344
	15	—	—	—	—	—	—	1.068
初选潜在特征数量		13	6	9	5	6	12	15
原始过程变量数量		79	14	19	53	6	115	286

0.0595)、(30.70, 2.080)、(5.950, 0.5950)、(1520800, 22816) 和 (1362400, 158.5), 对应的测试均方根误差 (Root mean square error, RMSE) 分别为 0.01676、0.02302、0.01348、0.01943、0.01475、0.02261 和 0.02375. 可见, 采用 MSWI 全流程系统的再选潜在特征构建的子模型预测误差最大 (0.02375), 其对应的超参数对具有较大范围 (1362400, 158.5), 表明蕴含全局信息的潜在特征的数据分布范围较宽; 相对而言, 依据工艺流程划分的焚烧、烟气处理和烟气排放子系统所对应的子模型的预测误差分别为 0.01676、0.01348 和 0.01475, 其相应的超级参数分别为 (109, 109), (5.950, 0.0595) 和 (5.950, 0.5950); 从机理上讲, 上述 3 个子系统是与 DXN 的生成与燃烧、吸收、排放等工艺密切先关, 该结果进一步表明了本文所提方法的合理性; 此外, 具有最大潜在特征 MI 值的公用工程子系统所构建的子模型的预测误差仅为 0.02261, 这也是与 DXN 机理和该子系统相关性很小的事实相符合.

针对基于上述各个子系统及全流程系统潜在特征的候选子模型, 采用基于分支定界和预测误差信息熵加权算法的耦合策略, 可得到集成尺寸为 26 时的 SEN 模型, 其测试误差相应地分别为 0.01345、0.01332、0.01401、0.01460 和 0.01560. 可见, 在集成尺寸为 2 和 3 时, 具有强于最佳候选子模型的预测性能.

最终构建得到的 DXN 软测量模型的集成尺寸为 3, 其集成子模型所对应的子系统为烟气处理、烟气排放和焚烧. 由前文分析可知, 此 3 个子系统所对应的工艺流程与 DXN 的吸收、排放和生成与燃烧相关. 目前焚烧厂为降低 DXN 排放浓度, 除在烟气处理阶段均采用较为先进的末端处理装置外, 对能够吸附 DXN 的活性炭供应量往往为上限值; 但需要提出的是, 该方法只是改变了 DXN 的存在方式, 并未从根本上予以消除. 因此, 在未来工作中, 非常有必要对烟气处理装置之前的 DXN 含量进行测量, 为优化控制 DXN 的生成与燃烧以及吸附提供支撑.

由上述仿真结果可知, 从国内实际 MSWI 过程获取的 DXN 检测数据具有一定程度的可用性, 本文所提算法也是有效的, 但更具实用性的研究还需要结合更多定制化、高成本的实验才能深入进行.

4.3 比较结果

由前文可知, 此处所提方法与文献 [22, 24, 37] 相比的优点在于: 1) 采用全部 MSWI 过程变量建模, 之前方法是采用依据机理或经验选择的部分与 DXN 相关的过程变量建模, 丢弃的过程变量会造成信息损失; 2) 依据工艺流程划分为不同的阶段子系统, 分别提取和选择各自潜在特征构建候选子模型, 并将蕴含全局信息的 MSWI 全流程系统也作为一个广义概念下的子系统, 最终构建 DXN 排放浓

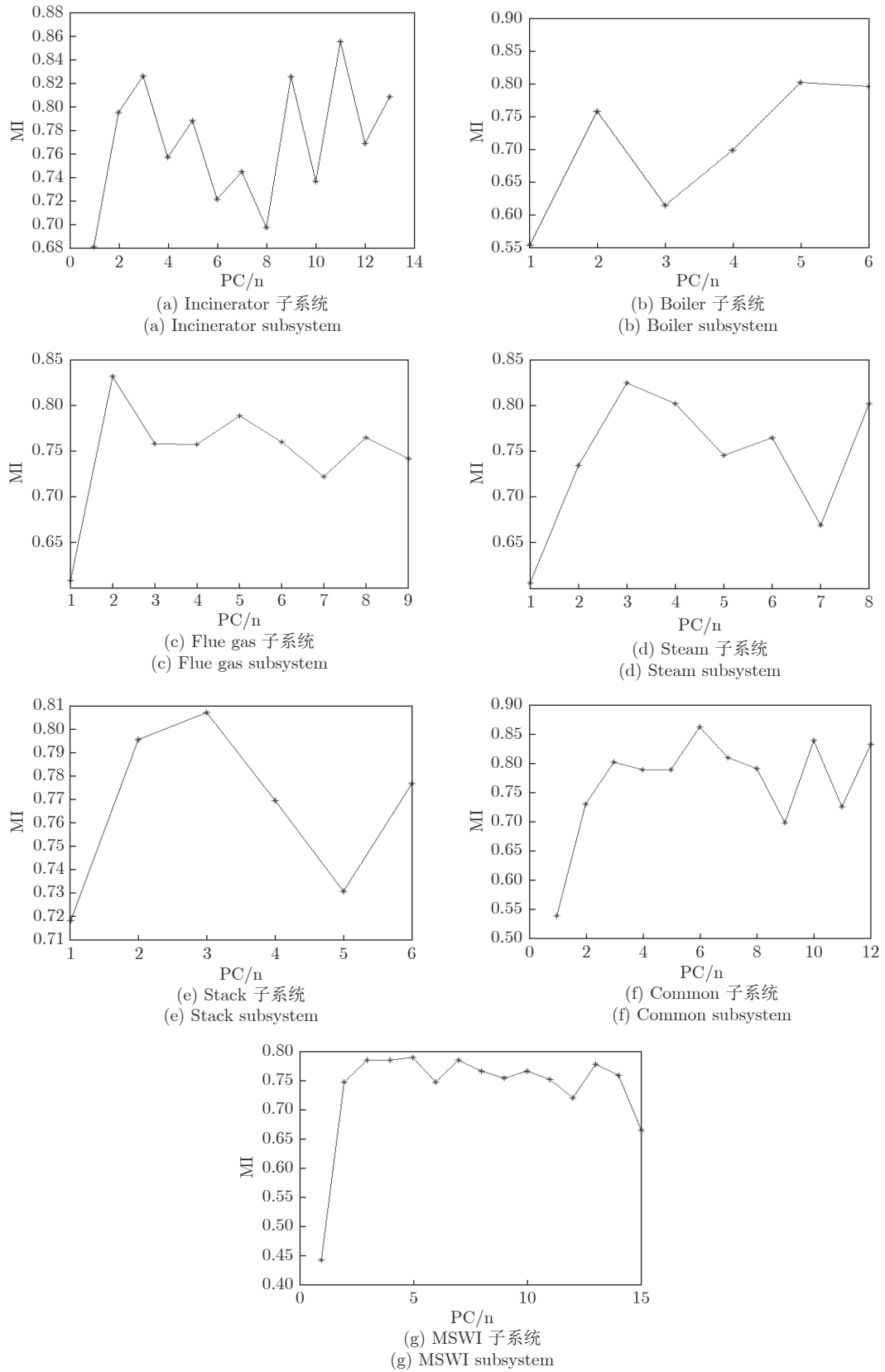


图 4 全部子系统及 MSWI 全流程系统的初选潜在特征与 DXN 间的 MI 值

Fig.4 MI value between DXN and primary potential characteristics of all subsystems and MSWI whole process systems

表 3 全部子系统及 MSWI 全流程系统初选潜在特征 MI 值的极值统计表
Table 3 Extremum statistical table of potential characteristic MI values for primary selection latent feature of all Subsystems and MSWI whole process system

子系统	最大值集合			最小值集合		
	MI 值	贡献率 (%)	PC 编号	MI 值	贡献率 (%)	PC 编号
Incinerator	0.8559	1.514	11	0.6814	29.90	1
Boiler	0.8019	3.036	5	0.5527	70.99	1
Flue gas	0.8316	10.42	2	0.6084	54.57	1
Steam	0.8249	7.691	3	0.6059	63.34	1
Stack	0.8067	17.30	3	0.7182	42.91	1
Common	0.8613	4.221	6	0.5400	46.33	1
MSWI	0.7882	4.822	5	0.4429	43.58	1

表 4 再选潜在特征数量和 MI 值统计表
Table 4 Statistical table of re-selected latent feature's number and MI value

子系统	数量	MI 值					
Incinerator	5	0.7952	0.8267	0.8258	0.8559	0.8088	—
Boiler	2	0.8019	0.7952	—	—	—	—
Flue gas	1	0.8316	—	—	—	—	—
Steam	3	0.8249	0.8022	0.8019	—	—	—
Stack	2	0.7952	0.8067	—	—	—	—
Common	6	0.8019	0.8613	0.8088	0.7904	0.8383	0.8316
MSWI	1	0.7882	—	—	—	—	—

度 SEN 软测量模型, 实现具有互补特性候选子模型的自适应选择性融合, 而之前方法多为单模型, 即使 SEN 模型也未进行多源信息的融合。

因此, 为与已有方法进行比较, 基于本文所用的近几年实际工业过程数据, 选择与文献 [22, 24, 37] 所采用的相同或类似的过程变量及方法进行比较: 1) 借鉴文献 [22], 选择“炉膛温度、二次燃烧室出口温度左、二次燃烧室出口温度右、一级过热器出口蒸汽温度、二级过热器出口蒸汽温度、三级过热器出口蒸汽温度、活性炭储仓给料量、FGD 出口烟气温度 A、FGD 出口烟气温度 B、以及烟气中 SO₂、HCL 和 O₂ 浓度”共 12 个变量; 2) 借鉴文献 [24], 选择“炉膛温度、一级过热器出口蒸汽温度、二级过热器出口蒸汽温度、三级过热器出口蒸汽温度、烟道入口烟气流量、以及烟气中 SO₂、HCL、灰尘浓度”共 8 个变量; 3) 借鉴文献 [37], 选择“炉膛温度、锅炉出口主蒸汽流量、烟道入口烟气温度、烟道入口烟气流量、HCL 浓度、PM 浓度”共 6 个变量。此外, 也与采用 PLS 和 AWF 进行合并加权融合策略的集成建模 (EN) 建模方法进行了比较。

不同建模方法的统计结果如表 5 所示。由表 5 可知:

1) 从用于构建软测量模型的过程变量数量的视角: 文献 [22, 24, 37] 所采用输入特征是从 286 个

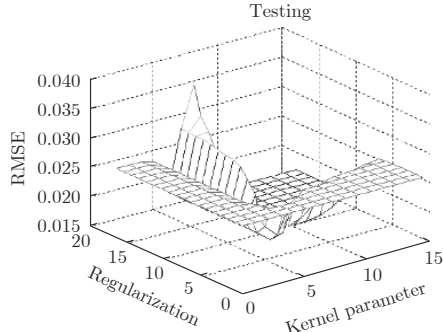
变量中选择的 12、8 和 6 个机理或经验特征, 采用经验风险神经网络模型、结构最小化风险 SVM 模型和基于多核参数 SEN 模型构建的 DXN 模型在预测性能上均弱于以 286 全部变量为输入的 3 类模型 (PLS 模型、PCA-LSSVM 模型、基于 PCA-MI-LSSVM 子模型的 EN 模型), 以及实际上以 104 个过程变量为输入的基于 PCA-MI-LSSVM 子模型的 SEN 模型, 表明前 3 种方法在丢弃大部分过程变量的过程中导致了有价值信息的损失, 也说明了采用适当数量的过程变量建模是必要的。

2) 从采用建模方法的视角: 文献 [22] 的结果表明神经网络在面向小样本数据建模时具有较大的随机性, 文献 [24] 的结果表明 SVM 适合构建小样本数据模型, 通过集成多个核参数的 SEN 策略 [37] 可进一步提高预测性能; PLS 方法通过提取与 286 个输入变量和输出 DXN 排放浓度均相关的潜在特征构建线性回归模型, 也获得了比上述 3 种仅采用部分过程变量建模更好的性能; PCA-LS-SVM 通过提取 MSWI 全流程系统的潜在特征, 构建了预测误差进一步降低的软测量模; 基于 PCA-MI-LSSVM 子模型的 EN 和 SEN 方法也进一步提高建模精度, 表明选择性融合局部与全局信息的合理性。总体上, 结构风险最小化的建模算法性能较佳, SEN 建模策略更优。

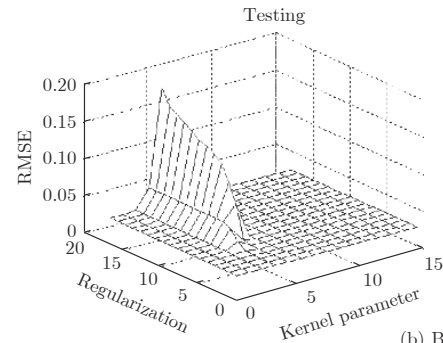
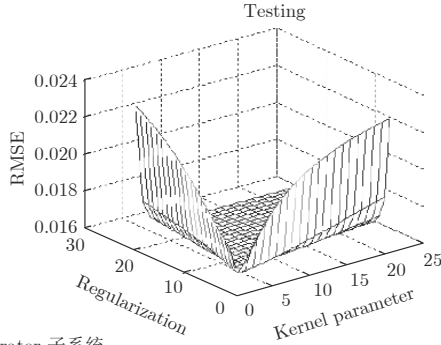
3) 从集成子模型融合策略的视角: 基于 PLS 的 EN 建模方法具有仅次于本文所提方法的预测性能, 表明 PLS 能够有效对 7 个候选子模型进行较为合适的融合. 此处, PLS 融合相当于构建了候选

子模型输出与最终 EN 输出间的映射关系^[38]. 考虑到仅构建了线性融合模型, 还需进行深入研究.

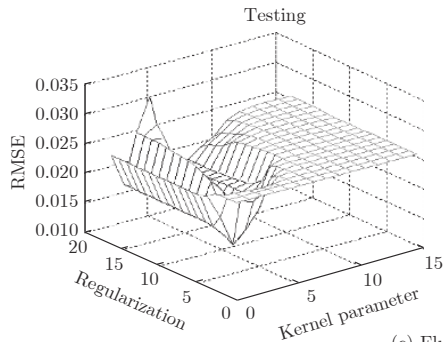
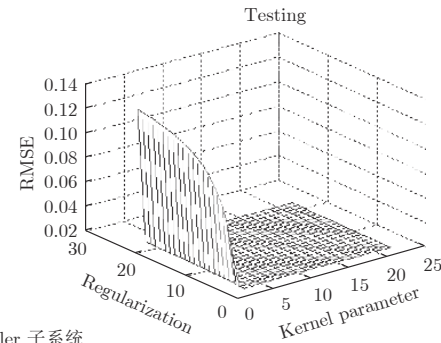
4) 从 DXN 软测量模型的集成子模型自适应选择与合并的视角: 本文所提方法无需对全部过程变



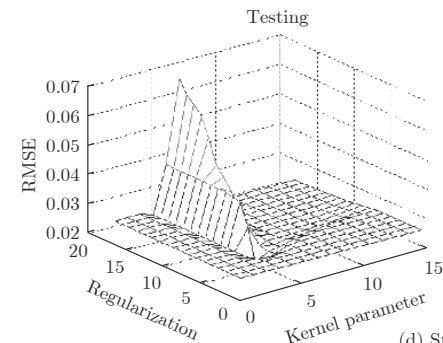
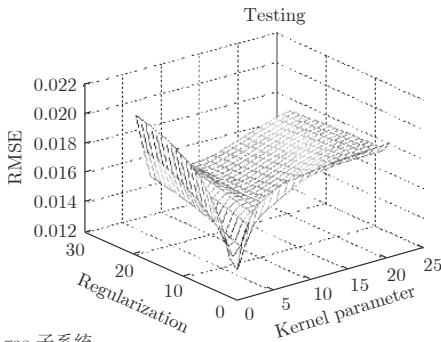
(a) Incinerator 子系统
(a) Incinerator subsystem



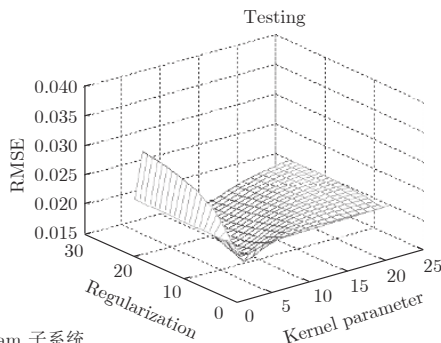
(b) Boiler 子系统
(b) Boiler subsystem



(c) Flue gas 子系统
(c) Flue gas subsystem



(d) Steam 子系统
(d) Steam subsystem



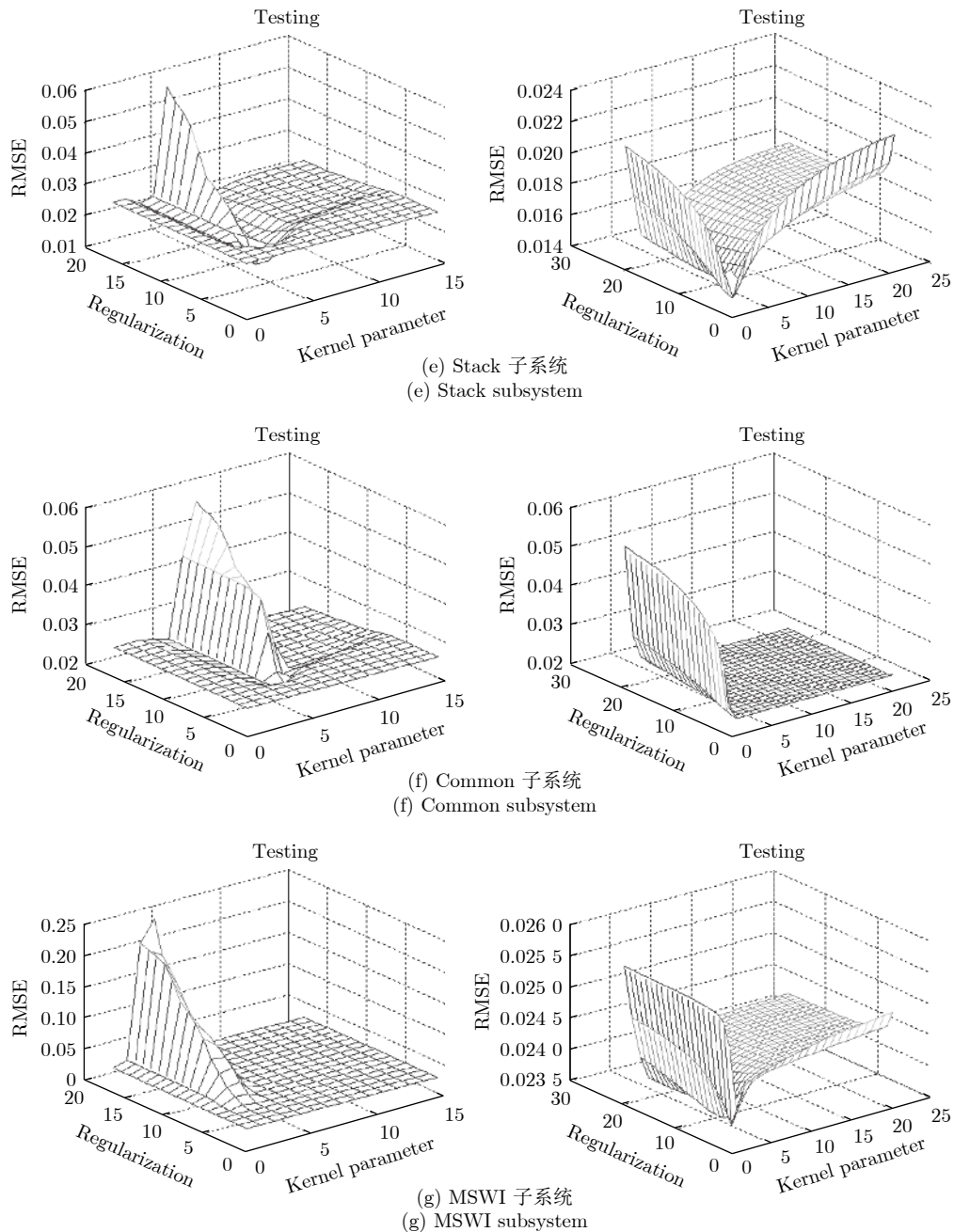


图 5 子模型超参数自适应寻优的第 1 次和第 2 次的曲线

Fig. 5 Curves of the 1st and 2nd curves for adaptive hyperparametric optimization of submodels

量进行选择, 只是通过对依据工艺阶段的划分的包含局部信息的 6 个阶段子系统及包含全局信息的全流程系统的潜变量特征候选子模型的优化选择与加权, 构建得到面向多源信息的选择性融合模型, 结果表明所选择的集成子模型是基于与 DXN 的生成、燃烧、吸附和排放机理相关的阶段子系统构建, 进一步验证了建模数据的可用性和算法的有效性。

综上所述, 本文所提方法具有最佳预测性能, 能够对表征局部信息和全局信息的阶段子系统和

MSWI 全流程系统所蕴含的潜在特征进行有效的选择性融合。同时, 由于不同焚烧系统的差异性, 在子系统划分粒度与方式、全流程系统融合效果等方面, 本文所提方法还需要在更多实际工业数据上进行进一步验证。

5 结论

本文针对北京某焚烧企业的工业实际, 依据工艺流程将 MSWI 过程划分为多个阶段子系统, 针对

表 5 不同建模方法统计结果
Table 5 Statistical results of different modeling methods

方法	过程变量数量	加权方法	RMSE	参数 (LV/PC) (K_{cr} , R_{cg})	备注
文献 [22]	12	—	0.08869 ± 0.3000	(—) (—)	单模型, RWNN
文献 [24]	8	—	0.02695	(—) (21, 21)	单模型, SVM
文献 [37]	6	AWF	0.02306	(—) (0.1, 1; 400, 6400; 12800, 25600; 51200, 102400)	SEN, 基于多核参数
PLS	286	—	0.01790	(13) (—)	单模型, MSWI系统
PCA-LS-SVM	286	—	0.01563	(18) (36240, 83904)	单模型, MSWI系统
集成建模 (EN)	286	PLS	0.01420	(5, 2, 1, 3, 2, 6, 1) (109, 109; 10000, 25.75; 5.950, 0.0595; 30.70, 2.080; 5.950, 0.5950; 1520800, 22816; 1362400, 158.5)	PCA-MI-LSSVM子模型, EN, 全部子模型
		AWF	0.01851		
		Entropy	0.01625		
选择性集成建模 (SEN) (本文方法)	286 (104)	BB-AWF	0.01348	(5, 1, 2) (109, 109; 5.950, 0.0595; 5.950, 0.5950)	PCA-MI-LSSVM子模型, SEN, Incinerator, Flue gas, Stack共3个子模型
		BB-Entropy	0.01332		

划分得到的表征局部信息的阶段子系统和表征全局信息的全流程系统, 采用基于 PCA 和先验知识的潜在特征提取和初选、基于 MI 和先验知识的初选潜在特征度量和再选、面向再选潜在特征的自适应 SEN 建模机制共 3 个阶段的有效融合, 提出了基于多源潜在特征 SEN 建模的 DXN 排放浓度软测量方法, 并基于实际运行数据仿真验证了所提方法的有效性. 本文在自适应调整贡献率阈值、MI 再选阈值、超参数和 SEN 模型结构等方面还有待于加强. 此外, 如何结合数值仿真的模型与工艺流程进行 DXN 排放机理的分析也还有待于深入进行. 所提方法也需要在具有不同特性的焚烧系统上进一步验证.

References

- Chai Tian-You. Operational optimization and feedback control for complex industrial processes. *Acta Automatica Sinica*, 2013, **39**(11): 1744–1757 (柴天佑. 复杂工业过程运行优化与反馈控制. *自动化学报*, 2013, **39**(11): 1744–1757)
- Chai T Y, Ding J L, Yu G, Wang H. Integrated optimization for the automation systems of mineral processing. *IEEE Transactions on Automation Science and Engineering*, 2014, **11**(4): 965–982
- Chai T Y, Qin S J, Wang H. Optimal operational control for complex industrial processes. *Annual Reviews in Control*, 2014, **38**(1): 81–92
- Arafat H A, Jijakli K, Ahsan A. Environmental performance and energy recovery potential of five processes for municipal solid waste treatment. *Journal of Cleaner Production*, 2015, **105**: 233–240
- Hu Y A, Cheng H A, Tao S. The growing importance of waste-to-energy (WTE) incineration in China's anthropogenic mercury emissions: Emission inventories and reduction strategies. *Renewable and Sustainable Energy Reviews*, 2018, **97**: 119–137
- Huang T, Zhou L L, Liu L F, Xia M. Ultrasound-enhanced electrokinetic remediation for removal of Zn, Pb, Cu and Cd in municipal solid waste incineration fly ashes. *Waste Management*, 2018, **75**: 226–235
- Jones P H, Degerlache J, Marti E, Mischer G, Scherrer M C, Bontinck W J, Niessen H J. The global exposure of man to dioxins — A perspective on industrial-waste incineration. *Chemosphere*, 1993, **26**: 1491–1497
- Li X M, Zhang C M, Li Y Z, Zhi Q. The status of municipal solid waste incineration (MSWI) in China and its clean development. *Energy Procedia*, 2016, **104**: 498–503
- Phillips K, Longhurst P J, Wagland S T. Assessing the perception and reality of arguments against thermal waste treatment plants in terms of property prices. *Waste Management*, 2014, **34**(1): 219–225
- Zhang H J, Ni Y W, Chen J P, Zhang Q. Influence of variation in the operating conditions on PCDD/F distribution in a full-scale MSW incinerator. *Chemosphere*, 2008, **70**(4): 721–730
- Mukherjee A, Debnath B, Ghosh S K. A review on technologies of removal of dioxins and furans from incinerator flue gas. *Procedia Environmental Sciences*, 2016, **35**: 528–540
- Stannmore B R. Modeling the formation of PCDD/F in solid waste incinerators. *Chemosphere*, 2002, **47**: 565–773
- Qiao Jun-Fei, Guo Zi-Hao, Tang Jian. Dioxin emission concentration measurement approaches for municipal solid wastes incineration process: A survey. *Acta Automatica Sinica*, 2020, **46**(6): 1063–1089 (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. *自动化学报*, 2020, **46**(6): 1063–1089)
- Pandelova M, Lenoir D, Schramm K W. Correlation between PCDD/F, PCB and PCBz in coal/waste combustion Influence of various inhibitors. *Chemosphere*, 2006, **62**: 1196–1205
- Gullett B K, Oudejans L, Tabor D, Touati A, Ryan S. Near-real-time combustion monitoring for PCDD/PCDF indicators by GC-REMPI-TOFMS. *Environmental Engineering Science*, 2012, **46**: 923–928
- Wang W, Chai T Y, Yu W, Wang H, Su C Y. Modeling component concentrations of sodium aluminate solution via hammerstein recurrent neural networks. *IEEE Transactions on Control Systems Technology*, 2012, **20**(4): 971–982
- Tang J, Chai T Y, Yu W, Zhao L J. Modeling load parameters of ball mill in grinding process based on selective ensemble multisensor information. *IEEE Transactions on Automation Science and Engineering*, 2013, **10**(3): 726–740
- Li D C, Liu C W. Extending attribute information for small data set classification. *IEEE Transactions on Knowledge and Data Engineering*, 2010, **24**(3): 452–464
- Tang Jian, Qiao Jun-Fei, Chai Tian-You, Liu Zhuo, Wu Zhi-

- Wei. Modeling multiple components mechanical signals by means of virtual sample generation technique. *Acta Automatica Sinica*, 2018, **44**(9): 1569–1589
(汤健, 乔俊飞, 柴天佑, 刘卓, 吴志伟. 基于虚拟样本生成技术的多组分机械信号建模. *自动化学报*, 2018, **44**(9): 1569–1589)
- 20 Chang N B, Huang S H. Statistical modelling for the prediction and control of PCDDs and PCDFs emissions from municipal solid waste incinerators. *Waste Management and Research*, 1995, **13**: 379–400
- 21 Chang N B, Chen W C. Prediction of PCDDs/PCDFs emissions from municipal incinerators by genetic programming and neural network modeling. *Waste Management and Research*, 2000, **18**(4): 41–351
- 22 Bunsan S, Chen W Y, Chen H W, Chuang Y H, Gridanurak N. Modeling the dioxin emission of a municipal solid waste incinerator using neural networks. *Chemosphere*, 2013, **92**: 258–264
- 23 Gomes T A F, Prudêncio R B C, Soares C, Rossi A L D, Carvalho A. Combining meta-learning and search techniques to select parameters for support vector machines. *Neurocomputing*, 2012, **75**(1): 3–13
- 24 Xiao Xiao-Dong, Lu Jia-Wei, Hai Jing, Liao Li. Prediction of dioxin emissions in flue gas from waste incineration based on support vector regression. *Renewable Energy Resources*, 2017, **35**(8): 1107–1114
(肖晓东, 卢加伟, 海景, 廖利. 垃圾焚烧烟气中二噁英类浓度的支持向量回归预测. *可再生能源*, 2017, **35**(8): 1107–1114)
- 25 Tang J, Chai T Y, Yu W, Zhao L J. Feature extraction and selection based on vibration spectrum with application to estimate the load parameters of ball mill in grinding process. *Control Engineering Practice*, 2012, **20**(10): 991–1004
- 26 Soares C. A hybrid meta-learning architecture for multi-objective optimization of SVM parameters. *Neurocomputing*, 2014, **143**(143): 27–43
- 27 Yu G, Chai T Y, Luo X C. Multiobjective production planning optimization using hybrid evolutionary algorithms for mineral processing. *IEEE Transactions on Evolutionary Computation*, 2011, **15**(4): 487–514
- 28 Yin S, Yin J P. Tuning kernel parameters for SVM based on expected square distance ratio. *Information Sciences*, 2016, **370–371**: 92–102
- 29 Tang J, Liu Z, Zhang J, Wu Z W, Chai T Y, Yu W. Kernel latent feature adaptive extraction and selection method for multi-component non-stationary signal of industrial mechanical device. *Neurocomputing*, 2016, **216**(C): 296–309
- 30 Tang Jian, Tian Fu-Qing, Jia Mei-Ying. *Soft Measurement of Rotating Machinery Equipment Load Based on Spectrum Data Drive*. Beijing: National Defense Industry Press, 2015.
(汤健, 田福庆, 贾美英. 基于频谱数据驱动的旋转机械设备负荷软测量. 北京: 国防工业出版社, 2015.)
- 31 Brown G, Wyatt J, Harris R, Yao X. Diversity creation methods: A survey and categorisation. *Information Fusion*, 2005, **6**: 5–20
- 32 Tang J, Chai T Y, Yu W, Liu Z, Zhou X J. A comparative study that measures ball mill load parameters through different single-scale and multi-scale frequency spectra-based approaches. *IEEE Transactions on Industrial Informatics*, 2016, **12**(6): 2008–2019
- 33 Zhou Z H, Wu J X, Tang W. Ensembling neural networks: Many could be better than all. *Artificial Intelligence*, 2002, **137**(1–2): 239–263
- 34 Ma G K, Wang Y, Wu L G. Subspace ensemble learning via totally-corrective boosting for gait recognition. *Neurocomputing*, 2017, **224**: 119–127
- 35 Tang J, Qiao J F, Wu Z W, Chai T Y, Zhang J, Yu W. Vibration and acoustic frequency spectra for industrial process modeling using selective fusion multi-condition samples and multi-source features. *Mechanical Systems and Signal Processing*, 2018, **99**: 142–168
- 36 Soares S, Antunes C H, Araújo R. Comparison of a genetic algorithm and simulated annealing for automatic neural network ensemble development. *Neurocomputing*, 2013, **121**(18): 498–511
- 37 Tang Jian, Qiao Jun-Fei. Dioxin emission concentration soft measuring approach of municipal solid waste incineration based on selective ensemble kernel learning algorithm. *Journal of Chemical Industry and Engineering (China)*, 2019, **70**(2): 696–706
(汤健, 乔俊飞. 基于选择性集成核学习算法的固废焚烧过程二噁英排放浓度软测量. *化工学报*, 2019, **70**(2): 696–706)
- 38 Tang J, Chai T Y, Liu Z, Yu W. Selective ensemble modeling based on nonlinear frequency spectral feature extraction for predicting load parameter in ball mills. *Chinese Journal of Chemical Engineering*, 2015, **23**(12): 2020–2028



汤健 北京工业大学教授。主要研究方向为小样本数据建模, 城市固废处理过程智能控制。本文通信作者。

E-mail: freeflytang@bjut.edu.cn

(**TANG Jian** Professor at Beijing University of Technology. His research interest covers small sample

data modeling and intelligent control of municipal solid waste treatment process. Corresponding author of this paper.)



乔俊飞 北京工业大学信息学部教授。主要研究方向为污水处理过程智能控制, 神经网络结构设计与优化。

E-mail: junfeq@bjut.edu.cn

(**QIAO Jun-Fei** Professor at the Faculty of Information Technology, Beijing University of Technology.

His research interest covers intelligent control of wastewater treatment process, and structure design and optimization of neural networks.)



郭子豪 北京工业大学信息学部硕士研究生。主要研究方向为高维小样本数据的特征建模, 固废处理过程难测参数软测量。

E-mail: miller94@163.com

(**GUO Zi-Hao** Master student at the Faculty of Information Techno-

logy, Beijing University of Technology. His research interest covers feature modeling of high-dimensional small sample data and soft measurement of difficulty-to-measure parameters in municipal solid waste treatment process.)