

城市固废焚烧过程炉温与烟气含氧量多目标鲁棒预测模型

胡开成^{1,2} 严爱军^{1,2,3} 汤健¹

摘要 为实现城市固废焚烧 (Municipal solid waste incineration, MSWI) 过程炉温与烟气含氧量的准确预测, 提出一种基于改进随机配置网络的多目标鲁棒建模方法 (Multi-target robust modeling method based on improved stochastic configuration network, MRI-SCN). 首先, 设计了一种并行方式增量构建 SCN 隐含层, 通过信息叠加与跨越连接来增强隐含层映射多样性, 并利用参数自适应变化的监督不等式分配隐含层参数; 其次, 使用 F 范数与 $L_{2,1}$ 范数正则项建立矩阵弹性网对模型参数进行稀疏约束, 以建模炉温与烟气含氧量间的相关性; 接着, 采用混合拉普拉斯分布作为每个目标建模误差的先验分布, 通过最大后验估计重新评估 SCN 模型的输出权重, 以增强其鲁棒性; 最后, 利用城市固废焚烧过程的历史数据对所提建模方法的性能进行测试. 实验结果表明, 所提建模方法在预测精度与鲁棒性方面具有优势.

关键词 城市固废焚烧, 炉温, 烟气含氧量, 随机配置网络, 隐含层并行构造, 多目标鲁棒建模

引用格式 胡开成, 严爱军, 汤健. 城市固废焚烧过程炉温与烟气含氧量多目标鲁棒预测模型. 自动化学报, 2024, 50(5): 1001-1014

DOI 10.16383/j.aas.c230430

Multi-target Robust Prediction Model for Furnace Temperature and Flue Gas Oxygen Content in Municipal Solid Waste Incineration Process

HU Kai-Cheng^{1,2} YAN Ai-Jun^{1,2,3} TANG Jian¹

Abstract To achieve accurate prediction of furnace temperature and flue gas oxygen content in municipal solid waste incineration (MSWI) process, a multi-target robust modeling method based on improved stochastic configuration network (MRI-SCN) is proposed. First, a parallel method is designed to incrementally build SCN hidden layers, which enhances the diversity of hidden layer mapping through information superposition and spanning connection, and assign hidden layer parameters using the supervised inequality with adaptive parameter changes. Second, a matrix elastic net is established by using F-norm and $L_{2,1}$ -norm regularization terms to sparsely constrain the model parameters to model the correlation between furnace temperature and flue gas oxygen content. Then, the mixture Laplace distribution is used as the prior distribution of each target modeling error, and the output weights of the SCN model are re-evaluated by maximum a posteriori estimation to enhance its robustness. Finally, the performance of the proposed modeling method is tested on the historical data of municipal solid waste incineration process. The experimental results show that the proposed modeling method has advantages in prediction accuracy and robustness.

Key words Municipal solid waste incineration (MSWI), furnace temperature, flue gas oxygen content, stochastic configuration network (SCN), hidden layer parallel construction, multi-target robust modeling

Citation Hu Kai-Cheng, Yan Ai-Jun, Tang Jian. Multi-target robust prediction model for furnace temperature and flue gas oxygen content in municipal solid waste incineration process. *Acta Automatica Sinica*, 2024, 50(5): 1001-1014

收稿日期 2023-07-13 录用日期 2024-01-08

Manuscript received July 13, 2023; accepted January 8, 2024
国家自然科学基金 (62373017, 62073006), 北京市自然科学基金 (4212032) 资助

Supported by National Natural Science Foundation of China (62373017, 62073006) and Beijing Natural Science Foundation of China (4212032)

本文责任编辑 谢永芳

Recommended by Associate Editor XIE Yong-Fang

1. 北京工业大学信息学部 北京 100124 2. 数字社区教育部工程研究中心 北京 100124 3. 城市轨道交通北京实验室 北京 100124

1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124 2. Engineering Research Center of Digital Community, Ministry of Education, Beijing 100124 3. Beijing Laboratory for Urban Mass Transit, Beijing 100124

城市固废焚烧 (Municipal solid waste incineration, MSWI) 具有占地面积小、处理时间短、资源回收利用率高等优势, 是我国目前处理城市固废的首选方案^[1]. 从 MSWI 过程控制的角度来看, 将焚烧炉的一燃室烟气温度均值 (以下简称为炉温) 稳定控制在工艺要求的 850 °C 以上、余热锅炉出口的烟气含氧量控制在 6% ~ 9%, 才有可能使得炉内固废与可燃性烟气充分燃烧, 从而保障二次污染物 (如: 氮氧化物、强致癌物二噁英等) 的排放浓度达标^[2-3]. 因此, 建立准确的炉温与烟气含氧量预测模

型有利于操作人员或自动控制系统及时发现焚烧炉运行状况的变化,从而采取相应的措施来避免异常情况的发生,对实现 MSWI 过程的平稳、高效、环保运行具有重要的现实意义。

目前,针对 MSWI 过程的炉温与烟气含氧量建模方法主要有机理建模与数据驱动建模两种。例如:从能量守恒与物料守恒的角度出发,利用物理、化学方程式建立固废焚烧炉与余热锅炉的机理模型,通过分析进料量、风量、炉排速度等参数对炉温与烟气含氧量的影响,可为 MSWI 过程的优化控制提供指导^[4-5]。然而,由于焚烧炉类型不同、固废组分复杂、模型简化等因素的影响,基于机理的炉温与烟气含氧量建模方法难以推广应用,且模型精度往往无法满足实际应用需求。值得注意的是,MSWI 过程中存在丰富的运行数据,这些离线或在线数据中隐含了设备与工艺参数的变化信息,在这种情况下,研究基于数据驱动的炉温与烟气含氧量建模方法可有效避免机理建模的局限性。

由于神经网络具有较好的非线性学习能力,在数据驱动参数建模领域受到研究人员的广泛关注^[6-7]。在 MSWI 过程参数建模方面,文献 [8] 利用基于梯度下降算法的自组织 T-S 模糊神经网络来预测炉温的变化趋势。文献 [9] 利用卷积神经网络对数据增强后的焚烧火焰图像进行智能识别,其识别结果可为现场操作人员修正焚烧控制策略提供参考。文献 [10] 则采用改进的长短时记忆神经网络建立烟气含氧量预测模型,并在控制过程中根据实际工况在线更新模型参数。但梯度下降算法在进行网络参数学习时易陷入局部最优,且模型训练周期较长。为提升神经网络的建模效率,以随机向量函数链网络(Random vector functional link, RVFL)^[11]为代表的随机神经网络相继被提出,但这类网络一般在固定区间内随机生成隐含层节点参数,其通用逼近性有时难以保证,且网络结构需预先指定。于是,文献 [12] 提出了随机配置网络(Stochastic configuration network, SCN),SCN 在可变区间内通过监督机制配置隐含层新增节点参数以增量构建学习网络,并利用最小二乘法计算输出权值。与基于梯度下降算法的误差反传网络、径向基神经网络等相比,SCN 的建模效率与精度相对较高,在污水处理^[13]、金属热轧^[14]、原油提炼^[15]等工业过程中得以成功应用。因此,文献 [16] 将 SCN 与递推最小二乘法相结合,构建了炉温动态预测模型,并将其用于炉温非线性模型预测控制中。文献 [17] 以正则化 SCN 为基模型,提出一种异构特征与负相关学习策略的集成建模方法来提升炉温预测模型的准确性与

训练效率。然而,上述均是针对炉温或烟气含氧量单目标进行建模,难以为 MSWI 过程的多目标协同优化控制提供全面的预测信息。文献 [18] 针对 MSWI 过程构建了基于 T-S 模糊神经网络的多目标被控对象模型,但其未利用炉温与烟气含氧量等多目标间的相关性来提升建模精度。此外,固废焚烧环境十分复杂,从现场采集到的数据中不可避免地包含噪声或离群点,上述方法还未充分考虑这些异常数据对模型准确性的负面影响。

综上所述,为实现对 MSWI 过程炉温与烟气含氧量的准确预测,本文以建模效率高且具有通用逼近性的 SCN 为基础,提出一种基于改进 SCN 的多目标鲁棒建模方法(Multi-target robust modeling method based on improved SCN, MRI-SCN)以促进 MSWI 过程的平稳高效运行,进而提升固废焚烧发电量、降低二次污染物排放浓度。本文主要工作如下:1) 在参数自适应的不等式监督机制下,采用并行构造的方式配置隐含层节点,以增强模型的非线性映射能力;2) 利用矩阵弹性网将炉温与烟气含氧量间的相关性假设转化为对模型参数的稀疏约束,以提高建模精度;3) 将混合拉普拉斯(Laplace)分布作为每个目标建模误差的先验分布,并使用最大后验估计来优化 SCN 模型的输出权值,以降低其对异常数据的敏感性。通过实验验证了所提建模方法的有效性。

本文余下部分的组织结构安排是:第 1 节是 MSWI 工艺流程描述与建模分析;第 2 节介绍 MRI-SCN 建模方法的实现过程;第 3 节是实验测试与结果分析;最后总结全文并提出展望。

1 MSWI 工艺流程描述与建模分析

本节主要介绍 MSWI 的工艺流程与 SCN 的参数学习过程,并分析将 SCN 用于构建数据驱动的炉温与烟气含氧量多目标鲁棒预测模型时需要解决的问题。

1.1 MSWI 工艺流程

以北京市某固废焚烧厂的炉排炉为例,MSWI 的工艺流程如图 1 所示,主要包括固废储运、固废焚烧、余热锅炉、蒸汽发电、烟气净化 5 个子系统。固废焚烧过程大致如下:固废吊车将堆放发酵后的固废投入进料斗,并由进料器推送至炉排上;固废经过干燥炉排时,在高温一次风对流干燥、炉膛内侧壁辐射干燥以及炉排热传导干燥的共同作用下蒸发水分,挥发份析出并不断升温达到着火点;固废通过燃烧炉排时,二次风由上而下吹入,促进炉排

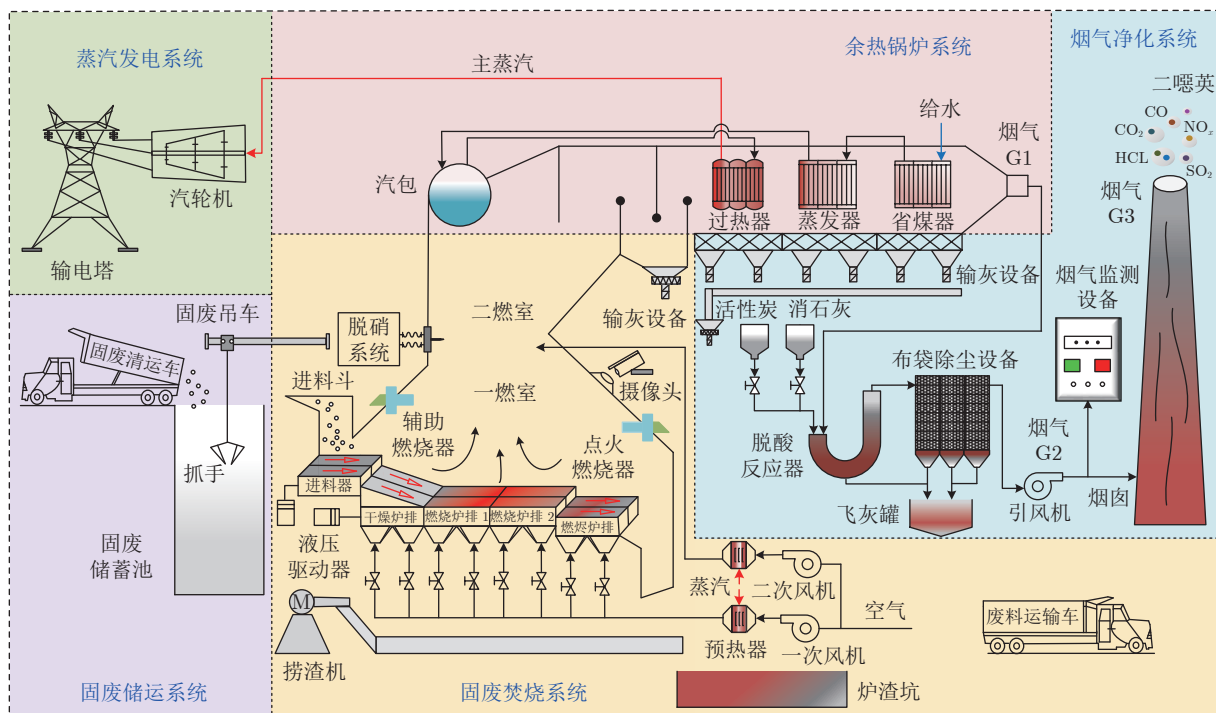


图 1 MSWI 工艺流程

Fig.1 MSWI process flow

上的固废与炉内可燃性烟气充分燃烧; 最后, 固废在燃烬炉排上完全燃烧, 冷却后形成稳定的固体残渣. 固废焚烧产生的高温烟气经余热锅炉系统回收热能, 产生饱和蒸汽用于推动汽轮机发电, 焚烧尾气经过脱酸反应、布袋除尘等处理后排入大气.

在 MSWI 过程中, 炉温与烟气含氧量是反映固废焚烧质量的关键指标^[2], 具体来说: 固废焚烧工艺要求将炉温稳定控制在 850 °C 以上, 否则炉温过低, 炉内固废与烟气无法充分燃烧, 从而导致固废焚烧效率下降, 二次污染物排放浓度超标. 就烟气含氧量而言, 通过优化一、二次风量及其配比, 使炉内烟气具有良好的湍流度, 在保证炉内固废充分燃烧的同时, 降低余热锅炉的排烟热损失以增加主蒸汽产量, 这有利于提高固废焚烧发电厂的经济效益. 然而, 目前已有的针对 MSWI 过程的炉温与烟气含氧量预测大多是单目标模型, 即分别对每个目标建立预测模型. 这种单目标模型只关注某个指标的变化情况, 并未考虑到炉温与烟气含氧量的整体变化趋势, 所能提供的预测信息相对较少. 由于炉温与烟气含氧量均为反映固废焚烧质量的指标, 二者在变化机理与趋势上具有一定的关联性. 为促进 MSWI 过程的平稳、高效运行, 同时兼顾模型的简洁性与鲁棒性, 本文力求在一个整体框架内实现炉温与烟气含氧量预测, 从而构建出炉温与烟气含氧

量多目标鲁棒预测模型.

由固废焚烧机理可知, 下一时刻的炉温与烟气含氧量主要受固废组分、入炉固废量、固废在炉内的停留时间、风量、风温以及当前时刻的炉温与烟气含氧量等因素的影响. 在固废热值、一次风压等诸多干扰中, 固废热值的波动对炉温与烟气含氧量的影响较大, 但该参数难以直接测量. 考虑到炉排与固废直接接触, 炉排温度在一定程度上可以反映固废热值的变化, 所以可将炉排温度作为炉温与烟气含氧量预测建模的辅助变量. 因此, 本文选取进料器速度、炉排速度、风量、炉排温度等 47 个变量作为预测模型的输入, 变量明细表如附录中的表 A1 所示.

1.2 SCN 参数学习过程

SCN 是一种单隐含层的随机学习模型, 其隐含层参数在不等式监督机制下随机生成, 并采用最小二乘法求解隐含层输出权值, 从而保证了 SCN 模型的通用逼近性. 下面简介 SCN 的参数学习过程, 关于 SCN 的详细介绍可阅读文献 [12].

对于目标函数 $f: \mathbf{R}^d \rightarrow \mathbf{R}^m$, 假设 SCN 训练集的输入为 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, 输出为 $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$, 当 SCN 隐含层已配置完成 $L-1$ 个节点时, 其输出表达式如下所示:

$$f_{L-1}(\mathbf{x}) = \sum_{j=1}^{L-1} \beta_j g_j(\mathbf{w}_j^T \mathbf{x} + b_j) \quad (1)$$

其中, β_j 表示隐含层第 j 个节点的输出权值; $g_j(\cdot)$ 表示隐含层第 j 个节点的激活函数; \mathbf{w}_j 与 b_j 分别表示隐含层第 j 个节点的输入权值与偏置. 当网络残差大于容忍误差 τ 且隐含层节点数量未超过最大预设值 L_{\max} 时, 根据式 (2) 所示的监督机制在可变区间 $\Upsilon = [\gamma_{\min} : \Delta\gamma : \gamma_{\max}]$ 内随机分配 T_{\max} 个候选节点, 每个候选节点对应一组输入权值 \mathbf{w}_L 与偏置 b_L . 为提升网络收敛速度, SCN 将 ξ_L 最大值对应的候选节点作为最优节点添加到隐含层中, 并利用式 (4) 重新计算隐含层输出权值 β .

$$\xi_L = \frac{(\mathbf{e}_{L-1}^T(\mathbf{X})\mathbf{h}_L(\mathbf{X}))^2}{\mathbf{h}_L^T(\mathbf{X})\mathbf{h}_L(\mathbf{X})} \quad (2)$$

$$(1 - r - \mu_L)\mathbf{e}_{L-1}^T(\mathbf{X})\mathbf{e}_{L-1}(\mathbf{X}) \geq 0$$

$$\mathbf{h}_L(\mathbf{X}) = [g_L(\mathbf{w}_L^T \mathbf{x}_1 + b_L), g_L(\mathbf{w}_L^T \mathbf{x}_2 + b_L), \dots, g_L(\mathbf{w}_L^T \mathbf{x}_N + b_L)]^T \quad (3)$$

$$[\beta_1^*, \beta_2^*, \dots, \beta_L^*] = \operatorname{argmin}_{\beta} \|\mathbf{H}_L \beta - \mathbf{Y}\|_{\text{F}}^2 = \mathbf{H}_L^{\dagger} \mathbf{Y} \quad (4)$$

其中, $\mathbf{e}_{L-1}(\mathbf{X})$ 表示隐含层节点为 $L-1$ 时的网络残差; $\mathbf{h}_L(\mathbf{X})$ 表示隐含层第 L 个节点的输出; $0 < r < 1$; $\{\mu_L\}$ 表示非负实数序列, $\lim_{L \rightarrow +\infty} \mu_L = 0$ 且 $\mu_L \leq (1-r)$; $\mathbf{H}_L = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_L]$ 表示隐含层输出矩阵; $(\cdot)^{\dagger}$ 表示伪逆运算; $\|\cdot\|_{\text{F}}$ 表示矩阵的 Frobenius 范数 (简称 F 范数), 其计算方式如下所示:

$$\|\mathbf{A}\|_{\text{F}} = \sqrt{\sum_{i=1}^m \sum_{j=1}^n \mathbf{A}_{ij}^2}, \quad \mathbf{A} \in \mathbf{R}^{m \times n} \quad (5)$$

1.3 SCN 多目标建模问题分析

SCN 在参数学习速度与建模精度等方面具有优势^[12], 其单目标建模性能已在 MSWI、污水处理等工业过程中得以验证^[13-17], 但多目标建模需同时对输入输出之间的非线性与多目标之间的相关性进行建模^[19], 因此将 SCN 用于数据驱动的炉温与烟气含氧量多目标预测建模时, 以下问题仍需进一步研究:

1) SCN 隐含层映射多样性存在不足. SCN 隐含层在增量构建过程中通常采用单一构造方式, 使得隐含层映射多样性存在不足, 从而限制了模型的非线性映射能力. 通过增加隐含层节点数量在一定程度上可以避免映射单一化, 但会导致模型存在过拟合的风险, 同时模型复杂度也相对较高, 因而有

必要从映射多样性方面考虑如何提升 SCN 的建模性能.

2) 多目标间的相关性建模. SCN 在进行多输入多输出建模时, 并未考虑到利用多目标间的相关性来提升模型精度, 仅是对输入输出间的非线性映射关系进行建模. 若直接将其用于构建炉温与烟气含氧量多目标预测模型, 则模型性能可能无法满足实际应用需求.

3) SCN 学习算法难以识别训练数据集中的异常样本. 由于固废焚烧环境十分复杂, 从现场采集到的数据中会包含噪声与离群点, 而 SCN 在进行参数学习时, 难以区分正常样本与异常样本, 即认为所有训练数据同等重要. 若用于训练 SCN 模型的数据受到污染, 则学习到的模型不具备鲁棒性, 建模质量较低.

2 MRI-SCN 建模方法

本节针对 SCN 在隐含层映射多样性、多目标间的相关性建模以及模型鲁棒性 3 个方面的不足进行研究并给出相应的解决策略, 从而提出一种基于改进 SCN 的多目标鲁棒建模方法 (MRI-SCN), 以实现 MSWI 过程炉温与烟气含氧量的准确预测.

2.1 隐含层并行构造

利用神经网络进行建模时, 增强隐含层映射多样性往往可以提升模型的学习能力^[20-21]. SCN 属于一种单隐含层的前馈神经网络, 拓扑结构如图 2(a) 所示, 其隐含层采用标准前馈构造, 即隐含层节点只接受上一层节点的输出, 隐含层节点之间并无信息传递, 模型结构相对简单. 标准前馈构造的隐含层第 L 个节点的输出可表示为:

$$\tilde{\mathbf{h}}_L(\mathbf{X}) = g_L(\mathbf{w}_L^T \mathbf{X} + \mathbf{b}_L)^T, \quad L \geq 1 \quad (6)$$

其中, \mathbf{b}_L 表示元素均为 b_L 的向量.

文献 [22] 指出级联构造的神经网络在解决某些问题时具有更好的信息处理能力. 级联构造允许神经网络隐含层节点间进行信息叠加与跨越连接^[23], 其拓扑结构如图 2(b) 所示. 但随着隐含层节点间信息叠加与跨越连接次数的增加, 模型复杂度也会逐渐上升. 级联构造的隐含层第 L 个节点的输出可表示为:

$$\tilde{\mathbf{h}}_L(\mathbf{X}) = \begin{cases} g_L(\mathbf{w}_L^T \mathbf{X} + \mathbf{b}_L)^T, & L = 1 \\ g_L(\mathbf{w}_L^T \mathbf{X} + \mathbf{b}_L + \sum_{j=1}^{L-1} \mathbf{q}_j \tilde{\mathbf{h}}_j^T(\mathbf{X}))^T, & L \geq 2 \end{cases} \quad (7)$$

其中, \mathbf{q}_j 表示隐含层第 j 个节点与其他隐含层节点之间的连接权值.

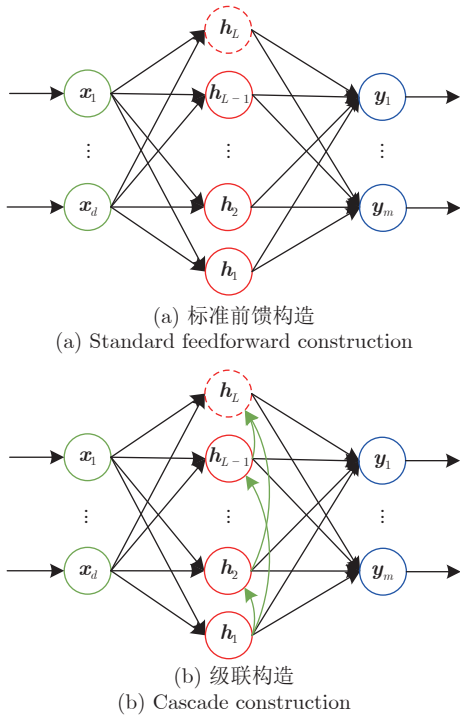


图2 前馈神经网络隐含层构造方式

Fig.2 The hidden layer construction methods of feedforward neural network

为综合考虑 SCN 模型的非线性映射能力与复杂度, 本节从隐含层映射多样性的角度出发, 将标准前馈构造与级联构造相融合, 采用并行构造的方式增量构建 SCN 隐含层. 采用并行构造时, SCN 隐含层第 L 个节点的输出可表示为:

$$\mathbf{h}_L(\mathbf{X}) = \begin{cases} \tilde{\mathbf{h}}_L(\mathbf{X}), & \xi(\tilde{\mathbf{h}}_L(\mathbf{X})) \geq \xi(\bar{\mathbf{h}}_L(\mathbf{X})) \\ \bar{\mathbf{h}}_L(\mathbf{X}), & \xi(\tilde{\mathbf{h}}_L(\mathbf{X})) < \xi(\bar{\mathbf{h}}_L(\mathbf{X})) \end{cases} \quad (8)$$

在 SCN 建模过程中, 式 (2) 中的参数 r 用于控制监督不等式的约束强度^[12]. 随着隐含层新增节点的配置难度逐渐增加, 监督不等式的约束强度应逐渐降低. 标准 SCN 利用线性搜索获得最佳参数 r , 但基于多次循环的线性搜索会降低 SCN 隐含层参数的配置效率, 甚至会出现新增节点配置失败的情况. 因此, 为提升 SCN 的参数配置效率, 本文在满足 SCN 模型通用逼近性的前提下, 基于监督不等式约束先紧后松的思想, 提出自适应变化的参数 r , 具体如下:

定理 1. 假设 $\text{span}(\Gamma)$ 稠密于 L_2 空间, $\forall \mathbf{g} \in \Gamma$, $0 < \|\mathbf{g}\|_2 < b_g$, $b_g \in \mathbf{R}^+$. 对于 $L = 1, 2, \dots$, 令 $r = 1 - \ln(1 + 1/(L+1)^P)$, $P > 1$, $0 < \mu_L \leq (1-r)/(L+1)$. 如果隐含层新增节点满足式 (2) 所示的监督不等式, 并根据式 (4) 求解隐含层输出权值, 则 $\lim_{L \rightarrow +\infty} \|\mathbf{e}_L^*\|_F = 0$.

证明. 当 SCN 隐含层节点个数为 L 时, 其残差可表示为 $\mathbf{e}_L^* = \mathbf{e}_{L-1}^* - \beta_L^* \mathbf{h}_L$. 令 $\tilde{\beta}_L = \langle \mathbf{e}_{L-1}^*, \mathbf{h}_L \rangle / \|\mathbf{h}_L\|_2^2$, 于是:

$$\begin{aligned} \|\mathbf{e}_L^*\|_F^2 &\leq \|\tilde{\mathbf{e}}_L^*\|_F^2 = \\ &\langle \mathbf{e}_{L-1}^* - \tilde{\beta}_L \mathbf{h}_L, \mathbf{e}_{L-1}^* - \tilde{\beta}_L \mathbf{h}_L \rangle = \\ \|\mathbf{e}_{L-1}^*\|_F^2 - \frac{\langle \mathbf{e}_{L-1}^*, \mathbf{h}_L \rangle^2}{\|\mathbf{h}_L\|_2^2} &\leq \|\mathbf{e}_{L-1}^*\|_F^2 \quad (9) \end{aligned}$$

即残差 $\|\mathbf{e}_L^*\|_F^2$ 单调递减. 当 $x > 0$ 时, 由于

$$\frac{x}{x+1} < \ln(1+x) < x \quad (10)$$

所以:

$$1-x < 1 - \ln(1+x) < 1 - \frac{x}{x+1} \quad (11)$$

令 $\|\mathbf{e}_0\|_F^2 = \|f\|_F^2$, 当 $r = 1 - \ln(1 + 1/(L+1)^P)$ 时, 可得:

$$\begin{aligned} \|\mathbf{e}_L^*\|_F^2 - (r + \mu_L) \|\mathbf{e}_{L-1}^*\|_F^2 &< \\ \|\mathbf{e}_L^*\|_F^2 - r \|\mathbf{e}_{L-1}^*\|_F^2 &= \\ \|\mathbf{e}_L^*\|_F^2 - \left(1 - \ln\left(1 + \frac{1}{(L+1)^P}\right)\right) \|\mathbf{e}_{L-1}^*\|_F^2 &< \\ \|\mathbf{e}_L^*\|_F^2 - \left(1 - \frac{1}{L+1}\right) \|\mathbf{e}_{L-1}^*\|_F^2 &< \\ \prod_{i=1}^L \left(1 - \frac{1}{i+1}\right) \|\mathbf{e}_0\|_F^2 &\leq \\ \exp\left(-\sum_{i=1}^L \frac{1}{i+1}\right) \|f\|_F^2 &\quad (12) \end{aligned}$$

所以, 当 $L \rightarrow +\infty$ 时, $\lim_{L \rightarrow +\infty} \exp(-\sum_{i=1}^L \frac{1}{i+1}) = 0$, 于是 $\lim_{L \rightarrow +\infty} \|\mathbf{e}_L^*\|_F^2 = 0$, 即 $\lim_{L \rightarrow +\infty} \|\mathbf{e}_L^*\|_F = 0$. \square

2.2 矩阵弹性网

从信息处理的角度来看, 炉温与烟气含氧量以不同的视角展现了炉内固废的焚烧状态, 利用炉温与烟气含氧量之间的相关性进行建模可以充分挖掘建模数据中的隐含信息; 从机器学习归纳偏置的角度来看, 多目标建模倾向于获得可以同时处理多个相关任务的解决方案, 这有利于降低模型过拟合的风险^[24-25]. 由于炉温与烟气含氧量多目标建模共享同一组输入特征且输出之间具有关联性, 这使得预测模型的参数具有稀疏性. 因此, 将炉温与烟气含氧量间的相关性假设转化为对 SCN 模型输出权值的稀疏约束可进一步提升建模质量.

向目标函数中引入由 L_1 范数与 L_2 范数正则项组合构成的矩阵弹性网^[26], 可让学习到的模型参数

具有稀疏性且模型泛化性能较好,但包含 L_1 正则项的目标函数一般无法直接进行参数求导,使得模型参数的优化求解比较困难.相对于 L_1 范数与 L_2 范数而言, $L_{2,1}$ 范数首先求矩阵每一行的 L_2 范数,再将得到的结果求 L_1 范数^[27],将基于 $L_{2,1}$ 范数的正则项加入目标函数中也可保证模型参数具有良好的稀疏性^[28-29]且便于参数求导, $L_{2,1}$ 范数正则项已广泛用于机器学习^[30]、模式识别^[31]等领域中. SCN 在进行多目标建模时,输出权值 β 为矩阵形式,其对应于向量空间的 L_2 范数称为 F 范数^[29].因此,本节采用 F 范数与 $L_{2,1}$ 范数正则项构建矩阵弹性网对 SCN 模型的输出权值进行稀疏约束,以建模炉温与烟气含氧量之间的相关性.输出权值 β 的 $L_{2,1}$ 范数计算公式如下所示:

$$\|\beta\|_{2,1} = \sum_{i=1}^L \sqrt{\sum_{j=1}^m \beta_{ij}^2} \quad (13)$$

其中, $\|\cdot\|_{2,1}$ 表示矩阵的 $L_{2,1}$ 范数.此时,输出权值 β 的目标函数可表示为:

$$J(\beta) = \frac{1}{2} \|\mathbf{H}_L \beta - \mathbf{Y}\|_{\text{F}}^2 + \lambda \left(\alpha \|\beta\|_{2,1} + \frac{1-\alpha}{2} \|\beta\|_{\text{F}}^2 \right) \quad (14)$$

其中, λ 与 α 表示正则化系数.令 $\frac{\partial J(\beta)}{\partial \beta} = 0$, 可得:

$$\mathbf{H}_L^T \mathbf{H}_L \beta + \lambda \alpha \frac{\partial \|\beta\|_{2,1}}{\partial \beta} + \lambda(1-\alpha)\beta = \mathbf{H}_L^T \mathbf{Y} \quad (15)$$

其中, $\frac{\partial \|\beta\|_{2,1}}{\partial \beta}$ 可表示为^[28]:

$$\frac{\partial \|\beta\|_{2,1}}{\partial \beta} = \text{diag} \left\{ \frac{1}{\|\beta_1\|_2}, \dots, \frac{1}{\|\beta_L\|_2} \right\} \beta = \mathbf{U} \beta \quad (16)$$

其中, \mathbf{U} 表示系数矩阵; β_L 表示输出权值 β 的第 L 个行向量; $\|\cdot\|_2$ 表示向量的 L_2 范数.综合式 (15)、式 (16), 输出权值 β 迭代求解的表达式如下所示:

$$\beta^{*(t+1)} = [\mathbf{H}_L^T \mathbf{H}_L + \lambda \alpha \mathbf{U}^{(t+1)} + \lambda(1-\alpha)\mathbf{I}]^{-1} \mathbf{H}_L^T \mathbf{Y} \quad (17)$$

$$\mathbf{U}^{(t+1)} = \text{diag} \left\{ \frac{1}{\|\beta_1^{(t)}\|_2}, \dots, \frac{1}{\|\beta_L^{(t)}\|_2} \right\} \quad (18)$$

其中, \mathbf{I} 表示单位矩阵.迭代终止条件如下所示:

$$\|\beta^{*(t+1)} - \beta^{*(t)}\|_{\text{F}} < \kappa \quad (19)$$

其中, κ 表示趋于 0 的正实数, 本文取 10^{-3} .

2.3 鲁棒建模

当训练数据中的噪声或离群点为非高斯分布

时,基于最小二乘算法求解的模型参数并不具备鲁棒性^[32].由于 Laplace 分布具有良好的尖峰、重尾特性,因此其常被作为噪声或离群点的先验分布^[33-34],但实际生产环境中的噪声十分复杂,采用混合概率分布作为噪声的先验分布可以提升模型的鲁棒性^[35].为避免异常数据对 SCN 模型准确性的负面影响,本节采用混合 Laplace 分布作为每个目标建模误差的先验分布,并通过最大后验估计来优化 SCN 模型的输出权值.

根据文献 [36],通过引入潜变量 v , Laplace 分布可以写成高斯分布与指数分布的乘积形式,如下所示:

$$\mathcal{L}(x, v | \mu, \sigma) = \mathcal{N}(x | v, \mu, \sigma) \mathcal{G}(v) = \frac{1}{v^2 \sigma \sqrt{\pi}} \exp \left(-\frac{v^2(x-\mu)^2}{\sigma^2} - \frac{1}{2v^2} \right) \quad (20)$$

其中, μ 与 σ 分别表示高斯分布的均值与标准差. $\mathcal{N}(x | v, \mu, \sigma)$ 与 $\mathcal{G}(v)$ 的表达式为:

$$\mathcal{N}(x | v, \mu, \sigma) = \frac{v}{\sigma \sqrt{\pi}} \exp \left(-\frac{v^2(x-\mu)^2}{\sigma^2} \right) \quad (21)$$

$$\mathcal{G}(v) = \frac{1}{v^3} \exp \left(-\frac{1}{2v^2} \right) \quad (22)$$

对于给定的训练数据集 $\mathbf{D} = \{\mathbf{X}, \mathbf{y}\} = \{(\mathbf{x}_n, y_n)\}_{n=1}^N$, SCN 模型的预测值 \hat{y}_n 与真实值 y_n 之间的关系可表示为:

$$y_n = \hat{y}_n + \varepsilon_n = \mathbf{h}(\mathbf{x}_n) \beta + \varepsilon_n \quad (23)$$

其中, ε_n 表示建模误差.假设 ε_n 服从混合 Laplace 分布,其概率分布可表示为:

$$p(\varepsilon) = \sum_{k=1}^K \omega_k \mathcal{L}(\varepsilon | 0, \sigma_k) \quad (24)$$

其中, $k = 1, 2, \dots, K$; K 表示混合 Laplace 分布的数量; $\omega = \{\omega_1, \omega_2, \dots, \omega_K\}$ 表示混合 Laplace 分布的权值, $\omega_k \geq 0$ 且 $\sum_{k=1}^K \omega_k = 1$; $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_K\}$ 表示混合 Laplace 分布的参数.此时, y_n 的概率分布可表示为:

$$p(y_n | \mathbf{x}_n, \beta, \omega, \sigma) = \sum_{k=1}^K \omega_k \mathcal{L}(y_n | \mathbf{h}(\mathbf{x}_n) \beta, \sigma_k) \quad (25)$$

令 $\mathbf{v} = \{v_n\}_{n=1}^N$, 基于新数据集 $\mathbf{R} = \{\mathbf{X}, \mathbf{y}, \mathbf{v}\}$, 式 (25) 可以改写为:

$$p(y_n, v_n | \mathbf{X}, \beta, \omega, \sigma) = \sum_{k=1}^K \omega_k \mathcal{L}(y_n, v_n | \mathbf{h}(\mathbf{x}_n) \beta, \sigma_k) \quad (26)$$

假设 \mathbf{R} 中的数据均为独立同分布, 其似然函数可表示为:

$$p(\mathbf{y}, \mathbf{v}|\mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}) = \prod_{n=1}^N \prod_{k=1}^K \omega_k \mathcal{L}(y_n, v_n | \mathbf{h}(\mathbf{x}_n) \boldsymbol{\beta}, \sigma_k) \quad (27)$$

此外, 假设每个目标的输出权值服从高斯分布^[35]:

$$p(\boldsymbol{\beta}|\sigma_{\boldsymbol{\beta}}^2) = \frac{1}{(2\pi\sigma_{\boldsymbol{\beta}}^2)^{L/2}} \exp\left(-\frac{\|\boldsymbol{\beta}\|_2^2}{2\sigma_{\boldsymbol{\beta}}^2}\right) \quad (28)$$

基于贝叶斯理论, 输出权值 $\boldsymbol{\beta}$ 的后验分布可表示为:

$$p(\boldsymbol{\beta}|\mathbf{R}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2) \propto p(\mathbf{y}, \mathbf{v}|\mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}) p(\boldsymbol{\beta}|\sigma_{\boldsymbol{\beta}}^2) \quad (29)$$

对上式两边取对数可得:

$$\ln p(\boldsymbol{\beta}|\mathbf{R}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2) = \ln p(\mathbf{y}, \mathbf{v}|\mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}) + \ln p(\boldsymbol{\beta}|\sigma_{\boldsymbol{\beta}}^2) + c \quad (30)$$

其中, c 表示常数. 参数 $\{\boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2\}$ 可通过 EM 算法^[37] 计算式 (31) 的最大后验估计获得:

$$\{\boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2\}^* = \arg \max_{\boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2} \{\ln p(\boldsymbol{\beta}|\mathbf{R}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2)\} \quad (31)$$

EM 算法求解式 (31) 主要包括 E-step 与 M-step 两步, 如下所示:

E-step:

通过引入潜变量 $z_n = \{z_{kn}\}_{k=1}^K$, 在完整数据集 $\mathbf{G} = \{\mathbf{X}, \mathbf{y}, \mathbf{v}, \mathbf{z}\}$ 下, 式 (27) 可改写为:

$$p(\mathbf{y}, \mathbf{v}, \mathbf{z}|\mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}) = \prod_{n=1}^N \prod_{k=1}^K [\omega_k \mathcal{L}(y_n, v_n | \mathbf{h}(\mathbf{x}_n) \boldsymbol{\beta}, \sigma_k)]^{z_{kn}} \quad (32)$$

于是, 式 (30) 可改写为:

$$\ln p(\boldsymbol{\beta}|\mathbf{G}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2) = \ln p(\mathbf{y}, \mathbf{v}, \mathbf{z}|\mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}) + \ln p(\boldsymbol{\beta}|\sigma_{\boldsymbol{\beta}}^2) + c \quad (33)$$

综合式 (28) 与式 (32), 式 (33) 的对数似然函数期望可表示为 (省去了与参数 $\{\boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2\}$ 无关项):

$$\begin{aligned} \mathbb{E}[\ln p(\boldsymbol{\beta}|\mathbf{G}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2)|\mathbf{D}] = & \sum_{n=1}^N \sum_{k=1}^K \left\{ \mathbb{E}[z_{kn} | (\mathbf{x}_n, y_n)] \left(\ln \omega_k - \frac{1}{2} \ln \sigma_k^2 \right) \right\} - \\ & \sum_{n=1}^N \sum_{k=1}^K \left\{ \mathbb{E}[z_{kn} | (\mathbf{x}_n, y_n)] \times \right. \end{aligned}$$

$$\left. \frac{\mathbb{E}[v_n^2 | (\mathbf{x}_n, y_n)] (y_n - \mathbf{h}(\mathbf{x}_n) \boldsymbol{\beta})^2}{\sigma_k^2} \right\} - \frac{L}{2} \ln \sigma_{\boldsymbol{\beta}}^2 - \frac{\|\boldsymbol{\beta}\|_2^2}{2\sigma_{\boldsymbol{\beta}}^2} + c_1 \quad (34)$$

其中, $\mathbb{E}[\cdot]$ 表示期望算子; c_1 表示常数. $\mathbb{E}[z_{kn} | (\mathbf{x}_n, y_n)]$ 与 $\mathbb{E}[v_n^2 | (\mathbf{x}_n, y_n)]$ 的计算公式如下所示^[36-37]:

$$\begin{aligned} \gamma_{kn} = \mathbb{E}[z_{kn} | (\mathbf{x}_n, y_n)] = & \frac{\omega_k \mathcal{L}(y_n | \mathbf{h}(\mathbf{x}_n) \boldsymbol{\beta}, \sigma_k)}{\sum_{k=1}^K \omega_k \mathcal{L}(y_n | \mathbf{h}(\mathbf{x}_n) \boldsymbol{\beta}, \sigma_k)} \quad (35) \end{aligned}$$

$$\chi_{kn} = \mathbb{E}[v_n^2 | (\mathbf{x}_n, y_n)] = \frac{\sigma_k}{\sqrt{2} |y_n - \mathbf{h}(\mathbf{x}_n) \boldsymbol{\beta}|} \quad (36)$$

其中, γ_{kn} 与 χ_{kn} 分别表示 z_{kn} 与 v_n^2 的期望.

M-step:

令 $\frac{\partial \mathbb{E}[\ln p(\boldsymbol{\beta}|\mathbf{G}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2)|\mathbf{D}]}{\partial \{\boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2\}} = 0$, 参数 $\{\boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2\}$ 的更新公式如下所示:

$$\omega_k^{(l+1)} = \frac{\sum_{n=1}^N \gamma_{kn}^{(l)}}{N} \quad (37)$$

$$\sigma_k^{(l+1)} = \sqrt{\frac{2 \sum_{n=1}^N \gamma_{kn}^{(l)} \chi_{kn}^{(l)} (y_n - \mathbf{h}(\mathbf{x}_n) \boldsymbol{\beta}^{(l)})^2}{\sum_{n=1}^N \gamma_{kn}^{(l)}}} \quad (38)$$

$$\sigma_{\boldsymbol{\beta}}^{2(l+1)} = \frac{\|\boldsymbol{\beta}^{(l)}\|_2^2}{L} \quad (39)$$

$$\boldsymbol{\beta}^{(l+1)} = [\mathbf{H}^T(\mathbf{X}) \boldsymbol{\Phi}^{(l+1)} \mathbf{H}(\mathbf{X}) + \mathbf{I}]^{-1} \mathbf{H}^T(\mathbf{X}) \boldsymbol{\Phi}^{(l+1)} \mathbf{y} \quad (40)$$

$$\phi^{(l+1)} = 2\sigma_{\boldsymbol{\beta}}^2 \sum_{k=1}^K \frac{\gamma_{kn}^{(l+1)} \chi_{kn}^{(l+1)}}{\sigma_k^{2(l+1)}} \quad (41)$$

其中, $\boldsymbol{\Phi}^{(l+1)} = \text{diag}\{\phi_n^{(l+1)}\}_{n=1}^N$ 表示惩罚矩阵. EM 算法的收敛条件如下所示:

$$\left| \frac{\mathbb{E}[\ln p(\boldsymbol{\beta}^{(l+1)}|\mathbf{G}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2)|\mathbf{D}]}{\mathbb{E}[\ln p(\boldsymbol{\beta}^{(l)}|\mathbf{G}, \boldsymbol{\omega}, \boldsymbol{\sigma}, \sigma_{\boldsymbol{\beta}}^2)|\mathbf{D}]} - 1 \right| < \kappa \quad (42)$$

2.4 MRI-SCN 算法伪代码

基于上述, 本节给出 MRI-SCN 算法伪代码, 如下所示:

算法 1. MRI-SCN 算法伪代码

输入. 训练集 $\mathbf{D} = \{\mathbf{X}, \mathbf{Y}\} = \{\mathbf{x}_n, \mathbf{y}_n\}_{n=1}^N$.

输出. MRI-SCN 模型与参数 $\boldsymbol{\beta}^*$, \mathbf{w}^* , b^* , \mathbf{q}^* .

- 1) 设置隐含层最大节点数量 L_{\max} , 最大配置次数 T_{\max} , 容忍误差 τ , 隐含层参数选择区间 Υ , 正则化系数 λ 与 α , 约束参数 P , 混合 Laplace 分布数量 K 以及参数 $\{\omega, \sigma, \sigma_{\beta}^2\}$;
- 2) 初始化 $\mathbf{e}_0 = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]^T$, $\Omega, \mathbf{Q}, \mathbf{W} = []$;
- 3) **while** $L \leq L_{\max}$ and $\|\mathbf{e}_0\|_F > \tau$ **do**
- 4) 在区间 Υ 内随机生成权值 $\mathbf{w}_k, \mathbf{q}_k$ 与偏置 b_k ;
- 5) 根据式 (2), 式 (6) ~ 式 (8) 计算 ξ_L , 分别将 \mathbf{w}_L, b_L 存入 \mathbf{W} , 将 \mathbf{q}_L 存入 \mathbf{Q} , 将 ξ_L 存入 Ω ;
- 6) 将 Ω 中 ξ_L 最大值对应的 $\mathbf{w}_L^*, b_L^*, \mathbf{q}_L^*$ 作为新增节点参数, 并令 $\mathbf{H}_L = [\mathbf{h}_1^*, \mathbf{h}_2^*, \dots, \mathbf{h}_L^*]$;
- 7) 根据式 (4) 计算输出权值 β^* ;
- 8) **while** 式 (19) 所示条件未被满足 **do**
- 9) 根据式 (17)、式 (18) 更新 β^* 与 \mathbf{U} ;
- 10) **end while**
- 11) 更新 $\mathbf{e}_0 = \mathbf{H}_L \beta^* - \mathbf{Y}$, $L = L + 1$;
- 12) **end while**
- 13) **for** $i = 1$ to m **do**
- 14) **while** 式 (42) 所示条件未被满足 **do**
- 15) E-step: 根据式 (35)、式 (36) 计算 γ_{kn} 与 χ_{kn} ;
- 16) M-step: 根据式 (37) ~ 式 (39) 计算参数 $\{\omega, \sigma, \sigma_{\beta}^2\}$;
- 17) 根据式 (40)、式 (41) 更新 β^* 与 Φ ;
- 18) **end while**
- 19) **end for**
- 20) 返回 MRI-SCN 模型与参数 $\beta^*, \mathbf{w}^*, b^*, \mathbf{q}^*$.

3 实验测试与结果分析

为便于表述, 将 2.1 节改进的方法记为 I-SCN, 将 2.1 节与 2.2 节共同改进的方法记为 MI-SCN. 实验数据来自北京市某固废焚烧厂 2021 年的历史运行数据, 分别从春、夏、秋、冬 4 个季节中选取, 从而获得 4 组数据集. 每组数据集中训练集的样本数量为 1000, 验证集与测试集的样本数量均为 200, 样本采样时间为 1 min. 所有数据均归一化至 $[0, 1]$ 区间内. 本节实验主要包括 3 部分: 实验一是在原始数据集上, 将 MRI-SCN 与误差反传网络 (BP)、径向基神经网络 (RBF)、随机向量函数链网络 (RVFL)^[11] 以及多输出最小二乘支持向量回归机 (MLS-SVR)^[38] 等不同类型的建模方法进行对比实验; 实验二是在原始数据集上, 将 MRI-SCN 与多目标 SCN (MT-SCN)^[39], 混合高斯-拉普拉斯分布的鲁棒 SCN (MoGL-SCN)^[35], MI-SCN 以及 SCN 等同类型的建模方法进行对比实验; 实验三是在实验二的基础上, 向训练集中引入比例为 $\zeta = \{10\%, 15\%, 20\%, 25\%, 30\%\}$ 的异常值, 即将所选样本对应的输出替换为 $[-0.5, 1.5]$ 上的随机数, 从而获得 4 组

噪声数据集以验证 MRI-SCN 在不同异常值比例下的建模性能. 每种方法独立运行 30 次, 采用平均均方根误差 (Average root mean squared error, aRMSE)、平均绝对误差 (Mean absolute error, MAE) 以及平均确定系数 (Average R^2 , a R^2) 评价模型性能, 计算公式如下所示:

$$\text{aRMSE} = \frac{1}{m} \sum_{i=1}^m \sqrt{\frac{\sum_{j=1}^{N_t} (y_{ij} - \hat{y}_{ij})^2}{N_t}} \times 100\% \quad (43)$$

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m \frac{\sum_{j=1}^{N_t} |y_{ij} - \hat{y}_{ij}|}{N_t} \times 100\% \quad (44)$$

$$\text{aR}^2 = \frac{1}{m} \sum_{i=1}^m \left(1 - \frac{\sum_{j=1}^{N_t} (y_{ij} - \hat{y}_{ij})^2}{\sum_{j=1}^{N_t} (y_{ij} - \bar{y}_{ij})^2} \right) \times 100\% \quad (45)$$

其中, N_t 表示测试集数量; y_{ij} 表示真实值; \hat{y}_{ij} 表示预测值; \bar{y}_{ij} 表示 y_{ij} 的均值.

实验参数设置如下: 在实验一中, BP 的训练次数为 200, 学习率为 0.01; RBF 采用 MATLAB 工具箱中的 newrb 函数实现, 隐含层最大节点数量 $L_{\max} = 60$, 其余参数默认设置; RVFL 隐含层参数区间 $\Upsilon = 1$; MLS-SVR 的正则化参数在区间 $[2^{-10}:1:10]$ 内选择. 每种建模方法的容忍误差 $\tau = 10^{-3}$, BP 与 RVFL 的隐含层最大节点数量与 RBF 一致, 激活函数为 sigmoid. 在实验二中, 最大配置次数 $T_{\max} = 50$, 隐含层参数区间 $\Upsilon = [1:1:10]$, 隐含层最大节点数量、激活函数以及容忍误差与实验一保持一致; MT-SCN 的正则化参数选择区间与 MLS-SVR 一致; MRI-SCN 与 MI-SCN 的正则化参数 $\lambda = 2^{-5}$, $\alpha = 0.2$, 约束参数 $P = 10$; MRI-SCN 与 MoGL-SCN 的混合分布数量 $K = 3$, 对应的权重与分布参数分别为 $\omega_1 = 0.40$, $\omega_2 = 0.30$, $\omega_3 = 0.30$, $\sigma_1 = 0.10$, $\sigma_2 = 0.13$, $\sigma_3 = 0.12$, $\sigma_{\beta}^2 = 0.15$. 在实验三中, MRI-SCN 与 MI-SCN 的正则化参数 $\lambda = 1.00$, $\alpha = 0.40$, 其余参数设置保持不变. 以上超参数由验证集获得.

表 1 与表 2 统计了 MRI-SCN 与其他建模方法在原始数据集上的实验结果. 从表 1 中可以看出, MLS-SVR 与 MRI-SCN 通过对模型参数进行稀疏约束, 利用多目标间的相关性提升了建模精度, 但 MLS-SVR 未考虑到异常数据对模型性能的负面影响, 所以其模型精度低于 MRI-SCN. 从表 2 中可以看出, 由于 SCN 未利用多目标间相关性且对训练

表 1 MRI-SCN 与不同类型建模方法在原始数据集上的对比实验结果

Table 1 Results of experiments comparing MRI-SCN with the different type of modeling methods on the original dataset

数据集	BP	RBF	RVFL	MLS-SVR	MRI-SCN
春季	5.80; 4.57; 86.57	5.40; 3.74; 89.14	4.55; 3.59; 92.36	3.41; 2.38; 95.70	3.12; 2.34; 96.38
夏季	5.63; 4.33; 87.48	4.85; 3.75; 91.49	4.60; 3.63; 92.14	3.37; 2.52; 95.90	3.16; 2.26; 96.25
秋季	5.39; 4.27; 89.03	4.92; 3.73; 91.25	4.52; 3.55; 92.44	3.44; 2.69; 95.41	3.08; 2.31; 96.48
冬季	5.30; 4.19; 89.93	5.72; 4.42; 89.14	4.96; 3.91; 91.46	3.52; 2.53; 95.61	3.10; 2.33; 96.68

表 2 MRI-SCN 与同类型建模方法在原始数据集上的对比实验结果

Table 2 Results of experiments comparing MRI-SCN with the same type of modeling methods on the original dataset

数据集	SCN	MI-SCN	MT-SCN	MoGL-SCN	MRI-SCN
春季	3.67; 2.86; 95.07	3.45; 2.64; 95.62	3.46; 2.67; 95.59	3.61; 2.68; 94.50	3.12; 2.34; 96.38
夏季	3.70; 2.80; 95.40	3.43; 2.55; 95.65	3.32; 2.57; 95.84	3.40; 2.67; 95.63	3.16; 2.26; 96.25
秋季	3.63; 2.73; 95.17	3.31; 2.56; 95.91	3.34; 2.61; 95.92	3.52; 2.76; 95.42	3.08; 2.31; 96.48
冬季	3.74; 2.94; 95.23	3.56; 2.73; 95.33	3.49; 2.81; 95.84	3.55; 2.31; 95.82	3.10; 2.33; 96.68

集中的正常样本与异常样本进行无差别学习, 所以 MT-SCN、MoGL-SCN 等建模方法的表现均优于 SCN. 综合对比表 1 与表 2 可知, 与 BP、RBF 相比, SCN 的预测精度相对较高; 与 RVFL 相比, SCN 通过监督不等式选择优质节点添加至隐含层, 从而提升了建模质量. 而 MRI-SCN 在进行多目标建模的同时兼顾模型鲁棒性, 其在 3 个评价指标上的表现均优于其他对比方法, 以春季数据集为例, MRI-SCN 在 aRMSE 评价指标上比 RVFL、MLS-SVR 分别下降 31.43%、8.50%; 比 SCN、MT-SCN、MoGL-SCN 分别下降 14.99%、9.83%、13.57%.

为测试不同范数约束对模型性能的影响, 以 2.1 节改进的方法 I-SCN 为基础, 将 F 范数约束的 I-SCN 记为 I-SCN-F, 将 $L_{2,1}$ 范数约束的 I-SCN 记

为 I-SCN- $L_{2,1}$, MI-SCN 则表示 F 范数与 $L_{2,1}$ 范数共同约束的 I-SCN. 由图 3 可知, 相对于 F 范数约束, $L_{2,1}$ 范数约束由于具有稀疏性, 其利用炉温与烟气含氧量间的相关性来提升模型质量, 所以 I-SCN- $L_{2,1}$ 的建模精度高于 I-SCN-F. 而将上述两种范数约束相结合构成矩阵弹性网, 可使模型输出权重具有稀疏性且相对较小, 从而进一步降低了模型过拟合的风险, 所以 MI-SCN 的预测精度相对较高.

为进一步验证 MRI-SCN 的鲁棒性, 表 3 统计了 SCN、MI-SCN、MT-SCN、MoGL-SCN 以及 MRI-SCN 等建模方法在 4 组噪声数据集上的实验结果. 由表 3 可知, 由于 SCN 对正常样本与异常样本进行无差别学习, 所以模型质量相对较差. MI-SCN 与 MT-SCN 利用炉温与烟气含氧量间的相关

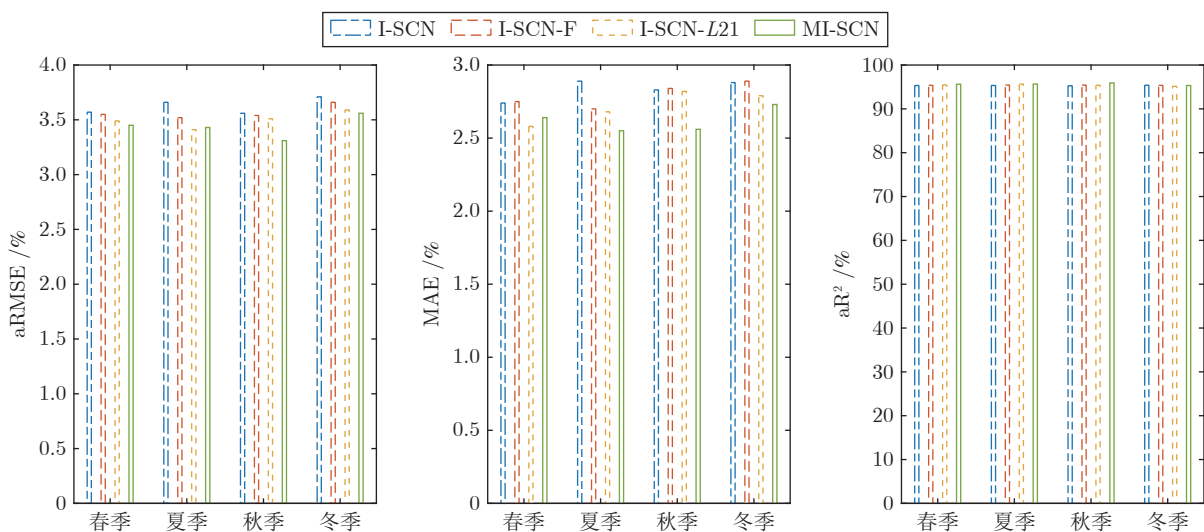


图 3 不同范数约束在原始数据集上的实验结果

Fig. 3 Results of experiments with different paradigm constraints on the original dataset

性来提升模型精度, 所以它们的建模性能均优于 SCN. 但随着异常值比例的增加, MI-SCN 与 MT-SCN 的模型性能退化较大, 缺乏鲁棒性. 虽然 MoGL-SCN 采用基于混合分布的建模策略提升了模型鲁棒性, 但其未利用多目标间的相关性来提升建模精度, 而 MRI-SCN 利用矩阵弹性网将炉温与烟气含氧量间的相关性假设转化为对模型参数的稀疏约束, 并在此基础上利用混合 Laplace 分布作为

每个目标建模误差的先验分布, 通过最大后验估计优化模型输出权值, 从而有效提升了模型精度与鲁棒性, 所以 MRI-SCN 在 4 组噪声数据集上的表现优于其他对比方法.

从图 4 的曲线中可以看出, MI-SCN 与 MT-SCN 由于缺乏鲁棒性, 随着异常值比例的增加, MI-SCN 与 MT-SCN 的建模误差增长幅度较快, 难以实现炉温与烟气含氧量的准确预测. 而 MoGL-

表 3 四组噪声数据集上的实验结果
Table 3 Results of experiments on the four noisy datasets

数据集	ζ	SCN	MI-SCN	MT-SCN	MoGL-SCN	MRI-SCN
春季	10%	6.62; 5.15; 83.95	5.80; 4.57; 87.67	5.59; 4.40; 88.55	3.86; 2.87; 94.43	3.65; 2.74; 95.07
	15%	7.59; 5.97; 78.84	6.39; 5.07; 84.97	6.35; 5.05; 85.16	4.00; 2.95; 94.02	3.87; 2.89; 94.45
	20%	8.96; 7.11; 70.38	7.33; 5.80; 80.14	7.36; 5.88; 79.99	4.26; 3.11; 93.21	4.03; 3.02; 93.95
	25%	10.38; 8.05; 60.18	8.26; 6.46; 74.83	8.58; 6.74; 72.80	4.48; 3.26; 92.53	4.34; 3.24; 92.95
	30%	11.17; 8.71; 53.98	8.70; 6.86; 72.21	9.40; 7.41; 67.46	4.82; 3.50; 91.25	4.56; 3.38; 92.28
夏季	10%	6.31; 4.90; 85.43	5.77; 4.56; 87.74	5.50; 4.30; 88.88	4.00; 2.93; 93.81	3.76; 2.78; 94.66
	15%	7.38; 5.71; 79.98	6.48; 5.13; 84.56	6.24; 4.89; 85.75	4.38; 3.14; 92.66	3.96; 2.92; 94.07
	20%	8.91; 7.00; 70.13	7.45; 5.94; 79.13	7.49; 5.96; 78.73	4.38; 3.17; 92.55	4.22; 3.13; 93.10
	25%	9.40; 7.48; 66.62	7.89; 6.34; 76.58	7.93; 6.35; 76.22	4.68; 3.33; 91.83	4.45; 3.27; 92.43
	30%	10.39; 8.21; 59.58	8.61; 6.90; 72.39	8.85; 7.09; 70.72	5.08; 3.65; 89.94	4.65; 3.41; 91.83
秋季	10%	6.40; 5.04; 85.18	5.98; 4.79; 86.83	5.57; 4.43; 88.69	3.73; 2.75; 94.76	3.46; 2.60; 95.53
	15%	7.33; 5.72; 80.40	6.33; 5.01; 85.27	6.10; 4.78; 86.37	3.80; 2.80; 94.63	3.61; 2.72; 95.11
	20%	8.90; 6.85; 71.15	7.19; 5.66; 80.85	7.40; 5.74; 79.96	4.03; 2.95; 94.02	3.86; 2.88; 94.45
	25%	9.82; 7.52; 65.08	7.62; 6.00; 78.77	8.00; 6.20; 76.81	4.11; 2.98; 93.71	3.96; 2.96; 94.16
	30%	10.79; 8.25; 57.89	8.28; 6.44; 74.97	8.75; 6.73; 72.24	4.50; 3.22; 92.52	4.26; 3.18; 93.28
冬季	10%	6.86; 5.34; 83.29	6.55; 5.16; 84.72	6.22; 4.86; 86.29	4.10; 3.04; 94.20	3.93; 2.98; 94.56
	15%	7.94; 6.20; 78.14	7.27; 5.76; 81.77	6.97; 5.50; 83.25	4.40; 3.21; 93.30	4.27; 3.18; 93.73
	20%	9.40; 7.36; 69.16	8.05; 6.32; 77.44	7.91; 6.20; 78.23	4.55; 3.33; 92.83	4.37; 3.28; 93.32
	25%	10.56; 8.18; 60.47	8.81; 6.94; 72.48	8.90; 6.97; 71.96	4.88; 3.56; 91.40	4.62; 3.45; 92.41
	30%	11.26; 8.74; 55.94	9.50; 7.48; 68.70	9.65; 7.57; 67.57	5.02; 3.64; 91.25	4.83; 3.60; 91.72

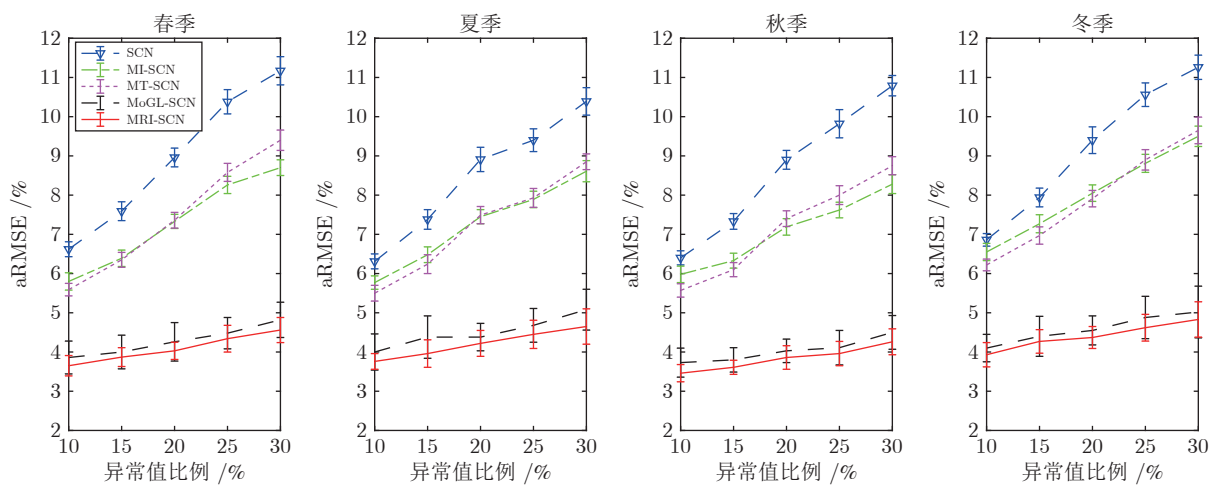


图 4 不同异常值比例下的 aRMSE 变化曲线

Fig.4 aRMSE change curves with different outlier percentages

SCN 与 MRI-SCN 采用鲁棒建模策略提升了模型鲁棒性, 所以建模误差增长幅度较为缓慢, 尤其是 MRI-SCN 在不同异常值比例下的模型精度相对较好, 这充分说明了在异常数据影响下建立炉温与烟气含氧量鲁棒预测模型的必要性。

从图 5 中可以看出, 当 $\zeta=20\%$ 时, MRI-SCN 的预测结果比 SCN 更加靠近对角线, 这验证了本文所提改进策略的有效性. 从预测误差概率分布曲线中可以看出, 由于 MoGL-SCN 与 MRI-SCN 采用鲁棒建模策略来抑制异常数据对模型准确性的负面影响, 所以它们预测误差小于 SCN、MI-SCN 以及 MT-SCN. 此外, MRI-SCN 的预测误差概率分布曲线在 0 附近的尖峰特征比 MoGL-SCN 更加明

显, 所以 MRI-SCN 的预测精度相对较高。

图 6 是在 $\zeta=30\%$ 的情况下, SCN 与 MRI-SCN 在 4 组噪声数据集上的炉温与烟气含氧量预测误差曲线及其模型输出权值. 从图 6 中可以看出, 在异常数据影响下, 采用 MRI-SCN 建立的模型对炉温与烟气含氧量的预测精度相对较高且模型输出权值较小, 这在一定程度上可以提升模型的泛化能力。

表 4 统计了 BP、SCN 以及 MRI-SCN 等建模方法运行 30 次所花费的时间. 从表 4 中可以看出, MI-SCN、MRI-SCN 的建模速度均比 MT-SCN 与 MoGL-SCN 快, 这说明利用参数自适应变化的监督机制可以提升 MI-SCN 与 MRI-SCN 隐含层参数的配置效率。

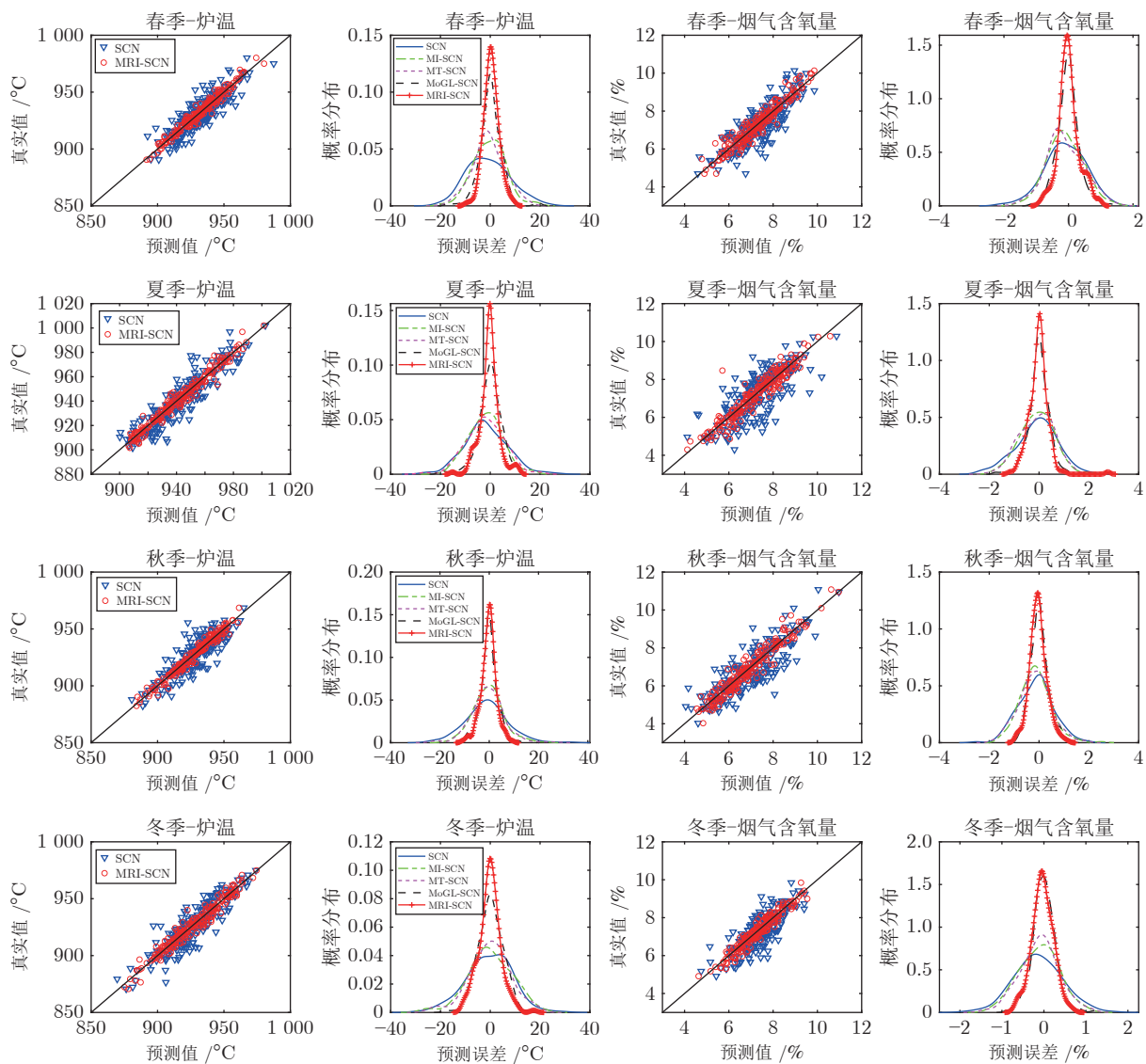


图 5 炉温与烟气含氧量散点图及预测误差概率分布曲线 ($\zeta = 20\%$)

Fig.5 Scatterplot of furnace temperature, flue gas oxygen content and probability distribution curves of prediction error ($\zeta = 20\%$)

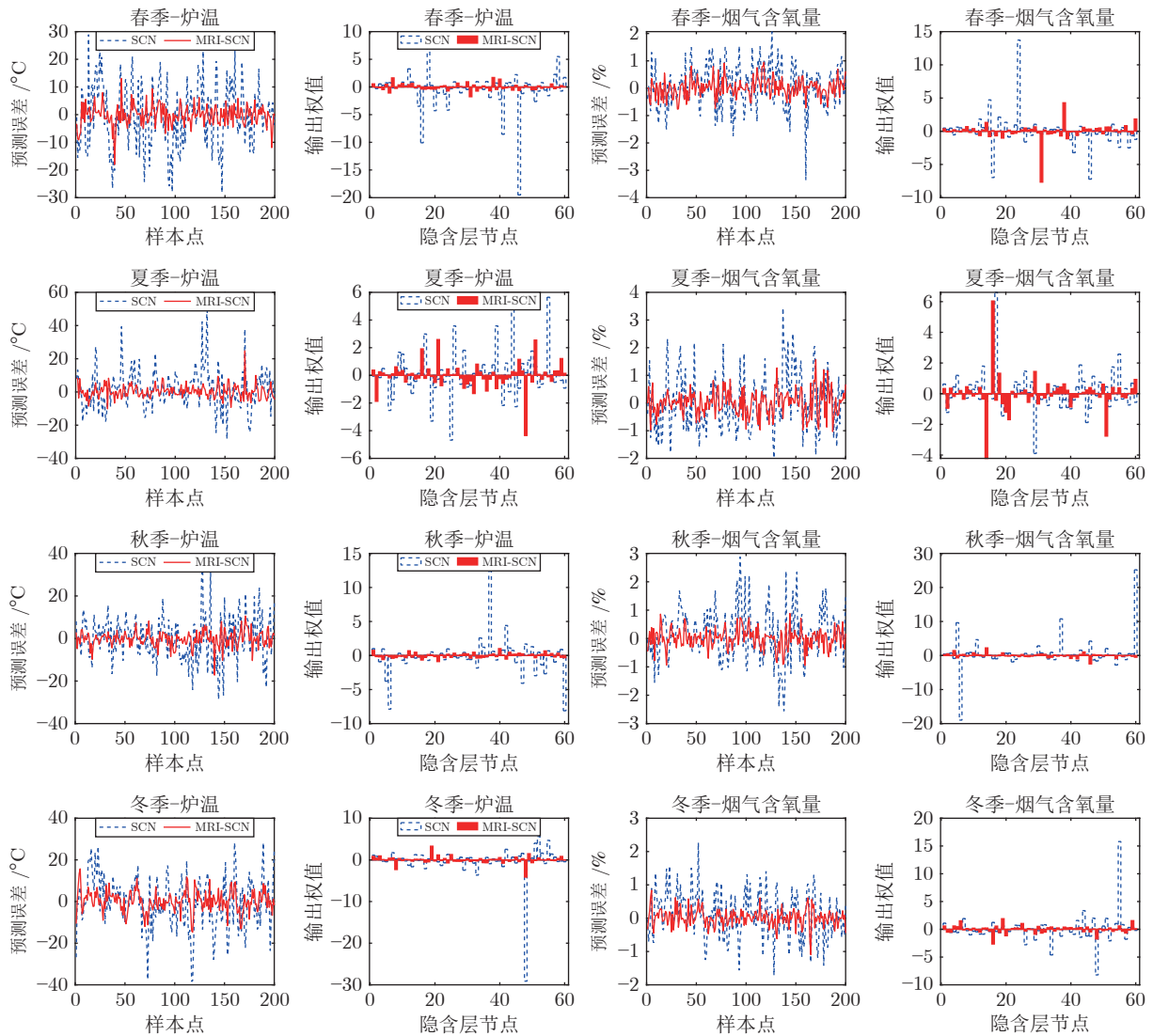
图 6 炉温与烟气含氧量预测误差曲线及模型输出权值 ($\zeta = 30\%$)Fig.6 Furnace temperature, flue gas oxygen content prediction error curves and model output weights ($\zeta = 30\%$)

表 4 不同建模方法运行 30 次的时间对比

Table 4 Comparison of time for 30 runs of different modeling methods

方法	BP	RBF	SCN	MI-SCN	MT-SCN	MoGL-SCN	MRI-SCN
时间 (s)	21.43	31.70	5.23	16.29	36.22	37.78	28.17

4 结束语

为准确预测 MSWI 过程的炉温与烟气含氧量,提出了一种基于改进 SCN 的多目标鲁棒建模方法 (MRI-SCN), 并通过实验验证了 MRI-SCN 建模方法的有效性与优越性. 主要贡献总结如下:

第一, 将标准前馈与级联构造相融合, 在参数自适应的不等式监督机制下分配隐含层参数, 从而增强了 SCN 隐含层映射多样性, 并提高了模型参

数的配置效率.

第二, 利用 F 范数与 $L_{2,1}$ 范数正则项构建矩阵弹性网对模型输出权值进行稀疏约束, 利用炉温与烟气含氧量之间的相关性进一步提升了建模精度.

第三, 将混合 Laplace 分布作为每个目标建模误差的先验分布, 并采用最大后验估计重新评估 SCN 模型的输出权值, 使学习到的模型具有良好的鲁棒性.

实验结果表明, 在异常数据影响下, 基于 MRI-

SCN 建立的预测模型可以较为准确地预测炉温与烟气含氧量的变化趋势, 从而为 MSWI 过程的多目标协同优化控制奠定了良好的基础, MRI-SCN 建模方法在工业过程参数预测建模领域具有一定的应用潜力. 考虑到 MSWI 过程的参数变化具有时变性, 提升 MRI-SCN 模型对复杂工况的自适应能力是后续研究工作的重点.

附录 A

表 A1 多目标鲁棒预测模型输入变量明细
Table A1 Input variable details of multi-target robust prediction model

序号	变量名称	单位
1	进料器左内侧速度	%
2	进料器左外侧速度	%
3	进料器右内侧速度	%
4	进料器右外侧速度	%
5	干燥炉排左内侧速度	%
6	干燥炉排左外侧速度	%
7	干燥炉排右内侧速度	%
8	干燥炉排右外侧速度	%
9	干燥炉排左 1 空气流量	km ³ N/h
10	干燥炉排右 1 空气流量	km ³ N/h
11	干燥炉排左 2 空气流量	km ³ N/h
12	干燥炉排右 2 空气流量	km ³ N/h
13	燃烧炉排左 1-1 段空气流量	km ³ N/h
14	燃烧炉排右 1-1 段空气流量	km ³ N/h
15	燃烧炉排左 1-2 段空气流量	km ³ N/h
16	燃烧炉排右 1-2 段空气流量	km ³ N/h
17	燃烧炉排左 2-1 段空气流量	km ³ N/h
18	燃烧炉排右 2-1 段空气流量	km ³ N/h
19	燃烧炉排左 2-2 段空气流量	km ³ N/h
20	燃烧炉排右 2-2 段空气流量	km ³ N/h
21	燃烬炉排左空气流量	km ³ N/h
22	燃烬炉排右空气流量	km ³ N/h
23	二次风量	km ³ N/h
24	一次风机出口空气压力	kPa
25	一次空气加热器出口空气温度	℃
26	干燥炉排左内侧温度	℃
27	干燥炉排左外侧温度	℃
28	干燥炉排右内侧温度	℃
29	干燥炉排右外侧温度	℃
30	燃烧炉排 1-1 段左内侧温度	℃
31	燃烧炉排 1-1 段左外侧温度	℃
32	燃烧炉排 1-1 段右内侧温度	℃
33	燃烧炉排 1-1 段右外侧温度	℃
34	燃烧炉排 1-2 段左内侧温度	℃

表 A1 多目标鲁棒预测模型输入变量明细 (续表)

Table A1 Input variable details of multi-target robust prediction model (continued table)

序号	变量名称	单位
35	燃烧炉排 1-2 段左外侧温度	℃
36	燃烧炉排 1-2 段右内侧温度	℃
37	燃烧炉排 1-2 段右外侧温度	℃
38	燃烧炉排 2-1 段左内侧温度	℃
39	燃烧炉排 2-1 段左外侧温度	℃
40	燃烧炉排 2-1 段右内侧温度	℃
41	燃烧炉排 2-1 段右外侧温度	℃
42	燃烧炉排 2-2 段左内侧温度	℃
43	燃烧炉排 2-2 段左外侧温度	℃
44	燃烧炉排 2-2 段右内侧温度	℃
45	燃烧炉排 2-2 段右外侧温度	℃
46	当前时刻的炉温	℃
47	当前时刻的烟气含氧量	%

References

- Li Y, Zhao X G, Li Y B, Li X. Waste incineration industry and development policies in China. *Waste Management*, 2015, **46**(8): 234–241
- Tang Jian, Xia Heng, Yu Wen, Qiao Jun-Fei. Research status and prospects of intelligent optimization control for municipal solid waste incineration process. *Acta Automatica Sinica*, 2023, **49**(10): 2019–2059
(汤健, 夏恒, 余文, 乔俊飞. 城市固废焚烧过程智能优化控制研究现状与展望. *自动化学报*, 2023, **49**(10): 2019–2059)
- Yan Ai-Jun, Hu Kai-Cheng. Multi-objective optimization setting method for temperature control of municipal solid waste incinerator. *Control Theory & Applications*, 2023, **40**(4): 693–701
(严爱军, 胡开成. 城市生活垃圾焚烧炉温控制的多目标优化设定方法. *控制理论与应用*, 2023, **40**(4): 693–701)
- Sun R, Ismail T M, Ren X H, Abd El-Salam M. Numerical and experimental studies on effects of moisture content on combustion characteristics of simulated municipal solid wastes in a fixed bed. *Waste Management*, 2015, **39**(5): 166–178
- Magnanelli E, Tranàs O L, Carlsson P, Mosby J, Becidan M. Dynamic modeling of municipal solid waste incineration. *Energy*, 2020, **299**(10): Article No. 118426
- Jiang Ke, Jiang Zhao-Hui, Xie Yong-Fang, Pan Dong, Gui Wei-Hua. Online prediction method for silic on content of molten iron in blast furnace based on dynamic attention deep transfer network. *Acta Automatica Sinica*, 2023, **49**(5): 949–963
(蒋珂, 蒋朝辉, 谢永芳, 潘冬, 桂卫华. 基于动态注意力深度迁移网络的高炉铁水硅含量在线预测方法. *自动化学报*, 2023, **49**(5): 949–963)
- Zhou X F, Zhai N J, Li S A, Shi H B. Time series prediction method of industrial process with limited data based on transfer learning. *IEEE Transactions on Industrial Informatics*, 2023, **19**(5): 6872–6882
- He H J, Meng X, Tang J, Qiao J F. A novel self-organizing TS fuzzy neural network for furnace temperature prediction in MSWI process. *Neural Computing and Applications*, 2022, **34**(12): 9759–9776
- Guo Hai-Tao, Tang Jian, Ding Hai-Xu, Qiao Jun-Fei. Combustion states recognition method of mswi process based on mixed data enhancement. *Acta Automatica Sinica*, 2024, **50**(3): 560–575
(郭海涛, 汤健, 丁海旭, 乔俊飞. 基于混合数据增强的 MSWI 过程燃烧状态识别. *自动化学报*, 2024, **50**(3): 560–575)
- Qiao J F, Sun J, Meng X. Event-triggered adaptive model predictive control of oxygen content for municipal solid waste incineration process. *IEEE Transactions on Automation Science and Engineering*, 2024, **21**(1): 463–474

- 11 Pao Y H, Park G H, Sobajic D J. Learning and generalization characteristics of random vector functional-link net. *Neurocomputing*, 1994, **6**(2): 163–180
- 12 Wang D H, Li M. Stochastic configuration networks: Fundamentals and algorithms. *IEEE Transactions on Cybernetics*, 2017, **47**(10): 3346–3479
- 13 Li K, Yang C C, Wang W, Qiao J F. An improved stochastic configuration network for concentration prediction in wastewater treatment process. *Information Sciences*, 2023, **622**(4): 148–160
- 14 Li X, He Y, Ding J, Luan F, Zhang D. Predicting hot-strip finish rolling thickness using stochastic configuration networks. *Information Sciences*, 2022, **611**(9): 677–689
- 15 Lu J, Ding J L. Construction of prediction intervals for carbon residual of crude oil based on deep stochastic configuration networks. *Information Sciences*, 2019, **486**(6): 119–132
- 16 Hu Kai-Cheng, Yan Ai-Jun, Wang Dian-Hui. Nonlinear model predictive control of furnace temperature for a municipal solid waste incineration process [Online], available: <http://kns.cnki.net/kcms/detail/44.1240.TP.20230330.0900.006.html>, March 22, 2024
(胡开成, 严爱军, 王殿辉. 城市固废焚烧过程炉温非线性模型预测控制 [Online], available: <http://kns.cnki.net/kcms/detail/44.1240.TP.20230330.0900.006.html>, 2024-03-22)
- 17 Yan A J, Guo J C, Wang D H. Heterogeneous feature ensemble modeling with stochastic configuration networks for predicting furnace temperature of a municipal solid waste incineration process. *Neural Computing and Applications*, 2022, **34**(18): 15807–15819
- 18 Ding H X, Tang J, Qiao J F. MIMO modeling and multi-loop control based on neural network for municipal solid waste incineration. *Control Engineering Practice*, 2022, **127**(10): Article No. 105280
- 19 Borchani H, Varando G, Bielza C, Larrañaga P. A survey on multi-output regression. *Wiley Interdisciplinary Reviews Data Mining & Knowledge Discovery*, 2015, **5**(5): 216–233
- 20 Kiliçarslan S, Közkurt C, Baş S, Elen A. Detection and classification of pneumonia using novel superior exponential (SupEx) activation function in convolutional neural networks. *Expert Systems With Applications*, 2023, **217**(5): Article No. 119503
- 21 Gao Z, Yu W, Yan J. Neuro adaptive fault-tolerant control embedded with diversified activating functions with application to auto-driving vehicles under fading actuation. *IEEE Transactions on Neural Networks and Learning Systems*, DOI: 10.1109/TNNLS.2023.3248100
- 22 Li F J, Qiao J F, Han H G, Yang C L. A self-organizing cascade neural network with random weights for nonlinear system modeling. *Applied Soft Computing*, 2016, **42**(5): 184–193
- 23 Luo H Y, Han G L, Wu X T, Liu P X, Yang H, Zhang X. Cascaded hourglass feature fusing network for saliency detection. *Neurocomputing*, 2021, **428**(3): 206–217
- 24 Li J P, Hua C C, Qian J L, Guan X P. Low-rank based multi-input multi-output Takagi-Sugeno fuzzy modeling for prediction of molten iron quality in blast furnace. *Fuzzy Sets and System*, 2021, **421**(9): 178–192
- 25 Arashloo S R, Kittler J. Multi-target regression via non-linear output structure learning. *Neurocomputing*, 2022, **492**(7): 572–580
- 26 Tak N, İnan D. Type-1 fuzzy forecasting functions with elastic net regularization. *Expert Systems With Applications*, 2022, **199**(8): Article No. 116916
- 27 Nie F P, Huang H, Cai X, Ding C. Efficient and robust feature selection via joint $L_{2,1}$ -norms minimization. In: Proceedings of Advances in Neural Information Processing Systems. Vancouver, Canada: MIT Press, 2010. 1813–1821
- 28 Xiang S M, Nie F P, Meng G F, Pan C H, Zhang C S. Discriminative least squares regression for multiclass classification and feature selection. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, **23**(11): 1738–1754
- 29 Li Y H, Hu L, Gao W F. Multi-label feature selection via robust flexible sparse regularization. *Pattern Recognition*, 2023, **134**(2): Article No. 109074
- 30 Lv S H, Zhao H Q, Zhou L J. Robust proportionate normalized least mean M-estimate algorithm for block-sparse system identification. *IEEE Transactions on Circuits and Systems-II Express Briefs*, 2022, **69**(1): 234–238
- 31 Wang Q, He X, Jiang X, Li X L. Robust bi-stochastic graph regularized matrix factorization for data clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, **44**(1): 390–403
- 32 Yang X Y, Mu Y F, Cao K, Lv M Z, Peng B, Zhang Y, et al. Robust kernel recursive adaptive filtering algorithms based on M-estimate. *Signal Processing*, 2023, **207**(6): Article No. 108952
- 33 Duong N C, Speyer J L, Idan M. Laplace estimation for scalar linear systems. *Automatica*, 2022, **144**(10): Article No. 110301
- 34 Liang Z Z, Zhang L. L_1 -norm discriminant analysis via Bhattacharyya error bounds under Laplace distributions. *Pattern Recognition*, 2023, **141**(9): Article No. 109609
- 35 Lu J, Ding J L. Mixed-distribution based robust stochastic configuration networks for prediction interval construction. *IEEE Transactions on Industrial Informatics*, 2020, **16**(8): 5099–5109
- 36 Song W X, Yao W X, Xing Y R. Robust mixture regression model fitting by Laplace distribution. *Computational Statistics and Data Analysis*, 2014, **71**(3): 128–137
- 37 Phillips R F. Least absolute deviations estimation via the EM algorithm. *Statistics and Computing*, 2002, **12**(3): 281–285
- 38 Xu S, An X, Qiao X D, Zhu L J, Li L. Multi-output least-squares support vector regression machines. *Pattern Recognition Letters*, 2013, **34**(9): 1078–1084
- 39 Wang Q J, Hong Q Q, Wu S, Dai W. Multi-target stochastic configuration network and applications. *IEEE Transactions on Artificial Intelligence*, 2023, **4**(2): 338–348



胡开成 北京工业大学信息学部博士研究生。主要研究方向为复杂过程建模与智能优化控制。

E-mail: hukaicheng@emails.bjut.edu.cn
(HU Kai-Cheng Ph.D. candidate at the Faculty of Information Technology, Beijing University of Technology. His research interest covers complex process modeling and intelligent optimization control.)



严爱军 北京工业大学信息学部教授。主要研究方向为复杂过程建模与智能优化控制。本文通信作者。

E-mail: yanaijun@bjut.edu.cn
(YAN Ai-Jun Professor at the Faculty of Information Technology, Beijing University of Technology.

His research interest covers complex process modeling and intelligent optimization control. Corresponding author of this paper.)



汤健 北京工业大学信息学部教授。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。

E-mail: freeflytang@bjut.edu.cn
(TANG Jian Professor at the Faculty of Information Technology, Beijing University of Technology.

His research interest covers small sample data modeling and intelligent control of municipal solid waste treatment process.)