

基于AI集成机器学习的微小通道流动沸腾传热预测

Prediction of flow boiling heat transfer in micro/mini-channels based on integrated machine learning

期刊：	中国科学: 技术科学
稿件ID：	SST-2025-0403.R3
稿件栏目：	论文
作者提交日期：	2026-04-02
参与作者列表：	李蔚, 何欢欢, 董源
关键词：	流动沸腾传热, 集成机器学习, 传热系数, 谱聚类, XGBoost
英文关键词：	low boiling heat transfer, integrated machine learning, heat transfer coefficient, spectral clustering, XGBoost
学科领域：	能源与动力工程
专题：	

For Review Only

中国科学: 技术科学 2025 年 第 55 卷 第*期: ** ~**
SCIENTIA SINICA Technologica tech.scichina.com
《中国科学》杂志社
SCIENCE CHINA PRESS

论 文

基于 AI 集成机器学习的微小通道流动沸腾传热预测

李蔚[1*], 何欢欢[1], 董源[2]

1. 浙江大学能源工程学院, 杭州 310027
2. 杭州电子科技大学机械工程学院, 杭州 310018
* E-mail: weili96@zju.edu.cn

国家自然科学基金国际合作研究重点项目(批准号: 52320105001) 资助项目

摘要 针对微小通道流动沸腾传热系数预测问题, 开发了一种融合谱聚类与极端梯度提升(eXtreme gradient boosting, XGBoost)算法的集成机器学习模型. 通过整合来自 42 篇文献的 8309 个实验数据点, 构建了涵盖 24 种流体, 通道水力直径范围为 0.19 - 6.5mm 的综合数据库, 并采用最大信息系数法从 47 个初始特征中筛选出 37 个参数作为模型输入. 鉴于流动状态对传热特性的影响, 模型首先利用气液相表观速度通过谱聚类将数据划分为代表五类流动状态的类别, 进而训练 XGBoost 分类模型(准确率达 99%)预测样本属于各类别的概率, 并为每个类别单独建立 XGBoost 回归子模型. 最终通过加权融合分类概率与回归结果实现集成预测. 测试表明, 集成模型平均绝对相对误差（*MARD*）和决定系数(R^2)分别为 7.41%和 0.979, 显著优于传统经验关联式. 该方法在保持高精度的同时, 通过流型识别增强了模型的可解释性, 为流动沸腾传热预测提供了新思路.

关键词 流动沸腾传热, 集成机器学习, 传热系数, 谱聚类, XGBoost

1 引言

两相微通道换热器能够利用冷却工质的相变潜热实现高热流密度传热与良好的表面温度均匀性, 因而被广泛视为一种高效的热管理方案, 目前已普遍应用于电力电子系统、电动汽车、医疗设备、聚变反应堆及航空航天等多个领域[1,2]. 尽管关于该类换热器的研究已取得丰富成果, 但其性能参数的可靠预测工具的开发仍是当前研究的重点之一.

与传统通道相比, 微通道中粘性力、表面张力、重力与惯性力的相对重要性发生显著变化, 使其流动与传热行为更为复杂[3]. 研究表明, 微通道内流动

引用格式: 李蔚, 何欢欢, 董源. 基于 AI 集成机器学习的微小通道流动沸腾传热预测. 中国科学: 技术科学,
Li W, He H H, Dong Y. Prediction of flow boiling heat transfer in micro/mini-channels based on integrated machine learning (in Chinese). Sci Sin Tech, 2026,

© 2025《中国科学》杂志社

For Review Only

沸腾传热强烈依赖于流型[4-7]. 如图 1 所示, 微通道中常见观测到的流动状态包括泡状流、段塞流、环状流和雾状流[7,8]; 而在常规通道中, 低气液流速下可能出现分层流, 且随流速升高再转变为段塞流或环状流[10,11]. 泡状流与雾状流通常出现在极端工况下, 通常不易被观测[12]. 流型识别对构建基于物理机制的预测模型具有重要意义, 因此已发展了多种流型图, 常用坐标参数包括气液相表观速度[13,14]、质量流量与蒸汽质量[10,12,15], 以及雷诺数、韦伯数[16]等无量纲数. 然而, 目前仍没有一个适用于宽泛工况的通用流型图. 尽管部分研究者认为基于流型的传热模型有望提高传热预测精度, 并尝试建立流型与传热预测间的关联, 但微通道流动沸腾传热往往存在更加复杂的流型转变和热行为, 要准确预测传热系数(Heat transfer coefficients, HTCs)并非易事.

在流动沸腾 HTCs 的预测中, 经验或半经验关联式是目前较为常用的工具之一. 这类关联式通常基于特定实验数据建立, 往往仅适用于有限的工况范围, 一旦超出该范围, 预测性能会明显下降. 其中, 基于多来源实验数据的大型数据库构建的通用型关联式具有更广的适用性和更高的可靠性, 因为其覆盖了更多流体类型和工况条件. 例如, 针对微小通道的流动沸腾, Berstch 等人[17]基于 14 项研究的 3899 个数据点提出了一个复合关联式, 平均预测误差低于 30%; Li 和 Wu[18]利用 3700 个数据点, 发现努塞尔数与无因次参数 $BdRe^{0.36}$ 有很强的相关性, 并考虑了沸

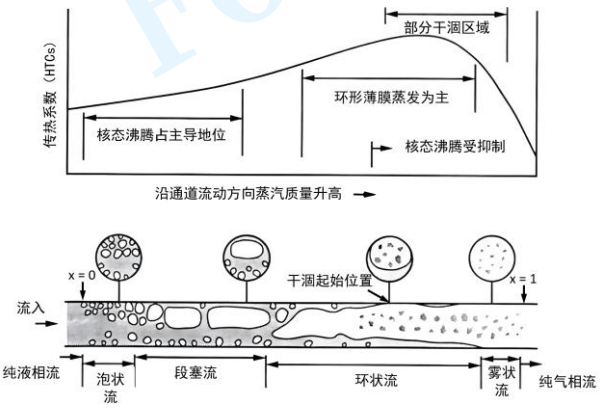


图 1 水平管内流动沸腾的流动状态和传热机理示意图及传热系数的定性变化[9]

Figure 1 Schematic of flow patterns and the corresponding heat transfer mechanisms and qualitative variation of the HTCs for flow boiling in a horizontal tube[9].

腾数的影响提出了一个简洁的关联式, 平均误差为 26.13%; Sun 和 Mishima[19]基于 11 种流体的 2505 个数据点所建关联式的平均误差为 30.78%; 而 Kim 和 Mudawar[20]则从 31 个数据源汇总了 10805 个预干涸数据, 所提出的关联式平均误差为 20.3%, 表现出良好的预测性能. 这些关联式多基于对流动行为的物理理解或实验观测结果构建, 具有一定的理论依据和较好的解释性.

随着人工智能技术的发展, 机器学习(Machine learning, ML)在传热预测领域中的应用日益广泛. Hughes 等[21]回顾了过去用于传热问题的机器学习研究, 指出机器学习不仅能作为提升两相流传热预测能力的实用工具, 还有助于揭示其中潜在的物理机制. Qiu 等人[22]基于来自 50 项研究的 16853 个数据点, 构建了用于微通道沸腾 HTCs 预测的人工神经网络模型, 测试集上的平均绝对相对误差(*MARD*)为 14.3%, 显示出良好的预测能力. 随后, Qiu 等人[23]进一步结合物理引导(传统的关联式分析)与数据驱动(皮尔逊相关系数、互信息特征选择)方法优化模型参数选择, 将 *MARD* 降低至 8.48%. Bard 等人[24]基于相同的数据库, 比较了线性建模、随机森林、支持向量机和神经网络等多种模型, 发现支持向量机回归模型表现最优, *MARD* 为 11.3%, 但在一些极端工况下预测表现不佳. Zhu 等人[25,26]基于 R134a 在锯齿翅片微通道内的实验数据分别建立了流动沸腾与冷凝的人工神经网络模型, *MARD* 分别为 11.41%与 6.06%, 并提出了一种迁移学习框架, 用于将传统的 ML 模型扩展到不同的数据域中, 误差低于 7.2%. 但这项研究的数据仅是针对 R134a 一种流体进行实验获得的. Li 等人[27]则针对强化管冷凝构建了含 1937 个数据点的数据库, 系统评估了不同 ML 模型及输入特征的影响. 最近, Zhang 等人[28]基于包含 5332 个数据点的数据库, 提出一种用于预测冷凝 HTCs 的集成模型：先通过 K-means 聚类将数据划分为五类, 再结合随机森林分类器预测样本类别概率, 并针对每类训练一个人工神经网络回归器, 最终通过加权融合分类概率与回归结果, 将 *MARD* 控制在 8%. 除以上研究外, 其他研究[29-31]也进一步验证了 ML 在两相流传热预测中的潜力, 尤其在处理强非线性关系时, 其性能常优于传统经验关联式.

目前, 多数研究已证实 ML 在预测流动沸腾

HTCs 方面具有潜力, 但现有工作多集中于单一 ML 模型的应用. 考虑到通道内流动状态对传热的影响, 本研究构建了一个微小通道内流动沸腾传热的综合数据库, 并提出一种融合谱聚类与 XGBoost 算法的集成 ML 模型. 该模型首先利用谱聚类将数据划分为不同类别, 以表征潜在的流动状态; 随后分别构建 XGBoost 分类模型与回归模型, 用于预测数据点属于各流动状态的概率及其传热系数值, 最终通过加权求和得到 HTCs 的集成预测结果. 鉴于收集的实验数据质量不一, 在回归训练中引入实验相对不确定度作为样本权重, 使模型更聚焦于高质量数据. 在输入特征选择上, 采用最大信息系数法从全面特征集中筛选出代表性特征子集. 测试集验证表明, 所提集成模型预测精度显著优于传统关联式, 并在保持较高预测性能的同时, 较单一 ML 模型具备更好的可解释性.

2 方法

2.1 单一机器学习模型

谱聚类(Spectral clustering, SC)属于无监督学习算法, 适用于对未标记数据进行聚类分组; 而极端梯度提升(XGBoost)则为有监督学习算法, 可用于处理分类与回归预测任务. 两种方法的原理简要介绍如下:

谱聚类是一种基于图论的聚类方法, 旨在将数据点划分为若干类别, 使得类内相似度最大化而类间相似度最小化. 该模型首先将数据点看作图节点, 并根据点之间的相似性构建连接边, 从而形成一个相似性图. 然后通过图切割的方式, 找到一种分组方法, 使得连接不同组的边的权重尽可能低, 而组内的边的权重尽可能高. 它特别擅长处理那些形状复杂的数据(非凸形状), 而传统的 K-means 算法在这种数据上易失败[32,33].

极端梯度提升是一种基于梯度提升框架改进的高效机器学习算法, 其核心思想是通过迭代地构建多个弱决策树, 并逐步优化一个结合了损失函数和正则化项的目标函数来提升预测精度. 与随机森林模型的装袋策略不同, XGBoost 采用 boosting 方法, 每棵树都致力于纠正前一棵树的错误, 从而形成一个强大且精准的集成模型[34]. 它非常适用于各种分类和回归任务.

2.2 集成模型

本研究基于 Python 3.10 及 NumPy、Pandas、Scikit-learn 和 XGBoost 等库, 构建了一种融合谱聚类与 XGBoost 的集成预测模型, 以考虑流动状态对传热的影响. 该方法首先利用谱聚类将数据划分为代表不同流动状态的类别, 再基于聚类结果分别建立 XGBoost 分类与回归模型. 具体流程包含以下五个步骤:

1. 数据库的构建和特征选择. 详细过程见 2.3 和 2.4 节.

2. 基于谱聚类的数据划分. 数据库包含广泛工质、水力直径范围 0.19-6.5mm 和蒸气干度范围 0.001-0.999, 因此根据其他研究者[4-12]总结出的水平通道中流动沸腾的典型流动状态, 假设存在 5 类流动状态. 选用表观液相速度(U_{ls})和表观气相速度(U_{vs})作为聚类特征将数据库划分为 5 类(聚类 0 - 4), 并分配类别标签. 聚类结果及各类别 HTCs 分布如图 2 所示.

3. 基于 XGBoost 的分类. 基于聚类结果, 将类别标签作为目标变量, 训练 XGBoost 多分类模型, 用于预测数据点属于各类别的概率值(P_0 - P_4).

4. 基于 XGBoost 的回归. 基于聚类结果, 针对每个聚类分别训练一个 XGBoost 回归模型, 得到五个 HTCs 回归预测值(h_0 - h_4).

5. 加权集成预测. 将分类模型输出的概率值作为权重, 对五个回归模型的预测结果进行加权求和, 得到最终的沸腾传热系数预测值: $HTC = \sum_{i=0}^{4} P_i \cdot h_i$.

2.3 数据库与数据划分

本研究从 42 篇公开文献中收集了 8309 个微小通道流动沸腾传热的实验数据点, 构建了一个综合数据库. 在整合过程中已对缺失及可疑数据予以剔除. 该数据库涵盖 24 种流体, 覆盖了较宽的工况范围, 并包含少量微翅管数据. 数据库涵盖多种通道结构(圆形管、矩形管、光滑管和强化管), 由于表面粗糙度参数仅少部分文献提供, 因此表面改性等参数未被考虑, 仅区分是否为强化管. 数据库的基本概况总结于表 1, 而各子数据库的关键信息则列于表 2.

图 3 展示了数据库中部分运行参数的分布情况. 从图 3(a)可以看出, 数据库包含了纯质、混合物及天然工质, 其中 R134a 占比最高, 约为总量的三分之一, 其余主要工质为 CO_2、R1234yf 和 R1234ze(E)等工质. He 等人[35]的研究表明, ML 模型在预测新数据时, 对于训练过程中学习的工质种类的数据具有更好表现, 而未学习过的工质的数据则预测表现明显变差. 因此本研究在收集数据时尽可能地收集了多样的工质种类, 并且在特征选择中纳入了工质的多种热物性参数(如密度、粘度和表面张力等), 以期提高模型的外推能力. 图 3(b)和(c)分别展示了水力直径与质量流量的分布. 虽然目前对微小通道尚无统一定义, 但多数数据点水力直径小于 3 mm, 大部分数据满足小通道条件. 质量流量的分布主要集中在小于 $600\ kg/\left(m^2s\right)$, 大于 $1000\ kg/\left(m^2s\right)$ 范围也存在少量数据点, 主要是 CO_2 作为工质的实验数据. 传热系数分布较为均匀, 如图 3(d)所示, 其最大值不超过 $25\ kW/\left(m^2K\right)$.

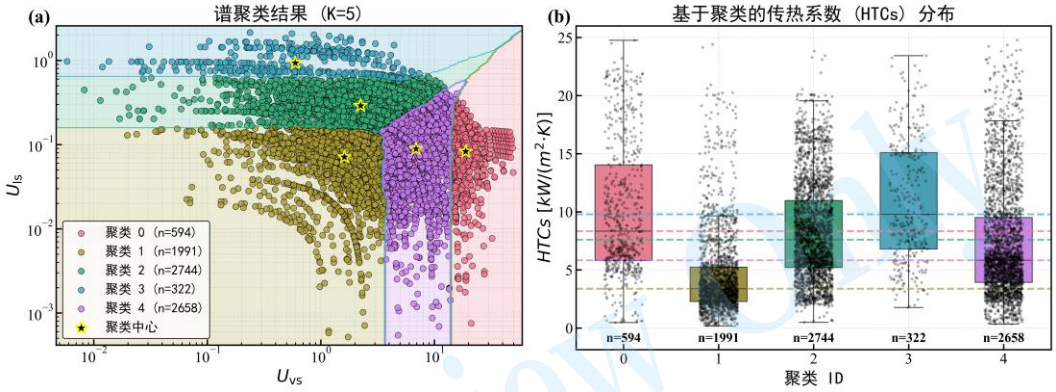


图 2 (网络版彩图) 基于谱聚类的数据划分. (a) 谱聚类结果; (b) 各类别传热系数分布

Figure 2 (Color online) Data partitioning based on SC. (a) Clustering distribution; (b) Distribution of HTCs of clusters.

表1 数据库的部分参数工况范围

Table 1 the working condition range of parameters of the database

工况	范围
工作流体	R134a, CO_2, R1234yf, R1234ze (E), R448A, R452A, R290, R513A, Water, R22, R454C, R245fa, R32, R123, R404A, R449A, R455A, R454B, R32/R1234yf (质量比: 0.15/0.85, 0.5/0.5, 0.85/0.15), R290/R601a (质量比: 0.3/0.7, 0.5/0.5, 0.75/0.25)
水力直径 (D)	$0.19 \le D \le 6.5mm$
热通量 (q)	$2 \le q \le 1400 kW/m^2$
质量流量 (G)	$25 \le G \le 1800 kg/\left(m^2s\right)$
蒸气干度 (x)	$0.001 \le x \le 0.999$
传热系数 (h)	$0.181 \le 24.786 kW/\left(m^2K\right)$

为保障模型评估的严谨性, 在聚类后采用分层抽样方法, 按 8:2 的比例从五个类别中分别划分训练集与测试集, 确保训练与测试数据相互独立, 避免信息泄露. 训练集与测试集的数据量及分布情况如图 4 所示. 训练集用于建立特征与目标之间的映射关系, 并通过交叉验证优化模型超参数; 测试集则用于评估模型对未知数据的预测性能, 即泛化能力.

2.4 特征与超参数选择

特征在机器学习模型中具有关键作用, 特征数量不足易导致欠拟合, 而过多的特征则可能引起过拟合. 本研究尽可能系统地纳入了多种相关特征, 如表 3 所示. 主要包括三类: a)从文献中直接提取的运行参数; b)基于原始参数通过流体物性查询工具 REFPROP 9.0[77] 获取的热物性参数; c)根据上述参数进一步推导出的流动沸腾相关无量纲数. 虽然管道形状结构和表面改性等参数会对流动换热产生影响, 但考虑到数据库的实际情况, 只考虑了水力直径和是否为强化管“*Tube_shape*”(强化管如微翅管手动标记为 1，非强化管为 0)两种相关特征. 其中, 蒸气干度(x)为直接从文献中提取的参数, 为测试段入口和出口处的平均蒸气干度值. 此外, 为便于聚类分析, 还引入了表观液相速度(U_{ls})和表观气相速度(U_{vs}). 共计 47 个特征被纳入初步特征集.

为确定合适的输入特征组合, 首先采用最大信

息系数(Maximum information coefficient, MIC)法[78]评估各特征与HTCs之间的关联强度，并按MIC值降序排序，结果如图5所示. 在此基础上，以前10个特征为起点，采用默认参数的XGBoost模型进行训练，并基于测试集上的决定系数(R^2)评估模型性能，逐步增加特征数量以观察其影响，结果如图6所示. 分析表明，当使用前46个特征时模型性能最优，但为平衡模型复杂度与精度，最终选定前34个特征作为主要输入. 尽管“*tube_shape*”(是否强化管)这一特征的MIC值较低，可能因数据库中强化管样本较少所

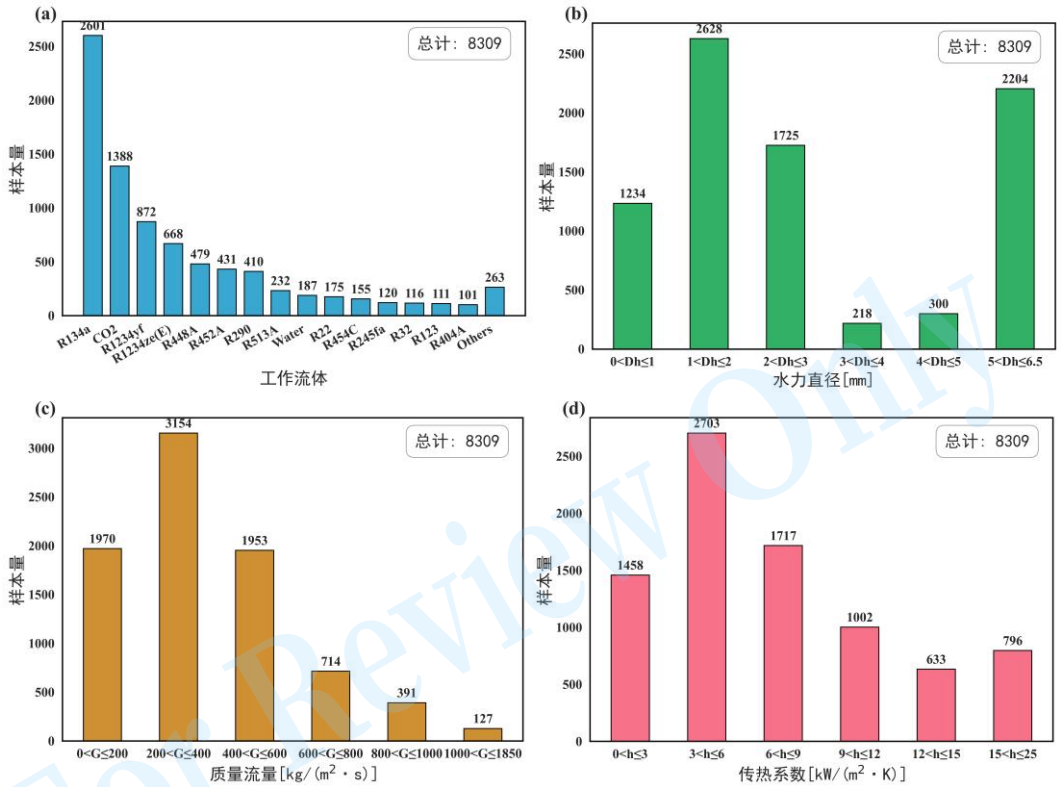


图 3 (网络版彩图) 数据库部分参数分布情况. (a) 工作流体; (b) 水力直径; (c) 质量流量; (d) 传热系数

Figure 3 (Color online) Distribution of some parameters in the database. (a) Working fluid; (b) Hydraulic diameter; (c) Mass flux; (d) Heat transfer coefficient.

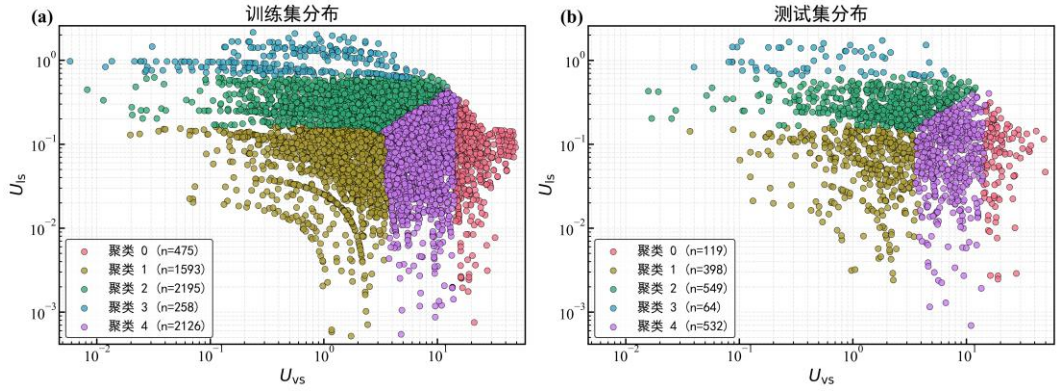


图 4 (网络版彩图) 训练集和测试集的分布. (a) 训练集; (b) 测试集

Figure 4 (Color online) Distribution of the training set and test set. (a) Training set; (b) Test set.

表2 数据库中的流动沸腾传热数据

Table 2 Flow boiling heat transfer datas in the database

作者	几何形状	工作流体	水力直径(mm)	热通量 (kW/m^2)	质量流量 (kg/m^2s)	不确定度	数据点
Yan 和 Lin (1998)[36]	C, Multi, H, ST	R134a	2	5-20	50 - 200	6%	143
Wang 等(1998)[37]	C, Single, H, ST	R22	6.5	2.5-20	100 - 400	12.4%	63
Yun 等 (2003)[38]	C, Single, H, ST	CO2	6	6-20	170-340	8%	126
Huo 等 (2004)[39]	C, Single, VU, ST	R134a	2.01, 4.26	13-130	100-500	6%	389
Zhao 和 Bansal (2007)[40]	C, Single, H, ST	CO2	4.57	12.6-19.3	139.5-230.9	10%	22
Bertsch 等 (2009)[41]	R, Multi, H, ST	R134a, R245fa	0.54, 1.09	26-196	42-334	7.9%	173
In 和 Jeong (2009)[42]	C, Single, H, ST	R123, R134a	0.19	10-20	34-470	13.8%	213
Mastrullo 等 (2009)[43]	C, Single, H, ST	R134a, CO2,	6	9.9-20.2	202-349	3.5%	161
Wang 等 (2009)[44]	C, Single, H, ST	R134a,	1.3	21.1-50	32-836	12%	365
Tibiriçá 等(2010)[45]	C, Single, H, ST	R134a, R245fa	2.3	5-35	200-700	30%	149
Bang 等 (2011)[46]	C, Single, H, ST	Water	1.73	50-160	100	14.8%	58
Copetti 等 (2011)[47]	C, Single, H, ST	R134a	2.62	5-100	240-930	9.5%, 26.2%, 34.2%	305
Kyu 和 Son (2011)[48]	C, Single, H, ST	R22, R134a	1.77, 3.36, 5.35	10-30	30-500	13.5%	151
Wu 等 (2011)[49]	C, Single, H, ST	CO2	1.42	7.5-29.8	300-600	10%	420
Balasubramanian 等 (2013)[50]	R, Multi, H, ST	Water	0.504	1400	88-156	19%	129
Col 等 (2013)[51]	C, Single, H, ST	R1234yf	0.96	30-50	200-600	10%	58
Grauso 等 (2013)[52]	C, Single, H, ST	R1234ze(E), R134a	6	5-20.2	146-502	18%	582
Wang 等 (2014)[53]	C, Single, H, ST	R290	6	11.7-53.2	70.8-74	17%	55
Xu 等 (2015)[54]	C, Single, H, ST	R134a	1.002, 2.168, 4.065	19-28.5	185-910	10%	169
Wen 等 (2018)[55]	R, Multi, H, ST	R134a	0.63, 0.72	17.3-41.7	146-290	7.1%	189
Diani 和 Rossetto (2018)[56]	R, Multi, H, EHT	R134a, R1234ze(E), R1234yf	2.4	10-50	375-940	6.7%	201
Li 等 (2018)[57]	R, Multi, H, ST	R32, R134a,	0.643	2-8	100-250	20%	62
Lillo 等 (2018)[58]	C, Single, H, ST	R290	6	2.48-40.3	150-499	9%	310
Zhang 等 (2018)[59]	C, -, H, ST	CO2	1.5	7.5-30	300-600	14.5%	386
Mastrullo 等 (2019)[60]	C, Single, H, ST	R452A,	6	9.96-40.1	149-600	7.3%	227
Wen 等 (2019)[61]	R, Multi, H, EHT	R134a	1.28, 1.59	15-47.7	70-190	6.8%	302
Mastrullo 等 (2019)[62]	R, Single, H, ST	R452A	6	9.96-40.1	149-600	14.1%	228
Lillo 等 (2019)[63]	C, Single, H, ST	R448A, R404A	6	10-40	150-600	17%	273
Li 和 Hrnjak (2019)[64]	R, Multi, H, ST	R1234ze(E)	0.643	2-6	100-200	20%	48
Kim 等 (2021)[65]	R, Multi, H, ST	R404A, R448A, R449A, R454C, R455A	1.61	2.8-6.5	200-400	20.4%	168
Xu 等 (2021)[66]	C, Single, H, ST	R1234ze(E)	1.88	55-65	540-870	4.01	153
Hellenschmidt 等 (2021)[67]	C, Single, H, ST	CO2	0.5, 1, 2.15	5-35	500-1800	10%	292
Hoang 等 (2022)[68]	R, Multi, H, ST	R448A, R404A	0.969	3-15	100-500	14.1%	117
Arcasi 等 (2022)[69]	C, Single, H, ST	R454C	6	99-40.1	151.2-597	16.48%	121
Xu 等 (2022)[70]	C, Single, H, ST	R1234yf	1.88	40-65	400-870	3.17%	337
Liu 等 (2023)[71]	C, Single, H, ST	R1234yf	1, 2	2-15	50-200	25.26%	418
Agustiarini 等 (2023)[30]	R, Multi, H, ST	R32	0.969	3-6	50-500	10%	54
Bediako 等 (2024)[72]	C, Single, H, ST	R1234ze(E)	2.07	30-50	300-500	15.6%	108
Diani 和 Rossetto (2024)[73]	C, Single, H, ST/EHT	R513A	2.5, 2.4	12-60	200-800	9.8%, 13.1%	232
Wu 等 (2024)[74]	R, Multi, H, ST	R454B, R32, R1234yf, R32/R1234yf (质量比: 0.15/0.85, 0.5/0.5, 0.85/0.15)	0.643	2-8	100-250	25.7%	85
Ren 等 (2025)[75]	C, Single, H, ST	R290, R290/R601a (质量比: 0.3/0.7, 0.5/0.5, 0.75/0.25)	2	30	100-500	7.49%	152
Hoang 等 (2025)[76]	R, Multi, H, ST	R448A	0.969, 0.831, 1.077, 1.141, 1.164	3-12	100-300	15.6%	115
总计							8309

C: 圆形; R: 矩形; H: 水平; VU: 垂直向上; ST: 光滑管; EHT: 强化管.

For Review Only

致，但基于已知的传热增强作用，仍将其纳入特征集. 此外，聚类所用的两个特征“U_{ls}”和“U_{vs}”也一并加入. 因此，本研究最终选用 37 个参数作为模型输入，具体包括: $Su_l, Su_v, Bd, Ga, P_r, Ca_{lo}, Ca_{vo}, \rho_l, \rho_v, G$, $Cp_l, Cp_v, \mathrm{Pr}_l, \mathrm{Pr}_v, \lambda_l, \lambda_v, \sigma, \mu_l, \mu_v, h_{lat}, F_{lo}, F_{vo}, F_l, P_{sat}, q, Re_{lo}$, $Re_{vo}, We_{lo}, We_{vo}, Fa, T_{sat}, D_h, Bo, P_{crit}, U_{ls}, U_{vs}, tube_shape$. XGBoost 模型本身通过正则化参数惩罚项可对特征系数进行稀疏约束，进一步降低过拟合风险，提升模型稳健性.

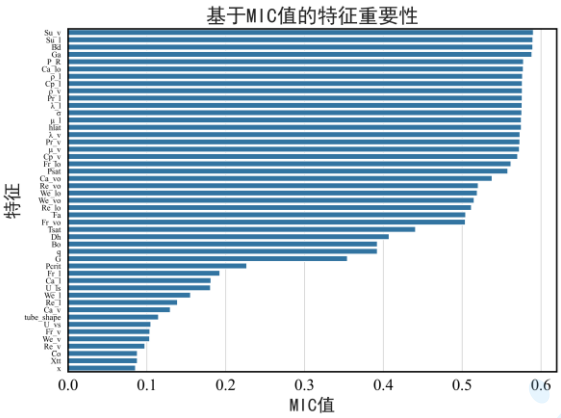


图 5 (网络版彩图)基于最大信息系数法的特征重要性排序
Figure 5 (Color online)Feature importance ranking based on the MIC

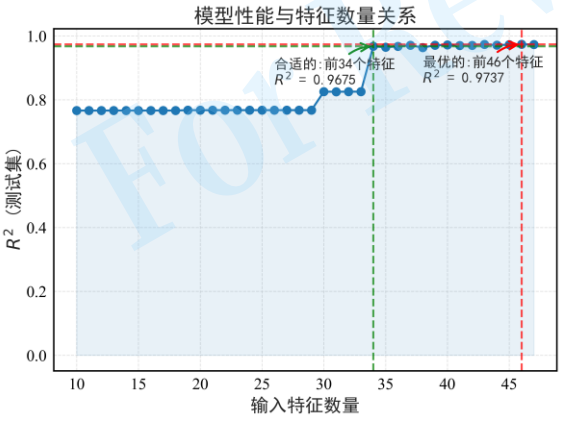


图 6 (网络版彩图) 使用默认参数的 XGBoost 模型在添加重要特征后的性能
Figure 6 (Color online) Performance of the XGBoost model with default hyperparameters after adding important features.

针对 XGBoost 模型在分类与回归任务中的超参数优化，选择采用贝叶斯优化(Bayesian optimization, BO)法[79]并结合五折交叉验证，以提高调优效率与模型性能. 在回归模型中，训练阶段使用均方误差(*MSE*)作为损失函数，因其处处可微，便于梯度下降优化；而在贝叶斯优化过程中，则选用平均绝对误差(*MAE*)作为评价指标，以降低异常值对参数选择的影响，增强鲁棒性. 分类模型与所有回归模型采用相同的超参数搜索空间，具体范围见表 4，优化结果汇总于表 5.

表3 原始输入特征选择
Table 3 Original input feature selection

原始参数特征	流体热物性特征	无量纲特征
tube_shape, q, T_{sat}, D, G, x, U_{ls}, U_{vs}.	$P_{sat}, P_{crit}, Cp_l, Cp_v$, $\rho_l, \rho_v, \lambda_l, \lambda_v$, $\mu_l, \mu_v, \sigma, h_{lat}$.	$Re_l, Re_{lo}, Re_v, Re_{vo}, Fr_l, Fr_{lo}$, $Fr_v, Fr_{vo}, We_l, We_{lo}, We_v, We_{vo}$, $Su_l, Su_v, Ca_l, Ca_{lo}, Ca_v, Ca_{vo}$, $Pr_l, Pr_v, Bd, Bo, Fa, P_R$, X_{tt}, Co, Ga.

表 4 XGBoost 模型的超参数与搜索空间
Table 4 Hyperparameters and search space of the XGBoost model

模型	超参数	搜索范围
XGBoost	Number of trees	100-400
	Learning rate	0.05-0.2
	Maximum depth	4-10
	Minimum child weight	1-10
	Subsample ratio	0.5-1.0
	Feature subsample ratio	0.5-1.0
	Gamma	0-2
	L1 regularization	0.01-1.0
	L2 regularization	0.1-2.0

3 结果与讨论

3.1 传统经验关联式预测

为评估 ML 模型的预测性能，基于数据库利用几组传统的关联式对传热系数进行预测. 表 6 提供了这些关联式的公式和条件，均是针对小通道流动沸腾传热的通用关联式. 为了评估模型的预测能力，采用平均绝对相对误差(*MARD*)、平均相对误差(*MRD*)和决定系数(R^2)指标评估预测模型的性能. 前者主要反映预测准确度和误差分布，后者反映模型拟合质量:

$$MARD = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|h_{i,pred} - h_{i,exp}\right|}{h_{i,exp}} \times 100\%, \quad (1)$$

For Review Only

$$MRD = \frac{1}{N}\sum_{i=1}^{N}\frac{h_{i,pred} - h_{i,exp}}{h_{i,exp}} \times 100\%, \quad (2) \qquad R^2 = 1 - \frac{\sum_{i=1}^{N}(h_{i,pred} - h_{i,exp})^2}{\sum_{i=1}^{N}(\bar{h}_{i,exp} - h_{i,exp})^2}. \quad (3)$$

表 5 XGBoost 模型超参数的优化结果

Table 5 Optimization results of hyperparameters of the XGBoost model

超参数/模型	分类器	回归器 I	回归器 II	回归器 III	回归器 IV	回归器 V	单一模型	单一加权模型
Number of trees	318	400	400	400	400	400	400	400
Learning rate	0.1124	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Maximum depth	7	10	7	10	4	9	10	10
Minimum child weight	1	1	10	10	3	1	1	1
Subsample ratio	0.647	0.5	0.82	0.5	0.95	0.68	0.613	0.58
Feature subsample ratio	0.615	0.5	0.692	0.882	0.909	0.932	1	1
Gamma	0.324	0	0	0	0.527	0	0	0
L1 regularization	0.038	0.01	0.58	0.01	0.126	0.01	0.01	0.01
L2 regularization	0.427	2	0.911	2	1.216	0.1	0.1	2

表6 传统关联式和建立条件

Table 6 Traditional correlation expressions and establishment conditions

作者	关联式	建模范围
Bertsch 等人(2009)[17]	$h_{tp} = h_{nb}\cdot S + h_{cb}\cdot F$, $S = 1 - x, F = 1 + 80\left(x^2 - x^6\right)e^{-0.6C}, C = \sqrt{\frac{\sigma}{g(\rho_l - \rho_v)D^2}}$, $h_{nb} = 55P_R^{0.12-0.2\log_{10}R_P}\left(-\log_{10}P_R\right)^{-0.55}M^{-0.5}q^{0.67}$, $h_{cb} = h_{cb,lo}(1-x) + h_{cb,vo}x$, $h_{cb,lo} = \left[3.66 + \frac{0.0668\frac{D}{L}Re_{lo}Pr_l}{1+0.04\left(\frac{D}{L}Re_{lo}Pr_l\right)^{2/3}}\right]\frac{\lambda_l}{D}$, $h_{cb,vo} = \left[3.66 + \frac{0.0668\frac{D}{L}Re_{vo}Pr_v}{1+0.04\left(\frac{D}{L}Re_{vo}Pr_v\right)^{2/3}}\right]\frac{\lambda_v}{D}$.	12 种流体: 水, 氮, 甲醇, FC-77, R11, R123, R134a, R245fa 等, . $D = 0.16 - 2.92mm$, $G = 20 - 3000kg/m^2s$, 3899 个数据点.
Li 和 Wu (2009)[18]	$h_{tp} = 334Bo^{0.3}\left(BdRe_l^{0.36}\right)^{0.4}\frac{\lambda_l}{D}$.	13 种流体: 水, 乙醇, FC-77, CO_2, R12, R134a 等, $D = 0.16 - 3.1mm$, $G = 23.4 - 3500kg/m^2s$, 3744 个数据点.
Sun 和 Mishima (2009)[19]	$h_{tp} = \frac{6Re_{lo}^{1.05}Bo^{0.54}}{We_l^{0.191}\left(\frac{\rho_l}{\rho_v}\right)^{0.142}}\frac{\lambda_l}{D}$.	11 种流体: 水, CO_2, R11, R134a, R410A 等, $G = 44 - 1500kg/m^2s$, 2505 个数据点.
Kim 和 Mudawar (2013)[20]	$h_{tp} = \left(h_{nb}^2 + h_{cb}^2\right)^{0.5}$, $h_{nb} = \left[2345\left(Bo\frac{P_H}{P_F}\right)^{0.70}P_R^{0.38}(1-x)^{-0.51}\right]\left(0.023Re_l^{0.4}\frac{\lambda_l}{D}\right)$, $h_{cb} = \left[5.2\left(Bo\frac{P_H}{P_F}\right)^{0.08}We_{lo}^{-0.54} + 3.5\left(\frac{1}{X_{tt}}\right)^{0.94}\left(\frac{\rho_v}{\rho_l}\right)^{0.25}\right]\left(0.023Re_l^{0.8}Pr_l^{0.4}\frac{\lambda_l}{D}\right)$.	18 种流体: 水, CO_2, R113, R134a, R22, R245fa, R404A, R410A, R1234yf, R32 等, $D = 0.19 - 6.5mm$, $G = 19 - 1608kg/m^2s$, 10805 个数据点.

For Review Only

由于数据库中大部分样本缺乏 R_P(通道粗糙度)和 L(通道长度)信息，未对 Bertsch 等人[17]提出的模型进行评估，其他关联式中涉及缺失关键参数的少量数据亦予以排除. 图 7 展示了三种关联式在数据库上的预测结果，横坐标为实验值，纵坐标为预测值. Li 和 Wu[18]模型的预测 *MARD* 为 70.49%，仅 62.46%的样本误差低于 50%，相对而言表现较差. 其原因可能在于该关联式在直径范围为 0.16-3.1mm 内建立，而数据库中有相当一部分数据的直径大于 3mm. Sun 和 Mishima[19]关联式所使用的原始数据库工况与本数据库较为接近，*MARD* 为 50.12%，误差低于 50%的样本占比为 75.25%，表现有所提升. Kim 和 Mudawar[20]模型表现最佳，*MARD* 为 46.51%，误差小于 50%的样本占 78%，这可能得益于其建模时基于大规模数据库，且适用工况范围与本数据库重合度高. 值得注意的是，该模型原本基于预干涸数据建立，而本研究未区分，因为数据库大部分数据处于预干涸区域. 预干涸区域表示干涸起始处上游直至蒸气干度 $x=0$ 的区域，具体可见图 1 和 Kim 和 Mudawar[8]的研究. 总之，传统关联式在数据库上的预测效果存在差异. 可以得出结论，当目标数据库的工况越接近关联式的原始建立范围，且关联式本身所依托的数据库规模越大时，其预测性能通常越好.

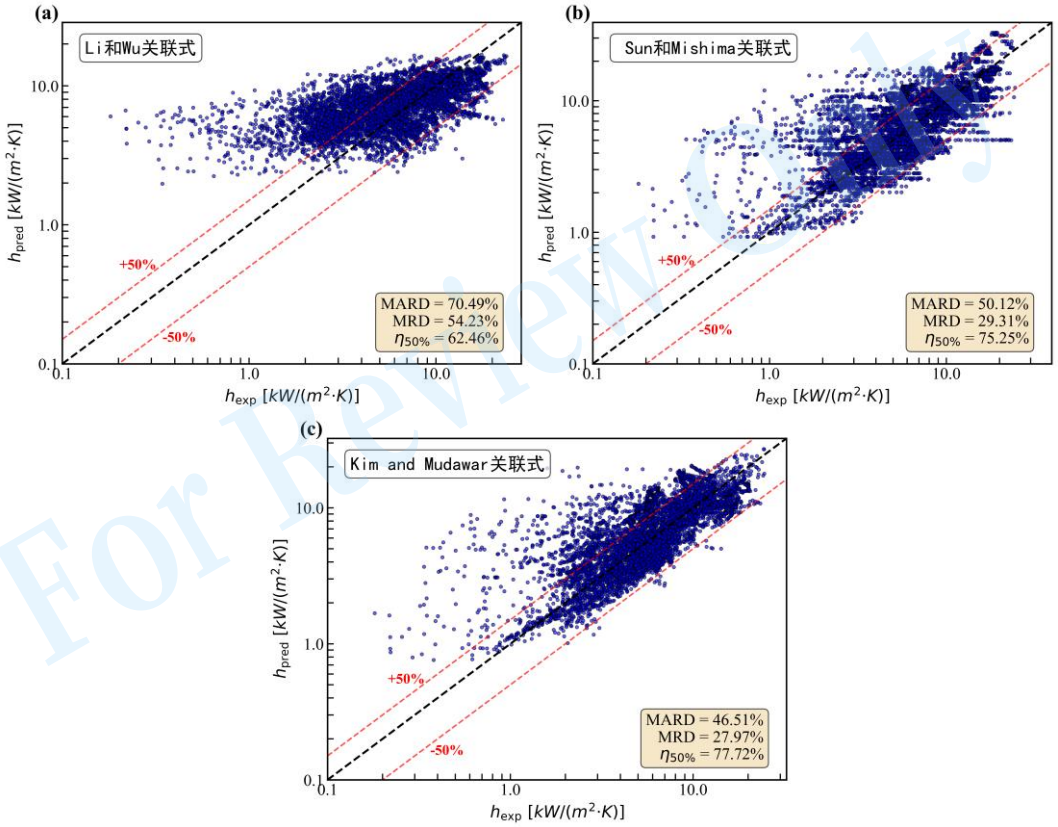


图7 (网络版彩图) 传统关联式预测沸腾传热系数. (a) Li和Wu [18]; (b) Sun和Mishima[19]; (c) Kim和Mudawar[20]

Figure 7 (Color online) Prediction of boiling HTCs by traditional correlations. (a) Li and Wu[18]; (b) Sun and Mishima[19]; (c) Kim and Mudawar[20].

3.2 单一机器学习模型预测

考虑到数据库来源于多项独立研究，各研究在实验仪器与主观操作上存在差异，在收集数据时整理了各数据点对应的实验相对不确定度. 并将不确定度作为样本权重引入模型的训练过程，具体通过在均方误差(*MSE*)损失函数中为每个样本赋予不确定度倒数的平方的权重，即不确定度小的样本权重更大. 其中，样本的相对不确定度来源于各文献报告值，且基本计算方式一致. 权重策略的目的是让模型在训练过程中更侧重于高置信度的数据，并非期望通过权重来消除数据的实验误差. 相对于 3.1 节提出的

评估标准，进一步提出加权指标. 加权指标考虑了样本权重的差异，能够更合理地反映模型在高质量数据上的拟合优度，而原始指标则对所有样本平等对待. 权重、加权损失函数及加权指标的定义如下:

$$weight = \frac{1}{uncertainty^2}, \tag{4}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(h_{i,pred} - h_{i,\exp}\right)^2, \tag{5}$$

$$WMSE = \frac{1}{N}\sum_{i=1}^{N} weight_i \left(h_{i,pred} - h_{i,\exp}\right)^2, \tag{6}$$

$$WMARD = \frac{\sum_{i=1}^{N}\left(weight_i \frac{\left|h_{i,pred} - h_{i,exp}\right|}{h_{i,exp}}\right)}{\sum_{i=1}^{N} weight_i} \times 100\%, \tag{7}$$

$$WMARD = \frac{\sum_{i=1}^{N}\left(weight_i \frac{h_{i,pred} - h_{i,exp}}{h_{i,exp}}\right)}{\sum_{i=1}^{N} weight_i} \times 100\%. \tag{8}$$

图 8 展示了单一 XGBoost 模型在测试集上的预测性能，其表现直观上显著优于传统经验关联式. 单一未加权 XGBoost 模型(图 8(a))和考虑数据不确定度的单一加权 XGBoost 模型(图 8(b))均表现出优秀的预测性能，*MARD* 分别为 7.60%和 7.18%，93.92%和 93.74%的样本预测误差小于 20%，整体精度相当. 但对于加权模型，如图 8(b)中加权指标表现明显优于原始指标，*WMARD*(4.92%)优于 *MARD*(7.18%)，表明经过加权训练后的模型在预测高置信度，即不确定度小的样本时表现更好.

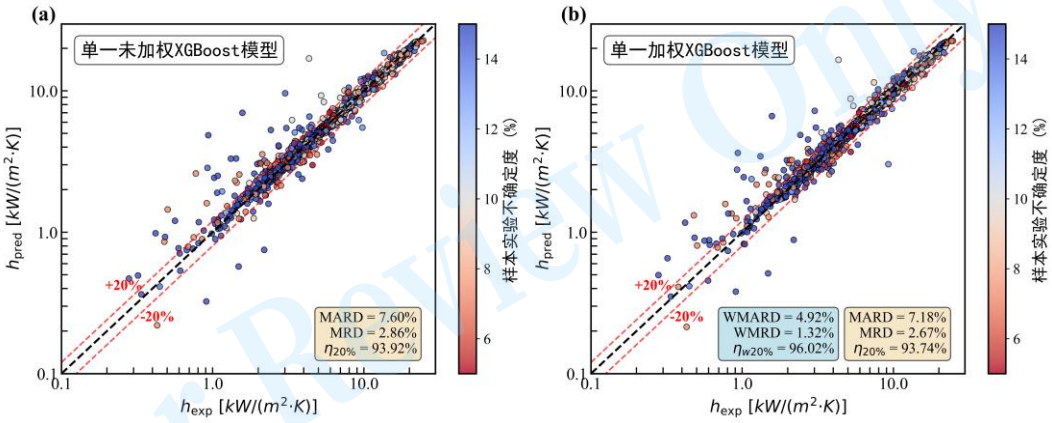


图 8 (网络版彩图) 单一 XGBoost 模型在测试集上预测沸腾传热系数. (a) 未加权模型; (b) 加权模型

Figure 8 (Color online) Prediction of boiling HTCs by a single XGBoost model on the test set. (a) Unweighted model; (b) Weighted model.

3.3 集成模型预测

在传热系数预测关联式的开发中，部分研究者曾尝试针对不同流动状态分别建立模型[4-6]，这是由于各流型具有独特的热物理特性，会显著影响传热机制的建模准确性[7]. 受此启发，在机器学习预测中，一种值得探索的思路是：先通过聚类方法将数据划分为若干类别以近似代表不同流动状态，再对每一类别分别建立预测模型.

参照 Thome[7]与 Kim 和 Mudawar[8]等对水平通道中流动沸腾的典型流动状态的分类以及结合本数据库宽泛的工况范围，本研究将数据库中的样本对应的主要流动状态归纳为泡状流、段塞流、分层流、环状流与雾状流五种基本类型. 基于该假设，采用谱聚类方法将整个数据库划分为五个簇. 聚类特征选用表观气液相速度，该类参数常用于流型图的构建. 聚类结果见图 2.

除了为每个聚类单独训练回归模型外，还构建了一个 XGBoost 分类模型，用于预测样本属于各流动状态的概率. 如表 7 所示，该模型在包含 1662 个样本的测试集上分类准确率达 99%，仅有少量样本存在误分类. 以类别 3 为例，其精确率为 100%，表明所有被预测为该类的样本均正确；召回率为 98%，说明模型识别出了 98%的真实类别 3 样本，混淆矩阵显示仅有 1 个样本被误判为类别 2. F1 分数是精确率和召回率的调和平均数，是一个综合指标. 从混淆矩阵中可知，错误分类主要出现在类别 1、2、4 之间以及类别 4 与 0 之间，反映出这些类别在特征空间中存在一

For Review Only

定的交界或相似性, 从图 2 中可以得到证明. 总之, 各类别的精确率、召回率和 F1 分数均超过 98%, 平均指标均达到 99%, 表明模型在所有类别上均实现了令人满意的预测.

表7 XGBoost模型分类结果评估

Table 7 Evaluation of classification results of XGBoost model

指标/类别	0	1	2	3	4	平均值
精确率	0.98	0.99	0.99	1.00	0.99	0.99
召回率	1.00	0.99	0.99	0.98	0.99	0.99
F1-分数	0.99	0.99	0.99	0.99	0.99	0.99
准确率	-	-	-	-	-	0.991
support	119	398	549	64	532	1662
混淆矩阵	119.	0	0	0	3	
	0	396	4	0	1	
	0	0	544	1	3	
	0	0	0	63	0	
	0	2	1	0	525	

在应用集成模型之前, 针对每个聚类分别构建了 XGBoost 回归模型. 图 9(a)–(e)展示了五个子回归模型在各自测试集上的预测效果, 对应聚类 0 到 4. 总体来看, 各子模型均表现出优秀的预测性能. 其中, 回归器 III 表现最佳, *MARD* 为 4.37%, R^2达 0.981, 98.5% 的样本预测误差低于 20%, 显示出对数据规律的良好捕捉能力(图 9(c)). 回归器 IV 为次优模型, 其系统性偏差最小, *MRD* 为 0.80%, 绝大多数样本误差控制在 20% 以内, 在图 9(d) 中能够直观地感受到. 回归器 I、II 和 V 的 *MARD* 均在 9%左右, 虽略低于前两者, 但仍属优秀. 根据图 2 的聚类结果, 表现最好的回归器 III 和 IV 对应聚类-2 和 3, 为液相表观速度较大的区域, 流动相对稳定. 随着气相表观速度增大, 蒸气干度提高, 流动热行为更复杂, 使模型预测更困难. 此外, 各数据簇之间的质量和分布差异也可能是造成各个回归器预测精度差异的原因. 集成模型的预测结果通过加权融合分类概率与各回归输出得到(图 9(f)). 其综合性能良好, *MARD* 为 7.41%, $\eta_{20\%}$ 为 93.6%, R^2为 0.979. 尽管未超越回归器 III 和 IV, 集成模型仍显示出显著优于传统关联式的预测能力, 结果令人满意.

表 8 汇总了所有回归模型在各自测试集上的预测结果. 尽管集成模型的预测精度略低于部分子回归模型, 但其测试范围覆盖了所有子模型数据, 因而在实际应用中具有更好的泛化能力. 表中也列出了单一 XGBoost 模型的结果. 整体上, 集成模型性能未能优于单一加权模型, 而与未加权模型相当. 三者的 *MRD*、$\eta_{20\%}$ 和 R^2 指标差异甚微, 主要区别体现在 *MARD* 上: 单一加权模型的 *MARD* 为 7.18%, 略优于未加权模型(7.60%), 说明在训练中引入不确定度权重可能有助于提升对高质量数据的学习效果, 但其普遍适用性仍需进一步验证. 此外, 集成模型相比单一加权模型精度没有明显提升, 可能的原因包括: a)XGBoost 即为一种强大的集成学习算法, 在大型数据库上已能实现全局优化. 两种模型的 *MARD* 分别为 7.18%和 7.41%, 预测精度已非常理想, 均难有更好突破, 使所提出的集成模型优势未得到凸显; b)当前数据库的覆盖范围有限, 部分数据簇的数据量不足，数据分布表现差, 影响了子回归模型的学习. 尽管如此, 本文所提出的考虑流动状态的集成模型仍展现出优秀的预测性能, 相比于预测精度有限的传统关联式和可解释性有限的单一“黑盒”模型, 集成模型在降低过拟合风险、提高预测稳健性与可解释性方面具有实用价值.

表8 所有回归模型的预测性能

Table 8 Prediction performance of all regression models

模型/指标	*MARD*	*MRD*	$\eta_{20\%}$	R^2
回归器 I	9.08%	1.92%	91.6%	0.972
回归器 II	9.04%	3.34%	92.0%	0.968
回归器 III	4.37%	1.11%	98.5%	0.981
回归器 IV	6.15%	0.80%	98.4%	0.974
回归器 V	9.55%	4.14%	88.5%	0.966
集成模型	7.41%	2.71%	93.6%	0.979
单一加权模型	7.18%	2.67%	93.7%	0.981
单一未加权模型	7.60%	2.86%	93.9%	0.980

4 结论与讨论

考虑到流动状态对沸腾传热的重要影响, 本研究开发了一种基于谱聚类与 XGBoost 的集成机器学习模型, 用于预测微小通道流动沸腾传热系数. 该模型依托于一个包含 8309 个实验数据点的综合数据库,

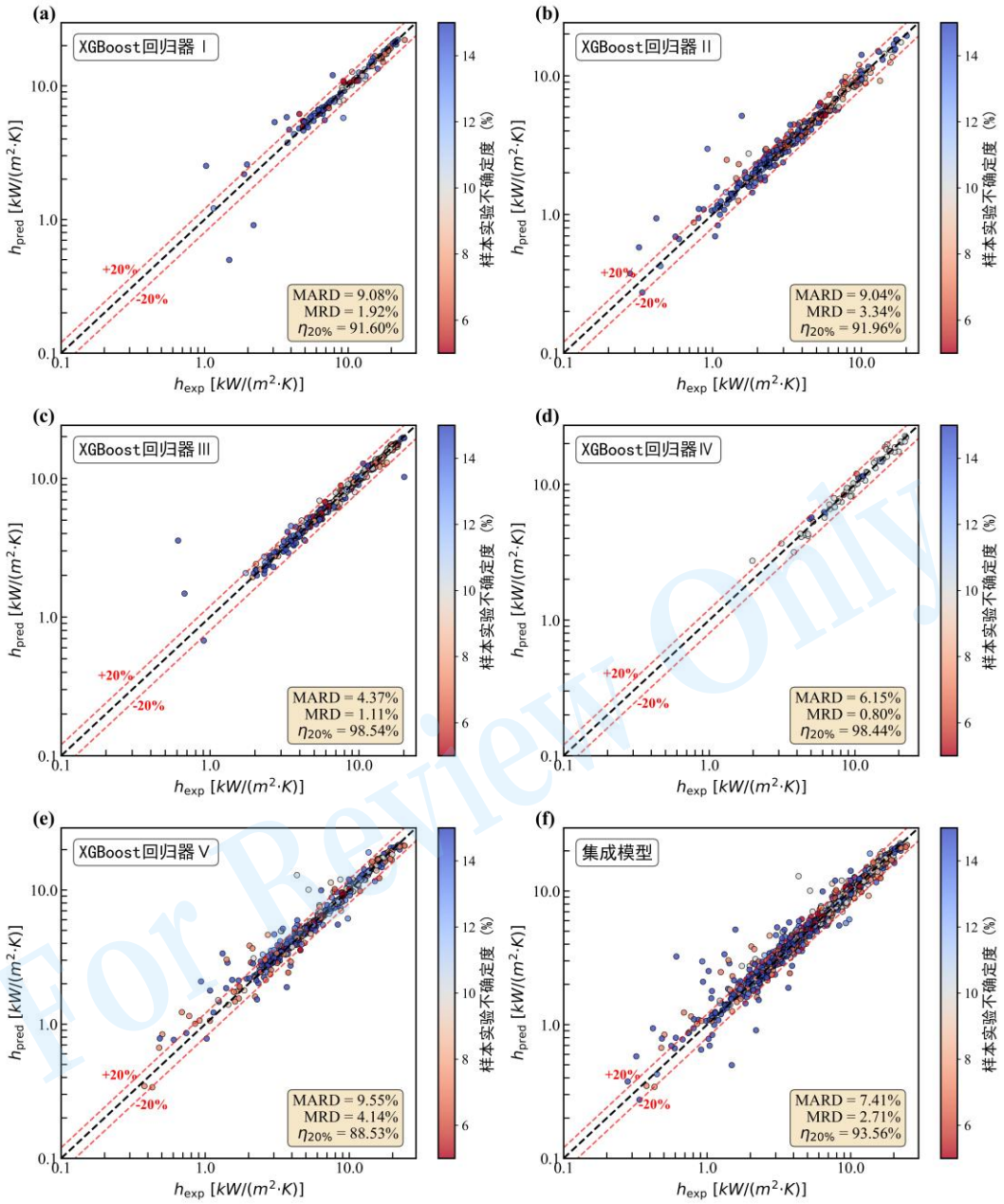


图 9 (网络版彩图) 子回归模型和集成模型在测试集上预测传热系数. (a)-(e) 子回归模型; (f) 集成模型

Figure 9 (Color online) Prediction of boiling HTCs by the sub-regressors and integrated model on the test set. (a)-(e) Sub-regressors; (f) Integrated model.

数据来源于 42 篇公开文献, 涵盖 24 种工质和广泛的运行工况. 主要结果如下:

(1)基于 MIC 法与 XGBoost 模型进行特征筛选, 将初始 47 个特征优化为 37 个. 在相同超参数搜索空间下, 采用贝叶斯优化分别训练了两个单一 XGBoost 模型: 一个引入以数据不确定度倒数平方加权的损失函数, 另一个为等权训练. 两者性能均显著优于传统关联式, 其中加权模型 *MARD* 为 7.18%, 略优于未加权模型的 7.60%, 表明加权策略可能有助于提升预测精度, 未来将考虑获取更多高质量数据进一步探索不同加权策略对模型的影响.

(2)以表观气液相速度为输入特征, 利用谱聚类

For Review Only

将数据库划分为五类流动状态，并训练 XGBoost 分类模型预测样本属于各聚类结果的概率. 该模型整体分类准确率达 99%，各类别的精确率、召回率与 F1 分数均超过 98%，表现出优异的分类性能. 本研究的局限性在于数据库中的样本所对应的真实流动状态未知，因此无法验证聚类结果是否能够很好地对应典型的流动状态，未来将进一步探索机器学习在流型识别和传热机制分析中的应用.

(3)基于聚类结果，分别为五类流动状态训练子 XGBoost 回归模型，各模型预测的 *MARD* 介于 4.37%–9.55%. 通过以分类概率为权重加权五个回归预测值的集成预测，最终集成模型的 *MARD* 为 7.41%，R^2 为 0.979. 该模型在保持高精度的同时，通过引入流动状态识别增强了物理可解释性，为沸腾传热预测提供了可靠工具.

符号说明

T	温度，℃
P	压力，*MPa*
G	质量流量，$kg/\left(m^2 s\right)$
x	蒸气干度, [-]
q	热通量，kW/m^2
D	水力直径, *mm*
g	重力加速度, m/s^2
Gp	定压比热容，$kJ/\left(kgK\right)$
h_{lat}	汽化潜热，kJ/kg
Re_l	液相雷洛数，$Re_l = G(1-x)D/\mu_l$, [-]
Re_{lo}	纯液相雷洛数，$Re_{lo} = GD/\mu_l$, [-]
Re_v	气相雷洛数，$Re_v = GxD/\mu_v$, [-]
Re_{vo}	纯气相雷洛数，$Re_{vo} = GD/\mu_v$, [-]
Fr_l	液相弗劳德数，$Fr_l = \left[G\left(1-x\right)\right]^2/\rho_l^2 gD$, [-]
Fr_{lo}	纯液相弗劳德数，$Fr_{lo} = G^2/\rho_l^2 gD$, [-]
Fr_v	气相弗劳德数，$Fr_v = \left(Gx\right)^2/\rho_v^2 gD$, [-]
Fr_{vo}	纯气相弗劳德数，$Fr_{vo} = G^2/\rho_v^2 gD$, [-]
We_l	液相韦伯数，$We_l = \left[G\left(1-x\right)\right]^2 D/\left(\rho_l\sigma\right)$, [-]
We_{lo}	纯液相韦伯数，$We_{lo} = G^2 D/\left(\rho_l\sigma\right)$, [-]
We_v	气相韦伯数，$We_v = \left(Gx\right)^2 D/\left(\rho_v\sigma\right)$, [-]
We_{vo}	纯气相韦伯数，$We_{vo} = G^2 D/\left(\rho_v\sigma\right)$, [-]
Ca_l	液相毛细数，$Ca_l = \mu_l G\left(1-x\right)/\left(\rho_l\sigma\right)$, [-]
Ca_{lo}	纯液相毛系数，$Ca_l = \mu_l G/\left(\rho_l\sigma\right)$, [-]
Ca_v	气相毛系数，$Ca_v = \mu_v Gx/\left(\rho_v\sigma\right)$, [-]
Ca_{vo}	纯气相毛系数，$Ca_{vo} = \mu_v G/\left(\rho_v\sigma\right)$, [-]
Su_l	液相 Suratman 数，$Su_l = \rho_l\sigma D/\mu_l^2$, [-]
Su_v	气相 Suratman 数，$Su_v = \rho_v\sigma D/\mu_v^2$, [-]
Pr_l	液相普朗特数，$Pr_l = Cp_l\mu_l/\lambda_l$, [-]
Pr_v	气相普朗特数，$Pr_v = Cp_v\mu_v/\lambda_v$, [-]
Bd	邦德数，$\left(\rho_l-\rho_v\right)gD^2/\sigma$, [-]
Bo	沸腾数，$q/\left(Gh_{lat}\right)$, [-]
Co	限制数，$\left(1/x-1\right)^{0.8}\left(\rho_l-\rho_v\right)^{0.5}$, [-]
Ga	伽利略数，$\left(\rho_l-\rho_v\right)\rho_l gD^3/\mu_l^2$, [-]
X_{tt}	Martinelli数，$\left(\mu_l/\mu_v\right)^{0.1}\left(\rho_v/\rho_l\right)^{0.5}\left(1/x-1\right)^{0.9}$, [-]
Fa	Fang数，$\left(\rho_l-\rho_v\right)\sigma/\left(G^2 D\right)$, [-]
P_R	对比压力, P_{sat}/P_{crit}
U_{ls}	表观液相速度, $G\left(1-x\right)/\rho_l$
U_{vs}	表观气相速度, Gx/ρ_v
R_P	表面粗糙度参数，μm
L	通道流动方向长度, *m*
P_F	通道湿润周长, *mm*
P_H	通道加热周长, *mm*通道加热周长, *mm*
h	沸腾传热系数，$kW/\left(m^2 K\right)$
R^2	决定系数, [-]
MARD	平均绝对相对误差, *%*
MRD	平均相对误差, *%*

希腊符号

ρ	密度，kg/m^3
λ	导热系数，$mW/\left(m^2 K\right)$
μ	动力粘度，$\mu Pa/s$
σ	表面张力，mN/m
$\eta_{20\%}$	预测误差在±20%范围内的样本所占百分比
$\eta_{50\%}$	预测误差在±50%范围内的样本所占百分比

下标

exp	实验值
pred	预测值
l, v	纯液相, 纯气相
lo, vo	纯液相, 纯气相
sat,crit	饱和, 临界
lat	潜热

For Review Only

致谢 过增元院士创立场协同与 Entransy 理论，举办国际传热大会，建立《International Journal of Heat and Mass Transfer》中国影响，首倡成立亚洲热科学与工程学会；为我国学者开拓研究方向、构筑学术平台作出历史性卓越贡献；作者们有幸参与其中．在过老师的指导与关心下，作者们开展了一系列研究，因而撰成此论文．在过老师九秩华诞之际，谨致以崇高敬意．感谢梁新刚、曹炳阳、陈群等三位教授的建设性意见．

参考文献

1 Mudawar I. Two-phase microchannel heat sinks: theory, applications, and limitations. J Electron Packag, 2011, 133: 041002
2 Mudawar I. Recent advances in high-flux, two-phase thermal management. Therm Sci Eng Appl, 2013, 5: 021012
3 Li W, Wu Z. A general correlation for evaporative heat transfer in micro/mini-channels. Int J Heat Mass Transf, 2010, 53: 1778–1787
4 Kattan N , Thome J R, Favrat D. Flow boiling in horizontal tubes: Part 3—development of a new heat transfer model based on flow pattern. J Heat Transf, 1998, 120: 156–165
5 Wojtan L, Ursenbacher T, Thome J R. Investigation of flow boiling in horizontal tubes: Part II—development of a new heat transfer model for stratified-wavy, dryout and mist flow regimes. Int J Heat Mass Transf, 2005, 48: 2970–2985
6 Cheng L, Ribatski G, Thome J R. New prediction methods for CO2 evaporation inside tubes: Part II—an updated general flow boiling heat transfer model based on flow patterns. Int J Heat Mass Transf, 2008, 51:125–135
7 Thome J R, Bar-Cohen A, Revellin R, et al. Unified mechanistic multiscale mapping of two-phase flow patterns in microchannels. Exp Therm Fluid Sci, 2013, 44: 1–22
8 Kim S M, Mudawar I. Universal approach to predicting saturated flow boiling heat transfer in mini/micro-channels–Part I. dryout incipience quality. Int J Heat Mass Transf, 2013, 64: 1226–1238
9 Cheng L, Ribatski G, Thome J R. Two-phase flow patterns and flow-pattern maps: Fundamentals and applications. Appl Mech Rev, 2008, 61: 050802
10 Thome J R, Hajal J E. Two-phase flow pattern map for evaporation in horizontal tubes: Latest version. Heat Transf. Eng, 2003, 24: 3–10
11 Rahim E, Revellin R, Thome J, et al. Characterization and prediction of two-phase flow regimes in miniature tubes. Int J Multiphase Flow, 2011, 37: 12–23
12 Kattan N, Thome J R, Favrat D. Flow boiling in horizontal tubes: Part 1—Development of a diabatic two-phase flow pattern map. J Heat Transf, 1998, 120: 140–147
13 Taitel Y, Dukler A E. A model for predicting flow regime transitions in horizontal and near horizontal gas-liquid flow. AIChE J, 1976, 22: 47–55
14 Triplett K A, Ghiaasiaan S M, Abdel-Khalik S I, et al. Gas–liquid two-phase flow in microchannels Part I: two-phase flow patterns. Int J Multiphase Flow, 1999, 25: 377–394
15 Wojtan L, Ursenbucher T, Thome J R. Investigation of flow boiling in horizontal tubes: Part I—A new diabatic two-phase flow pattern map. Int J Heat Mass Transf, 2005, 48: 2955–2969
16 Akbar M K, Plummer D A, Ghiaasiaan S M. On gas–liquid two-phase flow regimes in microchannels. Int J Multiphase Flow, 2003, 29: 855–865
17 Bertsch S S, Groll E A, Garimella S V. A composite heat transfer correlation for saturated flow boiling in small channels. Int J Heat Mass Transf, 2009, 52: 2110–2118
18 Li W, Wu Z. A general correlation for evaporative heat transfer in micro/mini-channels. Int J Heat Mass Transf, 2010, 53: 1778–1787
19 Sun L, Mishima K. An evaluation of prediction methods for saturated flow boiling heat transfer in mini-channels. Int J Heat Mass Transf, 2009, 52: 5323–5329
20 Kim S M, Mudawar I. Universal approach to predicting saturated flow boiling heat transfer in mini/micro-channels – Part II. Two-phase heat transfer coefficient. Int J Heat Mass Transf, 2013, 64: 1239–1256
21 Hughes M T, Kini G, Garimella S. Status, challenges, and potential for machine learning in understanding and applying heat transfer phenomena. J Heat Transf, 2021, 143: 120802
22 Qiu Y, Garg D, Zhou L, et al. An artificial neural network model to predict mini/micro-channels saturated flow boiling heat transfer coefficient based on universal consolidated data. Int J Heat Mass Transf, 2020, 149: 119211
23 Qiu Y, Vo T, Garg D, et al. A systematic approach to optimization of ANN model parameters to predict flow boiling heat transfer coefficient in mini/micro-channel heatsinks. Int J Heat Mass Transf, 2023, 202: 123728
24 Bard A, Qiu Y, Kharangate C R, et al. Consolidated modeling and prediction of heat transfer coefficients for saturated flow boiling in mini/micro-channels using machine learning methods. Appl Therm Eng, 2022, 210: 118305

For Review Only

25 Zhu G, Wen T, Zhang D. Machine learning based approach for the prediction of flow boiling/condensation heat transfer performance in mini channels with serrated fins. Int J Heat Mass Transf, 2021, 166: 120783

26 Zhu G, Liu S, Zhang D, et al. Transfer learning model to predict flow boiling heat transfer coefficient in mini channels with micro pin fins. Int J Heat Mass Transf, 2024, 220: 125020

27 Li W, Zhang G, Yang D. Comparative study for flow condensation heat transfer in horizontal enhanced tubes based on machine learning. Int J Heat Mass Transf, 2024, 224: 125330

28 Zhang G, Li W, Yang D, et al. Integrated machine learning model for condensation flow heat transfer in smooth and enhanced tubes. Energy, 2025, 317: 134592

29 Zhang J, Ma Y, Wang M, et al. Prediction of flow boiling heat transfer coefficient in horizontal channels varying from conventional to small-diameter scales by genetic neural network. Nucl Eng Technol, 2019, 51: 1897–1904

30 Agustiarini N, Hoang H N, Oh J T, et al. Application of machine learning to the prediction of the boiling heat transfer coefficient of R32 inside a multiport mini-channel tube. J Therm Anal Calorim, 2023, 148: 3137–3153

31 Sajjad U, Raza W, Hussain I, et al. On the prediction and optimization of the flow boiling heat transfer in mini and micro channel heat sinks. Prog Nucl Energy, 2024, 177: 105466

32 Luxburg U. A tutorial on spectral clustering. Stat Comput, 2007, 17: 395–416

33 Ding L, Li C, Jin D, et al. Survey of spectral clustering based on graph theory. Pattern Recogn, 2024, 151: 110366

34 Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Krishnapuram B, Shah M, eds. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 2016. 785–794

35 He H H, Li W, Zhu Y, et al. Machine learning analysis for condensation flow heat transfer in mini/micro-channels. Int J Heat Mass Transf,2026,255(P1):127775-127775

36 Yan Y Y, Lin T F. Evaporation heat transfer and pressure drop of refrigerant R-134a in a small pipe. Int J Heat Mass Transf, 1998, 41: 4183–4194

37 Wang C C, Chiang C S, Yu J G. An experimental study of in-tube evaporation of R-22 inside a 6.5-mm smooth tube. Int J Heat Fluid Flow, 1998, 19: 259–269

38 Yun R, Kim Y, Kim M S, et al. Boiling heat transfer and dryout phenomenon of CO2 in a horizontal smooth tube. Int J Heat Mass Transf, 2003, 46: 2353–2361

39 Huo X, Chen L, Tian Y S, et al. Flow boiling and flow regimes in small diameter tubes. Appl Therm Eng, 2004, 24: 1225–1233

40 Zhao X, Bansal P K. Flow boiling heat transfer characteristics of CO2 at low temperatures. Int J Refrig, 2007, 30: 937–945

41 Bertsch S S, Groll E A, Garimella S V. Effects of heat flux, mass flux, vapor quality, and saturation temperature on flow boiling heat transfer in microchannels. Int J Multiphase Flow, 2009, 35: 142–154

42 In S, Jeong S. Flow boiling heat transfer characteristics of R123 and R134a in a micro-channel. Int J Multiphase Flow, 2009, 35: 987–1000

43 Mastrullo R, Mauro A W, Rosato A, et al. Carbon dioxide local heat transfer coefficients during flow boiling in a horizontal circular smooth tube. Int J Heat Mass Transf, 2009, 52: 4184–4194

44 Wang L, Chen M, Groll M. Flow boiling heat transfer characteristics of R134a in a horizontal mini tube. J Chem Eng Data, 2009, 54: 2638–2645

45 Tibiriçá C B, Ribatski G. Flow boiling heat transfer of R134a and R245fa in a 2.3mm tube. Int J Heat Mass Transf, 2010, 53: 2459–2468

46 Bang K H, Kim K K, Lee S K, et al. Pressure effect on flow boiling heat transfer of water in minichannels. Int J Therm Sci, 2011, 50: 280–286

47 Copetti J B, Macagnan M H, Zinani F, et al. Flow boiling heat transfer and pressure drop of R-134a in a mini tube: an experimental investigation. Exp Therm Fluid Sci, 2011, 35: 636–644

48 Kyu H, Son C H. Evaporation flow pattern and heat transfer of R-22 and R-134a in small diameter tubes. Heat Mass Transfer, 2011, 47: 703–717

49 Wu J, Koettig T, Franke C, et al. Investigation of heat transfer and pressure drop of CO2 two-phase flow in a horizontal minichannel. Int J Heat Mass Transf, 2011, 54: 2154–2162

50 Balasubramanian K, Jagirdar M, Lee P S, et al. Experimental investigation of flow boiling heat transfer and instabilities in straight microchannels. Int J Heat Mass Transf, 2013, 66: 655–671

51 Col D D, Bortolin S, Torresin D, et al. Flow boiling of R1234yf in a 1mm diameter channel. Int J Refrig, 2013, 36: 353–362

52 Grauso S, Mastrullo R, Mauro A W, et al. Flow pattern map, heat transfer and pressure drops during evaporation of R-1234ze(E) and R134a in a horizontal, circular smooth tube: Experiments and assessment of predictive methods. Int J Refrig, 2013, 36: 478–491

53 Wang S, Gong M Q, Chen G F, et al. Two-phase heat transfer and pressure drop of propane during saturated flow boiling inside a horizontal tube. Int J Refrig, 2014, 41: 200–209

54 Xu Y, Fang X, Li G, et al. An experimental study of flow boiling heat transfer of R134a and evaluation of existing correlations. Int J Heat Mass Transf, 2016, 92: 1143–1157

55 Wen T, Zhan H, Lu L, et al. Experimental investigation and development of new correlation for flow boiling heat transfer in mini-channel. Int J Therm Sci, 2018, 129: 209–217

56 Diani A, Rossetto L. Experimental analysis of refrigerants flow boiling inside small sized microfin tubes. Heat Mass Transfer, 2018, 54:

For Review Only

2315–2329
57 Li H, Hrnjak P. Heat transfer and pressure drop of R32 evaporating in one pass microchannel tube with parallel channels. Int J Heat Mass Transf, 2018, 127: 526–540
58 Lillo G, Mastrullo R, Mauro A W, et al. Flow boiling heat transfer, dry-out vapor quality and pressure drop of propane (R290): Experiments and assessment of predictive methods. Int J Heat Mass Transf, 2018, 126: 1236–1252
59 Zhang L, Jiang L L, Liu J H, et al. Investigation of flow boiling heat transfer characteristics of CO2 in horizontal mini-tube. Int J Therm Sci, 2019, 138: 109–115
60 Mastrullo R, Mauro A W, Viscito L. Flow boiling of R452A: Heat transfer data, dry-out characteristics and a correlation. Exp Therm Fluid Sci, 2019, 105: 247–260
61 Wen T, Zhan H, Zhang D. Flow boiling heat transfer in mini channel with serrated fins: Experimental investigation and development of new correlation. Int J Heat Mass Transf, 2019, 128: 1081–1094
62 Mastrullo R, Mauro A W, Viscito L. Flow boiling of R452A: heat transfer data, dry-out characteristics and a correlation. Exp Therm Fluid Sci, 2019, 105: 247–260
63 Lillo G, Mastrullo R, Mauro A W, et al. Experimental thermal and hydraulic characterization of R448A and comparison with R404A during flow boiling. Appl Therm Eng, 2019, 161: 114146
64 Li H, Hrnjak P. Heat transfer coefficient, pressure drop, and flow patterns of R1234ze(E) evaporating in microchannel tube. Int J Heat Mass Transf, 2019, 138: 1368–1386
65 Kim C H, Kim N H. Evaporation heat transfer and pressure drop of low GWP R-404A alternative refrigerants in a multiport tube. Int J Heat Mass Transf, 2022, 184: 122386
66 Xu Y, Yan Z, Li L. Experimental investigation on flow boiling heat transfer and pressure drop of a low-GWP refrigerant R1234ze(E) in a horizontal minichannel. Energies, 2021, 14: 5972
67 Hellenschmidt D, Petagna P. Effects of saturation temperature on the boiling properties of carbon dioxide in small diameter pipes at low vapour quality: Heat transfer coefficient. Int J Heat Mass Transf, 2021, 172: 121094
68 Hoang H N, Agustiarini N, Oh J T. Experimental investigation of two-phase flow boiling heat transfer coefficient and pressure drop of R448A inside multiport mini-channel tube. Energies, 2022, 15: 4331
69 Arcasi A, Mauro A W, Napoli G, et al. Heat transfer coefficient, pressure drop and dry-out vapor quality of R454C. Flow boiling experiments and assessment of methods. Int J Heat Mass Transf, 2022, 188: 122599
70 Xu Y, Yan Z, Li L. Flow boiling heat transfer, pressure drop and flow patterns of the environmentally friendly refrigerant R1234yf for cooling avionics. Appl Therm Eng, 2022, 209: 118301
71 Liu H, Yuan C, Li H, et al. Experimental study of R1234yf flow boiling in small tubes at saturation temperature of 5 °C and low mass flux. Appl Therm Eng, 2023, 229: 120653
72 Bediako E G, Rullière R, Marchetto D B, et al. Flow boiling heat transfer of R1234ze(E) in a horizontal mini-channel at medium and high saturation temperatures. Int J Heat Mass Transf, 2024, 226: 125495
73 Diani A, Rossetto L. Characteristics of R513A evaporation heat transfer inside small-diameter smooth and microfin tubes. Int J Heat Mass Transf, 2020, 162: 120402
74 Wu Z, Yuan C, Li H, et al. Flow boiling heat transfer of R454B in a 24-port microchannel tube. Appl Therm Eng, 2024, 248: 123150
75 Ren C, Song Q, Sun S, et al. Experimental study on flow boiling heat transfer of zeotropic mixture R290/R601a in a mini-channel. Energies, 2025, 18: 2074
76 Hoang H N, Oh J T, Kim J K. Flow boiling of R448A inside mini-channel: an integrated experimental and computational investigation. Int J Air Cond Refrig, 2025, 33: 2345
77 NIST standard reference database 23: reference fluid thermodynamic and transport properties-REFPROP, version 9.0. Gaithersburg, MD: National Institute of Standards and Technology. 2010
78 Reshef D N, Reshef Y A, Finucane H K, et al. Detecting novel associations in large data sets. Science, 2011, 334: 1518–1524
79 Nguyen V. Bayesian optimization for accelerating hyper-parameter tuning. In: Proceedings of IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Sardinia, Italy, 2019. 302–305

Prediction of flow boiling heat transfer in micro/mini-channels based on integrated machine learning

LI Wei[1], HE HuanHuan[1] & DONG Yuan[2]

[1] *Department of Energy Engineering, Zhejiang University, Hangzhou 310027, China*
[2] *School of Mechanical Engineering, Hangzhou Dianzi University, Hangzhou 310018, China*

94

For Review Only

To address the prediction of flow boiling heat transfer coefficients in micro/mini-channels, an integrated machine learning model combining spectral clustering and eXtreme Gradient Boosting (XGBoost) was developed. By integrating 8309 experimental data points from 42 literature sources, a comprehensive database was constructed, covering 24 working fluids and channel hydraulic diameters ranging from 0.19 to 6.5 mm. The Maximum Information Coefficient method was employed to screen 37 parameters from 47 initial features as model inputs. Given the significant influence of flow patterns on heat transfer characteristics, the model first utilizes gas and liquid superficial velocities with spectral clustering to categorize the data into five classes representing distinct flow patterns. Subsequently, an XGBoost classification model (achieving 99% accuracy) was trained to predict the probability of a sample belonging to each flow pattern category. A dedicated XGBoost regression sub-model was established for each category. The final integrated prediction is achieved by a weighted fusion of the classification probabilities and the regression results. Testing demonstrates that the integrated model achieves a mean absolute relative difference (*MARD*) of 7.41% and a coefficient of determination (R^2) of 0.979, significantly outperforming traditional empirical correlations. This method maintains high prediction accuracy while enhancing model interpretability through flow pattern recognition, providing a novel approach for flow boiling heat transfer prediction.

flow boiling heat transfer, integrated machine learning, heat transfer coefficient, spectral clustering, XGBoost

For Review Only