

极端环境下碱激发模拟月壤：基于数据驱动的微观结构表征与性能预测

姚羿舟^{1,2}, 朱 超^{1,2,3}, 张 韦^{1,2}, 刘 超^{1,2,3}

(1. 西安建筑科技大学土木工程学院, 西安 710055; 2. 西安建筑科技大学数字化绿色建筑研究院, 西安 710055;
3. 陕西省智能建造未来产业创新研究院, 西安 710055)

摘 要: 碱激发月壤作为极具应用前景的月基建筑材料, 准确表征、理解和预测其在极端环境下的性能演化规律是确保月基建筑结构在施工与服役过程中安全可靠的基础。本工作制备了氢氧化钠基和硅酸钠基碱激发模拟月壤(AALRS)净浆并在真空和热真空环境下原位养护。采用降噪、聚类、堆叠算法与超参数优化算法, 结合多尺度试验表征, 提出基于数据驱动的“算法解构-性能预测-机制解析”的 AALRS 跨尺度智能分析范式, 分别构建了可解释的特征 X 射线能谱(EDS)图像定量分析框架与强度预测模型, 并基于此开展极端环境下 AALRS 微观结构表征与宏观性能预测。结果表明, 与真空环境相比, 热真空环境下 AALRS 有更高的产物相占比与更低的孔隙率, 其物相分布更均匀, 基体展现出更高的力学强度。而真空环境下 AALRS 出现孔隙率增加、碱溶液残留与物相分布异质性加剧的现象, 基体力学强度明显下降。基于物相定量分析结果揭示了极端环境下 AALRS 的微观结构特征与反应机理。在性能预测模型中, 基于随机搜索算法优化的 Stacking-RS 模型展现出最优的预测精度、泛化能力和计算效率。进一步基于全局特征重要性分析、特征依赖分析与特征局部贡献分析解构 Stacking-RS 模型的预测逻辑并分析强度影响机制。研究成果推动了极端环境下 AALRS 材料表征和性能预测由经验试错驱动向智能数据驱动的研究范式转变。

关键词: 碱激发月壤; 极端环境; 机器学习; 微观结构; 性能预测

中图分类号: TU528; P184 文献标志码: A 文章编号: 0454-5648(2026)08-2673-29

网络出版时间: 2026-07-21



月基建造作为深空探测的核心战略方向, 亟须研发能够适应极端环境的月基建筑材料以实现月面长期驻留与原位资源开发。月地间的遥远距离及月表高真空(约 $10^{-12} \sim 10^{-10}$ Pa)、高低温宽温域($-180 \sim 120$ °C)和小重力(1.62 m/s^2)等严苛环境条件, 使传统建筑材料因运输成本高、资源消耗大和环境适应性差, 而难以满足月基结构的工程需求^[1,2]。利用月球当地资源并原位制备建筑材料成为破解“地外建造依赖地面制造”难题的关键路径。碱激发月壤通过碱激发剂与月壤中活性硅铝酸盐反应生成高强度凝胶网络, 兼具较好的力学性能、优异的耐久性能、较低的资源能耗和简便的制备工艺等优势, 已成为未来大规模月基建造的潜在候选材料^[3-5]。

为研发适应月面极端环境的高性能碱激发月

壤材料, 国内外学者普遍以模拟月壤作为真实月壤的替代材料, 并围绕模拟月壤物化性质、活化技术、配合比设计、成形工艺和环境适应性等关键因素开展了系统研究^[6-13]。Geng 等^[6]研究了氢氧化钠、硅酸钠、碳酸钠和硫酸钠等碱激发剂对碱激发模拟月壤(AALRS)力学性能和微观结构的影响。Wu 等^[7]研究了不同粒径分布和矿物组成的模拟月壤原料对 AALRS 性能的影响, 并基于统计学分析揭示了影响 AALRS 硬化过程的关键因素。Xue 等^[8]发现热活化处理能够显著提升模拟月壤中 Si 和 Al 的溶解度, 促使 AALRS 形成更致密的微观结构, 进而提高强度。Davis 等^[9]发现真空暴露会削弱 AALRS 的力学性能, Mills 等^[10]也报道了类似现象, 而 Zhou 等^[11]指出在真空暴露下 AALRS 具有

收稿日期: 2025-11-13。 修订日期: 2025-12-01。

基金项目: 国家自然科学基金资助项目(52178251); 陕西省秦创原“科学家+工程师”队伍建设项目(2023KXJ-242); 陕西省教育厅产业化项目(23JC047); 陕西高校青年创新团队(2023-2026)。

第一作者: 姚羿舟(1999—), 男, 博士研究生。

通信作者: 刘 超(1982—), 男, 博士, 教授。

Received date: 2025-11-13. Revised date: 2025-12-01.

First author: YAO Yizhou (1999—), male, Doctoral candidate.

E-mail: yaoyizhou@xauat.edu.cn

Correspondent author: LIU Chao (1982—), male, Ph.D., Professor.

E-mail: chaoliu@xauat.edu.cn

更高强度, 并认为这归因于风化效应的减弱。Zhang 等^[12]指出高低温循环(−180~100 °C)会导致 AALRS 强度显著降低, 而 Xiong 等^[13]通过分子动力学模拟和试验表明 AALRS 在高温(120 °C)、低温(−30 °C)和大温差循环(−196~25 °C)下其 Si—O—Al 网络结构及强度仍保持稳定。

尽管现有研究为 AALRS 在月球极端环境中的性能演化规律提供了宝贵见解, 但研究结果和结论仍存在差异和争议。这主要源于以下挑战: 1) 月壤固有的高比例复杂矿物成分(如长石、辉石、橄榄石和钛铁矿)及其物化性质的地域波动性; 2) 碱激发月壤性能发展受多维因素耦合影响, 其作用机制复杂; 3) 对微观结构特征的认识普遍依赖于局部定性判断, 缺乏系统的全局定量描述; 4) 极端环境下微观结构特征尚不明晰, 宏观性能演化规律仍缺乏微观机理层面的支撑; 5) 多数研究局限于有限样本数据的传统统计分析方法, 尚未构建特征变量与性能的定量关联, 难以做到对性能演化规律的全面认知与可靠预测。因此, 亟须提出可解决上述挑战的研究范式, 以深入揭示 AALRS 的微观结构特征并精准预测其宏观性能, 从而推动月基建筑材料的智能设计, 并以此确保月基建筑结构在短期施工与长期服役过程中的安全可靠。

机器学习(ML)凭借其强大的数据特征挖掘能力与非线性映射优势, 正加速推动材料科学领域研究范式的革新^[14]。Naser^[15]指出 ML 算法可用于预测地外建筑材料在极端环境下的性能响应, 有助于

加深对其复杂演化机制的理解。以监督学习为核心的 ML 算法能够捕捉输入参数间的隐式关联, 可有效预测如强度、渗透性、吸水性、触变性及耐久性等宏观性能指标^[16]。在微观结构表征方面, 李粒琿等^[17]利用机器学习算法实现物相的动态演化追踪。Yu 等^[18]利用卷积神经网络(CNN)实现了背散射电子(BSE)图像的像素级物相分割, 推动复合胶凝材料的微观结构定量表征与分析。尽管 ML 技术已在传统水泥基材料研究领域展现出巨大潜力, 但在组分复杂、服役环境极端的 AALRS 领域, 研究仍多依赖传统经验试错范式, 鲜见系统引入先进 ML 算法以解决上述关键挑战的相关研究报道。

本工作利用聚类算法、堆叠集成算法与超参数优化算法, 结合多尺度试验表征, 提出了基于数据驱动的“算法解构—性能预测—机制解析”的 AALRS 跨尺度智能分析范式(图 1)。在微观层面, 构建基于无监督聚类算法的特征 X 射线能谱(EDS)图像分析框架, 实现了极端环境下 AALRS 微观结构特征的算法解构、自动识别、定量表征和机制理解。在宏观层面, 构建堆叠集成强度预测模型, 并通过多种超参数优化算法进行超参数寻优, 结合沙普利可加性解释(SHAP)分析, 实现了极端环境下 AALRS 强度预测与关键特征的贡献量化。在此基础上, 进一步建立微观结构定量表征与宏观性能预测之间的关联, 阐明二者在该范式下的协同分析机制。研究成果为面向月面极端环境的月基建筑材料智能设计、性能调控与优化提供了理论基础与方法支撑。

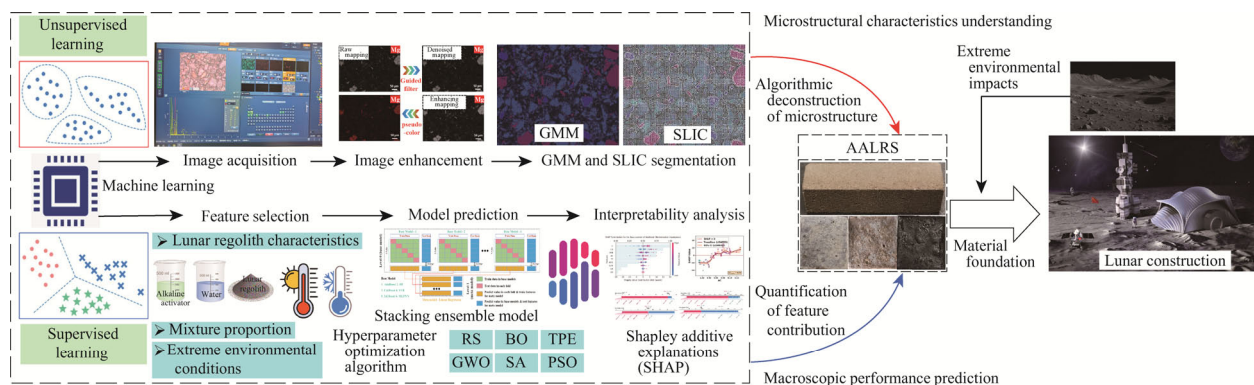


图 1 基于数据驱动的 AALRS 跨尺度智能分析范式^[19]

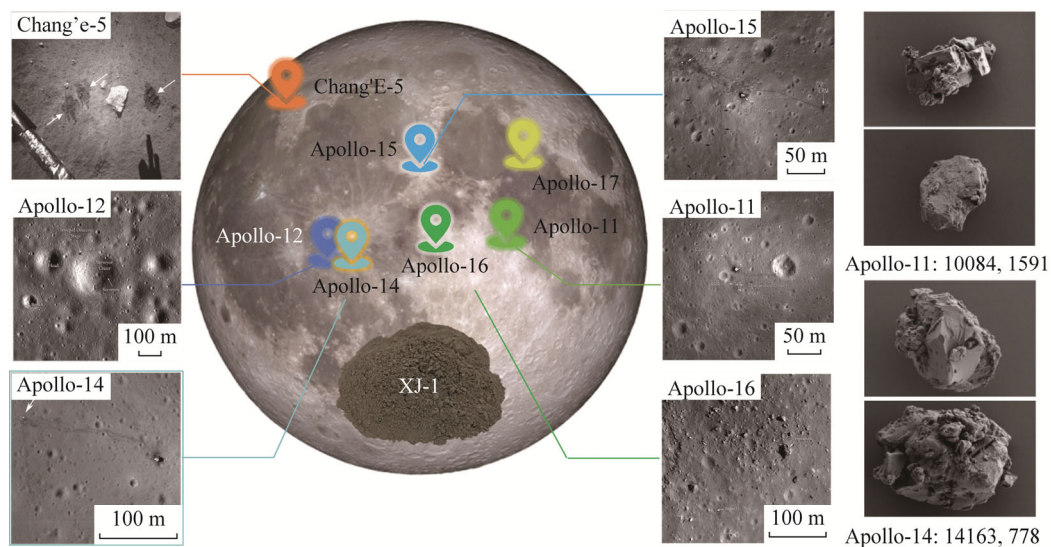
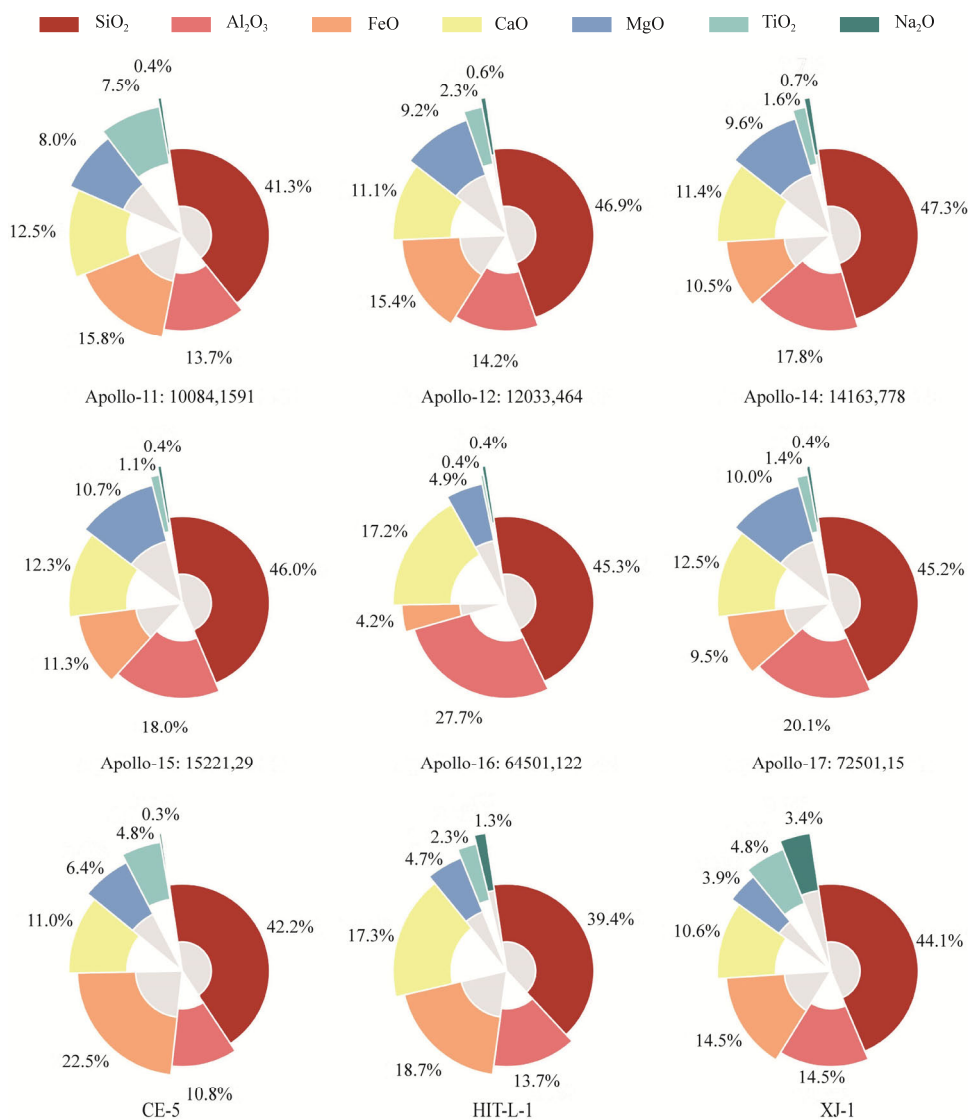
Fig. 1 Data-driven AALRS cross-scale intelligent analysis paradigm^[19]

1 实验

1.1 模拟月壤

模拟月壤 XJ-1 源自中国吉林火山渣矿床, 模拟对象为月海低钛型月壤。为验证 XJ-1 作为模拟

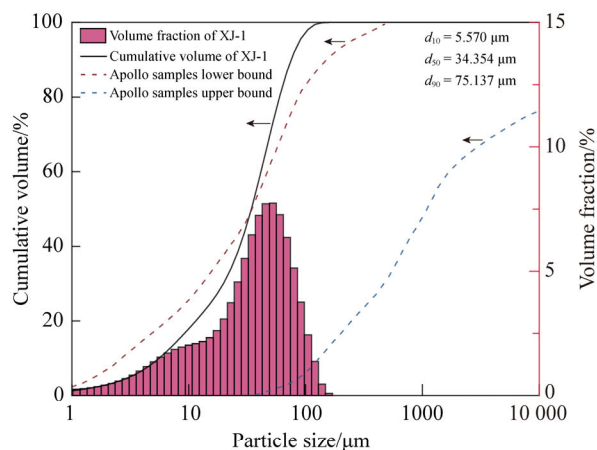
月壤的合理性和科学性, 在组成、形貌和性质方面将其与真实月壤进行对比。图 2 展示了月球探测影像及真实月壤的微观形貌^[20, 21]。由图 2 可见, 月壤颗粒展现出不规则形状及熔融玻璃质层表面特征。图 3 比较了真实月壤与模拟月壤的主要化学成

图 2 月球探测图像与月球风化层微观形貌^[20, 21]Fig. 2 Lunar exploration image and micromorphology of lunar regolith^[20, 21]图 3 真实月壤与模拟月壤的化学成分^[8, 22]Fig. 3 Chemical compositions of real lunar regolith and lunar regolith simulant^[8, 22]

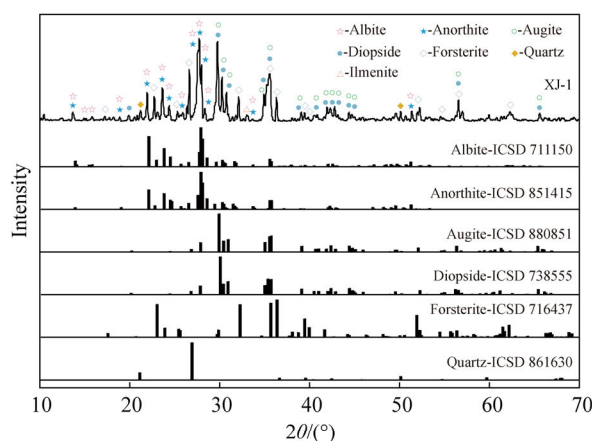
分^[8, 22]。结果表明, 真实月壤以 SiO_2 为主, 含量均大于 40%。 Al_2O_3 、 FeO 、 CaO 、 MgO 和 TiO_2 含量占比同样较高, 并表现出显著的区域波动性。模拟月壤(如 HIT-L-1 和 XJ-1)在主要氧化物含量上与真实月壤有较高的相似性。其中 XJ-1 的化学组成与 Apollo12-12033,464 样品高度相似, SiO_2 、 Al_2O_3 和 CaO 的含量差异分别仅为 2.8%、0.3%和 0.5%。

图 4 展示了模拟月壤 XJ-1 的基本性质, 其粒径级配大致处于 Apollo 样品粒径级配范围之内^[6],

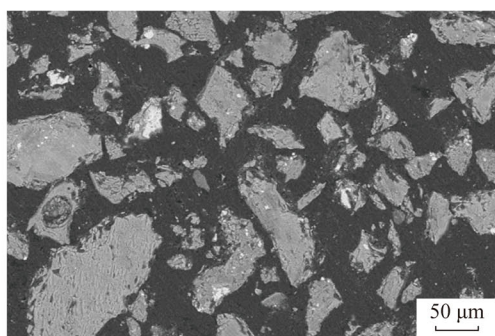
中值粒径为 $34.35 \mu\text{m}$ 。XJ-1 的主要矿物成分为钠长石、钙长石、辉石、镁橄榄石和石英。BSE 成像显示, XJ-1 颗粒具有长条状、多棱角、次棱角和钩角等典型不规则形态特征。尽管 XJ-1 含多种物相, 但其在 BSE 照片中展现出重叠的灰度范围, 难以依赖单一维度的灰度阈值精确识别物相。通过 EDS 面扫描能谱表征元素分布并构建元素-物相映射, 检测到长石、辉石、镁橄榄石和钛铁矿等矿物相存在。综合上述表征结果可见, XJ-1 在化学组成、



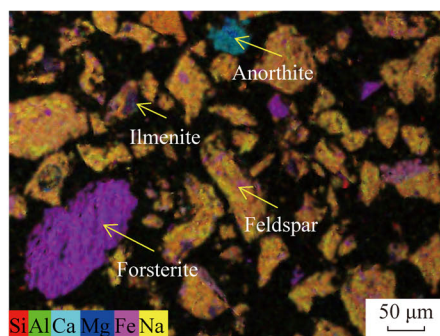
(a) Particle size distribution of XJ-1



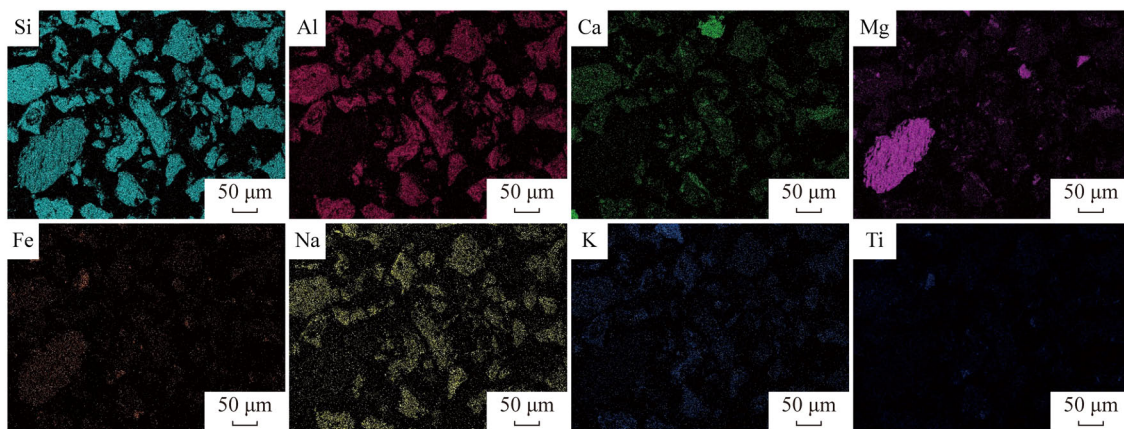
(b) XRD patterns of XJ-1



(c) BSE image of XJ-1



(d) Element to phase mapping



(e) EDS mapping images of XJ-1

图 4 模拟月壤 XJ-1 的性质
Fig. 4 Properties of the lunar regolith simulant XJ-1

矿物成分、粒径级配和微观形貌方面与真实月壤样本高度相似, 为 AALRS 研究提供可靠的材料基础。另外, 碱激发剂采用氢氧化钠(SH)溶液和模数为 1 的硅酸钠(SS)溶液。

1.2 试验方案

碱当量设定为 7.5%, 水与 XJ-1 的质量比设定为 0.3。基于月球表面温度的变化规律^[23], 高温环境温度设定为 90 °C, 并设置 20 °C 的常温对照组以研究温度影响。环境气压恒定为 100 Pa 用于模拟真空环境。试验设计共考虑环境、激发体系和龄期 3 个控制因素。暴露环境分为真空(V)与热真空(HV), 激发体系包括氢氧化钠激发体系与硅酸钠激发体系, 龄期设置为 1、3、7、14 d 和 28 d。试验组命名中的 3 个字段依次代表激发体系, 环境和龄期。例如, SH-V-1 代表采用氢氧化钠激发体系并在真空环境下养护 1 d 的 AALRS 试样。

AALRS 试样的制备流程如下: 将 XJ-1、碱激发剂与水混合, 并恒速搅拌 5 min, 随即将拌和物倒入 40 mm×40 mm×40 mm 模具, 并振动 30 s 以消除气泡。为避免试样在真空暴露过程中发生膨胀变形, 将其先置于 60 °C 常压环境预养护 1 d, 使其初步硬化后脱模, 此阶段不计入龄期。随后试样分别置于真空(20 °C, 100 Pa)和热真空(90 °C, 100 Pa)环境中养护至规定龄期。到达龄期后对试样进行抗压强度测试。在微观表征中, 选取 7 d 龄期试样, 从中心部位切取边长约为 10 mm 的立方体块, 并浸泡于异丙醇溶液 5 d 以终止反应, 随后在 40 °C 下干燥至质量恒定。

1.3 抗压强度测试

抗压强度采用 WAW-1000 电子万能试验机进行测定。加载速率设定为 480 N/s, 以确保 AALRS 在加载过程中的稳定性。每个试验组测定 6 个试样, 并计算其强度的平均值和标准差。

1.4 微观结构表征

采用配备 UltimMax 100 能谱仪的扫描电子显微镜(ZEISS, Gemini SEM 500)对经镶嵌、抛光和喷金处理后的样品进行 EDS 面扫描能谱采集。面扫描能谱分辨率设置为 1024×768 pixels, 对应的实际视场范围为 568.89 μm×426.67 μm。加速电压设为 15 kV, 工作距离保持在 10 mm, 每像素扫描时间设为 100 μs/pixel、处理时间设为 6、通道数量设为 2048、收集帧数设为 3。在数据采集过程中, 特征 X 射线每秒输入计数率稳定在 30 000~50 000 cps 之间, 探测器死时间控制在 20%~30%之间。每组

样品选取 15 个不同区域进行 EDS 能谱面扫描, 每个区域同时获取硅(Si)、铝(Al)、钠(Na)、钙(Ca)、镁(Mg)和铁(Fe)元素的面扫描能谱。

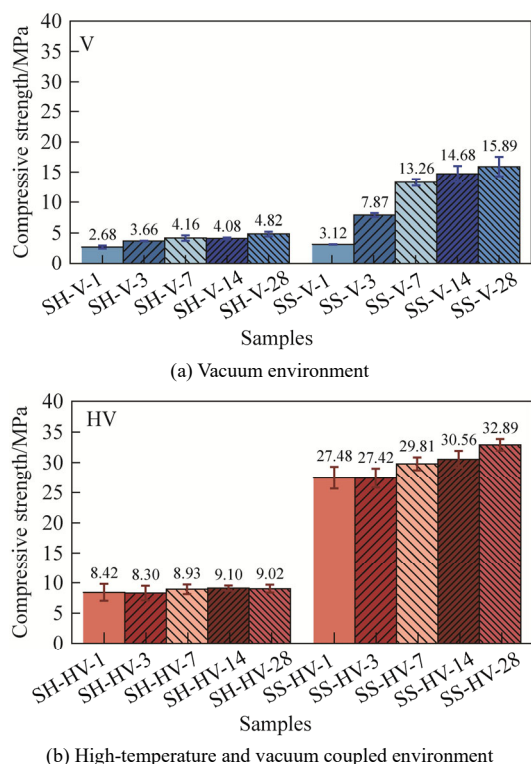
2 结果与讨论

2.1 抗压强度

图 5 展示了不同龄期和环境下由 SH 与 SS 激发的 AALRS(分别记为 AALRS-SH 与 AALRS-SS)抗压强度结果。整体来看, AALRS-SS 在各龄期均表现出显著高于 AALRS-SH 的强度水平, 且 HV 环境对强度的提升作用明显优于 V 环境。以 28 d 强度为例, SH-V-28、SS-V-28、SH-HV-28 和 SS-HV-28 的抗压强度分别为 4.82、15.90、9.02 MPa 和 32.90 MPa。其中 SS-HV-28 的强度为 SH-HV-28 的 3.65 倍, 进一步说明 SS 体系具有更强的激发能力。从龄期变化趋势来看, SH 体系强度增长较为缓慢, 在 V 环境中, 其强度由 2.68 MPa 上升至 4.82 MPa, 增长率为 79.9%。而在 HV 环境下, 其强度仅由 8.42 MPa 增至 9.02 MPa, 增长率仅为 7.1%。相较之下, SS 体系的强度增长更为显著, 其中 SS-V 的 28 d 强度增长率高达 409.2%, SS-HV 的 28 d 强度增长率为 19.86%。以上结果表明, 激发体系与极端环境均会对强度发展产生显著影响, 其差异主要源于微观结构演化机制的不同。为进一步阐明其内在规律, 后续将基于微观层面的 EDS 面扫描元素分布图定量分析结果, 对各组 AALRS 的微观结构特征及其与强度演化之间的内在联系进行系统分析。

2.2 基于聚类算法的 EDS 面扫描元素分布图定量分析框架

针对 AALRS 这一复杂胶凝材料体系, 传统微观表征方法难以全局量化决定其强度的关键物相演化特征。为此, 本工作借助 EDS 的二维元素分布成像能力, 提出基于聚类算法驱动的 EDS 面扫描元素分布图定量分析框架^[24]。该框架以元素分布图为原始数据, 采用自编写的 Python 脚本, 调用 OpenCV、scikit-learn 和 scikit-image 等库, 实现元素分布数据自动化处理与物相分割。通过智能算法与材料学知识的融合, 可实现 AALRS 体系微观结构的客观识别、自动分割与定量表征, 进而挖掘物相分布的隐含信息与潜在反应机制。该框架为地外复杂胶凝体系的微观结构分析和反应机制理解提供技术支撑, 并为月基材料研发与工艺优化提供微观层面的理论依据。



Note: First, second and third characters of the sample name represent the alkali-activated system, environment and curing age

图 5 AALRS 的抗压强度
Fig. 5 Compressive strength of AALRS

图 6 展示了 EDS 面扫描元素分布图定量分析框架，其主要包括以下关键步骤：

1) EDS 面扫描元素分布图预处理：由于硅漂移能谱探测器在短时间采集下像素信号计数偏低^[25]，且特征 X 射线与样品间的作用体积会引入高频纹理噪声，原始 EDS 面扫描元素分布图常表现出非均质亮度分布且分辨率不足，从而影响特征提取与分类。为此，采用引导滤波进行降噪处理，其在移除高频纹理的同时完整保留边缘特征与灰度整体变化^[26]。针对降噪后灰度范围压缩的问题，引入对数校正算法增强边缘对比度，通过非线性拉伸提升各元素的视觉分离度，以提高不同特征的可分性。

2) 复合 EDS 面扫描能谱构建与聚类分析：将特定元素分配至 RGB 颜色通道并合成复合 EDS 面扫描能谱。采用高斯混合模型(GMM)算法对像素进行聚类划分。综合聚类结果并结合材料学知识确定聚类数为 6。原始复合 EDS 面扫描能谱受噪声干扰导致元素分布的空间连续性破坏，物相边界模糊。经步骤(1)增强处理后，复合 EDS 面扫描能谱信噪比显著提高，相边界更加清晰。

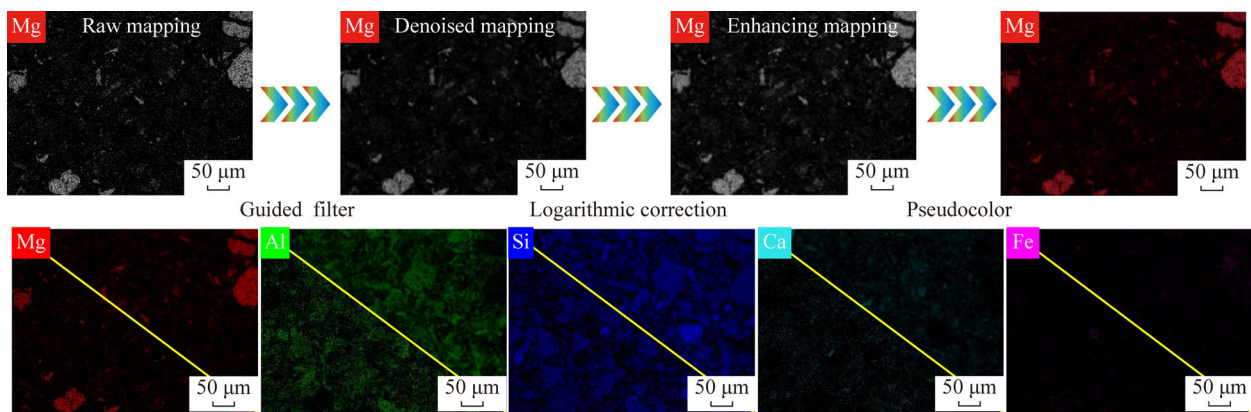
3) 物相掩模提取与占比计算：将聚类结果经

二值化处理后生成物相掩模，用于确定各物相的分布范围。通过计算掩模面积可获得物相的二维面积占比。

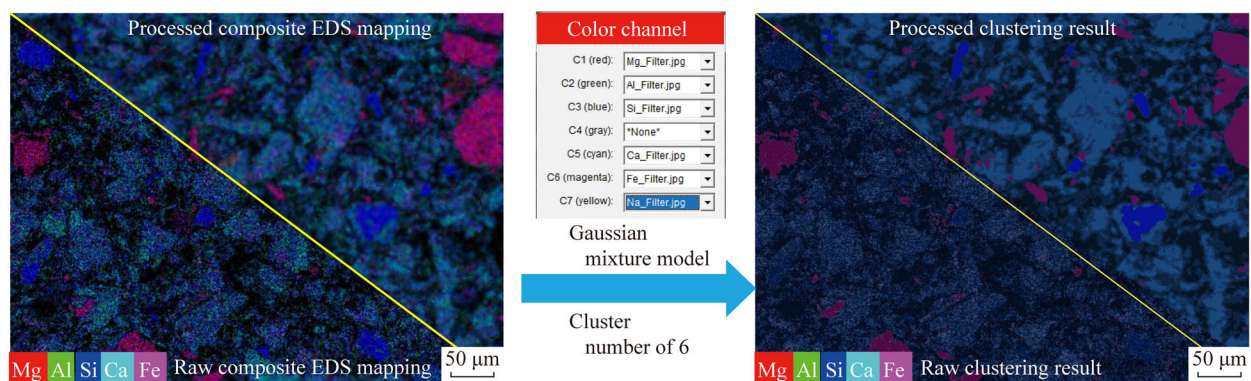
4) 超像素聚类分割及元素组成分析：为探究 AALRS 中元素的空间分布与组成特征，利用简单线性迭代聚类算法(SLIC)对复合 EDS 面扫描元素分布图进行超像素分割。该算法通过聚合颜色相似且空间邻近的像素生成超像素，不仅显著降低数据冗余，且其尺度与 EDS 作用体积尺度相匹配，从而提高元素信息的可靠性，有效解决像素级分辨率与 EDS 信号作用体积的尺度失配问题。为精准表征元素含量变化，移除孔隙掩模内的超像素以避免孔隙分布干扰。综合考虑超像素尺寸与计算效率，将超像素数设定为 10 000 个。以超像素为基本单元，提取其平均元素信号强度与空间质心坐标，进而构建包含空间模态和化学模态的超像素数据表。在此基础上，基于元素组成特征建立分析策略，生成超像素基元素散点图与三元相图，实现元素组成特征及其演化规律的可视化表征。

图 7 展示了各组 AALRS 的复合 EDS 面扫描元素分布图。如图 7 所示，复合 EDS 面扫描元素分布图通过元素信号的色彩可视化，直观揭示了微观区域内多元素分布及其相互作用特征。为突出关键物相，统一设定颜色与元素的对应关系如下：红色-Mg、绿色-Al、蓝色-Si、青色-Ca、洋红色-Fe。由于 Na 信号强度差异较小且分布广泛，通常不纳入颜色通道，以避免干扰聚类结果。然而，在 SH-V 组中 Na 信号表现出特征性富集，因此将其纳入黄色通道以表征该特征相。通道合并后，各物相均呈现独立的颜色特征，可据此识别 AALRS 体系的主要物相分布。高 Si 高 Al 区域对应长石相，高 Si 低 Al 区域对应石英相，高 Ca 区域代表钙长石相，高 Mg 区域为镁橄榄石或辉石相，高 Fe 区域为钛铁矿相，高 Na 区域则代表孔隙中残留的氢氧化钠溶液。

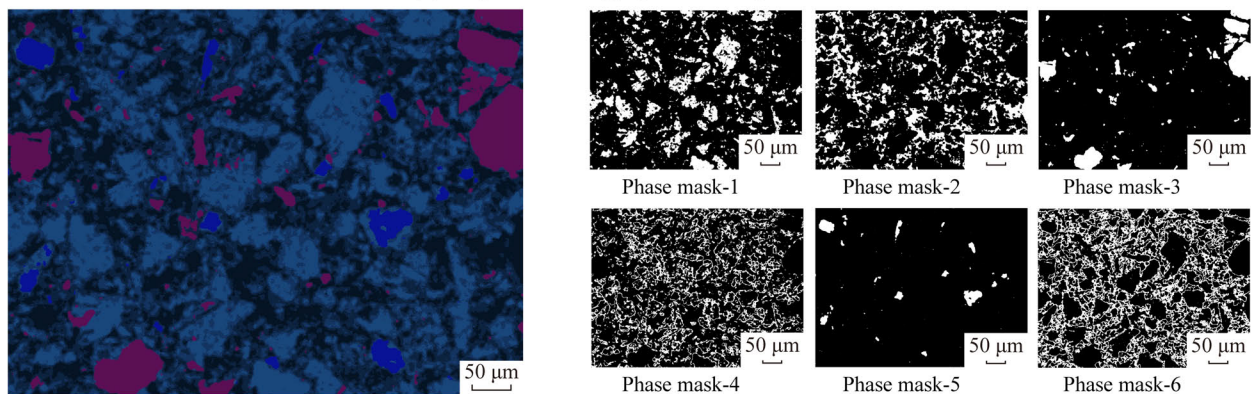
基于复合 EDS 面扫描元素分布特征，进一步采用 GMM 算法进行聚类分析，并获得对应的物相掩模(图 8)，其与实际矿物成分之间的对应关系如表 1 所示。由图 8 可见，AALRS-SH 的物相掩模包括活性 XJ-1 相、惰性 XJ-1 相、产物相、富镁相、孔隙相与富钠(硅)相。具体而言，高 Al 区域被划分为活性 XJ-1 相，而低 Al 高 Ca 或高 Fe 区域被划分为惰性 XJ-1 相，其反应活性显著低于前者。高 Mg 区域被划分为富镁相，即镁橄榄石或辉石。XJ-1 相外部区域 Si、Al、Na 元素的富集特征表明外部反



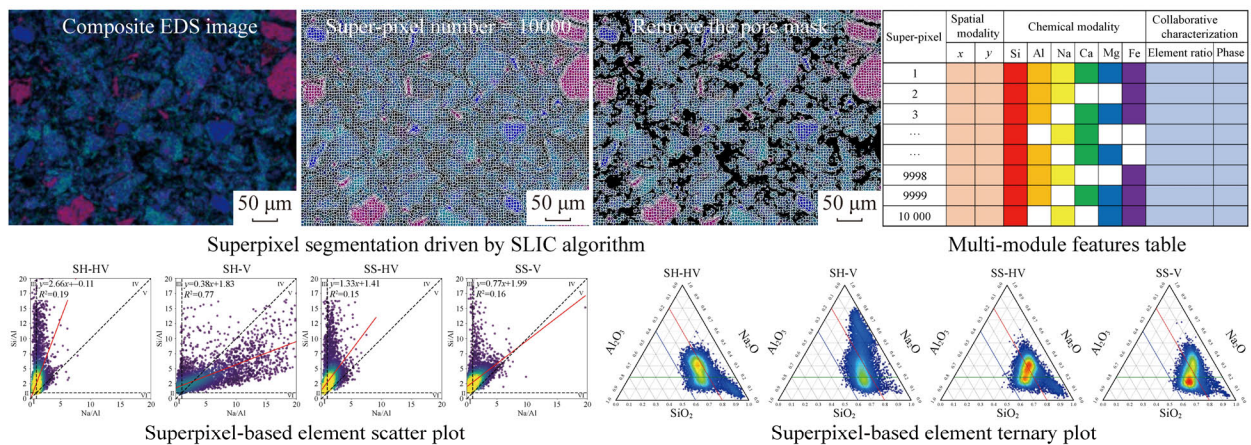
(a) Raw and denoised EDS mappings



(b) Composite EDS mapping and clustering result



(c) Phase mask extraction and area calculation



(d) Superpixel segmentation and elemental analysis

图 6 EDS 面扫描图定量分析框架
Fig. 6 Quantitative analysis framework of EDS mappings

应产物的生成, 其被划分为反应产物相。GMM 算法进一步将 SH-V 中的高 Na 区域划分为独立的富钠相。而在无 Na 富集特征的 SH-HV 中, GMM 算法重新根据化学特征自适应地捕捉规律, 界定出富硅相。AALRS-SS 的物相掩模包括 XJ-1 相、产物相、富镁相、孔隙相、富硅相与产物边缘相。在 AALRS-SS 中, GMM 算法未能区分活性与惰性 XJ-1 相, 而是将二者统一划为 XJ-1 相。此外, 由于 AALRS-SS 的孔隙直径与面积较 AALRS-SH 局部增大, 导致无元素信号区域的局部面积显著增加, 使产物相与孔隙相之间的元素信号强度差异更加明显, 因此 GMM 算法在两者之间进一步划分出产物边缘相。

图 9 展示了各物相掩模面积分数, 可近似代表物相占比, 并统计了各组物相占比均值和四分位距

(IQR)。由图 9a~图 9f 可见, 环境对物相占比影响显著。无论在 SH 体系还是 SS 体系中, HV 环境相较 V 环境均表现出更高产物相占比和更低孔隙率。例如, SH-HV 的产物相占比高达 31.35%, 较 SH-V 提高 7.87%。此外, AALRS-SS 与 AALRS-SH 展现出不同的物相分布特征。前者普遍具有更高的产物相占比和更低的 XJ-1 相占比及孔隙率, 说明 AALRS-SS 的整体反应程度更高, 结构更加致密, 而 AALRS-SH 的微观结构则相对疏松多孔。此外, SH-V 和 SH-HV 中存在 12.49% 和 6.03% 的惰性 XJ-1 相, 表明 XJ-1 活性较低, 一定程度上阻碍了 AALRS 的强度发展。从数据的离散度来看, HV 环境下各物相的 IQR 值普遍低于 V 环境, 说明前者在不同区域的物相分布波动更低, 表明 HV 环境有利于提高成分匀质性, 而 V 环境则会加剧成分异质性。

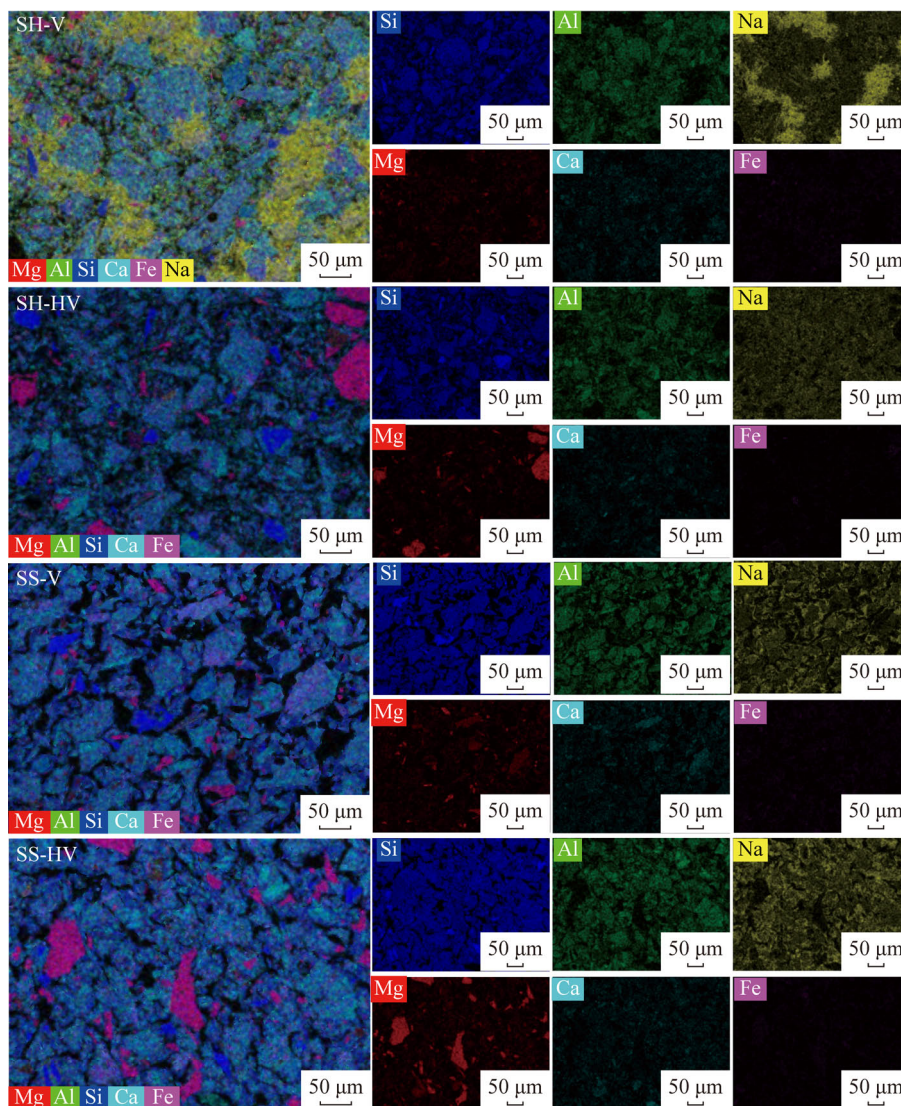


图 7 复合 EDS 面扫描图
Fig. 7 Composite EDS mappings

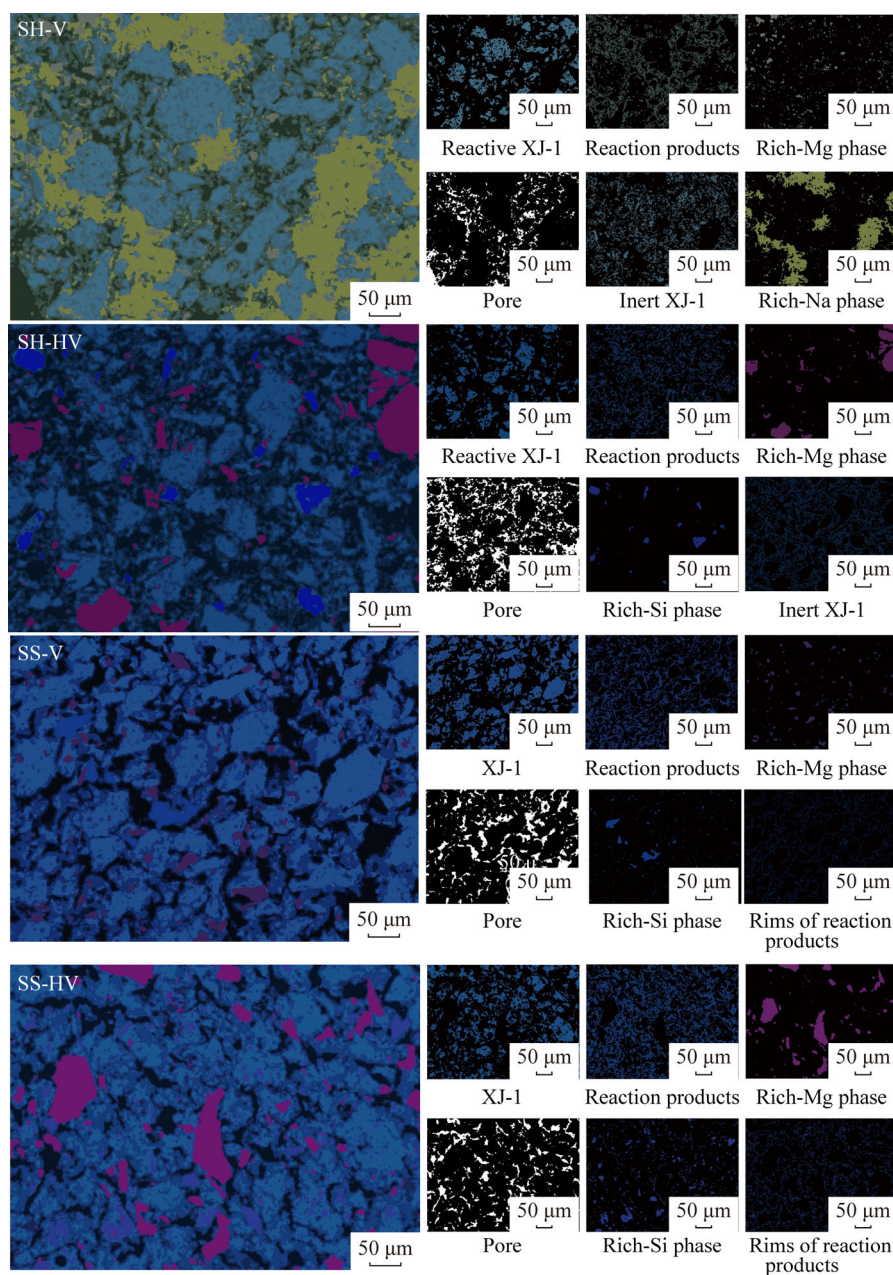


图 8 GMM 聚类结果及对应物相掩模

Fig. 8 GMM clustering results and corresponding phase masks

AALRS 本质上可视为由 XJ-1 相、产物相和孔隙相构成的复合材料，各试验组的物相掩模可分类归纳为上述三相：富硅、富镁、活性 XJ-1 相与惰性 XJ-1 相归入 XJ-1 相；产物边缘相归入产物相；富钠相与孔隙分布高度重合，归入孔隙相。由图 9g 可见，尽管 HV 环境与 SS 体系均可促进 XJ-1 相溶解和产物相生成，但激发体系对反应过程的调控作用更具决定性。SS 体系中产物相占比(面积)增量(21.0%)显著高于 HV 环境下的产物相占比增量(7.8%)，表明 SS 体系与 SH 体系的物相分布特征存在显著差异。

表 1 物相掩模与实际矿物成分之间的关系

Table 1 Relationship between phase masks and actual mineral compositions

Phase mask	Mineral composition
Reactive XJ-1	Albite (Plagioclase)
Inert XJ-1	Anorthite (Plagioclase), Ilmenite
Rich-Mg phase	Pyroxene/Forsterite
Rich-Na phase	Sodium hydroxide
Rich-Si phase	Quartz
Reaction products	Amorphous gel (Bulk)
XJ-1	Plagioclase
Rims of reaction products	Amorphous gel (Rim)

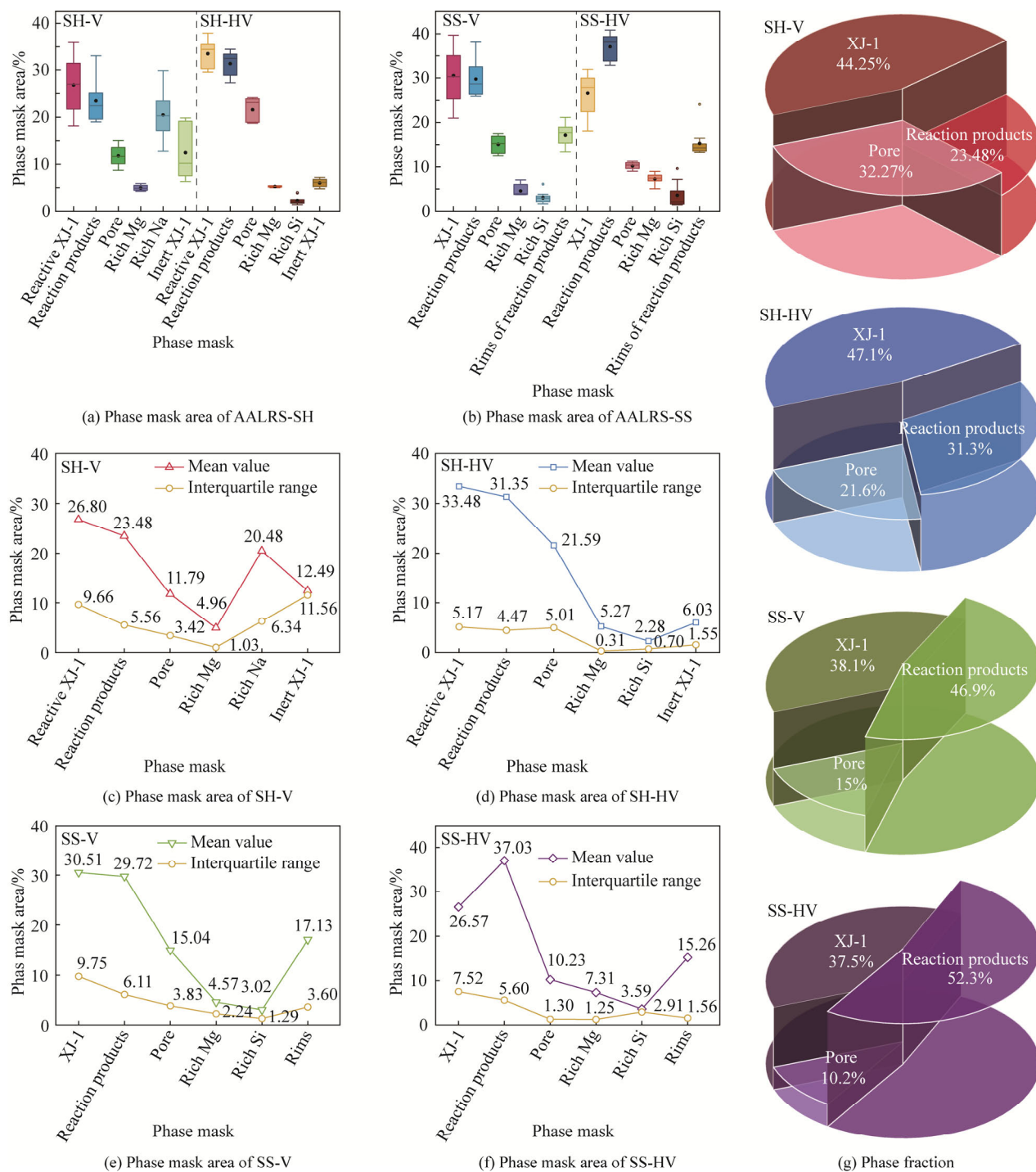


图9 物相掩模面积分数
Fig. 9 Area fraction of phase mask

为比较不同环境和激发体系下的元素组成差异,根据各元素在硅铝酸盐聚合网络中的作用及反应活性,将其划分为3类。分别为网络骨架元素(Si、Al)、网络修饰元素(Na)和惰性元素(Mg、Ca、Fe)。其中惰性元素统一记为R,即 $R = \text{Mg} + \text{Ca} + \text{Fe}$ 。图10展示了AALRS超像素特征元素比的散点分布情况。选取3组关键元素比,分别为 $\text{Si}/\text{Al}-\text{Na}/\text{Al}$ 、

$\text{R}/\text{Al}-\text{Si}/\text{Al}$ 和 $\text{Si}/\text{Al}-\text{Ca}/\text{Al}$ 。图10中横坐标值(X)与纵坐标值(Y)分别代表对应元素信号强度比值,利用 $X=1$ 、 $Y=1$ 及 $Y/X=1$ 的3条特征线将坐标平面划分为6个特征区域,并采用颜色梯度表示数据点的密度分布,颜色越深代表点分布越密集。同时,对每组数据进行线性回归拟合并计算决定系数(R^2),图10中红色实线为拟合特征线,拟合斜率在物理

意义上对应 2 组元素比值之比。由图 10a 可见, 除 SH-V 外, 其余各组数据点的分布趋势具有较高一致性, 而 AALRS-SS 的拟合线斜率显著低于 AALRS-SH, 这表明 AALRS-SS 物相的 Si/Na 显著降低, 其原因在于

产物相中起平衡电荷作用的 Na 元素含量显著增加。相比之下, SH-V 中数据点表现出极高的 Na/Al 且分布更加离散, 反映其中存在大量未参与凝胶网络构建的游离 Na。由图 10b 可见, HV 环境下的数据

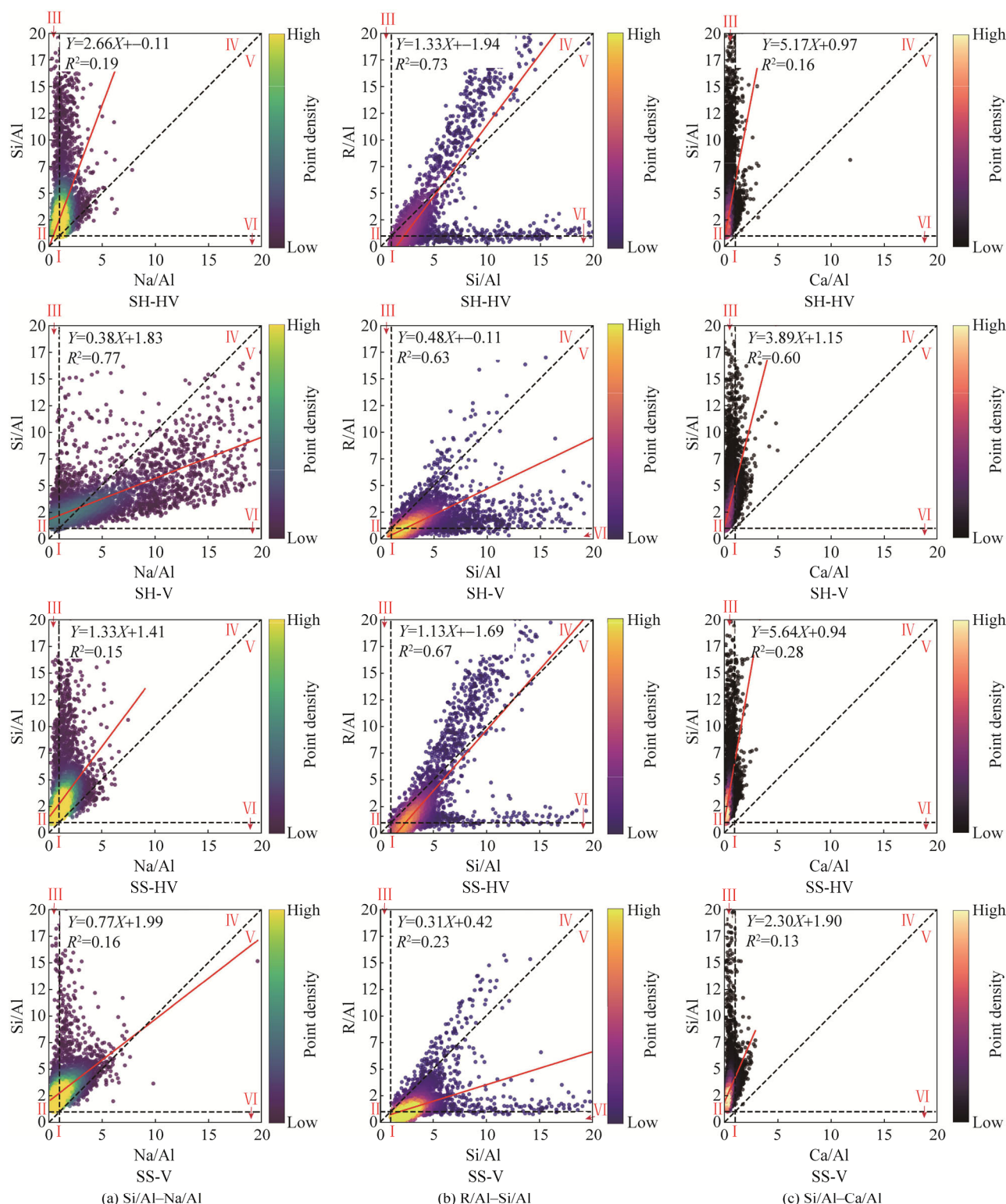


图 10 AALRS 超像素基特征元素比散点图

Fig. 10 AALRS Superpixel-based scatter plot of feature element ratios

点分布较 V 环境更为分散, 且拟合斜率更大。HV 环境下 R/Al 显著增大说明原始硅酸盐聚合网络中 Al 溶解过程更为强烈, 而在 V 环境下该过程受到明显抑制。由图 10c 可见, AALRS 体系中 Ca/Al 显著低于 Si/Al。更重要的是, 尽管 Si/Al 发生较大变化, Ca/Al 仍基本保持稳定, 数据点呈近乎垂直的簇状分布。这一特征表明, Ca 在反应过程中基本保持惰性, 未深入参与凝胶网络的重组缩聚。

图 11 展示了不同特征区域中超像素的数量占比。这些区域根据其独特的元素比例, 可近似对应于不同的代表性物相。由图 11a 所示, 超像素主要分布在区域 III、区域 IV 和区域 V, 区域 III 元素丰度关系为 $x(\text{Si}) > x(\text{Al}) > x(\text{Na})$ (x 为归一化后的 EDS 信号强度), 区域 IV 元素丰度关系为 $x(\text{Si}) > x(\text{Na}) > x(\text{Al})$ 。区域 V 元素丰度关系为 $x(\text{Na}) > x(\text{Si}) > x(\text{Al})$, Na 元素占据主导地位, 结合前述结果可知, 这对应氢氧化钠溶液中的游离钠。由图 11b 可见, 与真空环境相比, 热真空环境下区域 IV [$x(\text{R}) > x(\text{Si}) > x(\text{Al})$] 和区域 VI [$x(\text{Si}) > x(\text{Al}) > x(\text{R})$] 占比均增加, 而区域 V [$x(\text{Si}) > x(\text{R}) > x(\text{Al})$] 占比则明显降低。由

图 11c 可见, 绝大多数超像素(>90%)集中于区域 III [$x(\text{Si}) > x(\text{Al}) > x(\text{Ca})$], 仅有少量分布于区域 IV [$x(\text{Si}) > x(\text{Ca}) > x(\text{Al})$], 表明 Ca 在 AALRS 体系中含量极低。据此推断, AALRS 体系中形成的凝胶相主要为 N-A-S-H 凝胶, 而非 C-(A)-S-H 凝胶。

图 12 展示了 AALRS 超像素基元素三元相图, 包括 Si-Al-Ca 和 Si-Al-Na 体系, 三元素组分均已归一化至 100%。图中由蓝到红的颜色渐变表示点密度的逐渐升高。由图 12a 可见, 各组数据点分布整体相似。相较于 AALRS-SH, AALRS-SS 的数据点分布更为集中, 而 AALRS-SH 在高 Al 区域呈现出更高的分布密度。这表明后者含有更多未反应硅铝酸盐或富铝凝胶相。由图 12b 可见, AALRS-SH 和 AALRS-SS 的数据点分布呈现出显著差异。AALRS-SH 的数据点簇集中于中 Al 含量区域, 而 AALRS-SS 的数据点簇则在中 Al 和低 Al 含量区域形成双重分布中心。这表明 AALRS-SS 生成了更多高 Si、低 Al 的物相, 其凝胶产物逐渐由富铝凝胶向富硅凝胶转变。另外, SH-V 在高 Na 区域显示出广泛的分布, 与前述结果一致。

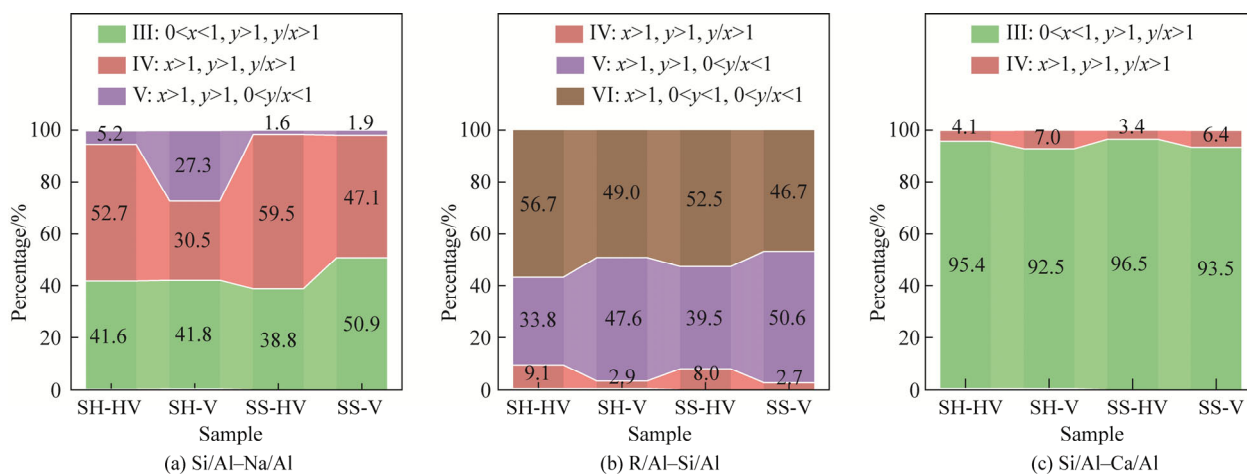


图 11 不同特征区域的超像素数量占比

Fig. 11 Proportion of the number of superpixels in different feature regions

2.3 AALRS 微观结构特征与反应机理

基于 EDS 图像定量分析结果, 图 13 展示了 AALRS 在不同环境和激发体系下的微观结构特征与反应机理。对于 AALRS-SH, 在 V 环境中, XJ-1 的溶解与硅(铝)单体的缩聚均受到环境抑制。由内外压力差驱动的表面脱水增加了碱溶液的迁移路径, 导致孔隙率升高并引起孔隙内部的 Na 富集(图 7)。在 HV 环境中, 尽管仍存在迁移效应, 但高温环境成为主导因素, 其蒸发速率和蒸发量远高于真空环境。高温加速碱对 XJ-1 结构的解聚, 释放大量可溶

性硅(铝)单体, 这些单体经扩散与缩聚生成 Al 含量较高的 N-A-S-H 凝胶, 并在 XJ-1 颗粒表面沉淀形成阻碍后续溶解的产物层。因此, AALRS-SH 的产物相占比较低、孔隙相占比较高, 形成相对多孔的微观结构, 是其强度偏低的主要原因(图 5)。对于 AALRS-SS, 其碱溶液中含有大量通过共氧连接形成的链状与环状的硅酸盐多聚体^[27]。在 HV 环境中, 硅酸盐多聚体发生解聚^[27], 使反应介质中成核位点密度显著增加。XJ-1 溶解产生的硅(铝)单体与成核位点缩聚, 生成 Si 含量较高

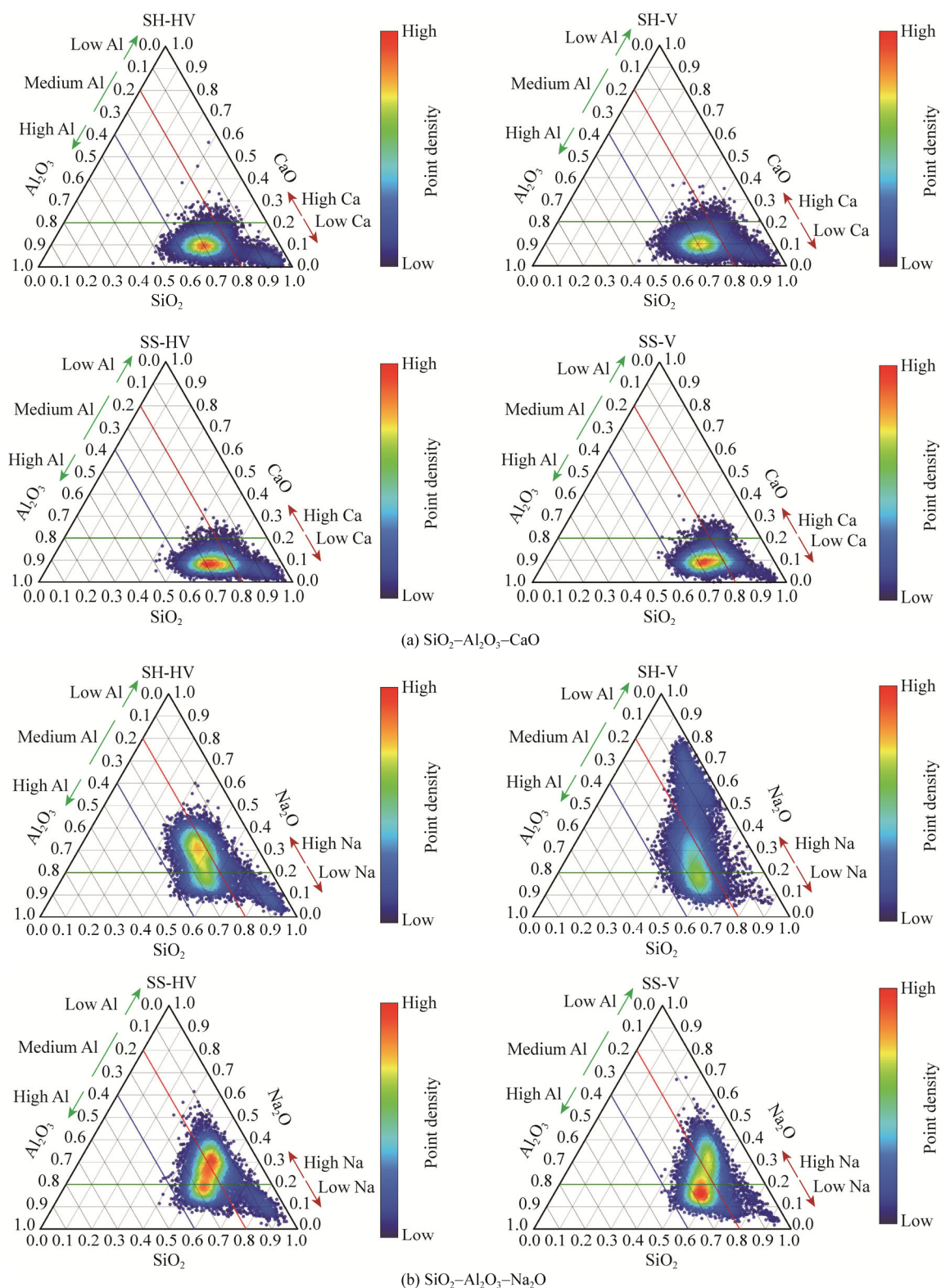


图 12 Si--Al--Ca 和 Si--Al--Na 超像素基元素三元图
Fig. 12 Superpixel-based element ternary plots of Si--Al--Ca and Si--Al--Na

的 N-A-S-H 凝胶, 从而显著提高产物相占比与反应面积(图 9)。整体上, AALRS-SS 呈局部致密、整体孔隙率较低但大直径孔比例较高的结构特征, 从而展现出最高的强度(图 5)。在 V 环境中, 尽管

体系成分异质性增强、产物相占比降低, 但持续性脱水会诱导硅酸盐多聚体通过化学键与氢键自组装并发生凝胶化^[28]。形成的硅酸盐团簇可有效黏结 XJ-1 颗粒, 且随龄期增长, 硅酸盐团簇数量与内部

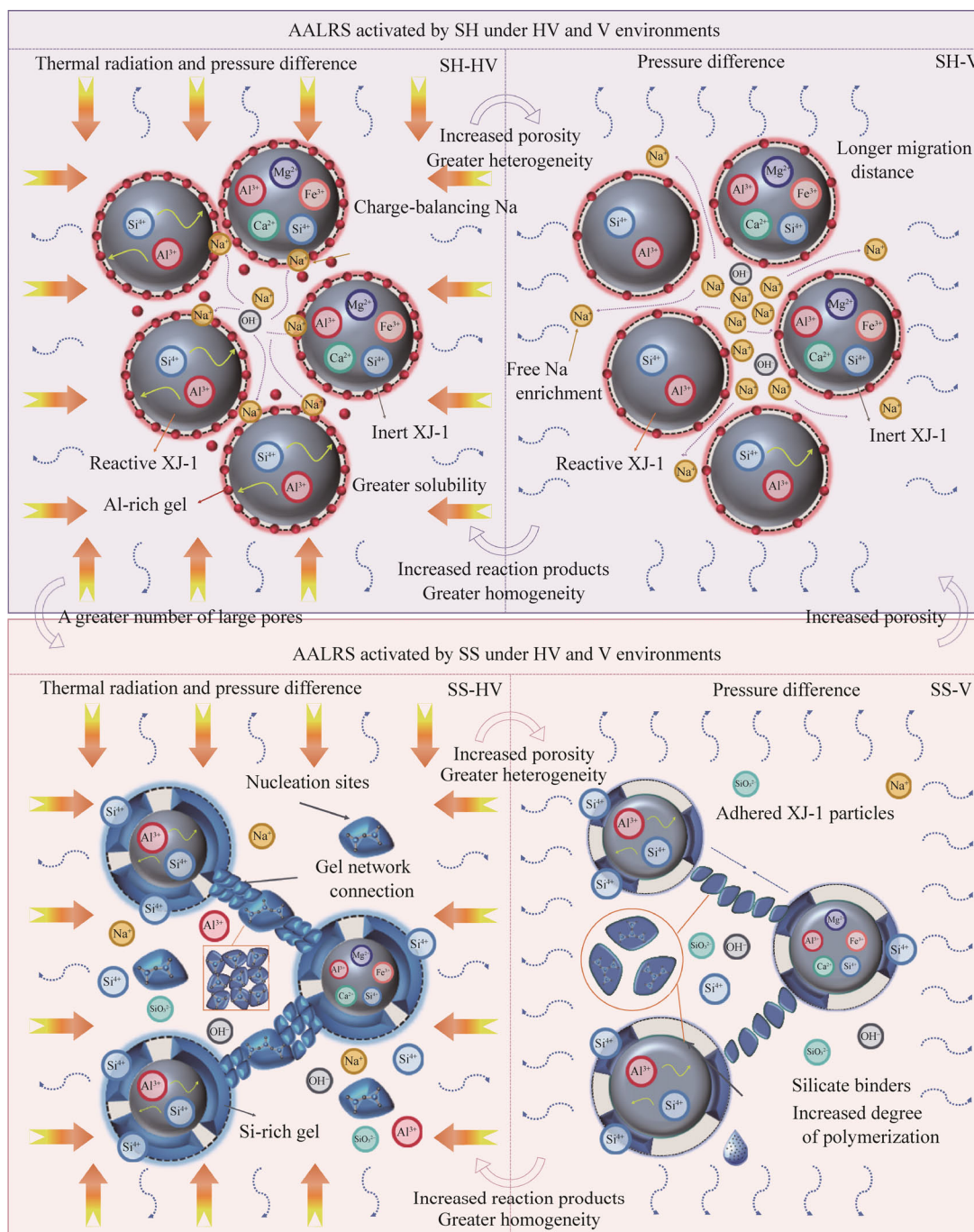


图 13 AALRS 微观结构特征与反应机理

Fig. 13 Microstructural characteristics and reaction mechanism of AALRS

聚合度持续提高, 因此 SS-V 的强度呈显著增长趋势(图 5)。类似现象也在其他研究中得到报道^[6]。综上, 极端环境与激发体系协同调控 AALRS 的反应路径、相组成、相空间分布与孔隙结构, 进而主导其力学性能的发展。

3 AALRS 强度预测模型

3.1 数据集的建立与分析

本工作整合大量已发表的文献数据^[3, 6-9, 11-13, 28-44],

根据文献中公开报道的材料组成、试验参数与性能指标, 构建了包含 375 组样本的数据集, 可通过在线公开仓库 <https://github.com/lyizhouyao/AALRS-yyz> 获取。模型的输入变量选取了影响强度发展的关键因素, 包括模拟月壤的中值粒径(MPD)、模拟月壤的硅铝比(Si/Al)、水胶比(W/B)、碱含量(AC)、硅酸盐模数(SM)、龄期(CA)、环境初始温度(IT)、环境终止温度(TT)与真空度(VD)。上述变量可归纳为 3 类: 模拟月壤性质、配合比参数与极端环境特

征。模型的输出变量为 AALRS 的抗压强度(CS)，各特征变量汇总于表 2，其统计指标汇总于表 3。为进行模型训练与性能验证，该数据集按照 7 : 3 的比例随机划分为训练集和测试集。测试集独立于模型训练过程，用于评估模型的泛化能力。另外，为消除特征量纲差异对模型的影响及加快模型收敛速度，训练前对数据进行归一化处理[式(1)]。

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

式中： X' 是归一化后的值； X 是给定观测值； $X_{\max}$ 和 $X_{\min}$ 分别为该特征的最大值和最小值。

为避免出现多重共线性问题，采用皮尔逊相关系数评估各特征间的线性相关性，并以绝对相关系数大于 0.7 作为存在强相关性的判断标准。图 14 展现了特征变量的分布与皮尔逊相关系数。其中，对角线为各特征的数据分布直方图；非对角线左侧为变量间的散点图以展现相关趋势；右侧标注变量间的皮尔逊相关系数及统计显著性水平。蓝色背景表示正相关，且系数绝对值越接近 1，线性相关性越强。结果显示，大多数变量间的

相关性较低，最大正相关系数为 0.311，最大负相关系数为-0.508。同时，各变量对的散点图也未显示出明显的线性关联趋势，图中“***”表示在极高置信水平下($p < 0.001$)，较低线性相关性依然具有统计显著性。综上，所选特征变量间不存在强线性相关性，可作为模型特征变量。

3.2 堆叠(Stacking)集成模型

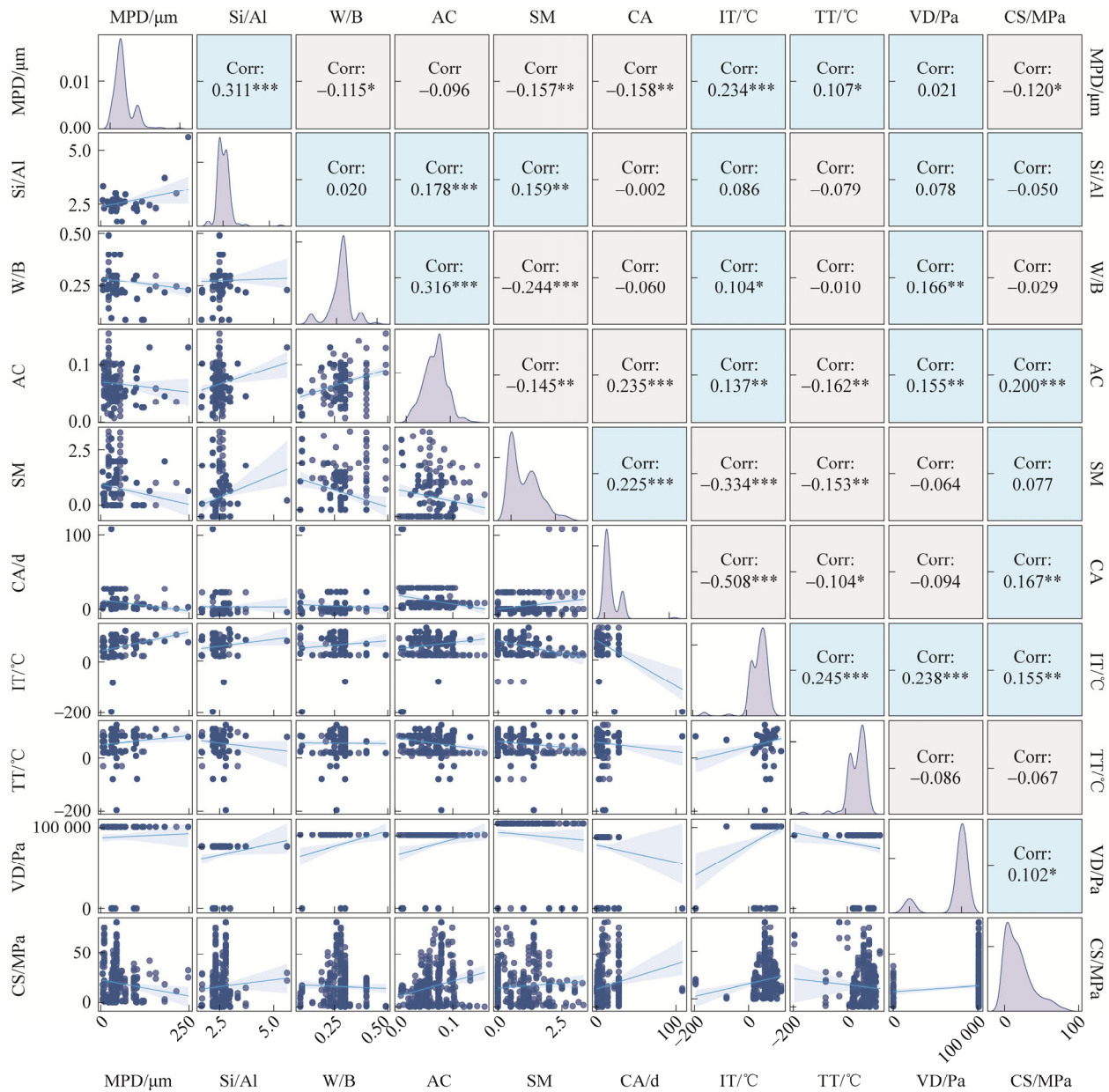
基础 ML 算法在探索复杂假设空间时，易陷入欠拟合或过拟合困境，从而导致预测性能受限。Stacking 集成模型通过构建双层学习器框架，融合多种异构模型的互补优势，从而有效提升泛化能力和预测精度，其结构如图 15 所示。第 1 层(Level-0)为基学习器层，包含 6 种异构模型：自适应提升(AdaBoost)、极端梯度提升(XGBoost)、分类梯度提升(CatBoost)、支持向量回归(SVR)、多层感知器神经网络(MLPNN)和随机森林(RF)。该组合涵盖 Boosting、Bagging、核方法及神经网络等多种学习机制，旨在从不同维度捕捉数据的复杂模式与非线性特征。同时，采用 5 折交叉验证机制为第二层提供可靠的输入特征。第 2 层(Level-1)选用线性回归

表 2 模型特征变量
Table 2 Feature variables of model

Categories	Feature variables	Abbreviation	Unit
Properties of Lunar regolith simulant	Median particle diameter	MPD	μm
	Si/Al ratio of LRS	Si/Al	
Mixture proportion parameter	Water-to-binder ratio	W/B	
	Alkali content	AC	
	Silicate modulus	SM	
Extreme environmental conditions	Initial temperature	IT	$^{\circ}\text{C}$
	Termination temperature	TT	$^{\circ}\text{C}$
	Vacuum degree	VD	Pa
	Curing age	CA	d
Mechanical properties	Compressive strength	CS	MPa

表 3 特征变量统计指标
Table 3 Statistic indicators of feature variables

Feature variables	Minimum	Maximum	Mean	Standard deviation	Coefficient of variation
MPD	5.760	247.450	48.650	35.440	0.730
Si/Al	1.670	5.610	2.530	0.370	0.140
W/B	0.090	0.490	0.280	0.070	0.250
AC	0.007	0.155	0.070	0.030	0.380
SM	0	3.340	0.760	0.810	1.070
CA	0.190	108.000	9.690	13.220	1.370
IT	-196.000	120.000	56.340	44.490	0.790
TT	-196.000	120.000	54.140	43.960	0.810
VD	10.000	101 325.000	87 823.000	34 468.000	0.390
CS	0.030	81.500	19.070	18.540	0.970



p indicates the probability of observing a sample correlation coefficient at least as extreme as the observed value, assuming the null hypothesis is true. * means $p < 0.05$, ** means $p < 0.01$, *** means $p < 0.001$.

图 14 特征变量的分布与皮尔逊相关系数

Fig. 14 Distribution of feature variables and Pearson correlation coefficient

作为元学习器。其核心任务是基于基学习器的预测结果进行学习，从而自动确定各基学习器的最优加权系数并输出最终预测值。

3.3 模型评估指标

采用决定系数(R^2)、相关系数(R)、均方根误差(RMSE)、平均绝对误差(MAE)和标准差(SD)评估模型预测性能。计算方法如下：

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_a - X_p)^2}{\sum_{i=1}^n (\bar{X}_a - X_a)^2} \quad (2)$$

$$R = \frac{\sum_{i=1}^n (X_a - \bar{X}_a)(X_p - \bar{X}_p)}{\sqrt{\left[\sum_{i=1}^n (X_a - \bar{X}_a)^2 \right] \left[\sum_{i=1}^n (X_p - \bar{X}_p)^2 \right]}} \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (X_a - X_p)^2}{n}} \quad (4)$$

$$\text{MAE} = \frac{\sum_{i=1}^n |X_a - X_p|}{n} \quad (5)$$

$$SD = \sqrt{\frac{\sum_{i=1}^n (X_a - \bar{X}_a)^2}{n}} \quad (6)$$

式中： n 代表样本总数； X_a 代表实际值； X_p 代表预测值； $\bar{X}_a$ 代表实际值的平均值， $\bar{X}_p$ 代表预测值的平均值。

3.4 超参数优化算法

超参数优化旨在寻找最优超参数配置，以提升模型的训练效果与泛化能力。本工作采用 6 种代表

性的超参数优化算法：随机搜索(RS)、贝叶斯优化(BO)、树结构帕森估计(TPE)、粒子群优化(PSO)、模拟退火(SA)和灰狼优化(GWO)。这些算法基于各自独特的搜索机制与优化逻辑，在 Stacking 模型的超参数空间进行寻优，以最大化模型预测精度。具体而言，RS 通过在超参数空间中随机采样，快速搜索高维参数组合^[45]。BO 基于概率模型与期望改进筛选高性能超参数^[46]。TPE 利用树状分层划分超参数空间，以非参数密度估计识别高性能区域^[47]。图 16

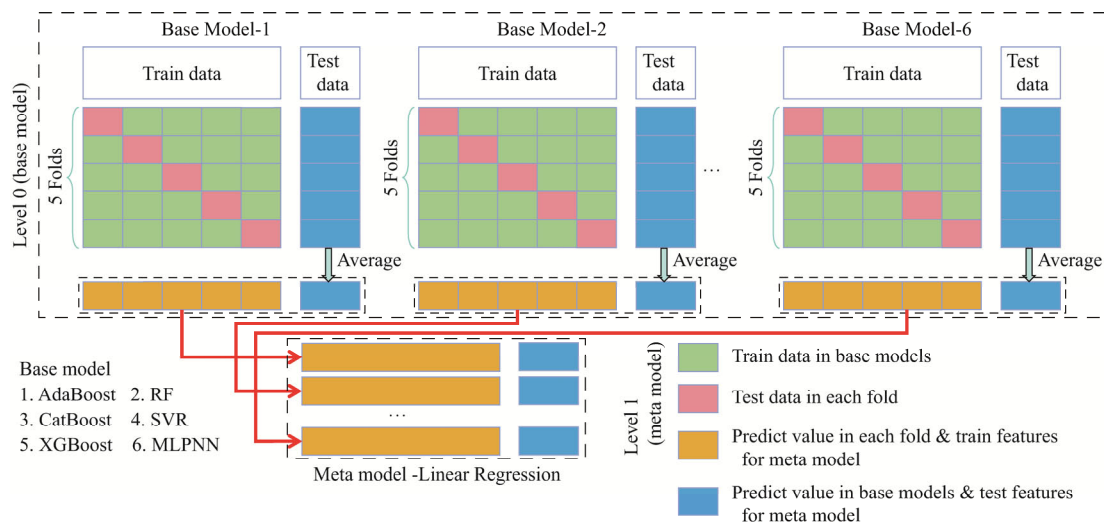


图 15 Stacking 集成模型框架

Fig. 15 Framework of stacking ensemble model

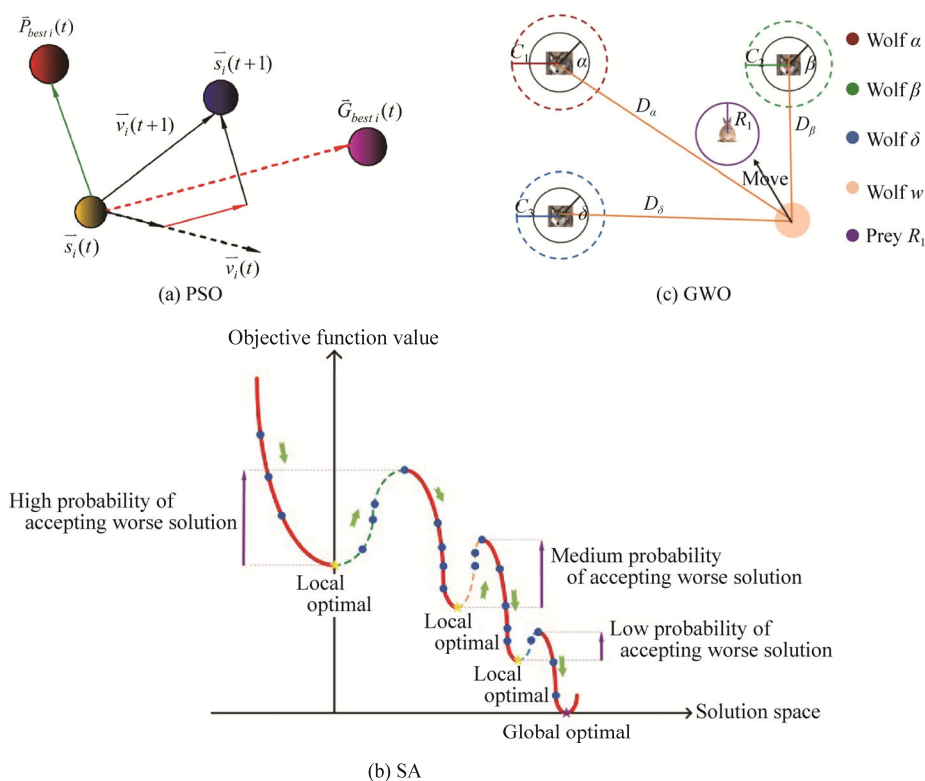


图 16 启发式优化算法示意图

Fig. 16 Schematic illustration of heuristic optimization algorithms

展示了启发式优化算法示意图。PSO 借鉴群体协作机制，通过个体与全局最优解的交互来迭代寻优^[48]。SA 借鉴固体退火过程，通过温度衰减函数与 Metropolis 准则，以随机扰动和概率接受机制规避局部最优^[49]。GWO 模拟狼群的等级狩猎机制，通过精英策略引导搜索方向，模拟包围、捕猎过程以实现高效收敛^[50]。6 种优化算法在相同的超参数搜索范围内，对 Stacking 模型的基学习器与元学习器执行最优超参数搜索。超参数搜索空间见表 4。不同优化算法获得的最佳超参数组合见表 5。

3.5 模型预测性能评估

图 17~图 20 展示了不同超参数优化算法作用下的 Stacking 模型的预测结果与性能指标。结果表

明，所有 Stacking 集成模型均展现出良好的预测能力，其中 Stacking-RS 模型在所有评估维度中表现最为出色和稳定。首先，其在测试集上的预测值与真实值的统计指标接近，差异无统计学显著性(图 17)。其次，由图 18 可见，其在训练集的 Pearson r 为 0.995，接近于 1，表明预测值与实际值的正线性相关性极高。在测试集上，Pearson r 同样有 0.949，为所有模型最高值。詹森-香农散度(JSD)用于量化预测值与实际值的分布差异，数值越小代表分布越相近。其 JSD 值为 0.214，仅次于 Stacking-SA 和 Stacking-GWO。误差分布图显示其标准差(SD=5.52)最小，异常点最少(图 19)，体现了模型预测的高度稳定性。这些优势最终反映在性能指标上，如

表 4 超参数搜索空间
Table 4 Hyperparameter search space

Models	Hyperparameter	Search space
RF	n estimators	Random integer from 50 to 200
	Max depth	Random integer from 4 to 8
SVR	C	Uniform distribution from 1 to 50
	Epsilon	Uniform distribution from 0.05 to 0.3
MLPNN	Hidden layer sizes	Choice of (30,), (30, 20), or (50,)
	Alpha	Uniform distribution from 0.001 to 0.01
XGBoost	n estimators	Random integer from 50 to 200
	Max depth	Random integer from 4 to 6
	Learning rate	Uniform distribution from 0.05 to 0.15
CatBoost	Iterations	Random integer from 50 to 200
	Depth	Random integer from 4 to 6
	Learning rate	Uniform distribution from 0.05 to 0.15
AdaBoost	n estimators	Random integer from 50 to 200
	Learning rate	Uniform distribution from 0.05 to 0.15
Linear Regression	Final estimator fit intercept	Choice of [True, False]

表 5 最优超参数组合
Table 5 Optimal hyperparameter configurations

Hyperparameter	RS	BO	TPE	PSO	SA	GWO
RF: n estimators	107	108	75	75	50	70
RF: max depth	5	8	6	4	8	5
SVR: C	34.3	39.8	10.3	10.4	7.8	14.9
SVR: ϵ	0.251	0.182	0.288	0.136	0.300	0.252
MLPNN: hidden layer sizes	(30, 20)	(30, 20)	(30, 20)	(30, 20)	(50,)	(30, 20)
MLPNN: α	0.007 13	0.004 94	0.004 17	0.008 72	0.001 00	0.008 26
XGBoost: n estimators	178	61	146	187	200	200
XGBoost: max depth	5	6	5	5	4	4
XGBoost: learning rate	0.081 6	0.107 0	0.101 0	0.119 0	0.050 0	0.102 0
CatBoost: iterations	182	132	161	164	199	200
CatBoost: depth	5	5	6	6	6	6
CatBoost: learning rate	0.150	0.092	0.130	0.129	0.150	0.127
AdaBoost: n estimators	67	157	178	160	200	200
AdaBoost: learning rate	0.142	0.105	0.112	0.145	0.150	0.060
Linear Regression: fit intercept	True	True	True	True	False	True

图 20 所示。Stacking-RS 在测试集上的统计指标为： $R^2=0.896$ 、RMSE=5.58 MPa 和 MAE=4.07 MPa。该模型取得了最高的 R^2 及最低的 RMSE 和 MAE。这些结果表明 RS 算法在寻找最优超参数组合上的优势。此外，Stacking-GWO 和 Stacking-PSO 同样展

现出具有竞争力的性能。一个值得注意的共性趋势是，所有 Stacking 模型的大误差样本主要出现在预测强度约 15~40 MPa 的区间，这表明模型对该强度范围的样本预测难度相对更大。总之，Stacking-RS 在预测精度、稳定性和泛化能力上具有明显优势。

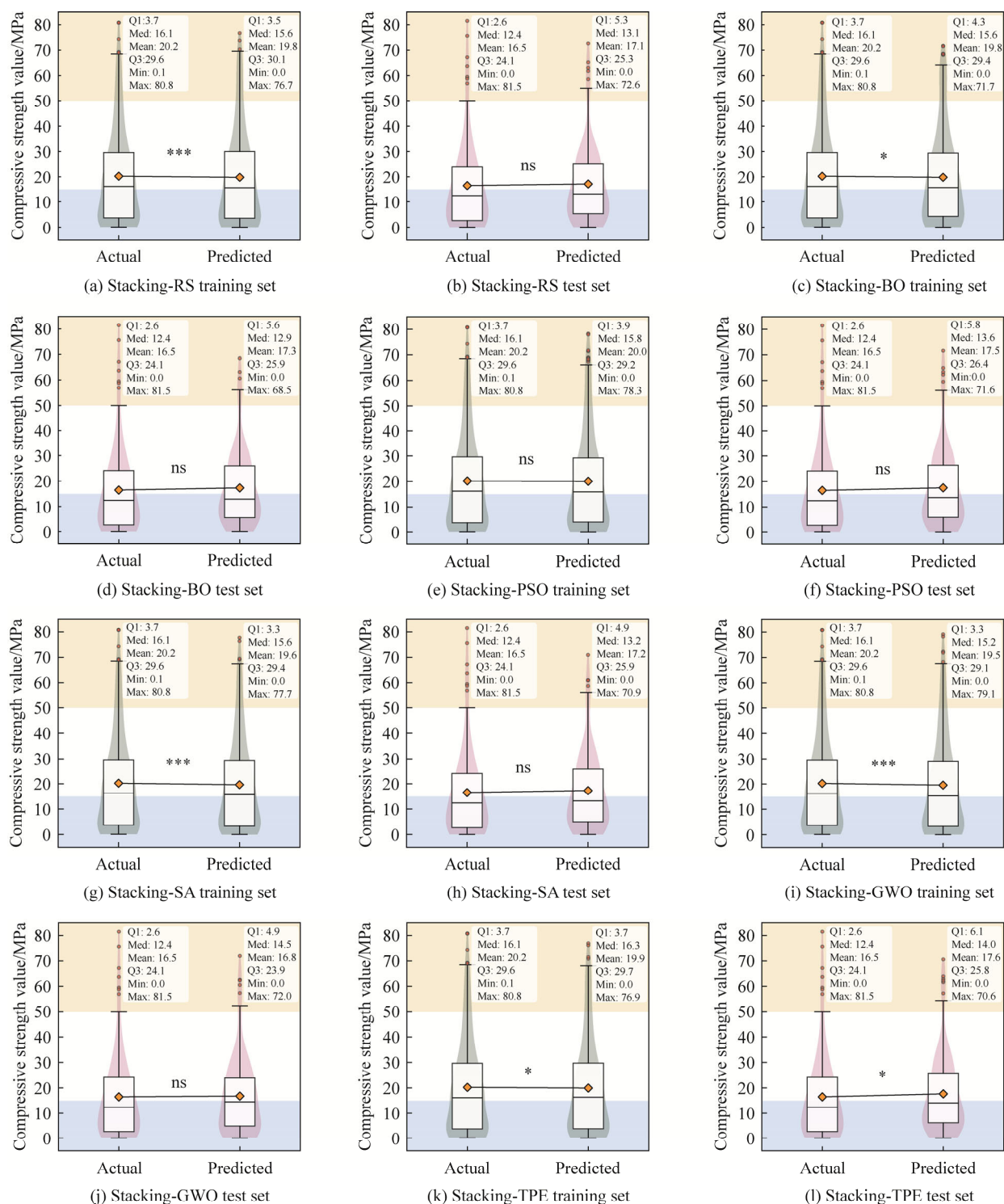


图 17 不同 Stacking 模型实测值与预测值的箱线图对比

Fig. 17 Boxplot comparison of actual and predicted values for different Stacking models

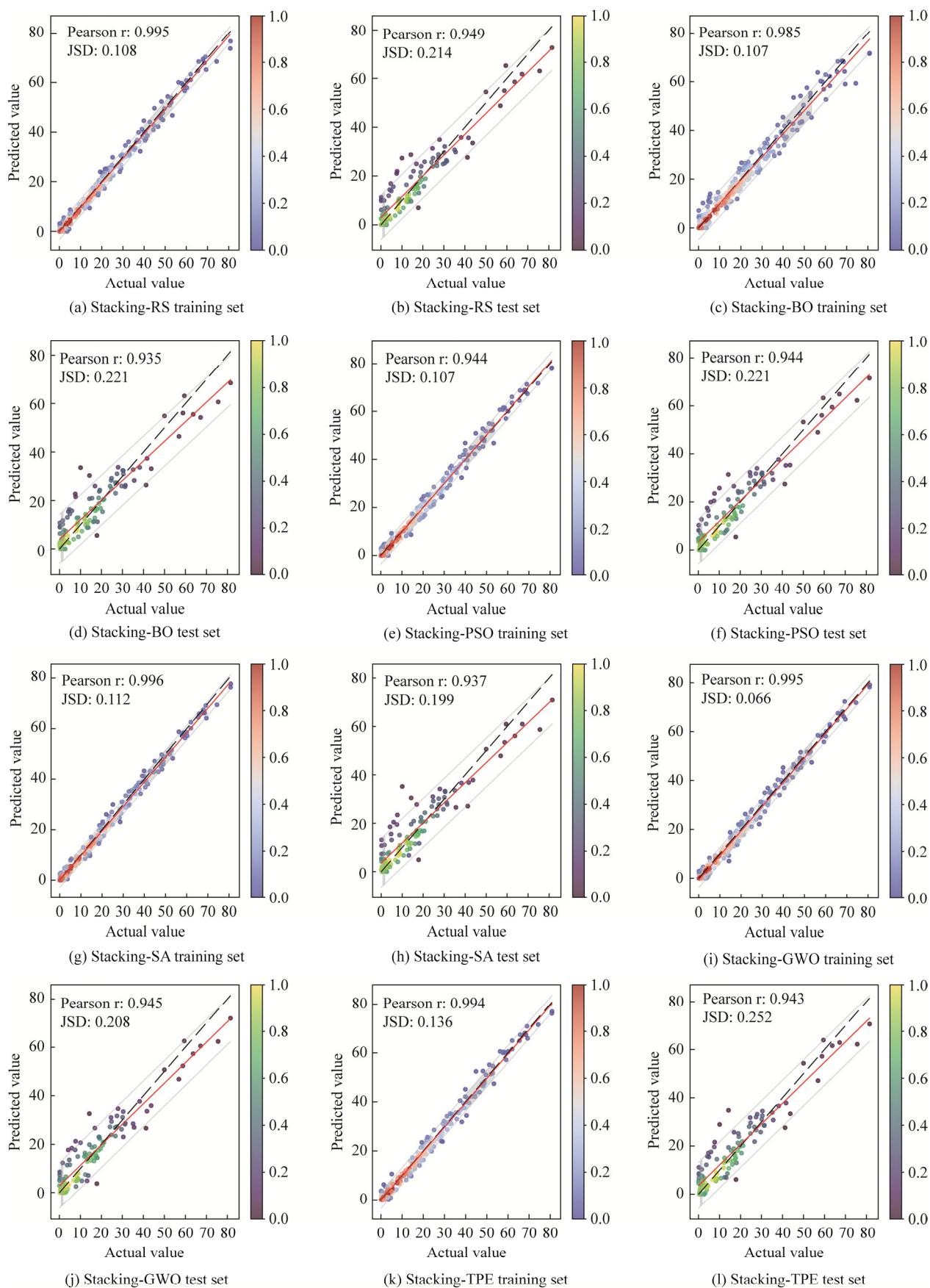


图 18 不同 Stacking 模型实测值与预测值的散点图对比

Fig. 18 Scatter plot comparison of actual and predicted values for different Stacking models

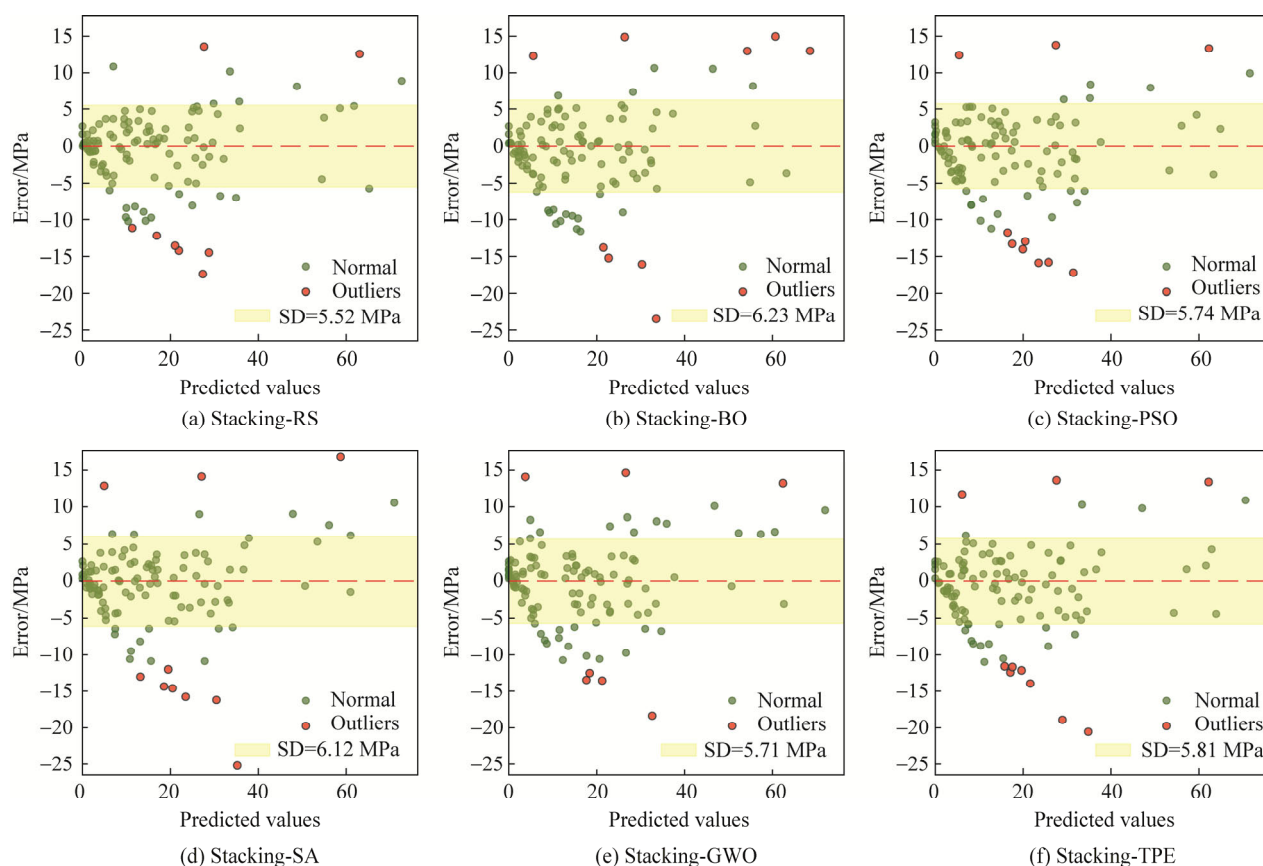


图 19 不同 Stacking 模型预测误差分布对比

Fig. 19 Comparison of prediction error distributions for different Stacking models

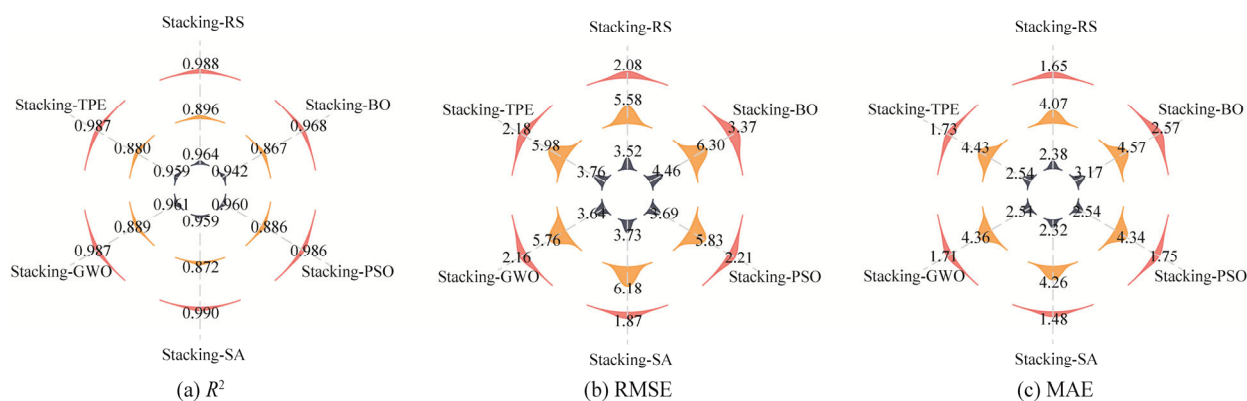


图 20 不同 Stacking 模型统计评价指标对比

Fig. 20 Comparison of statistical evaluation metrics for different stacking models

基于 R 、SD、RMSE 和时间 t 指标构建泰勒图, 比较不同 Stacking 模型在测试集上的预测性能, 结果如图 21 所示。观测点的 SD、RMSE 和 R 分别为 17.35、0 和 1。结果表明, Stacking-RS 表现最优, 其 R 最高 (0.9482), RMSE 最低 (5.58 MPa), SD 与观测点高度接近 (15.60 MPa), 且计算时间最短 ($t=0.77$ min), 体现其兼具高预测精度、优异泛化能力与计算效率。Stacking-GWO 性能次之 ($R=0.9439$, RMSE=5.76 MPa,

SD=15.63 MPa), 但计算时间最长 (73.18 min), 这会制约其在实际工程场景中的应用效率。Stacking-PSO 存在类似局限, 性能较好但计算时间较长 ($t=68.06$ min)。Stacking-TPE 的 SD 与观测点最接近 (15.79 MPa), 但预测精度略低于前述模型。Stacking-BO 预测性能最低, 表明 BO 优化算法在当前任务中的适用性存在局限。综合评估结果表明, Stacking-RS 以最优预测精度、泛化能力和计算效率成为整体性能最佳的模型。

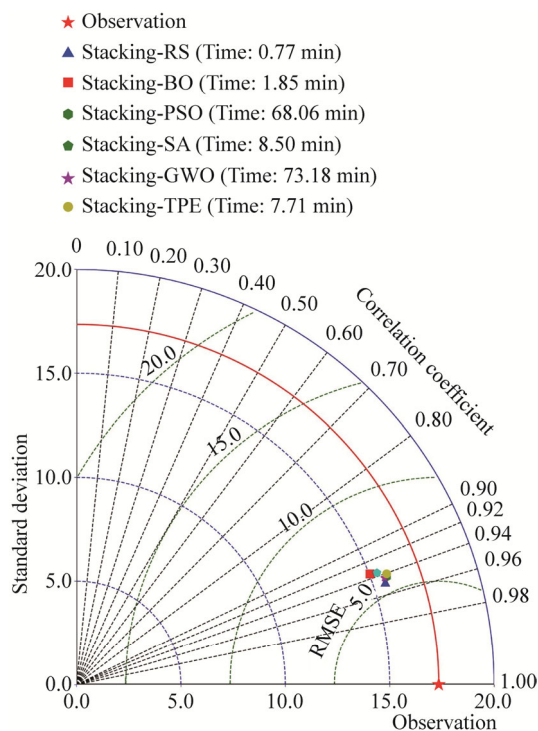


图 21 模型性能指标泰勒图

Fig. 21 Taylor diagram of model performance indicators

3.6 模型可解释性分析

尽管 Stacking-RS 展现了优异的预测精度，但其固有的“黑箱”特性限制了对模型内部决策机制的深刻理解，这成为揭示 AALRS 性能演化机制的关键难点。为突破这一局限，本工作引入基于博弈论的 SHAP 分析，并与 Stacking-RS 模型相结合，使其由“黑箱”工具升维为可解释性模型，从而为极端环境下 AALRS 材料设计与性能调控提供直接的科学指导与决策支持。

图 22 通过 SHAP 概要图与特征重要性排序，展示了模型的全局可解释性。结果表明，在 Stacking 的元学习器中，CatBoost 是贡献最大的基学习器。通过分析排序较高的基学习器的 SHAP 值分布与特征重要性排序，AC、CA 与 IT 被识别为正向贡献最显著的特征。AC 通过提高 XJ-1 的溶解程度以形成更多的凝胶产物^[38]，进而提高基体强度。IT 则通过加速 XJ-1 中 Si—O 和 Al—O 键的断裂及硅铝单体的缩聚以提高基体强度^[6]。而 MPD 的增大会产生显著的负向影响，其原因在于粒径增大导致比表面积降低，不利于碱溶液溶解模拟月壤，并易在其表面形成产物层，从而阻止模拟月壤进一步溶解。W/B 展现了非单调影响特征，其 SHAP 值随特征值增加先升后降。Si/Al 与 SM 的 SHAP 值则呈现无规

律波动。总体来看，较低的 Si/Al 和较高的 SM 通常有利于强度发展，但其参数达到极值反而对基体强度产生不利影响。

图 23 展示了 9 个特征的 SHAP 依赖图。散点颜色梯度变化反映了特征值的高低。局部加权平滑趋势线(LOWESS)反映了 SHAP 值随特征值的变化趋势。同一特征值下 SHAP 值的垂直离散分布反映了该特征与其他特征的交互效应。分析结果表明，AC 与 CA 的影响呈单调正相关性，其 SHAP 值随特征增加而持续上升。与之相反，MPD 的影响以 50 μm 为临界值，超过该值后其负面影响愈发显著。W/B 的影响具有非单调性，其正向贡献被限制在 0.20~0.28 区间内。IT 与 TT 表现出相似的趋势，均在 60 $^{\circ}\text{C}$ 附近实现由负转正的影响，但在 120 $^{\circ}\text{C}$ 附近略有回落，这与高温引发的微裂缝有关。SM 的 SHAP 值分布最为宽泛，表明其对预测值的贡献显著受其他特征交互作用影响。VD 在低特征值条件下对强度发展普遍产生正向影响，但同样受其他特征交互作用影响。

图 24 展示了 3 个典型示例样本的 SHAP 力图，旨在剖析 Stacking-RS 模型对高、中、低强度样本的具体预测逻辑。基准值为 17.14 MPa，代表模型对所有样本的平均预测值。分析以基准值为起点，量化了各输入特征对单次预测的贡献。对于高强度样本，其强度预测值为 48.79 MPa，AC(0.08)、CA(28)、Si/Al(2.79)、IT(60)、MPD(45) 和 SM(1.2)均表现为正向贡献，仅有 W/B(0.3)产生一定的负向影响，其 SHAP 值合计为 31.65 MPa。相反，对于预测值为 16.89 MPa 的低强度样本，尽管 AC、CA 等配合比特征以及 MPD、Si/Al 材料特征提供正向贡献，但 IT(20)、TT(20)环境特征产生显著负向影响，成为限制 AALRS 强度发展的关键因素。另外，中等强度样本的强度预测值为 31.34 MPa，其预测逻辑展现了模型对特征间复杂关系的捕捉能力。尽管较高的 AC(0.13)等特征提供了显著的正向贡献，但 XJ-1 过大的 MPD(247.45)和 Si/Al(5.61)产生了明显负向抵消作用，使预测值低于高强度样本。这些案例表明，模型的预测并非基于单一特征的简单叠加，而是源于对各参数边际贡献的复杂整合。Stacking-RS 模型的决策逻辑可为解析 AALRS 强度发展规律和影响机制提供见解。

3.7 强度预测图形用户界面

为提高 Stacking-RS 模型的可访问性和实用性并推动其工程化应用，本工作开发了基于 Stacking-RS 模

型的图形用户界面(GUI), 为非专业用户提供直观且便捷的 AALRS 强度预测平台。GUI 对输入参数设置范围约束, 其中 MPD、Si/Al、W/B、AC、SM、CA、VD 等参数为非负, IT 和 TT 允许输入负值, 且 VD 仅能输入离散值 10、100 和 101 325。如图 25 所示, 用户仅需

在交互面板输入特征参数, 点击预测按钮即可实时获取强度预测值。选取文献数据^[12]进行案例验证, 结果显示, 对于 32.40 MPa 的实际强度, 模型的预测值为 30.97 MPa, 相对误差仅为 4.41%。相关代码可通过 <https://github.com/lyizhouyao/AALRS-yyz> 获取。

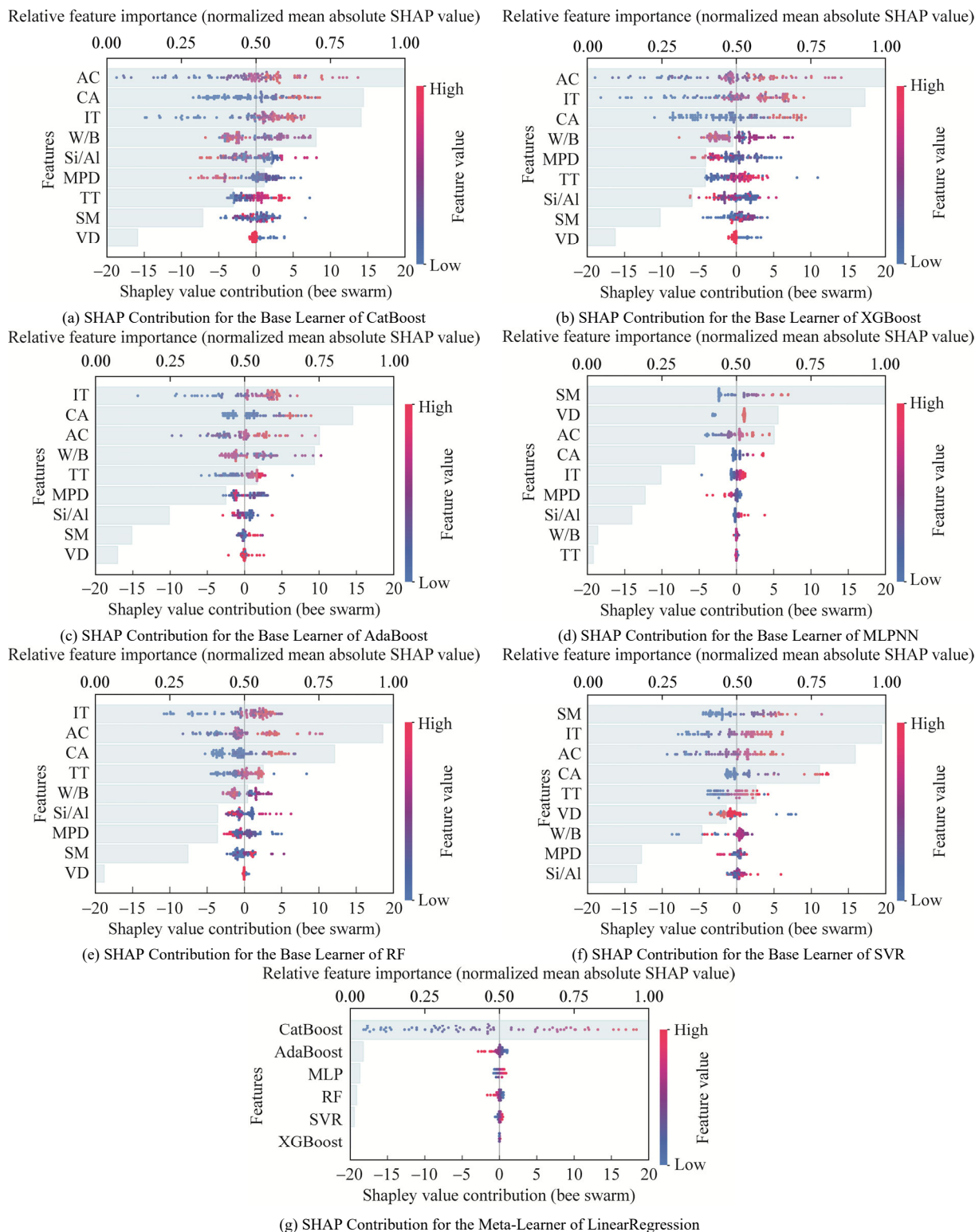


图 22 SHAP 值分布与重要性排序
Fig. 22 SHAP value distribution and importance ranking

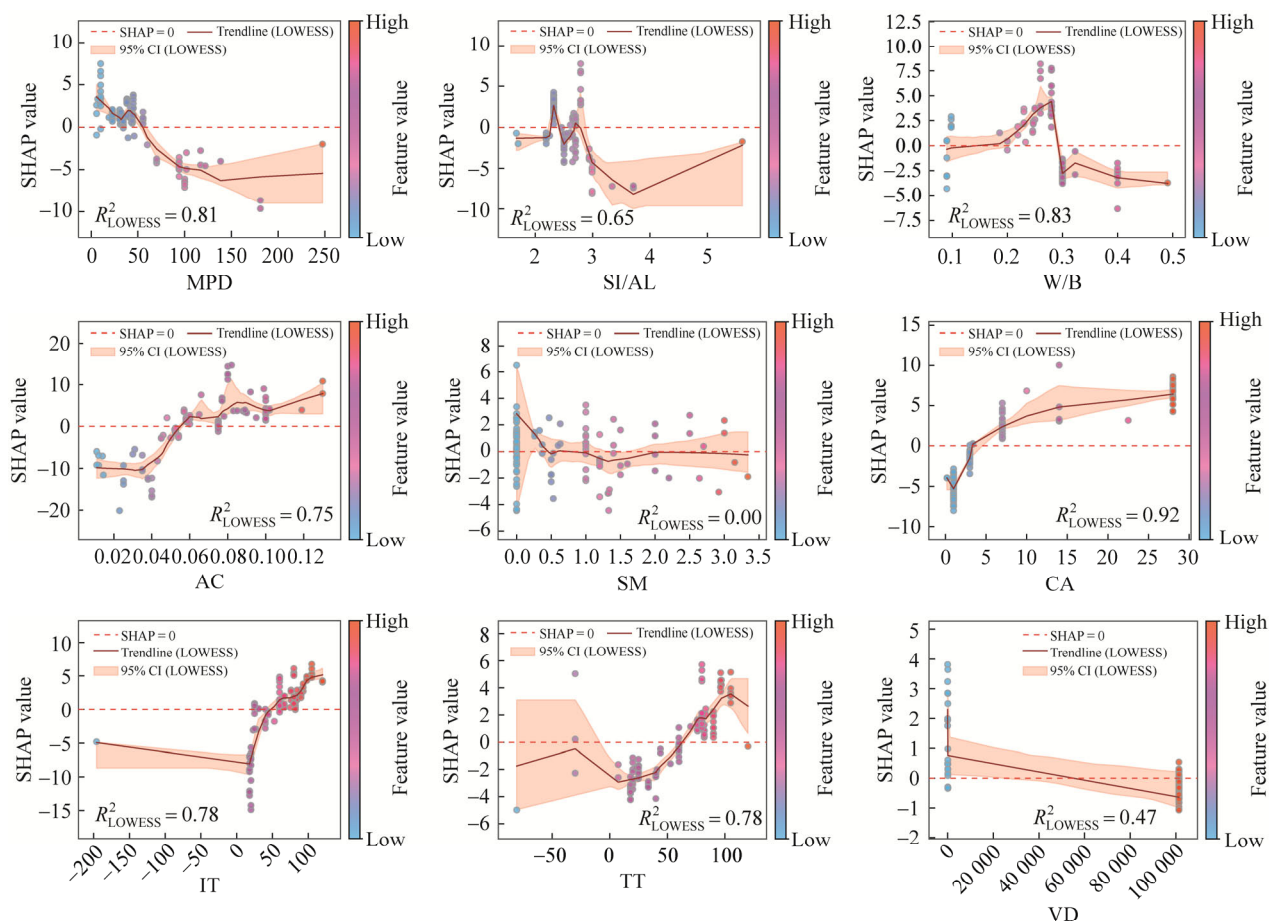


图 23 SHAP 依赖性分析
Fig. 23 SHAP dependence analysis

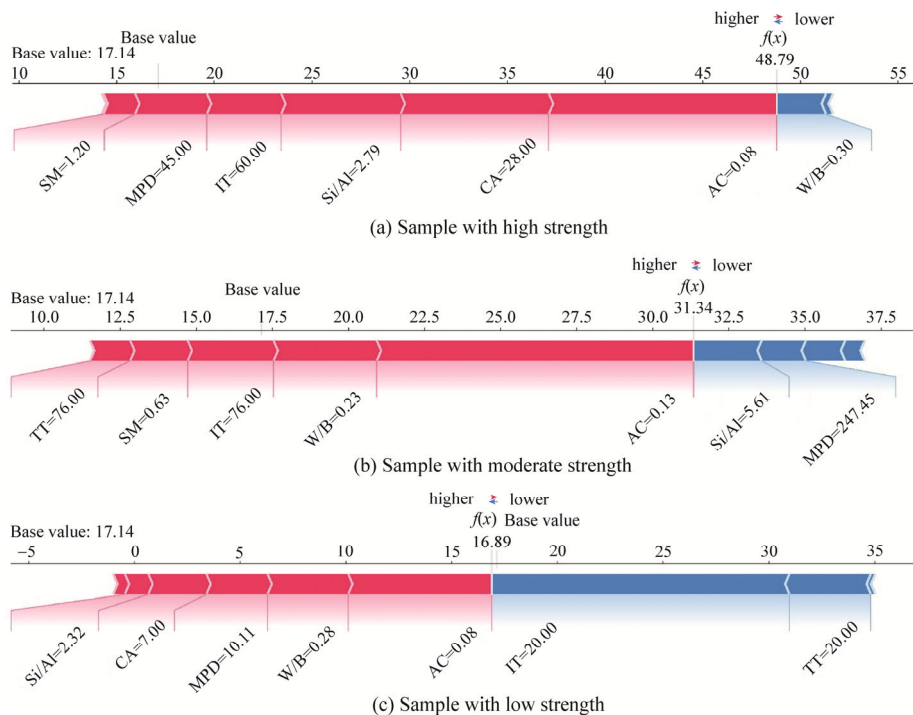


图 24 SHAP 局部解释性分析
Fig. 24 Local interpretability analysis of SHAP

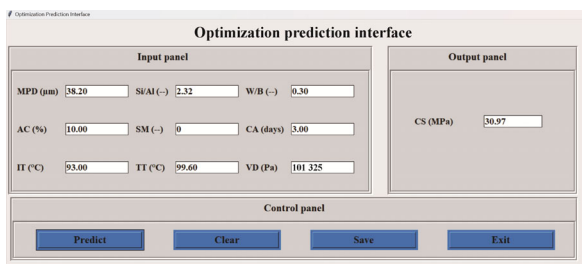


图 25 用于 AALRS 强度预测的图形用户界面设计

Fig. 25 Graphical user interface design for predicting strength of AALRS

4 “算法解构-性能预测-机制解析”

范式的协同分析机制

本工作构建了基于数据驱动的 AALRS 跨尺度智能分析框架，提出了“算法解构-性能预测-机制解析”的研究范式，其协同分析机制如图 26 所示。其中包含 2 个核心模块：宏观性能预测模块(第 3 章)利用 Stacking-RS 模型实现快速“性能预测”，并结合 SHAP 可解释性分析从宏观层面识别影响强度的关键因素及其贡献；微观结构解析模块(第 2 章)通过 GMM 和 SLIC 聚类算法驱动的 EDS 面谱定量分析框架，实现了微观结构特征的“算法解构”与定量表征。此模块并非预测模型的数据输入，而是为其提供了至关重要的可解释性参考和物理机制支撑，两者通过“机制解析”环节实现协同分析。例如，SHAP 分析在统计层面揭示特征 IT 对

强度预测有显著正向影响，而 EDS 定量分析则在物理层面验证并阐释其作用机制。即提高 IT 能显著增加 N-A-S-H 凝胶相的占比与成分匀质性，从而提升强度，两者结论相互印证。综上所述，宏观预测模块用于强度预测与识别关键因素，而微观结构模块提供直接的试验证据以阐明其背后的强度发展机制。在实际应用中，可首先通过宏观预测模块筛选关键材料、工艺与环境因素，再针对性地设计试验，借助微观结构模块以验证影响规律并深化机理认知，从而为 AALRS 在极端环境下性能的准确表征、深入理解和可靠预测提供技术支撑。该范式可进一步拓展至逆向优化设计，即通过预设月基工程的目标性能需求，反向优化月壤性质与工况配比，并结合地外环境条件的原位利用，有望支持月壤基材料的“材料成本最低、输入能耗最小、环境适配最高、宏观性能最优”的多目标协同设计。

本工作仍存在若干局限：1) 试验采用的 100 Pa 环境属于实验室真空环境，与月表高真空环境仍有显著差异，且辐射、小重力及月震等极端环境未被同时复现；2) 微观结构分析基于二维 EDS 面扫描元素分布图，其面积分数和相对元素信号不能完全等同于三维物相体积分数和绝对化学组成；3) 强度模型基于不同文献汇集的数据建立，可能受到材料来源、试验标准和报告方式差异的影响，后续需基于更接近月表条件的试验结果进一步验证。

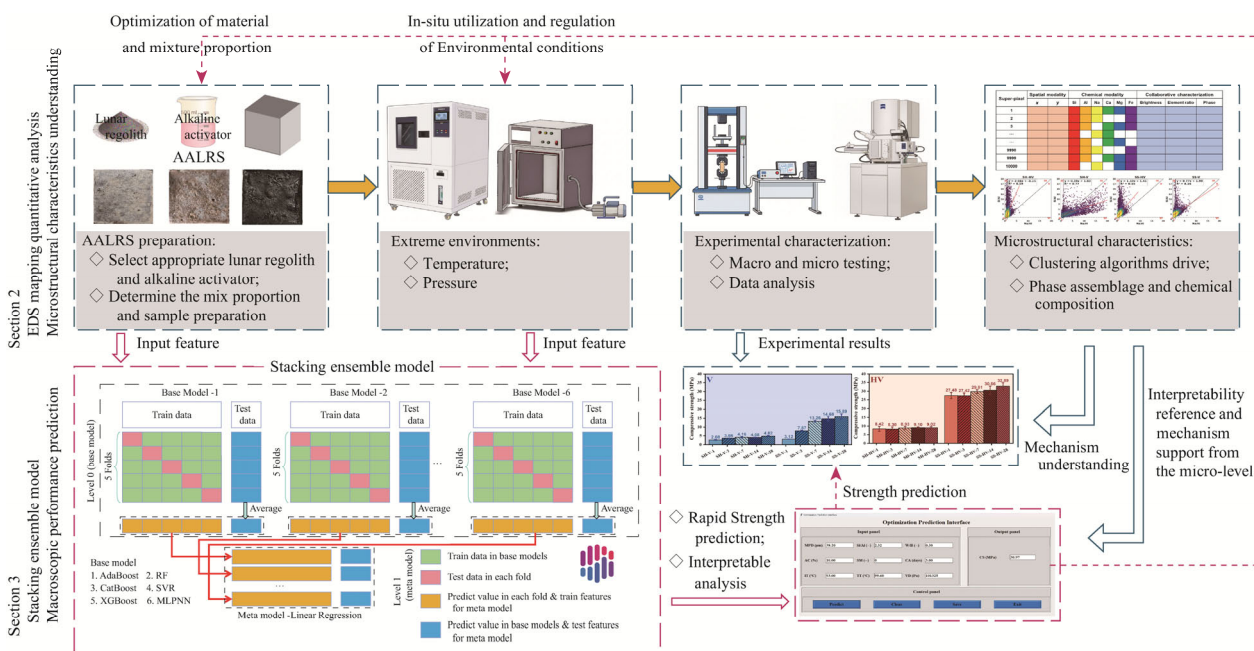


图 26 AALRS 跨尺度智能分析框架协同分析机制

Fig. 26 Collaborative analysis mechanism of AALRS cross-scale intelligent analysis framework

5 结论

1) 在 HV 环境下, AALRS-SS 展现出最高强度, 其 28 d 强度高达 32.9 MPa, 是 AALRS-SH (9.0 MPa) 的 3.65 倍。这归因于 AALRS-SS 具有更高占比的凝胶产物。SH-V-28 的强度仅为 4.8 MPa, 源于基体中残留大量碱溶液, 导致反应程度较低。SH-HV、SH-V 和 SS-HV 强度总体呈现平稳增长趋势, 而 SS-V 强度增长显著, 其 28 d 强度增长率高达 409.2%。

2) 构建了基于 GMM 和 SLIC 聚类算法的 EDS 图像定量分析框架, 实现了 AALRS 微观结构特征的“算法解构”与定量表征。结果表明, 相比 V 环境, HV 环境能显著提高反应产物占比与成分均匀性, 并降低孔隙率。AALRS-SS 的反应程度显著高于 AALRS-SH。AALRS 体系的 Ca/Al 远低于 Si/Al, 产物以 N-A-S-H 凝胶为主, Ca 元素基本保持惰性。AALRS-SH 凝胶产物铝含量偏高, 而 AALRS-SS 凝胶产物硅含量偏高。

3) 揭示了极端环境下 AALRS-SH 与 AALRS-SS 的反应机理。AALRS-SH 在 V 环境下孔隙率显著增加, 且孔隙内部出现 Na 富集。在 HV 环境下, 凝胶在颗粒表面沉淀形成产物层, 阻碍后续溶解, 导致后续反应受限、结构多孔与强度偏低。AALRS-SS 因硅酸盐多聚体存在, 在 HV 环境下经解聚产生大量成核位点, 形成致密的微观结构。而在 V 环境下则通过硅酸盐多聚体自组装凝胶化产生黏结作用。

4) 使用 6 种异构模型作为基学习器, 以线性回归作为元学习器建立 Stacking 集成模型, 并利用 6 种超参数优化算法在模型配置空间进行高效寻优。综合评估表明, RS 算法优化的 Stacking-RS 模型展现出最优的预测精度、泛化能力和计算效率 ($R=0.948\ 2$, $RMSE=5.58\ \text{MPa}$, $MAE=4.07\ \text{MPa}$, $t=0.77\ \text{min}$)。

5) SHAP 可解释性分析(全局、局部及依赖性分析)结果表明, AC、CA 与 IT 为贡献最显著的正向影响特征; MPD 的负向影响在 $50\ \mu\text{m}$ 后愈发显著; W/B 的影响呈非单调性; SM 的 SHAP 值分布范围较宽, 受其他特征交互作用影响。模型的决策逻辑为解析 AALRS 强度发展规律和影响机制提供了重要洞见。此外, 基于 Stacking-RS 模型开发了 GUI, 提升了模型的实用性并推动其工程化应用。

6) 提出了基于数据驱动的“算法解构-性能预测-机制解析”的 AALRS 跨尺度智能分析新范式。其核心在于宏观预测模块用于强度预测与识别关键因素, 而微观结构模块提供直接的试验证据以阐明其背后的强度发展机制。研究范式为 AALRS 在极端环境下性能的准确表征、深入理解和可靠预测提供支撑, 推动月基建筑材料表征和性能预测由传统经验试错向智能数据驱动的科学范式转变。

参考文献:

- [1] 丁烈云, 周诚, 高玉月, 等. 地外建造研究进展与科学技术挑战[J]. 土木工程学报, 2024, 57(6): 26–42.
DING Lieyun, ZHOU Cheng, GAO Yuyue, et al. China Civ Eng J, 2024, 57(6): 26–42.
- [2] 周诚, 高玉月, 覃文波, 等. 月球基地建造的发展趋势、目标体系与政策路径[J]. 科学通报, 2025, 70(21): 3460–3469.
ZHOU Cheng, GAO Yuyue, QIN Wenbo, et al. Chin Sci Bull, 2025, 70(21): 3460–3469.
- [3] CHEN L C, WANG T, LI F, et al. Preparation of geopolymer for *in-situ* pavement construction on the Moon utilizing minimal additives and human urine in lunar regolith simulant[J]. Front Mater, 2024, 11: 1413432.
- [4] BAO C R, ZHANG D B, WANG Q Y, et al. Lunar *in situ* large-scale construction: Quantitative evaluation of regolith solidification techniques[J]. Engineering, 2024, 39: 204–221.
- [5] 姚羿舟, 朱超, 张伟, 等. 极端环境下碱激发模拟月壤孔隙结构特征研究[J/OL]. 土木工程学报, 2025: 1–14. (2025–11–03).
<https://doi.org/10.15951/j.tmgcxb.25070477>.
YAO Yizhou, ZHU Chao, ZHANG Wei, et al. China Civ Eng J, 2025: 1–14. (2025–11–03). <https://doi.org/10.15951/j.tmgcxb.25070477>.
- [6] GENG Z F, ZHANG L Z, PAN H, et al. *In-situ* solidification of alkali-activated lunar regolith: Insights into the chemical and physical origins[J]. J Clean Prod, 2023, 391: 136147.
- [7] WU Z W, GENG Z F, PAN H, et al. Determinants of lunar geopolymer synthesis based on experimental and statistical analysis: Towards lunar regolith's diversity[J]. Constr Build Mater, 2024, 450: 138557.
- [8] XUE G J, QIAO G F. Impacts of thermal activation on lunar regolith simulant-based precursor and resulting geopolymer: Composition, structure, solubility, and reactivity[J]. Cem Concr Compos, 2025, 155: 105840.
- [9] DAVIS G, MONTES C, EKLUND S. Preparation of lunar regolith based geopolymer cement under heat and vacuum[J]. Adv Space Res, 2017, 59(7): 1872–1885.
- [10] MILLS J N, KATZAROVA M, WAGNER N J. Comparison of lunar and Martian regolith simulant-based geopolymer cements formed by alkali-activation for *in-situ* resource utilization[J]. Adv Space Res, 2022, 69(1): 761–777.
- [11] ZHOU S Q, YANG Z N, ZHANG R R, et al. Preparation and evaluation of geopolymer based on BH-2 lunar regolith simulant under

- lunar surface temperature and vacuum condition[J]. Acta Astronaut, 2021, 189: 90–98.
- [12] ZHANG R R, ZHOU S Q, LI F. Preparation of geopolymers based on lunar regolith simulant at *in-situ* lunar temperature and its durability under lunar high and cryogenic temperature[J]. Constr Build Mater, 2022, 318: 126033.
- [13] XIONG G Y, GUO X L, YUAN S T, et al. The mechanical and structural properties of lunar regolith simulant based geopolymer under extreme temperature environment on the Moon through experimental and simulation methods[J]. Constr Build Mater, 2022, 325: 126679.
- [14] 张文生, 曹传荔, 鄧晓, 等. 机器学习方法用于水泥基材料的研究进展[J]. 硅酸盐学报, 2024, 52(11): 3617–3630.
- ZHANG Wensheng, CAO Fuli, ZHI Xiao, et al. J Chin Ceram Soc, 2024, 52(11): 3617–3630.
- [15] NASER M Z. Extraterrestrial construction materials[J]. Prog Mater Sci, 2019, 105: 100577.
- [16] 刘晓, 王思迈, 卢磊, 等. 机器学习预测混凝土材料耐久性的研究进展[J]. 硅酸盐学报, 2023, 51(8): 2062–2073.
- LIU Xiao, WANG Simai, LU Lei, et al. J Chin Ceram Soc, 2023, 51(8): 2062–2073.
- [17] 李粒琰, 王相权, 姚羿舟, 等. 基于背散射电子–X射线能谱图像智能分析的胶凝材料微观多模态多视图表征[J]. 硅酸盐学报, 2025, 53(8): 2362–2373.
- LI Lihui, WANG Xiangquan, YAO Yizhou, et al. J Chin Ceram Soc, 2025, 53(8): 2362–2373.
- [18] YU Y, GENG G Q. Deep learning methods for phase segmentation in backscattered electron images of cement paste and SCM-blended systems[J]. Cem Concr Compos, 2025, 155: 105810.
- [19] ZHOU C, CHEN R, XU J, et al. *In-situ* construction method for lunar habitation: Chinese Super Mason[J]. Autom Constr, 2019, 104: 66–79.
- [20] CHEN Z, ZHANG L X, TANG Y C, et al. Pioneering lunar habitats through comparative analysis of *in-situ* concrete technologies: A critical review[J]. Constr Build Mater, 2024, 435: 136833.
- [21] GOGUEN J, SHARITS A, CHIARAMONTI A, et al. Three-dimensional characterization of particle size, shape, and internal porosity for Apollo 11 and Apollo 14 lunar regolith and JSC-1A lunar regolith soil simulant[J]. Icarus, 2024, 420: 116166.
- [22] HEIKEN G H, VANIMAN D T, FRENCH B M. Lunar sourcebook: A user's guide to the Moon[M]. Cambridge, Cambridge University Press 1991: 285–356.
- [23] YAO Y Z, LIU C, LIU H W, et al. Multiscale study of microstructural evolution in alkali-activated lunar regolith simulant under *in-situ* lunar temperatures: Insight into the reaction mechanism[J]. J Build Eng, 2025, 103: 112209.
- [24] LI L H, MAO L X, YANG J. A review of principles, analytical methods, and applications of SEM–EDS in cementitious materials characterization[J]. Adv Mater Technol, 2025, 10(7): 2401175.
- [25] LI L H, DING S, CAI Y X. Phase–Glue: An interpretable and efficient backscattered electron-energy dispersive X-ray spectroscopy (BSE-EDS) mapping analysis approach to dissect the phase assemblage of cementitious material systems[J]. Expert Syst Appl, 2025, 283: 127923.
- [26] LI L H, YANG J, LI H Y. Quantifying the microstructure and phase assemblage of alkali-activated fly ash/slag materials by EDS mapping analysis[J]. Mater Des, 2023, 234: 112320.
- [27] MATINFAR M, NYCHKA J A. A review of sodium silicate solutions: Structure, gelation, and syneresis[J]. Adv Colloid Interface Sci, 2023, 322: 103036.
- [28] GENG Z F, ZHANG L Z, WU Z W, et al. Silicate-lunar regolith composite: A vacuum self-hardened and comprehensively durable extraterrestrial construction material[J]. Constr Build Mater, 2024, 449: 138517.
- [29] ALEXIADIS A, ALBERINI F, MEYER M E. Geopolymers from lunar and Martian soil simulants[J]. Adv Space Res, 2017, 59(1): 490–495.
- [30] COLLINS P J, EDMUNSON J, FISKE M, et al. Materials characterization of various lunar regolith simulants for use in geopolymer lunar concrete[J]. Adv Space Res, 2022, 69(11): 3941–3951.
- [31] DEBBARMA S, SHI X J, TORRES A, et al. Fiber-reinforced lunar geopolymers synthesized using lunar regolith simulants[J]. Acta Astronaut, 2024, 214: 593–608.
- [32] EGNACZYK T M, HARTT V W H, MILLS J N, et al. Composition-property relationships of BP-1 lunar regolith simulant geopolymers for *in-situ* resource utilization[J]. Adv Space Res, 2024, 73(1): 885–917.
- [33] GU J H, MA Q Y. Experimental study on geopolymerization of lunar soil simulant under dry curing and sealed curing[J]. Materials, 2024, 17(6): 1413.
- [34] JAVED U, AHMED SHAIKH F U, RAHMAN A K M S. Synthesizing lunar regolith-geopolymer emulating lunar positive temperature regime[J]. Planet Space Sci, 2024, 244: 105890.
- [35] MA Q Y, ZHANG B B. Study on the performance mechanism of alumina-alkali activator synergistic solidified lunar soil simulant[J]. Case Stud Constr Mater, 2024, 21: e03680.
- [36] MONTES C, BROUSSARD K, GONGRE M, et al. Evaluation of lunar regolith geopolymer binder as a radioactive shielding material for space exploration applications[J]. Adv Space Res, 2015, 56(6): 1212–1221.
- [37] PILEHVAR S, ARNHOF M, PAMIES R, et al. Utilization of urea as an accessible superplasticizer on the Moon for lunar geopolymer mixtures[J]. J Clean Prod, 2020, 247: 119177.
- [38] XUE G J, QIAO G F. Coupling thermodynamic modelling with experimental study to reveal the evolutionary relationship of pore solutions, products, and compressive strength for lunar regolith simulant geopolymers[J]. Compos Part B Eng, 2025, 289: 111949.
- [39] YAO Y Z, LIU C, ZHANG W, et al. Influence of vacuum and high-temperature on the evolution of mechanical strength and microstructure of alkali-activated lunar regolith simulant[J]. J Build Eng, 2024, 97: 110709.
- [40] YAO Y Z, LIU C, ZHANG W, et al. Influence of *in-situ* lunar temperature on the pore structure and solid phases of alkali-activated lunar regolith simulant[J]. J Build Eng, 2024, 98: 111162.
- [41] ZHAN J H, XUE X Y, HUA J M, et al. Effects of curing temperature on early-age mechanical property and microstructure of lunar regolith simulant geopolymer[J]. Case Stud Constr Mater, 2025, 22: e04329.
- [42] ZHOU S Q, LU C H, ZHU X Y, et al. Preparation and characterization

- of high-strength geopolymer based on BH-1 lunar soil simulant with low alkali content[J]. *Engineering*, 2021, 7(11): 1631–1645.
- [43] ZHOU S Q, ZHU X Y, LU C H, et al. Synthesis and characterization of geopolymer from lunar regolith simulant based on natural volcanic scoria[J]. *Chin J Aeronaut*, 2022, 35(1): 144–159.
- [44] 姚羿舟, 刘超, 刘化威, 等. 热真空环境下碱激发模拟月壤物相分布特征与固化反应机理研究[J]. *土木工程学报*, 2025, 58(5): 105–118.
YAO Yizhou, LIU Chao, LIU Huawei, et al. *China Civ Eng J*, 2025, 58(5): 105–118.
- [45] BERGSTRA J, BENGIO Y. Random search for hyper-parameter optimization[J]. *J Mach Learn Res*, 2012, 13: 281–305.
- [46] LI Q F, WANG X. Bayesian optimization of stacking ensemble learning model for HPC compressive strength prediction[J]. *Expert Syst Appl*, 2025, 288: 128281.
- [47] JIANG B N, ZHANG Y Y, ZHANG Z Y, et al. Tree-structured parzen estimator optimized-automated machine learning assisted by meta-analysis for predicting biochar-driven N₂O mitigation effect in constructed wetlands[J]. *J Environ Manag*, 2024, 354: 120335.
- [48] ŞENER R, KOÇ M A, ERMIŞ K. Hybrid ANFIS-PSO algorithm for estimation of the characteristics of porous vacuum preloaded air bearings and comparison performance of the intelligent algorithm with the ANN[J]. *Eng Appl Artif Intell*, 2024, 128: 107460.
- [49] KIRKPATRICK S, GELATT C D Jr, VECCHI M P. Optimization by simulated annealing[J]. *Science*, 1983, 220(4598): 671–680.
- [50] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer[J]. *Adv Eng Softw*, 2014, 69: 46–61.

作者贡献声明:

姚羿舟: 论文构思、设计研究方法/实验方案/模型算法、分析/综合数据并推导结论、绘制图表、撰写初稿

朱超: 论文构思、调查研究、验证

张伟: 论文构思、软件使用、数据管理、设计模型算法

刘超: 论文构思、提供资源、获取资助、项目管理、监督指导、审阅与修订

Alkali-Activated Lunar Regolith Simulant Under Extreme Environments: Data-Driven Microstructural Characterization and Performance Prediction

YAO Yizhou^{1,2}, ZHU Chao^{1,2,3}, ZHANG Wei^{1,2}, LIU Chao^{1,2,3}

(1. School of Civil Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China; 2. Digital Green Construction Research Institute, Xi'an University of Architecture and Technology, Xi'an 710055, China; 3. Shaanxi Intelligent Construction Future Industry Innovation Research Institute, Xi'an 710055, China)

Extended Abstract

Introduction Lunar construction represents a strategic frontier in deep space exploration. Alkali-activated lunar regolith simulant (AALRS) is a highly promising *in-situ* construction material. Accurately characterizing, understanding and predicting its performance under extreme environments are critical to the safety of lunar structures during construction and service. However, the existing research shows significant discrepancies and controversies. These arise from a series of challenges, *i.e.*, 1) the inherent compositional complexity and variability of lunar regolith simulant, 2) the coupled effects of multiple factors, 3) a general reliance on qualitative and localized understanding of microstructure, 4) the lack of an underlying foundation for understanding the mechanisms that influence macroscopic performance, and 5) the reliance on simple statistical analysis of experimental data without establishing quantitative association models between parameters and performance. To address these challenges, this study was to propose an interpretable, cross-scale and data-driven research paradigm of “algorithmic deconstruction-performance prediction-mechanism analysis”.

Methods AALRS pastes were prepared using lunar regolith simulant XJ-1, alkaline activators (*i.e.*, sodium hydroxide or sodium silicate), and water at a specified alkali content of 0.075 and a water-to-binder ratio of 0.3. To simulate extreme environments, the samples were exposed to a vacuum (*i.e.*, 20 °C and 100 Pa) and a thermal vacuum (*i.e.*, 90 °C and 100 Pa) for 1, 3, 7, 14 d and 28 d, respectively. Macroscopically, the compressive strength was measured in a WAW-1000 testing machine at a loading rate of 480 N/s. The energy dispersive X-ray spectroscopy (EDS) mappings of key elements (*i.e.*, Si, Al, Na, Ca, Mg, and Fe) were acquired in an accelerating voltage of 15 kV, a working distance of 10 mm, a dwell time of 100 µs/pixel, a process time of 6, and a channel count of 2048.

A quantitative analysis framework for EDS mappings driven by unsupervised clustering algorithms was developed. Implemented in Python, this framework automates the establishment of an interpretable analysis paradigm using libraries such as OpenCV, scikit-learn, and scikit-image. To address underfitting, overfitting, and poor performance of basic ML algorithms in complex hypothesis spaces, a two-level Stacking ensemble model was proposed. The first level integrates six heterogeneous base learners like AdaBoost, XGBoost, CatBoost, SVR, MLPNN, and RF. The predictions from these base learners served as input features for the second level by a 5-fold cross validation strategy. This second level employs a linear regression model as a meta-learner. To further enhance the predictive accuracy of the model, its configuration was optimized *via* applying six hyperparameter tuning algorithms (*i.e.*, RS, BO, PSO, GWO, SA, and TPE) to efficiently search the hyperparameter space.

Results and discussion The results of compressive strength reveal the disparities governed by both the activation system and environment. The AALRS-SS consistently outperforms the AALRS-SH. The AALRS-SS under HV environment achieves a peak 28 d strength of 32.9 MPa, which is 3.65 times higher than its AALRS-SH counterpart under the same conditions. In addition, the HV environment proves crucial for the strength development, compared to the V environment. To understand the mechanisms behind these macroscopic differences, a quantitative EDS mapping analysis is employed. The results show that AALRS-SS generally contains a higher area fraction of reaction products and correspondingly lower proportions of unreacted XJ-1 and porosity. This indicates a higher overall degree of reaction and a denser microstructure in the AALRS-SS, providing a direct explanation for its superior strength. The HV environment facilitates a more homogeneous phase assemblage, compared to the greater heterogeneity observed in the V environment. The elemental analysis based on the superpixel clustering further reveals that calcium remains largely inert. The gels in the AALRS-SH exhibit a higher aluminum content, while those in the AALRS-SS are characterized at a higher silicon content. Based on the quantitative EDS analysis, the microstructural characteristics and reaction mechanisms of AALRS under different environments and activation systems are elucidated.

To capture the complex, non-linear relationships among material composition, environmental conditions, and mechanical performance, a two-level Stacking ensemble model is proposed. The model tuned with RS is identified as an optimal choice for its excellent balance of high performance and rapid computation after comparing multiple optimization algorithms on both accuracy and computational efficiency. This selected Stacking-RS model is interpreted based on the SHAP analysis to reveal the underlying strength development mechanisms. The analysis confirms the strong, positive, and monotonic effects of AC, CA and IT. In contrast, it identifies a critical negative threshold for MPD around 50 μm and reveals a non-monotonic influence of W/B. The effect of SM is complex, with significant interactions with other features. A graphical user interface (GUI) is developed based on the Stacking-RS model, thus facilitating the efficient design of AALRS. Furthermore, this study establishes a cross-scale, data-driven analytical framework that integrates algorithmic deconstruction, performance prediction, and mechanistic interpretation, wherein the microstructural analysis and macroscopic prediction modules mutually complement and validate each other to enable an intelligent understanding and accurate prediction of AALRS performance under extreme environments.

Conclusions AALRS-SS under the HV environment exhibited the maximum strength. Its 28 d strength reached 32.9 MPa, which was 3.65 times greater than that of AALRS-SH (9.0 MPa). This was attributed to the SS system forming a higher proportion of gel products and a denser microstructure.

A quantitative EDS image analysis framework based on GMM and SLIC clustering algorithms was established. Compared with the V environment, the HV environment significantly increased the proportion of reaction products and compositional homogeneity, while reducing porosity. The reaction degree of AALRS-SS was significantly higher than that of AALRS-SH.

The reaction mechanisms of AALRS-SH and AALRS-SS under extreme environments were revealed. Under the V environment, AALRS-SH showed a significantly increased porosity with Na enrichment inside the pores. Under the HV environment, gels precipitated on the XJ-1 particle surfaces, forming a product layer that hindered subsequent dissolution, leading to a porous structure, and low strength. In contrast, AALRS-SS, due to the presence of silicate polymers, provided numerous nucleation sites through depolymerization under the HV environment, forming a relatively dense microstructure. Under the V environment, it provided a bonding force through self-assembly gelation of silicate polymers.

A Stacking ensemble model optimized by RS was identified as the best-performing model, showing excellent accuracy, generalization, and efficiency (*i.e.*, $R=0.9482$, $\text{RMSE}=5.58\text{ MPa}$, $\text{MAE}=4.07\text{ MPa}$, and $t=0.77\text{ min}$).

The SHAP analysis (global, local, and dependence) showed that AC, CA, and IT were the most significant positive contributing features. The negative impact of MPD became more significant beyond 50 μm . The effect of W/B was non-monotonic. A GUI was developed based on the Stacking-RS model to enhance its practicality.

A data-driven, cross-scale intelligent analysis paradigm of “algorithmic deconstruction-performance prediction-mechanism analysis” was proposed. Its core lay in the macro-prediction module, which was used for strength prediction and identifying key factors, and the microstructure module, providing direct experimental evidence to elucidate the underlying strength development mechanism.

Keywords alkali-activated lunar regolith; extreme environment; machine learning; microstructure; performance prediction