

组成驱动的无机材料XPS O1s结合能与化学位移智能预测

付铭凯¹, 吴璇¹, 于丹凤^{2*}, 吴正龙¹

(1. 北京师范大学分析测试中心, 北京 100875; 2. 北京师范大学珠海校区理工实验平台, 广东 珠海 519087)

摘要: O 1s XPS结合能是识别无机材料氧化化学环境的核心指标, 在新能源、催化及合成监测等领域应用广泛, 但跨体系峰位指认仍依赖大量验证性实验。该文基于514个含氧无机化合物O 1s结合能数据, 构建仅依赖化学式输入的组成驱动机器学习预测框架, 用于实验前低成本峰位预估和趋势判断。原始组成描述符下, 梯度提升回归(GB)模型测试的平均绝对误差(MAE)为0.505 eV。引入阳离子场强、过渡金属比例等物理启发描述符后, MAE降至0.481 eV, 决定系数 R^2 升至0.716, 平均阳离子半径被识别为关键变量, 并得到 $\text{Ca}_{14}\text{Mg}_5\text{M}_4\text{O}_{25}$ (M=Al, Sc, Y) XPS实验结果的支持。在盐/氧化物配对体系中, 模型对相对化学位移预测达到MAE=0.339 eV和 $R^2=0.822$ 。该方法可为复杂含氧无机材料XPS分析提供低成本、可解释的辅助参考。

关键词: 绿色高效分析; X射线光电子能谱; 化学位移; 机器学习; 组成描述符

中图分类号: O657.3 **文献标识码:** A **文章编号:** 1004-4957(2026)10-0001-09

Composition-driven Intelligent Prediction of XPS O 1s Binding Energy and Chemical Shift in Inorganic Materials

FU Ming-kai¹, WU Xuan¹, YU Dan-feng^{2*}, WU Zheng-long¹

(1. Analytical and Testing Center, Beijing Normal University, Beijing 100875, China; 2. Instrumentation and Service Center for Science and Technology, Beijing Normal University, Zhuhai 519087, China)

Abstract: The O 1s XPS binding energy is a key indicator for identifying oxygen chemical environments in inorganic materials with broad applications in energy materials, catalysis, and synthesis monitoring. However, the cross-system peak assignment relies heavily on extensive validation experiments. In this study, we developed a composition-driven machine learning framework based on O 1s binding energy data from 514 oxygen-containing inorganic compounds. This framework exclusively uses chemical formula inputs for low-cost pre-experimental peak-position estimation and trend assessment. Using the original compositional descriptors, the Gradient Boosting model achieved a test MAE of 0.505 eV. After introducing physically inspired descriptors, including the cation field strength and transition metal fraction, the MAE decreased to 0.481 eV and R^2 increased to 0.716. The mean cation radius was identified as a key variable, a finding further supported by the XPS measurements of $\text{Ca}_{14}\text{Mg}_5\text{M}_4\text{O}_{25}$ (M=Al, Sc, Y). For the paired salt/oxide systems, the model predicted the relative chemical shifts with an MAE of 0.339 eV and a R^2 of 0.822. This method provides a low-cost and interpretable auxiliary reference for XPS analysis of complex oxygen-containing inorganic materials.

Key words: green and efficient analysis; X-ray photoelectron spectroscopy; chemical shift; machine learning; compositional descriptors

绿色化与智能化是当前分析测试技术的核心发展趋势。X射线光电子能谱(XPS)可精准表征材料表面化学组成及元素化学态, 是无机材料研究的关键技术^[1-2]。通过测量芯能级结合能的化学位移, XPS能够区分同一元素在不同化学环境中的电子态, 为氧化物功能材料^[3]提供关键化学信息。在含氧无机材料中, 氧可以以多种形式存在, 包括晶格氧、表面吸附氧、羟基、碳酸盐、硝酸盐及硅酸盐等, 准

收稿日期: 2026-06-10; 修回日期: 2026-07-20

基金项目: 国家自然科学基金资助项目(U2444219, 42372292)

* 通讯作者: 于丹凤, 博士, 工程师, 研究方向: 光谱与能谱分析, E-mail: yudanfeng@bnu.edu.cn

确识别这些氧化物对于理解氧化还原机理^[4]和表面反应过程^[5-6]具有重要意义。

然而, O 1s XPS 峰位的解释仍面临较大挑战^[7], 不同氧化化学环境的峰位区间存在明显重叠, 简单金属氧化物中晶格氧的 O 1s 结合能通常位于较低结合能区域, 含氧酸根中的氧通常位于较高结合能区域, 而表面羟基、吸附水和缺陷相关氧化物可能进一步增加谱峰解释的复杂性^[8]。O 1s 化学位移会受到阳离子电负性、离子半径、氧局域配位和成键性质等多重因素共同影响^[9]。基于此, 从跨体系数据中挖掘可解释的组成-峰位关联规律, 是提升 O 1s 峰位指认可靠性的核心挑战。

密度泛函理论(DFT)和机器学习已成为材料性质预测与谱学数据分析的重要工具^[10-12]。对于 XPS 结合能, DFT 等基于结构的计算方法需要以具体原子结构为输入。常规静态计算通常反映 0 K 条件下的性质。要描述热振动和热膨胀等有限温度效应, 则需要进行额外计算, 而含掺杂或缺陷的大超胞体系会进一步增加这类计算的成本。现有机器学习在 XPS 领域的应用主要集中于实验谱图的自动解析^[13-15]。例如, Vakhrushev 等^[13]采用基于 U-Net 架构的卷积神经网络识别高分辨 XPS 谱图中的峰区和峰顶区域, 并结合背景扣除与峰函数拟合, 实现了峰位、峰面积等参数的自动提取。相关研究表明, 机器学习可有效支持 XPS 谱图的自动处理与参数提取。除实验谱图解析外, 机器学习还可基于原子结构预测谱学化学位移。例如, Paruzzo 等^[16]基于原子结构构建机器学习模型, 实现了分子固体核磁共振化学位移的高效预测。有别于依赖实验谱图或具体原子结构信息的现有方法, 本文以材料化学式构建组成描述符, 建立组成驱动的回归模型, 预测跨体系含氧无机材料的 O 1s 结合能及相对化学位移, 以为 XPS 测试前的峰位初筛和变化趋势判断提供低成本参考。

1 计算模型与方法

1.1 数据集与描述符

本研究使用美国国家标准与技术研究院 National Institute of Standards and Technology (NIST) 数据库中 514 个含氧无机化合物的 O 1s XPS 结合能实验数据^[17], 筛选标准为化学式明确且化合物类别可判定。考虑到 NIST 中的结合能通常已按原文实验流程完成荷电校正, 数据库的结合能不再进行跨文献的二次统一校正。

鉴于原始数据仅包含化学式和平均结合能信息, 本研究采用组成层面的描述符构建预测模型。各描述符的中英文名称及图中缩写见表 1。基线模型使用 10 个组成描述符, 包括平均阳离子电负性、阳离子电负性标准差、平均阳离子半径、阳离子半径标准差、O 原子分数、H 原子分数、非 OH 阴离子比例、阳离子种类数、平均阳离子价态代理和化合物族编码。非 OH 阴离子比例表征羟基以外含氧酸根的占比, 多阳离子体系的平均值与标准差均按化学计量数加权计算。平均阳离子价态是依据化学式电荷平衡和常见氧化态规则得到的组成层面近似量。化合物族编码用于提供氧化物、碳酸盐、磷酸盐等一级化学环境信息。此外, 增强模型引入 6 个物理启发描述符, 以补充阳离子场强、过渡金属占比、半径变异性、和含氧酸根相关相互作用等信息。

表 1 组成描述符和增强描述符的中英文名称及缩写

Table 1 Chinese and English names of compositional and enhanced descriptors and their abbreviations

Category	Chinese name	English name	Abbreviation
基线描述符	平均阳离子电负性	Mean cation electronegativity	Mean EN
基线描述符	阳离子电负性标准差	Standard deviation of cation electronegativity	EN Std
基线描述符	平均阳离子半径	Mean cation radius	Mean Radius
基线描述符	阳离子半径标准差	Standard deviation of cation radius	Std Radius
基线描述符	O 原子分数	Oxygen atomic fraction	O Frac
基线描述符	H 原子分数	Hydrogen atomic fraction	H Frac
基线描述符	非 OH 阴离子比例	Non-OH anion fraction	Anion Frac
基线描述符	阳离子种类数	Number of cation types	Cation Cnt
基线描述符	平均阳离子价态代理	Mean cation valence proxy	Mean Val
基线描述符	化合物族编码	Compound family code	Family
增强描述符	阳离子场强	Cation field strength	CFS
增强描述符	阳离子半径变异性	Cation radius variability	CRV
增强描述符	过渡金属比例	Transition metal fraction	TM Frac
增强描述符	阴离子-价态相互作用	Anion-valence interaction	AVI
增强描述符	阳离子比例	Cation fraction	Cation Frac
增强描述符	非整数化学计量标记	Non-integer stoichiometry flag	Non-int Flag

1.2 模型训练与评价

本研究对比 4 种回归模型：梯度提升回归(GB)、随机森林回归(RF)、径向基函数支持向量回归(SVR-RBF)和岭回归(Ridge)。其中，GB 和 RF 用于捕捉描述符与 O 1s 结合能的非线性关系，SVR-RBF 作为核函数非线性对照，Ridge 作为线性模型对照。

数据集按化合物族进行分层划分，训练集、验证集和独立测试集比例约为 80%、10% 和 10%。模型超参数在训练集上通过五折交叉验证优化，优化指标为平均绝对误差(MAE)。最终模型性能在独立测试集上用 MAE、均方根误差(RMSE)和决定系数(R^2)评价。

机器学习源代码和全部数据集可经作者同意后提供。

1.3 外部实验验证

为评估模型在真实实验条件下的泛化能力，选取 3 种典型混合氧化物($\text{Ca}_{14}\text{Mg}_5\text{Al}_4\text{O}_{25}$ 、 $\text{Ca}_{14}\text{Mg}_5\text{Sc}_4\text{O}_{25}$ 、 $\text{Ca}_{14}\text{Mg}_5\text{Y}_4\text{O}_{25}$)。通过溶胶-凝胶法制备，在北京师范大学分析测试中心 ESCALAB 250Xi 型 XPS 仪上测试，所有样品以 $\text{C } 1s=284.8 \text{ eV}$ 为基准进行荷电校正，得到的 O 1s 结合能作为独立外部验证集。

2 结果与讨论

2.1 O 1s 结合能分布

图 1 展示了数据库中 O 1s 结合能的整体分布。O 1s 结合能覆盖范围为 527~534 eV，平均值为 531.07 eV，中位数为 531.20 eV，标准差约为 1.27 eV。MgO 参考线位于分布的较低结合能侧，这与简单离子型氧化物中晶格氧通常具有较低 O 1s 结合能的规律一致。

图 2 按化合物族展示 O 1s 结合能分布。简单氧化物整体位于较低结合能区域，含氧酸盐体系整体向高结合能移动。上述分布特征表明，O 1s 结合能不仅取决于氧本身的化学态，更与其局域配位环境及阳离子性质密切相关。此外，混合氧化物中可能存在多种阳离子—氧局域环境，结合能分布最宽。

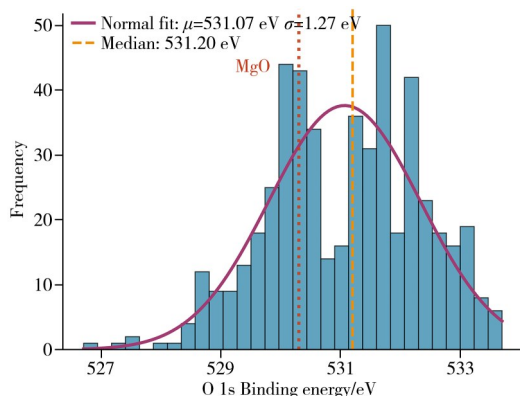


图 1 数据库中 O 1s 结合能分布

Fig. 1 Distribution of O 1s binding energies in the dataset

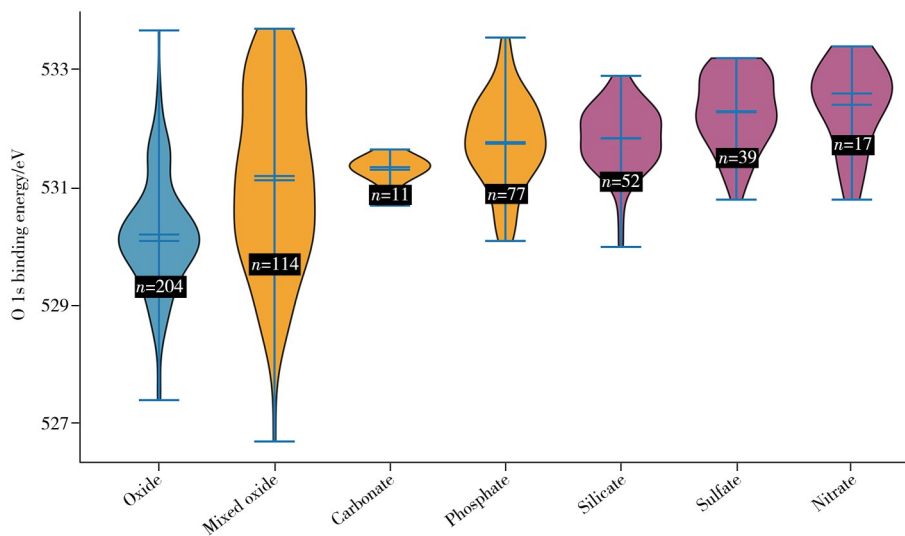


图 2 不同化合物族中 O 1s 结合能分布

Fig. 2 Distribution of O 1s binding energies by compound family

2.2 描述符相关性分析

为了探究原始组成描述符是否包含与 O 1s 结合能相关的化学信息，本研究首先对 10 个基线描述符进行主成分分析。图 3 展示了各主成解释释方差比例、累计解释方差比例以及前 5 个主成分中贡献最大

的描述符。如图 3 所示, 前几个主成分已经解释了描述符空间中的大部分方差, 主成分贡献主要来自化合物族、平均阳离子价态、阳离子种类数和平均阳离子电负性。

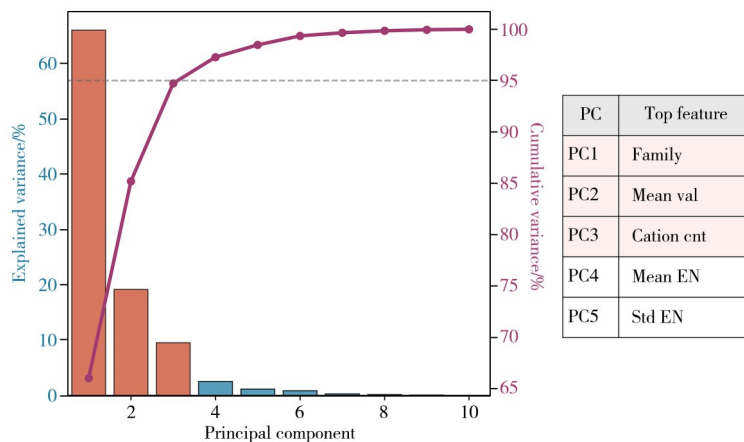


图 3 基线组成描述符的主成分分析

Fig. 3 Principal component analysis of baseline compositional descriptors

图 4 显示了主要描述符与 O 1s 结合能的相关性。其中, 平均阳离子半径与 O 1s 结合能的负相关最强($r=-0.64$), 散点图呈现出 O 1s 结合能随平均阳离子半径增大而整体降低的趋势。平均阳离子电负性则与 O 1s 结合能呈正相关, 说明阳离子尺寸及电子吸引能力均与氧局域化学环境密切相关。从局域电势与成键性质分析: 大半径阳离子局域场强较弱, 易形成离子性更强的 M—O 键, 对应较低的 O 1s 结合能; 高电负性阳离子对氧电子云的吸引力更强, 可增强 M—O 键的共价性, 使 O 1s 结合能升高。相比之下, 非 OH 阴离子比例和化合物族编码与 O 1s 结合能呈正相关($r=+0.57$ 和 $+0.56$), 与含氧酸盐整体位于较高结合能区域的结果一致。

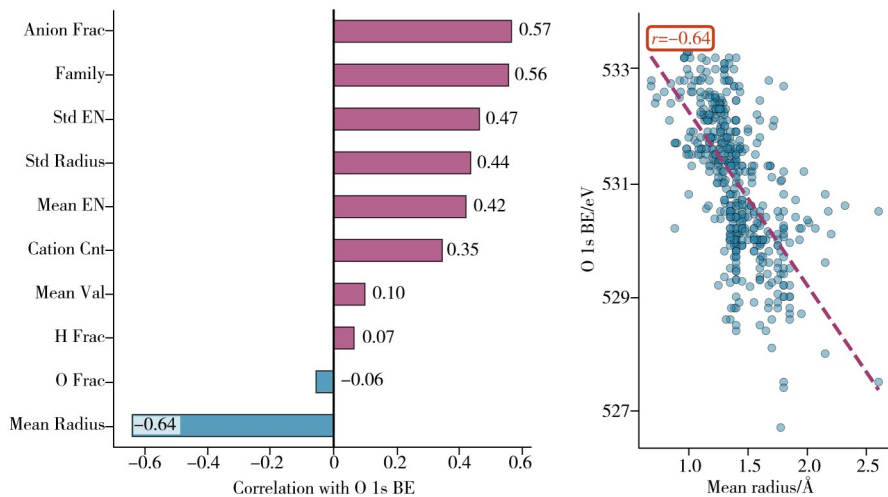


图 4 组成描述符与 O 1s 结合能的相关性

Fig. 4 Correlation between compositional descriptors and O 1s binding energies

图 5 展示了主要组成描述符之间的 Pearson 相关系数。如图所示, 这些组成描述符之间并非相互独立。平均阳离子半径与平均阳离子电负性显著负相关, 与非 OH 阴离子比例也呈负相关, 而阳离子半径标准差与阳离子电负性标准差高度相关。这些结果说明描述符空间中存在化学耦合关系, 需要非线性模型从部分冗余但具有化学意义的描述符中提取组合信息。

2.3 基线模型性能

基于上述 10 个组成描述符, 本研究对比了“1.2”中 4 种回归模型的预测性能。从图 6A 可以看出, GB 和 RF 模型在五折交叉验证和验证集上均表现更优。GB 模型的五折交叉验证 MAE 为 0.465 eV, 验证集 MAE 为 0.442 eV; RF 模型的交叉验证 MAE 为 0.474 eV, 验证集 MAE 为 0.416 eV。两者性能均显著优于线性的 Ridge 模型, 表明 O 1s 结合能与组成特征之间存在明显的非线性关系, 更适合采用非线性

模型进行描述。独立测试集结果进一步证实了这一点。图 6B~D 显示，GB 是最佳基线模型，测试 MAE 为 0.505 eV，RMSE 为 0.792 eV， R^2 为 0.679。因此，GB 被选为后续误差分析和增强特征建模的基线模型。

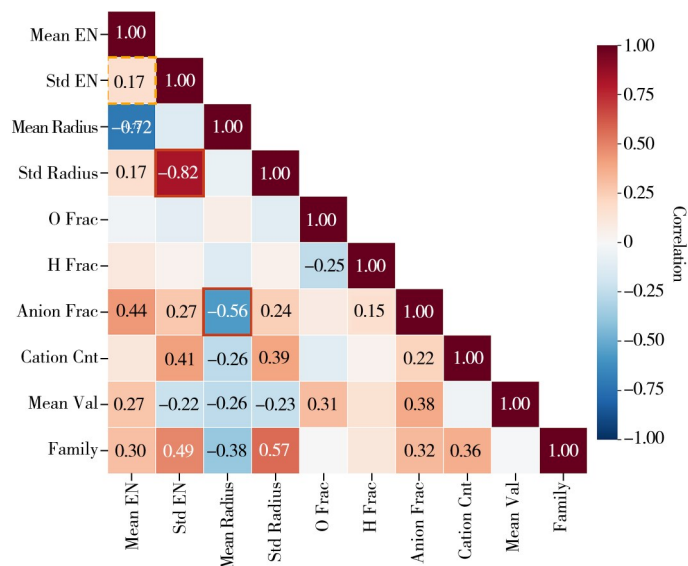


图 5 基线组成描述符相关矩阵

Fig. 5 Correlation matrix of baseline compositional descriptors

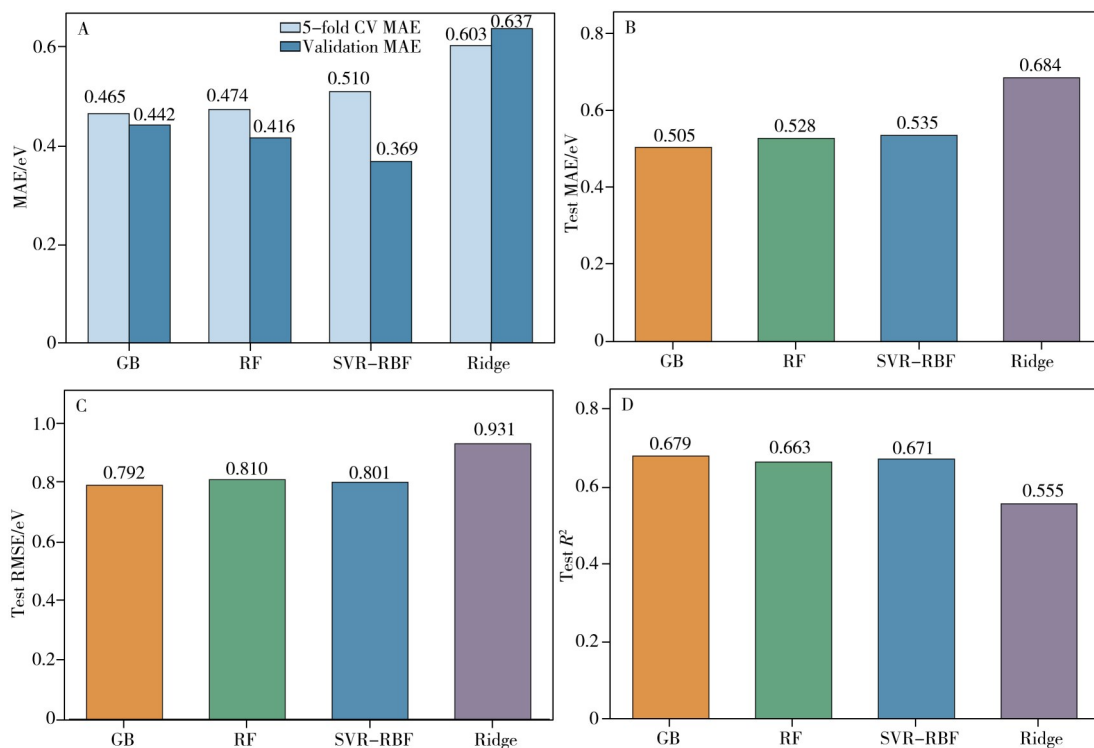


图 6 基线模型性能比较

Fig. 6 Performance comparison of baseline models

A. Five-fold cross-validation MAE and validation-set MAE for different models; B. test-set MAE; C. test-set RMSE; D. test-set R^2 (A. 不同模型的五折交叉验证 MAE 和验证集 MAE; B. 不同模型在独立测试集上的 MAE; C. 不同模型在独立测试集上的 RMSE; D. 不同模型在独立测试集上的 R^2)

图 7 进一步给出了 GB 模型在不同化合物族中的误差分布及代表性大误差样本。图 7A 显示，硝酸盐、硅酸盐、硫酸盐以及多数简单氧化物样本的绝对误差整体较小，箱体范围也相对集中，说明对于具有相对明确一级化学环境的含氧酸盐和氧化物体系，组成描述符能够较好捕捉 O 1s 结合能的主要变化趋势。相比之下，混合氧化物的误差分布明显更宽，整体离散程度更高，表明多阳离子氧化物中更复杂的局部配位与成键环境会增加仅基于组成预测的难度。图 7B 进一步显示， BaTiO_3 和 CaAl_2O_4 等混合氧化物样本的绝对误差超过 2 eV，是测试集中主要的大误差来源。 BaTiO_3 中 Ti-O 配位多面体的结构畸

变可改变氧周围的局域电势和成键环境, CaAl_2O_4 中 O 与 Ca、Al 阳离子的连接方式及局部配位差异也可形成不同的氧化学环境。由于当前模型仅基于组成特征, 未包含晶体结构、氧配位数以及 M—O 键长/键角等局部结构与键合信息, 因此难以区分这类由结构差异导致的特殊氧环境, 从而在上述样本中产生较大的预测偏差。这些结果说明, 化学式描述符能够有效表征含氧酸盐和简单氧化物等体系中的主要组成差异, 但对于存在多种阳离子配位环境或复杂晶体结构的材料, 仍难以完全替代局部结构信息。

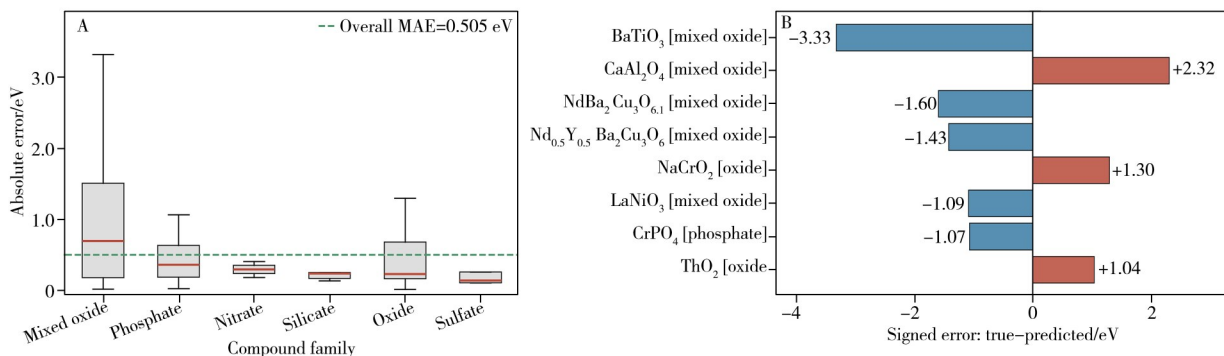


图7 基线GB模型误差分析

Fig. 7 Error analysis of the baseline GB model

A. distribution of absolute errors grouped by compound family; B. representative samples with the largest absolute errors and their signed errors
(A. 按化合物族分类的绝对误差分布; B. 绝对误差最大的代表性样本及其符号误差)

2.4 特征增强

为更精准地表征氧局域化学环境, 在基线特征基础上引入6个物理启发型增强描述符: 阳离子场强, 定义为平均阳离子价态与平均阳离子半径平方的比值, 表征阳离子对氧电子云的极化能力; 阳离子半径变异性, 为阳离子半径标准差与平均值之比, 反映多阳离子体系的尺寸不均一性; 过渡金属比例, 即过渡金属阳离子在总阳离子中的化学计量占比, 体现过渡金属对氧环境的影响; 阴离子-价态相互作用, 由非OH阴离子比例与平均阳离子价态组合得到, 描述含氧酸根环境的价态差异; 阳离子比例, 表征基本化学计量关系; 非整数化学计量标记, 识别非整数配比或组成复杂体系。

图8显示, 引入增强描述符后, GB模型在独立测试集上的3个评价指标均优于基线模型: 测试MAE从0.505 eV降至0.481 eV, R^2 从0.679提高至0.716。虽然MAE仅降低0.024 eV, 但多项指标的一致改善表明, 增强描述符为模型提供了有价值的补充信息。由于模型仅依赖化学式输入, 该结果说明其可作为一种低成本的组成驱动方法, 用于实验前快速评估O 1s峰位区间及相对偏移趋势。

图9A、B对比了基线模型和增强模型的预测-实验散点图。两个模型均能捕捉O 1s结合能的整体变化趋势, 但增强模型的点分布相对更接近 $y=x$ 对角线。图9C、D的残差分析显示, 增强模型残差大多数集中在零附近, 残差均值为-0.115 eV, 标准差为0.735 eV。

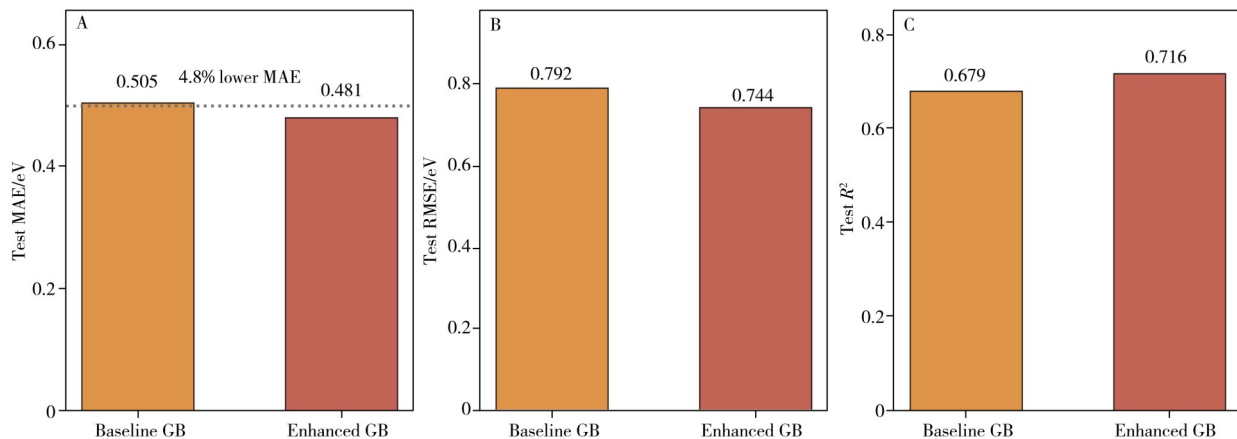


图8 物理启发增强描述符对模型性能的影响

Fig. 8 Effect of physically motivated enhanced descriptors on model performance

A. test-set MAE comparison between baseline and enhanced models; B. testset RMSE comparison; C. test-set R^2 comparison (A. 基线模型与增强模型的测试集MAE对比; B. 基线模型与增强模型的测试集RMSE对比; C. 基线模型与增强模型的测试集 R^2 对比)

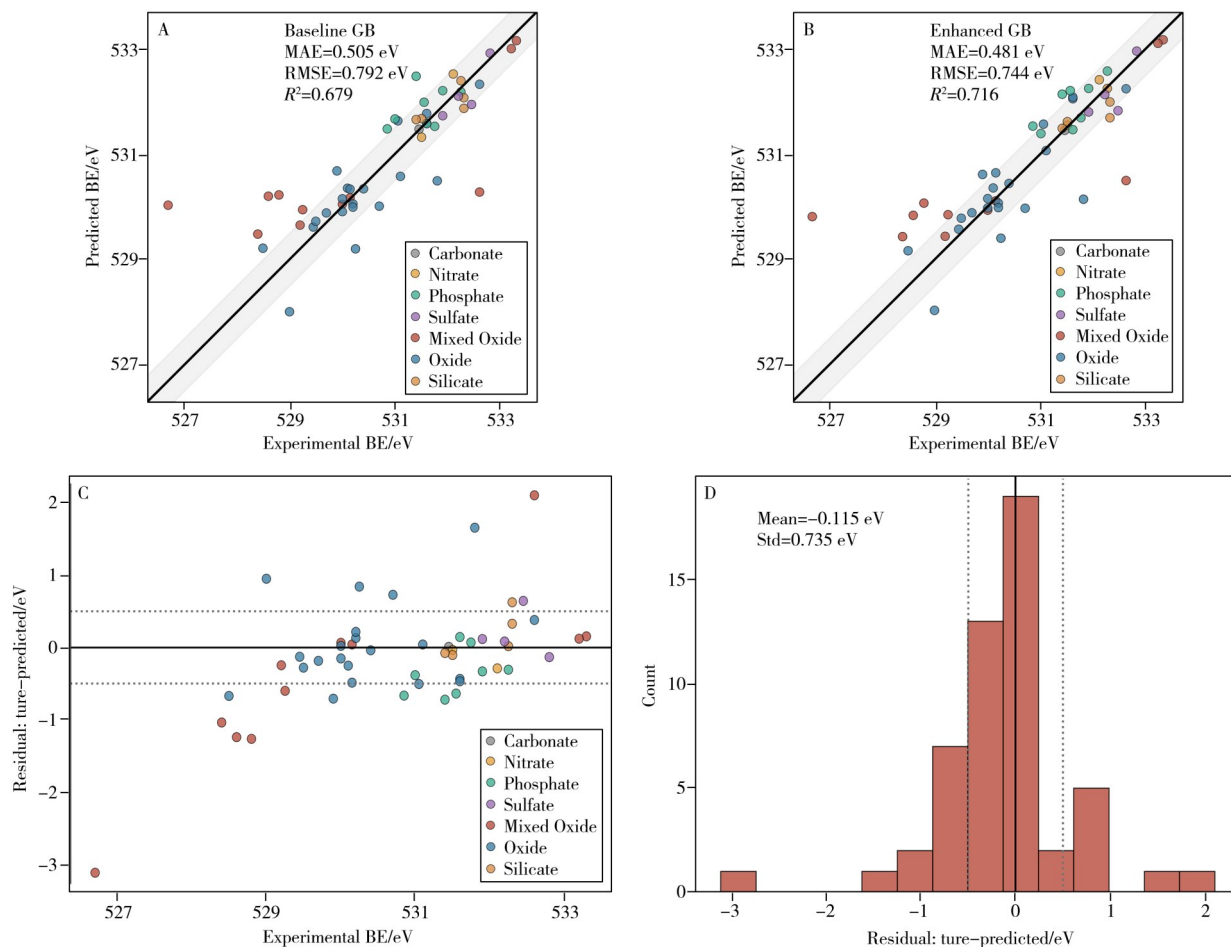


图9 基线模型与增强模型的预测结果和残差分析

Fig. 9 Prediction results and residual analysis for baseline and enhanced models

A. predicted versus experimental binding energies for the baseline GB model; B. predicted versus experimental binding energies for the enhanced GB model; C. residuals of the enhanced model as a function of experimental binding energy; D. residual histogram of the enhanced model (A. 基线GB模型的预测值与实验值对比; B. 增强GB模型的预测值与实验值对比; C. 增强模型残差随实验结合能的分布; D. 增强模型残差直方图)

为了判断增强模型的性能提升是否具有化学可解释性,本研究计算了置换特征重要性。图10横坐标表示对单个特征进行置换后模型MAE的增加量,误差线表示重复置换得到的标准差。结果显示,平均阳离子半径是最重要的特征,置换后MAE增加量远高于其他描述符。这一结论与图4中平均阳离子半径和O 1s结合能的强负相关趋势吻合。除平均阳离子半径外,阳离子场强、过渡金属比例和阳离子半径变异性也具有较高重要性。阳离子场强整合了价态与半径的协同作用,可更精准地表征氧周围的局域静电场。过渡金属比例则体现了过渡金属在体系中的占比,可作为其成键及可变价态效应的组成层面标识。阳离子半径变异性反映混合阳离子体系中的组成和局域环境不均一性。

2.5 盐/氧化物相对位移预测

在原位XPS研究中,研究者通常更关注反应过程中峰位的相对变化,以此判断化学环境变化。与绝对结合能分析相比,相对位移分析在某些场景下对系统偏差的敏感度更低。因此,本研究进一步使用增强模型预测成对盐/氧化物体系中的O 1s相对位移。

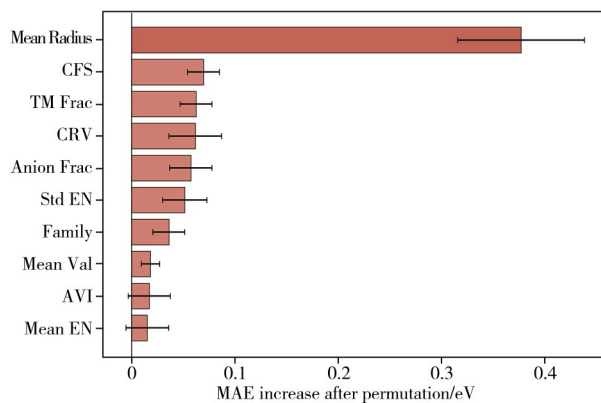


图10 增强模型的置换特征重要性

Fig. 10 Permutation feature importance of the enhanced model

图 11A 显示, 预测位移与数据库实验位移具有较好相关性, 配对体系中 MAE 为 0.339 eV, R^2 为 0.822。这一精度优于绝对 O 1s 结合能预测的 MAE 0.481 eV。图 11B 展示了代表性盐/氧化物配对的实验位移和预测位移, 大多数盐到氧化物的转化表现为负位移, 即氧化物 O 1s 结合能低于对应盐。这与图 2 中含氧酸盐整体位于高结合能区域、简单氧化物整体位于低结合能区域的趋势一致。化学上, 这反映了氧从含氧酸根环境转变为氧化物型晶格氧环境时, 其 O 1s 结合能通常降低。图 11C 的位移误差显示, 多数代表性配对的误差处于 ± 0.5 eV 参考范围内。

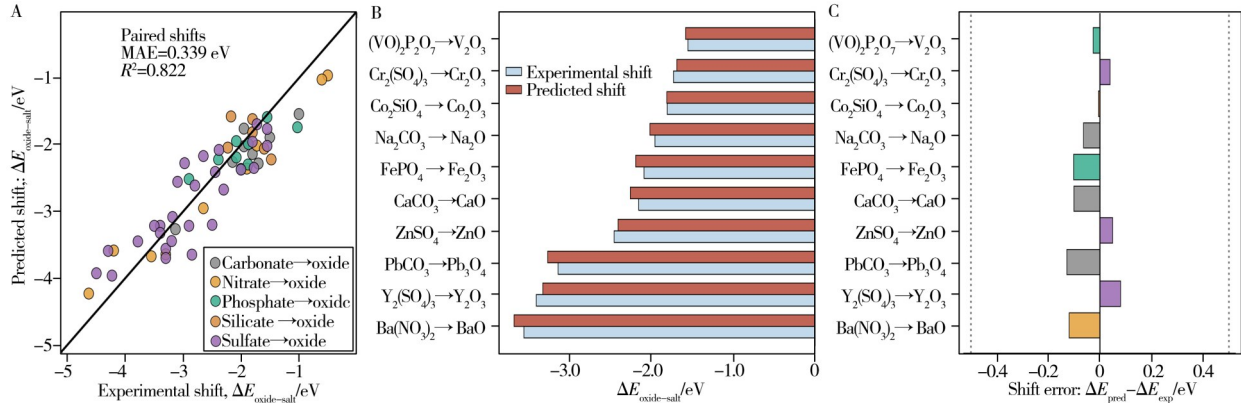


图 11 盐/氧化物配对体系的 O 1s 相对位移预测

Fig. 11 Prediction of O 1s relative shifts in paired salt/oxide systems

A. comparison between experimental and predicted relative shifts; B. experimental and predicted shifts for representative salt/oxide pairs; C. shift-prediction errors for representative pairs (A. 实验相对位移与模型预测相对位移对比; B. 代表性盐/氧化物配对体系的实验位移和预测位移; C. 代表性配对体系的位移预测误差)

2.6 外部实验验证

为评估模型的泛化能力, 选取 $\text{Ca}_{14}\text{Mg}_5\text{M}_4\text{O}_{25}$ ($\text{M}=\text{Al}, \text{Sc}, \text{Y}$) 混合氧化物作为外部验证样品。如图 12A 所示, Al、Sc 和 Y 样品的实验 O 1s 结合能分别为 531.37、531.17、531.12 eV, 呈现出随平均阳离子半径增大而降低的趋势, 这与图 4 中整体数据集的相关性规律一致。该结果说明平均阳离子半径不仅与数据库样本分布相关, 也反映了阳离子尺寸及其相关成键环境对 O 1s 结合能的影响。图 12B 比较了 GB 和 RF 模型对三种外部样品的预测结果。引入 RF 模型的主要目的是考察增强描述符在不同树模型中的适用性。结果显示, 基线模型已能给出接近实验值的峰位范围, 引入增强描述符后预测整体进一步接近实验结果, 其中增强 RF 在 3 种样品上的误差均较小。以 Y 样品为例, 基线 GB 模型误差为 +0.37 eV, 而增强 RF 误差仅为 +0.07 eV, 可见增强描述符显著提升了模型对复杂混合氧化物的预测能力。

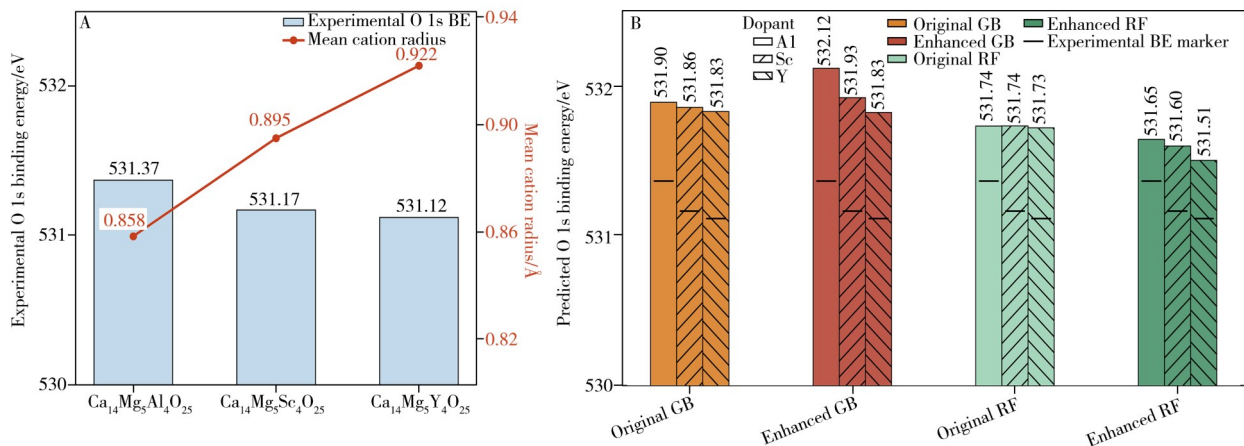


图 12 $\text{Ca}_{14}\text{Mg}_5\text{M}_4\text{O}_{25}$ ($\text{M}=\text{Al}, \text{Sc}, \text{Y}$) 混合氧化物的实验验证

Fig. 12 Experimental validation using $\text{Ca}_{14}\text{Mg}_5\text{M}_4\text{O}_{25}$ ($\text{M}=\text{Al}, \text{Sc}, \text{Y}$) mixed oxides

A. relationship between experimental O 1s binding energies and mean cation radius; B. predicted O 1s binding energies from the baseline GB, enhanced GB, baseline RF, and enhanced RF models (A. 实验 O 1s 结合能与平均阳离子半径的关系; B. 基线 GB、增强 GB、基线 RF 和增强 RF 模型对三种样品的预测结果)

3 结 论

本研究基于 514 个含氧无机化合物的 O 1s XPS 结合能数据，建立了仅以化学式组成描述符为输入的机器学习预测模型。GB 模型能够捕捉 O 1s 结合能与组成特征之间的非线性关系，在原始描述符上取得了 0.505 eV 的测试集 MAE。引入物理启发描述符后，MAE 降至 0.481 eV， R^2 提高至 0.716。相关性分析和置换特征重要性分析表明，平均阳离子半径是影响 O 1s 结合能的关键变量，阳离子场强、过渡金属比例、阳离子半径变异性等特征进一步增强了模型对复杂含氧体系的区分能力。 $\text{Ca}_{14}\text{Mg}_5\text{M}_4\text{O}_{25}$ (M=Al, Sc, Y) 混合氧化物实验显示，O 1s 结合能随平均阳离子半径增大而降低，与数据库相关性和模型解释结果一致，进一步验证了模型解释结果的合理性。

在相对位移预测方面，增强模型在盐/氧化物配对体系中表现更出色，MAE 低至 0.339 eV， R^2 高达 0.822，能够准确反映盐向氧化物转化过程中的峰位变化趋势。本研究构建的组成驱动机器学习模型可实现跨体系 O 1s 结合能及相对位移的有效预测，为含氧无机材料的 XPS 峰位初筛和变化趋势判断提供了低成本、可解释的辅助参考。考虑到化学式描述符难以充分反映局域氧配位环境，未来可进一步引入氧配位数、M—O 键长和键角等结构描述符，并结合结构感知机器学习方法，以提升模型对局域氧化学环境的表征能力和在复杂氧化物体系中的预测可靠性。

参考文献：

- [1] Wang Y, Yue R F. *J. Instrum. Anal.* (王燕, 岳瑞峰. 分析测试学报), **2001**, 20(1): 30–33.
- [2] Pan Y F, Li X J, Yang W C, Li D, Liao H P, Xu Y N. *J. Instrum. Anal.* (潘燕芳, 李晓静, 杨文超, 李东, 廖华平, 许燕娜. 分析测试学报), **2025**, 44(5): 956–967.
- [3] Li X L, Xie F Y, Gong L, Zhang W H, Yu X L, Chen J. *J. Instrum. Anal.* (李晓莉, 谢方艳, 龚力, 张卫红, 于晓龙, 陈建. 分析测试学报), **2013**, 32(5): 535–540.
- [4] Feng B, Chen J Y, Cheng M. *J. Instrum. Anal.* (冯波, 陈静允, 程萌. 分析测试学报), **2010**, 44(11): 2408–2413.
- [5] Chen J Y, Huang W. *J. Instrum. Anal.* (陈晋阳, 黄卫. 分析测试学报), **2002**, 3: 70–72.
- [6] Merino N A, Barbero B P, Eloy P, Cadús L E. *Appl. Surf. Sci.*, **2006**, 253(3): 1489–1493.
- [7] Wu Z. *J. Instrum. Anal.* (吴正龙. 分析测试学报), **2005**, 24(3): 45–47.
- [8] Dupin J C, Gonbeau D, Vinatier P, Levasseur A. *Phys. Chem. Chem. Phys.*, **2000**, 2(6): 1319–1324.
- [9] Bagus P S, Illas F, Pacchioni G, Parmigiani F. *J. Electron Spectrosc. Relat. Phenom.*, **1999**, 100(1/2/3): 215–236.
- [10] Zhang S, Man H Z, Guo C, Zhao Y B, Li Z P. *J. Instrum. Anal.* (张硕, 满瀚泽, 郭冲, 赵雅彬, 李展平. 分析测试学报), **2025**, 44(6): 1208–1218, 1226.
- [11] Huang J J, Xiang W X, Xia J Z, Pan J B. *J. Instrum. Anal.* (黄娟娟, 向万欣, 夏婧竹, 潘建斌. 分析测试学报), **2025**, 44(12): 2526–2534.
- [12] Butler K T, Davies D W, Cartwright H, Isayev O, Walsh A. *Nature*, **2018**, 559(7715): 547–555.
- [13] Vakhrushev A A, Matveev A V, Nartova A V. *Kinet. Catal.*, **2024**, 65(6): 788–796.
- [14] Greczynski G, Hultman L. *Prog. Mater. Sci.*, **2020**, 107: 100591.
- [15] Schmid M, Steinrück H P, Gottfried J M. *Surf. Interface Anal.*, **2014**, 46(8): 505–511.
- [16] Paruzzo F M, Hofstetter A, Musil F, De S, Ceriotti M, Emsley L. *Nat. Commun.*, **2018**, 9: 4501.
- [17] Lee A Y, Powell C J, Gorham J M, Morey A, Scott J H J, Hanisch R J. *Data Sci. J.*, **2024**, 23: 45.

(责任编辑：盛文彦)