

Interpretable Machine Learning Reveals Atom-Resolved VOCs Adsorption and Reactivity on Metal Oxides

Yanbing Guo

guoyanbing@mail.ccnu.edu.cn

Central China Normal University <https://orcid.org/0000-0002-5399-1739>

Biluan Zhang

Central China Normal University

Pan Yan

Wuhan Institute of Photochemistry and Technology

Lu Yan

Central China Normal University

Dongbo Zhou

Central China Normal University

Jinlong Wang

Central China Normal University

Article

Keywords: Interpretable machine learning; Dual-Tower framework; Transfer learning; Volatile organic compounds (VOCs)

Posted Date: October 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11192911/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Interpretable Machine Learning Reveals Atom-Resolved VOCs Adsorption and Reactivity on Metal Oxides

Biluan Zhang^{1,2}, Pan Yan³, Lu Yan^{1,2}, Dongbo Zhou^{4*}, Jinlong Wang^{1,2,3*}, Yanbing Guo^{1,2,3*}

¹*Institute of Environmental and Applied Chemistry, College of Chemistry, Central China Normal University, Wuhan 430079, P. R. China*

²*Engineering Research Center of Photoenergy Utilization for Pollution Control and Carbon Reduction, Ministry of Education, Central China Normal University, Wuhan 430079, P. R. China*

³*Wuhan Institute of Photochemistry and Technology, Wuhan 430083, P. R. China*

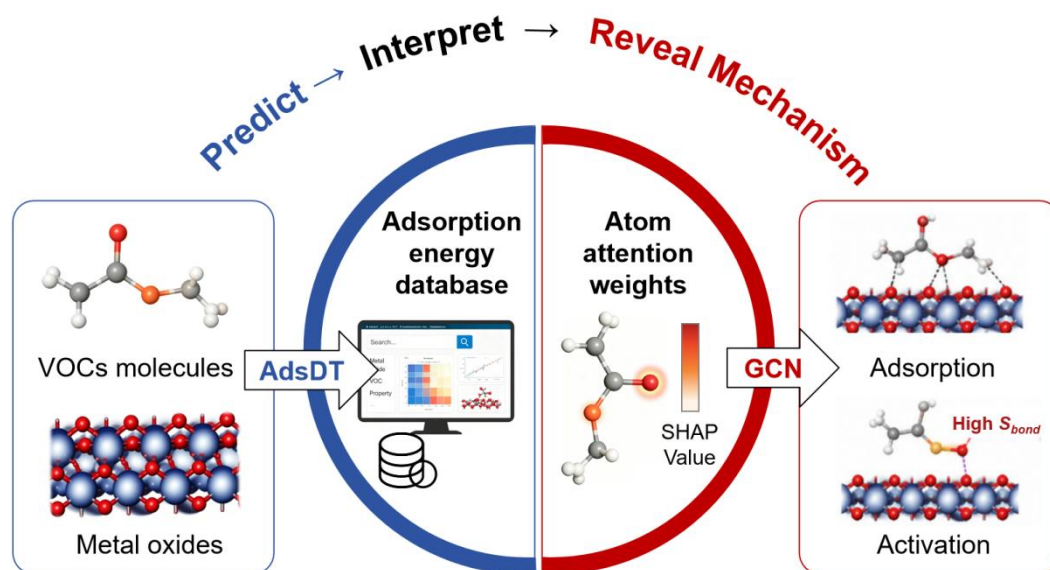
⁴*Laboratory for Artificial Intelligence and New Forms of Education, Central China Normal University, Wuhan 430079, China*

*E-mail: zhoudongbo@mail.ccnu.edu.cn; wjl2022@ccnu.edu.cn; guoyanbing@mail.ccnu.edu.cn

Abstract

The adsorption and reactivity of volatile organic compounds (VOCs) on metal oxides govern catalytic reaction pathways, yet predictive rules remain elusive because molecular functionality, surface chemistry, and adsorption geometry are strongly coupled. Here we report a transferable adsorption dual-tower framework (AdsDT) that independently encodes molecular and crystal features to predict adsorption energies directly from unrelaxed CIF structures. By pretraining on the Open Catalyst 2020 dataset and progressively transferring through intermediate adsorption systems, AdsDT enables rapid construction of metal oxides-VOCs interaction databases and achieves strong agreement with density functional theory calculations ($R^2 = 0.85$). Using 816 TiO_2 (anatase, 101)-VOCs adsorption configurations, we combine molecular fingerprints, SHAP analysis, and graph attention mapping to identify fragment-level and atom-resolved adsorption contribution. Saturated alkyl carbon environments increase predicted adsorption energies, whereas unsaturated and electronegative atoms, particularly carbonyl oxygen, favor stronger adsorption. Extension to eight representative metal oxides reveals material-dependent and functional-group correlation. Furthermore, a bond activation score integrating atomic attention, bond dissociation energy, and bond polarization links adsorption configurations to reaction pathways, showing that adsorption sites activate neighboring bonds through electronic redistribution rather than serving simply as cleavage centers. These results provide transferable atom-resolved descriptors for VOCs adsorption and catalytic reactivity across metal oxides.

Keywords: Interpretable machine learning; Dual-Tower framework; Transfer learning; Volatile organic compounds (VOCs)



42

43

Introduction

Volatile organic compounds (VOCs) are among the major atmospheric pollutants, posing serious threats to environmental sustainability and human health¹⁻³. Catalytic oxidation is an effective technology for VOCs elimination⁴⁻⁷, yet the rational design of efficient catalysts requires a fundamental understanding of the elementary processes governing VOCs conversion, particularly the initial adsorption and activation of reactant molecules on catalytic surfaces^{8,9}. Metal oxides are widely used as catalytic substrates owing to their tunable electronic structures, diverse surface coordination environments, and abundant defect chemistry¹⁰⁻¹³. The interactions between surface metal sites and VOCs adsorption centers determine adsorption configurations, surface coverage, and interfacial charge redistribution, thereby regulating subsequent bond activation and reaction pathways^{14,15}. Nevertheless, VOCs adsorption on metal oxide surfaces remains complex and strongly site-dependent. Therefore, establishing atom-resolved adsorption principles and identifying transferable molecular descriptors that govern VOCs and metal oxides interactions are critical for understanding and predicting catalytic oxidation processes^{16,17}.

Systematically understanding adsorption behavior of VOCs requires comprehensive descriptions of the adsorption spanning diverse molecular structures, metal oxide surfaces, and adsorption configurations¹⁸. Such datasets are essential for uncovering transferable chemical principles, including functional-group-dependent surface affinity and the molecular factors governing adsorption strength and subsequent activation pathways^{19,20}. Conventional density functional theory (DFT) approaches require extensive sampling of surface adsorption sites, molecular orientations, and structural relaxation for each adsorbate-surface combination, leading to prohibitive computational costs when exploring the vast chemical space^{21,22}. Although descriptor-based scaling relationships, such as oxygen charge descriptors for oxygenated molecule adsorption on MO₂ surfaces, have provided useful insights^{23,24}, its general applicability remains limited because adsorption energetics arise from the coupled effects of molecular functionality, surface coordination environments and interfacial electronic

interactions^{25,26}. Therefore, a scalable strategy capable of efficiently expanding adsorption databases is urgently needed.

Recent advances in artificial intelligence have provided new opportunities for accelerating molecular simulations by learning complex structure-property relationships from large-scale datasets²⁷⁻²⁹. Various machine learning (ML) models, such as random forests (RF)³⁰, multilayer perceptions (MLP)³¹, and graph neural networks (GNNs)^{32,33}, have been extensively explored to predict the adsorption energy of adsorbate-surface systems. However, most existing ML models remain highly dependent on explicitly defined and pre-relaxed three-dimensional adsorption configurations, requiring prior knowledge of adsorption geometries that are typically obtained through computationally expensive structural optimization^{34, 35}. To overcome these limitations, *Ulissi* and co-workers proposed the AdsorbML workflow, which heuristic structure searching with machine learning potentials to accelerate the calculation³⁶. Similarly, *Margraf et al.* proposed a global optimization protocol that employs on-the-fly ML potentials iteratively refined by DFT calculations, enabling efficient identification of stable adsorption configurations across diverse surface-adsorbate combinations^{37,38}. Nevertheless, these approaches still rely on extensive exploration of initial adsorption configurations and iterative DFT calculations, resulting in considerable computational costs for training data generation and adsorption geometry determination³⁹. Therefore, current ML strategies have not yet fully addressed the challenges of scalable adsorption database construction and transferable descriptor discovery for complex metal oxides-VOCs interaction systems.

In this work, we address these challenges by uncovering transferable atom-resolved principles that govern VOCs adsorption and activation on metal oxides surfaces through an interpretable artificial intelligence framework. We first develop a transferable dual-tower (AdsDT) architecture that independently encodes molecular and crystal structures and subsequently learns their interfacial interactions, enabling geometry-free adsorption energy prediction through progressive transfer learning from the Open Catalyst 2020 dataset to complex metal oxides-VOCs interaction systems (**Figure 1a**). We further establish an open high-throughput prediction platform for rapid

construction of a comprehensive metal oxides-VOCs adsorption database, providing the data foundation for large-scale knowledge discovery (**Figure 1b**). Building on this database, we integrate molecular fingerprints, SHAP analysis, and graph attention mapping into an interpretable learning framework that reveals fragment-level and atom-resolved descriptors governing VOCs adsorption across chemically diverse metal oxides (**Figure 1c**). Finally, we extend analysis across multiple representative metal oxides, revealing material-dependent responses. More importantly, we propose a Bond Activation Score (S_{bond}) by integrating atomic attention weights, bond dissociation energy, and bond polarization characteristics, thereby linking adsorption behavior with catalytic reactivity (**Figure 1d**). Overall, this work yields transferable atom-resolved descriptors and universal adsorption-activation principles, providing a unified atomistic framework for VOCs adsorption and catalytic reactivity across diverse metal oxide surfaces.

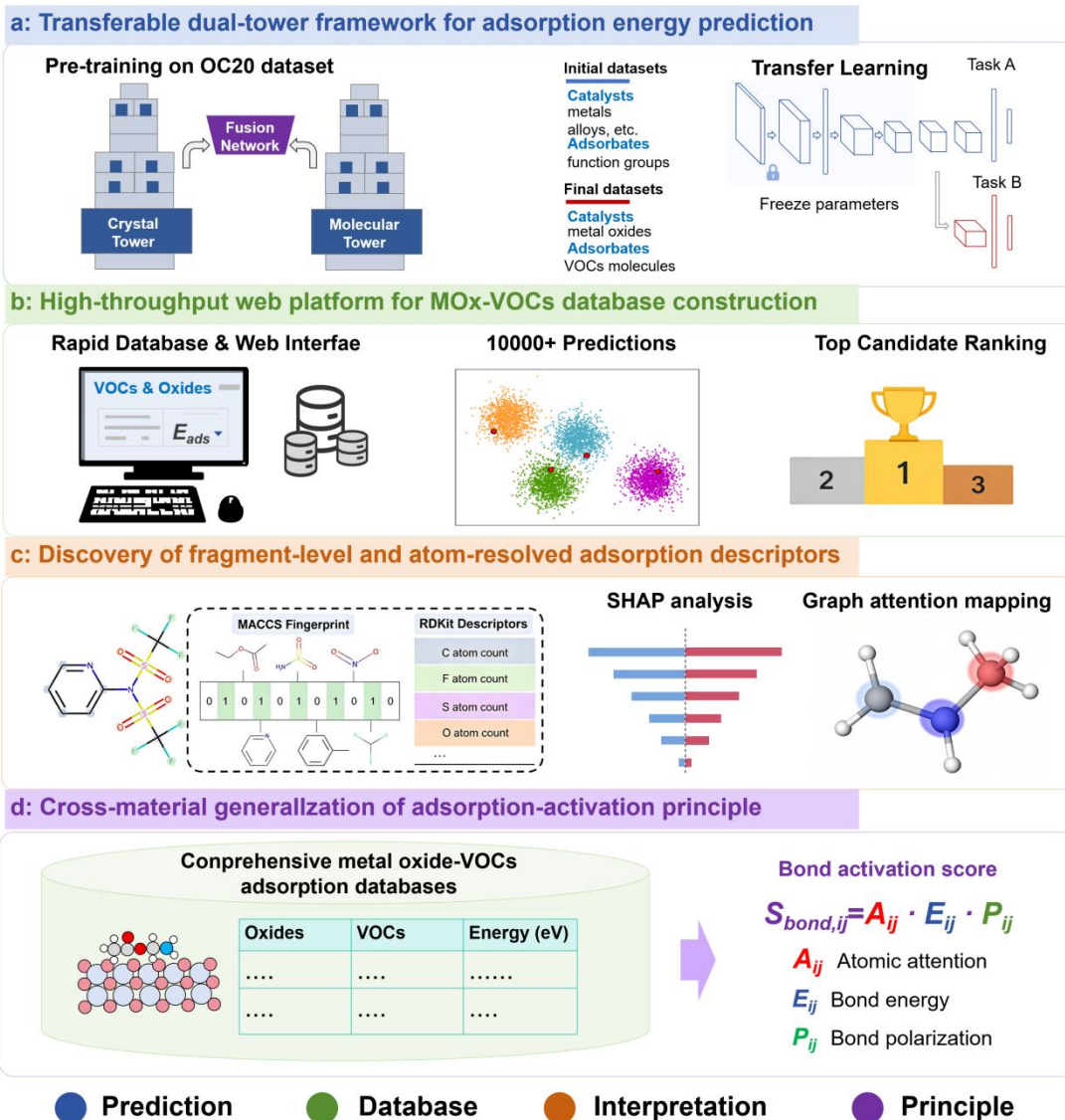


Figure 1. Workflow of the transferable and interpretable AI framework for VOCs adsorption and activation on metal oxide surfaces. (a) Transferable dual-tower architecture (AdsDT) for geometry-free adsorption energy prediction via progressive transfer learning from OC20 to metal oxides-VOCs systems. (b) High-throughput web platform for rapid construction of comprehensive metal oxide-VOCs adsorption database. (c) Data-driven discovery of fragment-level and atom-resolved adsorption descriptors. Interpretable machine learning workflow integrating molecular fingerprints, SHAP analysis, and graph attention mapping. (d) Generalization of transferable adsorption-activation principles through cross-material attention analyses and bond activation score.

Results

AdsDT architecture and performance for adsorption energy prediction

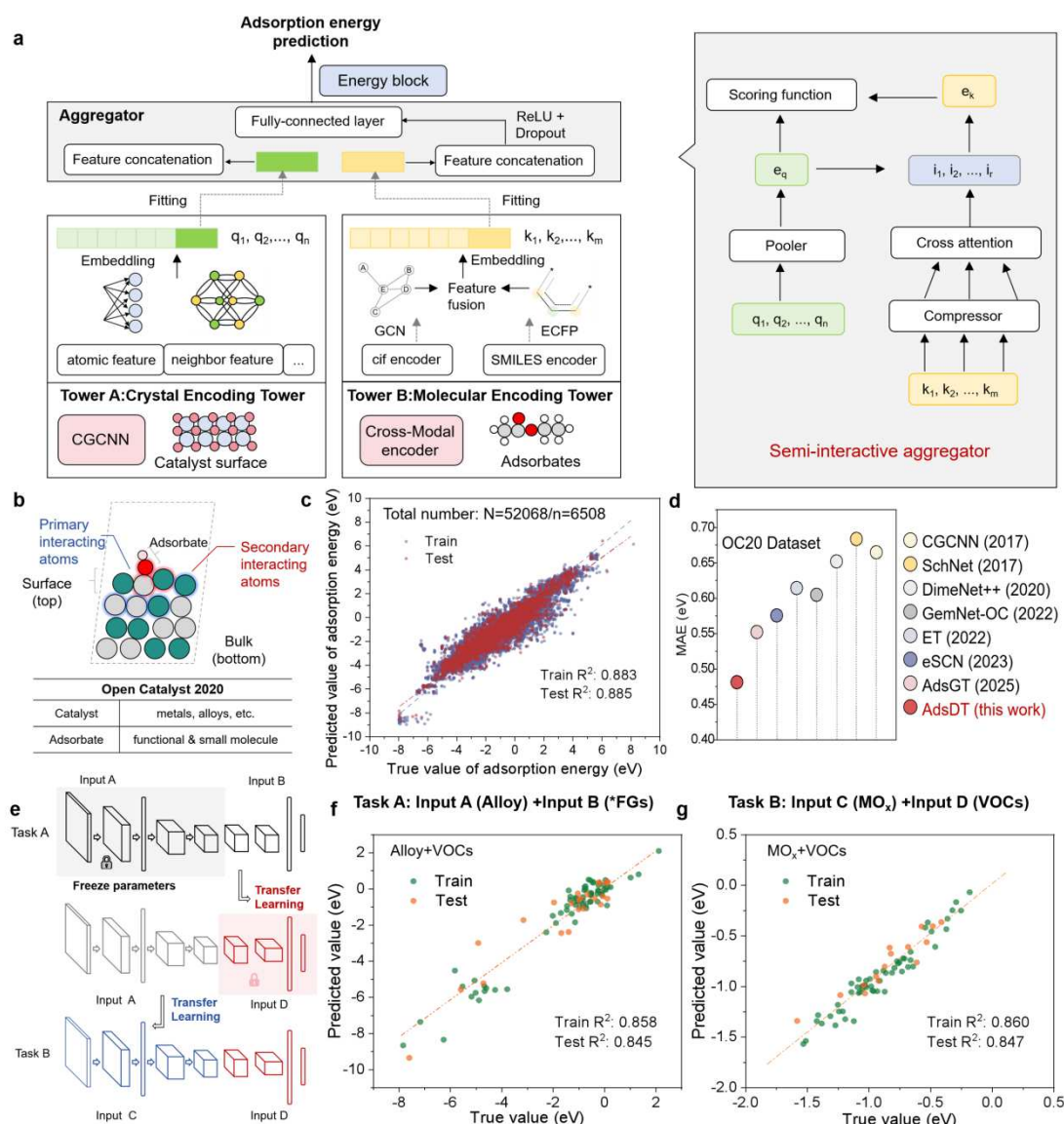


Figure 2. Architecture and performance of the transferable dual-tower AdsDT framework. (a) Overall architecture of the AdsDT framework. The framework consists of a crystal encoding tower (Tower A), a molecular encoding tower (Tower B), and a semi-interactive aggregator (b) Schematic illustration of adsorption systems in the Open Catalyst 2020 (OC20) dataset, highlighting catalyst surfaces, adsorbates, and primary/secondary interacting atoms used for model pretraining. (c) Parity plot comparing AdsDT-predicted adsorption energies with true values for the OC20 training and test sets. (d) Comparison of mean absolute error (MAE) with representative baseline models on the OC20 dataset. (e) Schematic illustration of the progressive transfer learning strategy. Task A transfers knowledge from alloy-functional group (FG) systems (Input A: alloy surfaces; Input B: FG molecules) to alloy-VOCs systems; Task B for transfer to metal oxides-VOCs systems (Input C: MO_x surfaces; Input D: VOCs molecules). Frozen blocks denote encoder parameters retained during fine-tuning. Parity plots comparing the AdsDT-predicted adsorption energies against with

DFT values for the downstream (f) Alloy-VOCs and, (g) metal oxides-VOCs target datasets.

We developed AdsDT, a transferable dual-tower deep learning framework that enables geometry-independent adsorption energy prediction by learning intrinsic representations of catalyst surfaces and adsorbate molecules. As illustrated in **Figure 2a**, AdsDT consists of three interconnected components: a crystal encoding tower (Tower A), a molecular encoding tower (Tower B), and a semi-interactive aggregator. The crystal encoder employs a crystal graph convolutional neural network (Crystal Graph Convolutional Neural Network, CGCNN)^{40,41} to extract structural representations from periodic catalyst surfaces, including atomic features and local coordination environments. Meanwhile, the molecular encoder utilizes a graph convolutional network (Graph Convolutional Network, GCN) to characterize adsorbate structures⁴². The detailed internal architecture of AdsDT and the corresponding feature propagation process are illustrated in **Figure S1 and Table S1**. Unlike conventional approaches that directly concatenate surface and molecular descriptors (**Figure S2**), AdsDT introduces a semi-interactive aggregation module based on cross-attention⁴³, enabling dynamic information exchange between catalyst and adsorbate representations (Aggregator). This hierarchical interaction strategy allows the model to capture adsorption-dependent interfacial features and establish nonlinear correlations between local chemical environments and adsorption energetics.

We first evaluated the predictive capability of AdsDT using the Open Catalyst 2020 (OC20) dataset, which contains a wide range of catalyst surfaces and adsorbate systems (**Figure 2b**). Since adsorption behavior is primarily governed by interactions between active surface atoms and adsorbed species, AdsDT incorporates both primary interacting atoms and their neighboring coordination environments into graph representations. The OC20-Dense dataset contains 58,576 adsorption configurations, including 52,068 samples for training and 6,508 samples for independent testing. As shown in **Figure 2c and Figure S3**, the predicted adsorption energies exhibit excellent agreement with DFT calculations, achieving R^2 values of 0.883 and 0.885 for the training and test sets, demonstrating excellent predictive accuracy and robust generalization ability. Furthermore, noise perturbation tests demonstrate the robustness

of AdsDT, with R^2 values remaining above 0.80 even under 25% noise, indicating stable adsorption representation learning against data uncertainties (**Figure S4**). Benchmark comparisons with representative state-of-the-art models further demonstrate the advantages of AdsDT (**Figure 2d, Table S2-S3**). Supported by these supporting results on model input requirements and performance of different graph encoder backbones, conventional structure-based models, including CGCNN and SchNet, exhibit relatively large prediction errors with MAEs exceeding 0.65 eV. Advanced multimodal models such as AdsMT improve prediction accuracy with an MAE of approximately 0.55 eV. In comparison, AdsDT achieves an MAE of approximately 0.48 eV on the OC20 benchmark, demonstrating enhanced capability in capturing complex catalyst-adsorbate interactions.

To extend adsorption prediction toward metal oxides-VOCs systems where high-quality training data remain limited, we constructed two customized DFT datasets covering Alloys-VOCs and metal oxides-VOCs systems for progressive transfer learning strategy (**Figure 2e, Table S4-S5**). Instead of directly optimizing models on small target datasets, AdsDT first learns general adsorption representations from chemically diverse alloy-functional group systems. During the first transfer stage, the pretrained surface encoder is retained while the molecular encoder is adapted toward Alloys-VOCs systems. Subsequently, the optimized molecular representation is transferred to metal oxides-VOCs systems and integrated with a newly initialized metal oxide surface encoder. This sequential transfer strategy enables efficient knowledge migration between chemically distinct catalytic domains. The effectiveness of this strategy was evaluated using downstream Alloys-VOCs and metal oxides-VOCs datasets. For the Alloys-VOCs system, AdsDT achieved R^2 values of 0.858 and 0.845 for training and testing sets, respectively (**Figure 2f**). For the more challenging metal oxides-VOCs system, the transferred model maintained comparable performance, achieving R^2 values of 0.860 and 0.847 for training and testing sets (**Figure 2g**). The predicted adsorption energies remain closely distributed around the ideal fitting line across the entire energy range, confirming the reliability and stability of the transferred representations.

Comprehensive ablation studies further confirm the importance of the proposed transfer pathway (**Figure S5-S6, Table S6**). Direct fine-tuning of all model parameters leads to reduced testing performance, suggesting that excessive adaptation to limited target data compromises generalization. In contrast, progressive transfer learning preserves transferable adsorption knowledge learned from the source domain while enabling adaptation to new surface chemistries. The improved transferability stems from the rich chemical diversity of alloy-based source systems, which span wide-ranging elemental compositions, metal coordination environments, and molecular adsorption configurations (**Table S7**). Detailed adsorbate statistics for OC20-Dense, Alloys-VOCs, and metal oxides-VOCs are documented in **Table S8-S9**, with their chemical-family distributions are summarized in **Table S10-S11**. Chemical-space analyses combining elemental-distribution mapping (**Figure S7-S9**) and UMAP dimensionality-reduction visualization (**Figure S10**) identify considerable feature overlap across the Alloys-VOCs, and metal oxides-VOCs systems. This observation validates the feasibility of knowledge transfer between these distinct catalytic domains. Collectively, these findings demonstrate that AdsDT functions as a transferable representation-learning framework to unify chemically heterogeneous adsorption systems. It lays a solid basis for large-scale adsorption database development and the subsequent extraction of atom-level adsorption principles.

221 Web platform and TiO₂(anatase, 101)-VOCs adsorption databases

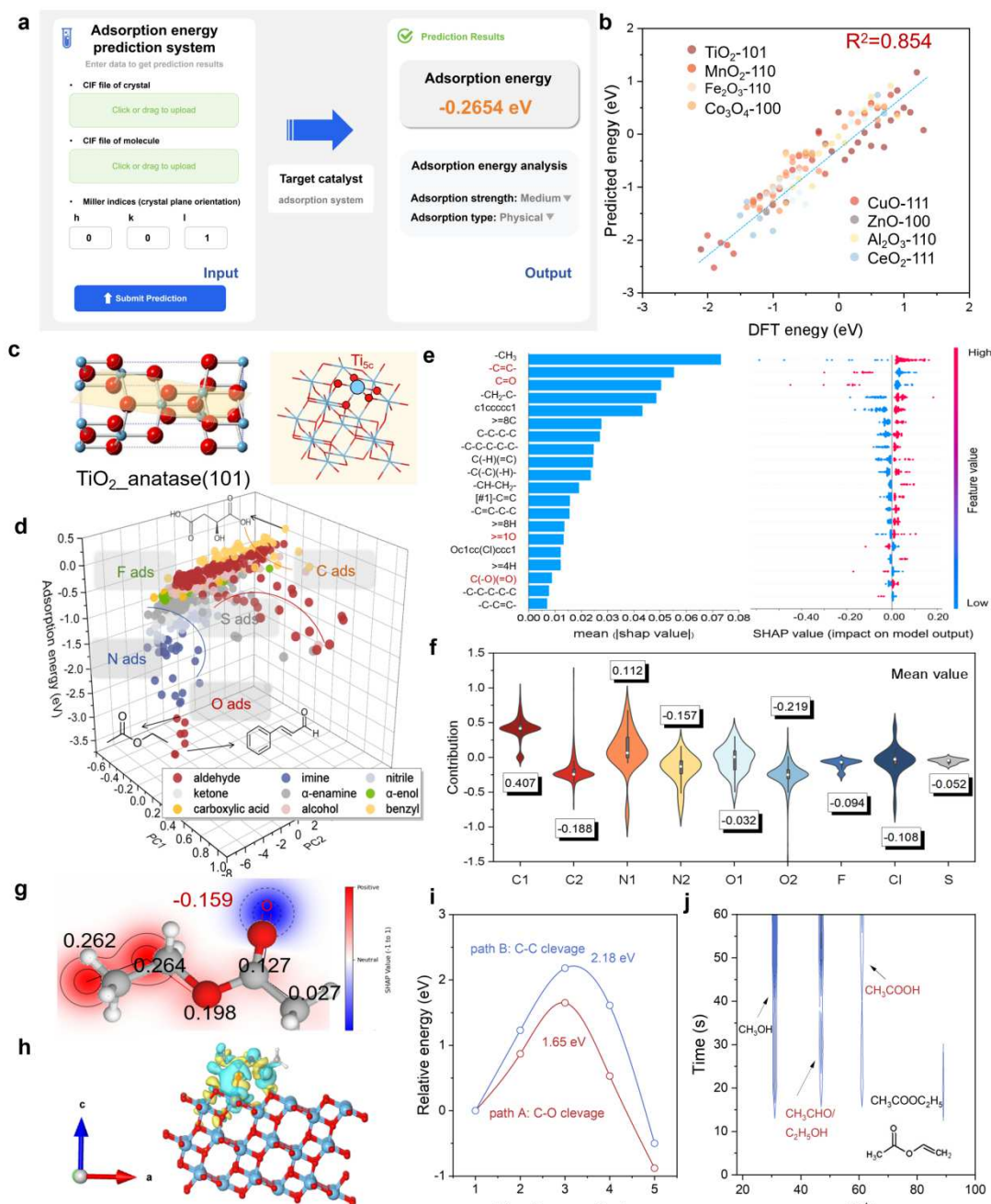


Figure 3. High-throughput web platform for database construction and data-driven discovery of fragment-level and atom-resolved adsorption descriptors. (a) Schematic illustration of the AdsDT web-based adsorption energy prediction platform. (b) Parity plot comparing AdsDT-predicted adsorption energies with DFT-calculated values across eight representative metal oxide surfaces. (c) Crystal structure configurations of anatase used for constructing the TiO₂ (anatase, 101)-VOCs adsorption databases. (d) Three-dimensional adsorption energy landscape of the TiO₂ (anatase, 101)-VOCs databases, illustrating the distribution of adsorption strengths associated with different molecular functional groups and chemical environments. (e) SHAP analysis based on MACCS molecular fingerprints to adsorption energy prediction. (f) Violin plots illustrating the statistical distribution of contributions from different atomic sites. (g) Atomic attention weight

visualization for ethyl acetate adsorption on TiO₂. (h) Charge density difference of ethyl acetate adsorption on TiO₂, where yellow and blue isosurfaces denote electron accumulation and depletion, respectively. (i) Reaction pathways comparing alternative bond cleavage mechanisms of ethyl acetate. (j) The time-of-flight mass (TOF-MS) evolution spectra of intermediates during ethyl acetate oxidation over TiO₂ catalyst.

To enable large-scale investigation of intrinsic metal oxides-VOCs interfacial interactions, we built an online platform for adsorption energy prediction based on the AdsDT model (**Figure 3a**). Users can directly upload unrelaxed crystallographic information (CIF) files of catalyst crystal structures and VOCs molecular structures, and corresponding surface orientations, after which the platform automatically performs structural encoding and outputs adsorption energies and adsorption characteristics. Compared with conventional DFT calculations, this platform accelerates adsorption energy evaluation by approximately five orders of magnitude (**Figure S11**), enabling rapid exploration of large adsorption chemical spaces. The predictive reliability of AdsDT was further validated across diverse metal oxide surfaces (**Figure 3b and Table S12**). The predicted adsorption energies show excellent agreement with DFT calculations, achieving an MAE of 0.095 eV and an RMSD of 0.118 eV (**Figure S12**). Representative oxides, including TiO₂, MnO₂, Fe₂O₃, Co₃O₄, CuO, ZnO, Al₂O₃, and CeO₂, all exhibit near-linear correlations between predicted and calculated adsorption energies, demonstrating that AdsDT effectively captures transferable metal oxides-VOCs interaction patterns across chemically distinct oxides systems.

Given its well-defined surface structure and abundant coordinatively unsaturated Ti sites (Ti_{5c}), anatase TiO₂ (101) was selected as a representative metal oxide surface for systematic investigation of VOCs adsorption (**Figure 3c**). The exposed Ti_{5c} sites act as primary Lewis acidic centers and play a critical role in coordinating VOCs molecules during adsorption⁴⁴. Leveraging the efficient prediction capability of AdsDT, we subsequently constructed a TiO₂ (anatase, 101)-VOCs adsorption datasets containing 816 adsorption configurations (**Figure S13 and Table S13**). This database covers diverse VOCs families, including aldehydes, ketones, alcohols, carboxylic acids, nitriles, imines, and aromatic compounds. Principal component analysis (PCA)-based

chemical space visualization reveals distinct adsorption distributions associated with different functional groups (**Figure 3d**). Carbon-, fluorine-, and sulfur-centered adsorption species are mainly distributed in weaker adsorption regions (approximately 0 to -1.0 eV), whereas oxygen- and nitrogen-containing adsorption configurations predominantly occupy regions with stronger adsorption energies (-1.5 to -3.5 eV). Particularly, O-centered adsorption exhibits the broadest energy distribution, suggesting that oxygen coordination not only enhances adsorption stability but also induces diverse interfacial electronic interactions.

To uncover the molecular origins governing VOCs adsorption, we established a hierarchical interpretable learning strategy combining MACCS fingerprint-based SHAP analysis and graph attention mapping. Among the eight classical machine learning regression models, XGB yields the best prediction performance for adsorption energy (**Figure S14**). The MACCS-SHAP analysis identifies fragment-level descriptors, revealing that saturated alkyl groups (-CH₃ and -CH₂-) contribute the largest fraction of feature importance (approximately 55-70%), whereas unsaturated C=C bonds and carbonyl groups exhibit negative contributions, indicating their tendency to stabilize adsorption configurations (**Figure 3e**). At the atomic scale, we performed weight analysis based on our cross-attention GCN framework to quantify the contribution of individual atoms within VOCs molecules. As illustrated in **Figure S15**, the multi-head cross-attention layer computes pairwise attention scores between adsorbate atoms and surface atoms. The averaged attention weights are projected onto the material surface to produce atom-wise importance mappings, enabling systematic identification of the critical sites that govern VOC adsorption behavior. Averaged over 816 VOC adsorption configurations on anatase TiO₂ (101), **Figure 3f and Table S14** resolve the distinct contribution patterns of heteroatoms with different bonding environments. Saturated single-bond carbon atoms (C1) exhibit the strongest positive contribution (+0.404). The positive contributions originate mainly from weak van der Waals interactions and geometric confinement effects at the oxide surface⁴⁵. In contrast, unsaturated carbon atoms in C=C bonds (C2) exhibit negative contributions (-0.188) due to π -electron interactions with oxide surfaces, indicating distinct roles of saturated

and unsaturated carbon environments⁴⁶⁻⁴⁷. For nitrogen-containing VOCs, atomic contributions also strongly depend on the local bonding environment. Single-bond nitrogen atoms (N1, +0.112) generally show positive contributions due to steric effects and weaker surface interactions, whereas nitro groups (N2, -0.157) preferentially interact through oxygen atoms because of resonance-induced charge delocalization⁴⁸. Single-bond oxygen atoms (O1, -0.032) in alcohols and ethers contribute negatively through polar interactions and surface coordination, while carbonyl oxygen atoms (O2) in aldehydes, ketones, acids, and esters exhibit the strongest negative contribution (-0.219), originating from direct coordination between oxygen lone-pair electrons and Lewis acidic metal sites⁴⁹. The details of visualized attention weight maps for representative VOCs structures are provided in **Figure S16-S23**. More evidences are given in **Figure S24**. Oxygen attention values across different oxygen-containing VOCs exhibit consistent correlations with adsorption energies and oxygen charge loss. Increasing attention values are accompanied by stronger adsorption and enhanced oxygen charge redistribution, demonstrating that the attention mechanism captures chemically meaningful electronic features rather than arbitrary model weights. Additionally, halogen contributions follow a F-Cl-Br trend governed by the balance between electronegativity, polarizability, and steric effects⁵⁰. Multifunctional VOCs exhibit competitive interactions among different reactive motifs, where the dominant adsorption center controls molecular activation.

To further determine whether the atom-resolved adsorption descriptors identified by AdsDT are associated with catalytic activation, ethyl acetate adsorption on TiO₂ was selected as a representative system. The atomic attention map reveals that the carbonyl oxygen exhibits the highest adsorption-related contribution (-0.159), identifying it as the dominant adsorption-contributing region predicted by the model, whereas the neighboring carbon framework displays opposite contribution trends, suggesting distinct roles of different atomic environments in regulating adsorption behavior (**Figure 3g**, **Figure S25**). Charge density difference analysis further confirms pronounced interfacial electron redistribution between carbonyl oxygen and Ti_{5c} sites, as evidenced by electron accumulation and depletion regions around the adsorption

interface (**Figure 3h**). Moreover, density of states (DOS) analysis reveals strong orbital coupling between Ti states and the carbonyl group near the Fermi level (**Figure S26**), indicating that Ti-O=C coordination induces localized electronic polarization within the adsorbed molecule. Importantly, the preferential adsorption center does not necessarily represent the direct bond cleavage site. Instead, adsorption-induced electronic redistribution associated with C=O extends toward neighboring chemical bonds, modifying their local electronic environment and facilitating selective activation. To verify this hypothesis, nudged elastic band (NEB) calculations were performed to compare possible decomposition pathways of ethyl acetate on TiO₂ (**Figure 3i, Figure S27**). The calculated barrier for C-O bond cleavage is 1.65 eV, which is substantially lower than that for C-C bond cleavage (2.18 eV), demonstrating that carbonyl-directed adsorption preferentially promotes adjacent C-O bond activation. To experimentally validate this adsorption-activation relationship, anatase TiO₂ with exposed (101) facets was synthesized and structurally characterized by X-ray diffraction (XRD) and high-resolution transmission electron microscopy (HRTEM), confirming the formation of anatase TiO₂ and the exposed (101) facets (**Figures S28-S29**). The resulting TiO₂ was further investigated using TOF-MS and *in situ* DRIFTS measurements (**Figure 3j, Figure S30 and Table S15**). During ethyl acetate oxidation, characteristic intermediates including ethanol and acetic acid were detected, supporting a reaction sequence involving carbonyl adsorption followed by neighboring C-O bond cleavage. Together, these computational and experimental results demonstrate that atomic attention weight provide not only adsorption-site identification but also mechanistic insight into how adsorption-induced electronic redistribution governs catalytic bond activation.

349 Cross-material generalization of adsorption-activation principles

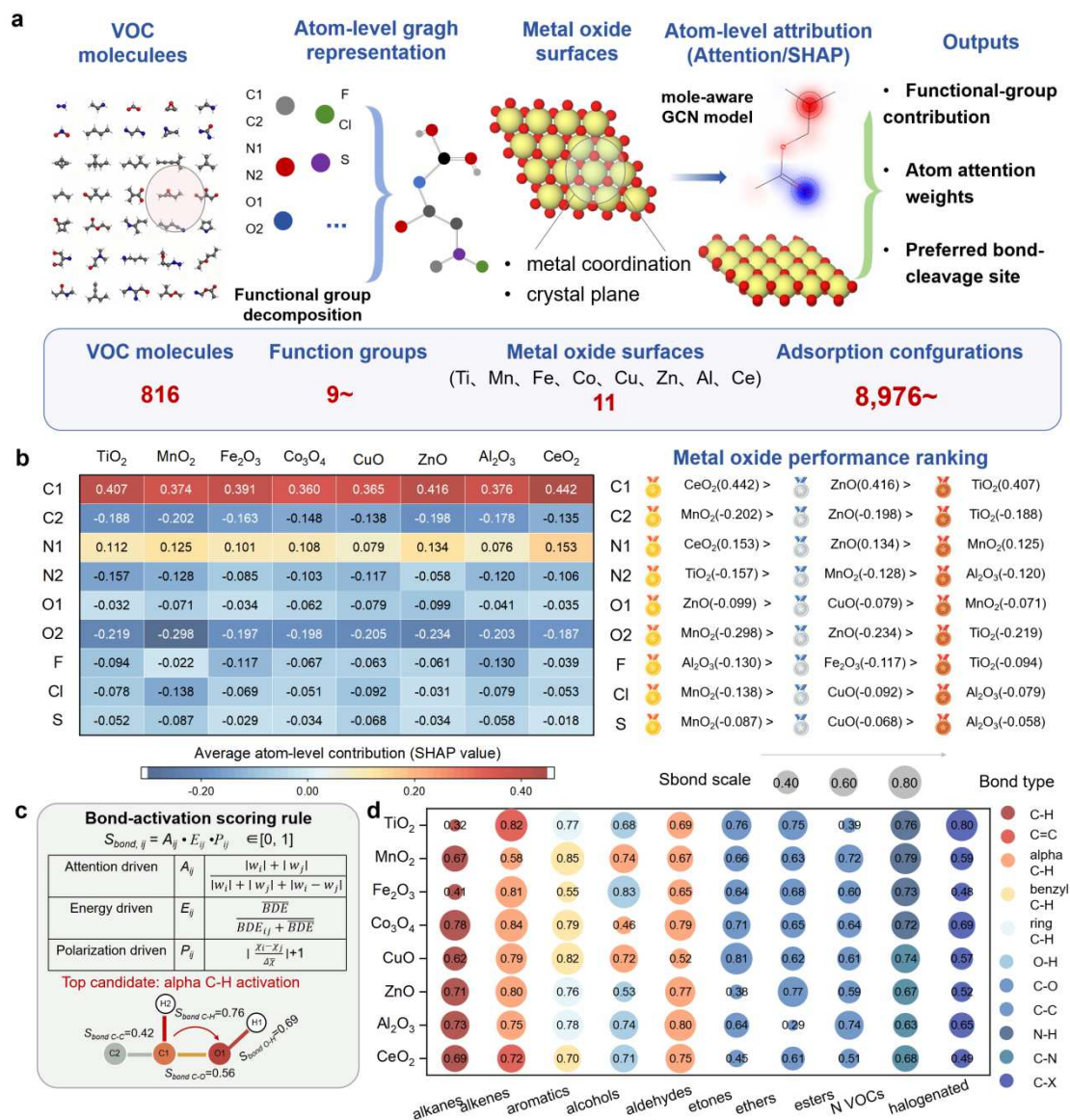


Figure 4. Molecule-aware graph attention framework for interpretable prediction of VOCs bond activation on metal oxides. (a) Workflow of the molecule-aware graph attention network (GCN) for interpretable prediction of VOCs adsorption on metal oxide surfaces. (b) Average atom-level SHAP values of representative atomic species on different metal oxide surfaces and the corresponding performance ranking of metal oxides for each atom type. (c) Bond-activation scoring strategy integrating attention weights, bond dissociation energy (BDE) and electronegativity (χ) to identify the most favorable bond-activation site. (d) Metal oxides-VOCs bond activation response map predicted by the proposed framework. Bubble color denotes the dominant activated bond type, whereas bubble size represents the bond-activation score.

To further extend the research framework from molecular to surface atomic scales and uncover transferable principles governing metal oxides-VOCs interactions, we constructed a comprehensive dataset comprising 816 VOCs molecules spanning nine

representative functional groups, 11 distinct crystal facets of eight metal oxides, and 8,976 heterogeneous adsorption configurations (**Figure 4a and Supplementary Data1**). The selected metal oxides surfaces cover broad ranges of surface-metal coordination numbers (3-6) and acid-base properties (**Figure S31 and Table S16**). VOCs molecules were characterized using atom-level descriptors derived from MACCS fingerprints and RDKit molecular features, including atomic types and local chemical environments such as C1, C2, N1, N2, O1 and O2. Meanwhile, the metal oxides surfaces were encoded according to their atomic coordination environments and crystal structures. These molecular and surface representations were subsequently converted into unified atomic-level graphs and used to construct the adsorption configurations for training a molecule-aware graph convolutional network (GCN)⁵¹. By integrating attention mechanisms with SHAP analysis, the framework quantitatively resolves the contributions of molecular motifs and surface atomic environments to adsorption, providing atom-level contribution weights and enabling subsequent identification of preferential bond-activation sites. The GCN achieved reliable adsorption-energy prediction across all eight metal oxides systems, with test R^2 values of 0.867-0.950 and MAE of 0.027-0.054 eV (**Figure S32 and Table S17**). This framework therefore establishes a unified connection from molecular and surface structural representations to adsorption prediction and microscopic mechanism interpretation.

The atomic-attention heatmap and statistical ranking reveal a largely conserved contribution pattern across different metal oxides surfaces (**Figure 4b and Table S18**). Although the magnitude of each contribution varies with surface chemistry, its sign is generally preserved, suggesting transferable structure-adsorption relationships. Saturated carbon (C1) consistently exhibits positive attention weights (+0.360 to +0.442), indicating that saturated carbon motifs generally weaken surface-VOCs interactions and lead to less-negative adsorption energies. This effect is most pronounced on CeO₂ (111), where C1 reaches the highest positive contribution. N1 overall also shows a positive contribution, with CeO₂ (0.153), ZnO (0.134), and MnO₂ (0.125) being the most significant. In contrast, unsaturated and heteroatom-containing

motifs, including C2, N2, O1, O2, F, Cl and S, predominantly show negative contributions, indicating enhanced surface interactions. In particular, C2 and O2 exhibit negative contributions across all investigated oxides, identifying unsaturated carbon and oxygen-containing motifs as broadly adsorption-promoting features. Different metal oxides show noticeable selectivity for various atoms. For example, MnO₂ has the strongest negative effects on C2, O2, Cl, and S, with corresponding SHAP values of -0.202, -0.298, -0.138, and -0.087. TiO₂ contributes most negatively to N2 (-0.157), ZnO has the clearest impact on O1 (-0.099), and Al₂O₃ has the highest negative effect on F (-0.130). The ranking on the right further summarizes the top three key metal oxides for each type of atom, showing that the composition of metal oxides can significantly influence how different atoms or functional groups contribute to the model's predictions. Overall, CeO₂ and ZnO are more favorable for enhancing the positive effects of C1 and N1, while MnO₂ is more sensitive to various negative features, which also reveals a clear matching dependency between metal oxide surfaces and different chemical structures⁵²⁻⁵⁴.

To further bridge adsorption interpretation with reactive site prediction, we developed an atomic-scale bond-activation scoring framework that integrates model attention, bond energetics and adsorption-induced polarization (**Figure 4c**). For each candidate bond (i-j), the activation score is defined as $S_{bond,ij} = A_{ij} \cdot E_{ij} \cdot P_{ij}$, where A_{ij} , E_{ij} , and P_{ij} represent the normalized attention, energetic and polarization contributions, while $S_{bond,ij}$ provides an integrated measure of the susceptibility of bond (i-j) to activation. The definition and detailed calculation of A_{ij} , E_{ij} , and P_{ij} are shown in **Supplementary Notes 15**. The resulting score is bounded between 0 and 1 and integrates model attention, intrinsic bond strength, and bond polarization into a single descriptor. For the representative adsorption configuration, the α -C-H bond exhibited the highest score ($S_{bond}=0.76$), exceeding the corresponding values for the C-C ($S_{bond}=0.42$), C-O ($S_{bond}=0.56$), and O-H ($S_{bond}=0.69$) bonds (**Table S19**). The carbonyl oxygen first acts as the primary adsorption anchor, whereas the neighboring α -C-H bond exhibits the highest activation propensity.

We next applied this scoring framework to 39 representative VOCs molecules across eight metal oxides, generating a comprehensive bond-reactivity map (**Figure 4d and Supplementary Data2**). The dataset covers alkanes, alkenes, aromatics, alcohols, aldehydes, ketones, esters and halogen-containing organics. The activation response depended strongly on both metal oxides composition and VOCs functionality. For example, aromatic VOCs showed a high activation score on MnO_2 ($S_{\text{bond}}=0.85$), with the response dominated by benzyl C-H activation, whereas halogenated VOCs reached a high score on Al_2O_3 ($S_{\text{bond}}=0.80$), consistent with preferential C-X activation. Similarly, alkenes on TiO_2 exhibited a high score of 0.82, with C=C identified as the dominant activated bond, while ketones on CuO showed a score of 0.81. These indicate that the most favorable bond-activation site cannot be determined solely from the VOCs structure or the metal oxides surface independently. Instead, it emerges from their coupled influence on interfacial attention, bond strength, and polarization. Collectively, these results establish a transferable framework connecting molecular structure, surface atomic environment, adsorption energetics and bond activation, providing quantitative guidance for catalyst screening and rational design.

Discussion

We developed AdsDT, a transferable dual-tower framework that enables high-throughput prediction of VOCs adsorption on diverse metal oxides surfaces directly from unrelaxed crystal structures. Beyond energy prediction, we employed SHAP analysis and graph attention mapping to extract chemically meaningful descriptors across including molecular fragment, atomic, and bond level. The results reveal that saturated alkyl fragments generally weaken adsorption, whereas unsaturated groups and heteroatoms strengthen surface interactions through electronic coupling. More importantly, we propose a bond activation score ($S_{\text{bond},ij}=A_{ij}\cdot E_{ij}\cdot P_{ij}$) that integrates atomic attention, bond energy, and bond polarization to quantitatively predict preferential bond cleavage sites. This descriptor bridges the gap between thermodynamic adsorption and kinetic reactivity, uncovering that strongly adsorbed functional groups do not serve as direct cleavage centers but instead induce electronic

redistribution to activate adjacent bonds with the highest S_{bond} scores. These atom-resolved insights provide a generalizable platform for understanding metal oxides-VOCs interactions, accelerating catalyst discovery, and developing interpretable machine-learning strategies for heterogeneous catalysis.

Methods

Dataset construction

The OC20-Dense dataset derived from the Open Catalyst 2020 project was employed as the source-domain dataset for transfer-learning. It contains DFT relaxation trajectories from IS2RE/IS2RS tasks with diverse adsorbate-surface configurations built from Materials Project crystal structures. Two target datasets, Alloys-VOCs and metal oxides-VOCs, were constructed in-house using consistent DFT settings. Surface structures were retrieved from the Materials Project database, and VOC adsorbate configurations were generated by heuristic and random sampling followed by geometry relaxation. Descriptor representation for surfaces was achieved using averaged smooth overlap of atomic positions (SOAP) descriptors implemented in Dscribe v2.1.0, with a local cutoff of 4.0 Å, 8 radial basis functions, and spherical-harmonic maximum degree 6. Adsorbate molecular descriptors were computed via RDKit. Dimensionality-reduction analysis was performed with UMAP (umap-learn v0.5.5) using a fixed random seed for reproducibility.

AdsDT model architecture

The AdsDT dual-tower network consists of two independent encoding branches for surface and adsorbate, projecting heterogeneous inputs into a shared embedding space for cross-modal interaction via cross-attention layers. Key hyperparameters were optimized by grid-search on a held-out OC20 subset and major settings are summarized in **Supplementary Table S1**. The model was trained with batch size 64, initial learning rate 0.001, ReduceLROnPlateau scheduler (factor 0.7, patience 10), and maximum epoch 200. Transfer-learning was implemented in a progressive fine-tuning workflow: pre-training on the source OC20-Dense/Alloys-VOCs dataset, then fine-tuning on metal-oxides-VOCs data with partial backbone freezing to preserve pre-learned

adsorption representations. To test geometric robustness, random atomic-coordinate perturbation (displacement $< 0.2 \text{ \AA}$) and in-plane lattice-vector distortion ($< 2\%$) were applied to input surface slabs while keeping atomic connectivity unchanged. Model-predicted uncertainties were calibrated following standard protocols using the Uncertainty Toolbox v0.1.1 to compute calibration curves and miscalibration area.

Model training

The model was trained by minimizing the mean absolute error (MAE) loss with the AdamW optimizer. A ReduceLROnPlateau scheduler was used to dynamically adjust the learning rate. The dataset was randomly partitioned into training, validation and test subsets at an 8:1:1 ratio. Target values were standardized using the mean and standard deviation calculated solely from the training set. Ten independent training runs with distinct random seeds were performed to obtain robust results. All models were trained on a single NVIDIA Tesla A100 GPU under float32 precision. For fair comparison, the models with and without transfer learning shared identical architectures and hyperparameters.

Computational methods

All density-functional theory (DFT) calculations were performed using the Vienna ab initio simulation package (VASP). The exchange-correlation functional adopted the generalized-gradient approximation (GGA) of Perdew-Burke-Ernzerhof (PBE). Projector-augmented-wave (PAW) pseudopotentials were used to describe core-valence electron interactions, and long-range van-der-Waals corrections were included via the zero-damping DFT-D3 method of Grimme. Plane-wave kinetic-energy cutoff was set to 450 eV. Brillouin-zone sampling was performed using a Monkhorst-Pack $2 \times 2 \times 1$ k-point mesh for slab models. Structural relaxation converged when atomic forces were less than $0.02 \text{ eV \AA}^{-1}$ and the electronic self-consistent-field energy change fell below 10^{-5} eV . A vacuum layer of $15 \text{ \AA}$ along the z-direction was introduced to eliminate spurious interactions between periodic slab images. Crystal-structure visualizations were generated using the Visualization for Electronic and Structural Analysis (VESTA) software.

SHAP atomic-contribution analysis

GraphSHAP was adopted to obtain atom-wise Shapley values quantifying atomic marginal contributions toward predicted adsorption energy. Monte-Carlo sampling was used to approximate exact Shapley values under chemical-connectivity constraints via atomic-node-feature masking. Obtained values were standardized by Z-score transformation followed by sigmoid mapping onto [-1,1] to enable cross-molecule comparison. Positive SHAP values indicate an atom promotes adsorption, negative values suppress adsorption. Global internal-layer attention weights were also extracted by forward-pass hooks for auxiliary atomic-importance visualization.

GCN-Ads model for VOC adsorption-energy prediction

The GCN-Ads graph-convolutional-network model treats each VOC molecule as an undirected graph built from SMILES strings using RDKit with explicit hydrogen atoms added. Atomic nodes were encoded as 29-dimensional feature vectors including element one-hot, atomic degree, hybridization, formal charge, ring membership, and adsorption-reactive-site markers identified by SMARTS sub-structure matching. The network contains five graph-convolution blocks followed by three fully-connected layers, using batch normalization and ReLU activation. A global-attention layer aggregated node embeddings for molecule-level representation to predict scalar adsorption energy. Training parameters: batch size 128, Adam optimizer, initial learning rate 10^{-4} , mean-squared-error loss, maximum 1000 epochs. Dataset was filtered to remove missing labels, physically unreasonable entries, molecules larger than 500 Da and exotic elements, then split into training-validation-test subsets.

Bond-activation score ($S_{bond,ij}$) construction

For each covalent bond (i-j) within VOC molecules, the bond-activation score $S_{bond,ij}=A_{ij}\cdot E_{ij}\cdot P_{ij}$ was defined to evaluate bond susceptibility toward activation upon adsorption. A_{ij} is the normalized atomic-weight interaction term derived from AdsDT/GCN atomic attention weights, incorporating an adsorption-reward hyperparameter $\lambda_{ads} \in [0,1]$; bonds involving hydrogen set $w_H=0$. E_{ij} is normalized homolytic bond-dissociation-energy contribution. P_{ij} quantifies bond polarization from Pauling electronegativity differences between bonded atoms. $S_{bond,ij}$ serves for intramolecular bond-ranking rather than direct prediction of kinetic activation barriers.

Experimental methods

Catalyst preparation. Pure anatase TiO₂ powder was synthesized via a sol-gel route. Briefly, surfactant (1 g) was dissolved in 1-butanol; titanium tetraisopropanolate (0.01 mol) precursor was added, followed by drop-wise addition of 1 g HNO₃ at room temperature. The mixture was kept at 70 °C for 4 h to form a pale-yellow transparent gel, which was calcined in air at 450 °C for 4 h to obtain the final catalyst powder.

In situ DRIFTS measurements. In-situ diffuse-reflectance Fourier-transform infrared spectroscopy (DRIFTS) was performed on a Thermo-Nicolet iS10 FTIR spectrometer equipped with an MCT-A detector and a Harrick in-situ cell with ZnSe window. Approximately 50 mg catalyst powder was loaded. Prior to gas exposure, the catalyst was pretreated at 300 °C for 1 h, and background spectra were collected under N₂ atmosphere before introducing ethyl-acetate-containing feed gas.

PTR-TOF-MS characterization. Proton transfer reaction time of flight mass spectrometry (PTR-TOF-4000, IONICON Analytik GmbH) was used to qualitatively and quantitatively detect gas-phase intermediates during ethyl-acetate catalytic oxidation. Outlet gas was delivered through a heated transfer line; H₃O⁺ reagent ions enabled soft proton-transfer ionization. Intermediates were assigned according to exact mass-to-charge ratios combined with reference mass-spectral data.

Data availability

Data supporting the findings of this study are available in the article and Supplementary Information.

Code availability

The source code of the AdsMT framework is available on GitHub at <https://github.com/iojsea/metal-oxide-vocs-adsorption>.

References

1. He, C., Cheng, J., Zhang, X., Douthwaite, M., Pattison, S. & Hao, Z. Recent advances in the catalytic oxidation of volatile organic compounds: a review based on pollutant sorts and sources. *Chem. Rev.* **119**, 4471-4568 (2019).

- 569 2. Lee, J. E., Ok, Y. S., Tsang, D. C. W., Song, J., Jung, S. C. & Park, Y. K. Recent advances
570 in volatile organic compounds abatement by catalysis and catalytic hybrid processes: a
571 critical review. *Sci. Total Environ.* **719**, 137405 (2020).
- 572 3. Guo, Y. et al. Recent advances in VOCs elimination by catalytic oxidation technology
573 onto various nanoparticles catalysts: a critical review. *Appl. Catal. B: Environ.* **281**, 119447
574 (2021).
- 575 4. Kamal, M. S., Razzak, S. A. & Hossain, M. M. Catalytic oxidation of volatile organic
576 compounds (VOCs)-a review. *Atmos. Environ.* **140**, 117-134 (2016).
- 577 5. Yang, C. et al. J. Abatement of various types of VOCs by adsorption/catalytic oxidation:
578 a review. *Chem. Eng. J.* **370**, 1128-1153 (2019).
- 579 6. Brummer, V., Teng, S. Y., Jecha, D., Skryja, P., Vavrcikova, V. & Stehlik, P. Contribution
580 to cleaner production from the point of view of VOCs emissions abatement: a review. *J.*
581 *Clean. Prod.* **361**, 132112 (2022).
- 582 7. Krishnamurthy, A., Adebayo, B., Gelles, T., Rownaghi, A. & Rezaei, F. Abatement of
583 gaseous volatile organic compounds: a process perspective. *Catal. Today* **350**, 100-119
584 (2020).
- 585 8. Lu, T. et al. P. Catalytic oxidation of volatile organic compounds over manganese-based
586 oxide catalysts: performance, deactivation and future opportunities. *Sep. Purif. Technol.*
587 **291**, 121436 (2022).
- 588 9. Sui, C. et al. Research progress of catalytic oxidation of volatile organic compounds over
589 Mn-based catalysts-a review. *Rev. Inorg. Chem.* **43**, 12 (2023).
- 590 10. Trovarelli, A. et al. Catalytic properties of ceria and CeO₂-containing materials. *Catal.*
591 *Rev.: Sci. Eng.* **38**, 439-520 (1996).
- 592 11. Montini, T., Melchionna, M., Monai, M. & Fornasiero, P. Fundamentals and catalytic
593 applications of CeO₂-based materials. *Chem. Rev.* **116**, 5987-6041 (2016).
- 594 12. Yang, C. et al.. Defect engineering on CeO₂-based catalysts for heterogeneous
595 catalytic applications. *Small Struct.* **2**, 2100058 (2021).
- 596 13. Gelles, T., Krishnamurthy, A., Adebayo, B., Rownaghi, A. & Rezaei, F. Abatement of
597 gaseous volatile organic compounds: a material perspective. *Catal. Today* **350**, 3-18 (2020).
- 598 14. Zhang, X. et al. Recent progress in the removal of volatile organic compounds by
599 mesoporous silica materials and supported catalysts. *Phys. Chem. Chem. Phys.* **17**,
600 19713-19727 (2015).
- 601 15. Wang, D. et al. Recent advances in catalytic removal of volatile organic compounds
602 over metal-organic framework-derived catalysts: a review. *Sep. Purif. Technol.* **326**,
603 124765 (2023).
- 604 16. Taylor, M. et al. Synergy between tungsten and palladium supported on titania for the
605 catalytic total oxidation of propane. *J. Catal.* **285**, 103-114 (2012).

606 17. Enterkin, J. A., Setthapun, W., Elam, J. W., Christensen, S. T., Rabuffetti, F. A., Marks,
607 L. D., Stair, P. C., Poeppelmeier, K. R. & Marshall, C. L. Propane oxidation over Pt/SrNiO₃
608 nanocuboids. *ACS Catal.* **1**, 629-635 (2011).

609 18. Chen, H. et al. AI for nanomaterials development in clean energy and carbon capture,
610 utilization and storage (CCUS). *ACS Nano* **17**, 9763-9792 (2023).

611 19. Hippalgaonkar, K. et al. Knowledge-integrated machine learning for materials: lessons
612 from gaming and robotics. *Nat. Rev. Mater.* **8**, 241-260 (2023).

613 20. Lei, Q. et al. Emerging directions for carbon capture technologies: a synergy of high-
614 throughput theoretical calculations and machine learning. *Environ. Sci. Technol.* **57**, 17189-
615 17200 (2023).

616 21. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy
617 calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169-11186 (1996).

618 22. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made
619 simple. *Phys. Rev. Lett.* **77**, 3865-3868 (1996).

620 23. Chen, C. et al. Descriptor-driven prediction of adsorption energy of oxygenates on
621 metal dioxide surfaces. *J. Phys. Chem. C* **129**, 6245-6253 (2025).

622 24. Sutton, J. E. & Vlachos, D. G. A theoretical and computational analysis of linear free
623 energy relations for the estimation of activation energies. *ACS Catal.* **2**, 1624-1634 (2012).

624 25. Ge, X. et al. Atomic design of alkyne semihydrogenation catalysts via active learning.
625 *J. Am. Chem. Soc.* **146**, 4993-5000 (2024).

626 26. Zheng, Y. et al. Online learning-based predictive control of crystallization processes
627 under batch-to-batch parametric drift. *AIChE J.* **68**, e17815 (2022).

628 27. Ciallella, H. L. et al. Machine learning and high-throughput in silico toxicity assessment
629 for environmental chemicals. *Environ. Sci. Technol.* **56**, 13879-13892 (2022).

630 28. Wu, T. et al. Small-sample learning for next-generation human health risk assessment:
631 harnessing AI, exposome data, and systems biology. *Environ. Sci. Technol.* **59**, 5-10 (2025).

632 29. Wang, H. et al. Transfer learning with a graph attention network and weighted loss
633 function for screening of persistent, bioaccumulative, mobile, and toxic chemicals. *Environ.*
634 *Sci. Technol.* **59**, 578-590 (2025).

635 30. Han, Z. et al. Interpretable machine learning for predicting toluene oxidation activity
636 over manganese oxide catalysts. *Appl. Catal. B: Environ.* **330**, 122624 (2023).

637 31. Fung, V. et al. Machine learned features from density of states for accurate adsorption
638 energy prediction. *Nat. Commun.* **12**, 88 (2021).

639 32. Pablo-García, S., Morandi, S., Vargas-Hernández, R. A., Jorner, K., Ivković, Ž., López,
640 N. & Aspuru-Guzik, A. Fast evaluation of the adsorption energy of organic molecules on
641 metals via graph neural networks. *Nat. Comput. Sci.* **3**, 433-442 (2023).

- 642 33. Schütt, K. T., Saucedo, H. E., Kindermans, P.-J., Tkatchenko, A. & Müller, K.-R.
643 SchNet—a deep learning architecture for molecules and materials. *J. Chem. Phys.* **148**,
644 241722 (2018).
- 645 34. Lan, J. et al. AdsorbML: a leap in efficiency for adsorption energy calculations using
646 generalizable machine learning potentials. *npj Comput. Mater.* **9**, 78 (2023).
- 647 35. Yin, J. et al. Generalized reactor neural ODE for dynamic reaction process modeling
648 with physical interpretability. *Chem. Eng. J.* **452**, 139487 (2023).
- 649 36. Price, C. C. et al.. Efficient catalyst screening using graph neural networks to predict
650 strain effects on adsorption energy. *Sci. Adv.* **8**, eabq5944 (2022).
- 651 37. Jung, H. et al. Machine-learning driven global optimization of surface adsorbate
652 geometries. *npj Comput. Mater.* **9**, 114 (2023).
- 653 38. Klimanova, O. et al. Accelerating the global search of adsorbate molecule positions
654 using machine-learning interatomic potentials with active learning. *Phys. Chem. Chem.*
655 *Phys.* **27**, 9201-9210 (2025).
- 656 39. Shapeev, A. V. Moment tensor potentials: a class of systematically improvable
657 interatomic potentials. *Multiscale Model. Simul.* **14**, 1153-1173 (2016).
- 658 40. Gubaev, K., Podryabinkin, E. V., Hart, G. L. & Shapeev, A. V. Accelerating high-
659 throughput searches for new alloys with active learning of interatomic potentials. *Comput.*
660 *Mater. Sci.* **156**, 148-156 (2019).
- 661 41. Kim, M. et al. Exploring optimal water splitting bifunctional alloy catalyst by Pareto
662 active learning. *Adv. Mater.* **35**, 2211497 (2023).
- 663 42. Kim, M. et al. Searching for an optimal multi-metallic alloy catalyst by active learning
664 combined with experiments. *Adv. Mater.* **34**, 2110890 (2022).
- 665 43. Ye, R. et al. Adidharma, H., Radosz, M. & Fan, M., Yao, Y.-G. Perspectives on the active
666 sites and catalyst design for the hydrogenation of dimethyl oxalate. *ACS Catal.* **10**, 3139-
667 3169 (2020).
- 668 44. Chen, X. et al. d_{z2} band links frontier orbitals and charge carrier dynamics of single-
669 atom cocatalyst-aided photocatalytic H₂ production. *J. Am. Chem. Soc.* **145**, 24789-24797
670 (2023).
- 671 45. Liu, J. et al. Van der waals interactions and geometric confinement effects of alkyl
672 chains on metal oxide surfaces. *J. Am. Chem. Soc.* **143**, 6572-6580 (2021).
- 673 46. Wang, Y. et al. π -electron interactions of unsaturated carbon-carbon bonds with
674 titanium dioxide surfaces. *J. Am. Chem. Soc.* **144**, 14215-14224 (2022).
- 675 47. Zhang, S. et al. Local bonding environment governs the adsorption behaviors of
676 nitrogen-containing volatile organic compounds on oxide catalysts. *J. Am. Chem. Soc.* **145**,
677 6789-6798 (2023).
- 678 48. Chen, L. et al. Resonance charge delocalization controls the anchoring mode of nitro
679 groups on metal oxide surfaces. *J. Am. Chem. Soc.* **142**, 12201-12209 (2020).

49. Wang, Y. et al. Adsorption mechanisms of typical carbonyl-containing volatile organic compounds on anatase TiO₂(001) surface: a DFT investigation. *ACS Catal.* **7**, 5441-5451 (2017).
50. Wang, X. et al. Adsorption of halogenated methanes on nitrogen-doped titania nanoparticles: insights from DFT calculations. *ACS Catal.* **13**, 12145-12154 (2023).
51. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. Preprint at arXiv:1609.02907 (2016).
52. Xie, X. et al. Haruta, M. & Shen, W. Low-temperature oxidation of CO catalysed by Co₃O₄ nanorods. *Nature* **458**, 746-749 (2009).
53. Luo, T. et al. Coote, M. L. & Liu, M. Unveiling tetrafluoromethane decomposition over alumina catalysts. *J. Am. Chem. Soc.* **146**, 35057-35063 (2024).
54. Liao, W. et al. Roles of oxygen vacancies in the bulk and surface of CeO₂ for toluene catalytic combustion. *ACS Catal.* **10**, 10204-10214 (2020).

Acknowledgements

The authors are grateful for the financial support from National Natural Science Foundation of China (No. 22376075, 22376074, 22525203, 62293553), the Central China Normal University project (CCNU24JCPT017).

Author contributions

B. Z. to methodology, model design, writing, visualization and assessment. P. Y. contributed to methodology and model creation. D. Zhou, J. W. and Y. G. contributed to conceptualization, methodology, model creation, writing, assessment, funding and project supervision.

Competing interests

The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryData2BondactivationScores.xlsx](#)
- [SupplementaryData1Atomattentionweight.xlsx](#)
- [Supportinginformation.docx](#)