

Unveiling the Activity Descriptors for Metal Nanoclusters-N4C as Electrocatalysts through Machine Learning and Experimental Validation

Arupjyoti Pathak

SRM University, Andhra Pradesh

Srijib Das

Central Mechanical Engineering Research Institute

Mukaddar SK

SRM University, Andhra Pradesh

Tapas Kuila

Central Mechanical Engineering Research Institute

Om Jadhav

Centre for Development of Advanced Computing

Samrit Kumar Maity

Centre for Development of Advanced Computing

Aniruddha Kundu

SRM University, Andhra Pradesh

Ranjit Thapa

ranjit.phy@gmail.com

SRM University, Andhra Pradesh

Article

Keywords: Machine learning, Descriptor, Nanocatalysts, Electrocatalyst, DFT

Posted Date: September 3rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10760815/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

Unveiling the Activity Descriptors for Metal Nanoclusters-N₄C as Electrocatalysts through Machine Learning and Experimental Validation

Arupjyoti Pathak^{1,2}, Srijib Das^{3,4}, Mukaddar Sk^{1,2}, Tapas Kuila^{3,4}, Om Jadhav⁵, Samrit Kumar Maity⁵, Aniruddha Kundu^{6*} and Ranjit Thapa^{1,2*}

¹ Department of Physics, SRM University - AP, Amaravati 522240, Andhra Pradesh, India

² Asian Consortium on Computational Materials Science - Global Research Centre, SRM University - AP, Amaravati 522240, Andhra Pradesh, India

³ Electric Mobility & Tribology Research Group, CSIR- Central Mechanical Engineering Research Institute, Durgapur 713209, India

⁴ Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, 201002, India

⁵ Centre for Development of Advanced Computing (C-DAC), Pune 411008, Maharashtra, India

⁶ Centre for Interdisciplinary Research, SRM University - AP, Amaravati 522240, Andhra Pradesh, India

*Corresponding Authors: ranjit.phy@gmail.com; write2aknow@gmail.com

Arupjyoti Pathak and Srijib Das contributed equally to this work.

Abstract

Metal nanoclusters are very difficult to explain as catalysts. Unlike bulk metals, their properties change significantly with every atom added or removed, making it nearly impossible to explain and predict the site-specific activity of the given catalytic reaction using traditional approaches. This unpredictability has made the explanation and rational design of nanocluster catalysts a long-standing open problem. Here, this complexity is resolved for the oxygen evolution reaction (OER) through a combined high-throughput DFT screening and machine learning-driven electronic descriptor analysis of 62 earth-abundant metal nanoclusters (Fe, Co, Ni; n=5-38)

supported on nitrogen-doped graphene. Evaluation of 211 distinct surface sites identifies Ni₁₉@N₄C as the most active catalyst, with optimal d-orbital occupancy and metal-adsorbate bonding strength (ICOHP), delivering a theoretical overpotential of 0.39 V. ΔG_{OOH} is identified as the key energy descriptor for predicting overpotential(η). Integration of these datasets into a machine learning framework reveals that the Random Forest Regression model ($R^2=0.87$) is the best one, with the d-band frontier down($d_f \downarrow$) as the primary electronic descriptor in the multi-descriptor framework, confirming that local frontier orbitals near the Fermi level affect reactivity in these confined systems. The computational predictions led to the synthesis of the NiNC@NGC catalyst, which underwent electrochemical testing. The experimental trends confirmed the theoretical findings with an IR-corrected overpotential(η_{10}) of 460 mV and a Tafel slope of 94 mV dec⁻¹, which is comparable to the performance of commercial RuO₂. This work provides an explainable, descriptor-based framework for decoding the electronic complexity of metal nanocluster catalysts and offers a reliable computational pathway for accelerating the discovery of scalable, earth-abundant OER electrocatalysts.

Keywords: Machine learning, Descriptor, Nanocatalysts, Electrocatalyst, DFT

41 1 Introduction

42 The properties of metals change in unexpected ways when confined to the sub-nanometer range¹.
43 Something fundamental happens when bulk metal is reduced down to a nanocluster of just tens of
44 atoms: its continuous electronic band structure collapses into discrete energy levels, and adding or
45 removing even a single atom can meaningfully change how the cluster reacts chemically, its
46 magnetic character, and coordination geometry^{2,3}. This sensitivity, at once what makes
47 nanoclusters scientifically exciting, also makes them extremely difficult to understand, since even
48 two clusters of identical metal and size can behave very differently depending on their shape,
49 support and local bonding environment⁴. Decoding this complexity for electrocatalysis is one of
50 the central open problems in nanoscience, and it is precisely this problem that motivates the present
51 work.

52 Within electrocatalysis, the oxygen evolution reaction (OER) provides a particularly demanding
53 testbed for nanocluster catalysts as it is one of the critical bottlenecks in clean energy technologies
54 like fuel cells and rechargeable metal-air batteries⁵. OER is intrinsically limited by sluggish
55 kinetics and high overpotentials (η), governed by the Suboptimal binding of the oxygenated
56 intermediates ($\ast\text{OH}$, $\ast\text{O}$, $\ast\text{OOH}$) in the catalysts⁶. Although IrO_2 and RuO_2 remain the benchmark
57 catalysts, their scarcity and cost necessitate earth-abundant alternatives^{7,8}. Their high cost and
58 limited availability (IrO_2 , RuO_2) have accelerated the transition toward Single-Atom Catalysts
59 (SACs), first introduced by Qiao et al.⁹, offering maximum atom utilisation via isolated TM- N_x/C
60 active sites whose OER activity is dictated by the metal's d-band character and local N-
61 coordination environments¹⁰. However, SACs suffer from inherent limitations: the spatial
62 isolation of active sites prevents cooperative intermediate stabilisation, and under anodic OER
63 potentials, isolated metal atoms become prone to migration and aggregation into nanoclusters. This
64 has shifted attention toward transition-metal nanoclusters (Co, Fe, Ni), where sub-nanometer
65 metal-metal clusters show excellent catalytic activity, multiple active sites enabling charge
66 redistribution, and superior structural stability relative to SACs¹¹. Inspired by these advantages,
67 several experimental studies have demonstrated the excellent performance of metal nanoclusters
68 for OER, but the origin behind the activity remained unexplained. For example, Li et al.
69 synthesised carbon-encapsulated Co clusters (Co_{x-2}) that exhibited an ultralow overpotential of
70 227 mV at a current density of 10 mA cm^{-2} , along with remarkable stability for more than 200
71 hours in 1.0 M KOH¹². Similarly, Lu et al. developed $\text{Co}_2\text{Fe}_2/\text{Fe}_5$ nanoclusters that exhibited
72 improved OER activity and stability compared with commercial Pt/C and RuO_2 catalysts¹³. Seong
73 et al. reported a high current density of 500 mA cm^{-2} for OER at a low overpotential of 0.39 V
74 using Ni_6/NF , demonstrating its strong catalytic potential¹⁴. In another study, Yu et al. achieved

75 superior OER performance using nitrogen-doped Ni nanoclusters confined in an amorphous
76 carbon matrix, delivering a low overpotential of 240 mV at 10 mA cm⁻² with good stability.
77 Moreover, Chen et al. synthesised Fe nanoclusters combined with single-atom sites on one-
78 dimensional carbon nanotubes (Fe₃C-NCNTs-90 min CC⁻¹).

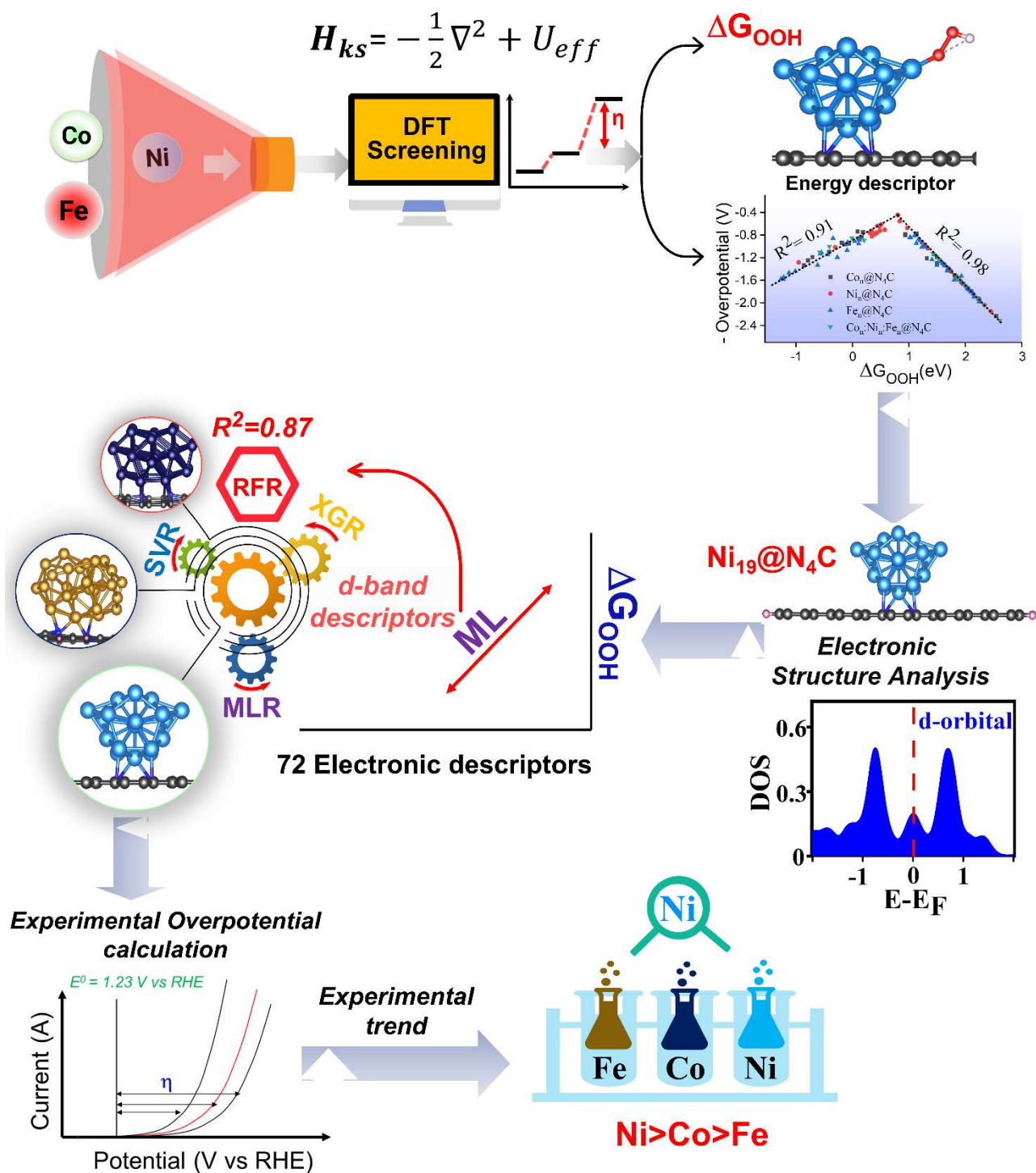
79 Although Co, Ni, and Fe-based nanoclusters have shown promising OER activity and stability,
80 further improvement in their catalytic efficiency is still necessary to meet the performance
81 requirements for large-scale industrial water splitting. Enhancing their activity and durability can
82 significantly reduce the energy consumption and operational cost, which is essential for the
83 practical and commercial deployment of sustainable hydrogen production technologies. But the
84 experimental synthesis relies on a trial-and-error approach, which is time-consuming and costly,
85 limiting its effectiveness in further increasing the OER activity of nanoclusters. Also, without
86 clearly understanding the electronic structure and the important electronic properties of these
87 nanoclusters, which influence the OER activity, one can never design a highly efficient
88 electrocatalyst in a cost-effective way. In this context, DFT-based electronic descriptors have
89 emerged as a promising approach, as they can directly correlate the electronic structure of materials
90 with their catalytic activity, thereby enabling the prediction of highly efficient candidates for the
91 oxygen evolution reaction (OER). For example, Hong et al. performed a comprehensive statistical
92 analysis and demonstrated that the number of d electrons, charge-transfer energy (covalency), and
93 the optimal occupancy of the e_g orbital play crucial roles in describing the OER activity of 51
94 perovskite materials¹⁵. Similarly, Zhou et al. identified the d-band centre and e_g filling as key
95 descriptors for predicting the activity of MOF-based electrocatalysts, offering valuable insights
96 for the rational design of highly efficient OER catalysts¹⁶. Furthermore, Cheng et al. established
97 correlations between electronic conductivity, oxygen vacancy concentration, flat-band potential
98 (E_{fb}), and OER activity across a wide range of perovskite compositions, including the [La_{1-x}
99 Sr_xCoO_{3-δ} series (with x = 0, 0.2, 0.4, 0.6, 0.8), LaMO_{3-δ} series (M = Cr, Mn, Fe, Co, Ni),
100 Ba_{0.5}Sr_{0.5}Co_{0.8}Fe_{0.2}O_{3-δ} (BSCF), and PrBaCo₂O_{6-δ} (PBCO)]¹⁷. These studies collectively highlight
101 the importance of electronic descriptors in understanding and predicting OER catalytic
102 performance. To provide a broader picture, **Table 1** summarises the key electronic descriptor-
103 based studies reported in the literature for metal-based catalysts, covering both DFT and machine
104 learning (ML) approaches across a range of electrocatalytic reactions, including OER, ORR, HER,
105 NRR, and CO₂RR. As seen in the table, the descriptors employed span from well-established ones
106 like the d-band centre and e_g orbital occupancy to less known ones such as the generalised
107 coordination number, O p-band centre, orbital-wise coordination number, work function combined
108 with Bader charge, and width-corrected band centre. The catalyst classes explored are equally

109 diverse, ranging from transition metals and alloys, perovskites, MOFs, and spinels to emerging
110 materials like MXenes, high-entropy alloys, single-atom catalysts (SACs), and dual-atom
111 catalysts. Notably, composite descriptors by combining two or more electronic parameters, such
112 as d-band filling paired with local electronegativity, or width-corrected band centre with d-band
113 frontier, have also been explored to capture the complexity of catalytic activity trends across
114 systems. Although these descriptor-based approaches have shown great promise for screening
115 metal-based catalysts, relying on a single electronic descriptor is generally not sufficient to
116 describe and predict the OER activity on metal nanoclusters. This is because nanoclusters possess
117 catalytic activity that is very complex and difficult to explain; activity is governed by multiple
118 interdependent factors rather than a single electronic parameter. In this case, nanoclusters exhibit
119 significant structural and electronic heterogeneity. Unlike bulk materials or extended surfaces,
120 atoms in nanoclusters occupy different coordination environments, including edge, corner, and
121 facet sites, and therefore, there is a presence of local electronic structure. A single descriptor,
122 including d-band centre or e_g occupancy, cannot adequately capture these site-dependent electronic
123 variations that strongly influence catalytic activity. Second, OER involves multiple proton-
124 electron transfer steps and several reaction intermediates, including $^*\text{OOH}$, $^*\text{O}$, and $^*\text{OH}$. The
125 adsorption strength of these intermediates is controlled by different aspects of the electronic
126 structure, such as orbital hybridisation, charge transfer, and metal-oxygen covalency and
127 interaction. As each reaction step depends on a different electronic feature, and electronic features
128 vary unexpectedly in different local environments, a single descriptor cannot accurately capture
129 the entire reaction activity. In this regard, machine learning (ML) has emerged as a highly effective
130 approach for catalyst design and explanation, as it can efficiently analyse large datasets and
131 automatically uncover the parameters governing catalytic activity¹⁸. In contrast to traditional
132 descriptor-based strategies, ML techniques are capable of incorporating numerous variables
133 simultaneously, which improves the reliability of predictions related to reaction energetics and
134 catalytic activity trends^{19,20}. For example, Siddharthan et al. employed several ML algorithms-
135 including Multiple Linear Regression (MLR), Ridge Regression (RR), Support Vector Regression
136 (SVR), Random Forest Regressor (RFR), XGBoost regression(XGR), Bayesian Ridge Regression
137 (BRR), Least Absolute Shrinkage and Selection Operator (LASSO), and K-Nearest Neighbours
138 (KNN), to establish multi-parameter relationships for evaluating the oxygen evolution reaction
139 (OER) performance of single-atom catalysts (SACs)²¹. Similarly, Kapse et al. utilised ML
140 techniques to construct predictive models based on π -electron descriptors, enabling rapid
141 identification of promising carbon-based OER catalysts while significantly reducing
142 computational expense²². In another study, Elcicek et al. assessed the predictive capability of

143 different ML methods, including multilayer perceptron artificial neural networks (MLP-ANN),
144 SVR, and random forest (RF). Their analysis indicated that the MLP-ANN model achieved the
145 highest predictive accuracy for the performance of Pt/C electrocatalysts when evaluated using
146 statistical indicators such as R^2 , MAE, MSE, and RMSE²³. However, an integrated machine
147 learning (ML)-descriptor framework for N₄-doped carbon-anchored transition-metal nanoclusters,
148 which is verified with synthesised experimental catalysts, has not yet been reported. This
149 framework could explain and predict the oxygen evolution reaction (OER) activity of transition-
150 metal nanoclusters, which is a much-needed study and would give key understanding for rapidly
151 and cheaply developing highly active OER catalysts.

152 In this article, high-throughput DFT screening of 62 earth-abundant metal nanoclusters (Fe, Co,
153 and Ni; n=5-38) Supported on nitrogen-doped graphene has been conducted. The overall
154 computational and experimental workflow of the study is illustrated schematically in Figure 1.
155 Analysis of 211 distinct surface sites identified Ni₁₉@N₄C as the most promising catalyst,
156 exhibiting a low theoretical overpotential of 0.39 V. The results further reveal that ΔG_{OOH} serves
157 as a key energy descriptor for predicting the overpotential (η). Notably, electronic structure
158 investigations reveal that traditional linear correlations with d-orbital parameters are insufficient
159 for these systems. The failure of simple linear scaling underscores the sophisticated electronic
160 perturbations inherent to nanoclusters, which are highly sensitive to their local coordination
161 environments. To address this complexity, we integrated these electronic parameters into a
162 machine learning framework. The resulting Random Forest Regression (RFR) model ($R^2 = 0.92$)
163 demonstrates the highest predictive fidelity. A composite multi-descriptor framework is
164 established in the RFR model, within which the d-band frontier down ($d_f \downarrow$) is identified as the
165 most influential electronic feature among the 14 descriptors considered. Leveraging these insights,
166 the NiNC@NGC catalyst was synthesized and subjected to experimental validation.
167 Electrochemical measurements confirmed the theoretical findings, delivering an iR-corrected
168 overpotential (η_{10}) of 460 mV and a Tafel slope of 94 mV dec⁻¹, comparable to commercial RuO₂.
169 Overall, this study demonstrates a scalable descriptor-driven strategy, integrating DFT and
170 machine learning, to unravel the electronic complexity of nanoclusters and accelerate the discovery
171 of efficient earth-abundant OER catalysts.

172



173

174 **Figure 1:** Schematic representation of the complete workflow adopted in this study.

175

176 2 Computational Methodology

177 Density functional theory (DFT) calculations were performed with the Vienna Ab initio
 178 Simulation Package (VASP)^{24–27}. The electron-ion interactions were described by the Projector
 179 Augmented Wave (PAW) method²⁸. The electron exchange-correlation was treated within the
 180 Generalised Gradient Approximation (GGA) using the Perdew-Burke-Ernzerhof (PBE)

functional²⁹. The Grimme DFT-D3 dispersion correction with Becke-Johnson damping was employed to consider van der Waals interactions, which is crucial for accurately describing molecular adsorption and surface-related phenomena. The kinetic energy cutoff of the plane-wave basis set was set to 450 eV after convergence tests, and the Brillouin zone was sampled using the Monkhorst-Pack scheme with a $1 \times 3 \times 1$ for accurate integration over the reciprocal space. Geometric optimisations were performed until the energy convergence criterion reached 10^{-6} eV and the forces on each atom were relaxed until they were below 0.01 eV/Å, ensuring a relaxed structure according to the Hellmann-Feynman forces. A vacuum layer of 20 Å was added in the z-direction to avoid interactions between periodic images. The Nanocluster catalyst library was constructed by starting with bulk crystal lattices of the individual elements in their ground-state phase (fcc for Ni, hcp for Co and bcc for Fe) and then systematically truncating them to create a total of 62 unique geometries from 5 to 38 atoms. In an attempt to mimic a more realistic catalytic interface, all the resulting clusters were later anchored on a 152-atom nitrogen-doped graphene nanoribbon scaffold represented as C₄N, where four graphitic carbons were replaced with four nitrogens. Also, the Gibbs free energy (G) of the system was calculated using the equation $G = E + \text{ZPE} - TS - n e U$, where E is the total energy obtained from DFT calculations, ZPE is the zero-point energy, TS is the entropic contribution, n is the number of electrons transferred, and U is the applied potential³⁰. The values of entropy and ZPE of free molecules were taken from the standard chemical database and are listed in the table 1, and for the catalyst is not considered, as for big systems, it is negligible. The van der Waals interactions were considered by using the DFT-D3 dispersion correction with Becke-Johnson damping, as proposed by Grimme³¹. Furthermore, the Hubbard U correction has also been employed in the best system. The U_{eff} represents the effective interaction term, which is the difference between the screened Coulomb and exchange parameters ($U_{\text{eff}} = U - J$)^{32,33}. The U value was considered to be 6 eV, which is taken from the literature for this study. In order to get more insights into metal-adsorbate bonding, Crystal Orbital Hamilton Population (COHP) analysis was carried out using the Local Orbital Basis Suite Toward Electronic-Structure Reconstruction (LOBSTER). Additionally, ab initio molecular dynamics (AIMD) simulations were performed at 350 K for a total of 10 ps.

Machine learning methods were implemented using scikit-learn, a Python library. To reduce the dimension of the feature space and focus on the most significant descriptors, Principal Component Analysis (PCA) was utilised. Several regression algorithms were developed and evaluated, including Multiple Linear Regression (MLR), Ridge Regression (RR),

214 LASSO, Bayesian Ridge Regression (BRR), Support Vector Regressor (SVR), K-Nearest
 215 Neighbours (KNN), Random Forest Regressor (RFR), and XGBoost Regressor (XGR). Prior
 216 to model training, the dataset was normalised using z-score standardisation, which is defined
 217 as

$$218 \quad z = \frac{x - \mu}{\sigma}$$

219 where x is the feature value, μ is the mean, and σ is the standard deviation of the
 220 corresponding feature. Electronic descriptors are calculated based on the following set of
 221 equations²¹, which enable a quantitative representation of the d-orbital characteristics that
 222 directly impact the adsorption energetics and reaction intermediates in electrocatalysis:

$$223 \quad \text{Band occupancy: } n = \int_{-\infty}^{E_F} \rho dE \quad (1)$$

$$224 \quad \text{d}_{xz} \text{ orbital's occupancy: } n_{d_{xz}} = \int_{-\infty}^{E_F} \rho d_{xz}(E) dE \quad (2)$$

$$225 \quad \text{d-band centre : } \varepsilon_d = \frac{\int_{-\infty}^{E_F} \rho E dE}{\int_{-\infty}^{E_F} \rho dE} \quad (3)$$

$$226 \quad \text{The occupancy of } dz^2 : n_{dz^2} = \int_{-\infty}^{E_F} \rho dz^2(E) dE \quad (4)$$

$$227 \quad \text{d}_{xz} \text{ (down) orbital's occupancy : } n \downarrow_{d_{xz}} = \int_{-\infty}^{E_F} \rho \downarrow d_{xz}(E) dE \quad (5)$$

$$228 \quad \text{The occupancy of d orbital: } n \downarrow_d = \int_{-\infty}^{E_F} \rho \downarrow d(E) dE \quad (6)$$

$$229 \quad \text{Width of d-band } W_d = \sqrt{\frac{\int_{-\infty}^{\infty} (E - \varepsilon_d)^2 \rho dE}{\int_{-\infty}^{\infty} \rho dE}} \quad (7)$$

$$230 \quad \text{Width corrected band centre } W = \varepsilon_d + \frac{W_d}{2} \quad (8)$$

$$231 \quad \text{d-band frontier : } d_f = \frac{\int_{-1}^1 \rho_d E dE}{\int_{-\infty}^{+\infty} \rho_d E dE} \quad (9)$$

$$\text{d-band frontier up : } d_f \uparrow = \frac{\int_{-1}^1 \rho_d \uparrow E dE}{\int_{-\infty}^{+\infty} \rho_d \uparrow E dE} \quad (10)$$

$$\text{d-band frontier down : } d_f \downarrow = \frac{\int_{-1}^1 \rho_d \downarrow E dE}{\int_{-\infty}^{+\infty} \rho_d \downarrow E dE} \quad (11)$$

In this relation, ρ represents the density of states, E_F refers to the Fermi level, and E signifies the energy associated with that density of states.

3 Results and Discussion

3.1 Structural details and Stability

In order to gain a systematic understanding of the size-, composition-, and local electronic environment-dependent kinetics of metal nanocluster electrocatalysts, a diverse library of 62 nanocluster architectures was assembled, termed $M_n@N_4C$. This library spans a wide material space ($M = \text{Fe, Co, Ni, FeCo, FeNi, CoNi}$) and size range ($n = 5\text{-}38$ atoms). It is worth noting that nanoclusters of Fe, Co, or Ni containing an even number of atoms tend to have closed-shell electronic structures, often exhibit weaker or less optimal adsorption and thus suppress catalytic activity. In contrast, nanoclusters with an odd number of atoms in some cases lead to adsorption energies closer to the Sabatier-optimal range. The Host is a quasi-one-dimensional graphene nanoribbon (GNR) with a central anchoring site, where four carbon atoms in the graphitic basal plane are replaced with nitrogen to yield a tetradentate N_4 moiety (N_4C). The catalyst geometries, denoted as $M_n@N_4C$ (e.g. Fe_{19} , $\text{Ni}_{14}@N_4C$), were constructed by truncating thermodynamically stable bulk lattices, specifically FCC-Ni, HCP-Co, and BCC-Fe, to generate initial nanocluster geometries. Each nanocluster and graphene layer was separately optimised using high-energy and force convergence criteria. The optimised nanocluster was then placed on the graphene sheet, where the N_4C site provides optimal coordination and high stability to anchor the nanocluster. The nanoclusters are effectively trapped by the basal N_4 -sites (Figure S1, S2, S3, S4), leaving accessible active sites for electrocatalytic turnover. The lone-pair electrons on the graphitic N sites act as σ -donors to the metal orbitals^[20], while the partially occupied transition metal's 3d states provide a back-donation of electron density to the π^* antibonding orbitals of the N-doped graphitic carbon³⁴⁻³⁷. This bidirectional electron flow facilitates robust orbital hybridisation, serving a dual structural and electronic purpose (consistent with calculations reported by He et al.^{38,39}). Strong metal-support interactions simultaneously suppress cluster migration via diffusion barriers and tune d-band electronic

structure through charge redistribution, optimising intermediate binding energetics for enhanced OER kinetics^{40–43}. To quantitatively validate this structural resilience, the thermodynamic stability was evaluated by analysing the energetic competition between cluster cohesion (E_{coh}) and interfacial binding energy per atom (E_b) as shown in Figure 2(a). The cohesive energy was determined by calculating the difference between the energy of the isolated nanocluster (E_{Cluster}) and the sum of the chemical potentials of its constituent atoms ($\sum N\mu M_n$) normalised by the atom count (N). Concurrently, the binding energy was derived from the ground-state energy of the anchored system ($E_{\text{Mn@N}_4\text{C}}$) relative to its isolated components ($E_{\text{N}_4\text{C}}$ and E_{Mn}):

$$E_b = E_{\text{Mn@N}_4\text{C}} - E_{\text{N}_4\text{C}} - E_{\text{Mn}}$$

$$E_c = \frac{E_c - \sum N\mu M_n}{N}$$

The difference between the binding energy (E_b) and the cohesive energy (E_{coh}) provides a qualitative measure of aggregation tendency^{44,45}. For the Fe-, Co-, and Ni-based systems, the $E_b - E_{\text{coh}}$ values are consistently negative, spanning a range from -0.24 to -13.80 eV. Physically, these negative values indicate that the metal-support interaction (MSI) is the dominant energetic force, effectively surmounting the inherent cohesive energy of the bulk-like metal phase. The N-doped graphene lattice acts as a thermodynamic trap, anchoring nanoclusters against thermally driven agglomeration and preserving the active surface area^{44–46}. To further elucidate the interfacial interactions, the Ni–N bond distances were analysed. In all configurations, these bond lengths are shorter than the sum of the ionic radii of Ni and N (Supplementary Table S2), indicating the presence of strong orbital interactions involving donation and back-donation between the nitrogen atoms of the support and the metal atoms within the nanocluster⁴⁷. The optimised structure of a representative Ni₁₉ cluster supported on an N₄C supercell ($6 \times 5 \times 1$) is depicted in Figure 2b. In this configuration, each nitrogen atom is coordinated to two adjacent carbon atoms within the graphene lattice and simultaneously interacts with two Ni atoms from the cluster. For clarity, the optimised geometries of all systems are presented in Figures S1–S3, where the graphene edges are hydrogen-passivated to eliminate spurious boundary effects (Figure 2c). Within this framework, the focus was on Fe-, Co-, and Ni-based nanoclusters with odd numbers of atoms, motivated by their superior catalytic performance relative to even-numbered counterparts^{40,48,49}. Also, from a structural perspective, these odd nuclearities correspond to highly symmetric structures built around a single central atom. This perfectly aligns with the Mackay icosahedral "magic-number" series (13, 55, 147, ...), which gives

structurally precise odd-numbered highly stable cluster geometries (Mackay, 1962)⁵⁰. On the other hand, from the electronic perspective, odd-numbered clusters typically possess a singly occupied molecular orbital (SOMO)⁵¹. This unpaired electron promotes spin-polarised charge transfer and improves charge redistribution during the oxygen evolution reaction (OER). As OER involves converting singlet-state reactants ($\text{OH}^-/\text{H}_2\text{O}$) into triplet-ground-state O_2 , it is an intrinsically spin-selective process^{52,53}. Notably, odd-numbered clusters tend to display moderate adsorption strengths, consistent with the Sabatier principle, enabling a balance between intermediate stabilisation and facile desorption, which is critical for optimising OER kinetics^{48,54}. To synergistically combine high intrinsic activity with structural robustness, N-doped graphene is employed as the supporting substrate, enhancing both nanocluster stability and electronic coupling. To probe these effects in detail, the charge density difference of the $\text{Ni}_{19}@\text{NC}$ system was computed (Figure 2c), with additional systems presented in Figure S5. The Bader charge analysis reveals a significant charge transfer from the support to the Ni_{19} nanocluster (Table S6), quantified as 1.5464 e. This interfacial charge transfer stabilises the nanocluster while tuning its electronic structure and intermediate binding energetics.

To further elucidate the electronic interactions, partial density of states (PDOS) analysis was performed. The PDOS of $\text{Ni}_{19}@\text{N}_4\text{C}$ is presented in Figure 2d, with additional systems shown in Figure S6. A pronounced overlap between the d-states of Ni/Co/Fe and the p-orbitals of the coordinated nitrogen atoms is clearly observed, indicating strong interfacial hybridisation. Such electronic coupling not only stabilises the supported nanoclusters but also enhances charge transport, thereby promoting catalytic activity⁵⁵. To quantify the bonding characteristics, crystal orbital Hamilton population (COHP) analysis was carried out for systems⁵⁶. The Ni–N COHP for $\text{Ni}_{19}@\text{N}_4\text{C}$ is shown in Figure 2e, while corresponding results for a few other systems are provided in Figure S7. The observed moderate negative ICOHP values (-0.488eV for $\text{Ni}_{19}@\text{N}_4\text{C}$) indicate sufficient orbital overlap between the metal nanocluster and nitrogen and therefore efficient bonding interactions, further enhancing stability^{57,58}. The thermal and dynamic stability of the supported nanoclusters was further evaluated using ab initio molecular dynamics (AIMD) simulations within the canonical (NVT) ensemble at 350 K for 10000 fs. As illustrated in Figure 2g for $\text{Ni}_{19}@\text{N}_4\text{C}$ and in Figures S8 for the remaining **systems**, only minor energy fluctuations are observed, attributable to intrinsic thermal vibrations. The absence of abrupt energy deviations and structural integrity throughout the analysis confirms the thermodynamic robustness and experimental synthesizability of these systems⁵⁹.

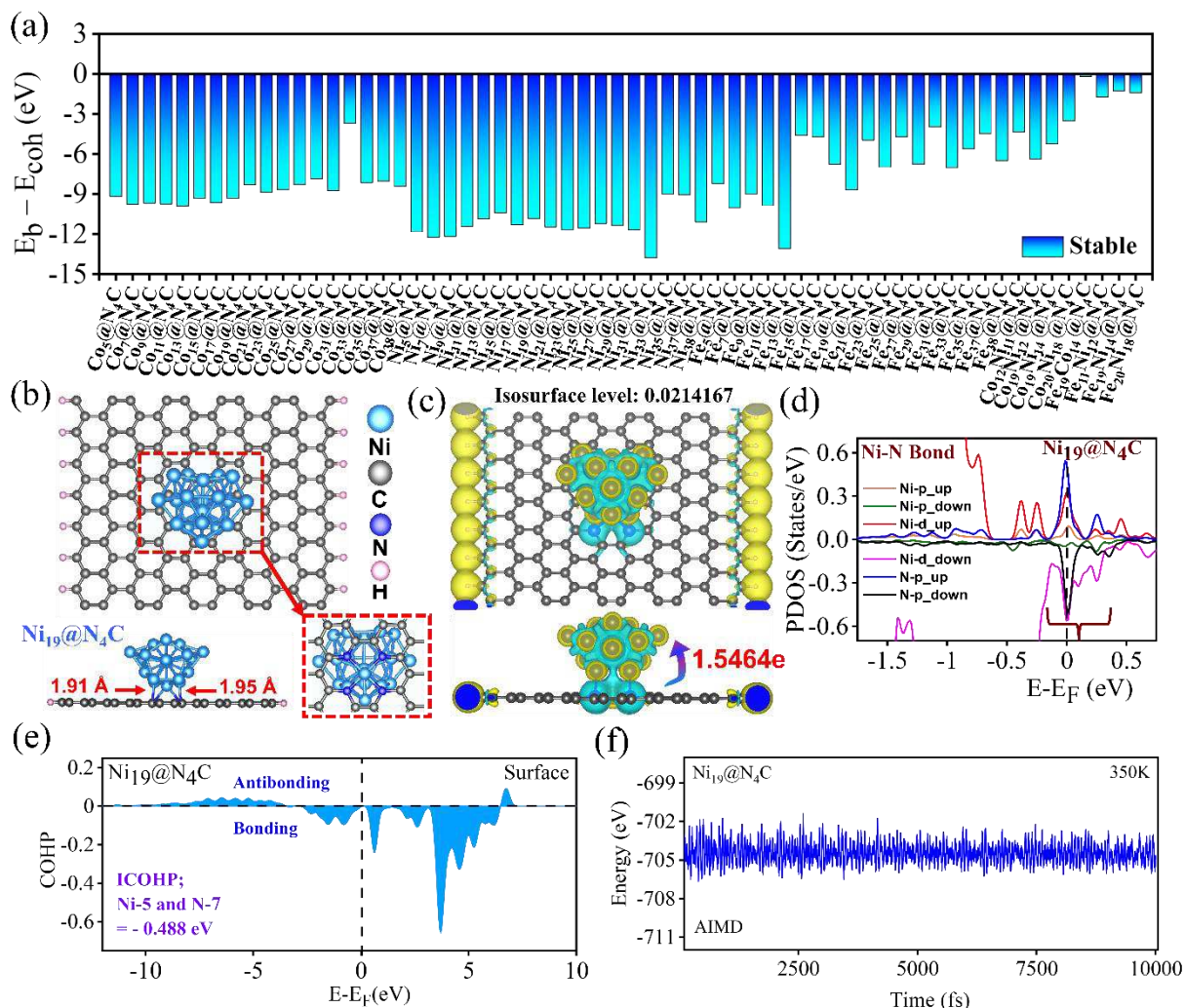


Figure 2: (a) The obtained binding energy per atom for the 62 $M_n@N_4C$ s; (b) Represents the optimised structure in a planner and side view of the $Ni_{19}@N_4C$; (c) Charge density difference plot of $Ni_{19}@N_4C$ interface; (d) Partial density of states of $Ni_{19}@N_4C$ showing orbital hybridisation near the fermi level; (e) presented the COHP plot between the Ni-5 atom and N-7 atoms of $Ni_{19}@N_4C$ interface; (f) AIMD simulation at 350K of the $Ni_{19}@N_4C$ calculated using canonical ensemble.

3.2 Free Energy Relationship

The oxygen evolution reaction (OER) activity of the $M_n@N_4C$ systems was systematically evaluated by constructing Gibbs free-energy profiles along the four-electron oxidation pathway, from which the corresponding overpotentials (η) were derived. These profiles capture the reaction energetics of the OER process proceeding via the elementary steps ($* + 4OH^- \rightarrow *OH \rightarrow *O \rightarrow *OOH \rightarrow * + O_2 + 2H_2O$). Representative free-energy profile of $Ni_{19}@N_4C$ at the 3-Ni site is presented in Figure 3a. This site exhibits the lowest overpotential among all 62 screened catalysts, highlighting its superior catalytic activity and selectivity toward OER. Free-

energy profiles for other highly active sites are provided in Figures S15–S20, with the reaction pathway schematically illustrated in Figure 3b^{60,61}. Analysis of the free-energy profile (Figure 3a) indicates that the formation of *O and *OOH, corresponding to the third and fourth proton-coupled electron transfer steps, constitutes the most energetically demanding processes among this class of catalysts, thereby identifying them as the rate-determining steps across these catalysts. This observation is consistent with prior reports; for example, *O and *OOH adsorption have been shown to govern the OER overpotential in CoFe/Fe₃C systems⁶². Notably, for Ni₁₉@N₄C, the *O→*OOH transformation shows the highest uphill free-energy change (0.44 eV), thereby making it the potential-determining step (PDS) at the most active site (3-Ni site, Figure S4). The optimised geometries of the key intermediates (*OH, *O, and *OOH) are shown in Figure S9. Consistently, other favourable sites also display low overpotentials, as reflected in their corresponding free-energy profiles in Figures S15–S20. The observed low overpotentials of the catalyst can be attributed to the favourable electronic structure of the nanoclusters, which promotes effective electronic coupling with reaction intermediates and enables optimal adsorption strengths of key reaction intermediates essential for efficient OER kinetics. Such behaviour is consistent with previous reports; for instance, Talib et al. demonstrated that Co₁/PMA achieves an overpotential of 0.45 eV due to balanced intermediate binding⁶³. Experimental studies further corroborate this principle – Chen et al. reported that Ni(OH)₂ exhibits an overpotential of 196 mV at 10 mA cm⁻² with excellent durability when intermediate adsorption is appropriately optimised⁶⁴. Similarly, Zhang et al. showed that suitable values of ΔG_{OOH} and $\Delta G_{\text{O}} - \Delta G_{\text{OOH}}$ lead to enhanced OER performance in Ni₂B₂ and Co₃B₄ MXenes⁶⁵. Motivated by these insights, we systematically examined the scaling relationships between ΔG_{OOH} , $\Delta G_{\text{O}} - \Delta G_{\text{OOH}}$ and negative overpotential (- η) across 211 active sites derived from 62 nanoclusters supported on N₄C. In line with the Sabatier principle, optimal catalytic performance requires a balanced interaction between the catalyst surface and reaction intermediates, wherein adsorption is neither too strong nor too weak^{66,67}. Consequently, OER activity follows a characteristic volcano-type dependence on adsorption energetics. Catalysts located on the weak-binding (left) side of the volcano exhibit insufficient adsorption and activation of oxygen intermediates, whereas those on the strong-binding (right) side suffer from hindered product desorption and sluggish surface regeneration^{49,68}. Maximum activity is therefore achieved near the volcano apex, where intermediate binding energies are optimally balanced⁶⁹. As shown in Figure 3c, $\Delta G_{\text{O}} - \Delta G_{\text{OOH}}$ exhibits relatively weak linear scaling with the overpotential, with correlation coefficients of $R^2 = 0.78$ for the weak-binding (left) region and $R^2 = 0.62$ for the strong-binding (right) region of the volcano. These modest correlations

indicate that $\Delta G_O - \Delta G_{OOH}$ is not a reliable descriptor for predicting OER activity in nanocluster-based systems. This deviation from conventional scaling behaviour can be attributed to the intricate electronic structures and quantum confinement effects inherent to nanoclusters, which fundamentally distinguish them from bulk catalysts⁷⁰. Consistently, ΔG_{OH} and ΔG_O alone also fail to capture the observed overpotential trends with R^2 of 0.15 and 0.30, respectively (Figure S10). In contrast, ΔG_{OOH} demonstrates a significantly improved correlation with the negative overpotential ($-\eta$), as shown in Figure 3d, with R^2 values of 0.91 and 0.98 for the weak- and strong-binding regions, respectively. It is worth noting here that ΔG_{OOH} shows the strongest correlation among the tested thermodynamic descriptors. Notably, active sites with ΔG_{OOH} values in the range of 0.5-1.5 eV exhibit the highest catalytic activity, reflecting an optimal balance between intermediate adsorption and product desorption. Deviations toward either excessively strong or weak adsorption result in increased overpotentials, consistent with the Sabatier principle and the requirement for balanced surface binding in efficient oxygen electrocatalysis⁶¹. Thus, it can be inferred that ΔG_{OOH} is a reasonable energetic descriptor to interpret and predict OER activity for Co, Ni, and Fe-based nanoclusters supported on N₄C. For any given catalyst space, these could be represented as region 1 (weak binding) and Region 2 (strong binding), with the peak point or “suitable spot = 0.8032” being the boundary between the two. The weak-binding Region-1 in volcano plot in Figure 3d follows the relation:

$$\eta = 0.9070 - 0.4734 \times \Delta G_{OOH} \quad \text{----- (E1)}$$

with an R^2 value of 0.91, and the strong-binding side (Region-2) is described by:

$$\eta = -0.19975 + 0.95142 \times \Delta G_{OOH} \quad \text{----- (E2)}$$

with an R^2 value of 0.98 (Figure S11). It is now possible to use equations 1 and 2 to predict the overpotential for this class of catalysts.

The Hubbard U correction ($U = 6$ eV) was used to account for the strong electron-electron correlation effects at Ni centres. This value is often used for Ni oxides and oxyhydroxides, but is also suitable here because the surface Ni atoms are partially oxidized by the adsorption of oxygenated intermediates (*OH, *O, and *OOH), resulting in a more oxide-like electronic environment than that of bulk metallic Ni. Previous computational studies have applied similar U values^{33,71–73}. The resulting free-energy profile for the 3-Ni active site of Ni₁₉@N₄C (Figure S12) reveals a further reduction in the OER overpotential to 0.39 V, indicating enhanced catalytic activity. This improvement arises from the increased localisation of Ni d-electrons induced by the U correction, which alleviates the overestimated Ni–O binding strength typically obtained in conventional GGA-PBE calculations⁷⁴. Consequently, the adsorption free energies of the *O and *OOH intermediates shift closer to the

thermodynamically optimal regime defined by the Sabatier principle, thereby lowering the potential-determining barrier and enhancing the overall reaction efficiency³².

To further elucidate the reaction kinetics, the nudged elastic band (NEB) method was employed to identify transition states and map the minimum energy pathways (MEPs) for each elementary step. By optimising a series of intermediate images that connect the initial and final states, the climbing image NEB approach minimises forces perpendicular to the reaction coordinate, thereby enabling an accurate description of the transition pathway. The optimised initial, intermediate, and final configurations for the 3-Ni active site of Ni₁₉@N₄C are presented in Figure S9. As shown in Figure S13, the steps (1→2, 2→3 and 3→4) are endothermic, with activation barriers of 0.72, 0.89 eV and 0.69 eV, respectively, relative to the preceding state. Among these, the 2→3 transition state possesses the highest activation barrier, thereby identifying it as the rate-determining step. Notably, the calculated barriers for all steps are lower than those reported for comparable systems, highlighting the intrinsic kinetic advantage of this catalyst for OER⁷⁵. The reduced activation energy associated with the rate-determining step directly facilitates faster reaction kinetics, leading to lower overpotentials and enhanced catalytic performance. Consistent with these theoretical predictions, experimental results demonstrate that Ni@N₄C outperforms Co@N₄C and Fe@N₄C in OER activity, in line with its more favourable reaction kinetics relative to the other catalysts.

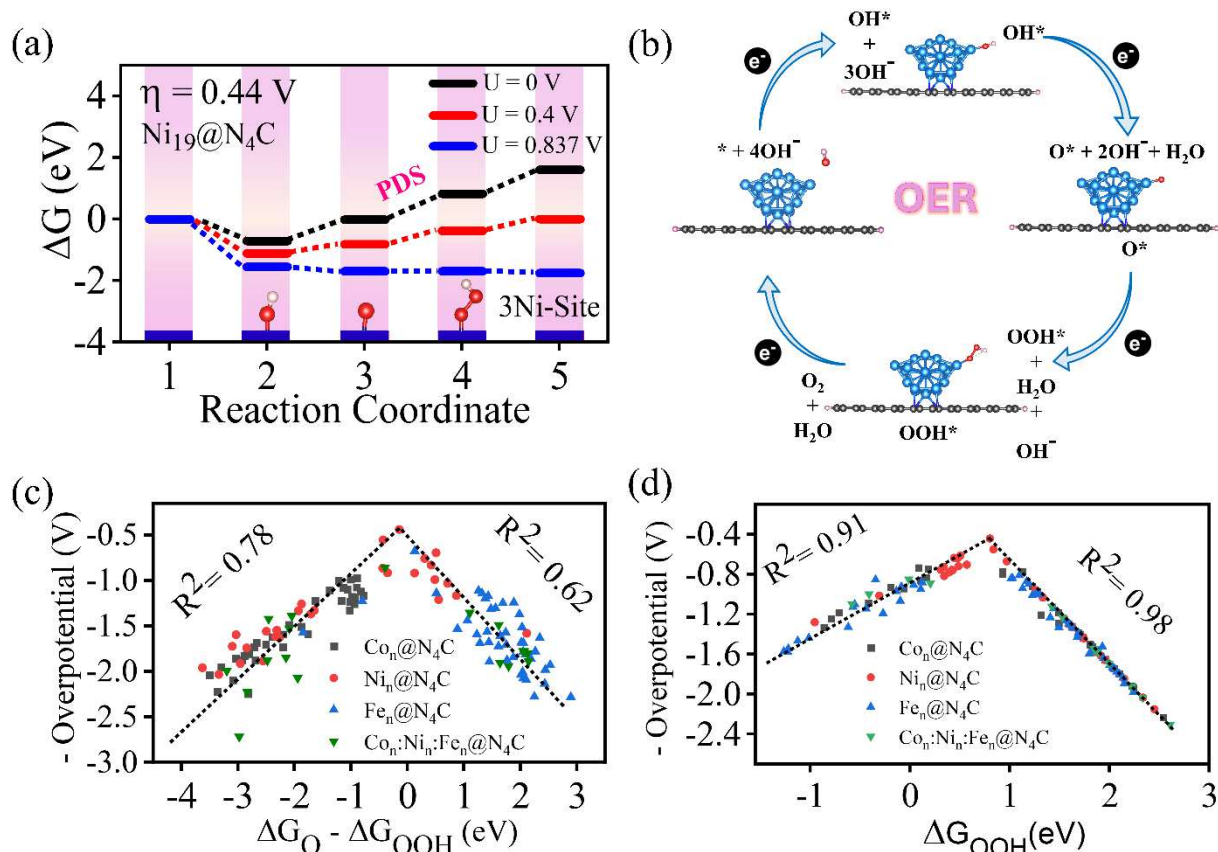


Figure 3: (a) The free energy profile for OER of 3-Ni active site of $\text{Ni}_{19}@\text{N}_4\text{C}$; (b) presents the schematic representation of steps involved in OER catalytic activity in metal nanoclusters; (c-d) represents the volcano plot between ΔG_{OOH} , $\Delta G_{\text{O}} - \Delta G_{\text{OH}}$ and negative overpotential, respectively.

3.3 Electronic Structure Analysis

Charge redistribution at the metal-adsorbate interface is a key parameter governing the activity of transition-metal nanoclusters toward the oxygen evolution reaction (OER)⁷⁶. To elucidate the electronic redistribution upon intermediate adsorption, charge density difference (CDD) plots were evaluated (Figure 4a)⁷⁵. The spatial separation of electron accumulation and depletion regions (green and yellow isosurfaces) evidences a net electron flow from the metal centres toward the adsorbates, indicative of metal-oxygen hybridisation of varying strength across intermediates. A quantitative assessment reveals that $\text{Ni}_{19}@\text{N}_4\text{C}$ transfers 0.287 e, 0.319 e, and 0.341 e to $^*\text{OH}$, $^*\text{O}$, and $^*\text{OOH}$, respectively. Notably, the magnitude of charge exchange at the potential-determining step ($^*\text{O} \rightarrow ^*\text{OOH}$) remains moderate relative to $\text{Co}_{31}@\text{N}_4\text{C}$ and $\text{Fe}_{29}@\text{N}_4\text{C}$, which exhibit more pronounced electron transfer. Specifically, $\text{Co}_{31}@\text{N}_4\text{C}$ shows charge transfer values of 0.294 e, 0.390 e, and 0.379 e to $^*\text{OH}$, $^*\text{O}$, and $^*\text{OOH}$, respectively, while $\text{Fe}_{29}@\text{N}_4\text{C}$ yields 0.277 e, 0.301 e, and 0.314 e for the same intermediates (Figure 4a; Tables S3–S5). The comparatively balanced charge redistribution in $\text{Ni}_{19}@\text{N}_4\text{C}$ during $^*\text{OOH}$

formation suggests neither excessive electron donation nor insufficient interaction, resulting in a more optimal adsorption of intermediates. This behaviour is in line with the Sabatier principle, which dictates that catalytic efficiency is maximised when the interaction strength between the catalyst and intermediates is optimally tuned, i.e., strong enough to enable intermediate formation, yet sufficiently weak to allow their efficient desorption⁷⁷. In this context, the moderate charge transfer associated with Ni₁₉@N₄C reflects an effective electronic coupling between the active site and oxygenated species, leading to an energetically balanced adsorption landscape and improved OER kinetics, in agreement with prior reports[53,62]. Among the key electronic descriptors, the degree of d-orbital occupancy, defined as $O_d = \int_{-\infty}^{E_F} D_d(E) dE$, exerts a dominant influence on catalytic activity. A partially filled d-band establishes a fine balance between bonding and antibonding states near the Fermi level, thereby enabling optimal electron exchange with adsorbates and yielding favourable adsorption energetics[61,63]. To quantify this effect, O_d , ΔG_{OOH} , and the associated overpotentials were systematically plotted across the most active Co-, Ni-, and Fe-based sites for a few systems as shown in Figure 4b. The resulting trend exhibits a characteristic volcano-type dependence, consistent with the Sabatier principle, where a moderate d-band occupancy, represented by the Ni-3 site in Ni₁₉@N₄C, corresponds to the lowest overpotential. In contrast, higher d-band occupancies, such as those observed at the Co-3 site in Co₂₀Ni₁₈@N₄C and the Fe-2 site in Fe₁₉Co₁₄@N₄C, lead to an increased population of occupied d-states below the Fermi level. This electronic configuration strengthens metal-oxygen bonding, resulting in overly stable intermediates that need more energy during desorption and elevated overpotentials (Figure S15a,b)⁷⁹. Conversely, insufficient d-state filling weakens adsorption, also reducing the catalytic efficiency. The Ni₁₉@N₄C system achieves an optimal balance of the metal-oxygen interaction, yielding a ΔG_{OOH} value close to the thermodynamic optimum and thereby enhancing catalytic performance. More broadly, tuning the d-state filling modulates the metal-oxygen interaction such that it remains sufficiently strong to enable bond formation while weak enough to permit facile desorption of intermediates. This optimal interaction regime is consistent with the Sabatier principle and directly contributes to lowering the overall OER overpotential. To further elucidate the underlying bonding characteristics, crystal orbital Hamilton population (COHP) analysis was performed for the Ni₁₉@N₄C system, focusing on the Ni-3 active site following *OH, *O, and *OOH adsorption (Figures 4e and S14)^{56,80}. The corresponding integrated COHP (ICOHP) values are uniformly negative, indicating that bonding interactions dominate over antibonding contributions across all intermediates. Specifically, the ICOHP values of Ni-3 site and O bond in Ni₁₉@N₄C are -0.840, -0.981, and -1.104 eV for *OH, *O, and *OOH

adsorption, respectively, signify a moderate interaction strength that is sufficient to stabilise intermediates while still allowing their efficient conversion. To place this in perspective, analogous analyses were carried out for the Fe-1 site in Fe₂₉@N₄C and the Co-3 site in Co₃₁@N₄C (Figure 4(c,f,d)). Upon *OOH adsorption, these systems exhibit substantially less negative and more negative ICOHP values of -0.93295 eV (Fe-O) and -1.5283 eV (Co-O), reflecting weaker and stronger metal-oxygen interactions respectively. Similarly for Co-3 site in Co₂₀Ni₁₈@N₄C and the Fe-2 site in Fe₁₉Co₁₄@N₄C ICOHP values are -1.8215 and -1.9303 respectively (Figures S15c,d). Such strong binding, although beneficial for intermediate stabilisation, tends to impede subsequent reaction steps by limiting intermediate desorption. In contrast, the Ni-3 site preserves an moderate bonding strength, effectively positioning the metal-oxygen interaction within an optimal window. This balanced interaction facilitates both stabilisation and turnover of intermediates, thereby promoting faster reaction kinetics and superior OER activity[61].

The evolution of d-band occupancy (O_d) upon adsorption of oxygenated intermediates was quantified from projected density of states (PDOS) analyses for representative active centres, 3-Ni (Ni₁₉@N₄C), 3-Co (Co₃₁@N₄C), and 1-Fe (Fe₂₉@N₄C) (Figures 4(g-i) and S4). For Ni₁₉@N₄C, the pristine Ni-3 site exhibits an O_d value of 8.426, which gradually decreases upon adsorption of *OH, *O, and *OOH to 8.383, 8.379, and 8.344, respectively. This controlled and moderate depletion of d-electron occupancy reflects balanced charge transfer to the adsorbates, indicative of optimised metal-oxygen interactions along the reaction pathway⁷⁵. A comparable trend is observed for Co₃₁@N₄C, where O_d decreases from 7.307, before adsorption, to 7.234 (*OH), 7.239 (*O), and 7.256 (*OOH). The overall shift indicates substantial electronic redistribution upon adsorption, with a magnitude slightly exceeding that of the Ni-based site. For Fe₂₉@N₄C, O_d decreases from 6.238 to 6.199 (*OH), 6.186 (*O), and 6.202 (*OOH), indicating a greater degree of electron depletion than in Ni₁₉@N₄C. The moderate degree of d-band depletion in the Ni-3 site relative to Co and Fe centres enables a favourable balance between adsorption and desorption steps, facilitating enhanced OER kinetics. Overall, the reduction in d-band occupancy across all M_n@N₄C systems confirms net electron transfer from the metal centres to oxygenated intermediates, promoting metal-oxygen bond formation⁶². This redistribution of electronic density stabilises key reaction intermediates and lowers the energetic barriers associated with elementary OER steps. Collectively, these results substantiate the critical role of d-band modulation in governing adsorption energetics and catalytic activity, reinforcing the established correlation between d-band electronic structure and OER performance.

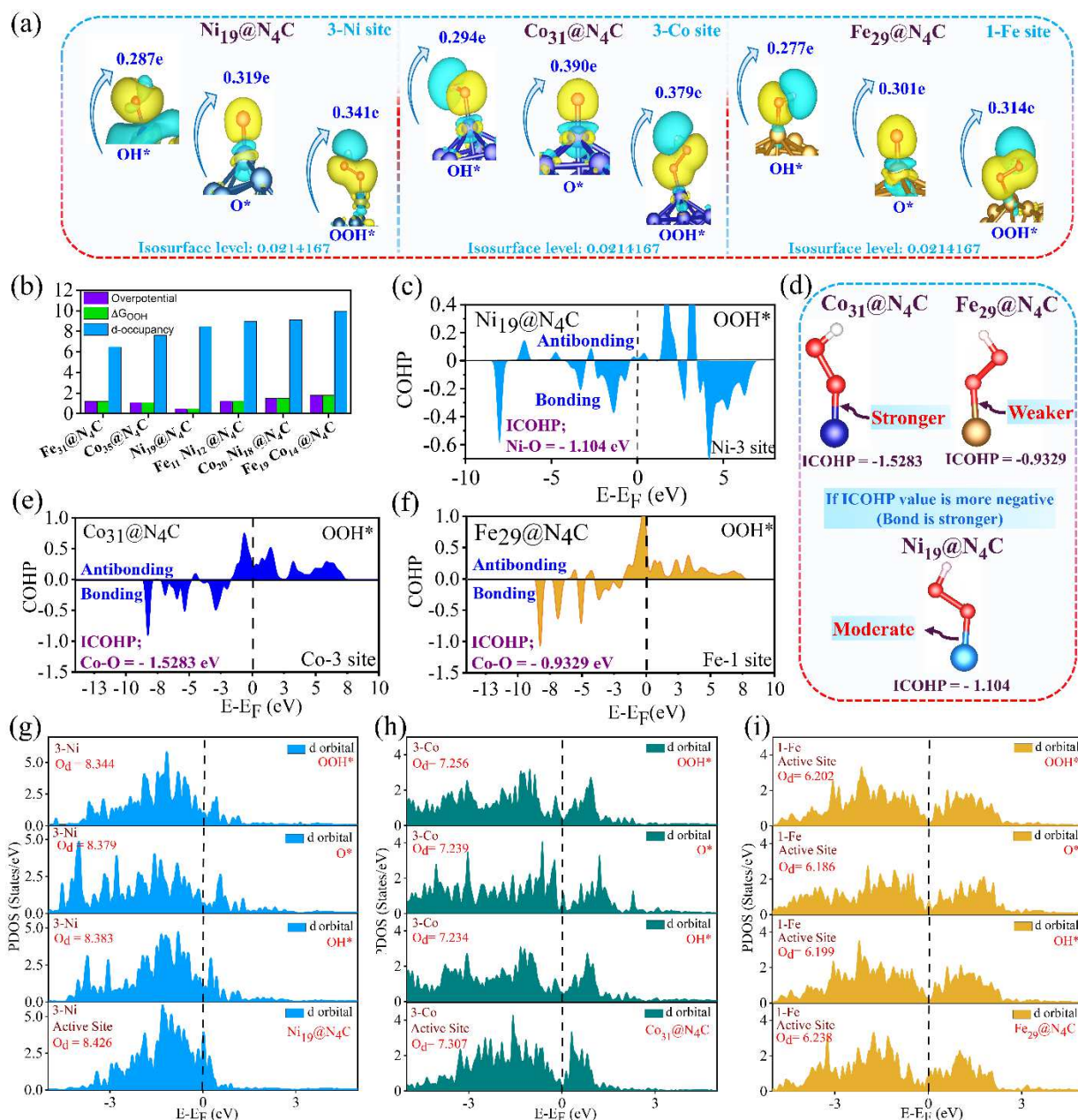


Figure 4: (a) Charge density difference plots for the best active sites of $\text{Ni}_{19}@\text{N}_4\text{C}$, $\text{Co}_{31}@\text{N}_4\text{C}$, and $\text{Fe}_{29}@\text{N}_4\text{C}$ after intermediate adsorption; (b) Comparison of the best catalysts in each class (Co, Ni, Fe, and their alloys) with corresponding overpotential, ΔG_{OOH} , and d-occupancy. (c) COHP plot of the Ni-3 site in $\text{Ni}_{19}@\text{N}_4\text{C}$ after OOH^* adsorption; (d) ICOHP values comparison for the best sites of $\text{Ni}_{19}@\text{N}_4\text{C}$, $\text{Fe}_{29}@\text{N}_4\text{C}$ and $\text{Co}_{31}@\text{N}_4\text{C}$ after OOH^* adsorption; (e) COHP plot of the Co-3 site in $\text{Co}_{31}@\text{N}_4\text{C}$ after OOH^* adsorption; (f) COHP plot of Fe-1 site in $\text{Fe}_{29}@\text{N}_4\text{C}$ after OOH^* adsorption; (g-i) PDOS analysis of active sites (3-Ni, 3-Co, 1-Fe) before and after intermediate adsorption.

3.4 Descriptor and Machine Learning

To elucidate the catalytic complexity of nanocluster-based Mn@N₄C systems toward OER, linear correlation analysis was performed across 211 distinct active sites using 72 PDOS-derived electronic descriptors correlated against ΔG_{OOH} (Section S3; Figures S21–S28). No descriptor yielded any meaningful linear correlation, confirming intrinsically non-linear reaction energetics. Among all 72, only 5 features, d-band occupancy ($R^2 = 0.0909$, $r = 0.3015$), $d_{z^2_occ}$ ($R^2 = 0.0977$, $r = 0.3126$), d_{xz_occ} ($R^2 = 0.0986$, $r = 0.3140$), $t2g_occ$ ($R^2 = 0.0928$, $r = 0.3046$), and z_c orbital occupancy ($R^2 = 0.1022$, $r = 0.3198$), showed Pearson coefficients barely exceeding 0.30, insufficient to rationalize activity variations. This breakdown of linear scaling relations stems from the tightly coupled interplay between local coordination environments and site-specific electronic structures, producing pronounced site-dependent variability that renders single-descriptor approaches inadequate for capturing the OER mechanism in this class of materials. An ML framework was developed by integrating multiple d-orbital-derived descriptors into a unified multi-descriptor framework²¹. This high-dimensional electronic structure representation explicitly captures inter-descriptor synergies and non-linearities inaccessible to single-parameter models, enabling physically grounded and reliable structure–activity descriptions for M_n@N₄C catalysts.

To establish a quantitative structure-property relationship between the electronic characteristics of M_n@N₄C catalysts and the OER-relevant descriptor ΔG_{OOH} , a systematic ML workflow encompassing data preprocessing, hierarchical feature selection, model training, optimisation, and validation was implemented on a dataset of 211 samples, each represented by 72 computed electronic descriptors. Before model construction, data preprocessing was performed to ensure numerical stability. Infinite values were converted to NaNs, and all entries containing missing values were removed, yielding a clean and robust dataset. All descriptors were standardised via StandardScaler (zero mean, unit variance) to prevent high-magnitude feature dominance and ensure efficient Principal Component Analysis (PCA) convergence. A two-stage feature selection strategy was then employed to obtain a compact and physically meaningful descriptor space. First, PCA-based filtering retained descriptors with significant contributions to the dominant variance (PC1 and PC2 capturing 62.96%) (Figure S29), reducing the feature set from 72 to 60. In the second stage, multicollinearity was eliminated using Pearson correlation and hierarchical clustering, where parameter pairs having a correlation of more than 0.90 ($|r| > 0.90$) were pruned by removing less informative descriptors (Figures S30, S33). This process yielded a final set of 14 non-redundant features for model development. The

curated dataset, including ΔG_{OOH} as the target, is provided in Table S7 of the Supplementary Information.

The reduced set of 14 non-redundant electronic descriptors was used to train and systematically benchmark eight regression models, namely MLR, RR, XGR, LASSO, BRR, SVR, KNN, and RFR. Detailed descriptions of these algorithms are provided in Supplementary Section 4, and the complete implementation workflow is documented in Supplementary Section 5. To ensure statistical reliability and eliminate bias arising from a single data split, each model was evaluated over 100 independent train–test iterations, employing randomised 80/20 partitions of the dataset²¹. Model performance was evaluated using the Root Mean Squared Error (RMSE) and the coefficient of determination (R^2). Among the models considered, the Random Forest Regressor (RFR) exhibits the highest predictive accuracy, achieving RMSE = 0.077 and $R^2 = 0.92$ on the training set, and RMSE = 0.094 with $R^2 = 0.87$ on the test set (Figure 5a). The XGR model demonstrated performance of test RMSE of 0.123 and R^2 values of 0.78 (training) and 1 (testing), indicating the model is overfitting (Figure 5c). The SVR model with RBF kernel also delivers competitive performance, with a test RMSE of 0.130 and R^2 values of 0.75 (testing) and 0.61 (training) (Figure S31(d)). In contrast, linear approaches such as MLR fail to capture the intrinsic non-linearity between electronic descriptors and ΔG_{OOH} , resulting in inferior performance (test RMSE = 0.182; training RMSE = 0.239 and $R^2 = 0.41$ training, 0.29 testing) (Figure 5b). The BRR model demonstrates poor predictive capability, with a test RMSE of 0.200 and R^2 values of 0.22 (training) and 0.22 (testing) (Figure 5d). Other models like LASSO, Ridge and KNN performed inadequately (Figure S31). Learning curve analysis (Figure S34) shows a systematic reduction in RMSE with increasing dataset size, accompanied by convergence of training and validation errors in the RFR model. This behaviour indicates that the model captures underlying physical trends rather than overfitting, confirming that the dataset size is sufficient for reliable generalisation. To elucidate the electronic factors governing OER energetics in $M_n@N_4C$ systems, normalised feature importance and SHAP analyses were performed on the models (Figures 5e, S32). The normalised feature importance of the RFR model reveals that the dominant descriptor is the density of low-energy spin-down d-frontier states near the Fermi level, with a normalised importance of 0.235. This highlights that the location and nature of the frontier d orbitals closest to E_F (Fermi level) play a decisive role in regulating adsorbate binding and facilitating interfacial charge transfer during the oxygen evolution reaction, which is also consistent with the SHAP analysis^{20,21} of RFR and XGR, where d-band frontier down is the 1st and 2nd dominant descriptors, respectively. This descriptor specifically measures the degree of localisation of the spin-down d orbitals in $\pm 1\text{eV}$ from the

Fermi energy, where their availability is optimal for interaction and charge exchange with OER intermediates. Mathematically (Equation 11, Supplementary information), the numerator reflects the projected density of these active d-states confined to the $\pm 1\text{eV}$ range around E_F , directly corresponding to states capable of engaging in catalytic bonding, while the denominator normalises this by the total spin-down d-band contribution, ensuring a properly weighted measure of frontier-state involvement. Efficient OER activity depends on how well the frontier d orbitals of the metal atoms overlap and align with the orbitals of the adsorbed species as captured collectively within the multi-descriptor framework. When their energy levels and spatial shapes match, strong hybridisation occurs, which helps in breaking and forming O-O bonds during the reaction. In nanocluster catalysts, where energy levels are discrete rather than continuous, the exact position and orientation of these frontier orbitals become even more important⁸¹. Therefore, for such systems, localised orbital properties provide a more reliable indicator of catalytic performance than general or averaged electronic descriptors. Notable contributions from d-occupancy (0.141) and the spin-resolved terms, d-up occupancy (0.139), demonstrate that both the total d-electron count and the spin asymmetry influence intermediate binding strength. Descriptors with comparatively lower, but still meaningful, correlation to the energy descriptor, including d-down centre (0.077), d-frontier up (0.071), Bader charge (0.060), $d_{xz_down_fermi}$ (0.054), Wd corrected d-band (0.043), reinforce that even subtle shifts in d-band energetics can influence the thermodynamics of OOH* adsorption. Additionally, the contribution of d-frontier up (0.071) highlights the role of local charge redistribution and metal-support electron transfer in modulating catalytic performance. Similarly, SHAP feature importance is also analysed to check which descriptors have the most influence on the models, e.g., RFR, XGR, SVR, MLR and BRR models. As shown in Figure S33, d-frontier down is the most dominant descriptor for the best model: RFR, holding the best SHAP feature importance value. Collectively, these trends highlight that OER activity in metal nanoclusters arises from a delicate interplay of localised frontier d-state energetics at the active sites.

Validation of the ML models was performed using an independent test set of 10 active sites selected from $\text{Co}_{17}@\text{N}_4\text{C}$, $\text{Ni}_{23}@\text{N}_4\text{C}$, $\text{Fe}_{27}@\text{N}_4\text{C}$, and $\text{Fe}_{19}\text{Co}_{14}@\text{N}_4\text{C}$ catalysts (Table S8), where the parity plot evaluates the predictive performance of eight ML models in estimating the adsorption free energy (ΔG_{OOH}) values against the corresponding DFT-calculated values (Figure 5f, Table S9). Among all the models, the Random Forest Regressor (RFR) exhibits the best agreement with DFT, yielding an R^2 of 0.80, with data points closely distributed along the parity line, indicating reliable predictive capability. The KNN and XGR models show moderate performance ($R^2=0.76$ and 0.74 , respectively), followed by SVR (RBF kernel) ($R^2=0.45$); these

627 nonlinear models capture general trends but exhibit noticeable scatter. In contrast, linear models
628 such as MLR, Ridge, LASSO, and BRR completely failed to capture any trend, as reflected in
629 negative R^2 values. LASSO showed a flat line with R^2 of -0.06 and simply guessed the average
630 data value (1.10 eV) for every site, creating a rigid vertical line on the parity plot. While the
631 MLR model catastrophically failed ($R^2 = 0.02$). Similarly, Ridge($R^2 = -0.22$) and BRR($R^2 = 0.25$)
632 show total failure. The linear models are structurally incapable of capturing these complex,
633 multi-dimensional variations, leading them to underfit severely. Overall, the validation results
634 unveil that OER energetics in $M_n@N_4C$ nanoclusters are governed by strongly non-linear
635 variations in the electronic parameters. The interactions between oxygen-containing
636 intermediates and the active sites of metal nanoclusters are governed by complex non-linear
637 variations in d-orbital occupancy, shifts in d-frontier electronic states, and local charge-transfer
638 processes. These non-linear variations cannot be accurately described using linear models. They
639 can only be accurately decoded by flexible, ensemble-based ML models like Random Forest.
640 This mechanistic understanding supports a concise, physically meaningful descriptor set that
641 facilitates efficient and reliable machine-learning-driven screening of OER electrocatalysts.

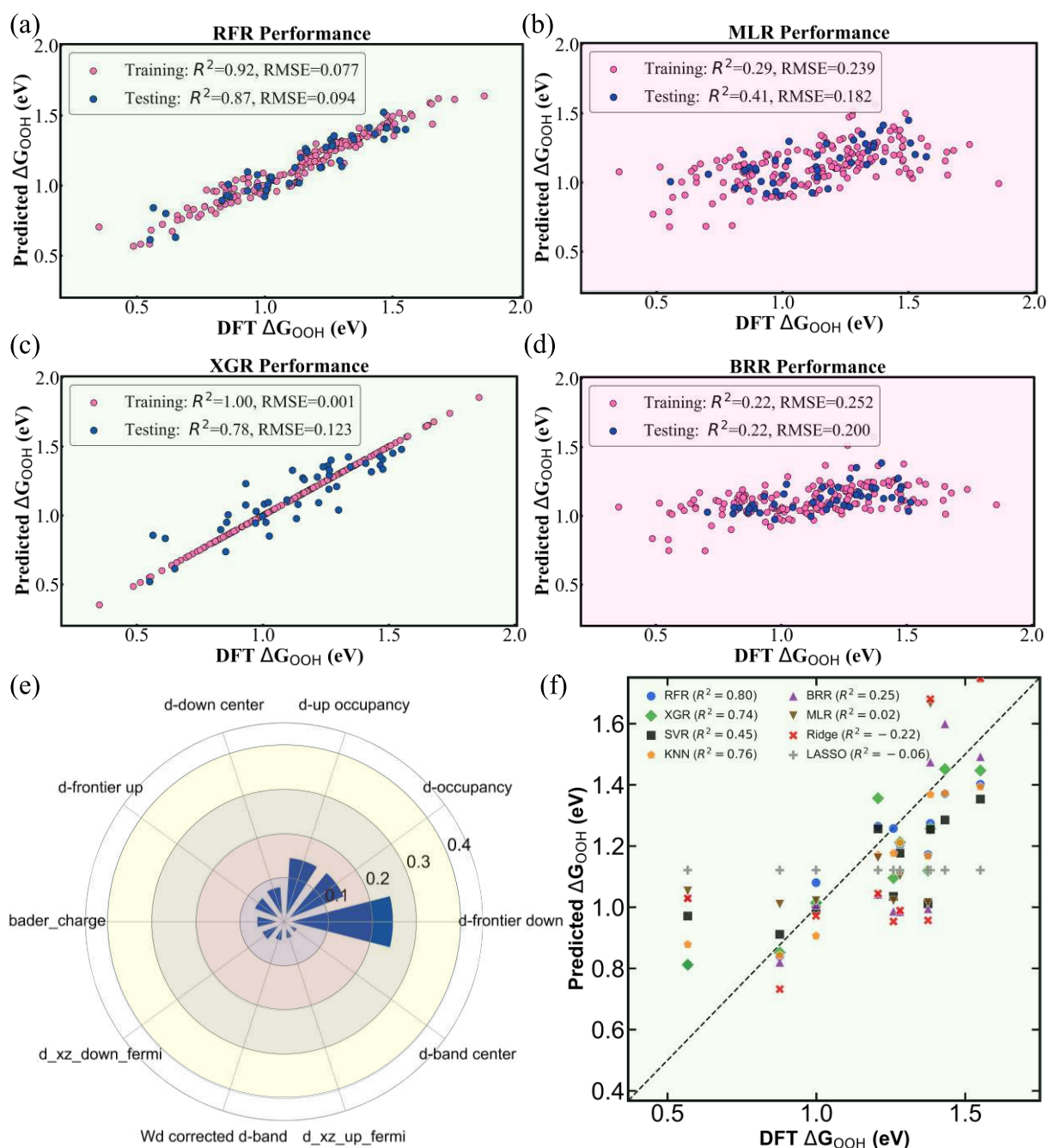


Figure 5: (a-b) Performance of non-linear machine-learning models including Random Forest Regression (RFR), XGBoost Regressor(XGR); (c-d) Performance of Linear machine Learning Models including: Multiple Linear Regression (MLR), and Bayesian Ridge Regression(BRR); (e) Feature-importance distribution radar plot, highlighting the ten most influential descriptors and their respective contributions to model prediction; (f) Parity plot for verification of DFT-calculated ΔG_{OOH} and corresponding ML predicted values of ΔG_{OOH} for 8 ML models in the 10 metal active sites.

3.5 Synthesis, Characterisation and Electrocatalytic Performance

To validate the structure-activity relationship, a strategy to rationally design the metal nanocluster on graphitic carbon supports was developed. Since atomically dispersed metal nanoclusters can be effectively stabilised through coordination with heteroatomic species,^{82–85} we adopted a one-step facile synthetic strategy (Figure S35a). In this strategy, the *in-situ* generation of C₃N₄ framework during pyrolysis serves as a stabilising scaffold for the nanocluster (for detailed synthesis procedure, please see Section 6. Supporting Information). The use of dicyandiamide (DCDA) ensures effective nitrogen incorporation, enabling strong metal-support interactions. As illustrated in Figure S35b, the PXRD patterns of all synthesised catalysts-NGC, FeNC@NGC, CoNC@NGC, and NiNC@NGC, display a prominent diffraction peak at ~26°, corresponding to the (002) plane of graphitic carbon. Additional diffraction peaks corresponding to metallic Ni, Co, and Fe nanoclusters verify successful nanocluster formation.⁸⁶ Detailed phase assignment and crystallographic analysis are provided in Supplementary Section S6: point 3.

Field emission scanning electron microscopy (FESEM) was employed to examine the surface morphology of the as-synthesised NiNC@NGC catalyst. As shown in Figures S35c and S35d, FESEM analysis reveals a uniform dispersion of Ni nanoclusters anchored on both the NGC sheets and tubular structures (Figure S35e-j). This is consistent with theoretical predictions. HR-TEM confirms lattice fringes corresponding to metallic phases along with graphitic carbon, indicating good crystallinity and structural integration. Additional morphological and diffraction analyses are discussed in Supplementary Section S6: point 4. X-ray photoelectron spectroscopy (XPS) was further utilised to determine the surface chemical environment and bonding configuration of the as-synthesised catalysts⁸⁷. The High-resolution XPS spectra in Figure S37 confirm the presence of C, N and O in all the samples, while Fe, Co, and Ni are also observed in FeNC@NGC, CoNC@NGC, and NiNC@NGC, respectively. The deconvoluted core-level spectra (Figure S36) further validate the chemical states of the respective metal nanoclusters and their interaction with the N-doped graphitic carbon support (Figure S38)^{88–91}. The deconvoluted C 1s peaks (Figure S36c) corresponding to sp² C=C, C–O/C–N, and C=O further verify the N-doped graphitic carbon support. Detailed spectral deconvolution and peak assignments are provided in Supplementary Section S6: point 5. In addition, the O 1s spectra (Figure S39) reveal two dominant oxygen species, adsorbed oxygen and chemisorbed H₂O, originating from partial aerobic oxidation. Overall, the XPS results

indicate the coexistence of metallic nanoclusters and their higher-valent states, stabilised by strong M-N_x coordination within the N-doped graphitic carbon framework.

All the electrochemical measurements were carried out using a three-electrode Rotating Disk Electrode (RDE) setup in 1 M KOH medium. The working electrode was prepared by drop-casting the catalyst ink onto an RDE and drying it under spinning conditions at 700-900 rpm until a uniform catalyst film was formed. Linear Sweep Voltammetry (LSV) was initially conducted at 1600 rpm, 10 mVs⁻¹ on NGC; however, the material was unable to attain the benchmark current density of 10 mA cm⁻², even under increased overpotential conditions, indicating limited electrocatalytic activity. In contrast, transition metal nanocluster-based systems (MNCs) demonstrated an enhancement in OER activity. This improvement can be attributed to the presence of metallic nanoclusters distributed over the nitrogen-doped graphitic carbon (NGC) matrix. Unlike the pristine NGC, the MNCs facilitated more efficient access of OER intermediates to the catalytically active sites on the metallic components, thereby promoting superior electrochemical performance. The *i*R corrected LSV polarisation curve of NiNC@NGC exhibited the lowest overpotential (η_{10}) of nearly 460 mV to reach the benchmark current density (Figure 6a, b and Figure S40), followed by CoNC@NGC (~500 mV), FeNC@NGC (~630 mV). Moreover, NiNC@NGC shows the OER performance which is closer to the state-of-the-art catalyst for example RuO₂ (η_{10} ~370 mV) (Figure S41). Additionally, as shown in bar diagram the NiNC@NGC shows better OER performance in term of overpotential at different current density. The corresponding Tafel slope (Figure 6c) of NiNC@NGC is 94 mV dec⁻¹ which is lower than that of the CoNC@NGC (108 mV dec⁻¹), FeNC@NGC (180 mV dec⁻¹), and NGC (404 mV dec⁻¹), suggesting the more facile OER kinetics at the surface of NiNC@NGC compared to other synthesized catalyst⁹⁰. Moreover, mass activity analysis (Figure 6d) revealed that NiNC@NGC exhibits nearly twice the activity of CoNC@NGC and approximately six times that of FeNC@NGC, indicating that NiNC@NGC possesses intrinsically superior OER activity among the studied systems. The fitted Nyquist plot (Figure 6e) obtained from Electrochemical impedance spectroscopy (EIS) confirms the lower charge-transfer resistance (R_{ct}) of NiNC@NGC (42.3 Ω) than other studied catalysts (Figure 6f, Figure S42 and Table S10). The double-layer capacitance (C_{dl}) was determined from the slope of the ΔJ versus scan rate plots, providing insight into the interfacial double-layer characteristics and the electrochemically active surface area of the electrode materials. Figure S43 and Figure 6g confirm the C_{dl} values for NiNC@NGC, CoNC@NGC, and FeNC@NGC were 6.395, 3.2, and 0.47 mF cm⁻², respectively.

Also, the electrochemically active surface area (EASA) was determined to check the surface available for adsorbing various oxygenated OER intermediates⁸⁹. Figure 6h confirms that NiNC@NGC exhibited a high EASA of 160 cm², significantly enhancing its OER activity compared to other electrocatalysts. The RRDE results (Figure 6i) reveal that the reversible OER process on NiNC@NGC is dominated by a desirable 4e⁻ pathway^{41,92}. In addition, the Faradaic efficiency (FE) of NiNC@NGC towards OER was determined by the RRDE test. The disc and ring current were maintained at 1600 rpm to evaluate the FE. The Pt ring was used with a constant potential to reduce the as-formed O₂ during OER. To check the robust performance of NiNC@NGC towards OER long-term chronoamperometric study was conducted (Figure 6j). The initial increase in the activity of the current is attributed to the activation of the exposed site of Ni nanoclusters and after prolonged time it shows very negligible performance decay (~20 %), suggesting the durability of Ni cluster towards OER performance. The dynamic stability of NiNC@NGC was further evaluated through 100 consecutive cyclic voltammetry (CV) cycles, a key metric for assessing the practical viability of electrocatalysts. Remarkably, NiNC@NGC exhibited an almost imperceptible change in terms of η_{10} (with a deviation of less than 20 mV), thereby confirming its outstanding long-term electrochemical stability as an OER catalyst (Figure 6k). The most important parameters for OER are summarised in Figure 6l, which clearly indicates the excellent activity of NiNC@NGC. The DFT calculations suggest that the overpotential of NiNC@NGC is 390 mV, which is in close agreement with the experimental one (460 mV). Also, the theoretically predicted activity trend of Ni > Co > Fe is validated. Therefore, it is observed that NiNC@NGC is a superior OER catalyst that is consistent with both theoretical and experimental results.

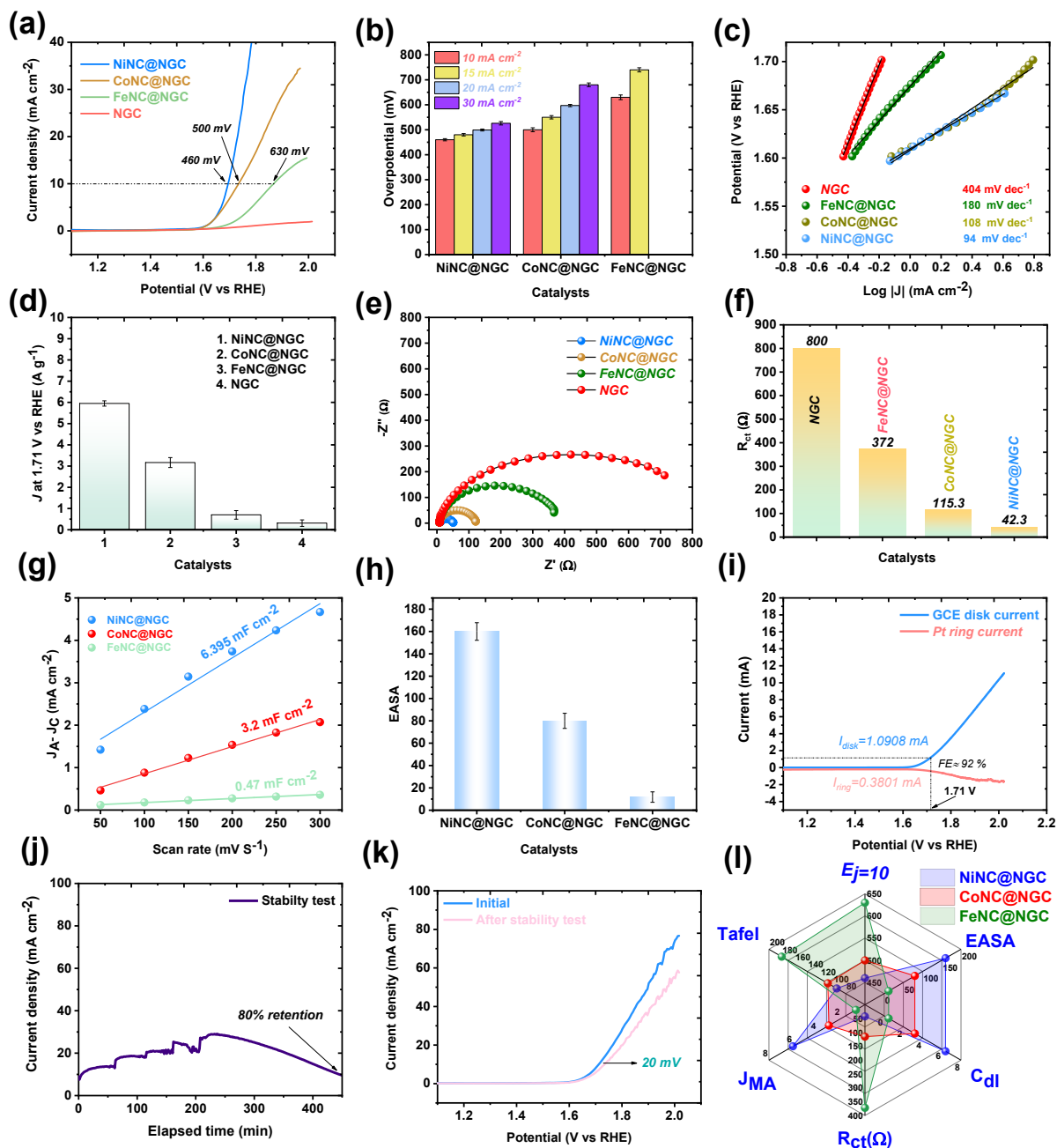


Figure 6: (a) OER LSV polarisation curves of different catalysts at 1600 rpm in 1 M KOH, (b) bar diagram of overpotential of the different catalysts at varying current densities, (c) Tafel slopes of three different catalysts, (d) OER mass activity of four synthesised catalysts, (e) Nyquist plot, (f) charge transfer resistance comparison plot, (g) C_{dl} value of the different catalysts, (h) corresponding EASA bar plot, (i) OER stability curve. (i) RRDE plot and corresponding Faradaic efficiency of NiNC@NGC at 1.71 V vs RHE, (j) OER chronoamperometry curve at 700 mV overpotential at 1600 rpm, (k) LSV polarisation after durability test, (l) radar plot showing the OER performance trend of three different catalysts.

4 Conclusions

This work set out to address a genuine open problem: why do metal nanoclusters behave the way they do as electrocatalysts, and can we predict and explain their activity from first principles? By systematically screening the activity of 62 Co, Fe, and Ni-based nanoclusters of varying size (5 to 38 metal atoms) supported on nitrogen-doped graphene, and evaluating 211 distinct surface sites through high-throughput DFT calculations, we show that this question can be answered in a physically interpretable and computationally tractable way. From this screen, Ni₁₉@N₄C emerges as the most active OER catalyst, giving the lowest overpotential of 0.39 V with the Hubbard U correction. These computational predictions were directly validated by experiment: the synthesised NiNC@NGC catalyst achieved an iR-corrected overpotential (η_{10}) of 460 mV and a Tafel slope of 94 mV dec⁻¹, performance on par with a commercial RuO₂ catalyst. These results clearly revealed Ni₁₉@N₄C as an alternative low-cost, high-stability, and high-activity OER catalyst, while experimentally validating the theoretically predicted activity trend of Ni > Co > Fe. The catalytic mechanisms of these 211 metal sites were also studied using eight different ML models: MLR, RR, SVR, RFR, BRR, XGR, LASSO, and KNN. Among the eight ML models, the RFR model performed the best with the highest R² of 0.87. In addition, ΔG_{OOH} was identified as the energy descriptor with R²>0.90, correlating with the catalytic overpotential(η). This work unravelled the electronic origins of OER catalytic activity in Co, Fe, and Ni-based Nanocluster catalysts anchored on N₄C graphene by finding d-band frontier down (df_↓) as the primary driver behind the complex nanocluster catalysis in the multi-descriptor framework and provided a rational design framework for advanced, efficient, and active electrocatalysts.

Author contributions

Arupjyoti Pathak: Conceptualisation, investigation, methodology, Coding, Machine learning, software, formal analysis, validation, visualisation, writing- original draft, writing-review & editing, Srijib Das: Experimental analysis, methodology, validation, writing-review & editing original draft (experimental part). Mukaddar SK: Formal analysis and validation of the DFT part, writing- original draft and review & editing. Tapas Kuila: Funding acquisition, Validation of the experimental part. Om Jadhav: Manuscript review editing and helping in analysis of ML data to Mr. Arupjyoti. Samrit Kumar Maity: Review and editing and helping in the analysis of ML data to Mr. Arupjyoti. Aniruddha Kundu: Experimental analysis, validation, Funding acquisition, writing - review & editing (experimental part). Ranjit Thapa: Conceptualisation,

Designed the Project, Supervision, Formal Analysis, Validation, Data Curation, writing-original draft, writing-review & editing, Project administration, funding acquisition.

Acknowledgments

The authors thank the National Supercomputing Mission (NSM) for the support for access to PARAM Rudra, for carrying out the computational work and the SRM University-AP, Andhra Pradesh for providing the central computational facilities. RT thanks the Anusandhan National Research Foundation (ANRF), India, for financial support (Grant No. CRG/2022/005423). This work got financial assistance by the Department of Science and Technology (DST), India, under the DST-INSPIRE Faculty Scheme, New Delhi (DST/INSPIRE/04/2020/000616). A.K. acknowledges financial support from the DST-INSPIRE Faculty Scheme (IFA20-MS-168). S.D. is thankful to the CSIR, Govt. of India, for the CSIR SRF-Direct fellowship (File No: 31/0019(23452)/2025-EMR-I).

Data Availability

The data that support the findings of this study are available from the corresponding author, R. T., upon reasonable request.

Ethics declarations

The authors declare no competing interests.

References

1. Jena, P. & Sun, Q. Super Atomic Clusters: Design Rules and Potential for Building Blocks of Materials. *Chem. Rev.* **118**, 5755–5870 (2018).
2. Tyo, E. C. & Vajda, S. Catalysis by clusters with precise numbers of atoms. *Nature Nanotech* **10**, 577–588 (2015).
3. Wang, L., Liu, H., Zhuang, J. & Wang, D. Small-Scale Big Science: From Nano- to Atomically Dispersed Catalytic Materials. *Small Science* **2**, 2200036 (2022).
4. Alonso, J. A. Electronic and Atomic Structure, and Magnetism of Transition-Metal Clusters. *Chem. Rev.* **100**, 637–678 (2000).

- 808 5. Wang, J. *et al.* Redirecting dynamic surface restructuring of a layered transition metal oxide
809 catalyst for superior water oxidation. *Nat Catal* **4**, 212–222 (2021).
- 810 6. Huang, Z.-F. *et al.* Chemical and structural origin of lattice oxygen oxidation in Co–Zn
811 oxyhydroxide oxygen evolution electrocatalysts. *Nat Energy* **4**, 329–338 (2019).
- 812 7. Garcés-Pineda, F. A., Blasco-Ahicart, M., Nieto-Castro, D., López, N. & Galán-Mascarós, J. R.
813 Direct magnetic enhancement of electrocatalytic water oxidation in alkaline media. *Nat*
814 *Energy* **4**, 519–525 (2019).
- 815 8. Shinde, S. S. *et al.* Ampere-hour-scale zinc–air pouch cells. *Nat Energy* **6**, 592–604 (2021).
- 816 9. Qiao, B. *et al.* Single-atom catalysis of CO oxidation using Pt₁/FeO_x. *Nature Chem* **3**, 634–641
817 (2011).
- 818 10. Yang, X.-F. *et al.* Single-Atom Catalysts: A New Frontier in Heterogeneous Catalysis. *Acc. Chem.*
819 *Res.* **46**, 1740–1748 (2013).
- 820 11. Li, X. *et al.* Tuning the Coordination Environment to Effect the Electrocatalytic Behavior of a
821 Single-Atom Catalyst toward the Nitrogen Reduction Reaction. *J. Phys. Chem. C* **125**, 11963–
822 11974 (2021).
- 823 12. Li, D., Li, H. & Lian, C. Dual-template designed porous cobalt nanoclusters achieve high-
824 efficiency alkaline oxygen evolution. *Chemical Engineering Journal* **522**, 168053 (2025).
- 825 13. Lu, Z. *et al.* Engineering CoN₄ and FeN₄ Dual Sites with Adjacent Nanoclusters on Flexible
826 Porous Carbon Fibers for Enhanced Electrocatalytic Oxygen Reduction and Evolution.
827 *Advanced Functional Materials* **35**, 2418489 (2025).
- 828 14. Seong, H. *et al.* Ni Foam-Supported Ni Nanoclusters for Enhanced Electrocatalytic Oxygen
829 Evolution Reaction. *J. Electrochem. Sci. Technol* **14**, 243–251 (2023).
- 830 15. Hong, W. T., Welsch, R. E. & Shao-Horn, Y. Descriptors of Oxygen-Evolution Activity for Oxides:
831 A Statistical Evaluation. *J. Phys. Chem. C* **120**, 78–86 (2016).

16. Zhou, J. *et al.* Discovery of Quantitative Electronic Structure-OER Activity Relationship in Metal-Organic Framework Electrocatalysts Using an Integrated Theoretical-Experimental Approach. *Adv Funct Materials* **31**, 2102066 (2021).
17. Cheng, X. *et al.* Oxygen Evolution Reaction on Perovskites: A Multieffect Descriptor Study Combining Experimental and Theoretical Methods. *ACS Catal.* **8**, 9567–9578 (2018).
18. Liu, X. *et al.* Building Catalyst Exploration Highways by Integrating High-Throughput and Machine Learning Technologies. *Advanced Energy Materials* **16**, e05497 (2026).
19. Singh, A. R., Rohr, B. A., Gauthier, J. A. & Nørskov, J. K. Predicting Chemical Reaction Barriers with a Machine Learning Model. *Catal Lett* **149**, 2347–2354 (2019).
20. Xing, L. *et al.* Combining machine learning with multi-physics modelling for multi-objective optimisation and techno-economic analysis of electrochemical CO₂ reduction process. *Carbon Capture Science & Technology* **9**, 100138 (2023).
21. Siddharthan, E. E., Ghosh, S. & Thapa, R. Site specific descriptors for oxygen evolution reaction activity on single atom catalysts using QMML. *J. Mater. Chem. A* **12**, 19176–19186 (2024).
22. Kapse, S., Janwari, S., Waghmare, U. V. & Thapa, R. Energy parameter and electronic descriptor for carbon based catalyst predicted using QM/ML. *Applied Catalysis B: Environmental* **286**, 119866 (2021).
23. Elçiçek, H. & Özdemir, O. K. Prediction of electrocatalyst performance of Pt/C using response surface optimization algorithm-based machine learning approaches. *Intl J of Energy Research* **46**, 21353–21372 (2022).
24. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
25. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science* **6**, 15–50 (1996).

- 857 26. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave
858 method. *Phys. Rev. B* **59**, 1758–1775 (1999).
- 859 27. Tan, X. *et al.* Zn-induced electron-rich Sn catalysts enable highly efficient CO₂
860 electroreduction to formate. <https://doi.org/10.1039/D3SC02790B> (2023)
861 doi:10.1039/D3SC02790B.
- 862 28. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
- 863 29. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple.
864 *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- 865 30. Wu, X. *et al.* Bifunctional Catalysts for Reversible Oxygen Evolution Reaction and Oxygen
866 Reduction Reaction. *Chemistry A European J* **26**, 3906–3929 (2020).
- 867 31. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio
868 parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu.
869 *J. Chem. Phys.* **132**, 154104 (2010).
- 870 32. Tripkovic, V., Hansen, H. A., Garcia-Lastra, J. M. & Vegge, T. Comparative DFT+U and HSE Study
871 of the Oxygen Evolution Electrocatalysis on Perovskite Oxides. *J. Phys. Chem. C* **122**, 1135–
872 1147 (2018).
- 873 33. Dudarev, S. L., Botton, G. A., Savrasov, S. Y., Humphreys, C. J. & Sutton, A. P. Electron-energy-
874 loss spectra and the structural stability of nickel oxide: An LSDA+U study. *Phys. Rev. B* **57**,
875 1505–1509 (1998).
- 876 34. Guba, M. & Höltzl, T. Stability and Electronic Structure of Nitrogen-Doped Graphene-
877 Supported Cu_n (n = 1–5) Clusters in Vacuum and under Electrochemical Conditions: Toward
878 Sensor and Catalyst Design. *The Journal of Physical Chemistry C*
879 <https://doi.org/10.1021/acs.jpcc.3c06475> (2024) doi:10.1021/acs.jpcc.3c06475.
- 880 35. Wang, H., Maiyalagan, T. & Wang, X. Review on Recent Progress in Nitrogen-Doped Graphene:
881 Synthesis, Characterization, and Its Potential Applications. *ACS Catal.* **2**, 781–794 (2012).

36. Liu, L. *et al.* Optimizing the electron donation and back-donation effect through the combination of d-block transition metal and s-block calcium atoms for efficient nitrogen fixation. *J. Mater. Chem. A* **13**, 20044–20053 (2025).
37. Zhang, Y., Ma, N., Wang, Y., Liang, B. & Fan, J. Theoretical design toward highly efficient single-atom catalysts for nitrogen reduction by regulating the “acceptance-donation” mechanism. *Applied Surface Science* **623**, 156827 (2023).
38. He, C., Yan, S. & Zhang, W. “Capture-activation-recapture” mechanism-guided design of double-atom catalysts for electrocatalytic nitrogen reduction. *Journal of Energy Chemistry* **96**, 437–447 (2024).
39. Humke, J. N. *et al.* Nickel binding enables isolation and reactivity of previously inaccessible 7-aza-2,3-indolynes. *Science* **384**, 408–414 (2024).
40. Lightcap, I. V., Kosel, T. H. & Kamat, P. V. Anchoring Semiconductor and Metal Nanoparticles on a Two-Dimensional Catalyst Mat. Storing and Shuttling Electrons with Reduced Graphene Oxide. *Nano Lett.* **10**, 577–583 (2010).
41. Yang, H. B. *et al.* Identification of catalytic sites for oxygen reduction and oxygen evolution in N-doped graphene materials: Development of highly efficient metal-free bifunctional electrocatalyst. *Sci. Adv.* **2**, e1501122 (2016).
42. Li, Q., Zhang, M., Wang, R., Pan, J. & Fu, H. TM-doping modulated p–d orbital coupling to enhance the oxygen evolution performance of Ni 3 S 2. *Nanoscale Advances* **6**, 5158–5163 (2024).
43. De Lucas-Consuegra, A. *et al.* Enhancing the catalytic activity and selectivity of the partial oxidation of methanol by electrochemical promotion. *Journal of Catalysis* **293**, 149–157 (2012).
44. Li, X. *et al.* Mechanistic study of CO₂ to C₁ reduction on single atom Pt synergized with transition metals: DFT study. *Applied Surface Science* **723**, 165669 (2026).

45. Yang, X. *et al.* Cohesive energy discrepancy drives the fabrication of multimetallic atomically dispersed materials for hydrogen evolution reaction. *Nat Commun* **15**, 8216 (2024).
46. Prats, H. & Stamatakis, M. Stability and reactivity of metal nanoclusters supported on transition metal carbides. *Nanoscale Adv.* **5**, 3214–3224 (2023).
47. Ma, J. *et al.* Electronic interaction between platinum nanoparticles and nitrogen-doped reduced graphene oxide: effect on the oxygen reduction reaction.
<https://doi.org/10.1039/C5TA01285F> (2015) doi:10.1039/C5TA01285F.
48. Li, H., Zhao, X., Wei, D., Li, L. & Li, S. Unexpected Odd–Even Oscillation in the Enhanced Chemical Activities of the Run ($n = 2–14$) Nanoclusters for H₂O Splitting. *J. Phys. Chem. C* **121**, 7188–7198 (2017).
49. Balandin, A. A. Modern State of the Multiplet Theor of Heterogeneous Catalysis. in *Advances in Catalysis* vol. 19 1–210 (Elsevier, 1969).
50. Mackay, A. L. A dense non-crystallographic packing of equal spheres. *Acta Crystallogr*
<https://doi.org/10.1107/S0365110X6200239X> (1962) doi:10.1107/S0365110X6200239X.
51. Nijamudheen, A. & Datta, A. Odd–even oscillations in structural and optical properties of gold clusters. *Journal of Molecular Structure: THEOCHEM* **945**, 93–96 (2010).
52. Ren, X. *et al.* Spin-polarized oxygen evolution reaction under magnetic field. *Nat Commun* **12**, 2608 (2021).
53. Gracia, J. *et al.* Analysis of the Magnetic Entropy in Oxygen Reduction Reactions Catalysed by Manganite Perovskites. *ChemCatChem* **9**, 3358–3363 (2017).
54. Miranda-Quintana, R. A., Adebar, N., Schulze, M. & Smiatek, J. Sabatier Principle Revisited: The Role of Electronic Properties in Simple Catalytic Reactions. *J. Phys. Chem. C* **129**, 9926–9934 (2025).
55. Tian, H. *et al.* p–d orbital hybridization induced by transition metal atom sites for room-temperature sodium–sulfur batteries. *National Science Review* **12**, nwaf241 (2025).

56. Maintz, S., Deringer, V. L., Tchougréeff, A. L. & Dronskowski, R. LOBSTER: A tool to extract chemical bonding from plane-wave based DFT. <https://doi.org/10.1002/jcc.24300> doi:10.1002/jcc.24300.
57. Luo, X., Liang, G., Li, Y., Yu, F. & Zhao, X. Regulating the Electronic Structure of Freestanding Graphene on SiC by Ge/Sn Intercalation: A Theoretical Study. *Molecules* **27**, 9004 (2022).
58. Wang, J. *et al.* Computational design of bimetallic TM₂@g-C₉N₄ electrocatalysts for enhanced CO reduction toward C₂ products. *Chinese Chemical Letters* **35**, 109050 (2024).
59. Lin, L. *et al.* Reversing sintering effect of Ni particles on γ-Mo₂N via strong metal support interaction. *Nat Commun* **12**, 6978 (2021).
60. Samantara, A. K. & Ratha, S. Mechanism and Key Parameters for Catalyst Evaluation. in *Metal Oxides/Chalcogenides and Composites: Emerging Materials for Electrochemical Water Splitting* (eds Samantara, A. K. & Ratha, S.) 11–29 (Springer International Publishing, Cham, 2019). doi:10.1007/978-3-030-24861-1_3.
61. Zhang, K. & Zou, R. Advanced Transition Metal-Based OER Electrocatalysts: Current Status, Opportunities, and Challenges. *Small* **17**, 2100129 (2021).
62. Das, S. *et al.* Harnessing the Trade-Off between CoFe/Fe₃C Interfacial Junction with Unparalleled Potential Gap of 0.58 V for Reversible Oxygen Electrocatalysis: Application toward Liquid and Solid-State Zn-Air Batteries. *Advanced Functional Materials* **34**, 2407078 (2024).
63. Talib, S. H. *et al.* Theoretical Inspection of M₁/PMA Single-Atom Electrocatalyst: Ultra-High Performance for Water Splitting (HER/OER) and Oxygen Reduction Reactions (OER). *ACS Catal.* **11**, 8929–8941 (2021).
64. Chen, X. *et al.* Achieving High-Performance Electrocatalytic Water Oxidation on Ni(OH)₂ with Optimized Intermediate Binding Energy Enabled by S-Doping and CeO₂-Interfacing. *Small* **20**, 2303169 (2024).

65. Zhang, Y. *et al.* Establishing theoretical landscapes for identifying basal plane active sites in MBene toward multifunctional HER, OER, and ORR catalysts. *Journal of Colloid and Interface Science* **652**, 1954–1964 (2023).
66. Chen, Z. W. *et al.* Unusual Sabatier principle on high entropy alloy catalysts for hydrogen evolution reactions. *Nat Commun* **15**, 359 (2024).
67. Huang, B. *et al.* Experimentally validating sabatier plot by molecular level microenvironment customization for oxygen electroreduction. *Nat Commun* **15**, 6077 (2024).
68. Cheng, J. & Hu, P. Utilization of the Three-Dimensional Volcano Surface To Understand the Chemistry of Multiphase Systems in Heterogeneous Catalysis. *J. Am. Chem. Soc.* **130**, 10868–10869 (2008).
69. Man, I. C. *et al.* Universality in Oxygen Evolution Electrocatalysis on Oxide Surfaces. <https://doi.org/10.1002/cctc.201000397> doi:10.1002/cctc.201000397.
70. Choudhary, K. Predicting anomalous quantum confinement effect in van der Waals materials. *Phys. Rev. Mater.* **5**, (2021).
71. Sanchez, J. *et al.* Isolating the Electrocatalytic Activity of a Confined NiFe Motif within Zirconium Phosphate. *Advanced Energy Materials* **11**, 2003545 (2021).
72. Wang, M. & Navrotsky, A. Enthalpy of formation of LiNiO₂, LiCoO₂ and their solid solution, LiNi_{1-x}Co_xO₂. *Solid State Ionics* **166**, 167–173 (2004).
73. Tkalych, A. J., Zhuang, H. L. & Carter, E. A. A Density Functional + *U* Assessment of Oxygen Evolution Reaction Mechanisms on β-NiOOH. *ACS Catal.* **7**, 5329–5339 (2017).
74. Jena, R. *et al.* In Situ Tracking of Ni-MOF Reconstruction into Active Ni(OH)₂ OER Catalysts. *Angewandte Chemie International Edition* **64**, e202510741 (2025).
75. Das, S. *et al.* Attaining the Sabatier Sweet Spot via d-Orbital Occupancy Engineering for High-Performance Zn–Air Batteries. *Advanced Energy Materials* e06489 (2026) doi:10.1002/aenm.202506489.

76. Ringe, S. The importance of a charge transfer descriptor for screening potential CO₂ reduction electrocatalysts. *Nat Commun* **14**, 2598 (2023).
77. Jin, H. *et al.* Sabatier Phenomenon in Hydrogenation Reactions Induced by Single-Atom Density. *J. Am. Chem. Soc.* **145**, 12023–12032 (2023).
78. Bhattacharjee, S., Waghmare, U. V. & Lee, S.-C. An improved d-band model of the catalytic activity of magnetic transition metal surfaces. *Sci Rep* **6**, 35916 (2016).
79. Liu, M. *et al.* In situ modulating coordination fields of single-atom cobalt catalyst for enhanced oxygen reduction reaction. *Nat Commun* **15**, 1675 (2024).
80. Dronskowski, R. & Bloechl, P. E. Crystal orbital Hamilton populations (COHP): energy-resolved visualization of chemical bonding in solids based on density-functional calculations. *J. Phys. Chem.* **97**, 8617–8624 (1993).
81. Fu, Z., Yang, B. & Wu, R. Understanding the Activity of Single-Atom Catalysis from Frontier Orbitals. *Phys. Rev. Lett.* **125**, 156001 (2020).
82. Ren, J., Antonietti, M. & Feller, T. Efficient Water Splitting Using a Simple Ni/N/C Paper Electrocatalyst. *Advanced Energy Materials* **5**, 1401660 (2015).
83. Wang, X. *et al.* Identifying the Key Role of Pyridinic-N–Co Bonding in Synergistic Electrocatalysis for Reversible ORR/OER. *Advanced Materials* **30**, 1800005 (2018).
84. Han, X. *et al.* Identifying the Activation of Bimetallic Sites in NiCo₂S₄@g-C₃N₄-CNT Hybrid Electrocatalysts for Synergistic Oxygen Reduction and Evolution. *Advanced Materials* **31**, 1808281 (2019).
85. Kundu, A., Kuila, T., Chandra Murmu, N., Samanta, P. & Das, S. Metal–organic framework-derived advanced oxygen electrocatalysts as air-cathodes for Zn–air batteries: recent trends and future perspectives. *Materials Horizons* **10**, 745–787 (2023).
86. Qiao, Y. *et al.* Sulfuration of an Fe–N–C Catalyst Containing Fe_xC/Fe Species to Enhance the Catalysis of Oxygen Reduction in Acidic Media and for Use in Flexible Zn–Air Batteries. *Advanced Materials* **30**, 1804504 (2018).

87. Jiang, M. *et al.* Integrated and Binder-Free Air Cathodes of Co₃Fe₇ Nanoalloy and Co_{5.47}N Encapsulated in Nitrogen-Doped Carbon Foam with Superior Oxygen Reduction Activity in Flexible Aluminum-Air Batteries. *Advanced Science* **7**, 2000747 (2020).
88. Ye, D. *et al.* Dual-sources directed construction of N-doped carbon nanotube arrays as superior Self-supported bifunctional air electrodes for Rechargeable/Flexible Zinc-air batteries. *Chemical Engineering Journal* **464**, 142601 (2023).
89. Jeon, S. S. *et al.* Active Surface Area and Intrinsic Catalytic Oxygen Evolution Reactivity of NiFe LDH at Reactive Electrode Potentials Using Capacitances. *ACS Catal.* **13**, 1186–1196 (2023).
90. Liu, Q. *et al.* Electronic Communication Between Single Atomic Nickel and Iron-Nitrogen Species Promote the Bifunctional Oxygen Evolution and Reduction for Efficient Rechargeable Zinc-Air Battery. *Adv Funct Materials* **35**, 2424597 (2025).
91. Das, S. *et al.* Pyridinic-N Seized Co in Biphasic Nanoarchitecture for Reversible Oxygen Electrocatalysis Enabling Longevous (>1200 h) Aqueous and Dual-Anion Kosmotropic Electrolyte Stabilized High Power Quasisolid-State Zn–Air Battery. *Small Methods* **9**, 2401874 (2025).
92. Tong, Y. *et al.* A Bifunctional Hybrid Electrocatalyst for Oxygen Reduction and Evolution: Cobalt Oxide Nanoparticles Strongly Coupled to B,N-Decorated Graphene. *Angew Chem Int Ed* **56**, 7121–7125 (2017).
93. Hammer, B. & Nørskov, J. K. Electronic factors determining the reactivity of metal surfaces. *Surface Science* **343**, 211–220 (1995).
94. Nørskov, J. K. *et al.* Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode. *J. Phys. Chem. B* **108**, 17886–17892 (2004).
95. Toyoda, K., Hinogami, R., Miyata, N. & Aizawa, M. Calculated Descriptors of Catalytic Activity for Water Electrolysis Anode: Application to Delafossite Oxides. *J. Phys. Chem. C* **119**, 6495–6501 (2015).

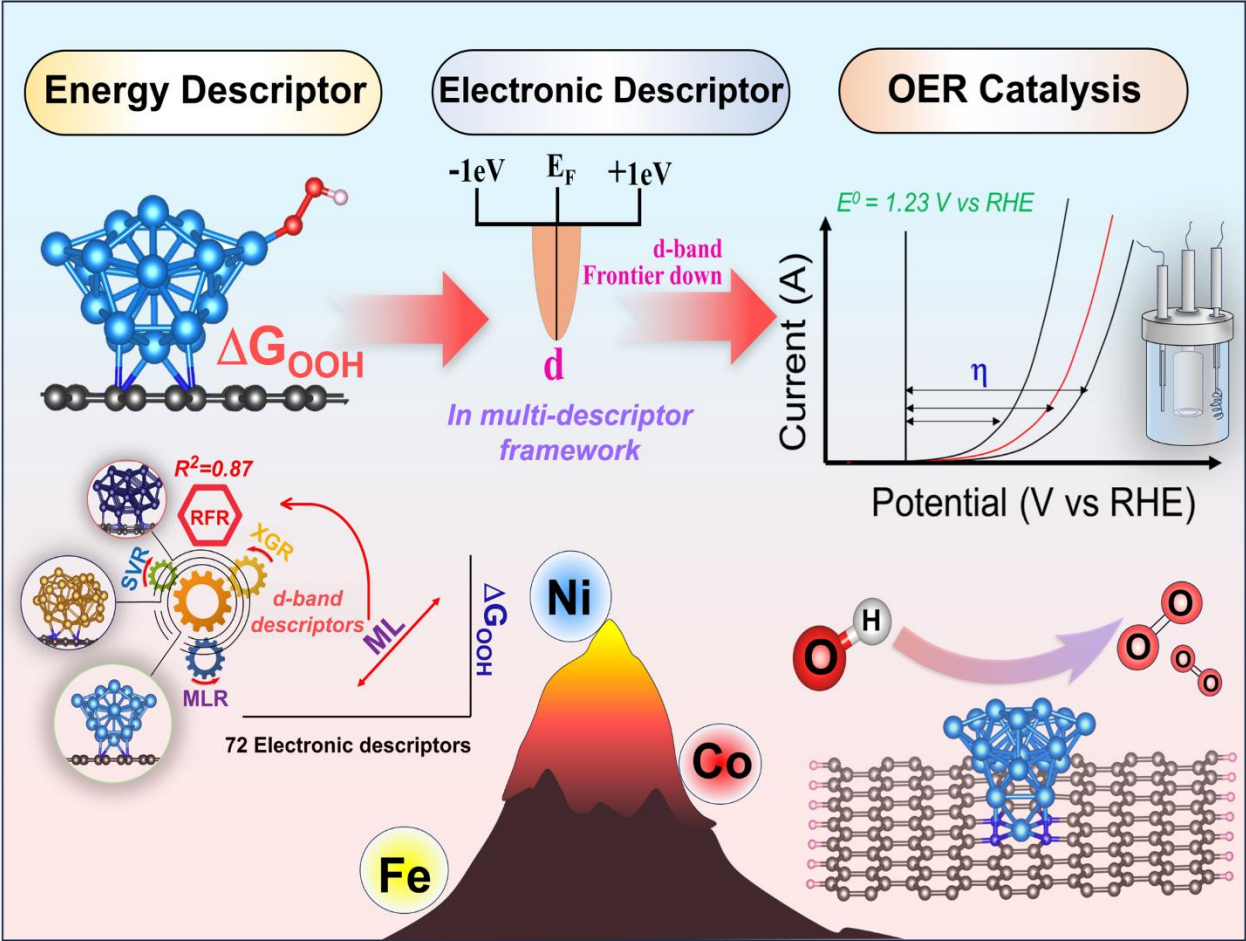
96. Wei, C. *et al.* Cations in Octahedral Sites: A Descriptor for Oxygen Electrocatalysis on Transition-Metal Spinel. *Advanced Materials* **29**, 1606800 (2017).
97. Wang, Y. *et al.* Computational design and descriptor development for metal single atoms anchored on CrP2 monolayer toward efficient electrocatalytic water splitting. *Journal of Physics and Chemistry of Solids* **208**, 113165 (2026).
98. Calle-Vallejo, F. *et al.* Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors. *Science* **350**, 185–189 (2015).
99. Grimaud, A. *et al.* Double perovskites as a family of highly active catalysts for oxygen evolution in alkaline solution. *Nat Commun* **4**, 2439 (2013).
100. Wu, D., Xi, C., Dong, C., Liu, H. & Du, X.-W. Bond-Energy-Integrated Coordination Number: An Accurate Descriptor for Transition-Metal Catalysts. *J. Phys. Chem. C* **123**, 28248–28254 (2019).
101. Kim, S., Kang, Y. & Ham, H. C. First-Principles Study of Pt-Based Bifunctional Oxygen Evolution & Reduction Electrocatalyst: Interplay of Strain and Ligand Effects. *Energies* **14**, 7814 (2021).
102. Suntivich, J., May, K. J., Gasteiger, H. A., Goodenough, J. B. & Shao-Horn, Y. A Perovskite Oxide Optimized for Oxygen Evolution Catalysis from Molecular Orbital Principles. *Science* **334**, 1383–1385 (2011).
103. Araujo, R. B. & Edvinsson, T. N₂ adsorption on high-entropy alloy surfaces: unveiling the role of local environments. *J. Mater. Chem. A* **11**, 12973–12983 (2023).
104. Cao, G., Yang, S., Ren, J.-C. & Liu, W. Electronic descriptors for designing high-entropy alloy electrocatalysts by leveraging local chemical environments. *Nat Commun* **16**, 1251 (2025).
105. Dutta, S. & Banerjee, R. Electronic structure engineering of Zr-doped Ti₃C₂ and Ti₃CN MXenes for efficient hydrogen evolution reaction. *Physica B: Condensed Matter* **707**, 417148 (2025).
106. Sk, M., Pathak, A. & Thapa, R. Integrating Machine Learning and DFT Descriptors for Screening Dual Metal-Site Catalysts for CO₂ Reduction to C₂ Products. *Small* **22**, e13368 (2026).

1059 **Table 1:** Electronic Descriptors reported for Metal-Based electrocatalysts using DFT and ML
 1060 approaches.

Electronic Descriptor(s)	Material Class/ Catalyst	Reaction(s)	Efficient Catalyst Identified
d-band centre (ϵ_d) ^{93,94}	Transition metals and alloys	ORR/HER	Pt(111) & Pt-alloy
t_{2g} occupancy ⁹⁵	CuBO ₂ (B = transition metal)	OER	CuCoO ₂ , PtCoO ₂
charge-transfer energy and the optimal occupancy of the e_g orbital ¹⁵	ABO ₃ perovskites	OER	LaCoO ₃
e_g occupancy in the octahedral site ⁹⁶	Transition-metal spinel oxides	OER/ORR	MnCo ₂ O ₄
d-band centre(ϵ_d) and e_g filling ¹⁶	Ni-M-MOFs (M = Fe, Co, Cu, Mn, and Zn)	OER	NiFe-MOF
electronic conductivity, oxygen vacancy concentration, flat-band potential (E_{fb}) ¹⁷	La _{1-x} Sr _x CoO _{3-δ} perovskites	OER	La _{0.8} Sr _{0.2} CoO _{3-δ}
Composite descriptors (χ for HER and ν for OER) ⁹⁷	Transition metal SACs on CrP ₂ monolayer	OER/HER	Fe@CrP ₂
Generalised coordination number ⁹⁸	Transition metals	ORR	Pt(111)
O p-band centre ⁹⁹	(Ln _{0.5} Ba _{0.5})CoO _{3-δ} Perovskite	OER	(Pr _{0.5} Ba _{0.5})CoO ₃
Orbital-wise coordination number (CN α) ¹⁰⁰	MnO ₂ (β -phase) & RuO ₂	CO ₂ RR	β -MnO ₂
Surface segregation energy ¹⁰¹	Pt-based near- surface alloys	OER/ORR	Pt ₃ M/Pt
e_g orbital occupancy ¹⁰²	BSCF Perovskite	OER/ORR	Ba _{0.5} Sr _{0.5} Co _{0.8} Fe _{0.2} O _{3-δ}
Width corrected band center(D_{wd}) + d- band frontier (d_f) ²¹	Transition metal(Co, Fe, Ni) SACs	OER	6aCoN ₄ S
Bader charge (q_B) + -pCOHP ¹⁰³	bcc/fcc High Entropy Alloys	NRR	PtNiCoFeCu
d-band filling + local electronegativity ($\epsilon_d + \Sigma\chi_{neighbor}$) ¹⁰⁴	PtRuPdCoNi High Entropy Alloys	ORR	Pd-Ag, Ir-Pt, Ir-Au compositions

Work function (Φ) and Bader charge (q_B) ¹⁰⁵	Ti ₃ C ₂ & Ti ₃ CN MXenes	HER	TZ ₇ CN
dz ² down occupancy ¹⁰⁶	Transition metal dual-atom catalysts	CO ₂ RR	MnCo@NC

1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105



Supplementary Files

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

- [SupportingInformation.docx](#)