International Journal of Quantum Chemistry

WILEY

International Journal of QUANTUM CHEMISTRY

RESEARCH ARTICLE

Interpretable Machine Learning for Benchmarking DFT Electron Affinity Predictions: A Generalized Additive Model Approach

Ismail Badran[1] | Abdelrahman Eid[2,3] | Motasem Far[1] | Nadeen Abbas[1] | Raghad Tayeh[1] | Sahar Salman[1] | Yasmeen Hamdan[1]

[1]Department of Chemistry, Faculty of Sciences, An-Najah National University, Nablus, Palestine | [2]Department of Mathematics, Faculty of Sciences, An-Najah National University, Nablus, Palestine | [3]Data Science Unit, An-Najah Innovation Park, An-Najah National University, Nablus, Palestine

Correspondence: Ismail Badran (i.badran@najah.edu) | Abdelrahman Eid (abed.eid@najah.edu)

Received: 17 February 2026 | **Revised:** 19 March 2026 | **Accepted:** 31 March 2026

Keywords: DFT | electron affinity | explainable AI (XAI) | feature influence | generalized additive model (GAM) | interpretable machine learning | SDG 12 | uncertainty quantification (UQ)

ABSTRACT

Electron affinity (EA) is vital for understanding charge transfer, redox reactions, and material design across chemistry and materials science. This study introduces a novel methodology that integrates interpretable machine learning with density functional theory (DFT) to provide a guideline for EA predictions. Using Generalized Additive Models (GAM), we reveal how functionals, basis sets, and molecular species interact in complex and nonlinear ways, offering insights that are not accessible through conventional analyses. A curated dataset of 47 atoms and molecules was used to assess the predictive performance of nine functionals across five basis sets. The results show that basis set selection has a more pronounced effect on EA prediction accuracy than functional choice, with the "*augmented*" ma-def2-SVP basis set, particularly in combination with B3LYP, providing the most reliable predictions. In contrast, the "*unaugmented*" def2-TZVP and def2-SVP basis sets frequently introduced large errors and high variability. The GAM framework identified species-specific challenges, especially for Cl_2, NH_3, and CN, and demonstrated that although statistically significant differences exist among functionals, their performance in predicting EA values is largely equivalent from a chemical perspective. The GAM-based ranking showed a spread of 1.6 kJ/mol among the average predicted deviations of the tested functionals, indicating that their practical differences are small from a chemical perspective under the studied conditions. This interpretable machine learning approach therefore provides a transparent and reproducible strategy for guiding the selection of DFT methods for EA calculations. To our knowledge, this is the first study to apply interpretable ML to systematically investigate and challenge generalized assumptions in EA benchmarking, establishing a data-driven framework for practical decision-making in computational chemistry.

1 | Introduction

Electron affinity (EA) is a key thermodynamic property that influences how atoms and molecules interact with thermal electrons in the gas, liquid, or solid phases. EA, in its adiabatic form, is defined as the energy change accompanying the addition of an electron to a neutral species ($M + e^- \rightarrow M^-$) [1, 2]. In addition to its fundamental significance in chemistry and physics, precise measurement of electron affinity (EA) values is crucial for understanding charge transfer reactions in voltaic cells, semiconductor fabrication processes [3–6], as well as DNA/RNA interactions, both within their structures and with other molecules [1, 7].

Electron affinity (EA) can be defined in two ways: adiabatic electron affinity (AEA) and vertical electron affinity, commonly known as vertical detachment energy (VDE). AEA is calculated from the total energy difference between fully optimized anion and neutral structures, while VDE corresponds to the energy difference with the neutral molecule in its frozen anionic geometry. AEA is generally more reliable and is the definition adopted in this study [7, 8].

It was reported that EA has been regarded "as one of the hardest atomic or molecular properties to reproduce using ab initio quantum mechanical calculations" [9–11]. With today's advances in computational chemistry, this raises the question whether modern ab initio methods can predict electron affinities within the limits of chemical accuracy. The coupled cluster theory with single, double, and perturbative triple excitations [CCSD(T)] is considered the "*gold standard*" for high-accuracy calculations. For practical applications, chemical accuracy is considered in the range of 1–3 kcal/mol (0.04–0.1 eV) [12–14]. However, CCSD(T) is time-consuming and highly sensitive to basis sets. Its accuracy comes from explicitly treating electron correlation through singles, doubles, and perturbative triples. Because correlation energy converges slowly with conventional Gaussian basis sets, achieving benchmark-quality electron affinities requires large, correlation-consistent basis sets or basis-set extrapolation [6, 15–17]. Density functional theory (DFT), on the other hand, has demonstrated strong performance in determining electronic properties such as bond energies, activation and reaction enthalpies, and ionization energies [12, 15, 18].

For the sake of EA estimations, two main approaches exist: the direct SCF method, where EA is obtained from the energy difference between the anion and the neutral species, and another approach based on molecular orbital energies [19–21]. The direct method is simple and has been benchmarked, but studies remain limited, especially those combining it with machine learning to analyze functional performance in EA calculations. For instance, Gu et al. benchmarked the EA of uracil and found HF, MP2, B3LYP, and M062X methods to be acceptable for electron-DNA/RNA interactions [1]. Another study showed that MP2 and CCSD(T), extrapolated to the CBS limit, performed well in the EA evaluation of methane. The authors also recommended the use of large and polarized basis sets like aug-cc-pVnZ [22]. Similarly, Richard et al. using a dataset of 24 small molecules, confirmed that both MP2 and CCSD(T)/CBS are reliable methods for calculating EA values [6]. Modern density functionals such as M062X, MN15-L, and ωB97X have demonstrated superior performance in calculating ionization potentials (IP), bond dissociation energies (BDE), and band gaps. However, they unfortunately lack consistency in predicting EA values [3, 23, 24]. Therefore, a conclusive assessment of DFT methods in obtaining reliable EA values is still missing.

At the edge of the fourth industrial revolution, machine learning (ML) is transforming quantum mechanical methods, not only by aiding data interpretation but also by enhancing predictive accuracy [25–29]. ML-driven approaches refine exchange-correlation functionals, accelerate quantum calculations, and improve data predictions [30, 31].

In this study, we analyzed 2115 computed AEA values generated using nine DFT functionals and five basis sets across 47 atoms and small molecules. While the number of unique compounds is limited, this selection captures a representative mix of atomic and molecular systems, which are modeled for ML as random effects to account for variability. The goal is not exhaustive coverage but to explore how DFT performance varies across different species types and modeling choices, making the dataset both sufficient and well-suited for interpretable machine learning analysis.

The DFT functionals were carefully chosen to represent all major types of density functional, covering local and gradient-corrected functionals, hybrid, meta, and range-separated GGA, and perturbatively corrected double-hybrid functionals. Further, to assess the impact of basis set size, the calculations were performed using five different basis sets.

This study employs an interpretable machine learning framework to investigate how different combinations of DFT functionals, basis sets, and species types affect the accuracy of electron affinity (EA) predictions. The analysis began with exploratory visualizations to examine overall patterns of absolute deviation from experimental values, providing initial insights into variability across computational settings and potential interactions.

A descriptive GAM was then used to model these relationships more formally. This flexible statistical tool enabled the detection of nonlinear effects and interactions, offering a deeper understanding of how functional–basis set–species combinations influence prediction errors. Species were modeled as a random effect to reflect inherent variability between chemical systems.

To translate these insights into practical recommendations, a predictive GAM model was constructed to estimate absolute deviations across various DFT configurations. These predictions formed the basis for a series of ranking analyses across functionals, basis sets, species, and their interactions, identifying optimal setups for EA calculations. Consequently, the ML steps in this study establish a reproducible, data-driven framework for guiding method selection in quantum chemistry, showing how interpretable machine learning can enhance benchmarking studies and promote informed decisions in electronic structure calculations.

2 | Calculations and Methods

2.1 | Computational Methods

In this work, the electron affinity (EA) is defined as the negative change of enthalpy at 0 K (H_0) for the reaction [1, 2]:

$$M + e^- \rightarrow M^- \tag{1}$$

where M and M^- represent a neutral species in its ground state and its singly charged anion, respectively. The enthalpy change (H_0) for a given specie, M, can be computed using ab initio methods as:

$$H_0 = \text{SPE} + \text{ZPE} \tag{2}$$

TABLE 1 | Description of the computational methods used in this work.

Functional	Description	Classification	References
PBE	Perdew-Burke-Erzerhoff GGA functional	Local and gradient corrected functionals	[35, 36]
PBE0	1-parameter version of PBE (25% HF exchange)	Hybrid GGA with no dispersion	[35, 36]
B3LYP	Becke's functional with the LYP expression	Hybrid GGA with no dispersion	[37]
M06L	Truhlar Minnesota functional	Hybrid meta GGA	[38, 39]
M062X	Truhlar Minnesota functional	Hybrid meta GGA	[38, 39]
ωB97X-D3	Head-Gordon's DF ωB97X with minimal Fock exchange with Grimme's dispersion	Range-separated hybrid functionals	[40, 41]
ωB97X-D4	Grimme's functional with DFT-D4 correction by Najibi and Goerigk	Range-separated hybrid functionals	[40, 41]
B2PLYP	Grimme's mixture of B88, LYP, and MP	Perturbatively corrected double-hybrid functionals	[42]
ωB2PLYP	Goerigk and Casanova-Páez's range-separated DHDF, with the correlation contributions based on B2PLYP, optimized for excitation energies	Range-separated double-hybrid functionals	[43]

where SPE is the single point energy at a given level of theory and ZPE is the zero-point energy at the same level. No scaling was done for ZPE in this project.

The approach described above represents the direct method, as stated in the introduction. This method offers an accurate and widely accepted means of calculating electron affinity (EA) values. It is worth noting that EA can also be estimated using molecular orbital (MO) energies through either Koopmans's [32] or Janak's theorems [33, 34], which provide approximate orbital-based estimates under restrictive conditions. However, the advantage of using the direct method is that it reflects the true thermodynamic definition of EA and is consistent with both experimental practice and theoretical principles [6, 7, 20].

The EA values for a dataset composed of 47 species were calculated using DFT methods. The dataset was constructed to cover different types of species, including main group elements (e.g., B, C, P), transition metals (e.g., Sc, Ti, Ni), diatomic molecules (e.g., F_2, Cl_2, NO), and polyatomic molecules (e.g., ClO, NH_3, SiH_2). The selection of these 47 species was based on reported EA values. A complete list of the species, along with their reported EA values, is provided in Table S1.

The calculations were performed using different types of density functionals, chosen to represent a range of computational complexities as mentioned earlier. A full list of the functionals, along with their descriptions and references, is shown in Table 1. This approach allows for the study of the effect of functional complexity on the accuracy of EA calculations.

In addition to examining the effect of functionals, calculations were also performed using five different basis sets: 6–31 + G(d,p), def2-SVP, def2-TZVP, ma-def2-SVP, and ma-def2-TZVP. The choice of these basis sets was based on their recent performance and recommendations [12, 15, 44]. All calculations were done using ORCA 6.01 software [44–49] and viewed in Avogadro [50].

A small number of benchmark species in this study were reported to have negative experimental electron affinity values, specifically NH_3 and CH_3OH. Negative EA values for such closed-shell species result from the excess electron leading to a temporary or metastable anion, and a standard bound-state ΔSCF treatment with finite Gaussian basis sets is not intended to provide a rigorous resonance description [11]. Therefore, these cases are included only as part of the overall comparative benchmark and are not used to support the physical interpretation of metastable-anion behavior. To examine whether they influenced the principal comparative findings of the study, an additional post hoc sensitivity analysis was performed in which NH_3 and CH_3OH were excluded from the screened benchmark summaries defined in Section 2.3 and summarized in the Supporting Information section.

2.2 | Metrics and Data Processing Definitions

To quantify and assess the quality of computed electron affinity (EA) values, several standard error measures were employed during data preprocessing and analysis:

2.2.1 | Absolute Error (AE)

The absolute error for a given observation was defined as the normalized difference between the calculated EA value and its corresponding experimental (reported) value: [51].

$$AE_i = \left| \frac{Calculated_i - Reported_i}{Reported_i} \right| \tag{3}$$

2.2.2 | Absolute Deviation (AD)

For comparative purposes across models, the absolute deviation is calculated as the simple absolute difference [51].

$$AD_i = \left| Calculated_i - Reported_i \right| \tag{4}$$

2.2.3 | Mean Absolute Deviation (MAD)

The mean absolute deviation over a group of observations (e.g., all results for a given functional or basis set) is computed as: [51].

$$MAD = \frac{1}{n}\sum_{i=1}^{n}\left|Calculated_i - Reported_i\right| \quad (5)$$

This metric is used to assess the average deviation from the reported value.

2.2.4 | Standard Deviation (SD) of Absolute Deviations

The variability of the absolute deviation values was measured using: [51].

$$SD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(AD_i - MAD\right)^2} \quad (6)$$

This measure describes the consistency or spread of prediction accuracy across molecular species or method combinations.

2.2.5 | Coefficient of Determination (R^2)

When assessing the overall fit between calculated and reported EA values, the coefficient of determination is used: [51].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Reported_i - Calculated_i\right)^2}{\sum_{i=1}^{n}\left(Reported_i - Reported\right)^2} \quad (7)$$

An R^2 value closer to 1 indicates better agreement between predicted and experimental values.

2.3 | Statistical Methods

To ensure the dataset was suitable for analysis and robust against noise such as missing data or extreme values, a comprehensive data cleaning process was performed. Missing values (due to basis sets not defined to given atoms) were removed to guarantee complete cases for statistical and machine learning models, ensuring reliability in subsequent analysis. Absolute errors were calculated as the normalized difference between the obtained (*calculated*) and experimental values (referred herein as *reported* values), and any value with absolute error equal to or exceeding 0.25 was excluded to minimize the influence of large deviations that could skew the results. This approach was applied with full awareness that normalized errors may appear large for compounds with very small EA values. However, such exclusions were necessary to preserve the stability and interpretability of the statistical models, particularly the smooth trends which could be captured in the used ML analysis.

The dataset was then used to compare the different methods and their effects on the reported values. Outliers were identified and removed using z-scores, calculated based on the mean and standard deviation of the calculated values within each species. This approach allowed for the retention of values within a statistically relevant range, enhancing the consistency of the dataset. Additionally, data from specific species, including ClO_3, F_2, BN, BH_2, O_3, PH_2, F, and Ti, were *excluded* due to their high errors. Furthermore, observations related to the "def2-SVP" method were *removed* to maintain consistency across methods under investigation.

Additional metrics, such as the mean and standard deviation of the calculated values, absolute deviations, were computed to support advanced analysis and visualization.

2.4 | Statistical Software and R Packages

The analysis was conducted using R programming language [52], which provides a robust environment for data manipulation, visualization, and statistical modeling. For data preprocessing, the packages dplyr [53], tidyr [54], and readxl [55] were used to clean, reshape, and import data efficiently. For statistical modeling, mgcv [56] was used to fit Generalized Additive Model (GAM) [57], and to capture both fixed and random effects in the dataset. For visualization, ggplot2 [58] was used to generate advanced plots, and reshape2 [59] was applied to transform data formats when needed.

3 | Results and Discussion

3.1 | Exploratory Analysis

The electron affinities (EA) of selected atoms and molecules investigated in this study were calculated as described in the computational methods section. The study includes 47 species—15 atoms and 32 molecules. EA values were determined using nine different DFT functionals, and each calculation was repeated with five distinct basis sets, resulting in a total of 2115 data points. Table 2 presents the raw, unprocessed EA values, both calculated and reported, obtained with the M062X functional across five basis sets: 6–31 + G(d,p), def2-SVP, def2-TZVP, ma-def2-SVP, and ma-def2-TZVP. These values are shown as originally computed, prior to any data cleaning or preprocessing steps. For clarity, only the M062X results are included in the main text; the full dataset is provided in Table S1 and the additional sensitivity analysis excluding NH_3 and CH_3OH is reported in the Supporting Information section.

As shown in the table, some of the calculated EA values closely match the reported data, while others show significant discrepancies. A preliminary analysis suggests that the source of these errors could be related to the species themselves, the choice of functional, the basis set, or a combination of these factors.

To better understand these differences, we performed a comprehensive statistical analysis aimed at achieving three objectives: (1) to examine how the level of theory, specifically the choice of functional and basis set, influences the accuracy of EA predictions; (2) to explore the relationships among the key variables, functional, basis set, and species type, and determine whether these relationships are linear, nonlinear, or uncorrelated; and (3) to evaluate the relative importance of the functional and basis set in determining model performance, and ultimately assess how

TABLE 2 | Raw, unprocessed electron affinity of atoms and molecules calculated using the M062X functional at five different basis sets.

Species	Exp. (kJ/mol)	6−31 + G(d,p)	def2-SVP	def2-TZVP	ma-def2-SVP	ma-def2-TZVP
Cr	64.3	117.7	97.8	98.4	111.9	119.1
Ti	7.6	1.5	8.4	2.4	6.8	0.4
BH_2	30.5	5.1	−37.6	−4.5	7.9	10.2
BN	304.5	380.2	352.5	379.0	391.7	384.8
CH_2	47.3	−221.8	−344.6	−281.5	−216.0	−171.8
CF_3	175.6	178.3	94.3	150.9	170.2	165.3
ClO_3	410.1	302.9	105.8	389.3	586.1	406.5
CO	63.9	−157.1	−277.7	−214.6	−161.8	−137.9
CS_2	56.0	117.6	71.9	106.1	127.0	118.8
HS	223.3	224.1	131.7	196.2	223.4	221.9
IO	229.6		171.5	227.1	246.6	244.7
SCN	340.8	347.9	303.3	337.8	352.0	346.8
SeH	213.3		157.0	197.2	215.9	214.3
SiH	123.2	35.5	−33.3	9.7	38.5	39.9
SiH_2	108.4	99.3	67.6	86.7	104.8	100.5
SO_2	106.8	149.2	81.5	103.7	160.3	125.4
B	27.0	17.6	−30.7	6.2	25.4	23.7
C	121.8	115.1	47.5	103.4	122.1	121.5
Cl	348.6	354.3	241.6	324.2	356.8	351.3
F	328.2	307.8	129.7	271.2	306.2	315.6
Fe	14.6	104.6	78.2	−31.9	−107.4	−3.3
H	72.8	−124.7	−56.6	−1.0	−56.6	−1.0
K	48.4	54.5	52.5	48.2	55.4	54.2
Na	52.9	56.7	37.7	47.3	55.6	58.5
Ni	111.5		−31.3	64.2	−18.5	104.3
O	141.0	126.1	−30.8	92.3	125.6	135.7
P	72.0	79.1	−26.7	48.8	77.7	82.3
S	200.4	204.4	99.3	176.4	205.1	204.3
Sc	18.1	−61.7	−89.5	−47.4	−79.6	−36.9
BF_3	256.0	−143.0	−420.6	−206.9	−136.7	−107.2
BO	242.1	233.5	193.7	226.7	234.0	233.9
CH_3	108.8	−5.9	−75.5	−26.1	−6.5	−0.4
CH_3O	151.4	136.6	80.4	125.4	140.8	143.4
CH_3OH	−155.0	−152.6	−349.9	−259.2	−148.3	−97.8
Cl_2	227.0	280.7	203.6	229.3	290.7	243.1
ClO	219.7	220.0	133.1	197.4	228.4	220.0
CN	368.6	389.2	339.0	380.8	392.8	390.8
F_2	297.0	320.4	198.3	285.5	326.8	308.1
HCC	286.4	284.8	235.1	274.2	283.9	284.4
HO_2	104.0	86.7	−20.2	−515.2	87.5	90.2
NH_2	74.4	54.1	−67.9	21.4	53.6	61.3
NH_3	−222.0	−222.8	−357.6	−247.8	−207.9	−130.5
NO	2.5	−79.7	−171.8	−107.3	−72.1	−73.5
NO_2	219.3	150.6	−52.6	−3.3	30.2	17.6
O_3	202.9	264.7	187.9	238.4	269.1	261.5
OH	174.0	149.5	−5.8	114.4	149.7	160.2
PH_2	146.1	116.1	45.4	91.6	116.0	115.1

Note: Missing data are due to the basis sets not being defined for the specified atoms. The rest of the data can be found in the Supporting Information section.

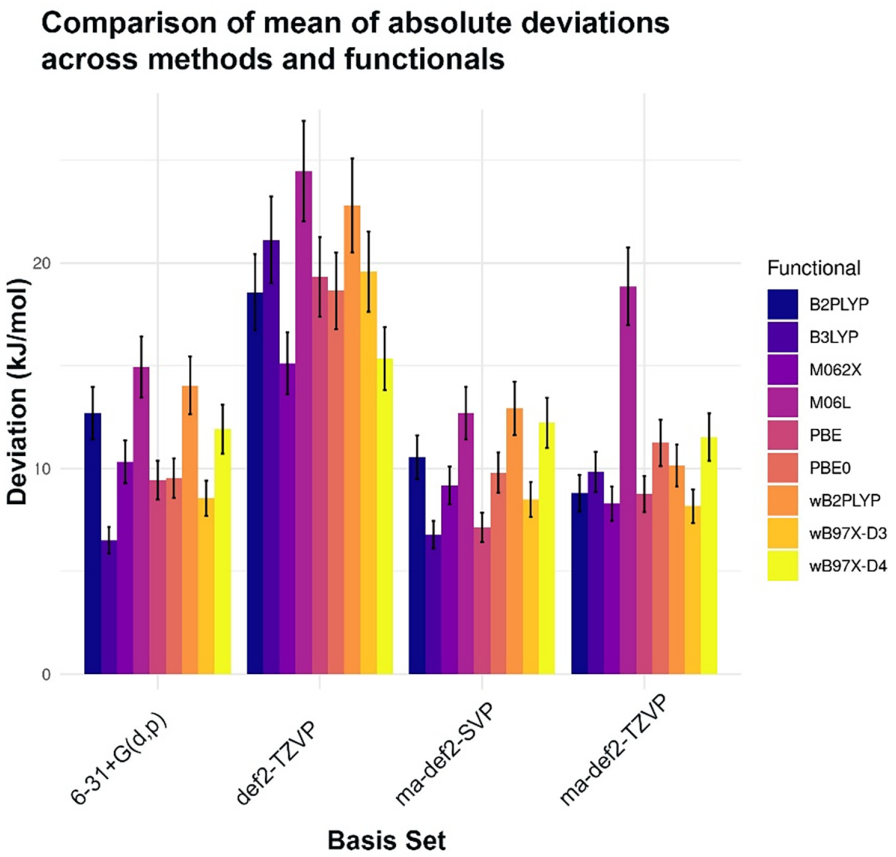

FIGURE 1 | Mean absolute deviation (MAD) in kJ/mol of EA values by basis set and functional after removing outliers. Results from the def2-SVP basis set are omitted due to large errors.

successful DFT is in approximating EA values with reasonable accuracy.

The bar plot in Figure 1 presents the mean absolute deviation (MAD) in units of kJ/mol for the four examined basis sets and functionals after eliminating noisy data. The plot provides insights into the methods' accuracy in reproducing EA values in the literature. Each bar represents the average deviation for a particular functional/basis set combination, while the error bars indicate the standard deviation, reflecting the variability across different molecular species. Note that the results from the def2-SVP basis set are not shown as mentioned earlier.

The *unaugmented* def2-TZVP method exhibits the highest MAD values, suggesting a systematic deviation from reported values compared to other basis sets. Additionally, the large error bars associated with def2-TZVP highlight significant variability, indicating inconsistent performance across different species. In contrast, the ma-def2-SVP and ma-def2-TZVP basis sets generally display lower MAD values, implying better approximations of the reported values. This comes as no surprise, as the minimally augmented (ma-def2-xx) family has recently demonstrated excellent performance in ab initio calculations and has been recommended as the basis set of choice for computational studies [44, 60, 61]. The good performance of the ma-def2 basis sets is attributed to their inclusion of a few diffuse functions, which better describe the extended electron density in anions and loosely bound electrons. This minimal augmentation improves accuracy for properties like electron affinities while keeping the stability and efficiency of the original def2 sets [44, 60, 61].

The 6−31 + G(d,p) basis set also shows acceptable deviations, falling between def2-TZVP and the ma-def2 basis set, indicating good reliability in predicting reported EA values. The good performance of the Pople-type basis set, 6−31 + G(d,p), was supported by a recent comment by Gray et al. [62], especially when the set includes proper polarization and diffuse functions.

Despite this good performance, the ma-def2 basis sets still produced discrepancies for some atomic systems. For example, all methods in this study failed to estimate an electron affinity for the hydride anion (H^-). Despite its simplicity, H^- is a known challenging case because the extra electron is very weakly bound and highly diffuse, with electron density extending far from the nucleus. In such cases, conventional basis sets may not adequately describe the diffuse electron, and more strongly augmented basis sets may be required (cf. Section 4, Limitation of the study).

Beyond the role of basis sets, density functionals play an important role in determining the accuracy of the computed values. Among the functionals examined, B3LYP consistently produces lower deviations. The low MAD values suggest good agreement between calculated and reported values, suggesting that B3LYP predicts EA values to a good accuracy. B3LYP was historically regarded as the standard hybrid functional in computational

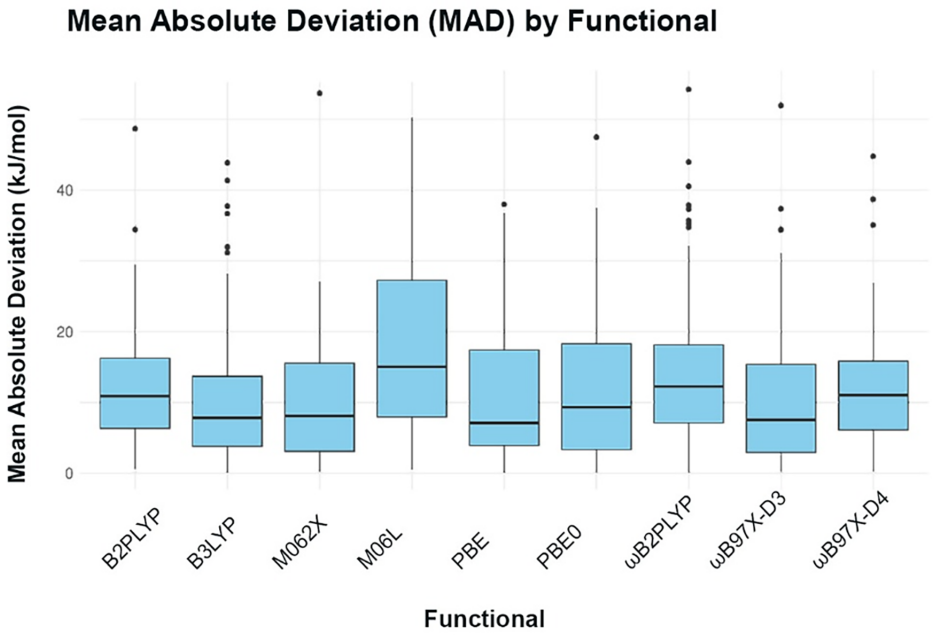

FIGURE 2 | Mean absolute deviation (MAD) in kJ/mol of EA values by functional, averaged across all basis sets and species after removing outliers.

chemistry. However, its performance has recently been heavily criticized [63, 64]. Some authors have even gone so far as to discourage its use for studying the thermochemistry of main-group elements, expressing hope that this would "*inspire a change in the user community's perception of common DFT methods*" [18]. Nevertheless, the good performance of B3LYP in this study highlights its effectiveness in estimating electron affinity (EA) values with the right choice of basis set, an aspect that may not have received sufficient attention in the literature. In the following sections, we will demonstrate how the choice of functional–basis set combination also plays an important role in accurately estimating EA values.

Another important observation from Figure 1 is the magnitude of the error bars, which reflect the consistency of each functional/basis set combination. Larger error bars, particularly noticeable in def2-TZVP, indicate substantial variability, suggesting that the performance of this basis set is highly dependent on the specific molecular system being studied. In contrast, the ma-def2-SVP and ma-def2-TZVP methods show smaller error bars, reinforcing their reputation for molecular systems. This result is consistent with previous studies showing that minimally augmented basis sets provide improved accuracy for systems involving diffuse electron density. This reduces superposition errors, without the computational overhead of full augmentation, ensuring reliable treatment of long-range electrostatics [15, 45].

To assess the predictive accuracy of different computational methods, we analyzed the MAD between computed and reported values. Unlike raw absolute deviations, MAD provides a more stable and representative measure of systematic deviations, minimizing the influence of extreme values. Figure 2 presents a boxplot showing the distribution of MAD across different functionals. A noticeable trend is that some functionals systematically produce lower deviations, while others exhibit broader distributions, indicating inconsistent performance.

Among the functionals analyzed, B3LYP once again shows relatively lower median MAD values with a compact interquartile range (IQR), indicating better and more consistent agreement with reported values.

Note that in Figure 2, MAD values are averaged over all basis sets and species for each functional. This averaging can make functionals like PBE0 appear to perform relatively well in this context. However, this differs from the results shown in Figure 1, where MAD was calculated for each functional-basis set combination. There, PBE0 demonstrated only moderate performance, particularly due to its poor results with the def2-TZVP basis set. This suggests that the choice of basis set can significantly influence the overall MAD and indicates why performance assessments may differ depending on how the data is aggregated.

The boxplot in Figure 3 visualizes the distribution of MAD for different basis sets. Clearly, def2-TZVP exhibits the highest MAD values, indicating that it systematically deviates more from reported values. Additionally, the spread of MAD values for def2-TZVP is larger compared to other basis sets, suggesting greater variability in accuracy across different species. The presence of multiple outliers further reinforces the inconsistency of this method.

In contrast, the ma-def2-SVP and ma-def2-TZVP basis sets show lower MAD values and narrower interquartile ranges (IQRs), reflecting greater stability and more consistent accuracy. The 6–31 + G(d,p) basis set also shows acceptable performance, though its slightly wider spread suggests moderate variability relative to the ma-def2 family. These observations align with earlier analyses and further support the robustness of the minimally augmented basis sets.

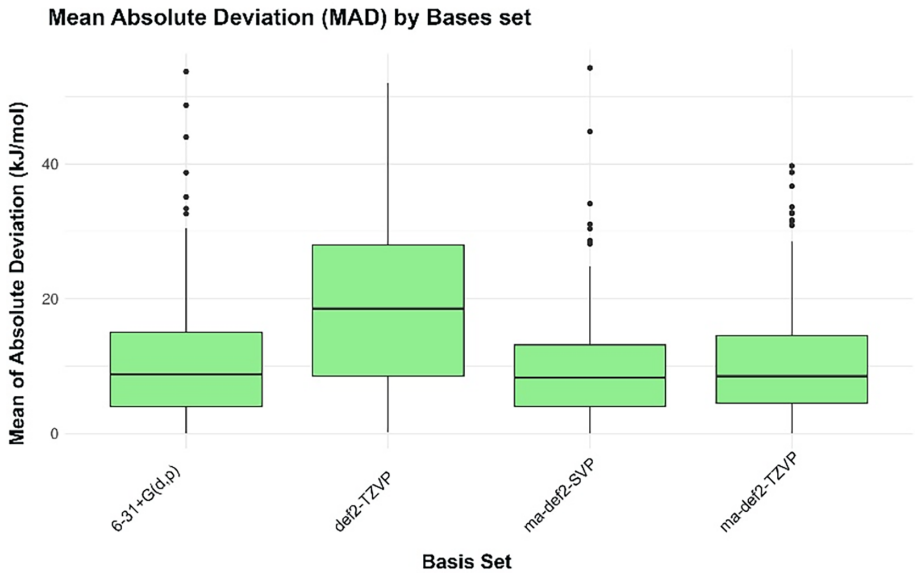

FIGURE 3 | Mean absolute deviation (MAD) in kJ/mol of EA values by basis set across all functionals and species after removing outliers. Results from the def2-SVP basis set are omitted due to large errors.

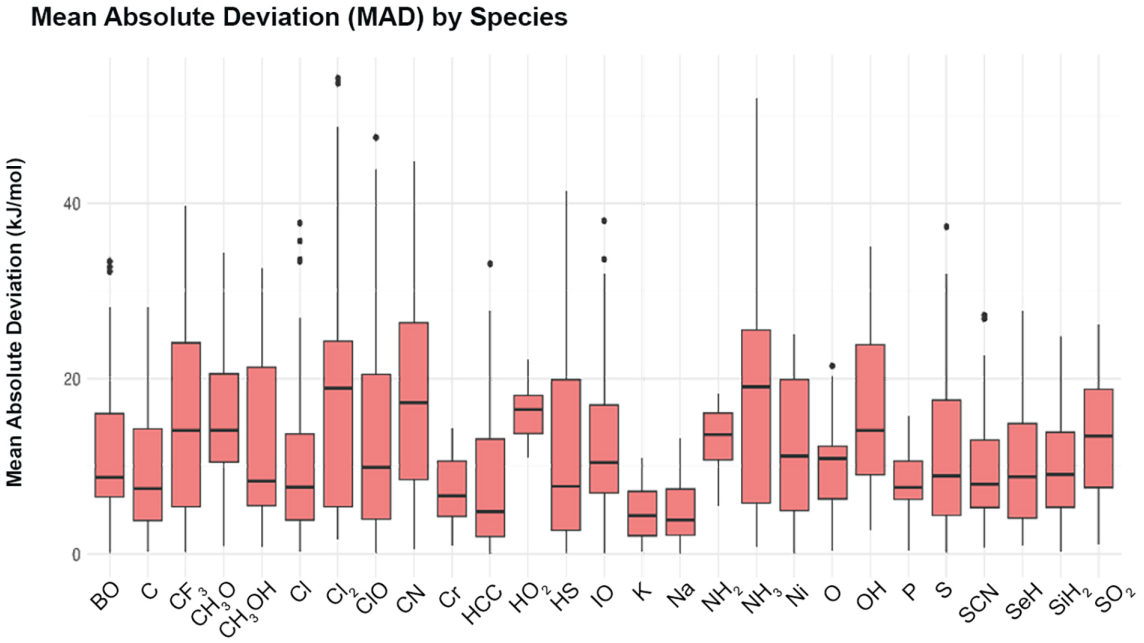

FIGURE 4 | Mean absolute deviation (MAD) in kJ/mol of EA values by species across all basis sets and functionals after removing outliers.

The distribution of MAD for different species is presented in Figure 4. The variation in MAD across species is an important factor in understanding the extent to which ab initio methods capture species-specific interactions accurately. A closer look at Figure 4 suggests that some species exhibit consistently lower MAD values, indicating that they are well-described by the level of theory being used. For example, species such as Na and NH_2 have lower deviations and a compact distribution, indicating that their electronic structure is more predictable across computational methods. In contrast, certain species display higher and more variable MAD values, such as HO_2, CN, and IO, suggesting that their unique electronic or structural properties make them more challenging to model accurately. The presence of a large IQR and outliers for these species further supports this observation, implying that computational methods struggle to achieve reliable predictions for such species.

In this study, heatmaps are employed to visually represent aggregated metrics, such as MAD, to assess how functional/basis sets combinations influence deviations from reported values. This helps to identify which pairings minimize deviations, hence offering better agreement with reported values. Additionally, this

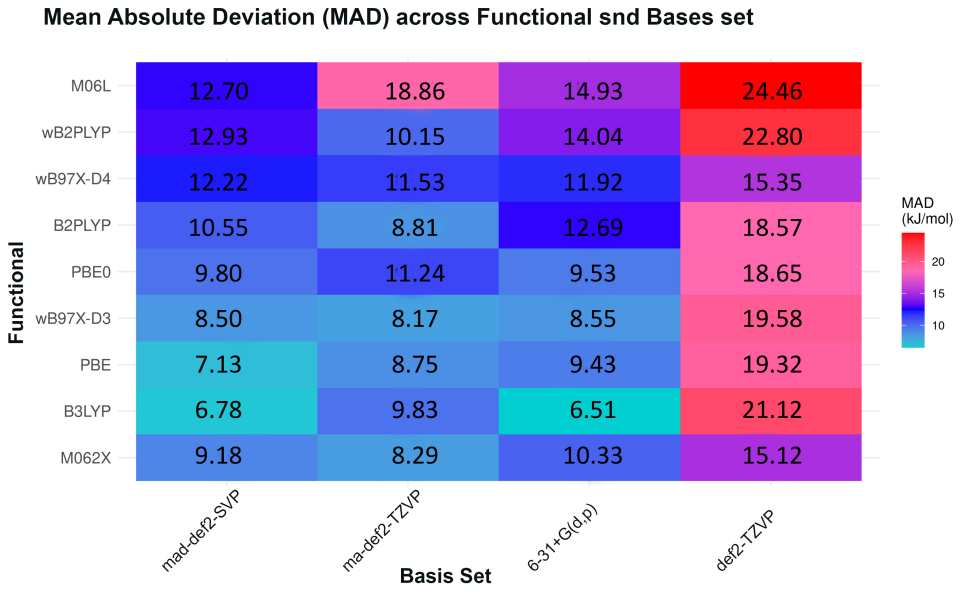

FIGURE 5 | Mean absolute deviation (MAD) of EA values across fontinales and basis sets after removing outliers.

approach helps to detect patterns in computational performance, enabling the detection of trends in stability and accuracy across different basis sets and functionals.

The heatmap in Figure 5 illustrates the MAD for each functional/basis set combination (across all species) after removing the outliers and data from the def2-SVP method as they produce severe noise. As seen in the figure, darker regions correspond to low deviation values, suggesting that these functional/basis set pairings yield better results. Specifically, the combination of ma-def2-SVP with B3LYP and PBE0 falls within the lower deviation range, reinforcing their reliability in this study.

In contrast, brighter regions correspond to higher MAD, indicating problematic functional/basis set combinations that introduce greater errors. The def2-TZVP basis set, for example, when paired with functionals such as M06L and B2PLYP, exhibits significant deviations, emphasizing its limited accuracy in reproducing reported EA values. Accordingly, this suggests that such combinations may lead to larger discrepancies and should be used with caution when high precision is required.

Because NH_3 and CH_3OH have negative experimental electron affinity values, we additionally examined whether their inclusion influenced the main comparative benchmark summaries. Excluding these two species from the screened benchmark summaries did not alter the principal conclusions of the exploratory analysis. In particular, the overall basis-set ordering remained unchanged, the minimally augmented basis sets continued to provide the most reliable overall performance, def2-TZVP remained the least reliable basis set, and the leading low-deviation combinations continued to involve B3LYP with ma-def2-SVP or 6−31 + G(d,p). These results indicate that the main comparative trends reported above are not driven by the two negative-electron-affinity species.

3.2 | Effect of Functional/Basis Sets Interactions

One goal of this study is to assess the performance of different functionals and basis sets in making accurate predictions of EA values. Interaction plots provide distinct but complementary perspectives on the relationship between different variables. The first interaction plot, Figure 6a, examines the reported values (Exp) across functionals, grouped by basis set, while the second plot (Figure 6b) explores the MAD of the computed values across functionals, also grouped by basis set. These plots help in understanding both absolute agreement with experimental (reported) data and the consistency of deviations across basis sets, which are critical in evaluating the reliability of computational approaches.

Figure 6a provides information about how different functionals influence the computed theoretical values relative to experimentally measured values. The lines represent trends for each basis set, showing how computational predictions vary across functionals. The presence of non-parallel lines in Figure 6a indicates strong interactions between functionals and basis sets. If functionals had consistent effects across all basis sets, the trends would be more parallel, reflecting a uniform influence of functionals on computational predictions. However, the significant divergence observed in this plot suggests that some basis sets exhibit greater sensitivity to specific functionals, leading to noticeable fluctuations in their computed values. This interaction implies that functionals are not universally applicable across all basis sets, reinforcing the need of carefully pairing functionals with appropriate computational techniques.

The non-parallel lines in Figure 6a confirm that functionals do not behave uniformly across all basis sets, reinforcing the idea that randomly applying functionals across multiple methods can introduce inconsistencies, which highlights the goal of the study

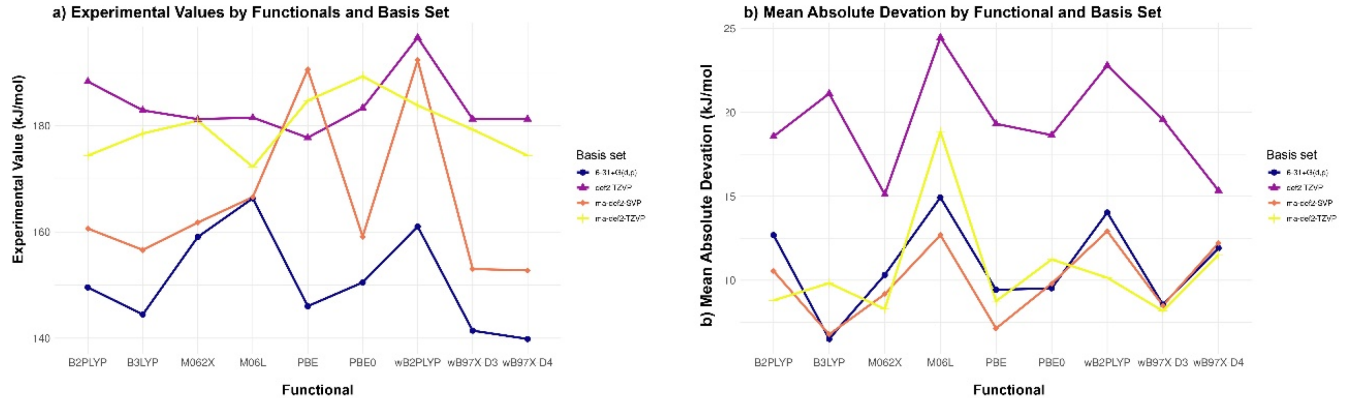

FIGURE 6 | Interaction plot of (a) experimental EA values (reported) as a function of functionals and basis sets and (b) experimental MAD as a function of functionals and basis sets.

regarding the need for basis set-dependent functional selection when comparing theoretical predictions with experimental reference values to ensure accuracy and reproducibility.

Figure 6b, on the other hand, visualizes how MAD of computed values varies across functionals, grouped by basis sets. It also shows which functionals and basis sets introduce the least or most deviations. For example, the M06L functional has a significant peak in MAD for multiple basis sets, which indicates that it is highly sensitive to basis sets. This suggests that M06L may require careful validation and calibration, as its deviation is not consistent across all basis sets. In contrast, B3LYP and PBE0 demonstrate lower MAD values across most basis sets, reinforcing their stability and reliability in minimizing deviations. Their smoother trends suggest that these functionals introduce less variation across different basis sets, making them more suitable for EA approximations that require high precision.

The non-parallel nature of the trends confirms that functionals interact differently with various basis sets, leading to varying levels of deviation. Certain basis sets exhibit sharp fluctuations in MAD for specific functionals, reinforcing that functional selection must be basis set-dependent rather than universally applied.

3.3 | Machine Learning Analysis

The complexity of functional/basis set interactions observed above requires an approach that can capture both linear relationships, where changes in one variable led to proportional changes in another, and nonlinear relationships, where the changes are not proportional, and the relationship is more complex. Traditional linear models may not fully describe how computed values align with experimental data, especially when multiple factors, such as functionals, basis set, and species, affect predictive accuracy. GAM provides a more flexible, data-driven approach by applying smooth, nonlinear transformations to predictor variables [56, 57].

At this phase of the study, the GAM smoothing plot shown in Figure 7 serves as an exploratory tool rather than a predictive model. Its primary objective is to examine whether the relationship between calculated and reported values exhibits nonlinearity, which could indicate underlying method-functional dependencies that affect computational accuracy. Unlike the inferential GAM model used later for factor influence (Section 3.4), Figure 7 is used only as an exploratory smoother to visualize the overall trend between computed and reported EA values. Here, a GAM-based smooth curve is drawn through the scatter to highlight any departure from linearity, without interpreting coefficients or decomposing effects by functional, basis set, or species.

The GAM plot in Figure 7 confirms that the relationship between the computed and reported values is not purely linear, as indicated by the curvature of the red smoothing curve. The presence of nonlinear behavior suggests that different functionals and basis sets may affect accuracy in different ways across reported values. This reinforces the need to further investigate basis set/functional interactions and their role in shaping EA calculations. The observed deviations motivate subsequent statistical and machine learning analysis to formally assess the impact of functionals and basis sets on predictive reliability.

To support machine learning modeling, faceted scatter plots are used to visually compare computed and experimental values across different methods and functionals. By organizing data into subplots for each method and functional, these plots reveal dependencies that might not be apparent in a combined analysis. This approach helps identify basis set– and functional-specific trends, allowing us to assess whether certain computational approaches systematically deviate from experimental values. Additionally, the use of different functionals enables us to determine whether specific functional/basis set combinations consistently yield values higher or lower than experimental references, revealing potential biases.

In the scatter plot of Figure 8, the general alignment of points along the diagonals suggests broad agreement between computed and reported EA values, with certain combinations exhibiting tighter clustering, indicative of better predictive reliability. However, deviations from the diagonal in some basis set/functional combinations suggest that computational accuracy is influenced by both the choice of functional and the basis set.

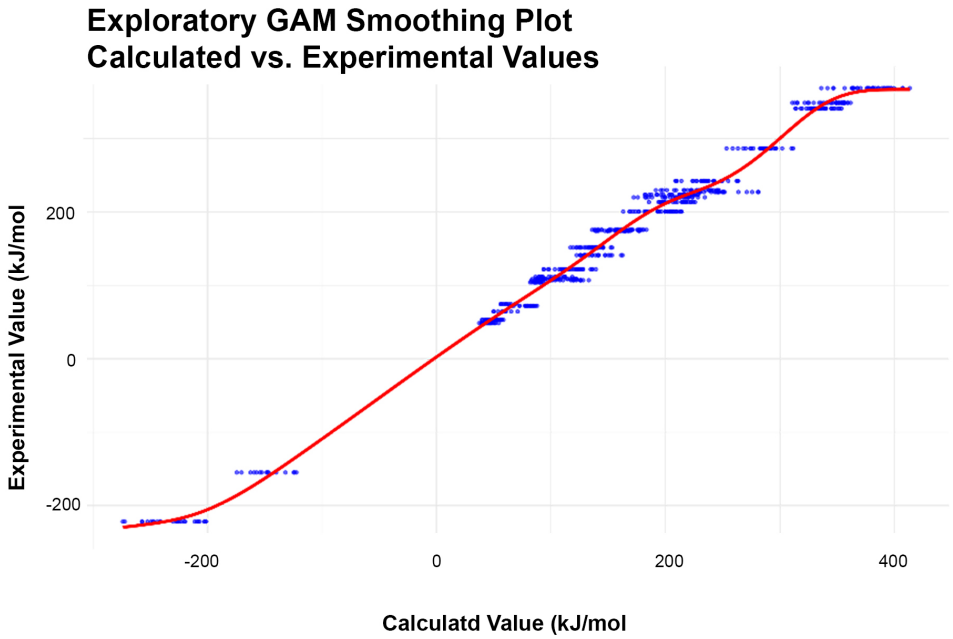

FIGURE 7 | Exploratory generalized additive models (GAM) smoothing plot.

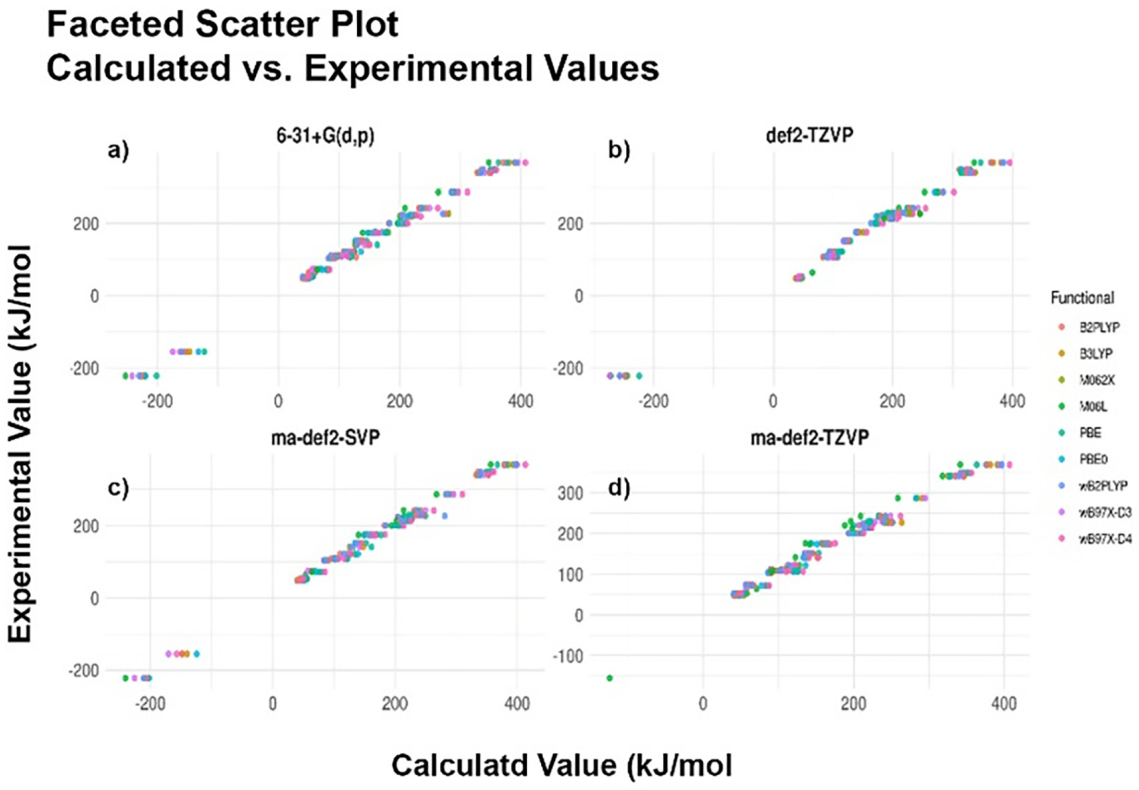

FIGURE 8 | Faceted scatter plot for calculated vs. reported EA values. Points represent the average of EA values for all species obtained for a given functional.

It is clear that the 6−31 + G(d,p) and ma-def2-SVP basis sets show the closest alignment with reported values across multiple functionals, suggesting comparatively stable and accurate performance. In contrast, def2-TZVP exhibits more noticeable spread and greater deviation from the diagonal, indicating that its predictive accuracy is more sensitive to the choice of functional. The ma-def2-TZVP basis set performs moderately well, though its spread is slightly broader than ma-def2-SVP. Within these methods, functionals such as B3LYP and M062X tend to demonstrate more consistent clustering, suggesting greater predictive stability. This observation is consistent with the findings reported in the previous section.

The presence of outliers and scattered deviations in certain basis set/functional combinations suggests complex interactions. For example, this is clear in def2-TZVP, where variability across functionals is more pronounced, reinforcing that functionals behave differently depending on the basis set used.

3.4 | Factor Influence Analysis

Unlike the exploratory smoother in Figure 7, the following GAM is fitted as an inferential model to quantify how basis set, functional, and species contribute to the deviation from experimental EA values.

Building on the descriptive analyses, this section introduces an interpretable machine learning approach to quantify and model the influence of key factors on computational accuracy. Specifically, we shift from purely exploratory statistics to predictive modeling using GAM.

The target variable in this analysis is the absolute deviation (in kJ/mol) from the reference (Exp) value. This outcome reflects the accuracy of EA predictions and is modeled as a function of three main predictors: basis set, functional, and species type.

We initially considered Analysis of Variance (ANOVA) to examine the contribution of these predictors. However, violations of key assumptions, including normality and homoscedasticity, made ANOVA unsuitable. Instead, GAM offers a more flexible, nonparametric framework that captures both linear and nonlinear effects using smooth functions.

In this GAM framework, the absolute deviation (Abs. deviation in kJ/mol) is the response variable, while basis set and functional are included as fixed effects. The species type is modeled as a random effect to account for inherent variability across different molecules. This interpretable ML model allows us to quantify how computational choices influence EA prediction errors and to explore complex dependencies that cannot be captured with traditional linear models.

The GAM used in this study can be expressed as:

$$\begin{aligned} AbsDeviation_i = {} & \beta_0 + \beta_1 \cdot BasisSet_i + \beta_2 \cdot FunctionalIndex_i \\ & + \sum f_j\left(Functional_i \vee BasisSet_j\right) \\ & + f_{Species}\left(Species_i\right) + \varepsilon_i \end{aligned} \tag{8}$$

where β_0 is the intercept, representing the starting point of average deviation when all other factors are at their reference level. β_1 BasisSet and β_2FunctionalIndex are adjustment factors that reflect how the error increases or decreases depending on the basis set or functional used. They are fixed linear effects for basis set and functional index. The smooth functions $f_j\left(Functional_i \vee BasisSet_j\right)$ are smooth nonlinear effects of functionals within each basis set (e.g., a curved trend in def2-TZVP). These smooth functions act like flexible curves; they allow the model to follow more complex, curved patterns in the data, especially when the relationship between functional and error does not follow a straight line. $f_{Species}(.)$is a smooth random effect accounting for species-specific deviations. This term accounts for the fact that some molecules are naturally harder or easier to predict than others, no matter what method is used. $\varepsilon_i$is the residual error, or the leftover error, and represents what's still unexplained by the GAM model.

Additionally, for the GAM used in this modeling step, smooth effects were fitted using penalized spline smoothing within the mgcv framework. Each smooth component is represented through a spline basis (a set of basis functions), and its coefficients are estimated from the data. To reduce overfitting, model fitting includes a roughness penalty that discourages unnecessary curvature, so the fitted curves only become more flexible when the data support it. The amount of smoothing is controlled by smoothing parameters that are estimated automatically during model fitting (using restricted maximum likelihood, REML, where applied). Model complexity is summarized by (i) the parametric coefficients for the fixed effects and (ii) the effective degrees of freedom (edf) reported for each smooth term, where larger edf indicates greater flexibility after penalization. The overall model rank indicates how many of the intended coefficients (parametric and spline-basis coefficients) were identifiable and stably estimated. Species identity was included as a random-effect term to account for systematic, species-specific differences not explained by basis set or functional.

Further, this model allows flexible, curved relationships between predictors and deviation, and accounts for molecule-level variability that is not explained by basis set or functional alone. Accordingly, it helps us understand which combinations of methods tend to give reliable results, and which ones tend to cause larger errors, without assuming that all relationships are simple or linear.

Table 3 shows a summary of the fixed effects according to our GAM model. GAM explained approximately 25.3% of the variation in absolute deviation (adjusted $R^2 = 0.253$) and 28.2% of the deviance. Although the R^2 value is relatively low, this does not indicate that the model is weak or unsuitable, especially since the model rank is 39/40. The model rank here tells how many parameters the model was able to estimate successfully, where each parameter represents part of the model that describes how a factor like a method, basis set, or molecule affects the prediction error. A rank of 39 out of 40 means nearly all model parameters were usable and stable, which suggests the model is well-specified and not overfitting or breaking due to complexity. This is important in scientific datasets, where variability across chemical systems is naturally high.

Based on the results, the def2-TZVP basis set showed a statistically significant increase in deviation compared to the baseline basis set (estimate = +8.96 kJ/mol, $p < 2e{-}16$). The 6–31 + G(d,p) basis set is used as the baseline reference by default in the above GAM model for comparison. This means the model compares the performance of other basis sets to it. An estimate of +8.96 for def2-TZVP means it tends to increase the error by nearly 9 kJ/mol compared to 6–31 + G(d,p), which performed more reliably.

The functional index had a significant linear effect (estimate = +2.16, $p < 2e{-}16$), suggesting that some functionals consistently produce higher deviations. Further, and as Table 4

TABLE 3 | GAM analysis, fixed effects summary.

Term	Estimate	Std. error	*p*	Significance
Intercept	0.0000	0.0000	NA	
def2-TZVP	+8.96	0.98	< 2e−16	Yes
ma-def2-SVP	−0.72	0.91	0.427	
ma-def2-TZVP	+0.11	0.91	0.905	
Functional (index)	+2.16	0.19	< 2e−16	Yes

shows, the smooth effects showed nonlinear interactions between functional and basis set, especially for def2-TZVP (edf = 1.73, $F = 30.33$, $p < 2e{-}16$), indicating performance varies across functionals. Further, the terms for ma-def2-SVP and ma-def2-TZVP were not statistically significant ($p > 0.05$). This means that, on average, these two basis sets did not introduce a clear, consistent increase or decrease in error compared to the baseline. In practical terms, their performance was neither worse nor significantly better than the reference basis set across all functionals and species. This suggests they are reasonably stable and reliable but don't stand out statistically.

While Table 3 focuses on average (linear) effects, Table 4 investigates how performance changes in a more flexible, curved way, especially when the effect of one variable depends on another, for example, how a functional performs differently across basis sets. These are modeled using smooth functions, which can bend and adapt to follow trends in the data. The smooth effect for def2-TZVP (edf = 1.73, $F = 30.33$) suggests that the impact of different functionals on deviation is not constant, but it changes in a nonlinear way.

In simpler terms, this infers that some functionals behave well with this basis set, while others perform poorly, leading to a curved or uneven pattern in error. This explains why def2-TZVP was not stable across all conditions. Similarly, the random smooth term for species (edf = 20.8) was highly significant. This confirms that some molecules are inherently more difficult to model accurately, regardless of the basis set or functional used. Recognizing species-specific behavior helps explain why some predictions are much closer to experimental values than others.

3.5 | Ranked Performance of Variables and Interactions

To identify the best-performing levels of theory in predicting absolute deviations of EA values, we ranked all functionals, basis sets, species, and functional-basis set combinations based on the mean predicted absolute deviation values obtained from the GAM model. The ranking is shown in Figure 9. The chart represents the ranking of density functionals based on the mean predicted absolute deviation (in kJ/mol) from reported values, as estimated by the fitted GAM. Lower values indicate better agreement with reported data and, therefore, better performance. Generally, the top functionals: PBE0, M06L, ωB97X-D3 are very close in accuracy, whereas ωB2PLYP stands out as an underperformer. A spread of 1.6 kJ/mol in mean deviation across functionals shows a significant main effect ($p < 2e{-}16$), indicating that they influence the deviation from experimental values even independently of the basis set.

It is important to note that this ranking reflects the average performance of each functional as estimated by the GAM, after accounting for the effects of basis sets and species. In other words, the functionals were not evaluated in isolation, but their performance was adjusted for and averaged across all tested conditions. The GAM-based ranking showed a spread of only 1.6 kJ/mol among the average predicted deviations of the tested functionals, which is small relative to the commonly accepted range of computational accuracy discussed in the introduction. This indicates that, although statistically significant differences exist among functionals, these differences remain minor from a chemical perspective under the conditions examined here.

Despite the significant difference predicted by the GAM model (1.6 kJ/mol), the bars in Figure 1 show that there is no meaningful difference from a chemical perspective, as a spread of 1.6 kJ/mol is well within the accepted margin of computational accuracy. This highlights an important point: all functionals in this study perform comparably well. However, as we will see in the following sections, this consistency does not hold when examining the influence of basis sets, species types, or their combinations.

The ranking presented in Figure 10 reflects the main effect of each basis set, meaning it represents their average contribution to prediction error after accounting for differences in functionals and species. This allows for a fair comparison of basis sets across all tested conditions. According to Figure 10, show the ranking of the basis sets based on their predicted absolute deviation from reported experimental values (in kJ/mol), as generated by the GAM model. Clearly, the ma-def2-SVP basis set demonstrates the lowest deviation, making it the best-performing basis set in terms of accuracy.

Additionally, the ma-def2-TZVP and 6−31 + G(d,p) basis sets show slightly lower performance but still offer competitive results, with deviations around ˜10.7−10.8 kJ/mol. Moreover, def2-TZVP shows a significantly higher deviation than all others, making it the least reliable method in this context.

The ranking in Figure 11 is based on the species-specific random effect estimated by the GAM model. It reflects how each species contributes to the prediction error on average, independently of the basis set or functional used. This provides insight

TABLE 4 | GAM analysis, smooth term summary.

Smooth term	edf	Ref.df	*F*	*p*	Significance
s(Species)	20.83	26	4.46	< 2e−16	Yes
s(Functional): basis set 6−31 + G(d,p)	1.00	1.00	30.49	< 2e−16	Yes
s(Functional): basis set def2-TZVP	1.73	1.93	30.33	< 2e−16	Yes
s(Functional): basis set ma-def2-SVP	1.00	1.00	35.41	< 2e−16	Yes
s(Functional): basis set ma-def2-TZVP	1.00	1.00	72.58	< 2e−16	Yes

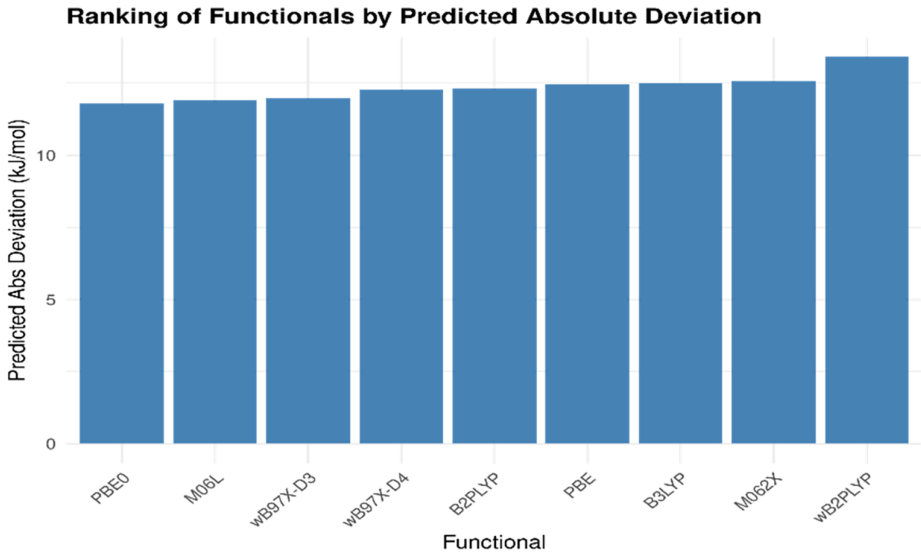

FIGURE 9 | Ranking of functional by predicted absolute deviations after removing outliers according to GAM.

into how inherently difficult each molecule is to model. The predicted absolute deviation between computed and reported values varies significantly across species, as shown in Figure 11. Species such as Na, K, and P exhibit the lowest mean deviations (approximately 6–9 kJ/mol), indicating that the computational models used in this work provided high accuracy. This variability supports the modeling of species as a random effect in the GAM framework, acknowledging species-specific influences on prediction error that are not fully captured by functional or basis set alone. The wide range of prediction errors (from 5.98 to 19.4 kJ/mol) highlights the importance of species identity in evaluating the reliability of computational chemistry methods and emphasizes the need for tailored approaches for more complex systems.

Unlike the previous figures, Figure 12 reflects the combined effects of functionals and basis sets, rather than their average performance alone. Figure 12 illustrates the interaction effects between functionals and basis sets based on the predicted absolute deviation from reported values. This interaction reveals critical performance differences that would not be observable when evaluating functionals or basis sets independently. For example, B3LYP, which is historically regarded as the standard hybrid functional, shows excellent accuracy with the 6−31 + G(d,p) method (mean deviation = 8.94 kJ/mol) and ma-def2-SVP (8.68 kJ/mol), but performs poorly with def2-TZVP, reaching the highest deviation among all combinations (21.5 kJ/mol). Similarly, the 6−31 + G(d,p) basis set generally shows low deviations for many functionals, supporting its reputation as a reliable basis set. However, the ranking of functionals changes significantly across basis sets. M06L and B3LYP perform best under ma-def2-SVP. These findings highlight the importance of considering basis set-functional interactions, demonstrating that a functional's performance cannot be universally ranked without accounting for the accompanying basis sets. This insight distinguishes our study from previous work that evaluated functionals in isolation, contributing a novel perspective to computational chemistry research.

4 | Limitation of the Study and Future Work

Despite the insights provided by this study, several limitations should be noted. First, the statistical framework, including the GAM analysis, depends on the size and composition of the dataset. The study used a curated dataset of 47 atomic and molecular species with 2115 calculated electron affinity values from combinations of nine density functionals and five basis sets. While this dataset captures a representative range of chemical systems, it does not encompass the full diversity of atoms, molecules, and electronic structures in computational chemistry. Therefore, the statistical relationships identified by the GAM should be interpreted as trends within the examined dataset, not as universal conclusions.

A second limitation stems from the inherent difficulty of predicting adiabatic electron affinities using the direct ΔSCF method. Some systems may yield negative calculated EA values under several levels of theory. This often occurs when the additional electron is weakly bound and requires highly diffuse basis functions, particularly for atomic systems. For example, the hydrogen atom binds an extra electron very weakly, producing a highly diffuse electronic distribution that conventional basis sets may inadequately represent. In such cases, standard basis sets can artificially destabilize the anionic state, leading to large errors or negative predicted values.

Future work should focus on expanding both the computational methodology and the statistical framework. Using larger and more diffuse basis sets, as well as alternative electronic structure methods, may improve the description of weakly bound anions.

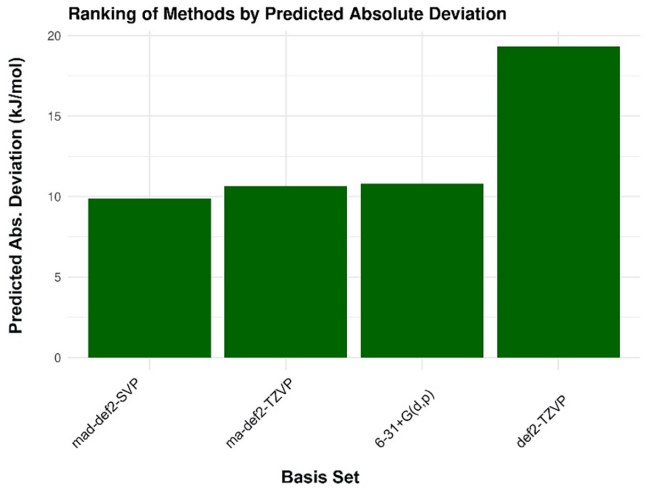

FIGURE 10 | Ranking of basis sets by predicted absolute deviation according to GAM.

5 | Conclusion

This study presents a comprehensive benchmarking study of electron affinity (EA) predictions using a broad set of density functionals and basis sets, evaluated through statistical modeling and machine learning. The EA values of 47 species were calculated using nine different density functionals and five different basis sets that cover different categories.

The statistical analysis revealed a strong correlation between the choice of density functional and basis set, and the species or their combinations. For example, although the historically standard density functional B3LYP has been heavily criticized in recent literature, its performance in predicting EA values in this study is notably high when paired with an appropriate basis set. Similarly, functionals such as M06L show a high sensitivity to the choice of basis set.

By integrating a flexible Generalized Additive Model (GAM), the results in this research reveal that the impact of basis sets is stronger and more diverse on prediction accuracy compared to that of functionals. Because of its small size ad lack of diffuse functions, the def2-SVP basis failed to most of the reported EA values for species under study, and therefore it is not recommended for EA calculations. The def2-TZVP basis set produced higher deviations owing to it is nonlinear interactions with the functionals. On the other hand, the most accurate and stable basis set is identified to be ma-def2-SVP, especially when combined with functionals such as M06L and B3LYP. Furthermore, species identity exhibited a statistically significant random

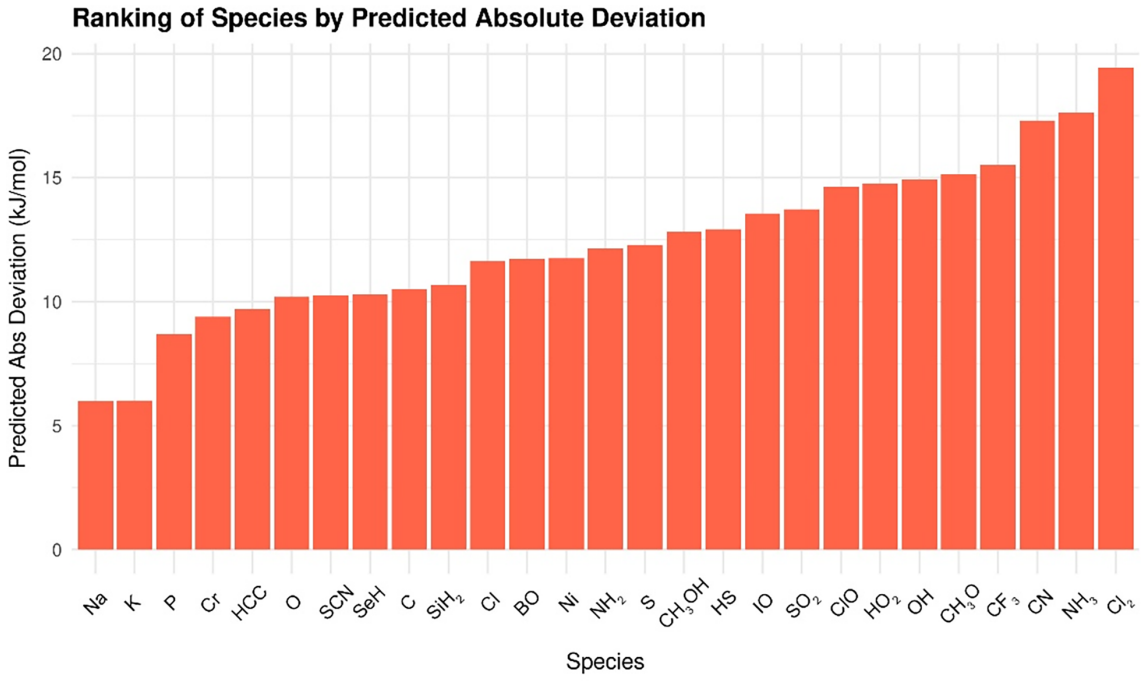

FIGURE 11 | Ranking of species by predicted absolute deviation according to GAM.

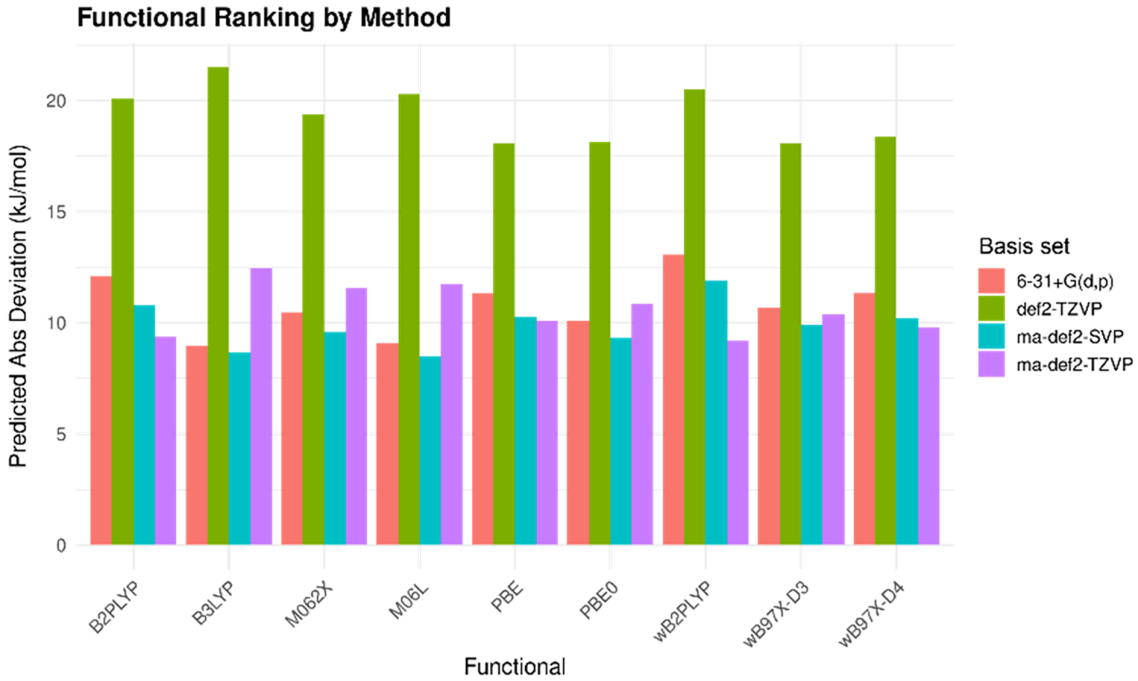

FIGURE 12 | Ranking of functional by basis sets according to GAM.

effect, with some molecules proving difficult to model, such as Cl_2 and CN.

An additional sensitivity check excluding the two negative-electron-affinity species, NH_3 and CH_3OH, did not change the main comparative conclusions regarding basis-set ranking and the leading functional–basis set combinations. At the same time, these systems remain a limitation of the present bound-state Δ_{SCF} framework, and future work should extend this interpretable benchmarking strategy using dedicated methodologies for temporary/metastable anions.

The key outcomes of this research are twofold. First, the final ranking provided by the machine learning model shows that although there are statistically significant differences in the performance of various density functionals, these differences are relatively minor in terms of chemical relevance. This suggests that all tested functionals perform comparably across the examined basis sets and species. In contrast, the choice of basis set has a much stronger impact on predictive accuracy.

The second key outcome is that the GAM-based ranking showed only a narrow spread of 1.6 kJ/mol among the average predicted deviations of the tested functionals. This indicates that, despite statistically significant differences, the practical differences among functionals are relatively small from a chemical perspective under the studied conditions.

Author Contributions

Abdelrahman Eid: visualization, formal analysis, writing – original draft, writing – review and editing, validation, methodology, software. **Ismail Badran:** conceptualization, methodology, validation, writing – original draft, project administration, software. **Sahar Salman:** validation, writing – review and editing, data curation. **Motasem Far:** data curation, validation, writing – review and editing. **Nadeen Abbas:** validation, writing – review and editing, data curation. **Raghad Tayeh:** data curation, writing – review and editing, validation. **Yasmeen Hamdan:** data curation, validation, writing – review and editing.

Acknowledgments

The authors are grateful to An-Najah National University for supporting this research. The authors are also thankful to the Centre of Excellence in High Energy Physics in Palestine (CEHEP2) hosted at An-Najah National University, Dr. Ahmed Bassalat, and Mr. Muhammad Adas for providing the computational resources used in this research.

Funding

The authors have nothing to report.

Ethics Statement

The authors have nothing to report.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Mendeley Data, V1 at https://doi.org/10.17632/wt9nwrpz9f.1

References

1. J. Gu, Y. Xie, and H. F. Schaefer, "Benchmarking the Electron Affinity of Uracil," *Journal of Chemical Theory and Computation* 10, no. 2 (2014): 609–612.

2. IUPAC, *Compendium of Chemical Terminology, ('Gold Book')* (Blackwell Scientific Publications, 2019).

3. T. Vikramaditya and S.-T. Lin, "Assessing the Role of Hartree-Fock Exchange, Correlation Energy and Long Range Corrections in Evaluating Ionization Potential, and Electron Affinity in Density Functional Theory," *Journal of Computational Chemistry* 38, no. 21 (2017): 1844–1852.

4. Y. Tao, K. Yuan, T. Chen, et al., "Thermally Activated Delayed Fluorescence Materials Towards the Breakthrough of Organoelectronics," *Advanced Materials* 26, no. 47 (2014): 7931–7958.

5. J.-C. Lee, J.-D. Chai, and S.-T. Lin, "Assessment of Density Functional Methods for Exciton Binding Energies and Related Optoelectronic Properties," *RSC Advances* 5, no. 123 (2015): 101370–101376.

6. R. M. Richard, M. S. Marshall, O. Dolgounitcheva, et al., "Accurate Ionization Potentials and Electron Affinities of Acceptor Molecules I. Reference Data at the CCSD(T) Complete Basis Set Limit," *Journal of Chemical Theory and Computation* 12, no. 2 (2016): 595–604.

7. P. Dedíková, P. Neogrády, and M. Urban, "Electron Affinities of Small Uracil–Water Complexes: A Comparison of Benchmark CCSD(T) Calculations With DFT," *Journal of Physical Chemistry. A* 115, no. 11 (2011): 2350–2358.

8. J. Schiedt, R. Weinkauf, D. M. Neumark, and E. W. Schlag, "Anion Spectroscopy of Uracil, Thymine and the Amino-Oxo and Amino-Hydroxy Tautomers of Cytosine and Their Water Clusters," *Chemical Physics* 239, no. 1 (1998): 511–524.

9. G. de Oliveira, J. M. L. Martin, F. de Proft, and P. Geerlings, "Electron Affinities of the First- and Second-Row Atoms: Benchmark Ab Initio and Density-Functional Calculations," *Physical Review A* 60, no. 2 (1999): 1034–1045.

10. G. S. Tschumper and H. F. Schaefer, III, "Predicting Electron Affinities With Density Functional Theory: Some Positive Results for Negative Ions," *Journal of Chemical Physics* 107, no. 7 (1997): 2529–2541.

11. J. M. Galbraith and H. F. Schaefer, III, "Concerning the Applicability of Density Functional Methods to Atomic and Molecular Negative Ions," *Journal of Chemical Physics* 105, no. 2 (1996): 862–864.

12. N. Mardirossian and M. Head-Gordon, "Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals," *Molecular Physics* 115, no. 19 (2017): 2315–2372.

13. M. Bogojeski, L. Vogt-Maranto, M. E. Tuckerman, K.-R. Müller, and K. Burke, "Quantum Chemical Accuracy From Density Functional Approximations via Machine Learning," *Nature Communications* 11, no. 1 (2020): 5223.

14. J. W. Knight, X. Wang, L. Gallandi, et al., "Accurate Ionization Potentials and Electron Affinities of Acceptor Molecules III: A Benchmark of GW Methods," *Journal of Chemical Theory and Computation* 12, no. 2 (2016): 615–626.

15. M. Bursch, J.-M. Mewes, A. Hansen, and S. Grimme, "Best-Practice DFT Protocols for Basic Molecular Computational Chemistry," *Angewandte Chemie (International Edition in English)* 61, no. 42 (2022): e202205735.

16. B. Chan and A. Karton, "Assessment of DLPNO-CCSD(T)-F12 and Its Use for the Formulation of the Low-Cost and Reliable L-W1X Composite Method," *Journal of Computational Chemistry* 43, no. 21 (2022): 1394–1402.

17. K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, "A Fifth-Order Perturbation Comparison of Electron Correlation Theories," *Chemical Physics Letters* 157, no. 6 (1989): 479–483.

18. L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi, and S. Grimme, "A Look at the Density Functional Theory Zoo With the Advanced GMTKN55 Database for General Main Group Thermochemistry, Kinetics and Noncovalent Interactions," *Physical Chemistry Chemical Physics* 19, no. 48 (2017): 32184–32215.

19. M. Amati, S. Stoia, and E. J. Baerends, "The Electron Affinity as the Highest Occupied Anion Orbital Energy With a Sufficiently Accurate Approximation of the Exact Kohn–Sham Potential," *Journal of Chemical Theory and Computation* 16, no. 1 (2020): 443–452.

20. J. K. Cooper, C. D. Grant, and J. Z. Zhang, "Ab Initio Calculation of Ionization Potential and Electron Affinity of Six Common Explosive Compounds," *Reports in Theoretical Chemistry* 1 (2012): 11–19.

21. C. P. Vibert and D. J. Tozer, "Simple DFT Scheme for Estimating Negative Electron Affinities," *Journal of Chemical Theory and Computation* 15, no. 1 (2018): 241–248.

22. A. Ramírez-Solís, "On the Accuracy of the Complete Basis Set Extrapolation for Anionic Systems: A Case Study of the Electron Affinity of Methane," *Computational Chemistry* 2, no. 02 (2014): 31–41.

23. I. Badran, K. Hashlamoun, and N. N. Nassar, "Bond Dissociation Energies of the Fifth-Row Elements (In—I): A Quantum Theoretical Benchmark Study," *International Journal of Quantum Chemistry* 123, no. 23 (2023): e27222.

24. I. Badran, "From Germolane to Germylenes: A Theoretical DFT Study of Thermal Decomposition Pathways and Reactivity," *Journal of Coordination Chemistry* 77, no. 20–21 (2024): 2440–2452.

25. K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, "Machine Learning for Molecular and Materials Science," *Nature* 559, no. 7715 (2018): 547–555.

26. K. Schwab, "The Fourth Industrial Revolution: What It Means, How to respond1," in *Handbook of Research on Strategic Leadership in the Fourth Industrial Revolution* (Edward Elgar Publishing, 2024), 29–34.

27. A. Eid, S. Jodeh, R. Eid, G. Hanbali, A. Chakir, and E. Roth, "Adaptive Data-Driven Framework for Unsupervised Learning of Air Pollution in Urban Micro-Environments," *Atmosphere* 17, no. 2 (2026): 125.

28. M. Assali, S. Sawalha, R. Hamad, I. Badran, and A. Eid, "Chitosan-Functionalized Amino Acids as Biostimulants for Advancing Sustainable Agriculture," *International Journal of Biological Macromolecules* 329 (2025): 147831.

29. P. Sahu, A. Kumar, R. Jain, K. Upreti, D. K. Yadav, and G. Radhakrishnan, "Explainable Hybrid Deep Learning Framework With Multimodal Inputs for Diabetic Retinopathy Detection," *An-Najah University Journal for Research-A (Natural Sciences)* 40, no. 3 (2025).

30. F. A. Faber, L. Hutchison, B. Huang, et al., "Prediction Errors of Molecular Machine Learning Models Lower Than Hybrid DFT Error," *Journal of Chemical Theory and Computation* 13, no. 11 (2017): 5255–5264.

31. G. R. Schleder, A. C. Padilha, C. M. Acosta, M. Costa, and A. Fazzio, "From DFT to Machine Learning: Recent Approaches to Materials Science–a Review," *Journal of Physics: Materials* 2, no. 3 (2019): 032001.

32. T. Tsuneda, J.-W. Song, S. Suzuki, and K. Hirao, "On Koopmans' Theorem in Density Functional Theory," *Journal of Chemical Physics* 133, no. 17 (2010): 1–49.

33. J. F. Janak, "Proof That $\frac{\partial E}{\partial n_i} = \epsilon$ in Density-Functional Theory," *Physical Review B* 18, no. 12 (1978): 7165–7168.

34. C.-G. Zhan, J. A. Nichols, and D. A. Dixon, "Ionization Potential, Electron Affinity, Electronegativity, Hardness, and Electron Excitation Energy: Molecular Properties From Density Functional Theory Orbital Energies," *Journal of Physical Chemistry. A* 107, no. 20 (2003): 4184–4195.

35. C. Adamo and V. Barone, "Toward Reliable Density Functional Methods Without Adjustable Parameters: The PBE0 Model," *Journal of Chemical Physics* 110, no. 13 (1999): 6158–6170.

36. A. Najibi, M. Casanova-Páez, and L. Goerigk, "Analysis of Recent BLYP- and PBE-Based Range-Separated Double-Hybrid Density Functional Approximations for Main-Group Thermochemistry, Kinetics, and Noncovalent Interactions," *Journal of Physical Chemistry. A* 125, no. 18 (2021): 4026–4035.

37. A. D. Becke, "Density-Functional Thermochemistry. III. The Role of Exact Exchange," *Journal of Chemical Physics* 98, no. 7 (1993): 5648–5652.

38. Y. Zhao and D. G. Truhlar, "The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals," *Theoretical Chemistry Accounts* 120, no. 1 (2008): 215–241.

39. M. Walker, A. J. A. Harvey, A. Sen, and C. E. H. Dessent, "Performance of M06, M06-2X, and M06-HF Density Functionals for Conformationally Flexible Anionic Clusters: M06 Functionals Perform Better Than B3LYP for a Model System With Dispersion and Ionic Hydrogen-Bonding Interactions," *Journal of Physical Chemistry. A* 117, no. 47 (2013): 12590–12600.

40. S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu," *Journal of Chemical Physics* 132, no. 15 (2010): 154104.

41. S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the Damping Function in Dispersion Corrected Density Functional Theory," *Journal of Computational Chemistry* 32, no. 7 (2011): 1456–1465.

42. L. Goerigk and S. Grimme, "Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals—Evaluation With the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions," *Journal of Chemical Theory and Computation* 7, no. 2 (2011): 291–309.

43. M. Casanova-Páez, M. B. Dardis, and L. Goerigk, "ωB2PLYP and ωB2GPPLYP: The First Two Double-Hybrid Density Functionals With Long-Range Correction Optimized for Excitation Energies," *Journal of Chemical Theory and Computation* 15, no. 9 (2019): 4735–4744.

44. F. Neese, F. Wennmohs, U. Becker, and C. Riplinger, "The ORCA Quantum Chemistry Program Package," *Journal of Chemical Physics* 152, no. 22 (2020): 224108.

45. F. Neese, "Software Update: The ORCA Program System—Version 5.0," *WIREs Computational Molecular Science* 12, no. 5 (2022): e1606.

46. G. L. Stoychev, A. A. Auer, and F. Neese, "Efficient and Accurate Prediction of Nuclear Magnetic Resonance Shielding Tensors With Double-Hybrid Density Functional Theory," *Journal of Chemical Theory and Computation* 14, no. 9 (2018): 4756–4771.

47. G. L. Stoychev, A. A. Auer, R. Izsák, and F. Neese, "Self-Consistent Field Calculation of Nuclear Magnetic Resonance Chemical Shielding Constants Using Gauge-Including Atomic Orbitals and Approximate Two-Electron Integrals," *Journal of Chemical Theory and Computation* 14, no. 2 (2018): 619–637.

48. A. Sen, B. de Souza, L. M. J. Huntington, M. Krupička, F. Neese, and R. Izsák, "An Efficient Pair Natural Orbital Based Configuration Interaction Scheme for the Calculation of Open-Shell Ionization Potentials," *Journal of Chemical Physics* 149, no. 11 (2018): 114108.

49. M. C. R. Melo, R. C. Bernardi, T. Rudack, et al., "NAMD Goes Quantum: An Integrative Suite for Hybrid Simulations," *Nature Methods* 15, no. 5 (2018): 351–354.

50. M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison, "Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform," *Journal of Cheminformatics* 4, no. 1 (2012): 17.

51. D. C. Montgomery, *Design and Analysis of Experiments* (John Wiley & Sons, 2017).

52. RCT, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2024), https://wwwR-projectorg/.

53. H. Wickham, "Dplyr: A Grammar of Data Manipulation," *R Package Version 04* 3 (2015): 156.

54. H. Wickham, D. Vaughan, and M. Girlich, *Tidyr: Tidy Messy Data [R Package Tidyr Version 1.3. 0]* (Comprehensive R Archive Network (CRAN), 2023).

55. H. Wickham and J. Bryan, readxl: Read Excel Files [Internet], 2023.

56. S. N. Wood, *Generalized Additive Models: An Introduction With R* (Chapman and Hall/CRC, 2017).

57. T. Hastie and R. Tibshirani, "Generalized Additive Models," *Statistical Science* 1, no. 3 (1986): 297–310.

58. H. Wickham and C. Sievert, *ggplot2: Elegant Graphics for Data Analysis* (springer New York, 2009).

59. H. Wickham, "reshape2: Flexibly Reshape Data: A Reboot of the Reshape Package," *R Package Version* 1, no. 4 (2020).

60. J. Zheng, X. Xu, and D. G. Truhlar, "Minimally Augmented Karlsruhe Basis Sets," *Theoretical Chemistry Accounts* 128, no. 3 (2011): 295–305.

61. B. Chan, "High-Level Quantum Chemistry Reference Heats of Formation for a Large Set of C, H, N, and O Species in the NIST Chemistry Webbook and the Identification and Validation of Reliable Protocols for Their Rapid Computation," *Journal of Physical Chemistry. A* 126, no. 30 (2022): 4981–4990.

62. M. Gray, P. E. Bowling, and J. M. Herbert, "Comment on "Benchmarking Basis Sets for Density Functional Theory Thermochemistry Calculations: Why Unpolarized Basis Sets and the Polarized 6-311G Family Should be Avoided"," *Journal of Physical Chemistry. A* 128, no. 36 (2024): 7739–7745.

63. L. Goerigk and S. Grimme, "A Thorough Benchmark of Density Functional Methods for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions," *Physical Chemistry Chemical Physics* 13, no. 14 (2011): 6670–6688.

64. I. Badran and Z. Talie, "Kinetics of Alizarin Dye Hydrolysis in Alkaline Medium for Wastewater Treatment," *Iran Journal of Chemistry and Chemical Engineering (IJCCE)* 40, no. 5 (2021): 1490–1501.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Electron affinity of atoms and molecules calculated using different functional sat five different basis sets. **Table S2:** Dataset counts used in the sensitivity check. **Table S3:** The main conclusions after removal of negative values.