Electronic Supplementary Information

Perovskite catalysts for carbon monoxide oxidation: a critical review of activity-associated parameters, mechanisms, and deactivation

Yuying Wang[a], Zhicheng Li[b], Jiaxin Song[a], Tong Yu[a], Binhong Zhao[a], Fan Zhou[a], Yiming He[a], Botao Qiao[c], Yunfei Gao[a,*], Fuchen Wang[a,*]

[a] Institute of Clean Coal Technology, East China University of Science and Technology, Shanghai, 200237, China

[b] School of Mechanical and Power Engineering, East China University of Science and Technology, Shanghai, 200237, China

[c] CAS Key Laboratory of Science and Technology on Applied Catalysis, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, 116023 Dalian, China

To construct a restricted subset with broadly comparable reaction conditions for quantitative analysis, studies were retained only when they met all of the following criteria:

(i) the target reaction was direct CO oxidation over ABO_3-type perovskite or doped perovskite catalyst;

(ii) the T_{50} value (reaction temperature corresponding to 50% CO conversion) was explicitly reported or could be reliably extracted from the literature;

(iii) no additional co-reactants including H_2O, CO_2, SO_2, hydrocarbons, NO_x, or soot were introduced into the feed gas;

(iv) the O_2/CO molar ratio in the feed was constrained within $0.5 \leq O_2/CO \leq 21$. The lower limit ensures that the feed was not oxygen-deficient relative to the stoichiometric requirement for complete CO oxidation, while the upper limit prevents excessively O_2-rich conditions from artificially facilitating CO oxidation and lowering

the apparent T_{50}. This upper limit also allows typical feeds such as 1% CO balanced with dry air to be included, for which the O_2/CO ratio is approximately 20.95.

(v) the reported space velocity or mass-normalized flow rate was not lower than the typical lower bound used in the selected literature, namely GHSV ≥ 6000 h^{-1} or mass-normalized flow rate ≥ 12000 mL $gcat^{-1}$ h^{-1}, depending on the unit reported in the original article. This criterion was adopted because a lower space velocity increases the gas-solid contact time and can artificially facilitate CO conversion, thereby lowering the apparent T_{50}. By excluding data obtained under excessively low space velocities, the comparison better reflects catalyst activity under relatively demanding and comparable reaction conditions. Entries with missing or uninterpretable space-velocity information were excluded.

A total of 38 valid catalyst samples were finally obtained after screening and standardized sorting.

1. Variable Encoding: Dummy Variable Encoding for Categorical Factors

Categorical variables must be transformed into numerically interpretable features via dummy variable encoding for regression modeling. For a categorical variable with k levels, $k-1$ dummy variables are constructed to avoid perfect multicollinearity in the design matrix. One level is set as the reference level, and the other levels are represented by binary (0/1) dummy variables: the dummy variable is assigned a value of 1 if the sample belongs to the corresponding level, and 0 otherwise. When all dummy variables for a factor are 0, the sample corresponds to the reference level.

2. Classification of Factor Levels and Dummy Variable Definition

In this study, 4 core categorical influencing factors and their levels were defined based on the physicochemical characteristics of perovskite catalysts and literature data. The reference level for each factor was selected following two principles: (i) the most classic and widely used formulation in the perovskite CO oxidation system; (ii) the level with the largest sample size to ensure the robustness of parameter estimation. The factor levels and reference levels are summarized in Table S1.

Table S1 Definition of categorical factors, levels and reference levels.

Factor	Levels	Reference Level
B-site Metal Type	A_1 (Fe-based)	A_1
	A_2 (Mn-based)	
	A_3 (Co-based)	
	A_4 (Dual-metal-based)	
A-site Doping Type	B_1 (La single component)	B_1
	B_2 (La-alkaline earth metal doping)	
	B_3 (La-Ce co-doping)	
	B_4 (Ba-based)	
Support Loading Condition	C_1 (Unsupported)	C_1
	C_2 (Supported)	
Preparation Method	D_1 (Sol-gel method)	D_1
	D_2 (Coprecipitation method)	
	D_3 (Other methods)	

3. Results of Dummy Variable Encoding

Taking the B-site metal type (Factor A) as an example, this factor contains 4 levels, which were encoded into 3 dummy variables as shown in Table S2.

Table S2 Dummy variable encoding rules for B-site metal type.

Category Level	x_{A_2}	x_{A_3}	x_{A_4}
Fe-based (A_1, Ref.)	0	0	0
Mn-based (A_2)	1	0	0
Co-based (A_3)	0	1	0

Dual-metal-based (A_4)	0	0	1

The other 3 categorical factors were encoded following the same rule. After encoding, a total of 9 main-effect dummy variables were obtained as the input features for subsequent modeling. The encoded dataset for partial representative samples is shown in Table S3.

Table S3 Encoded dataset for representative catalyst samples.

No.	Catalyst	x_{A_2}	x_{A_3}	x_{A_4}	x_{B_2}	x_{B_3}	x_{B_4}	x_{C_2}	x_{D_2}	x_{D_3}
1	$LaFeO_3$	0	0	0	0	0	0	0	0	0
2	$LaAl_{0.2}Mn_{0.8}O_3$	1	0	0	0	0	0	0	0	0
3	$LaCoO_3$	0	1	0	0	0	0	0	0	0
4	$LaCo_{0.5}Mn_{0.5}O_3$	0	0	1	0	0	0	0	0	0
…	…	…	…	…	…	…	…	…	…	…
38	$La_{0.7}Sr_{0.3}Mn0.25Co_{0.75}O_3$	0	0	1	1	0	0	0	0	0

4 Model Construction and Statistical Analysis Method

4.1 General Form of Huber Robust Regression Model

The Huber robust regression model was constructed to quantify the independent marginal effect of each factor on catalytic activity (T_{50}). The general form of the model is defined as:

$$y = \beta_0 + \sum_{i=1}^{p} \beta_i x_i + \sum_{i<j} \beta_{ij} x_i x_j + \varepsilon \#(4-1)$$

Where:

(i) y: Response variable, the T_{50} value of the perovskite catalyst (lower T50 represents higher catalytic activity);

(ii) p: Total number of dummy variables ($p = 9$ in this study);

(iii) β_0: Intercept term, representing the expected T_{50} value when all factors are at their reference levels (A_1, B_1, C_1, D_1, i.e., Fe-based, La single component, unsupported, sol-gel method). When all $x_i = 0$, $y = \beta_0 + \varepsilon$, thus $\beta_0 = E[y|A_1, B_1, C_1, D_1]$;

(iv) $\beta_i x_i$: Main effect term. β_i is the main effect coefficient, representing the difference in T_{50} between the corresponding level and the reference level. A negative β_i indicates that the level can reduce T_{50} (enhance catalytic activity) compared with the reference level, while a positive β_i indicates the opposite;

(v) ε: Random error term, satisfying $E[\varepsilon] = 0$, representing experimental error and uncontrollable random factors.

4.2 Response Variable Distribution Analysis and Normality Test

The distribution characteristics of the response variable (T_{50}) are the core basis for model selection. The number of samples in each subgroup is summarized in Table S4. Descriptive statistics and Shapiro-Wilk normality test were performed on the T_{50} data, and the results are shown in Table S5.

The skewness coefficient was 0.7109 (>0), indicating a significant right-skewed distribution of the data, with the mean value higher than the median, which confirmed the existence of high-value outliers. The Shapiro-Wilk normality test yielded a P-value of 0.0251, which is less than the significance level of 0.05. Thus, the null hypothesis that "the data follows a normal distribution" was rejected at the 5% significance level.

Table S4 Subgroup sample distribution of categorical descriptors in the restricted dataset.

Descriptor	Subgroup	Number of
B-site	Fe-based	8
	Mn-based	14

	Co-based	7
	Dual-metal-	9
A-site	La-based	24
	La-AEM	9
	La-Ce	3
	Ba-based	2
Support	Unsupported	31
	Supported	7
Synthesis	Sol-gel	22
	Co-	5
	Other methods	11

Table S5 Descriptive statistics and normality test results of T_{50}.

Sample Size	Mean (℃)	Median (℃)	Std. Dev.	Skewness	Kurtosis	Shapiro-Wilk W	Shapiro-Wilk P
38	191.29	173.50	76.77	0.7109	-0.3498	0.9331	0.0251

The distribution characteristics of T_{50} are visually presented in Fig. S1. The histogram (Fig. S1(a)) shows that the actual distribution of T50 deviates significantly from the fitted normal distribution curve, with an obvious long tail on the high value side, consistent with the right-skewed characteristic revealed by the skewness coefficient. The boxplot (Fig. S1(b)) further quantifies the dispersion of the data: the interquartile range (IQR) of T_{50} reaches 120.00 ℃, and the data presents a wide distribution range in the high-value interval, which directly confirms the non-normality and strong dispersion the response variable.

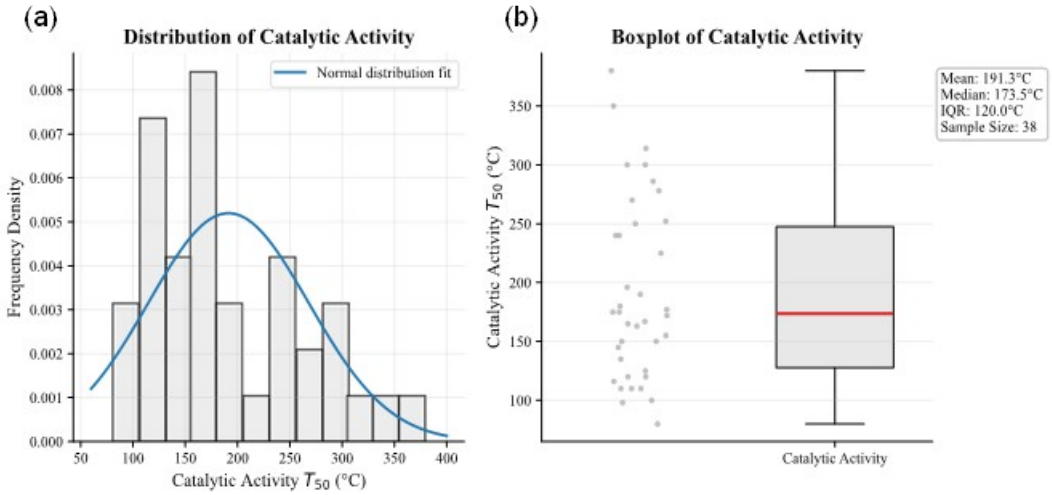
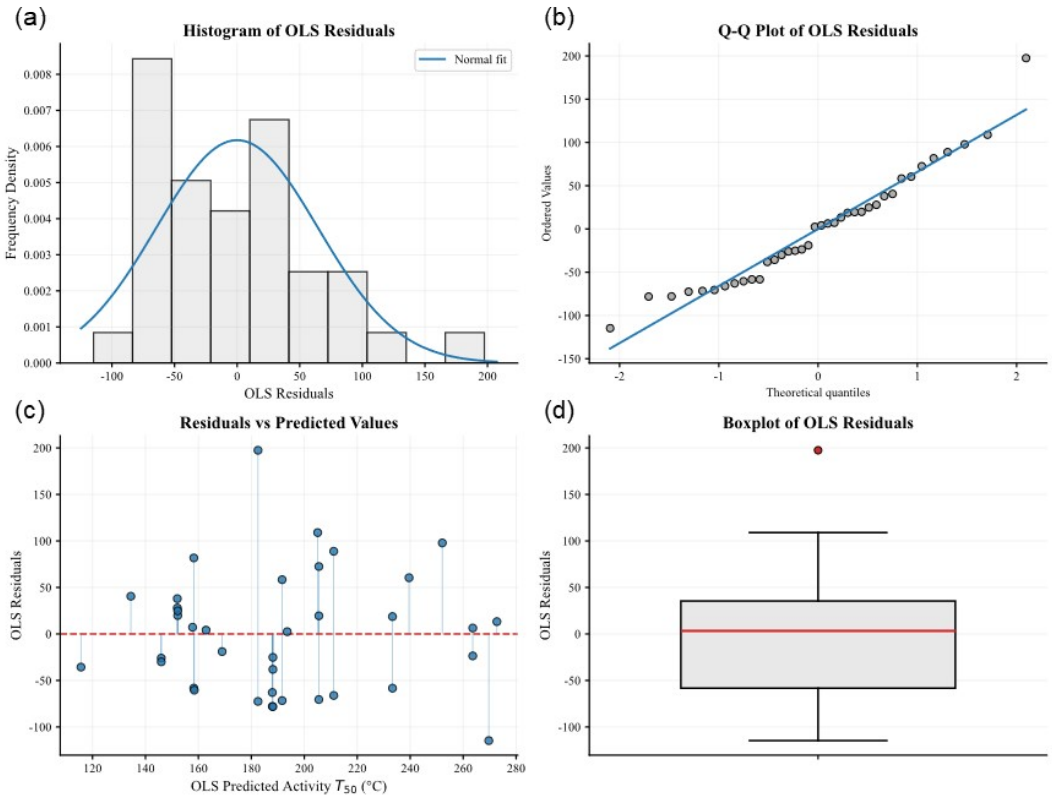


Fig. S1 Distribution characteristic of catalytic activity T_{50}. (a) Histogram of T_{50} with normal distribution fitting curve; (b) Boxplot of T_{50} with marked mean, median and IQR.

The non-normal marginal distribution of T_{50} does not by itself invalidate OLS regression. Therefore, OLS was first used as a baseline model, and residual diagnostics



were further performed to examine normality, heteroscedasticity and influential observations. The residual diagnostics showed deviations from ideal OLS assumptions, including non-normal residuals, heteroscedasticity and potential outliers. Therefore, Huber robust regression was adopted as the primary linear modeling approach. The use of OLS on this non-normal dataset would lead to biased parameter estimation, distorted significance test results, and unreliable confidence intervals. Residual diagnostic analysis of the OLS model was further performed to verify its inapplicability to this dataset, with the results shown in Fig. S2.

Fig. S2 Residual diagnostic analysis of the OLS regression model. (a) Frequency histogram of OLS residuals with normal fitting curve; (b) Q-Q plot of OLS residuals; (c) Scatter plot of residuals versus OLS predicted T50 values; (d) Boxplot of OLS residuals with extreme outlier distribution.

As shown in Fig. S2(a), the residuals of the OLS model present an obvious bimodal distribution, which deviates significantly from the fitted normal distribution curve. The Q-Q plot (Fig. S2(b)) further confirms the non-normality of residuals: most sample points deviate from the theoretical normal quantile diagonal, which cannot satisfy the normality assumption of OLS regression. The scatter plot of residuals versus predicted values (Fig. S2(c)) shows a clear funnel-shaped distribution: the dispersion of residuals expands significantly with the increase of predicted T_{50}, indicating severe heteroscedasticity of the OLS model, which violates the homoscedasticity assumption of classical linear regression. In addition, the boxplot (Fig. S2(d)) shows that the OLS residuals contain multiple extreme outliers, which will seriously bias the parameter estimation results of OLS.

Therefore, Huber robust regression was adopted in this study, which is insensitive to non-normal distribution, heteroscedasticity and outliers. The objective function of Huber regression is as follows:

$$\min_{\beta} \sum_{i=1}^{n} \rho\left(y_i - x_i^T \beta\right) \#(4-2)$$

Where the Huber loss function $\rho(r)$ is defined as:

$$\rho(r) = \begin{cases} \frac{1}{2} r^2, & |r| \le \delta \\ \delta\left(|r| - \frac{1}{2}\delta\right), & |r| > \delta \end{cases} \#(4-3)$$

In the formula, r is the model residual, and the tuning parameter δ is set to 1.345, which ensures 95% asymptotic estimation efficiency of OLS under normal distribution. The Huber loss function adopts squared loss for small residuals to ensure estimation efficiency, and switches to the linear loss for large residuals to reduce the weight of

outliers, thus achieving robust parameter estimation for non-normal and small-sample datasets.

4.3 Non-parametric Validation Methods

4.3.1 Permutation Test for Significance Validation

A nonparametric permutation test was performed to validate the statistical significance of each main effect, avoiding false positive result caused by non-normal data distribution. The core principle is: under the null hypothesis that "the factor has no significant effect on T_{50}", the feature sequence of the target factor is randomly permuted to destroy its correspondence with the response variable. The empirical null distribution of the test statistic is constructed via 1000 repeated permutations, and the permutation P-value is calculated by comparing the original observed statistic with the empirical null distribution. The factor was considered statistically significant when $P<0.05$. It should be noted that the permutation test and bootstrap confidence intervals evaluate different aspects of statistical robustness. The permutation test assesses whether the observed effect is unlikely under random label permutation, whereas bootstrap confidence intervals reflect the uncertainty of coefficient estimation under resampling. Therefore, the Mn-based and Co-based levels should be interpreted as showing a consistent activity-enhancing trend, but with higher coefficient uncertainty than the dual-metal-based level.

4.3.2 Bootstrap Uncertainty Quantification

Bootstrap resampling (1000 replicates with replacement) was used to quantify the uncertainty of model parameter estimation. The 95% confidence interval (CI) of each

coefficient was calculated via the quantile method. If the 95% CI of a coefficient does not contain 0, the effect of the corresponding factor is considered statistically robust at the 95% confidence level.

4.3.3 Random Forest Permutation Importance Analysis

A random forest regression model was constructed for cross-validation of factor importance ranking, which is a nonparametric ensemble learning method with strong adaptability to non-normal data and nonlinear relationships. The permutation importance index was used to quantify the contribution of each feature to catalytic activity, which is defined as:

$$Importance_j = \frac{1}{N}\sum_{i=1}^{N}\left(L\left(y_i,\hat{y}_i\right) - L\left(y_i,\hat{y}_i^{(j)}\right)\right)\#(4-4)$$

Where:

(i) $Importance_j$: Importance score of the j-th feature;

(ii) N: Number of samples;

(iii) $L(\cdot)$: Loss function (mean square error was adopted in this study);

(iv) $\hat{y}_i$: Predicted T50 value of the original sample;

(v) $\hat{y}_i^{(j)}$: Predicted T50 value after permutation of the j-th feature.

A larger $Importance_j$ value indicates a stronger regulatory effect of the feature on catalytic activity. The model hyperparameters were optimized via 5-fold cross validation to avoid overfitting.

4.4 Calculation Results

4.4.1 Huber Robust Regression Coefficients and Permutation Test Results

The Huber robust regression coefficients and permutation test results for each dummy variable are summarized in Table S6.

Table S6 Huber robust regression coefficients and permutation test results.

Feature	Corresponding Level	Regression Coefficient (°C)	Effect Direction	Permutation P-value	Significance (P<0.05)
x_{A_2}	Mn-based	-110.52	↑	0.0020	√
x_{A_3}	Co-based	-88.58	↑	0.0180	√
x_{A_4}	Dual-metal-based	-121.58	↑	0.0020	√
x_{B_2}	La-alkaline earth doping	13.34	↓	0.7313	
x_{B_3}	La-Ce co-doping	0.52	↓	0.9500	
x_{B_4}	Ba-based	32.13	↓	0.5754	
x_{C_2}	Supported	-9.20	↑	0.8122	
x_{D_2}	Coprecipitation method	9.59	↓	0.8282	
x_{D_3}	Other methods	31.39	↓	0.3337	

The results showed that B-site metal type was the dominant factor affecting catalytic activity. Compared with the reference Fe-based level, Mn-based, Co-based, and dual-metal-based all significantly reduced T_{50} (enhanced activity), with all three dummy variables reaching statistical significance (P<0.05). No significant independent main effects were detected for A-site doping, support loading condition, and preparation method in the current dataset.

4.4.2 Bootstrap 95% Confidence Intervals

The Bootstrap 95% confidence intervals for model coefficients are shown in Table S7.

Table S7 Bootstrap 95% confidence intervals for model coefficients.

Feature	Coefficient Mean (°C)	Coefficient Std. Dev. (°C)	95% Confidence Interval (°C)
x_{A_2}	-90.7191	54.1882	[-185.81, 15.02]
x_{A_3}	-78.2950	48.2154	[-165.41, 32.75]
x_{A_4}	-118.8048	49.1621	[-203.79, -0.08]
x_{B_2}	23.5406	54.3141	[-76.00, 145.01]
x_{B_3}	-1.4402	60.1449	[-132.05, 117.18]
x_{B_4}	30.0817	56.8723	[-98.51, 141.46]
x_{C_2}	-12.0526	40.3505	[-94.90, 64.13]
x_{D_2}	-0.4596	67.5497	[-156.38, 97.78]
x_{D_3}	22.9013	51.5336	[-84.16, 107.99]

Only the 95% confidence interval of x_{A_4} (dual-metal-based) was completely located in the negative half-axis and did not cross 0, confirming the robust T_{50} reduction effect of dual-metal-based B-site doping. The coefficient means of x_{A_2} (Mn-based) and x_{A_3} (Co-based) were also negative, showing a consistent trend of activity enhancement, while their confidence intervals crossed 0, indicating partial estimation uncertainty under the small sample size. All other factors had wide confidence intervals crossing 0, with no stable independent effects detected.

4.4.3 Random Forest Feature Importance Ranking

The cumulative permutation importance and overall priority ranking of each categorical factor are shown in Table S8, and the importance ranking of individual dummy variables is shown in Table S9.

Table S8 Cumulative importance and priority ranking of categorical factors.

Categorical Factor	Cumulative Importance	Overall Priority
B-site Metal Type	0.7874	1
A-site Doping Type	0.1309	2
Preparation Method	0.0750	3
Support Loading Condition	0.0623	4

Table S9 Permutation importance ranking of individual dummy variables.

Feature	Corresponding Level	Mean Permutation Importance
x_{A_4}	Dual-metal-based	0.3742
x_{A_2}	Mn-based	0.2515
x_{A_3}	Co-based	0.1616
x_{B_2}	La-alkaline earth doping	0.1048
x_{D_2}	Co-precipitation method	0.0749
x_{C_2}	Supported	0.0623
x_{D_3}	Other methods	0.0529
x_{B_3}	La-Ce co-doping	0.0259
x_{B_4}	Ba-based	0.0037