Article | Open Access

Energy Materials

Bi *et al. Energy Mater.* **2026**, *6*, 600121 **DOI:**10.20517/energymater.2026.216

Machine learning-driven optimization of plasma-coupled photocatalytic CO_2 reduction system for syngas production

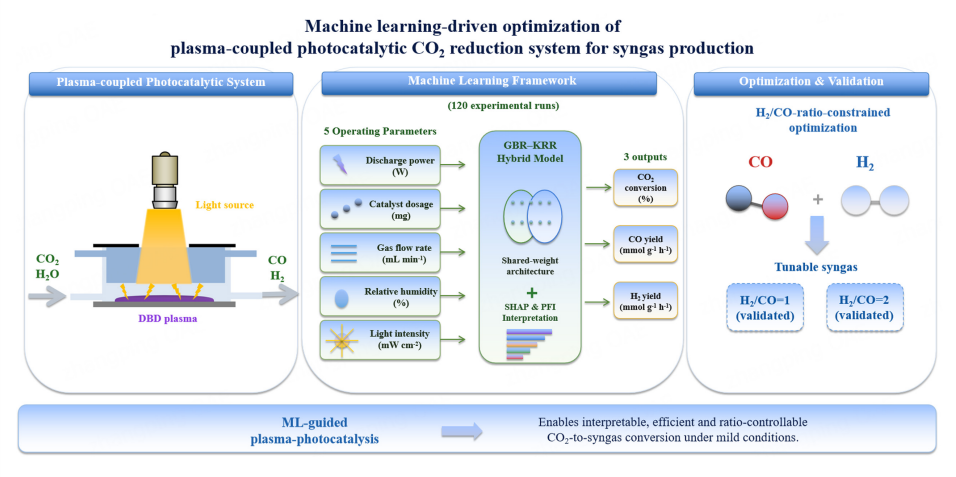
Runze Bi[1,2,#], Tingting Gao[3,#], Yanan Wang[1,4], Dong Wang[5], Huiwen Zhu[6], Ziyun Xi[1,2], Hongling Qin[1,2], Huanyi Wang[7], Ziqi Geng[2], Yijie Yao[2], Yiyong Guan[2], Honglei Zhang[1,2,*]

Keywords:
Plasma-photocatalysis, CO_2 conversion, syngas, machine learning

Citation: Bi, R.; Gao, T.; Wang, Y.; Wang, D.; Zhu, H.; Xi, Z.; Qin, H.; Wang, H.; Geng, Z.; Yao, Y.; Guan, Y.; Zhang, H. Machine learning-driven optimization of plasma-coupled photocatalytic CO_2 reduction system for syngas production. *Energy Mater.* **2026**, *6*, 600121. https://dx.doi.org/10.20517/energymater.2026.216

Received: 9 Jul 2026
First Decision: 30 Jul 2026
Revised: 27 Aug 2026
Accepted: 17 Sep 2026
Published: 20 Sep 2026

Academic Editor:
Ho Won Jang
Copy Editor:
Ping Zhang
Production Editor:
Ping Zhang



Abstract

Dielectric barrier discharge (DBD) plasma-coupled photocatalysis offers a promising route for CO_2-to-syngas conversion, but rational optimization remains challenging because operating variables are strongly coupled. Here, we develop a small-sample machine-learning framework to predict and optimize a plasma-coupled photocatalytic CO_2 conversion system using a Cu-Pd/TiO_2 photocatalyst. Using 120 experimental runs, five operating parameters, including discharge power, catalyst dosage, gas flow rate, relative humidity, and light intensity, were evaluated against CO_2 conversion, CO yield, and H_2 yield. Among the evaluated models, a shared-weight Gradient Boosting Regressor (GBR)-Kernel Ridge Regression (KRR) hybrid model achieved the great predictive performance, with test-set R^2 values of 0.947, 0.950, and 0.954 for CO_2 conversion, CO yield, and H_2

[1]Nottingham Ningbo China Beacons of Excellence Research and Innovation Institute, Ningbo 315100, Zhejiang, China.
[2]Department of Chemical and Environmental Engineering, The University of Nottingham Ningbo China, Ningbo 315100, Zhejiang, China.
[3]Beilun Branch of Ningbo Municipal Bureau of Ecology and Environment, Ningbo 315800, Zhejiang, China.
[4]Department of Mechanical, Materials and Manufacturing Engineering, The University of Nottingham Ningbo China, Ningbo 315100, Zhejiang, China.
[5]Beijing Perfectlight Technology Co., Ltd., Beijing 100080, China.
[6]College of Digital Technology and Engineering, Ningbo University of Finance & Economics, Ningbo 315175, Zhejiang, China.
[7]School of Materials Science and Chemical Engineering, Ningbo University, Ningbo 315211, Zhejiang, China
[#]These authors contributed equally to this work.

***Correspondence to:** Prof. Honglei Zhang, Nottingham Ningbo China Beacons of Excellence Research and Innovation Institute, Ningbo 315100, Zhejiang, China; Department of Chemical and Environmental Engineering, The University of Nottingham Ningbo China, Ningbo

315100, Zhejiang, China. E-mail: honglei-zhang@nottingham.edu.cn

yield, respectively. Permutation importance and SHapley Additive exPlanations (SHAP) analyses identified relative humidity, light intensity, and discharge power as the most influential variables and revealed distinct model-predicted trends across the three outputs. The trained model was further used for constrained optimization under target H_2/CO ratios of 1 and 2, followed by experimental validation of the selected operating conditions. Overall, this work establishes a data-driven strategy for interpreting nonlinear plasma-photocatalytic CO_2 conversion and provides practical guidance for syngas-ratio regulation under experimentally relevant conditions.

INTRODUCTION

Rising CO_2 emissions have intensified concerns about global warming and environmental sustainability in recent years[1]. Efficient CO_2 conversion and utilization are therefore essential for closing the carbon cycle and developing a low-carbon energy system[2,3]. The conversion of CO_2 into value-added fuels and chemicals not only mitigates CO_2 emissions but also provides alternative resources for the chemical and energy industries[4-7]. Among the possible products, syngas, a mixture of CO and H_2, is particularly attractive because it serves as an important intermediate for Fischer-Tropsch synthesis, methanol production, and other downstream chemical processes[8,9]. The H_2/CO ratio is also a key determinant of its suitability for different applications. Therefore, the efficient conversion of CO_2 and H_2O into syngas requires not only high reaction efficiency, but also effective control over product composition. However, conventional syngas-production routes, such as coal gasification and methane reforming, generally require harsh reaction conditions and are associated with substantial energy consumption and carbon emissions[10]. Developing mild, tunable, and low-carbon routes for CO_2-to-syngas conversion is therefore highly desirable[11,12].

Photocatalytic CO_2 reduction provides a promising route for solar-driven carbon utilization, but its practical performance is limited by the difficult activation of CO_2, rapid recombination of photogenerated charge carriers, and insufficient control over product selectivity[6,13-15]. To overcome these limitations, previous studies have demonstrated the potential of combining plasma with photocatalysts. Non-thermal plasma (NTP) generates energetic electrons, excited molecules, radicals, and other reactive species under relatively mild conditions, thereby providing an effective pathway for activating thermodynamically stable CO_2 molecules[16,17]. It can therefore be coupled with photocatalytic systems to enhance reaction efficiency[18]. For example, Huang *et al.* coupled a Cs_2SnCl_6 photocatalyst with a dielectric barrier discharge (DBD) plasma system and obtained a CO production rate of 294.47 μmol min^{-1}, which was approximately 6.5% higher than the sum of those achieved by plasma alone (276.38 μmol min^{-1}) and photocatalysis alone (0.0013 μmol min^{-1}). The introduction of H_2O further increased the CO_2 conversion by 50.6%, which was attributed to enhanced charge transfer resulting from the increased electrical conductivity of the photocatalyst surface during plasma discharge[19]. More recently, Feng *et al.* coupled a $CsPbBr_3$@ReS_2 heterojunction photocatalyst with a DBD plasma reactor and achieved a CO_2 conversion of 35.60% and an energy efficiency of 13.10%. They attributed the enhanced performance to improved optical absorption and interfacial charge migration, together with prolonged microdischarge duration and improved discharge uniformity[20]. These studies establish the feasibility of combining semiconductor photocatalysts with plasma. Nevertheless, these studies mainly focused on overall CO_2 conversion, CO production, or energy efficiency, while the simultaneous formation and systematic regulation of CO and H_2 from CO_2 and H_2O remain less explored.

A fully coupled plasma-photocatalytic system integrates plasma activation, external light irradiation, and photocatalytic surface reactions within the same reaction environment. This configuration may enable more flexible control of syngas formation, but it also introduces strong interactions among the operating variables. Discharge power, catalyst dosage, gas flow rate, relative humidity and light intensity may produce nonlinear

and output-dependent effects on CO_2 conversion, CO formation and H_2 evolution[21-23]. Conditions that favor CO_2 conversion or CO production may therefore not yield the desired H_2/CO ratio, making conventional one-factor-at-a-time optimization inadequate for this multivariable and multi-objective system. Therefore, machine learning provides a possible approach for describing such nonlinear relationships and identifying operating conditions within complex reaction spaces[24,25]. For instance, Shen *et al.* developed a Genetic Algorithm-Backpropagation Artificial Neural Network (GA-BPANN) model to predict CO_2 conversion in a DBD plasma system coupled with a Cs_2TeCl_6 photocatalyst, achieving R^2 values above 0.96 on both training and test sets and enabling quantitative analysis of process parameter effects[23]. Recently, Luo *et al.* applied machine learning to a DBD plasma system assisted by photocatalysts in CO_2 conversion using different halide-perovskite photocatalysts[26]. Process variables and catalyst material descriptors were used to establish separate predictive models for CO_2 conversion and energy efficiency. These studies demonstrate the value of machine learning for performance prediction and process optimization. Nevertheless, existing applications have generally focused on a single reaction metric or on separate performance targets. They have rarely addressed the simultaneous prediction of CO_2 conversion, CO yield, H_2 yield and H_2/CO ratio within a fully coupled plasma-photocatalytic CO_2/H_2O system. Moreover, optimization under a prescribed H_2/CO ratio followed by experimental validation remains insufficiently developed. Three specific knowledge gaps therefore motivate the present study. First, fully coupled plasma-photocatalytic systems for controllable syngas production from CO_2 and H_2O have not yet been systematically investigated. Second, the nonlinear and potentially competing effects of multiple operating parameters on CO_2 conversion, CO formation, H_2 evolution, and the resulting H_2/CO ratio remain unclear. Third, a data-efficient and interpretable machine-learning framework for multi-output prediction and H_2/CO-ratio-constrained optimization, followed by experimental validation, has not been established for this type of system.

Herein, to address these knowledge gaps, we developed a DBD plasma-coupled photocatalytic CO_2 reduction (P^2CR) system using Cu-Pd/TiO_2 as a representative photocatalyst. The P^2CR system integrates plasma activation, light irradiation, and photocatalytic surface reactions within the same reaction environment for the simultaneous conversion of CO_2 and H_2O into syngas. A dataset comprising 120 experimental runs was constructed using discharge power, catalyst dosage, gas flow rate, relative humidity, and light intensity as input variables, while CO_2 conversion, CO yield, and H_2 yield were selected as output targets. After comparing multiple machine-learning algorithms, a shared-weight hybrid model based on Kernel Ridge Regression and Gradient Boosting Regressor was developed to capture the nonlinear relationships between the operating parameters and reaction outputs. Permutation importance and SHapley Additive exPlanations (SHAP) analyses were further employed to identify the dominant variables and interpret their effects on CO_2 reduction, CO formation, and H_2 evolution. Finally, the trained model was used to screen operating conditions under prescribed H_2/CO ratios of 1 and 2, and the model-recommended conditions were experimentally validated. This work establishes a data-driven strategy for understanding and optimizing nonlinear plasma-photocatalytic CO_2 conversion and provides practical guidance for controllable syngas production.

EXPERIMENTAL

Experimental equipment, processes and photocatalyst preparation

Detailed information on the experimental equipment, experimental procedures and photocatalyst preparation is provided in the Supplementary Materials.

The procedures of the machine learning

The P^2CR system, as illustrated in Supplementary Figure 1, is influenced by many factors, including discharge power, catalyst dosage, gas flow rate, humidity, and light intensity. These five variables were selected because they were experimentally varied in all 120 runs and represent the main operation-level

descriptors of the present plasma-coupled photocatalytic reactor. Discharge power controls the plasma energy input, catalyst dosage determines the amount of available catalytic surface, gas flow rate affects reactant throughput and residence time, relative humidity regulates water participation and the plasma reaction atmosphere, and light intensity governs photocatalyst excitation. These variables were therefore selected as the most relevant operating parameters for model construction and optimization within the present fixed catalyst-reactor system. Other potentially influential factors, including catalyst composition, reactor geometry, electrode gap, plasma frequency, and the CO_2/Ar feed ratio, were kept constant in the present study and were not included as input variables. This design was adopted because the objective was to optimize the operating conditions of a fixed Cu-Pd/TiO_2 plasma-coupled photocatalytic system rather than to compare different catalysts or reactor configurations. In addition, the dataset contained 120 experimental runs. Introducing additional material-level and reactor-level variables would have increased the dimensionality of the problem and reduced the interpretability and statistical robustness of the small-sample machine-learning analysis.

Figure 1 illustrates the workflow of model construction, training, optimization, and evaluation in this study. The machine-learning workflow was designed to avoid information leakage during model development. The dataset consisted of 120 experimental samples and was divided by the Sample set Partitioning based on joint X-Y distances (SPXY) algorithm into a training set (n = 96) and an external test set (n = 24). The training set was used for model development and parameter optimization, whereas the external test set was used for final performance evaluation and subsequent permutation feature importance (PFI) analysis. Five experimentally measured operating variables, namely discharge power, catalyst dosage, gas flow rate, relative humidity, and light intensity, were used as the primary input features. To capture nonlinear interactions, nine engineered features were additionally constructed: power/flow, dosage/flow, light intensity/flow, humidity × light intensity, power × light intensity, humidity × flow, 1/flow, power/dosage, and dosage × light intensity. The hybrid model used five experimentally controllable variables together with nine engineered descriptors for nonlinear feature representation. For the Kernel Ridge Regression (KRR) model, the input features were standardized using StandardScaler within the modeling pipeline. The tree-based Gradient Boosting Regressor (GBR) model was fitted directly to the raw and engineered input features without feature standardization. No target transformation or target normalization was applied during model fitting; all predictions and evaluation metrics were calculated in the original physical units. During SPXY splitting, the input and output variables were temporarily standardized only for calculating distances in the combined X-Y space. This temporary scaling was not used to transform the model targets. In addition, mean squared error (MSE) was used during hybrid-weight selection. Five repeated five-fold cross-validation was used to screen the candidate base learners. Hyperparameter optimization was performed using four-fold inner cross-validation within the training set only. The optimized GBR parameters were $n_{estimators}$ = 280, learning rate = 0.06, max depth = 2, subsample = 0.85, and min samples leaf = 1. The optimized KRR parameters were alpha = 0.06 and gamma = 0.10 with an Radial Basis Function (RBF) kernel. After fixing the base-model hyperparameters, the shared hybrid weight was optimized using three-fold out-of-fold predictions generated only from the training set. The weight-search step size was 0.01. The random seeds were fixed at 42 for hyperparameter cross-validation and 2026 for the inner out-of-fold hybrid-weight optimization and GBR training. The repeated cross-validation seeds were also fixed in advance. All analyses were performed in Python 3.11.15 using NumPy 2.4.6, pandas 3.0.5, SciPy 1.17.1, scikit-learn 1.9.0, and joblib 1.5.3.

Description of the hybrid model

A single machine learning model may be sensitive to local data structures, noise, or train-test splitting[27-29]. To address these challenges, we developed a hybrid model with shared weights by combining GBR and KRR.

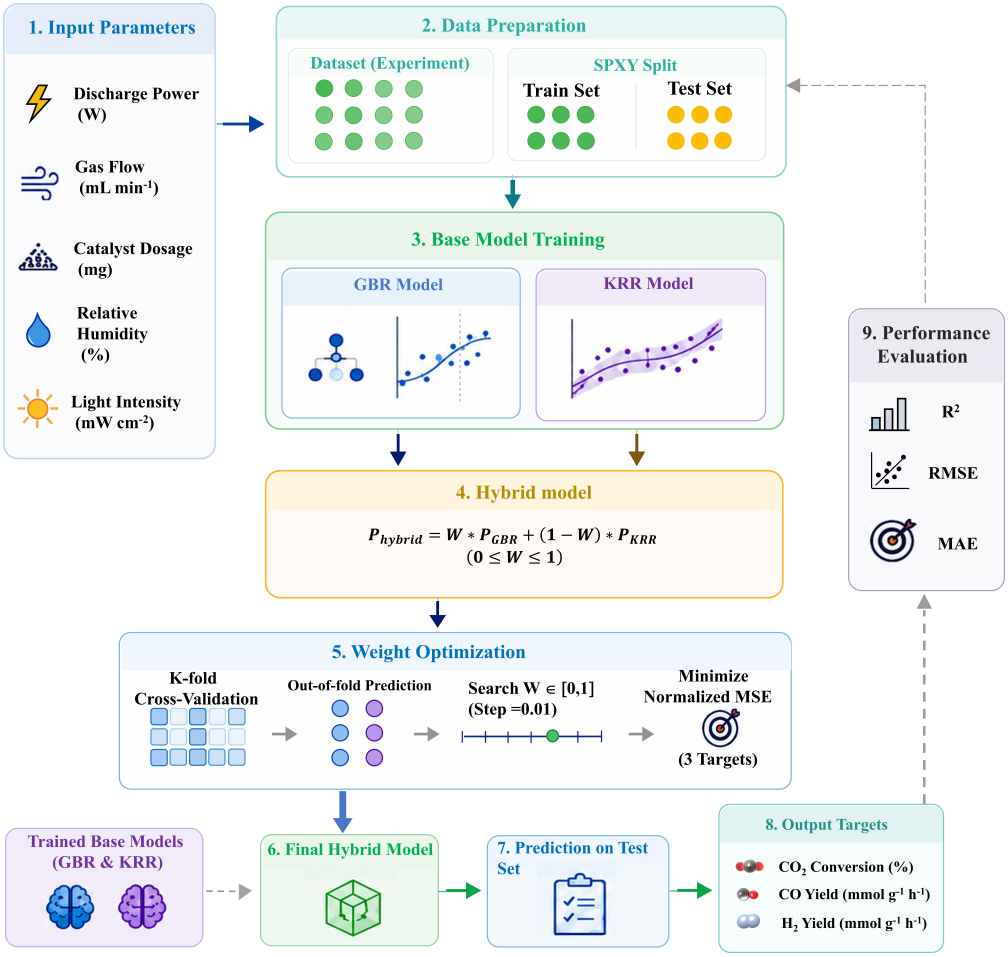


Figure 1. Process scheme of the optimization, evaluation and prediction framework for the hybrid machine-learning model. RMSE: Root mean square error; MAE: mean absolute error; MSE: mean squared error; SPXY: sample set partitioning based on joint X-Y distances; GBR: gradient boosting regressor; KRR: kernel ridge regression.

GBR is a tree-based ensemble method capable of capturing nonlinear relationships, local threshold effects, and interactions between variables[30]. In contrast, KRR performs smooth kernel-based regression and is suitable for modeling continuous nonlinear responses in small-sample datasets[31]. These two models have different learning mechanisms: GBR focuses on local nonlinear partitioning of the feature space, whereas KRR describes global smooth nonlinear trends based on kernel similarity. Their complementary characteristics motivated the construction of the hybrid model. The hybrid prediction was defined as:

$$P_{hybrid} = W * P_{GBR} + (1 - W) * P_{KRR} \ (0 \le W \le 1) \tag{1}$$

where P_{GBR} and P_{KRR} are the predictions of GBR and KRR, respectively, and W represents the weight assigned to GBR. Within the SPXY training set, cross-validation was used to generate out-of-fold predictions from GBR and KRR[32]. The hybrid weight W was then searched from 0 to 1 with a step size of 0.01. For each candidate weight, the hybrid prediction was calculated and evaluated by the mean squared error[33]:

$$MSE = \frac{1}{n} \sum_{j=1}^{n} \left(y_j^{pred} - y_j^{true} \right)^2 \tag{2}$$

where y_j^{pred} is the predicted value of the model; y_j^{true} is the true value.

After weight optimization, the optimal weights for the two models were determined to be 0.64 for GBR and 0.36 for KRR. The final hybrid model is therefore expressed as:

$$P_{hybrid} = 0.64 * P_{GBR} + 0.36 * P_{KRR} \tag{3}$$

In this study, the performance of various base machine learning models in predicting the catalytic reaction properties was evaluated. R^2 was used as the evaluation metric to measure the performance of each model and to make comparisons. Performance was also evaluated by mean absolute error (MAE) and root mean square error (RMSE), as determined using[33]:

$$R^2 = 1 - \frac{\sum_{j=1}^{n}\left(y_j^{pred} - y_j^{true}\right)^2}{\sum_{j=1}^{n}\left(y_j^{true} - \bar{y}^{true}\right)^2} \tag{4}$$

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left|y_j^{pred} - y_j^{true}\right| \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j^{pred} - y_j^{true}\right)^2} \tag{6}$$

where y_j^{pred} is the predicted value of the model; y_j^{true} is the true value; $\bar{y}^{true}$ is the mean true value.

Importance analysis and SHAP analysis

To evaluate how each experimental parameter influenced the hybrid model's predictions, permutation feature importance was applied. We chose PFI because it is model-agnostic and does not depend on the internal workings of a specific algorithm, which makes it particularly suitable for the GBR-KRR hybrid model[34,35]. In addition, PFI measures the change in model performance when a given feature is perturbed, providing a straightforward way to assess how sensitive the model is to each input variable. In this analysis, the trained model was first used to predict the external test set, and the original coefficient of determination, $R^2_{original}$, was recorded. Each input variable was then randomly permuted in turn while all other variables were kept unchanged. The model was used to make predictions on the permuted data, and the corresponding coefficient of determination, $R^2_{permuted}$, was calculated. The importance of each feature was expressed as the decrease in R^2:

$$R^2_{drop} = R^2_{original} - R^2_{permuted} \tag{7}$$

A larger R^2_{drop} indicates that permuting the corresponding feature causes a greater loss of model performance, suggesting that the model is more sensitive to that variable[34-36]. In this study, R^2_{drop} was calculated separately for CO_2 conversion, CO yield, and H_2 yield to compare the relative effects of 5 input parameters on each output. It is important to emphasize that PFI reflects the dependence of the trained model on each input variable, rather than establishing a direct causal relationship[37]. The results were therefore used to identify key operating parameters and to guide subsequent response-trend analysis and condition optimization.

SHAP analysis was further performed to gain a deeper understanding of feature contributions. Unlike PFI, which evaluates the overall importance of a variable, SHAP provides a sample-level decomposition of the prediction into contributions from individual features[38,39]. For a given output, a positive SHAP value indicates that the feature increases the predicted value relative to the baseline prediction, whereas a negative SHAP value indicates that it decreases the prediction. A larger absolute SHAP value corresponds to a stronger contribution to that specific prediction[38,39]. For the hybrid model developed in this study, SHAP analysis was used to interpret the prediction behavior of the weighted GBR-KRR model across the experimental domain. By comparing the SHAP distributions of different features for CO_2 conversion, CO

yield, and H_2 yield, the analysis helped determine whether key variables had different effects on CO_2 reduction and H_2 formation. This provided additional insights into the nonlinear responses and multi-objective competition in the P²CR system. For interpretability, PFI and SHAP were performed on the five original variables, while the nine engineered descriptors were used only internally in model construction. It is important to emphasize that permutation feature importance and SHAP quantify predictive associations within the experimental domain. These analyses were used to rank the input variables and interpret the behavior of the trained model. Mechanistic explanations based on these results are presented as plausible hypotheses guided by reactor physics and previous literature.

The details of the robustness-oriented strategy

To further assess the extrapolation risk of the generated candidate space, a convex-hull-based applicability-domain analysis was performed using the SPXY training set in the five-dimensional input space. For the feasible candidates satisfying the H_2/CO ratio constraints, the fractions of hull-included points were 50.49%, 49.78%, and 48.66% for H_2/CO = 1 under tolerance windows of ±0.05, ±0.10, and ±0.20, respectively, and 49.52%, 47.85%, and 48.59% for H_2/CO = 2. Importantly, the two final robust recommendations emphasized in this work, corresponding to H_2/CO = 1 (±0.05) and H_2/CO = 2 (±0.05), were both located within the convex hull of the training set. In contrast, some conversion-oriented recommendations under broader tolerance windows were outside the hull, indicating a higher likelihood of extrapolation. After generating the 100,000 candidate conditions, a two-step constrained optimization procedure was applied. First, a hard feasibility filter was used to retain only candidate points whose predicted H_2/CO ratio satisfied the selected tolerance window around the target ratio. Second, within this feasible candidate set, the remaining candidates were ranked using a robustness-oriented score:

$$Robust\ score = Y_{CO_2} - 0.45\sigma_{CO_2} - 0.15\sigma_{H_2/CO} - 0.08D_{train} \quad (8)$$

where Y_{CO_2} is the predicted CO_2 conversion, σ_{CO_2} is the disagreement between the GBR and KRR member models for CO_2 conversion, $\sigma_{H_2/CO}$ is the disagreement in the predicted H_2/CO ratio, and D_{train} is the nearest standardized distance from the candidate point to the training set. The penalty coefficients were selected empirically to preserve CO_2 conversion as the primary optimization objective while discouraging candidates with larger prediction disagreement and greater extrapolation risk.

RESULTS AND DISCUSSION

Model training, evaluation and comparison

The Cu-Pd/TiO_2 photocatalyst was selected based on preliminary catalyst screening and literature-guided catalyst design. Before constructing the machine-learning dataset, several support materials, including TiO_2, Al_2O_3, CeO_2, and ZSM-5, were evaluated. As shown in Supplementary Table 1, among different supports, TiO_2 showed the best overall performance, possibly due to its suitable photoresponse characteristics and compatibility with the plasma reaction environment, which is in good agreement with the results of previous studies[40]. Therefore, TiO_2 was selected as the model support for subsequent catalyst design. The selection of Cu and Pd was based on their complementary catalytic roles. Cu-containing sites have been widely reported to facilitate CO_2 activation and CO formation, whereas Pd can influence interfacial electron redistribution and hydrogen-involved elementary reaction steps[41-43]. In our preliminary comparison results [Supplementary Table 1], Cu/TiO_2 exhibited stronger plasma-photocatalytic coupling and higher CO yield, while Pd/TiO_2 promoted H-related product formation. Therefore, Cu and Pd were co-loaded on TiO_2 to combine the CO-forming ability of Cu with the H-related reaction regulation of Pd and Cu-Pd/TiO_2 was used as a representative model catalyst to establish and validate the machine-learning-guided optimization framework. To assess the synergistic effect in the P²CR system, control experiments were conducted to illustrate the synergistic effect of this system, as shown in Supplementary Figure 2. The light-only experiment

produced only trace amounts of CO (0.013 mmol g^{-1} h^{-1}) and negligible CO_2 conversion, indicating that under the current conditions, the contribution of photocatalysis to CO_2 reduction is negligible. In contrast, the CO yield using plasma alone was 13.93 mmol g^{-1} h^{-1} with 1.4% CO_2 conversion, while the plasma-photocatalysis combined conditions significantly increased the CO yield to 28.52 mmol g^{-1} h^{-1} (2.86% CO_2 conversion), far exceeding the yield achieved by plasma alone. These results clearly demonstrate that light irradiation can effectively promote the plasma-assisted CO_2 reduction process, thereby significantly increasing CO production, demonstrating the coupling and synergistic effect of this P^2CR system. Supplementary Figures 3-8 illustrate the detailed characterizations of Cu-Pd/TiO_2 photocatalyst. The results of Gas Chromatography (GC) chromatograms in Supplementary Figure 9 show that CO was identified as the only detectable carbon-containing product without any other carbon-containing products in this system. Consistent with this observation, the carbon balance remained between 99.70% and 99.93% across 20 independent experimental runs [Supplementary Table 2], indicating good carbon closure and supporting the reliability of the product quantification. In addition, the surface temperatures of the reactor during reaction were measured by infrared thermal images to investigate the thermal effect generated by light and plasma. As shown in Supplementary Figure 10, both light and plasma caused an increase in temperature. To eliminate the influence caused by thermal effects, a heat-only control experiment under comparable external temperature conditions (temperature from 30 to 100 °C) was conducted. However, no CO, H_2 or other products were detected, indicating that thermal heating alone was insufficient to drive the observed CO_2 conversion and syngas formation in this system. Moreover, Cu-Pd/TiO_2 was further compared with other state-of-the-art photocatalysts for syngas production [Supplementary Table 3]. Most reported systems exhibited relatively modest gas productivity, with CO and H_2 yields generally reported at the μmol g^{-1} h^{-1} level. In addition, their H_2/CO ratios are typically distributed within a comparatively narrow range of 0.31:1-6.00:1, limiting the flexibility of syngas composition control. By comparison, the present Cu-Pd/TiO_2 achieved maximum observed CO and H_2 yields of 56.42 and 53.71 mmol g^{-1} h^{-1}, respectively, and enabled a wide range of H_2/CO ratios (0.022:1-53.71:1). These results indicate that the Cu-Pd/TiO_2 photocatalyst in the present system provides both high syngas productivity and substantially broader flexibility in H_2/CO-ratio regulation.

Based on 120 experimental runs in Supplementary Tables 4 and 5, the hybrid model was trained and tested. As shown in Figure 2A, the proposed hybrid model achieved the highest R^2, with a value of approximately 0.996. It slightly outperformed KRR and GBR, which achieved R^2 values of approximately 0.992 and 0.987, respectively. Compared with traditional models like K-Nearest Neighbors (KNN) (0.84) and Decision Tree regressors (0.89), the hybrid model showed larger absolute R^2 improvements of approximately 0.16 and 0.11, respectively. These results indicate that the hybrid architecture more effectively captures the nonlinear relationships embedded in the experimental dataset. Figure 2B-D further compare the predictive performance of different models for CO_2 conversion, CO yield, and H_2 yield on both training and test sets. For CO_2 conversion [Figure 2B], the hybrid model achieves a test-set R^2 of 0.947, compared with 0.921 for KRR and 0.940 for GBR, corresponding to increases of 0.026 and 0.007, respectively. It also achieves the lowest test-set RMSE (0.271) and a tied-lowest test-set MAE (0.221, equal to GBR). For CO yield [Figure 2C], the hybrid model reaches an R^2 of 0.950 on the test set, slightly outperforming KRR (0.945) and GBR (0.934), demonstrating improved stability across both training and testing domains. For H_2 yield [Figure 2D], the hybrid model achieves a test-set R^2 of 0.954, higher than KRR (0.928) and slightly higher than GBR (0.950). The hybrid model has a lower test-set MAE than GBR; however, its test-set RMSE (1.60) is slightly higher than that of GBR (1.50). These results indicate a balanced error profile rather than consistently lower errors across all evaluation metrics. Parity plots in Figure 3A-C further support these results, with the training and test data distributed around the fitted lines. Together, these analyses indicate that the hybrid model achieves

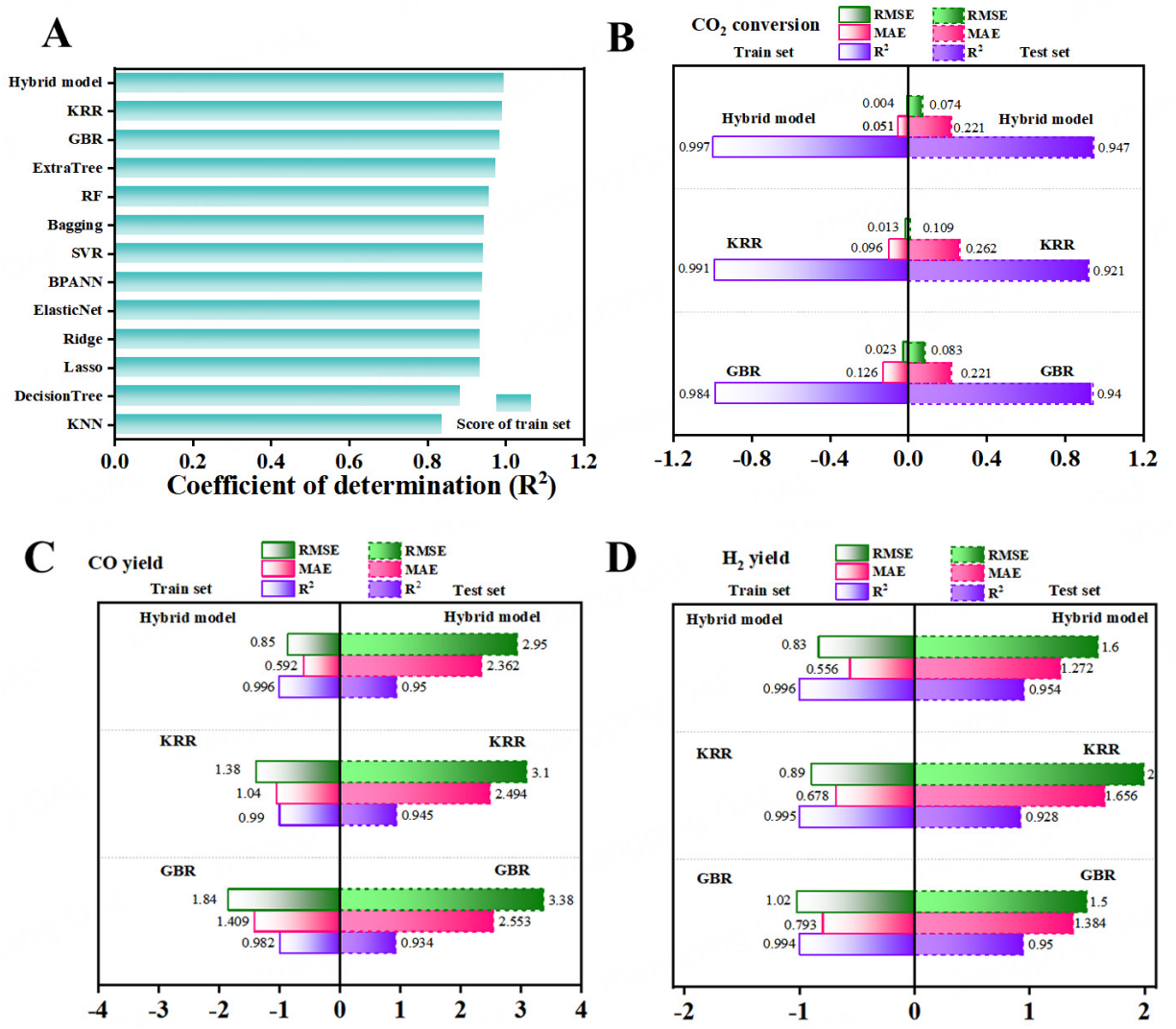


Figure 2. Performance comparison of the machine-learning models. (A) Overall comparison based on the training-set R^2; Performance of the hybrid model, KRR and GBR for (B) CO_2 conversion, (C) CO yield and (D) H_2 yield. RMSE: Root mean square error; MAE: mean absolute error; MSE: mean squared error; SPXY: sample set partitioning based on joint X-Y distances; GBR: gradient boosting regressor; KRR: kernel ridge regression; RF: random forest; SVR: support vector regression; BPANN: backpropagation artificial neural network; KNN: K-nearest neighbors.

a balance between training accuracy and test-set generalization, providing accurate and robust predictions for all three outputs. The model is therefore well suited to multi-objective prediction of the three reaction targets. To further assess the predictive performance of the hybrid model, 10 unseen experimental conditions were generated within the original experimental ranges and evaluated using the hybrid model. As shown in Supplementary Tables 6 and 7, the predicted values were in close agreement with the corresponding experimental results, further supporting the predictive capability of the hybrid model for this P²CR system. The correlation matrix of all variables is shown in Supplementary Figure 11. To assess whether the predictive performance of the models depended on catalyst-mass normalization, additional models were developed to predict the absolute CO and H_2 production rates in Supplementary Table 8.

Importance analysis of different parameters

To further assess the sensitivity of the model predictions to the operating variables, an R^2-drop permutation analysis was performed to quantify their relative importance, as shown in Figure 3D-F. Relative humidity was identified as the most influential variable across three outputs, followed by light intensity and discharge

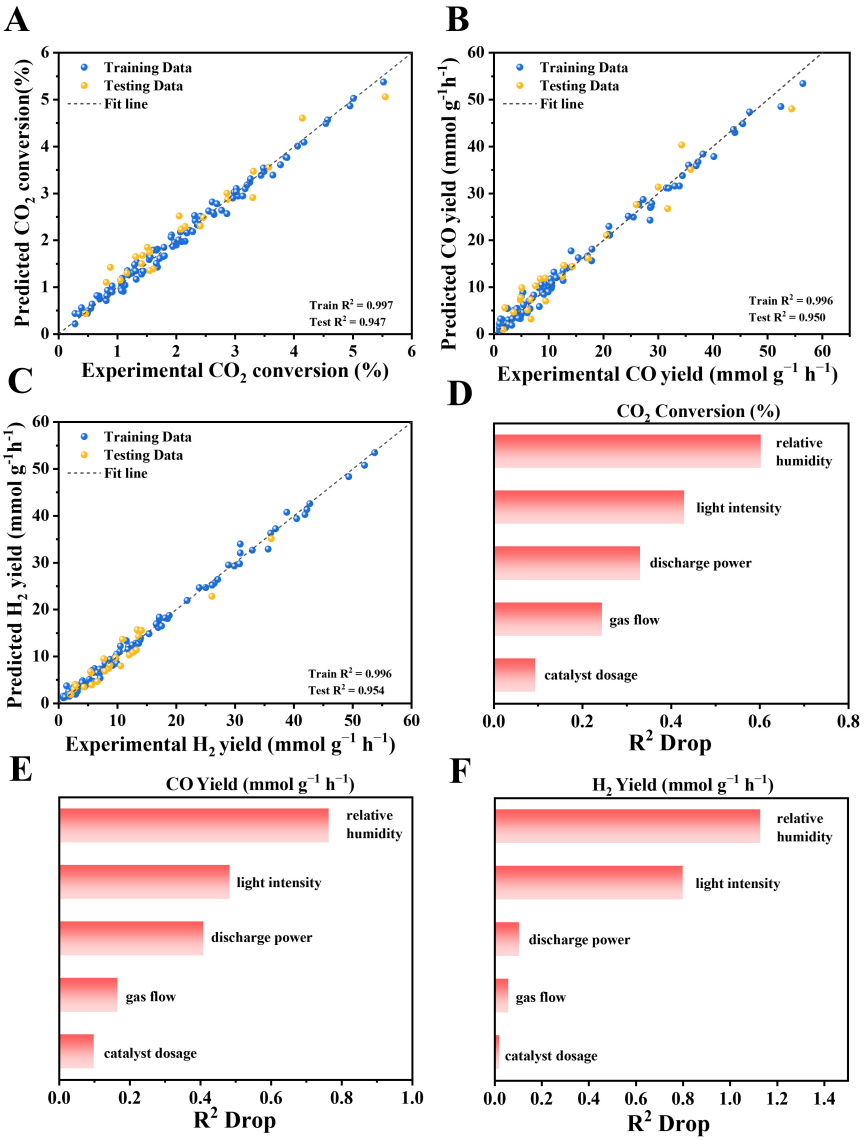


Figure 3. Predicted versus experimental values for (A) CO_2 conversion, (B) CO yield and (C) H_2 yield; Permutation importance of five input parameters for (D) CO_2 conversion, (E) CO yield and (F) H_2 yield.

power, whereas gas flow and catalyst dosage showed comparatively smaller contributions. These rankings describe the sensitivity of the trained model within the investigated domain and should not be interpreted as direct evidence of mechanistic importance. The importance of relative humidity may reflect the influence of water participation and the reaction environment on the predicted outputs. These effects could be associated with changes in surface reaction conditions and interfacial electron or proton transfer processes[44,45]. The importance of light intensity may reflect its association with photocatalyst excitation and the resulting changes in the model prediction[46]. Similarly, the importance of discharge power may reflect the influence of plasma energy input on the discharge environment and the associated reaction conditions[20,47]. Gas flow rate and catalyst dosage showed lower overall importance than relative humidity, light intensity, and discharge power. For gas flow rate, its influence was more pronounced for CO_2 conversion and CO yield than for H_2 yield. Figure 3D-F presents permutation-based model importance rather than direct causal response curves. Thus, the analysis identifies the variables to which the model predictions are most sensitive and provides guidance for operating-condition optimization.

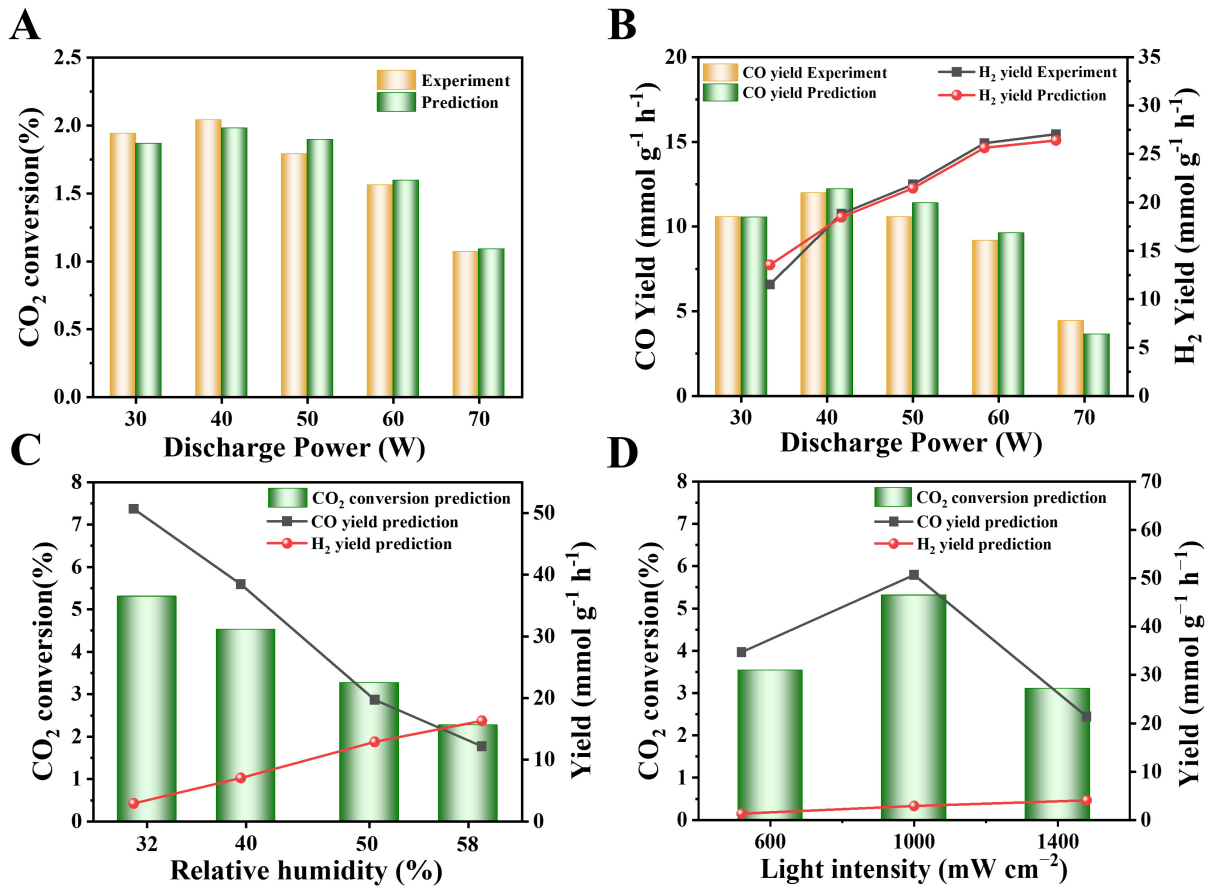


Figure 4. Comparison of model prediction with experimental data for the effect of discharge power on the plasma-photocatalytic CO_2 and H_2O conversion: (A) CO_2 conversion and (B) CO yield and H_2 yield (Relative humidity = 58%, light intensity = 1,000 mW cm^{-2}, gas flow = 30 mL min^{-1}, catalyst dosage = 40 mg); Predicted effects of (C) relative humidity (Discharge power = 50 W, light intensity = 1,000 mW cm^{-2}, gas flow = 20 mL min^{-1}, catalyst dosage = 60 mg) and (D) light intensity (Discharge power = 50 W, relative humidity = 32%, gas flow = 20 mL min^{-1}, catalyst dosage = 60 mg) on the plasma-photocatalytic CO_2 and H_2O conversion reaction performance.

Because the operating variables are strongly coupled, their combined effects on the reaction outputs must be considered when interpreting the P²CR system. We therefore used the hybrid machine-learning model to examine the response trends associated with the most influential variables. Based on the R^2-drop analysis, discharge power, relative humidity and light intensity were selected for detailed response-trend analysis. As shown in Figure 4A and B, the model predictions closely followed the experimental trends, consistent with the high R^2 values and low MAE and RMSE values. These results indicate that the hybrid model captured the nonlinear response behavior of the P²CR system. Under the conditions examined, CO_2 conversion and CO yield reached their maximum values at a discharge power of 40 W and decreased at higher power. One possible explanation is the non-uniform electric-field distribution within the square reactor. The reactor geometry may produce local electric-field enhancement or weakening near the corners and edges, thereby affecting the spatial distribution of the discharge and reactive species[48-51]. However, this explanation should be regarded as a hypothesis. Another possible explanation is that excessive discharge power may also induce local heating or alter the distribution of reactive species, thereby affecting the apparent catalytic performance[47]. In contrast, H_2 formation increased with increasing power. This trend may reflect the different responses of CO_2 reduction and hydrogen formation to increasing energy input. The greater thermodynamic stability of CO_2 and the potentially more accessible pathway for hydrogen formation possibly favor H_2 production under higher-power conditions[45,52]. These results suggest that moderate discharge power is a critical operating window in this system, reflecting the nonlinear role of power in regulating reactive species generation and reaction kinetics in the P²CR process.

Figure 4C further illustrates the effect of relative humidity on reaction outputs, while Supplementary Figure 12 compares the experimental and predicted values. Increasing the relative humidity from 32% to 58% decreased CO_2 conversion from 5.31% to 2.27% and CO yield from 50.67 to 12.19 mmol g^{-1} h^{-1}. In contrast, H_2 yield increased from 2.91 mmol g^{-1} h^{-1} to 16.32 mmol g^{-1} h^{-1}, which is consistent with the literature[45,53-55]. The increased water content may suppress the plasma discharge and thereby reduce the effective discharge density or modify electron-energy distribution[53]. A further possible explanation is that water-derived hydroxyl radicals could participate in secondary reactions involving CO intermediates and potentially promote CO reoxidation to CO_2[53,55]. This interpretation is proposed as a mechanistic hypothesis based on previously reported reaction pathways. Increased proton availability may also contribute to enhanced hydrogen formation, consistent with observations in other non-thermal plasma systems[20,54]. Overall, the model-predicted trends identify relative humidity as an important operating variable for regulating the balance between CO_2 reduction and H_2 formation, although the underlying elementary pathways remain to be established experimentally.

Figure 4D and Supplementary Figure 13 show the model-predicted influence of light intensity on the reaction outputs. A similarly nonlinear response was observed. Under the selected conditions, CO_2 conversion and CO yield reached their maximum values at an intermediate light intensity of approximately 1,000 mW cm^{-2}, with values of 5.31% and 50.67 mmol g^{-1} h^{-1}, respectively. This indicates that moderate illumination may enhance photocatalyst excitation and facilitate the participation of photogenerated charge carriers in CO_2 reduction in this P^2CR system, potentially in conjunction with plasma-induced activation[20,56]. However, further increasing the light intensity decreased both CO_2 conversion and CO yield, whereas H_2 yield increased from 1.33 to 4.07 mmol g^{-1} h^{-1}. This trend suggests that excessive illumination may favor water-related reduction and H_2 formation in this system[46]. Overall, these results indicate that light intensity has an optimal window in the P^2CR system with both insufficient and excessive illumination being unfavorable for maximizing CO_2 reduction. Supplementary Figure 14 shows the local model-predicted response to gas flow rate while discharge power, light intensity, relative humidity, and catalyst dosage were held constant. Increasing the gas flow rate from 20 to 60 mL min^{-1} caused an overall decrease in CO_2 conversion and CO yield. H_2 yield also decreased, but the decline was less pronounced than the declines observed for CO_2 conversion and CO yield. Gas flow rate affects the reaction through both residence-time and throughput effects. In the present reactor, the effective discharge volume was approximately 20 cm^3. Increasing the total gas flow rate from 20 to 60 mL min^{-1} therefore reduced the nominal gas residence time from approximately 60 to 20 s. The shorter residence time may reduce the interaction time among CO_2, plasma-generated energetic species, and subsequent surface reduction reactions. This effect could account for the decrease in CO_2 conversion. CO yield showed a similar response because CO formation is closely linked to the extent of CO_2 reduction and may require sufficient interaction time for the formation and desorption of CO intermediates. The weaker response of H_2 yield suggests that the H_2-forming pathway is less sensitive to gas flow rate than the CO-producing pathway. This is likely because H_2 formation does not require the same sequence of CO_2 activation and CO-intermediate formation as the CO-producing pathway. Consequently, H_2 yield is affected by gas flow rate to a lesser extent. The model-predicted response to catalyst dosage and the corresponding model-experiment correlations for catalyst dosage are provided in Supplementary Figure 15. This target-dependent behavior means that maximizing CO_2 conversion alone is insufficient to guarantee the desired syngas ratio. It is therefore necessary to combine model interpretation with constrained optimization for experimental condition screening.

SHAP analysis

To further examine how the operating parameters contributed to the model predictions, SHAP analysis was performed across the experimental domain[38,39]. As shown in Figure 5A-C, relative humidity, light intensity, and discharge power were identified as the most influential predictors for CO_2 conversion, CO yield and H_2

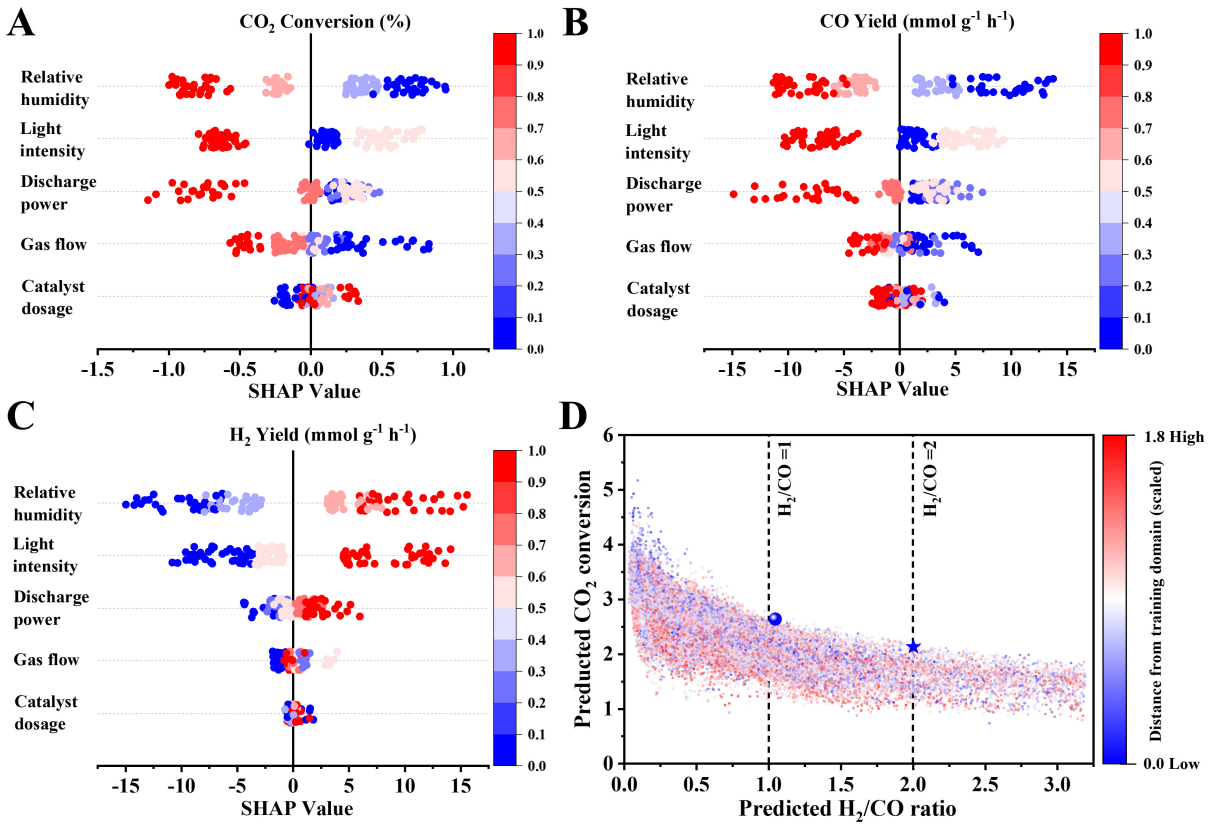


Figure 5. SHAP-value distribution for the five input parameters for (A) CO_2 conversion, (B) CO yield and (C) H_2 yield; (D) Optimization results of the hybrid model under H_2/CO-constraints. SHAP: SHapley Additive exPlanations.

yield, consistent with the R^2-drop analysis. SHAP analysis also revealed the direction and distribution of each feature's contribution across individual samples. For CO_2 conversion and CO yield, samples with high relative humidity were mainly located in the negative SHAP region, whereas samples with lower humidity were more associated with positive SHAP values. Thus, higher humidity was associated with lower predicted CO_2 conversion and CO yield. By contrast, high humidity was generally associated with positive SHAP values for H_2 yield. This output-dependent pattern is consistent with the response-trend analysis.

The SHAP distributions for light intensity and discharge power showed a pronounced nonlinear pattern. For CO_2 conversion and CO yield, moderate light intensity and discharge power were generally associated with positive SHAP contributions, whereas higher values did not further increase, and in some cases decreased, the model predictions. This indicates that the hybrid model captured an optimal operating window within the investigated domain. For H_2 yield, higher relative humidity, light intensity, and discharge power were generally associated with positive SHAP contributions, indicating that the model predictions were sensitive to these variables within the investigated range. By comparison, gas flow rate and catalyst dosage showed relatively small contributions to the model predictions between 30 and 60 mg catalyst dosage and across the investigated flow-rate range. Gas flow rate nevertheless showed a more noticeable contribution to the predictions of CO_2 conversion and CO yield than to H_2 yield. Overall, SHAP analysis provides an interpretable description of the positive and negative contributions of the input variables to the model predictions. These findings indicate that the model captured nonlinear input-output relationships within the P^2CR system while maintaining good predictive performance across the investigated experimental

conditions. Because global feature-importance measures do not fully describe conditional effects between operating variables, a low-high 2 × 2 conditional interaction analysis was additionally performed. The analysis procedure and complete results are provided in Supplementary Table 9.

Hybrid model application

The importance, response-trend, and SHAP analyses collectively indicate that the P²CR system exhibits strongly nonlinear, output-dependent behaviors. Relative humidity, light intensity, and discharge power produced different responses for CO_2 conversion, CO yield, and H_2 yield. In particular, conditions that favored CO_2 conversion and CO formation did not necessarily produce the desired H_2/CO ratio. Therefore, the trained hybrid model was applied to constrained optimization, aiming to identify operating conditions that satisfied prescribed H_2/CO ratios while maintaining relatively high CO_2 conversion. A total of 100,000 candidate conditions were generated by random sampling within the experimental ranges of the five operating variables. The hybrid model was used to predict CO_2 conversion, CO yield, and H_2 yield for each candidate, from which the corresponding H_2/CO ratio was calculated. H_2/CO ratios of 1 and 2 were selected as representative targets for downstream syngas applications. An H_2/CO ratio of 1 is relevant to oxo-synthesis, whereas a ratio of 2 is commonly associated with methanol synthesis and Fischer-Tropsch synthesis[57,58]. Among the 100,000 generated candidate conditions, 47,245 (47.2%) were located within the five-dimensional convex hull of the SPXY training set, whereas 52,755 (52.8%) fell outside the hull [Supplementary Figure 16]. Nevertheless, the large fraction of candidate conditions outside the hull highlights the potential risk of extrapolation when screening the broader hypothetical design space. Candidate conditions satisfying tolerance windows of ±0.05, ±0.10 and ±0.20 were screened. Among the feasible candidates, the condition with the highest predicted CO_2 conversion was selected as the conversion-oriented recommendation. Because predictions outside the training domain may be less reliable, the highest-conversion candidate was not necessarily considered the most robust choice[59,60]. We therefore introduced a robustness-oriented recommendation strategy that considered predicted CO_2 conversion, disagreement between the GBR and KRR models, deviation from the target H_2/CO ratio and distance from the training-data distribution. Smaller model disagreement, lower ratio error and greater proximity to the experimental data were considered indicators of lower prediction risk[60-62]. As shown in Figure 5D, each point represents a candidate reaction condition. A trade-off was observed between CO_2 conversion and syngas composition. Candidates with higher predicted CO_2 conversion were generally concentrated in the lower-H_2/CO region, whereas CO_2 conversion tended to decrease as the H_2/CO ratio increased. This indicates that high CO_2 conversion and a prescribed syngas composition cannot be optimized simultaneously without imposing an explicit ratio constraint[45,54]. The two final recommended conditions with a tolerance of ±0.05 were therefore selected under the specified H_2/CO constraint (1:1 and 2:1). Importantly, these two conditions were located within the training-set convex hull. The experimental validation results in Supplementary Tables 10 and 11 were consistent with the model predictions, supporting the use of the hybrid model for operating-condition screening within the investigated P²CR domain. In addition, CO productivity per unit plasma power was calculated for the two recommended conditions and is presented in Supplementary Table 12. The condition targeting an H_2/CO ratio of 1 achieved a CO productivity normalized by plasma power of 16.5 ± 0.9 mmol kWh^{-1}, whereas the condition targeting an H_2/CO ratio of 2 achieved 9.8 ± 0.3 mmol kWh^{-1}. These results indicate that the H_2/CO = 1 condition was more favorable for CO production per unit plasma power, whereas the H_2/CO = 2 condition produced a more H_2-rich syngas composition.

CONCLUSIONS

In this work, we developed and evaluated a small-sample machine-learning framework for a DBD plasma-coupled photocatalytic CO_2 reduction system using Cu-Pd/TiO_2. Based on 120 experimental runs, the shared-weight GBR-KRR model achieved test-set R^2 values of 0.947, 0.950, and 0.954 for CO_2 conversion, CO yield, and H_2 yield, respectively. Permutation importance and SHAP analyses identified relative humidity,

light intensity, and discharge power as the most influential predictors within the investigated operating domain. These rankings describe model sensitivity and should not be interpreted as direct evidence of mechanistic importance. The trained model was further used to screen operating conditions under target H_2/CO ratios of 1 and 2. The selected conditions were experimentally validated in the same Cu-Pd/TiO_2 system, yielding H_2/CO ratios of 1.06 ± 0.03 and 1.99 ± 0.07, respectively. Although the experimental mean value for the target of 1 is slightly above the nominal range of 1 ± 0.05, it remains very close to the desired value. These results indicate that the proposed framework can identify operating conditions close to the desired H_2/CO targets within the present P^2CR system. The conclusions of this study are limited to the Cu-Pd/TiO_2 photocatalyst, the square-quartz DBD reactor and the investigated ranges of discharge power, catalyst dosage, gas flow rate, relative humidity, and light intensity. Further work using broader catalyst systems, reactor configurations, plasma conditions, and independently collected datasets will be necessary to assess the generalizability and robustness of the framework.

DECLARATIONS

Acknowledgments

The authors wish to thank Prof. Aezid Hassan Najmi for help with language polishing.

Author's contributions

Substantial contributions to content discussion: Bi, R.; Gao, T.; Wang, Y.
Experimental equipment setup and maintenance: Wang, D.
Data curation, formal analysis: Zhu, H.; Xi, Z.; Qin, H.; Wang, H.; Geng, Z.; Yao, Y.; Guan, Y.
Manuscript revision before submission: Bi, R.; Zhang, H.; Wang, Y.

Availability of data and materials

The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author(s).

Financial support and sponsorship

The authors wish to thank the Ningbo Science and Technology Innovation Yongjiang 2035 Key R&D Project (No. 2025Z111) and the 2023 Ningbo Natural Science Funding Commonweal Programme Major Project (No. 2023S019). This work was also supported by the Ningbo Key Laboratory of Typical Solid Waste High-Value Utilisation and Low-Carbon Materials.

AI and AI-assisted tools statement

Not applicable.

Conflicts of interest

Wang, D. is affiliated with Beijing Perfectlight Technology Co., Ltd., which is associated with the DBD reactor system used in this study. The company had no involvement in the study design, data collection, data analysis, interpretation of the results, or the decision to publish the manuscript. The other authors declare that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Supplementary Materials

Supplementary Materials

REFERENCES

1. Friedlingstein, P.; O'sullivan, M.; Jones, M. W.; et al. Global carbon budget 2024. *Earth. Syst. Sci. Data.* **2025**, *17*, 965-1039. DOI
2. Grim, R. G.; Badgett, A.; Braunecker, W. A.; et al. The chemistry of CO_2 conversion: a review. *Chem. Rev.* **2026**, *126*, 5028-82. DOI
3. Garg, S.; Xie, Z.; Chen, J. G. Tandem reactors and reactions for CO_2 conversion. *Nat. Chem. Eng.* **2024**, *1*, 139-48. DOI
4. Yu, S.; Yamauchi, H.; Wang, S.; et al. CO_2-to-methanol electroconversion on a molecular cobalt catalyst facilitated by acidic cations. *Nat. Catal.* **2024**, *7*, 1000-9. DOI
5. Chen, H.; Xiong, Y.; Li, J.; et al. Epitaxially grown silicon-based single-atom catalyst for visible-light-driven syngas production. *Nat. Commun.* **2023**, *14*, 1719. DOI PubMed PMC
6. Zhou, W.; Wang, X.; Zhao, W.; et al. Photocatalytic CO_2 reduction to syngas using metallosalen covalent organic frameworks. *Nat. Commun.* **2023**, *14*, 6971. DOI PubMed PMC
7. Xie, Z.; Xu, S.; Li, L.; et al. Well-defined diatomic catalysis for photosynthesis of C_2H_4 from CO_2. *Nat. Commun.* **2024**, *15*, 2422. DOI PubMed PMC
8. Weber, J. L.; Mejía, C. H.; de Jong, K. P.; de Jongh, P. E. Recent advances in bifunctional synthesis gas conversion to chemicals and fuels with a comparison to monofunctional processes. *Catal. Sci. Technol.* **2024**, *14*, 4799-842. DOI PubMed PMC
9. Hou, X.; Jiang, Y.; Wei, K.; et al. Syngas production from CO_2 and H_2O via solid-oxide electrolyzer cells: fundamentals, materials, degradation, operating conditions, and applications. *Chem. Rev.* **2024**, *124*, 5119-66. DOI
10. Dossow, M.; Klüh, D.; Umeki, K.; Gaderer, M.; Spliethoff, H.; Fendt, S. Electrification of gasification-based biomass-to-X processes - a critical review and in-depth assessment. *Energy. Environ. Sci.* **2024**, *17*, 925-73. DOI
11. Zhang, Y.; He, T.; Chen, J.; et al. Recent progress in and future perspectives on high-density single-atom electrocatalysts. *Electrochem. Energy. Rev.* **2025**, *8*, 257. DOI
12. Xu, C.; Wu, N.; Ran, Y.; Hong, P.; Lei, Y. Two-in-one integrated CO_2/N_2 conversion and related systems: potential, status, and future. *Electrochem. Energy. Rev.* **2025**, *8*, 264. DOI
13. Fang, S.; Rahaman, M.; Bharti, J.; et al. Photocatalytic CO_2 reduction. *Nat. Rev. Methods. Primers.* **2023**, *3*, 243. DOI
14. He, H.; Ren, Y.; Zhang, L.; et al. Synergistic modulation of charge dynamics and mass transfer optimization via heterogeneous interface engineering in photothermal catalytic CO_2 reduction within continuous flow systems. *Nano. Energy.* **2025**, *142*, 111290. DOI
15. Choi, S. H.; Song, I.; Dong, W. J. Recent progress of photothermal catalysts for carbon dioxide conversion. *Energy. Mater.* **2025**, *5*, 500062. DOI
16. Ray, D.; Nepak, D.; Vinodkumar, T.; Subrahmanyam, C. g-C_3N_4 promoted DBD plasma assisted dry reforming of methane. *Energy* **2019**, *183*, 630-8. DOI
17. Gonzalez-Casamachin, D. A.; Hu, Y.; Rangarajan, S.; Baltrusaitis, J. Mechanistic insights for plasma-catalytic CO_2 reduction over TiO_2 in a dielectric barrier discharge reactor. *ACS. Eng. Au.* **2026**, *6*, 334-44. DOI
18. Yang, Y.; Murphy, A. B. CO_2 conversion using non-thermal plasmas: the path towards industrialisation. *Curr. Opin. Green. Sustain. Chem.* **2025**, *51*, 100994. DOI
19. Huang, Q.; Fu, C.; Shen, Y.; et al. Boosting CO_2 conversion by synergy of lead-free perovskite Cs_2SnCl_6 and plasma with H_2O. *J. Phys. Chem. Lett.* **2023**, *14*, 8922-9. DOI
20. Feng, X.; Cao, L.; Fu, C.; et al. Enhanced CO_2 conversion in dielectric barrier discharge plasma coupled with a heterojunction photocatalyst. *Chem. Commun.* **2024**, *60*, 8900-3. DOI
21. Bi, R.; Wang, Y.; Wang, D.; et al. Plasma-coupled photocatalytic CO_2/H_2O conversion to tunable syngas over Cu_xPd_y/TiO_2 with electronic modulation. *Appl. Catal. B. Environ. Energy.* **2027**, *402*, 127470. DOI
22. Wanten, B.; Vertongen, R.; De Meyer, R.; Bogaerts, A. Plasma-based CO_2 conversion: how to correctly analyze the performance? *J. Energy. Chem.* **2023**, *86*, 180-96. DOI
23. Shen, Y.; Fu, C.; Luo, W.; Liang, Z.; Wang, Z.; Huang, Q. Machine learning for CO_2 conversion driven by dielectric barrier discharge plasma and Cs_2TeCl_6 photocatalysts. *Green. Chem.* **2023**, *25*, 7605-11. DOI
24. Cai, Y.; Bai, Z.; Chen, C.; et al. Machine learning prediction of small molecule passivators and their impacts on the passivation and photocatalytic performance of organic-inorganic hybrid perovskite interfaces. *Energy. Mater.* **2025**, *5*, 500043. DOI
25. Han, N.; Su, B. L. AI-driven material discovery for energy, catalysis and sustainability. *Natl. Sci. Rev.* **2025**, *12*, nwaf110. DOI PubMed PMC

26. Luo, W.; Shen, Y.; Fu, C.; Feng, X.; Huang, Q. Exploring the CO_2 conversion activated by the dielectric barrier discharge plasma assisted with photocatalyst via machine learning. *J. Environ. Chem. Eng.* **2024**, *12*, 114428. DOI
27. Taniike, T.; Fujiwara, A.; Nakanowatari, S.; García-Escobar, F.; Takahashi, K. Automatic feature engineering for catalyst design using small data without prior knowledge of target catalysis. *Commun. Chem.* **2024**, *7*, 11. DOI PubMed PMC
28. Li, J.; Xu, J.; Rebrov, E.; Bogaerts, A. Machine learning-based prediction and optimization of plasma-catalytic dry reforming of methane in a dielectric barrier discharge reactor. *Chem. Eng. J.* **2025**, *507*, 159897. DOI
29. Hu, Y.; Sandt, R.; Spatschek, R. Practical feature filter strategy to machine learning for small datasets in chemistry. *Sci. Rep.* **2024**, *14*, 20449. DOI PubMed PMC
30. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Ann. Statist.* **2001**, *29*, 1189-232. DOI
31. Rupp, M.; Tkatchenko, A.; Müller, K. R.; von Lilienfeld, O. A. Fast and accurate modeling of molecular atomization energies with machine learning. *Phys. Rev. Lett.* **2012**, *108*, 058301. DOI PubMed
32. van der Laan, M. J.; Polley, E. C.; Hubbard, A. E. Super learner. *Stat. Appl. Genet. Mol. Biol.* **2007**, *6*, 25. DOI PubMed
33. Chicco, D.; Warrens, M. J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ. Comput. Sci.* **2021**, *7*, e623. DOI PubMed PMC
34. Williamson, B. D.; Gilbert, P. B.; Simon, N. R.; Carone, M. A general framework for inference on algorithm-agnostic variable importance. *J. Am. Stat. Assoc.* **2023**, *118*, 1645-58. DOI PubMed PMC
35. Fumagalli, F.; Muschalik, M.; Hüllermeier, E.; Hammer, B. Incremental permutation feature importance (iPFI): towards online explanations on data streams. *Mach. Learn.* **2023**, *112*, 4863-903. DOI
36. Altmann, A.; Toloşi, L.; Sander, O.; Lengauer, T. Permutation importance: a corrected feature importance measure. *Bioinformatics* **2010**, *26*, 1340-7. DOI PubMed
37. Bilodeau, B.; Jaques, N.; Koh, P. W.; Kim, B. Impossibility theorems for feature attribution. *Proc. Natl. Acad. Sci. U. S. A.* **2024**, *121*, e2304406120. DOI PubMed PMC
38. Štrumbelj, E.; Kononenko, I. Explaining prediction models and individual predictions with feature contributions. *Knowl. Inf. Syst.* **2014**, *41*, 647-65. DOI
39. Lundberg, S. M.; Erion, G.; Chen, H.; et al. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56-67. DOI PubMed PMC
40. Gao, B.; Feng, Y.; Li, C.; Zhao, J.; Fang, Y.; Cao, G. Defect-engineered Ru-Co/TiO_{2-x} catalysts for highly efficient plasma-catalytic ammonia synthesis under ambient conditions. *Green. Chem.* **2026**, *28*, 8235-48. DOI
41. Huang, Z.; Sun, S.; Ma, M.; et al. Facile synthesis of TiO_2 supported Pd nanoparticles for efficient photocatalytic CO_2 reduction to CH_4 with H_2O. *Sustain. Mater. Technol.* **2025**, *43*, e01247. DOI
42. Liu, S.; Meng, Y.; Luo, C.; et al. Green hydrogen and acetaldehyde production via photo-induced thermal dehydrogenation of bioethanol over a highly dispersed CuS/TiO_2 catalyst. *Int. J. Hydrogen. Energy.* **2024**, *86*, 1414-23. DOI
43. Liang, J.; Tian, F.; Li, Y.; Wang, Y.; Yan, X.; Zhang, H. Single-atom Cu anchored phosphorus-doped graphitic carbon nitride for selective photocatalytic CO_2 reduction. *Environ. Res.* **2025**, *285*, 122324. DOI
44. Huang, Q.; Fu, C.; Liang, Z.; et al. CO_2 conversion synergistically driven by radiofrequency inductively-coupled plasma and lead-free halide perovskite photocatalyst. *J. Phys. Chem. C.* **2023**, *127*, 11550-8. DOI
45. Han, Y.; Fan, G.; Guo, Y.; et al. Plasma-Driven Efficient conversion of CO_2 and H_2O into pure syngas with controllable wide H_2/CO ratios over metal-organic frameworks featuring in situ evolved ligand defects. *Angew. Chem. Int. Ed.* **2024**, *63*, e202406007. DOI
46. Haghshenas, Y.; Wong, W. P.; Gunawan, D.; et al. Predicting the rates of photocatalytic hydrogen evolution over cocatalyst-deposited TiO_2 using machine learning with active photon flux as a unifying feature. *EES. Catal.* **2024**, *2*, 612-23. DOI
47. Wu, X.; Shen, C.; Zhang, X.; Wang, Y.; Liu, C. Dielectric surface engineering via TiO_2 coating for enhanced plasma-driven efficient CO_2 decomposition. *Chem. Eng. J.* **2025**, *505*, 159743. DOI
48. Trampel, C. P.; Stieler, D. S. Control of electromagnetic edge effects in electrically-small rectangular plasma reactors. *J. Vac. Sci. Technol. A.* **2012**, *30*, 051305. DOI
49. Vertongen, R.; Bogaerts, A. How important is reactor design for CO_2 conversion in warm plasmas? *J. CO2. Util.* **2023**, *72*, 102510. DOI
50. Deo, J. V.; Madhukar, A. Study of electric field distribution on plasma and plasma catalysis reactor for different electrode configurations and pellet sizes. *PPT* **2024**, *11*, 54-60. DOI
51. Lisi, N.; Pasqual Laverdura, U.; Chierchia, R.; Luisetto, I.; Stendardo, S. A water cooled, high power, dielectric barrier discharge reactor for CO_2 plasma dissociation and valorization studies. *Sci. Rep.* **2023**, *13*, 7394. DOI PubMed PMC
52. Liang, C.; Zhang, Z.; Liu, Z.; Wang, L.; Lam, Y. M. Thermodynamic and microkinetic insights into potential-dependent competition between CO_2 reduction and HER on Mo_2CT_x MXene. *Chem. Eng. J.* **2025**, *520*, 165927. DOI

53. Chawdhury, P.; Chansai, S.; Conway, M.; et al. Enhancing the reaction of CO_2 and H_2O using catalysts within a nonthermal plasma. *ACS. Catal.* **2025**, *15*, 7053-65. DOI
54. Klein, M.; Jack, J. Non-thermal plasma upgrading of humidified CO_2 into syngas in a dielectric barrier discharge reactor: tuning H_2/CO ratios *via* specific energy input and gas flow rate. *Energy. Adv.* **2026**, *5*, 354-64. DOI
55. Snoeckx, R.; Ozkan, A.; Reniers, F.; Bogaerts, A. The quest for value-added products from carbon dioxide and water in a dielectric barrier discharge: a chemical kinetics study. *ChemSusChem* **2017**, *10*, 409-24. DOI
56. Behroozi, A. H.; Xu, R. Photocatalytic CO_2 reduction: photocatalysts, membrane reactors, and hybrid processes. *Chem. Catal.* **2023**, *3*, 100550. DOI
57. Alamdari, A.; Azarhoosh, M. J.; Aghaeinejad-Meybodi, A. Thermodynamic assessment of tri-reforming of methane with optimization of operating conditions to achieve suitable syngas for methanol production. *Sci. Rep.* **2026**, *16*, 14257. DOI PubMed PMC
58. Bonde, A.; Jakobsen, J. B.; Ahrens, A.; et al. Integrating hydroformylations with methanol-to-syngas reforming. *Chem* **2025**, *11*, 102396. DOI
59. Yoshizawa, T.; Ishida, S.; Sato, T.; Ohta, M.; Honma, T.; Terayama, K. A data-driven generative strategy to avoid reward hacking in multi-objective molecular design. *Nat. Commun.* **2025**, *16*, 2409. DOI PubMed PMC
60. Janet, J. P.; Duan, C.; Yang, T.; Nandy, A.; Kulik, H. J. A quantitative uncertainty metric controls error in neural network-driven chemical discovery. *Chem. Sci.* **2019**, *10*, 7913-22. DOI PubMed PMC
61. Haddadnia, M.; Grashoff, L.; Strieth-Kalthoff, F. *BoTier*: multi-objective Bayesian optimization with tiered objective structures. *Digit. Discov.* **2025**, *4*, 1417-22. DOI
62. Frömbgen, T.; Surzhikova, E.; Dölz, J.; Proppe, J.; Kirchner, B.; Jacob, C. R. Uncertainty quantification for in silico chemistry. *Chem. Rev.* **2026**, *126*, 4189-236. DOI PubMed PMC