Journal of Energy Chemistry 96 (2024) 153–163

Contents lists available at ScienceDirect

Journal of Energy Chemistry

journal homepage: www.elsevier.com/locate/jechem

http://www.journals.elsevier.com/journal-of-energy-chemistry/

Machine learning-driven optimization of plasma-catalytic dry reforming of methane

Yuxiang Cai [a,b], Danhua Mei [c], Yanzhen Chen [a], Annemie Bogaerts [b,*], Xin Tu [a,*]

[a] Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, UK
[b] Research Group PLASMANT, Department of Chemistry, University of Antwerp, Universiteitsplein 1, BE-2610 Wilrijk-Antwerp, Belgium
[c] College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing 211816, Jiangsu, China

ARTICLE INFO

Article history:
Received 28 December 2023
Revised 7 April 2024
Accepted 16 April 2024
Available online 25 April 2024

Keywords:
Plasma catalysis
Machine learning
Process optimization
Dry reforming of methane
Syngas production

ABSTRACT

This study investigates the dry reformation of methane (DRM) over Ni/Al_2O_3 catalysts in a dielectric barrier discharge (DBD) non-thermal plasma reactor. A novel hybrid machine learning (ML) model is developed to optimize the plasma-catalytic DRM reaction with limited experimental data. To address the non-linear and complex nature of the plasma-catalytic DRM process, the hybrid ML model integrates three well-established algorithms: regression trees, support vector regression, and artificial neural networks. A genetic algorithm (GA) is then used to optimize the hyperparameters of each algorithm within the hybrid ML model. The ML model achieved excellent agreement with the experimental data, demonstrating its efficacy in accurately predicting and optimizing the DRM process. The model was subsequently used to investigate the impact of various operating parameters on the plasma-catalytic DRM performance. We found that the optimal discharge power (20 W), CO_2/CH_4 molar ratio (1.5), and Ni loading (7.8 wt%) resulted in the maximum energy yield at a total flow rate of ~51 mL/min. Furthermore, we investigated the relative significance of each operating parameter on the performance of the plasma-catalytic DRM process. The results show that the total flow rate had the greatest influence on the conversion, with a significance exceeding 35% for each output, while the Ni loading had the least impact on the overall reaction performance. This hybrid model demonstrates a remarkable ability to extract valuable insights from limited datasets, enabling the development and optimization of more efficient and selective plasma-catalytic chemical processes.

1. Introduction

Methane (CH_4) and carbon dioxide (CO_2) are two of the most important greenhouse gases (GHGs), contributing significantly to climate change [1]. There has been an increasing demand to develop innovative approaches to transform these GHG emissions into useful products [2,3]. Dry reforming of methane (DRM) is a promising process to simultaneously reduce both CH_4 and CO_2 while producing valuable syngas (Eq. (1)), which can be used as a source of hydrogen or to synthesize higher value chemicals via the Fischer-Tropsch process [4].

$$CH_4 + CO_2 \rightarrow 2CO + 2H_2, \Delta H_0 = 247.3 \text{ kJ/mol} \quad (1)$$

The DRM process typically operates at temperatures over 800 °C in the presence of a catalyst to enhance the reaction kinetics. This requires substantial energy input, which can be a major cost barrier for commercialization. In recent decades, significant efforts have been devoted to exploring various catalysts, including noble [1,5] and transition metal-based catalysts [6–8], to lower the activation barriers of the reaction. Among these catalysts, nickel catalysts have emerged as a highly attractive alternative due to their affordability and superior catalytic activity [9–12]. Nevertheless, the rapid deactivation of nickel-based catalysts under harsh reaction conditions, induced by coking or sintering, represents a significant challenge to their practical application [13].

Non-thermal plasma (NTP) is a promising alternative to conventional methods for enabling the DRM reaction at low temperatures and ambient pressure. In a typical plasma-catalytic DRM process, plasma discharge is generated in the presence of a catalytic material. Such discharge provides energy to the reaction,

* Corresponding authors.
E-mail addresses: annemie.bogaerts@uantwerpen.be (A. Bogaerts), xin.tu@liverpool.ac.uk (X. Tu).

https://doi.org/10.1016/j.jechem.2024.04.022

facilitating the breakdown of methane and carbon dioxide and producing reactive species (e.g., CH_x, H, and O) that can participate in the DRM reaction. The catalytic material stabilizes these species and promotes the selective formation of syngas. The integration of plasma and catalyst has the potential to result in a synergistic effect, leading to enhanced overall efficiency [14,15]. The flexibility of the plasma-catalytic process for instant start-up and shutdown makes it a favorable choice for integration with intermittent renewable energy [16,17]. In recent decades, various types of NTP have been extensively studied for their potential in plasma-catalytic DRM, including glow discharge [18], corona discharge [19], gliding arc [20,21], microwave discharge [22], and dielectric barrier discharge (DBD) [23–26]. Alongside this research, the selection of Ni-based catalysts with different promoters [27–29] and supports [30–32] has also been explored. Despite its tremendous potential, further research is necessary to fully comprehend the interactions among the various reaction parameters, as well as the synergistic effect of plasma and catalytic material. Moreover, The commercial viability of plasma-catalytic DRM depends on overcoming its challenges in cost-effectiveness and scalability.

The plasma-catalytic DRM process is complex, with many interacting parameters [33,34]. A reliable model is crucial to qualitatively analyze these interactions and to comprehensively optimize reaction conditions. This can help to improve the overall efficiency of the process and address the challenges of this emerging technology [35]. Statistical models, such as the response surface method (RSM), have been used to analyze the effects of various reaction conditions on the plasma-catalytic DRM reaction [24] and acetone oxidation [36]. This has allowed researchers to identify optimal conditions for each reaction. However, the RSM is not always accurate or robust enough for complex and nonlinear plasma-catalytic systems. Kinetic models, on the other hand, provide a deeper understanding of the underlying mechanisms and reaction kinetics in a plasma reaction system [18,37,38]. These models are constructed using a series of mathematical equations that describe the time-evolution of all relevant plasma species based on production and loss rates, defined by chemical reactions. For example, Wanten et al. [18] developed a quasi-1D kinetics model to describe plasma-based DRM in an atmospheric pressure glow discharge reactor. The results of the model agreed well with experimental data, and the model provided insights into the reaction pathway and the significance of vibrationally excited CO_2 and CH_4 in the reactions. While kinetics models provide strong interpretability, their applicability is mainly limited to plasma-based chemical processes without a catalyst, although recently some models were developed for the integration of plasma chemistry and catalyst surface chemistry, for DRM [34]. Furthermore, kinetics models exhibit limitations in terms of predictability and present challenges as not all reaction rate coefficients are accurately known in literature.

To overcome these limitations, machine learning (ML) approaches have received increasing attention for plasma processes, including the synthesis of fuels and chemicals and pollution control. Prior research has predominantly employed artificial neural network (ANN) models. For instance, Liu et al. [39] developed a well-trained ANN model for plasma-based non-oxidative CH_4 coupling, while Zhu et al. [40] proposed a three-layer ANN model for methanol oxidation over Cu-Ce/Al_2O_3 catalysts in a post-plasma catalytic system. Similarly, Chang et al. [41] used an ANN model with non-dominated sorting genetic algorithm II (NSGA-II) to investigate the post-plasma catalytic removal of toluene. The model effectively provided a range of viable operating parameters, demonstrating its capability to achieve optimal reaction conditions. Recently, Pan et al. [42] proposed a deep learning (DL) model that cooperates with both experiments and kinetics simulations to optimize the plasma-catalytic CH_4 dissociation and ammonia synthesis. These models effectively elucidated the impact of various parameters on the reaction performance, predicted the optimal conditions, and identified the most significant factors. Despite recent advancements, the application of ML algorithms to optimize the plasma-catalytic DRM process remains a largely unexplored area in current research.

Support vector regression (SVR), regression trees (RT), and ANN algorithms have been extensively used in the field of ML to solve regression problems. The performance and accuracy of ML models depend on a substantial set of input data. The quality and quantity of data used for model training significantly impact the performance of the model. However, obtaining a sufficient amount of accurate experimental data for the plasma-catalytic process to train ML models is still a critical challenge. The limited sample size of plasma-catalytic reaction data sets and the intricate interdependence of operating parameters can easily lead to overfitting when using a single ML model [43].

To achieve cost-effectiveness plasma-catalytic DRM process, a thorough understanding and control of interconnected operating parameters is crucial. However, plasma catalysis is a complex nonlinear system lacking extensive datasets typically available in thermal catalysis and electrocatalysis. Recognizing this challenge (limited data in complex systems), we designed our model for small sample learning. This hybrid ML model integrates three algorithms (ANN, RT, and SVR) to address the challenges arising from system complexity and limited sample size. To enhance the robustness and prediction accuracy of the model, the K-fold cross-validation strategy was used in model training, and the hyperparameters of the three algorithms were optimized with a genetic algorithm (GA). Finally, the model was trained using representative experimental data from plasma-catalytic DRM and compared with models using a single algorithm. The well-trained model was then applied to assess the effects and significance of various operating parameters, including Ni loading, CO_2/CH_4 molar ratio, total gas flow rate, and discharge power, on critical performance indicators (CO_2 and CH_4 conversion, H_2 and CO yield). Furthermore, by introducing fuel production efficiency (FPE) and energy yield (EY) indices, the model can determine the optimal operating parameters that maximize the cost-effectiveness of the plasma-catalytic DRM process.

2. Experimental and method

2.1. Catalyst preparation

Ni/γ-Al_2O_3 catalysts with varying Ni loadings (5, 7.5, 10, 12.5, and 15 wt%) were synthesized using the wetness impregnation method. First, a specific quantity of $Ni(NO_3)_2{\cdot}6H_2O$ was dissolved in deionized water and stirred at 60 °C for 15 min to ensure complete dissolution. Subsequently, γ-Al_2O_3 beads (1 mm in diameter) were introduced into the solution and thoroughly stirred for 12 h. The resulting mixture was then evaporated at 80 °C for 4 h, followed by overnight drying at 110 °C. Finally, the catalyst samples were calcined under an air atmosphere at 400 °C for a duration of 4 h.

2.2. Experimental method

Fig. 1 depicts the schematic diagram of the experimental setup for DRM. The catalyst samples were placed within the discharge zone of a coaxial DBD reactor, whose structure was in detail described in our previous work [44]. The DBD reactor was powered by a high voltage alternating current (AC) source, operating at a fixed frequency of 10 kHz and a peak voltage range of 0–30 kV. Electrical signals were sampled using a four-channel digital oscillo-

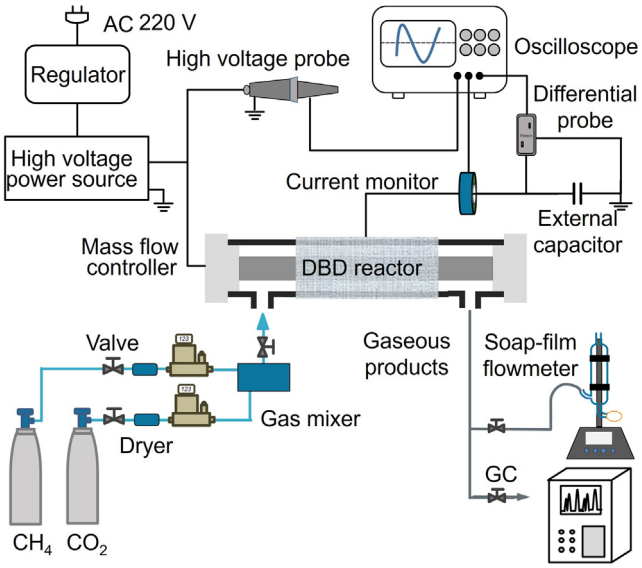

Fig. 1. Schematic diagram of the experimental setup.

scope (TDS2014), and the discharge power was calculated based on the measured Q-U Lissajous figure. Real-time monitoring of the discharge power was recorded using a custom-built online measurement system. A gas chromatograph (Shimadzu GC-2014) equipped with dual detectors was employed for product analysis. To ensure measurement accuracy, all experiments were conducted three times, showing remarkable reproducibility (with an error below 5%). The gas flow rate at the exit of the DBD reactor was measured using a soap-film flow meter to evaluate changes in the gas flow rate before and after the plasma reaction, i.e., gas expansion (or contraction), which is crucially important for determining the reaction performance [33].

The CO_2 conversion (X_{CO_2}) and CH_4 conversion (X_{CH_4}) were defined as follows

$$\chi_{CO_2}\,(\%) = \frac{CO_2\ \text{inlet}\ (\text{mol/s}) - CO_2\ \text{outlet}\ (\text{mol/s})}{CO_2\ \text{inlet}\ (\text{mol/s})} \times 100\% \quad (2)$$

$$\chi_{CH_4}(\%) = \frac{CH_4\ \text{inlet}\ (\text{mol/s}) - CH_4\ \text{outlet}\ (\text{mol/s})}{CH_4\ \text{inlet}\ (\text{mol/s})} \times 100\% \quad (3)$$

The H_2 yield (Y_{H_2}) and CO yield (Y_{CO}) were determined by

$$Y_{H_2}(\%) = \frac{2 \times H_2\ \text{outlet}\ (\text{mol/s})}{CH_4\ \text{inlet}\ (\text{mol/s})} \times 100\% \quad (4)$$

$$Y_{CO}(\%) = \frac{2 \times CO\ \text{outlet}\ (\text{mol/s})}{CH_4\ \text{inlet}\ (\text{mol/s}) + CO_2\ \text{inlet}\ (\text{mol/s})} \times 100\% \quad (5)$$

The EY and FPE [44] were calculated by

$$\text{EY}\ (\text{mmol/kJ}) = \frac{CH_4\ \text{inlet}\ (\text{mol/s}) - CH_4\ \text{outlet}\ (\text{mol/s}) + CO_2\ \text{inlet}\ (\text{mol/s}) - CO_2\ \text{outlet}\ (\text{mol/s})}{\text{Discharge power}\ (\text{kW})} \times 1000 \quad (6)$$

$$\text{FPE}\ (\%) = \frac{\sum \text{Fuel outlet}\ (\text{mol/s}) \times \text{LHV}\ (\text{kJ/mol})}{CH_4\ \text{inlet}\ (\text{mol/s}) \times \text{LHV}_{CH_4}\ (\text{kJ/mol}) + \text{Discharge power}\ (\text{kW})} \times 100\% \quad (7)$$

The lower heating value (LHV) of fuels is expressed in terms of the energy content per mole. For CH_4, H_2, and CO, the LHV values are 802.3, 241.8, and 283.2 kJ/mol, respectively [44].

For catalytic testing, 0.5 g Ni/Al_2O_3 pellets were placed within the discharge region of the DBD reactor, partially filling the discharge gap. Before the plasma-catalytic DRM experiment, the catalysts were reduced in situ with a H_2-Ar mixture for 30 min. During the reduction process, the discharge power was maintained at 40 W, while the total flow rate was fixed at 50 mL/min with a 20% H_2 composition. Following catalyst reduction, the plasma-catalytic DRM process was conducted. A feed gas mixture comprising CO_2 and CH_4, with CO_2/CH_4 molar ratios ranging from 0.5 to 1.5, was continuously supplied at total flow rates between 25 and 150 mL/min. The discharge power varied between 20 and 60 W. A comprehensive investigation involving 100 distinct reaction conditions was carried out. While the dataset comprised only 100 data points, these were meticulously collected across a broad range of testing conditions, ensuring the robustness and relevance of the model. The collected data, outlined in Tables S1 and S2, will form the foundation for training the ML model.

2.3. Description of the hybrid ML model

A hybrid model is proposed to comprehensively simulate and evaluate the plasma-catalytic DRM reaction. This model linearly combines ANN, SVR, and RT algorithms with individually assigned weights (W_1, W_2, and W_3), as shown in Eq. (8).

$$P = W_1 \times P_{ANN} + W_2 \times P_{SVR} + W_3 \times P_{RT}\ (0 \le W_1, W_2, W_3 \le 1) \quad (8)$$

The accuracy of the model is assessed using the mean squared error (MSE), which quantifies the average squared difference between the predicted values (P_i) and the corresponding experimental values (R_i).

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(P_i - R_i)^2 \quad (9)$$

To optimize the relative weight for each algorithm, the exhaustive method is employed to minimize the MSE of the model, with a step length of 0.01.

The developed hybrid model utilizes a framework for optimization, evaluation, and prediction based on four key input parameters: Ni loading, total flow rate, CO_2/CH_4 molar ratio, and discharge power (Fig. 2). The model predicts the conversion of the reactants (CO_2 and CH_4) and the yield of major products (H_2 and CO). To ensure compatibility between variables with different scales, input variables are normalized using the min-max normalization method before training the model. To prevent overfitting and ensure that the model generalizes well to unseen data (beyond the training set), we employed 10-fold cross-validation as a primary approach, given the relatively small dataset size. This method involves randomly dividing the dataset into ten parts, using nine for training and one for testing. This process is repeated ten times,

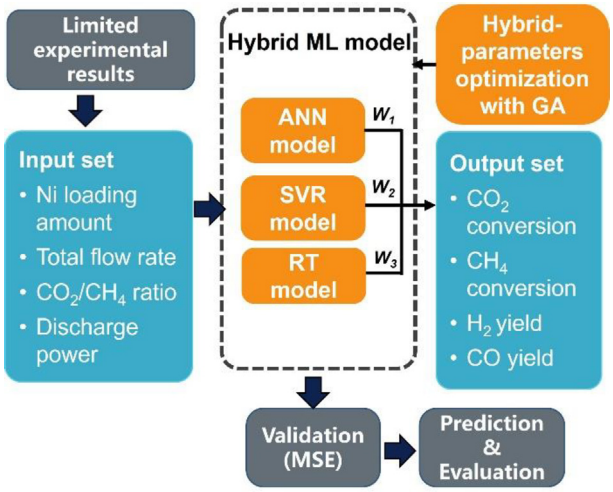

Fig. 2. Logical structure scheme of the framework for optimization, evaluation, and prediction within the hybrid ML model.

with each part being used as the testing set once. This approach not only helps in assessing the performance of the model but also ensures that it generalizes well across different subsets of the data.

ANN, inspired by the human brain's learning and decision-making processes, is widely used in time series prediction and modeling due to their exceptional predictive accuracy and adaptability to nonlinear and dynamic systems. SVR is an algorithm that excels in identifying hyperplanes in the feature space, demonstrating its superiority in nonlinear, small-sample, and high-dimensional regression tasks. The RT model, another classical supervised regression algorithm, recursively divides the dataset based on the most significant feature until a stopping criterion is met. RT algorithms are renowned for their efficiency and ability to handle missing data. In this study, a typical binary RT algorithm with a single root, multiple branches, nodes, and leaves is employed to further strengthen the robustness of the hybrid model.

A GA was chosen to optimize the hyperparameters for the three algorithms within the hybrid ML model due to its effectiveness in handling complex models with numerous parameters, as encountered in our study [45]. These parameters include the number of hidden layers, number of neurons per layer, and activation function for ANN, deviation (ε), punishment coefficient (C), kernel function for SVR, and max depth for RT. The optimized hyperparameters and their corresponding ranges are listed in Table 1; hyperparameters not explicitly mentioned were set to their default values. After GA optimization (Fig. S1), the hyperparameters are determined and listed in Table 1.

2.4. Methods of significance analysis

In this particular analysis, the Pearson's Correlation Coefficient (PCC) was utilized as a univariate analysis method to evaluate the strength of the relationship between each parameter and the performance [46]. The PCC quantifies the linear correlation between two variables and is bounded within the range of −1 to 1. A PCC of −1 represents the complete negative correlation, which means that as one variable increases, the other decreases. Conversely, a value of +1 signifies a full positive correlation, where both variables increase or decrease together. A value of 0 indicates the absence of a linear correlation between the two variables. The overall PCC between input and output variables is calculated as the quotient of the covariance and the standard deviation of the two variables.

Table 1
Hyperparameters and optimized ranges.

ML algorithm	Hyperparameter	GA optimization range	Optimized value
ANN	Hidden layer, n_{layer}	[1,2], $n_{layer} \in N$	1
	Neurons per layer, $n_{neuron,\ i}$	[2,15], $n_{neuron} \in N$	14
	Activation function	linear, sigmoid, ReLU[a], tanh	tanh
SVR	Deviation, ε	$\varepsilon = 10^{\alpha}$, $\alpha \in [-6, 0]$, $\alpha \in N$	10
	Punishment coefficient, C	$C = 10^{\beta}$, $\beta \in [-4, 2]$, $\beta \in N$	10
	Kernel function	linear, RBF[b], sigmoid, poly	RBF
RT	Max depth, n_{depth}	[1,9], $n_{depth} \in N$	7

[a] Rectified linear unit.
[b] Radial basis function.

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_x \sigma_y} = \frac{\sum(X-\mu_x)(Y-\mu_y)}{\sigma_x \sigma_y} \quad (10)$$

where cov(X, Y) represents the covariance between input variable X and output variable Y; while σ_x and σ_y correspond to the standard deviation of X and Y, respectively. After training and validating the hybrid ML model, we conducted a grid search, evaluating the model at 10,000 data points generated by selecting 10 values for each of the four input parameters within their respective ranges. Using the predicted values for these points, we calculated the PCC between each input parameter and each output variable (CO_2 conversion, CH_4 conversion, H_2 yield, and CO yield). To determine the relative significance of each input parameter (operating parameter) for each output variable (key performance metric), we calculated the relative importance scores based on the absolute values of the PCC.

3. Results and discussion

3.1. Hybrid model training and evaluation

The weights assigned to each algorithm significantly influence the MSE of the hybrid model, as shown in Fig. 3. When using a single algorithm, the MSEs of the ML models were found to be 0.0273 for ANN, 0.0125 for SVR, and 0.0176 for RT, respectively. The performance of the single-algorithm models followed the order SVR > RT > ANN, with the relatively lower performance of ANN being possibly attributed to the limited size of the training set. It is noteworthy that even after individually optimizing each model, their MSEs were still significantly higher than the minimum value of 0.0075 achieved by the hybrid model, demonstrating the superiority of the proposed hybrid model. Through an exhaustive optimization process, the optimal weights of ANN, SVR, and RT were determined to be 0.26, 0.44, and 0.30, respectively. Thus, the hybrid ML model can be expressed as

$$P_{Hybrid} = 0.26P_{ANN} + 0.44P_{SVR} + 0.30P_{RT} \quad (11)$$

The performance of the hybrid model was then evaluated by comparing the measured and predicted values for gas conversion and product yield, as presented in Fig. 4. Trained by optimizing connection weights to minimize MSE, the hybrid model demonstrated a commendable level of accuracy in predicting experimental data across the entire dataset. Notably, the measured collection efficiency was predicted with a confidence exceeding 98% and an MSE of only 0.0027. This remarkable performance indicates the effectiveness of the model in accurately predicting the performance of the plasma-catalytic DRM process. To further validate

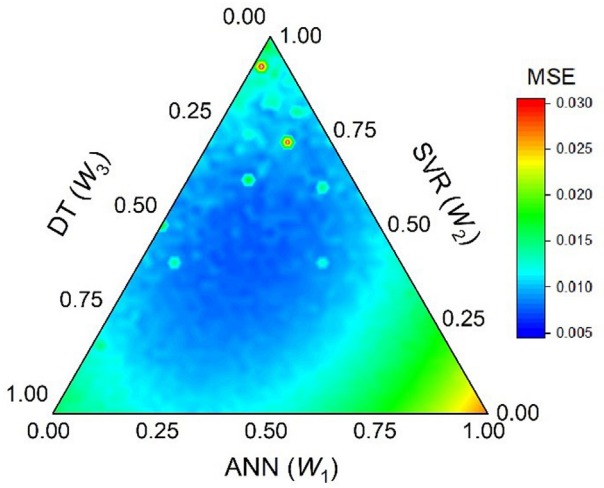

Fig. 3. MSE of the training and optimization process for the hybrid model. The lowest overall MSE is found when the weights of ANN, SVR, and RT are 0.26, 0.44, and 0.30, respectively.

the generalizability of the model, a series of experiments with new operating parameters within the defined input ranges (Ni loading = 7.5 wt%, CO_2/CH_4 = 1:1, discharge power = 30 W, and total flow rate = 25–125 mL/min) were performed. The prediction of the model for unseen data exhibited good agreement with experimental results, as presented in Fig. S2, demonstrating the reliability of the model.

3.2. Performance prediction

3.2.1. Effect of Ni loading

Fig. 5 shows the correlation between Ni loading and the performance of the plasma-catalytic DRM process. Solid scatter points represent the experimental data, while the lines depict predictions from the model. The hybrid model demonstrated exceptional accuracy in predicting gas conversions and major product yields, as evident from the excellent agreement between the model predictions and the experimental data.

The model predicted that the performance (gas conversion and product yield) of the plasma-catalytic DRM reaction would be increased with Ni loading up to around 8%, followed by a decrease at higher loadings. Consistent with this prediction, the experimental results also showed that the highest conversion of CO_2 (22.3%) and CH_4 (29.7%) was achieved using the 7.5 wt% Ni/Al_2O_3 catalyst, with other conditions kept constant. This finding aligns with the established trend of Ni loading impacting catalytic activity for CO_2 and CH_4 conversion: 7.5 wt% > 10.0 wt% > 5.0 wt% > 12.5 wt % > 15.0 wt%. Similar trends were observed for the yield of both H_2 and CO. Likewise, the predictions from the model effectively illustrate the performance as a function of Ni loading, with the optimal loading predicted to be around 7.8%. The performance of the plasma-catalytic DRM process is strongly affected by the physiochemical properties of Ni-based catalysts [27,29]. At lower Ni loading, the catalyst surface contains fewer active sites but a larger specific surface area [47], leading to an initial increase in CO_2 conversion with increasing Ni loading. However, excessive Ni loading can lead to metal particle aggregation, reducing the specific surface area and weakening metal dispersion, ultimately resulting in a

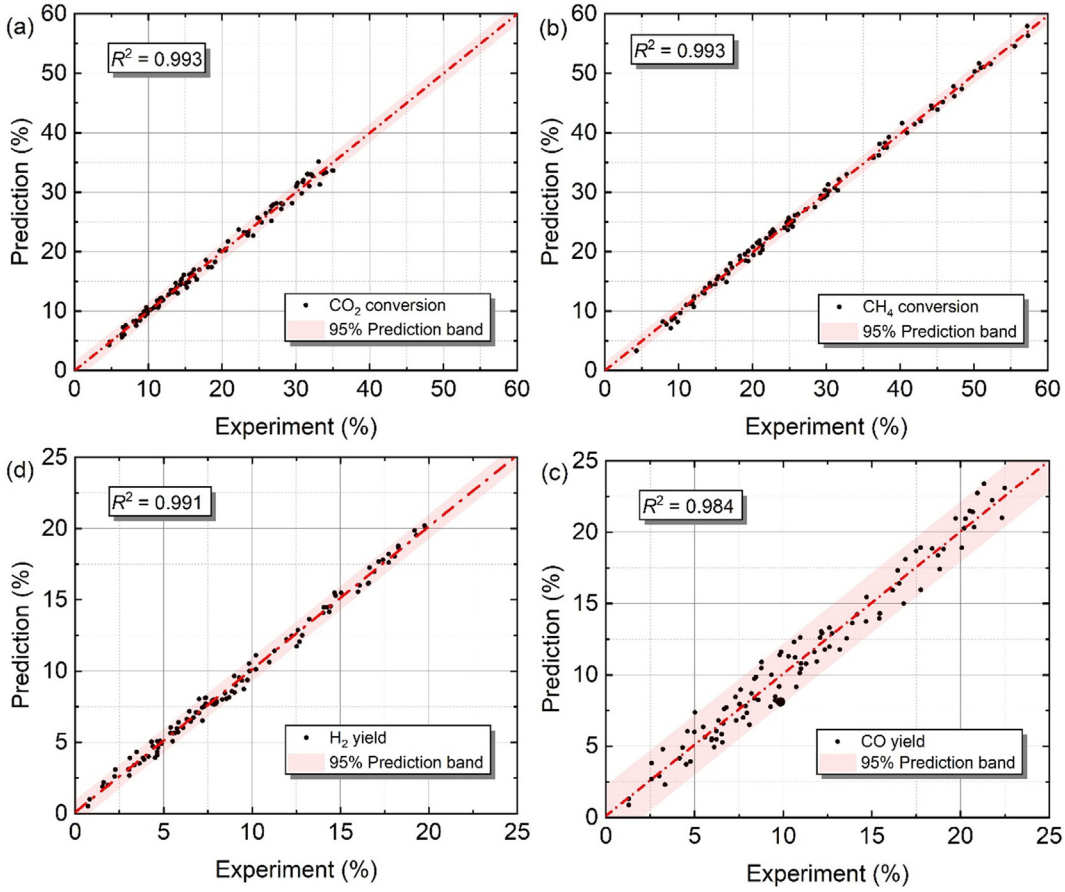

Fig. 4. Comparison of model predictions with experimental data. (a) CO_2 conversion; (b) CH_4 conversion; (c) CO yield; (d) H_2 yield.

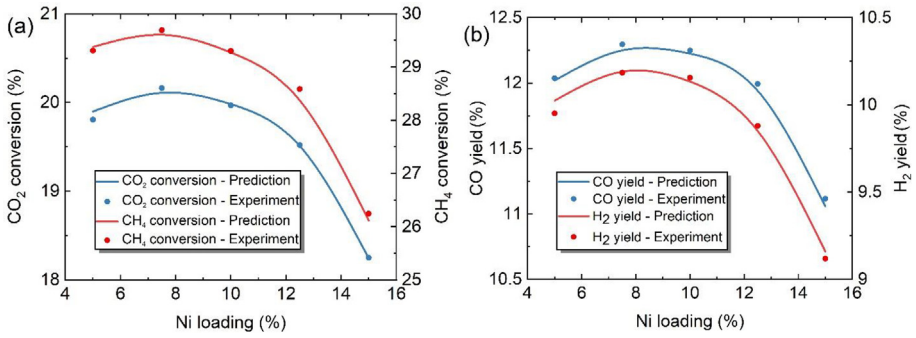

Fig. 5. Correlation between model predictions and experimental data for the effect of Ni loading on plasma-catalytic DRM performance. (a) CH_4 conversion and CO_2 conversion; (b) CO yield and H_2 yield (Discharge power = 40 W, CO_2/CH_4 = 1:1, and total flow rate = 75 mL/min).

decrease in the gas conversion and product yields [48]. Therefore, optimal CO_2 conversion is achieved at a moderate Ni loading under specific reaction conditions. This conclusion is supported by similar findings in literature [14–17]. Note that the discussion above is based on the characteristics of Ni/γ-Al_2O_3 catalysts prepared by similar methods in previous studies [47–49]. The hybrid model does not directly predict the physiochemical properties of the catalysts. This limitation highlights a common challenge with ML models: while they excel in prediction, their interpretability can be limited. This fact also underscores the importance of prior knowledge and empirical observations when discussing the outcomes of ML.

Despite its noticeable effect, the influence of Ni loading on the performance of plasma-catalytic DRM reaction cannot be considered as dominant. As the nickel loading increased from 5 wt% to 15 wt%, the variations in the CO_2 and CH_4 conversion remained within 5%, while the changes in the yields of CO and H_2 remained within 2%. In addition, according to PCC analysis, the effect of Ni loading amount shows very little interaction with the other parameters.

3.2.2. Effect of each reaction condition

In addition to Ni loading amount, the discharge power, CO_2/CH_4 ratio, and total flow rate also influence the plasma-catalytic DRM reaction. These parameters are recognized to have a profound impact on the reaction kinetics, thermodynamics, and mechanism [37]. Thus, a comprehensive understanding of these operating parameters is essential for optimizing the performance of the plasma-catalytic DRM process. The well-trained hybrid ML model enables us to study the overarching influence of the parameters beyond the constraints of limited data. Fig. 6 illustrates the predicted results for each parameter, while Figs. S3 and S4 present comparisons between predicted and (available) experimental data, demonstrating excellent agreement. This is further supported by the previously mentioned low MSE and high R^2 values.

Both CO_2 and CH_4 conversion exhibited a near-linear increase with discharge power, as predicted by the hybrid model (Fig. 6a). When using a 7.5 wt% Ni/Al_2O_3 catalyst, the highest CO_2 conversion of 20.1% and CH_4 conversion of 34.7% were predicted at a discharge power of 60 W. In this study, the DBD plasma employed for the DRM process exhibited characteristics of a filamentary discharge

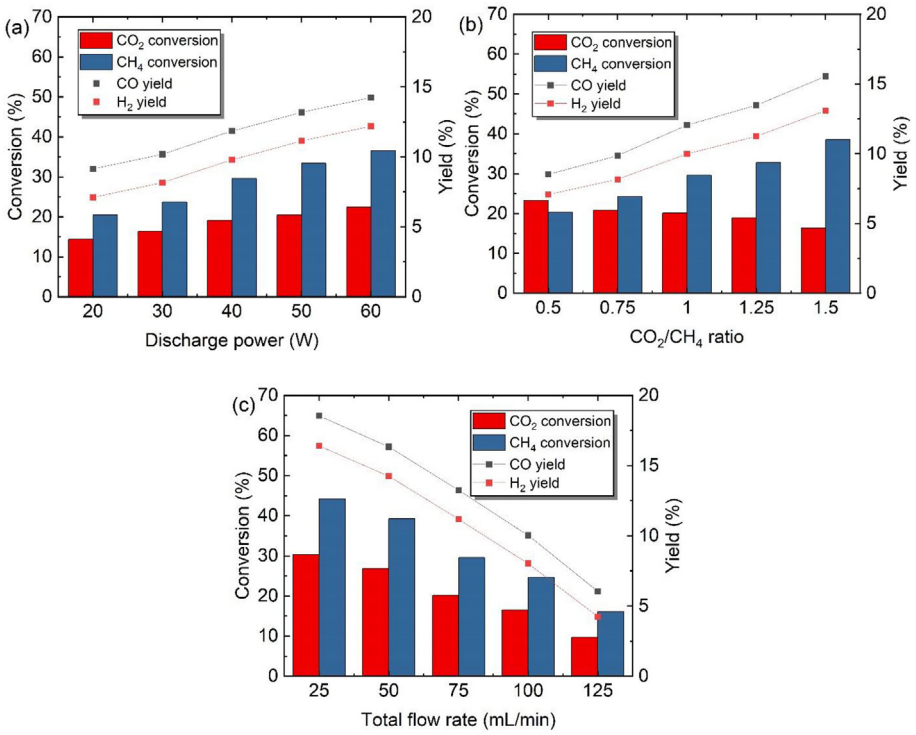

Fig. 6. Predicted effects of each operating parameter on the performance of plasma-catalytic DRM process using 7.5 wt% Ni/Al_2O_3. (a) Discharge power (CO_2/CH_4 = 1:1, total flow rate = 75 mL/min); (b) CO_2/CH_4 ratio (Discharge power = 40 W, total flow rate = 75 mL/min); (c) total flow rate (Discharge power = 40 W, CO_2/CH_4 = 1:1).

within the examined discharge power range (20–60 W). The discharge power was modulated by changing applied voltage at a fixed frequency (10 kHz); however, the increase of the applied voltage did not modify the breakdown voltage. Consequently, the average electric field and average electron energy remained constant in the tested range, as demonstrated in our previous work [50]. Nevertheless, previous studies have also demonstrated that a higher applied voltage produces more microdischarges and higher current intensity [51–53]. Increasing the discharge power in a typical DBD reactor promotes the formation of more microdischarges, generating a greater number of reaction channels and reactive species, including CH_x, O, and H radicals. These species actively participate in both plasma-induced gas-phase and catalytic surface reactions, thereby promoting the production of CO and H_2 [16].

Fig. 6(b) shows the effect of CO_2/CH_4 ratio on plasma-catalytic DRM reaction. Increasing the CO_2/CH_4 ratio enhanced the CH_4 conversion while reducing the CO_2 conversion due to chemical equilibrium. Both CO and H_2 yields increased with increasing CO_2/CH_4 ratio, but the rate of increase was not the same. At a CO_2/CH_4 ratio of 0.5, the CO yield surpassed the H_2 yield by 1.5%, but at a CO_2/CH_4 ratio of 1.5, the CO yield exceeded the H_2 yield by 2.8%. This phenomenon is likely due to the increased occurrence of the reverse water shift gas (RWGS) reaction at higher CO_2/CH_4 ratios. This reaction consumes hydrogen but also promotes the gasification of carbonaceous deposits, which reduces the possibility of catalyst deactivation. Indeed, previous studies have also shown that carbon deposition and catalyst deactivation can easily occur at CO_2/CH_4 ratios below 1 [54]. By combining these findings with the predictions from our hybrid model, a CO_2/CH_4 ratio of 1–1.5 is recommended. This ratio balances CH_4 conversion and H_2 selectivity, while also minimizing the possibility of carbon deposition. This optimal CO_2/CH_4 ratio aligns with the typical ratio found in biogas, which favors the application of plasma-catalytic DRM.

Fig. 6(c) illustrates the significant reduction in the conversion of CO_2 and CH_4 as the total flow rate increases. Specifically, the CO_2 conversion decreased from 30.2% to 9.7%, while the CH_4 conversion declined from 44.5% to 16.0% when increasing the total gas flow rate from 25 to 125 mL/min. This trend can be attributed to an 80% reduction in the residence time of the reactant gases in the discharge region as the total flow rate increases. The shorter residence time reduces the likelihood of collisions between the reactants (CH_4 and CO_2) and active species, including energetic electrons, excited species, and radicals (e.g., CH_x, O, and H), thereby lowering the conversion of CO_2 and CH_4. Consequently, the yields of CO and H_2 decrease by up to 70%.

3.2.3. Comprehensive effect of the parameters

The performance of plasma-catalytic DRM process is not only affected by individual parameters, such as discharge power, CO_2/CH_4 ratio, and total flow rate, but also by the combined effect of these parameters. The predicted results are shown in Fig. 7 and the comparison between predicted and (available) experimental values is presented in Fig. S5.

Fig. 7(a and b) indicates that a discharge power of over 50 W was required to achieve a CO_2 conversion of over 30% in the optimal range of CO_2/CH_4 ratio (0.5–1) and total flow rate (25–50 mL/min). The highest achieved CO_2 conversion was predicted to be 39.7% at a discharge power of 60 W, with a CO_2/CH_4 ratio of 0.5 and a total flow rate of 25 mL/min. On the other hand, CH_4 conversion of over 50% was achieved when the discharge power exceeded 50 W, within a range of CO_2/CH_4 ratio from 1.25 to 1.5 and a range of total flow rates from 25 to 50 mL/min. Notably, the highest CH_4 conversion of 58.6% was predicted under the condition of a discharge power of 60 W, a CO_2/CH_4 ratio of 1.5, and a total flow rate of 25 mL/min. The trend of CH_4 conversion followed a similar pattern to that of the CO_2 conversion with the discharge power and the total flow rate. Both CO_2 and CH_4 conversion are predicted to exhibit higher sensitivity to the CO_2/CH_4 ratio in the lower regions of the 3D contour representing total flow rate and discharge power. As discussed in Section 3.2.2, it is reasonable to maintain the CO_2/CH_4 ratio around 1–1.5, which is also a typical range for renewable biogas feedstock.

Fig. 7(c and d) shows that both the predicted CO yield and H_2 yield can exceed 20% in the optimal range of discharge power (50–60 W), CO_2/CH_4 ratio (1.25–1.5), and total flow rate (25–50 mL/min). Remarkably, the highest CO yield of 24.5% and H_2 yield of 21.1% were predicted at a discharge power of 60 W, a CO_2/CH_4 ratio of 1.5, and a total flow rate of 25 mL/min. Furthermore, the influence of the three parameters on both CO yield and H_2 yield follows a consistent trend. Additionally, the model prediction suggests that at lower total flow rates, increasing the discharge power has a more significant effect on enhancing CO and H_2 yields. Conversely, at higher discharge power, increasing the total flow rate leads to a more significant decrease in CO and H_2 yields. These findings indicate the significant interplay between total flow rate and discharge power within the investigated range.

To evaluate the cost-effectiveness of the process, the model employed two metrics: energy yield (EY) and fuel product efficiency (FPE). These two indicators reflect the efficiency of the plasma DRM process (Fig. 7e and f). Within the investigated range, the total flow rate emerged as the most significant factor affecting both indicators, while the effect of CO_2/CH_4 was relatively minor, especially at higher discharge power. Both EY and FPE were predicted to decrease with increasing discharge power, with a substantial drop observed at lower discharge power (20–40 W). At higher discharge power, the decline tended to plateau. The optimal flow rate was predicted to strongly depend on discharge power, with the optimal total flow rate increasing as discharge power increases. For instance, at a discharge power of 20 W, the optimal predicted EY and FPE values were achieved around a total flow rate of 50 mL/min. Conversely, for a discharge power exceeding 40 W, the optimal flow rate was predicted around 100 mL/min, further highlighting the interaction between discharge power and total flow rate. The optimal EY of 0.398 mmol/kJ was predicted at a discharge power of 20 W, CO_2/CH_4 ratio of 1.5, and total flow rate of 51.3 mL/min. Similarly, the highest FPE of 13.2% was predicted at a discharge power of 20 W, CO_2/CH_4 ratio of 1.5, and total flow rate of 53.3 mL/min.

These findings also suggest that there is a balance between reaction efficiency (conversion and yield) and cost-effectiveness (EY and FPE) for the optimal overall performance under the same operating parameters. Specifically, a high discharge power promotes CO_2 conversion, CH_4 conversion, CO yield, and H_2 yield when keeping the other parameters constant; however, the promotion comes at the cost of lower EY and FPE. Alternatively, a moderate or high total flow rate can lead to higher EY and FPE, but the corresponding reactant conversion and product yield are prohibited due to the reduction of the residence time.

3.3. Relative significance of different parameters

To understand the relative significance of various factors affecting the plasma-catalytic DRM process, we analyzed the predicted importance of each input parameter using the hybrid ML model. In conventional studies, sensitivity analysis is a valuable tool for understanding how different parameters influence the output of a process. While this traditional approach is undeniably useful and can be conducted independently of any model, the hybrid ML model offers distinct advantages, including computational efficiency, reliability with small sample sizes, the ability to identify complex parameter interactions, and the ability to directly mea-

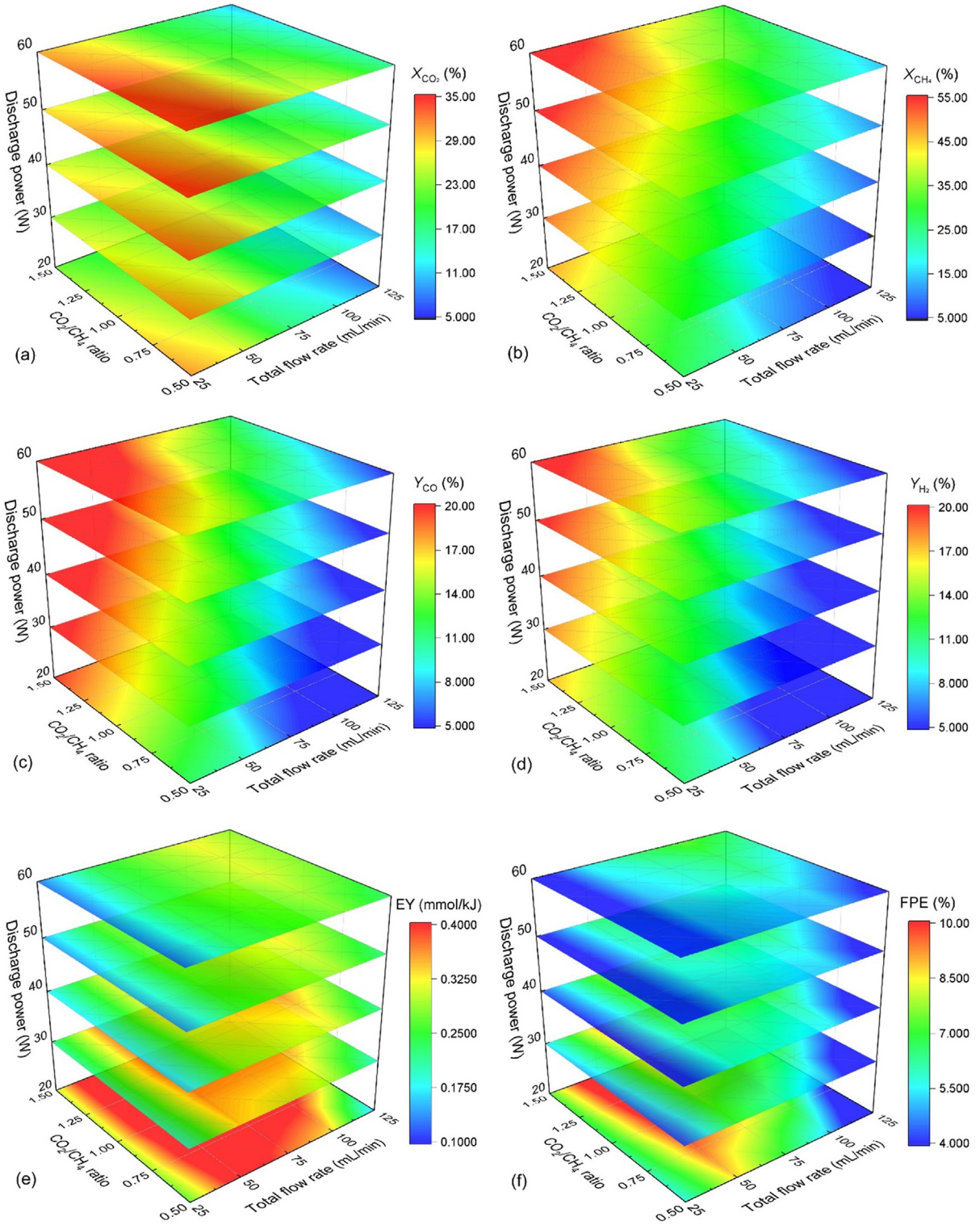

Fig. 7. Predicted interaction effects of discharge power, CO_2/CH_4 ratio, and total flow rate on the performance of plasma-catalytic DRM process using 7.5 wt% Ni/Al_2O_3. (a) CO_2 conversion; (b) CH_4 conversion; (c) CO yield; (d) H_2 yield; (e) EY; (f) FPE.

sure feature importance [55]. As proven in Section 3.1, the hybrid learning model provides robust predictions even in the context of small datasets, thereby enhancing the reliability of the conclusions drawn about parameter significance.

Fig. 8 shows that the total flow rate had the highest impact on CO_2 conversion (47.5%) and CH_4 conversion (37.6%), as well as CO yield (40.0%) and H_2 yield (47.5%). Additionally, both CO_2/CH_4 ratio and discharge power also had a considerable influence (>20%) on the yield of CO and H_2. In terms of EY, the total flow rate was again the most influential parameter (48.3%), compared to the other three process parameters. For the FPE, the discharge power, CO_2/CH_4 ratio, and total flow rate exhibited similar influences, all predicted within the range of 27.5%–34.7%, but the total flow rate again had the highest significance. Considering both EY

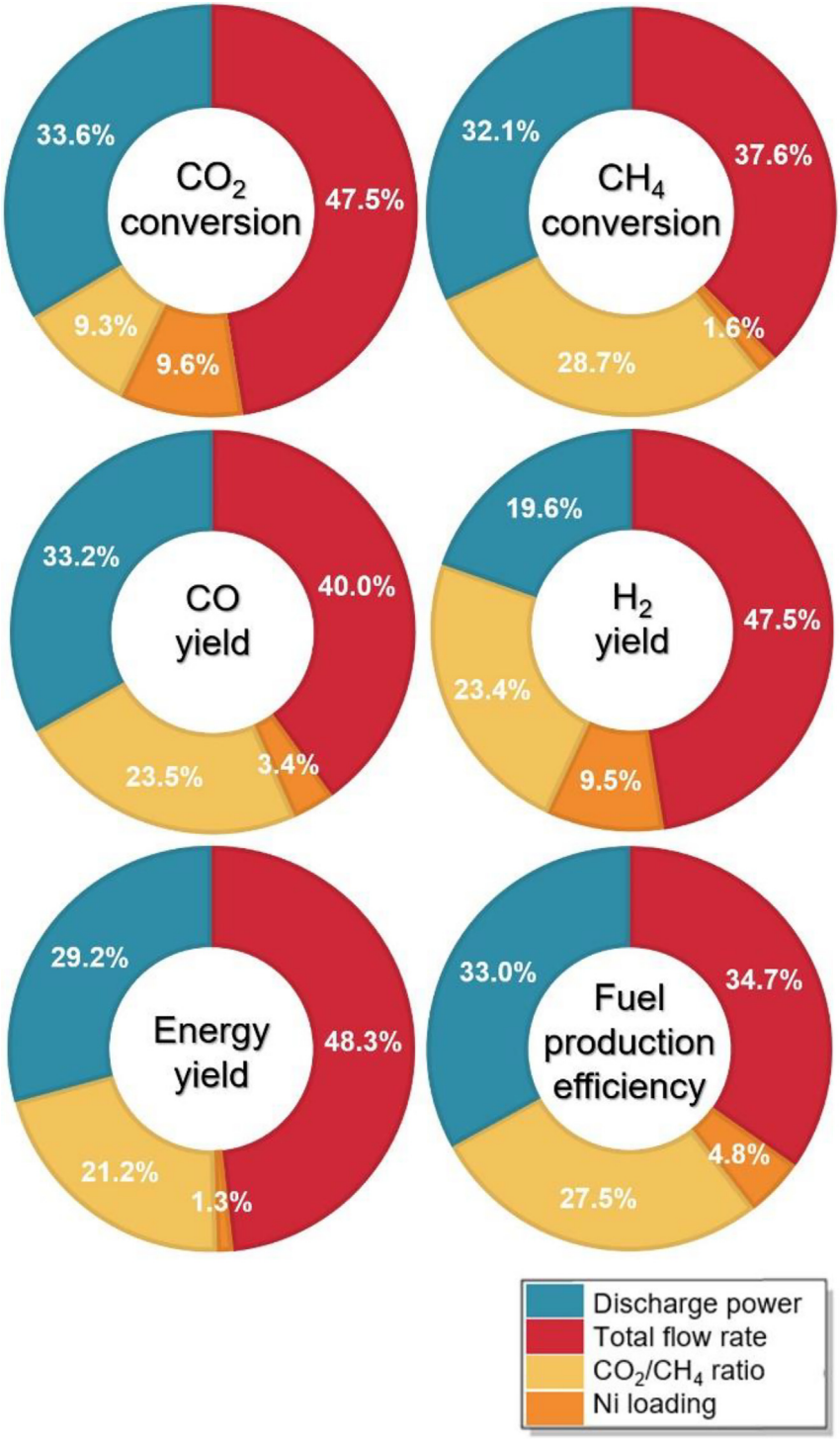

Fig. 8. Relative significance of operating parameters on the performance of plasma-catalytic DRM process.

and FPE, it is clear that the total flow rate plays a crucial role in determining the overall cost-effectiveness of the plasma-catalytic DRM process. As mentioned earlier, the parameters also involve a trade-off between reaction performance and cost-effectiveness, making the total flow rate a critical operating parameter to select for optimal performance. Notably, the significance of nickel loading amount was less than 10% for all performance indicators. These findings suggest that the plasma parameters and plasma properties exert a more significant influence within the investigated range, while minor changes in nickel loading amount (5–15 wt%) would not substantially affect the overall process performance.

4. Conclusions

This study developed a hybrid ML model to predict and optimize the plasma-catalytic DRM process. The model combined three common supervised learning algorithms (ANN, SVR, and RT) with GA for hyperparameter optimization, enabling effective

use of limited experimental data. The predictions of the hybrid model showed a strong correlation with the experimental data, with R^2 values above 0.98 for each output. A comprehensive analysis of the operating parameters revealed significant interactive effects between them, especially between discharge power and total flow rate. Furthermore, the hybrid ML model successfully identified the optimal conditions for achieving the highest EY and FPE within the continuous data space. According to the model, the highest EY of 0.398 mmol/kJ was achieved at a discharge power of 20 W, CO_2/CH_4 ratio of 1.5, and total flow rate of 51.3 mL/min. Similarly, the highest FPE of 13.2% was predicted at a discharge power of 20 W, CO_2/CH_4 ratio of 1.5, and total flow rate of 53.3 mL/min. The significance analysis revealed that total flow rate, discharge power, and CO_2/CH_4 ratio are the most significant parameters affecting the reaction performance, while the Ni loading is relatively less influential. The hybrid model demonstrates outstanding capability in extracting valuable insights relationships between input parameters and key performance metrics from limited datasets, enabling the development and optimization of complex nonlinear systems such as plasma-catalytic chemical processes within the continuous data space.

This study demonstrates the potential of the hybrid ML model as an accurate, effective, and fast tool for predicting and optimizing the plasma-catalytic DRM process, even with a relatively small training set. Notably, while this ML model shows excellent prediction ability, it cannot provide insights into the underlying reaction mechanisms, which require detailed plasma chemical kinetics models, as demonstrated in ref. [34]. Nevertheless, this ML model holds enormous potential for application in a broad spectrum of plasma-based chemical processes.

The ML strategy developed in this study also holds potential for application in predicting the performance of thermal catalytic DRM processes. While catalyst structure-activity relationships and condition optimization in thermal catalysis have been extensively investigated, leading to different model focal points (e.g., big data and meta-analysis with ML algorithm), our hybrid ML model approach provides valuable insights from limited experimental data with greater consistency and accuracy, particularly promising for the optimization of industrial-scale DRM process. This is an exciting and worthwhile direction for future research, underscoring the broad applicability and relevance of our ML application in catalysis.

CRediT authorship contribution statement

Yuxiang Cai: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Danhua Mei:** Formal analysis, Investigation, Writing – review & editing. **Yanzhen Chen:** Formal analysis, Writing – review & editing. **Annemie Bogaerts:** Funding acquisition, Project administration, Supervision, Writing – review & editing. **Xin Tu:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This project received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 813393. D. Mei acknowledges the funding from the National Natural Science Foundation of China (No. 52177149).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jechem.2024.04.022.

References

[1] D. Pakhare, J. Spivey, Chem. Soc. Rev. 43 (2014) 7813–7837.
[2] K. Jiang, H. Zhao, Y. Chen, B. Li, Z. Zhang, F. Cao, L. Wu, Y. Tang, T. Li, L. Tan, Catal. Sci. Technol. 14 (2023) 261–266.
[3] H. Zhao, R. Yu, S. Ma, K. Xu, Y. Chen, K. Jiang, Y. Fang, C. Zhu, X. Liu, Y. Tang, L. Wu, Y. Wu, Q. Jiang, P. He, Z. Liu, L. Tan, Nat. Catal. 5 (2022) 818–831.
[4] Y. Song, E. Ozdemir, S. Ramesh, A. Adishev, S. Subramanian, A. Harale, M. Albuali, B.A. Fadhel, A. Jamal, D. Moon, S.H. Choi, C.T. Yavuz, Science 367 (2020) 777–781.
[5] J. Requies, S. Rabe, F. Vogel, T.-B. Truong, K. Filonova, V.L. Barrio, J.F. Cambra, M. B. Güemez, P.L. Arias, Catal. Today 143 (2009) 9–16.
[6] U. Guharoy, T.R. Reina, J. Liu, Q. Sun, S. Gu, Q. Cai, J. CO2 Util. 53 (2021) 101728.
[7] Y. Xue, L. Xu, M. Chen, C. Wu, G. Cheng, N. Wang, X. Hu, Catal. Sci. Technol. 11 (2021) 6344–6368.
[8] Z. Li, Q. Lin, M. Li, J. Cao, F. Liu, H. Pan, Z. Wang, S. Kawi, Renew. Sust. Energ. Rev. 134 (2020) 110312.
[9] W.-J. Jang, J.-O. Shim, H.-M. Kim, S.-Y. Yoo, H.-S. Roh, Catal. Today 324 (2019) 15–26.
[10] N.A.K. Aramouni, J.G. Touma, B.A. Tarboush, J. Zeaiter, M.N. Ahmad, Renew. Sustain. Energy Rev. 82 (2018) 2570–2585.
[11] B. Fidalgo, A. Domínguez, J.J. Pis, J.A. Menéndez, Int. J. Hydrog. Energy 33 (2008) 4337–4344.
[12] J. Guo, H. Lou, H. Zhao, D. Chai, X. Zheng, Appl. Catal. a: Gen. 273 (2004) 75–82.
[13] M. Usman, W.M.A. Wan Daud, H.F. Abbas, Renew. Sustain. Energy Rev. 45 (2015) 710–744.
[14] A. Bogaerts, X. Tu, J.C. Whitehead, G. Centi, L. Lefferts, O. Guaitella, F. Azzolina-Jury, H.-H. Kim, A.B. Murphy, W.F. Schneider, T. Nozaki, J.C. Hicks, A. Rousseau, F. Thevenet, A. Khacef, M. Carreon, J. Phys. d: Appl. Phys. 53 (2020) 443001.
[15] A. George, B. Shen, M. Craven, Y. Wang, D. Kang, C. Wu, X. Tu, Renew. Sust. Energ. Rev. 135 (2021) 109702.
[16] A. Bogaerts, T. Kozák, K. van Laer, R. Snoeckx, Faraday Discuss. 183 (2015) 217–232.
[17] A. Bogaerts, E.C. Neyts, ACS Energy Lett. 3 (2018) 1013–1027.
[18] B. Wanten, S. Maerivoet, C. Vantomme, J. Slaets, G. Trenchev, A. Bogaerts, J. CO2 Util. 56 (2022) 101869.
[19] A. Aziznia, H.R. Bozorgzadeh, N. Seyed-Matin, M. Baghalha, A. Mohamadalizadeh, J. Nat. Gas Chem. 21 (2012) 466–475.
[20] J. Martin-del-Campo, M. Uceda, S. Coulombe, J. Kopyscinski, J. CO2 Util. 46 (2021) 101474.
[21] H. Zhang, L. Li, X. Li, W. Wang, J. Yan, X. Tu, J. CO2 Util. 27 (2018) 472–479.
[22] S.M. Chun, D.H. Shin, S.H. Ma, G.W. Yang, Y.C. Hong, Catalysts 9 (2019) 292.
[23] K. Pornmai, W. Ngamkala, T. Rirksomboon, P. Ouraipryvan, S. Chavadej, Ind. Eng. Chem. Res. 58 (2019) 6203–6217.
[24] D.H. Mei, S.Y. Liu, X. Tu, J. CO2 Util. 21 (2017) 314–326.
[25] B. Ashford, Y. Wang, C.-K. Poh, L. Chen, X. Tu, Appl. Catal. B: Environ. 276 (2020) 119110.
[26] M. Ronda-Lloret, Y. Wang, P. Oulego, G. Rothenberg, X. Tu, N.R. Shiju, ACS Sustainable Chem. Eng. 8 (2020) 17397–17407.
[27] Y.X. Zeng, L. Wang, C.F. Wu, J.Q. Wang, B.X. Shen, X. Tu, Appl. Catal. b: Environ. 224 (2018) 469–478.
[28] Y. Wang, L. Yao, S. Wang, D. Mao, C. Hu, Fuel Process. Technol. 169 (2018) 199–206.
[29] Z. Song, Q. Wang, C. Guo, S. Li, W. Yan, W. Jiao, L. Qiu, X. Yan, R. Li, Ind. Eng. Chem. Res. 59 (2020) 17250–17258.
[30] L. Brune, A. Ozkan, E. Genty, T.V. de Bocarmé, F. Reniers, J. Phys. d: Appl. Phys. 51 (2018) 234002.
[31] R. Vakili, R. Gholami, C.E. Stere, S. Chansai, H. Chen, S.M. Holmes, Y. Jiao, C. Hardacre, X. Fan, Appl. Catal. b: Environ. 260 (2020) 118195.
[32] E. Jwa, S.B. Lee, H.W. Lee, Y.S. Mok, Fuel Process. Technol. 108 (2013) 89–93.
[33] B. Wanten, R. Vertongen, R. De Meyer, A. Bogaerts, J. Energy Chem. 86 (2023) 180–196.
[34] B. Loenders, R. Michiels, A. Bogaerts, J. Energy Chem. 85 (2023) 501–533.
[35] A. Salden, M. Budde, C.A. Garcia-Soto, O. Biondo, J. Barauna, M. Faedda, B. Musig, C. Fromentin, M. Nguyen-Quang, H. Philpott, G. Hasrack, D. Aceto, Y. Cai, F.A. Jury, A. Bogaerts, P. Da Costa, R. Engeln, M.E. Gálvez, T. Gans, T. Garcia, V. Guerra, C. Henriques, M. Motak, M.V. Navarro, V.I. Parvulescu, G. Van Rooij, B. Samojeden, A. Sobota, P. Tosi, X. Tu, O. Guaitella, J. Energy Chem. 86 (2023) 318–342.
[36] X. Zhu, X. Tu, D. Mei, C. Zheng, J. Zhou, X. Gao, Z. Luo, M. Ni, K. Cen, Chemosphere 155 (2016) 9–17.
[37] J. Sun, Q. Chen, Y. Guo, Z. Zhou, Y. Song, J. Energy Chem. 46 (2020) 133–143.
[38] S. Li, J. Sun, Y. Gorbanev, K. van't Veer, B. Loenders, Y. Yi, T. Kenis, Q. Chen, A. Bogaerts, ACS Sustain. Chem. Eng. 11 (2023) 15373–15384.
[39] S.Y. Liu, D.H. Mei, Z. Shen, X. Tu, J. Phys. Chem. C 118 (2014) 10686–10693.

[40] X. Zhu, S. Liu, Y. Cai, X. Gao, J. Zhou, C. Zheng, X. Tu, Appl. Catal. b: Environ. 183 (2016) 124–132.
[41] T. Chang, J. Lu, Z. Shen, Y. Huang, D. Lu, X. Wang, J. Cao, R. Morent, Appl. Catal. b: Environ. 244 (2019) 107–119.
[42] J. Pan, Y. Liu, S. Zhang, X. Hu, Y. Liu, T. Shao, Energy Convers. Manag. 277 (2023) 116620.
[43] Y. Wang, Z. Liao, S. Mathieu, F. Bin, X. Tu, J. Hazard. Mater. 404 (2021) 123965.
[44] D. Mei, B. Ashford, Y.-L. He, X. Tu, Plasma Process. Polym. 14 (2017) 1600076.
[45] S.N. Sivanandam, S.N. Deepa, Introduction to Genetic Algorithms, Springer, Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 15–37.
[46] P. Sedgwick, BMJ 345 (2012) e4483.
[47] Sk. Mahammadunnisa, P. Manoj Kumar Reddy, B. Ramaraju, Ch. Subrahmanyam, Energy Fuels 27 (2013) 4441–4447.
[48] N. Sahli, C. Petit, A.C. Roger, A. Kiennemann, S. Libs, M.M. Bettahar, Catal. Today 113 (2006) 187–193.
[49] Q. Wang, B.-H. Yan, Y. Jin, Y. Cheng, Energy Fuels 23 (2009) 4196–4201.
[50] X. Tu, H.J. Gallon, M.V. Twigg, P.A. Gorry, J.C. Whitehead, J. Phys. d: Appl. Phys. 44 (2011) 274007.
[51] N. Rahemi, M. Haghighi, A.A. Babaluo, M.F. Jafari, S. Khorram, Int. J. Hydrog. Energy 38 (2013) 16048–16061.
[52] Y. Diao, X. Zhang, Y. Liu, B. Chen, G. Wu, C. Shi, Appl. Catal. B. Environ. 301 (2022) 120779.
[53] X. Dong, T. Ma, Q. Liu, Y. Yu, M. Zhang, Appl. Surf. Sci. 492 (2019) 843–848.
[54] J.-M. Lavoie, Front. Chem. 2 (2014) 81.
[55] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis: The Primer, John Wiley & Sons, 2008.