Journal of Hazardous Materials 404 (2021) 123965

Contents lists available at ScienceDirect

Journal of Hazardous Materials

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

# Prediction and evaluation of plasma arc reforming of naphthalene using a hybrid machine learning model

Yaolin Wang [a], Zinan Liao [a], Stéphanie Mathieu [a], Feng Bin [a,b,*], Xin Tu [a,*]

[a] Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, UK
[b] State Key Laboratory of High-Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

*Keywords:*
Machine learning
Non-thermal plasma
Biomass gasification
Tar reforming
Naphthalene

ABSTRACT

We have developed a hybrid machine learning (ML) model for the prediction and optimization of a gliding arc plasma tar reforming process using naphthalene as a model tar compound from biomass gasification. A linear combination of three well-known algorithms, including artificial neural network (ANN), support vector regression (SVR) and decision tree (DT) has been established to deal with the multi-scale and complex plasma tar reforming process. The optimization of the hyper-parameters of each algorithm in the hybrid model has been achieved by using the genetic algorithm (GA), which shows a fairly good agreement between the experimental data and the predicted results from the ML model. The steam-to-carbon (S/C) ratio is found to be the most critical parameter for the conversion with a relative importance of 38%, while the discharge power is the most influential parameter in determining the energy efficiency with a relative importance of 58%. The coupling effects of different processing parameters on the key performance of the plasma reforming process have been evaluated. The optimal processing parameters are identified achieving the maximum tar conversion (67.2%), carbon balance (81.7%) and energy efficiency (7.8 g/kWh) simultaneously when the global desirability index $I_2$ reaches the highest value of 0.65.

## 1. Introduction

Biomass has been regarded as one of the most important renewable energy sources for meeting the increasing energy demand as well as the mitigation of global climate change. Biomass gasification provides a sustainable and promising technical route for the production of high-value syngas (a mixture of $H_2$ and CO), which can be utilized to produce electricity, heat, chemicals, and fuels. However, the formation of undesirable contaminates (e.g. tars) in syngas remains a significant challenge to advance the gasification technology for large scale applications. Tar, consisting of a range of aromatic hydrocarbons with rings, causes serious operational problems, including blocking, corrosion, and crashing the whole system. These tars contain a range of polycyclic aromatic hydrocarbons (PAHs), some of which can be toxic and carcinogenic. The content of tars in the syngas can range from 1 g/Nm$^3$ up to 100 g/Nm$^3$, limiting the use of syngas for gas turbines, fuel cells and chemical/fuel synthesis. Therefore, effective removal of tars in biomass gasification to produce high-quality and clean syngas is critical for the biomass and bioenergy industries.

Non-thermal plasma (NTP) has been regarded as a promising and emerging technology for the synthesis of fuels and chemicals at low temperatures and ambient pressure (Bogaerts et al., 2020; George et al., 2021; Liu et al., 2019, 2017; Wang et al., 2019; Mei et al., 2019; Craven et al., 2020). Plasma processes can generate abundant energetic electrons and reactive species, which can break chemical bonds and rings of tar compounds, enabling thermodynamically unfavourable tar reforming to proceed at low temperatures. In addition, plasma processes can be operated responsively with a quick start-up and shut-down, which offers the flexibility for integration with renewable energy sources (e.g., wind and solar power) for distributed chemical energy storage, especially when using surplus or intermittent electricity from renewable energy during peak moments on the grid. Up till now, the reported experimental studies mainly focused on the reforming of a range of tar model compounds, including toluene (Zhang et al., 2019; Trushkin et al., 2013; Liu et al., 2017; Saleem et al., 2019a, 2019b; Zhu et al., 2020), benzene (Chun et al., 2012; Gao et al., 2016; Mao et al., 2018), naphthalene (Wang et al., 2019; Wu et al., 2017; Liu et al., 2019; Mei et al., 2019), anthracene (Chun et al., 2011), and phenanthrene (Kong et al., 2019), using different plasma sources (e.g., dielectric barrier discharge (DBD) (Liu et al., 2018; Xu et al., 2018; Wu et al., 2018), gliding arc discharge (GAD) (Liu et al., 2017;

* Corresponding authors.
*E-mail addresses:* binfeng@imech.ac.cn (F. Bin), xin.tu@liverpool.ac.uk (X. Tu).

https://doi.org/10.1016/j.jhazmat.2020.123965
Received 18 July 2020; Received in revised form 1 September 2020; Accepted 1 September 2020
Available online 16 September 2020

Zhang et al., 2019; Yang and Chun, 2011; Du et al., 2007; Yu et al., 2010; Nunnally et al., 2014; Zhu et al., 2016), and microwave (MW) plasma (Jamróz et al., 2018; Sun et al., 2018), etc.). However, plasma reforming of tars is a complex process and the process performance is influenced by a wide range of processing parameters (Młotek et al., 2020; Tao et al., 2013). Most of the previous works have simply investigated the effect of an individual processing parameter on the plasma tar reforming, while no efforts have been dedicated to investigating the coupling effects of various parameters on the plasma reforming of tars, which is critical for the further optimization and development of plasma reforming technologies. Although plasma modeling can be helpful to understand the plasma reforming process to some extents, developing a plasma chemical kinetic model to predict and understand the plasma tar reforming process remains a significant challenge as key reaction data (e.g., cross-section, rate constant) for large carbon molecules is not available. In addition, the multi-length-scale complexity in the plasma tar reforming process cannot be fully captured by plasma modeling.

Recently, machine learning (ML) has attracted increasing interest as a powerful tool for the prediction and optimization of a range of chemical processes. The use of ML for plasma gas conversions and air pollution control has also been explored. (Liu et al., 2014; Zhu et al., 2016; Istadi and Amin, 2007; Chang et al., 2019; Ye et al., 2020) Tu and co-workers developed a three-layer back-propagation (BP) artificial neural network (ANN) for the prediction of a complex plasma non-oxidative methane conversion process (Liu et al., 2014). The discharge power was found to be the most important parameter in the plasma conversion of methane, while the excitation frequency of the plasma system was identified as the least important parameter affecting the plasma process. Zhu et al. (2016) developed a three-layer ANN model to get new insights into the effect of different operating parameters on the post-plasma catalytic removal of methanol. The catalyst composition (i.e. Mn/Ce molar ratio of Mn-Ce oxide catalysts) was found to be the most critical parameter in determining the methanol removal efficiency. Chang et al. developed an ANN model to better understand the effect of four experimental parameters (discharge power, initial concentration of toluene, flow rate, and relative humidity) on the post-plasma catalytic removal of toluene over $MnCoO_x/\gamma$-$Al_2O_3$ catalysts (Chang et al., 2019). A multi-objective optimization model was proposed to determine the optimal experimental parameters using a genetic algorithm (GA). However, the use of ML for plasma chemical processes is still limited.

A variety of supervised ML algorithms are available for the prediction of chemical processes, including plasma-based processes. Each algorithm has its strengths and weaknesses, and there is no single ML algorithm that works best on all supervised learning problems, including regression and prediction problems. Artificial neural networks, are one of the classic supervised learning algorithms (Sun et al., 2016), with the advantages of self-adaptation, self-configuration, and self-learning. ANN usually requires a relatively large set of training data to avoid overfitting and enhance the generalization capability. However, the experimental data sets are often relatively small in plasma-based chemical processes as the process parameters in the experiments are often varied within a narrow range due to the complexity of the plasma chemical processes, and the relatively small data sets might lead toinstability of the predicted result of the plasma process when using the ANN.

Meanwhile, other supervised learning algorithms, such as support vector machine (SVM) and decision tree (DT), are also well-known for solving classification and regression problems (Smola and Schölkopf, 2004; Drucker et al., 1997). A SVM algorithm can create a hyperplane to classify the data sets into different categories, and then optimize the hyperplane by maximizing the margin between each data set and the hyperplane. This algorithm can efficiently handle both non-linear and high-dimensional regression problems, and is also known as support vector regression (SVR). It has excellent generalization capability to avoid overfitting of training data. However, the SVR algorithm often requires a relatively long training time with an unsatisfactory performance if the data sets are boisterous, which is likely to be the case for plasma chemical reactions. DT can solve regression problems by transforming a data set into a visualized tree representation, with an excellent noise tolerance character and strong robustness (Safavian and Landgrebe, 1991). Thus, the DT algorithm has the advantages required to deal with data sets with either limited samples or incomplete data. Compared to other ML algorithms, DT does not need pre-processing of data sets, such as normalization and standardization. However, the DT algorithm can easily cause over-fitting problems due to its weak generalization ability. It is also not effective at handling a complex system with large-scale data sets.

Clearly, choosing an appropriate algorithm is essential to achieve the best performance of the ML model in solving a specific problem (e.g. prediction, optimization), especially when dealing with multi-scale, non-linear, and complex chemical processes such as a plasma tar reforming process. In the practical experiments, we often encounter some difficulties, such as long cycle time, complicated procedures, and susceptibility to environmental interference, all of which cost us more time and resources to get reliable results. In addition, experimental complexity scales exponentially with the number of variables, restricting the number of experiments and narrowing the range of process parameters. In some cases, the experimental complexity might lead to issues, such as large random error, missing data, and bad points, etc., these make it difficult for the training of experimental data using a single ML algorithm. Hence, developing a hybrid ML model offers a promising solution for the fast and effective prediction and optimization of multi-scale and complex plasma processes with limited experimental data (Shao and Lunetta, 2012). However, up till now, only ANN has been used for plasma-based chemical processes, while the use of a hybrid ML model for the optimization of plasma chemical processes has not been reported yet.

In this work, a novel hybrid ML model combining ANN, SVR, and DT algorithms has been developed to predict and evaluate the critical performance of the plasma steam reforming of tar for the first time. Compared to the ML model using a single algorithm, the hybrid model combining three different algorithms can enhance the robustness and generalization ability of the ML model for fast and effective predictions. In addition, a genetic algorithm has been used to determine the optimal hyper-parameters of each algorithm (ANN, SVR, and DT) to enhance the adaptive ability and increase the accuracy of the prediction. Based on the well-trained hybrid model, the plasma tar reforming process has been analyzed to get new insights into the effect and relative importance of different process parameters, including discharge power, steam-to-carbon (S/C) ratio, and naphthalene concentration on the plasma reaction performance in terms of three key performance indicators (KPIs) including tar conversion, carbon balance and energy efficiency. The coupling effects of any two process parameters on the KPIs of the plasma process have also been evaluated. Furthermore, two indexes have been introduced to determine the optimal process parameters to maximize the key performance of the plasma tar reforming process.

## 2. Methodology and experiment

### 2.1. Experimental method

A GAD reactor with two stainless-steel electrodes has been developed for plasma tar reforming (Wang et al., 2019). The electrode gap was fixed at 2 mm for the initiation of the arc. Naphthalene (powder) was vaporized at 60–75 °C using a water bath and then mixed with nitrogen. The mixed naphthalene and nitrogen flow (with or without steam) was preheated to 200 °C before being injected into the GAD reactor. The total gas flow rate was maintained at 4 L/min, and the inlet concentration of naphthalene can be controlled between 1.1 mg/L and 1.7 mg/L. The S/C ratio can be varied between 0 and 4.0 to investigate the role of steam in the plasma reforming of naphthalene. The GAD reactor was connected to an alternating current high voltage transformer (10 kV/50 Hz). The

electrical signals (arc voltage U and arc current I) were recorded using a digital oscilloscope (Tektronix, MDO3024). The discharge power was determined through the integration of arc voltage and arc current in a cycle T.

$$\text{Discharge power(W)} = \frac{1}{T}\int_0^{t=T} U(t) \times I(t)dt \quad (1)$$

The gaseous products were measured and quantified using gas chromatography (Agilent 7820A) fitted with a flame ionization detector and a thermal conductivity detector. A Molecular Sieve 5A column (HP MOLESIEVE) was used for the measurement of CO and $H_2$, and an HP-PLOT/Q column was used to separate $CO_2$, $CH_4$, and $C_2$-$C_4$ hydrocarbons. The liquid products were condensed and dissolved in dichloromethane using a cold trap.

The conversion of naphthalene is defined as

$$X_{C_{10}H_8}(\%) = \frac{C_i - C_o}{C_i} \times 100\% \quad (2)$$

Where $C_i$ is the inlet concentration of naphthalene, and $C_o$ is the concentration of naphthalene after the plasma reaction.

The converted rate is defined as

$$R_{C_{10}H_8}(\text{mg/min}) = C_i \times X_{C_{10}H_8}(\%) \times \text{Total flow rate (L/min)} \quad (3)$$

The carbon balance of the plasma tar reforming process is given by.

Without steam

$$\text{Carbon balance}(\%) = \sum \frac{\text{Produced } C_xH_y(x = 1, 2, 3, 4)}{\text{Converted } C_{10}H_8} \times 100\% \quad (4)$$

With steam

$$\text{Carbon balance}(\%) = \sum \frac{\text{Produced } C_xH_y(x = 1, 2, 3, 4) \text{ and } CO_x(x = 1, 2)}{\text{Converted } C_{10}H_8} \times 100\% \quad (5)$$

The energy efficiency for tar conversion is defined as:

$$E(\text{g/kWh}) = \frac{\text{Mass of converted tar (g/h)}}{\text{Discharge power (kW)}} \quad (6)$$

Non-dimensional indexes, $I_1$ & $I_2$ ($0 < I_1, I_2 < 1$), are defined to evaluate the effectiveness of the plasma tar reforming process. Specifically, $I_1$ is defined as the product of the normalized (N) tar conversion and carbon balance to evaluate the effective conversion of $C_{10}H_8$ (Eq. 7). $I_2$ is defined as the product of three normalized KPIs to determine the optimal process parameters to maximize the conversion, carbon balance, and energy efficiency simultaneously (Eq. 8). All the normalized parameters are in the range of 0–1.

$$I_1 = \text{N (tar conversion)} \times \text{N (carbon balance)} \quad (7)$$

$$I_2 = \text{N (tar conversion)} \times \text{N (carbon balance)} \times \text{N (energy efficiency)} \quad (8)$$

### 2.2. Description of the hybrid model

The key reaction performance (P) of the plasma tar reforming process was simulated using a hybrid ML model via a linear combination of ANN, SVR, and DT algorithms, as shown in Eq. 9.

$$P = W_1 \times P_{ANN} + W_2 \times P_{SVR} + W_3 \times P_{RT} (0 \le W_1, W_2, W_3 \le 1) \quad (9)$$

Where $W_1$, $W_2$, and $W_3$ are the relative weight of ANN, SVR, and DT algorithms, respectively. The mean squared error (MSE), one of the most common criteria, is defined as the mean difference between the experimental data ($R_i$) and the predicted results ($P_i$),

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - R_i)^2 \quad (10)$$

In this study, the MSE can be used to evaluate the performance of the hybrid ML model in the optimization of the relative weight for each algorithm using an exhaustion method. The optimal relative weight for each algorithm in the hybrid model is achieved when the minimum MSE of the ML model is reached.

Fig. 1 shows the logical structure for the prediction, evaluation, and optimization using this hybrid ML model. Three key processing parameters, including discharge power, inlet $C_{10}H_8$ concentration, and S/C ratio are taken into account and used as the input in the model. The conversion of naphthalene, carbon balance, and energy efficiency are selected as the KPIs in this study. GA is applied to optimize the hyperparameters of these three algorithms (ANN, SVR, and DT) in the hybrid ML model, such as the learning rate for ANN, the variable C and γ (Radial Basis Function (RBF) kernel) for SVR, and the depth for DT.

#### 2.2.1. Artificial neural network

For the ANN model, Adaline (Adaptive linear neuron, Fig. 2), a single-layer neural network, was developed to simulate the plasma tar reforming process. The model includes three key process parameters as the inputs, including discharge power ($X_1$), inlet $C_{10}H_8$ concentration ($X_2$), and S/C ratio ($X_3$). Bias (b) is a basic parameter in the neural network, which can be used to adjust the output along with the weighted sum of the inputs to the neuron. The bias can be used to shift the activation function (σ(z)) and offset the predicted results, enhancing the flexibility and generalization of the ANN model. Z-score normalization has been used to pre-process all data sets, including the input process parameters and the predicted KPIs.

In this case, the cost function (J(w), Eq. 11) is determined as the sum of squared errors (SSE) between the predicted results ($\sigma(z^{(i)})$) and the experimental data ($y^{(i)}$), and can be minimized using a gradient descent algorithm to get the optimized weight for each processing parameter (input).

$$J(w) = \frac{1}{2}\sum_i \left(y^{(i)} - \sigma\left(z^{(i)}\right)\right)^2 \quad (11)$$

#### 2.2.2. Support vector regression

The SVR is a robust learning algorithm and has been widely used to solve regression problems with small samples, high-dimensions, and non-linearity. It can be formulated as an optimization problem (Eq. 12) to minimize the norm of the weight vector (w) with some slack variables ($\xi_i$ and $\xi_i^*$) introduced to increase the tolerance of regression errors. With a Lagrange dual formulation, the optimization (Eq. 12) can be done by solving its dual problem (Eq. 13).

$$\min_{w,b,\xi,\xi^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i^*),\ \xi_i,\ \xi_i^* \ge 0,\ i = 1, \ldots, m \quad (12)$$

$$\max \sum_{i=1}^{m} y_i(\alpha_i^* - \alpha_i) - \varepsilon\sum_{i=1}^{m}(\alpha_i^* + \alpha_i) - \frac{1}{2}\sum_{i,j=1}^{m}(\alpha_i^* - \alpha_i)(\alpha_i^* + \alpha_i)x_i^T x_j \quad (13)$$

Where C is the penalty term that determines the trade-off between the misclassifications of the training data and the width of the margin. $\alpha_i$ and $\alpha_i^*$ are the Lagrange multipliers, and ε is the tolerance of margin.

In this study, the RBF kernel function (Eq. 14) has been used in the SVR model to plot the process parameters and KPIs for effective and accurate prediction of the performance in the plasma tar reforming process.

$$K(x_i, x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right),\ \gamma > 0 \quad (14)$$

where γ defines the influence of a single sample on the entire classifi-

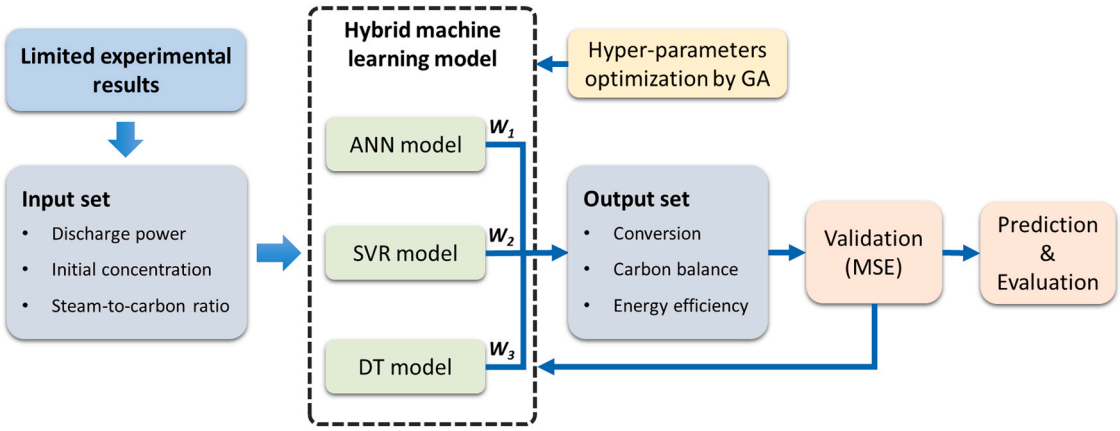


**Fig. 1.** Scheme of the logical structure for the prediction and evaluation using the hybrid ML model.

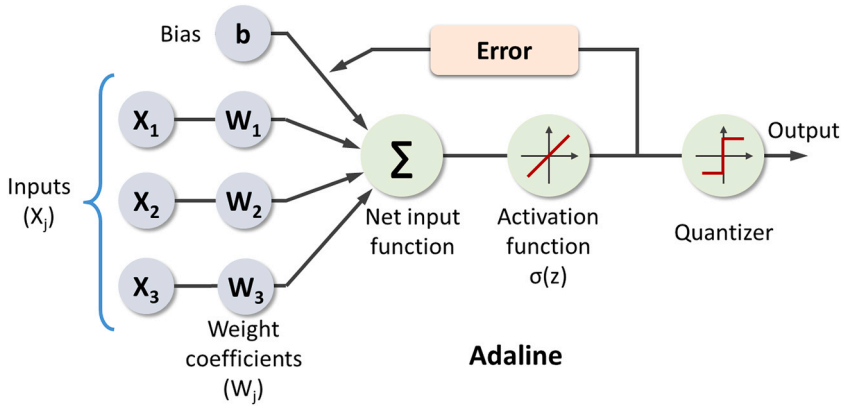


**Fig. 2.** Scheme of an adaptive linear neutron (Adaline).

cation hyperplane. Then, the generic equation of the SVR model based on the RBF kernel function can be described as:

$$f(x) = \sum_{i=1}^{m} (\alpha_i^* - \alpha_i)\, K(x_i, x) + b \quad (15)$$

Where b refers to the bias term.

GA is used to find the optimal hyper-parameters (γ and C) in the SVR training process when the minimum MSE for each KPI is reached. One hundred iterations are used in the optimization of the hyper-parameters using GA, with the hyper-parameters (γ and C) being chosen in the range of $0 < \gamma < 1$ and $0 < C < 10^4$, respectively. In this study, the optimized hyper-parameter (γ, C) of the SVR model is (1126, 0.56) for three KPIs. The SVR algorithm was developed using the "Scikit-learn" library in Python.

### 2.2.3. Decision tree

A standard binary decision tree for solving this regression problem is defined with several branches, one root, a few nodes, and leaves. Basically, one branch is a chain of nodes from root to a leaf, and each node refers to one attribute. The splitting criteria for this regression tree, which is also known as CART (classification and regression tree), is the MSE, which has been defined in Eq. 10. For each node, the algorithm will calculate the predicted value and calculate the MSE for each subset, and the regression tree will evolve by seeking the smallest MSE value. In our model, a maximum depth of the tree has been set to achieve high accuracy for the prediction of the three KPIs. The DT algorithm was also developed using the "Scikit-learn" library.

## 3. Results and discussion

### 3.1. Optimization of the hybrid ML model

Fig. 3 shows the influence of the relative weights for each algorithm on the MSE of the hybrid ML model. When the ML model uses a single algorithm, the MSE of the model is 0.110, 0.107, and 0.068 for ANN,

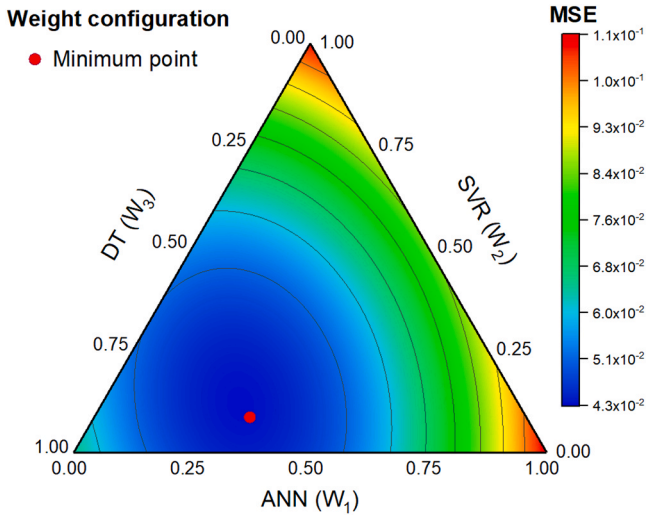


**Fig. 3.** Effect of the relative weights of ANN, SVR, and DT on the MSE for the predicted KPIs.

SVR, and DT, respectively. These values are much higher than the minimum MSE (0.043) of the hybrid model, indicating that the optimized hybrid ML model is far superior to the models using a single algorithm in this study. The optimal weight of ANN, SVR, and DT in this hybrid model is 0.31, 0.10, and 0.59, respectively. Thus, the optimal hybrid model can be expressed as:

$$\text{Hybrid model}: \ P = 0.31 \times P_{ANN} + 0.1 \times P_{SVR} + 0.59 \times P_{DT} \quad (16)$$

.

### 3.2. Output prediction and model validation

Fig. 4 shows a comparison between the experimental and simulated naphthalene conversion in the plasma reforming reaction without steam (S/C = 0). The predicted results obtained from the hybrid ML model agree reasonably well with those from the experiment. Hence, the model can be used to predict the reaction performance. For example, the conversion of naphthalene is predicted to be 72% at an inlet $C_{10}H_8$ concentration of 1.4 mg/L. At a higher tar concentration (e.g., > 1.4 g/Nm$^3$), the change of tar concentration has a weak effect on the conversion of naphthalene. By contrast, the converted rate of naphthalene was increased continuously from 3.7 to 5.6 mg/min with the increase of inlet $C_{10}H_8$ concentration from 1.1 to 2.0 g/Nm$^3$. In addition, the discharge power shows a limited effect on the conversion of naphthalene

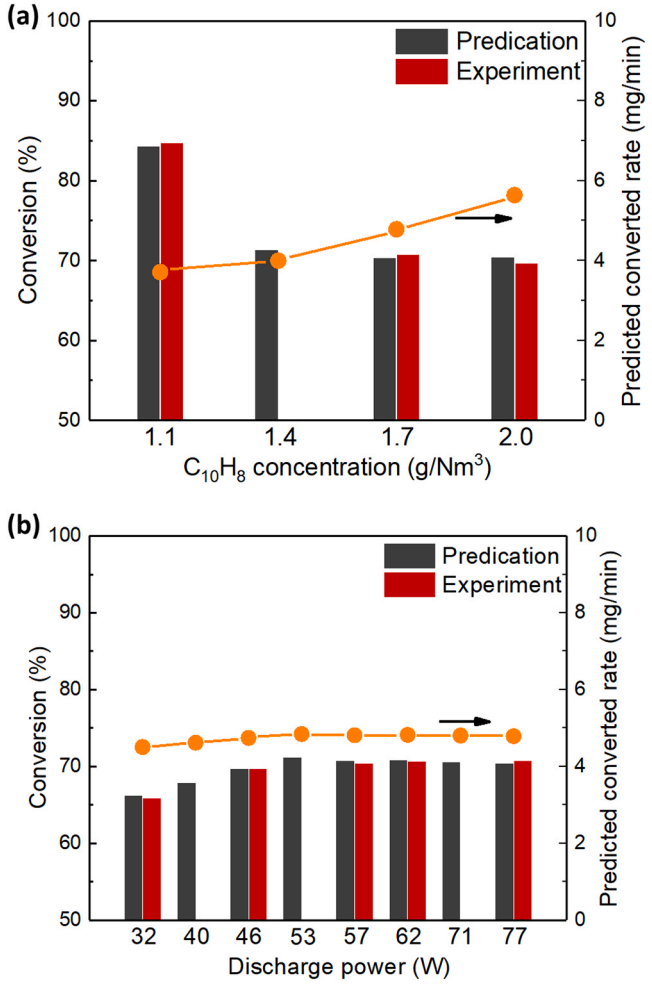


**Fig. 4.** (a) Effect of inlet $C_{10}H_8$ concentration on the conversion and predicted converted rate (discharge power = 77 W, S/C = 0); (b) Effect of discharge power on the conversion and predicted converted rate ($C_{10}H_8$ concentration = 1.7 g/Nm$^3$, S/C = 0).

in the reaction without steam. With the increase of discharge power, the predicted converted rate remained stable at around 4.7 mg/min, indicating that the discharge power had a limited effect on the performance of plasma processing of naphthalene.

As previously reported, the key reaction channels in the plasma reforming of naphthalene are dehydrogenation and one ring-open reaction. In these two reactions, the initial dissociation of naphthalene can proceed via the collisions of naphthalene with excited nitrogen species $N_2$*(e.g., $N_2$ ($A^3\Sigma^+$)) and energetic electrons (R1-R4) (Wang et al., 2019; Yu et al., 2010).

$$C_{10}H_8 + N_2^* \rightarrow C_{10}H_7 + H + N_2 \quad (R1)$$

$$C_{10}H_8 + N_2^* \rightarrow C_{10}H_7 + CH + N_2 \quad (R2)$$

$$C_{10}H_8 + e \rightarrow C_{10}H_7 + H + e \quad (R3)$$

$$C_{10}H_8 + e \rightarrow C_{10}H_7 + CH + e \quad (R4)$$

Fig. 5 shows the predicted values are in agreement with the experimental results at different S/C ratios and discharge powers. As shown in Fig. 5a, the conversion of naphthalene in the plasma steam reforming of tar gradually increases when increasing the S/C ratio, reaching a peak of 84.5% at an S/C ratio of 2.0 (Fig. 5a). However, the conversion of naphthalene drops by about 20% when further increasing the S/C ratio from 2.0 to 4.0. In this study, the optimal S/C ratio is found to be 2.0 to achieve the highest conversion of naphthalene. Due to the presence of steam in the plasma reforming process, abundant OH radicals can be produced via $H_2O$ dissociation by energetic electrons (R5) and excited $N_2$ species ($N_2$*) (R6) (Wang et al., 2019). The OH radicals are highly oxidative and can further oxidize naphthalene (R7) at a relatively high rate constant ($1 \times 10^{-11}$ cm$^3$ molecule$^{-1}$ s$^{-1}$).

$$H_2O + e \rightarrow H + OH + e \quad (R5)$$

$$H_2O + N_2^* \rightarrow H + OH + N_2 \quad (R6)$$

$$C_{10}H_8 + OH \rightarrow \text{products} \quad (R7)$$

.

However, adding excessive $H_2O$ (e.g. at a higher S/C ratio) to the plasma tar reforming process consumes a large number of energetic electrons due to the electron attachment effect of water, which in turn negatively affects the conversion of naphthalene. Thus, choosing the appropriate S/C molar ratio is critical to maximizing the performance of the plasma steam reforming process.

Moreover, the carbon balance increases significantly from 30% to 66% when adding steam into the reaction, but only weakly changes with the change of the S/C ratio (Fig. 5b). The presence of steam in the plasma reforming reaction enhances the oxidation of naphthalene and its intermediates due to the formation of OH radicals, thus reducing the carbon deposition. In addition, we found that the energy efficiency of the plasma tar reforming process as a function of the S/C ratio follows the same evolution as the conversion of naphthalene (Fig. 5c).

Figs. 5d-5f shows the effect of discharge power on the three KPIs with an S/C ratio of 2.0. Increasing the discharge power gradually enhances the conversion of naphthalene but decreases the energy efficiency of the plasma process. At a low discharge power (< 45 W), the influence of discharge power on the conversion of naphthalene is not as substantial.

In this study, 20% of the experimental data was selected randomly as a test set to validify the hybrid model using GA. The regression plots for the comparison between the experimental and simulated KPIs are presented in Fig. 6. The solid line represents the simulated results that are equal to the experimental data. The experimental data shows a fairly good agreement with the results simulated from the hybrid model for all three KPIs, as evidenced by the high correlation coefficients ($R^2$) of 0.997 (conversion), 0.998 (carbon balance), and 0.997 (energy efficiency).

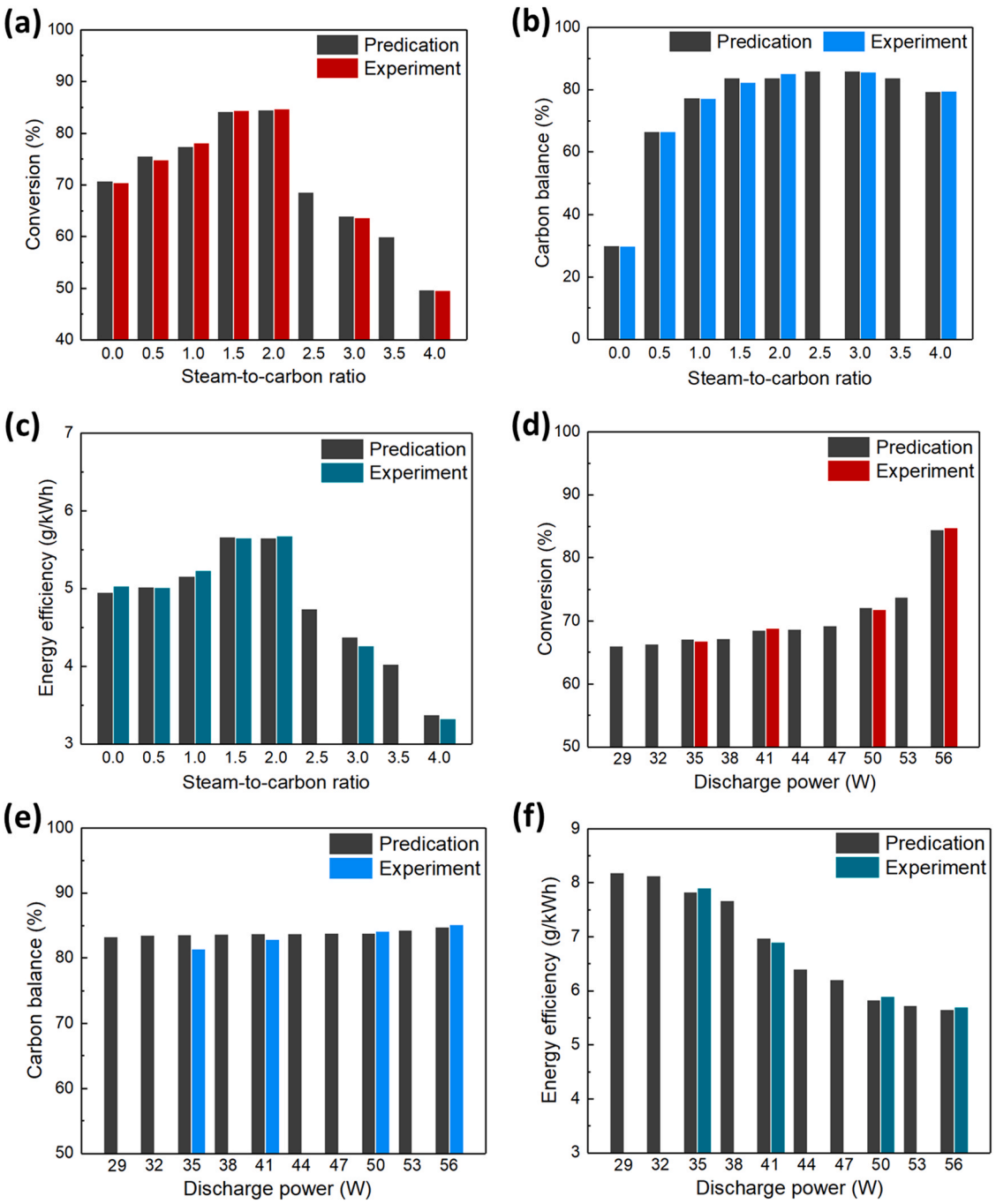


**Fig. 5.** Effect of S/C ratio on (a) tar conversion, (b) carbon balance, and (c) energy efficiency (discharge power = 57 W, concentration = 1.7 g/Nm$^3$). Effect of discharge power on (d) tar conversion, (e) carbon balance, and (f) energy efficiency ($C_{10}H_8$ concentration = 1.7 g/Nm$^3$, S/C ratio = 2.0).

### 3.3. *Importance of different process parameters*

Fig. 7 shows the relative importance of different operating parameters in the plasma steam reforming of naphthalene. The S/C ratio is found to be the most critical parameter affecting the conversion of naphthalene with a relative importance of 48%. Moreover, the relative importance of the S/C ratio and discharge power for the carbon balance is over 40%, which indicates that both parameters significantly affect the carbon balance. For the energy efficiency, the discharge power shows much higher importance of ~58% compared to other two process parameters, suggesting that the discharge power is the most crucial parameter in determining the energy efficiency of the plasma process for naphthalene conversion. Note that the relative importance of the inlet naphthalene concentration is 13–16% for all three KPIs. These findings show that the input concentration of naphthalene in the range of 1.0–2.0 g/Nm$^3$ makes the least contributions to this plasma steam reforming process.

### 3.4. *Coupling effect of process parameters*

The validation of the hybrid ML model enables us to investigate the interactions between the process parameters on the KPIs of the plasma reforming process. As shown in Fig. 8a, in the plasma reforming of naphthalene without steam (S/C = 0), the conversion of naphthalene can be enhanced by lowering the input concentration and increasing the discharge power in the range of 1.0–1.3 g/Nm$^3$ and 55–80 W,

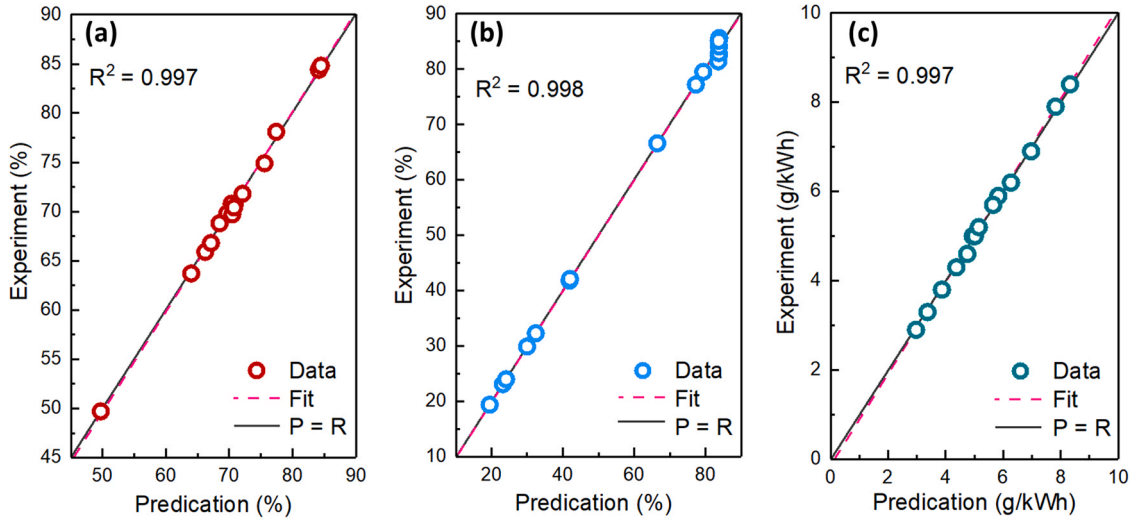


**Fig. 6.** Regression plots for (a) conversion, (b) carbon balance, and (c) energy efficiency using the hybrid prediction model.

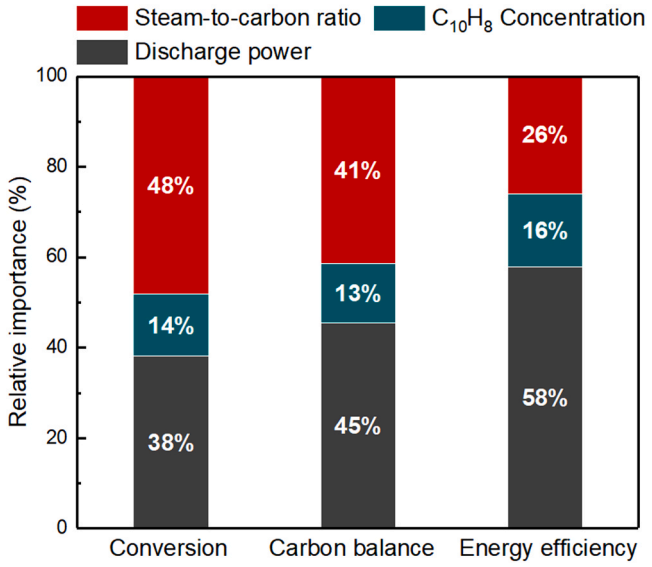


**Fig. 7.** Relative importance (%) of different process parameters in the plasma-driven steam reforming of naphthalene.

respectively, with the highest naphthalene conversion being achieved at 84%. Fig. 8b shows that the carbon balance is almost unchanged when the discharge power is lower than 50 W regardless of the change of the naphthalene concentration. However, at a discharge power of > 50 W (especially > 60 W), increasing the discharge power substantially enhances the carbon balance if the concentration is in the range of 1.8–2.0 g/Nm$^3$, but slightly decreases the carbon balance at a lower naphthalene concentration (1.0–1.4 g/Nm$^3$). Also, the predicted energy efficiency monotonically decreases with the increase of discharge power at a constant naphthalene concentration, as shown in Fig. 8c. The highest energy efficiency of the plasma process is predicted to be 8.2 g/kWh at 35 W with an inlet naphthalene concentration of 1.7 g/Nm$^3$.

In the plasma steam reforming of naphthalene, the interaction between the S/C ratio and discharge power on different KPIs is shown in Fig. 9. The effect of discharge power and S/C ratio on the conversion shows a ridge-shaped surface, with the maximum conversion being achieved at an S/C ratio of 1.0–2.0. In contrast, the highest conversion of naphthalene is predicted to be 89% at a discharge power of 70 W and an S/C ratio of 2.0. Compared to the discharge power, the S/C ratio has a more significant effect on the carbon balance, as shown in Fig. 9b. The carbon balance can reach the highest value of 85.5% at an S/C ratio of 2.0–3.0. Fig. 9c shows a higher energy efficiency can be expected when lowering both the discharge power and S/C ratio. The maximum energy efficiency can be around 8.0 g/kWh with the minimum discharge power (30 W) and S/C ratio (S/C = 0). Meanwhile, when the discharge power

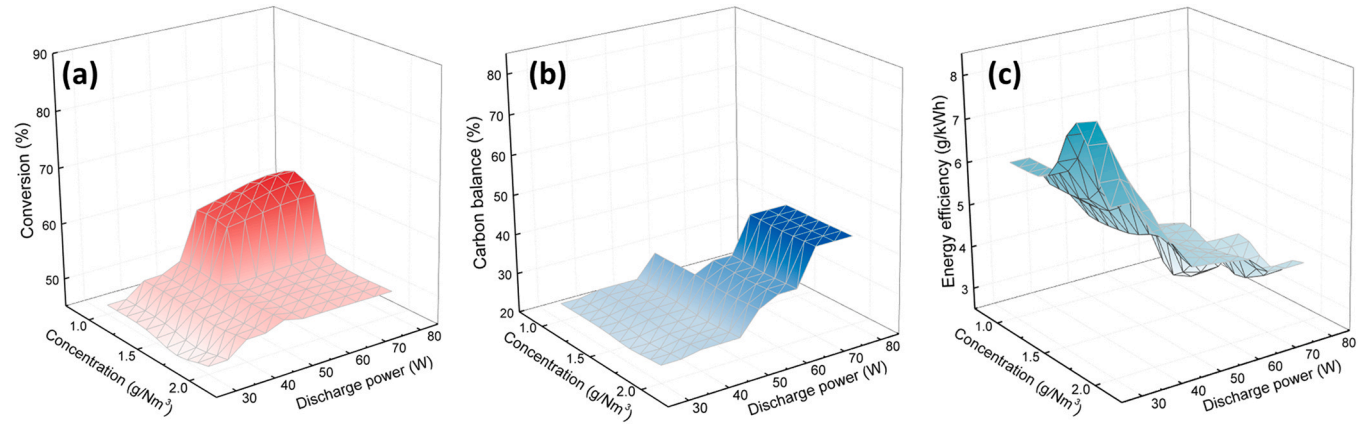


**Fig. 8.** Prediction effects of the interaction between naphthalene concentration and discharge power on (a) tar conversion, (b) carbon balance, and (c) energy efficiency without steam. (S/C = 0).

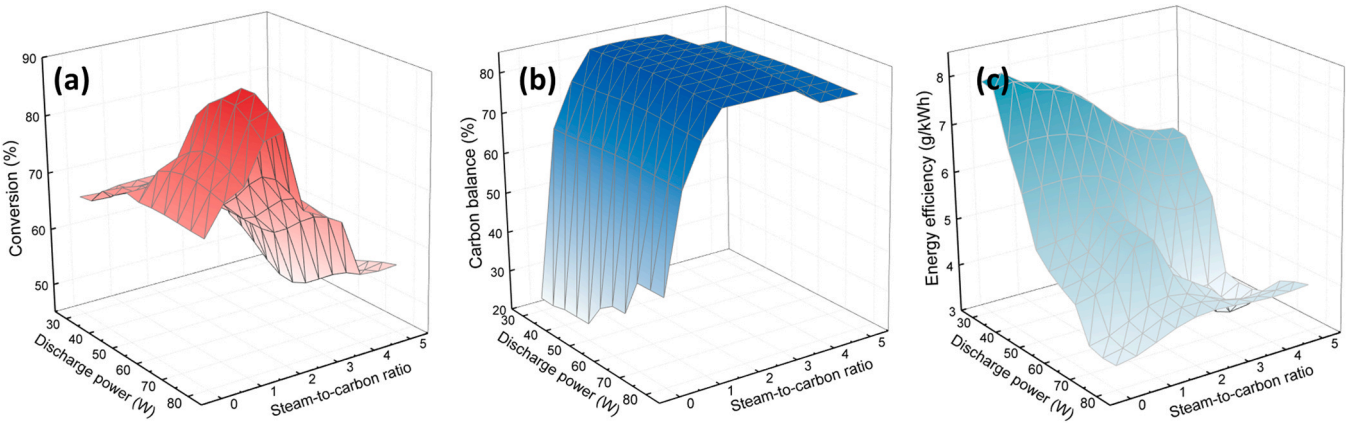


**Fig. 9.** Prediction effects of the interaction between discharge power and steam-to-carbon ratio on (a) conversion, (b) carbon balance and (c) energy efficiency in the steam reforming reaction (concentration = 1.7 g/Nm$^3$).

is greater than 50 W, the energy efficiency can be optimized and maximized at an S/C ratio of 2.0, indicating that adding a moderate amount of steam to the plasma process could enhance the energy efficiency.

Furthermore, the sliced graphs of the predicted results show the simultaneous effects of discharge power, S/C ratio, and inlet naphthalene concentration on the three KPIs (Figs. 10a-10c). When the inlet $C_{10}H_8$ concentration is lower than 1.3 g/Nm$^3$, the conversion can reach more than 80% in the optimal ranges of discharge power (60–80 W) and S/C ratio (0–2.0). To achieve a high naphthalene conversion (> 80%) at a high $C_{10}H_8$ concentration (> 1.4 g/Nm$^3$), the S/C ratio needs to be narrowed between 1.0 and 2.0 (Fig. 10a).

Carbon balance shows a similar trend under the coupling effects of these three operating parameters (discharge powers, S/C ratios, and tar concentration). A carbon balance of > 80% can be obtained if the tar concentration is in the range of 1.6–1.8 g/Nm$^3$ (Fig. 10b). As shown in Fig. 10c, the energy efficiency of the plasma reforming process is mainly affected by the discharge power, rather than the inlet $C_{10}H_8$ concentration and S/C ratio. Lower discharge power (30–40 W) is favourable to achieve higher energy efficiency. The highest energy efficiency (~8.4 g/

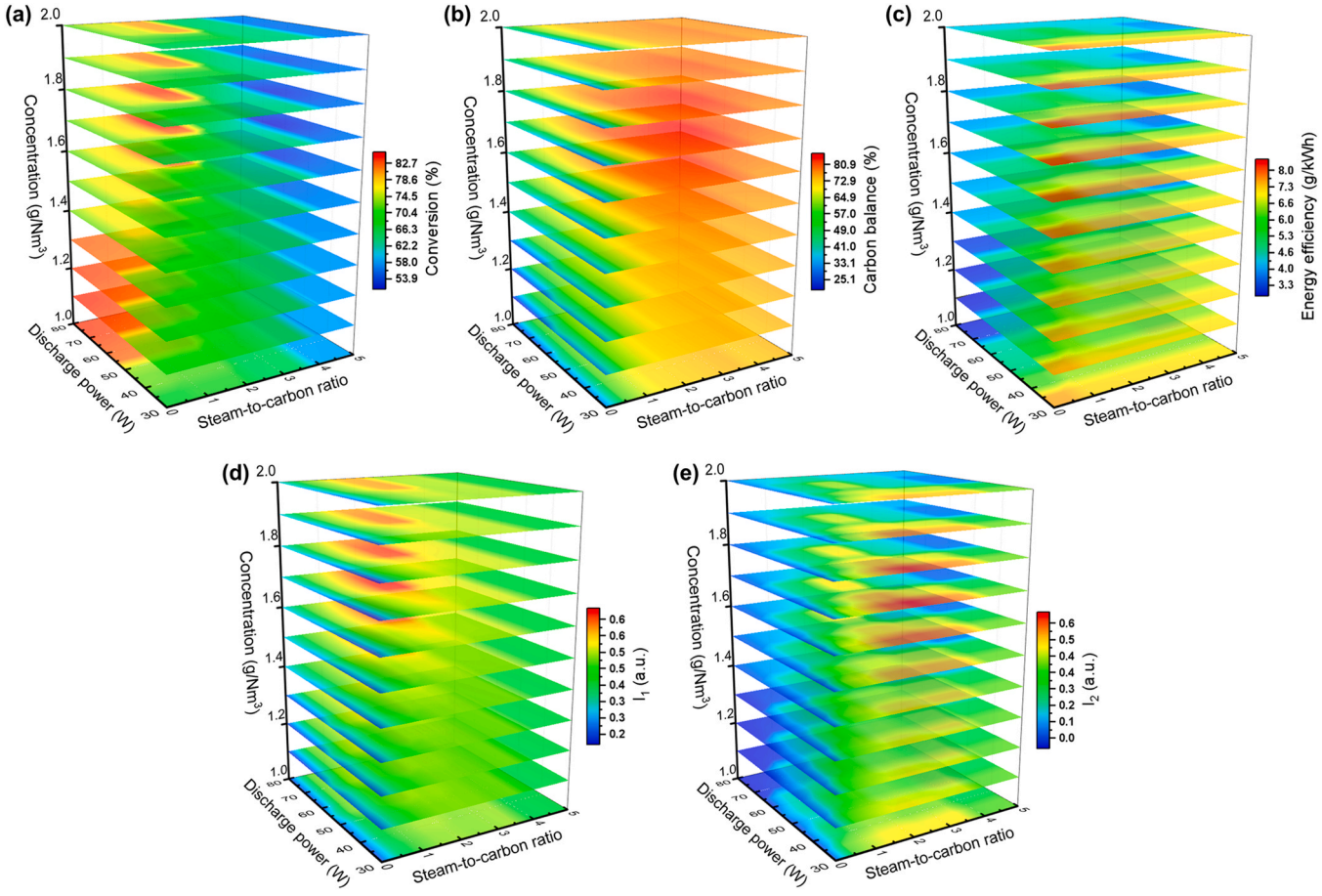


**Fig. 10.** Prediction effects of the three-body interactions of discharge power, steam-to-carbon ratio and inlet $C_{10}H_8$ concentration on (a) conversion, (b) carbon balance, and (c) energy efficiency; and the comprehensive evaluation indexes (d) $I_1$ and (e) $I_2$.

kWh) can be obtained when the discharge power, $C_{10}H_8$ concentration and S/C ratio are 30 W, 1.7 g/Nm$^3$, and 0.6, respectively (Fig. 10c).

Clearly, the overall performance of plasma steam reforming of naphthalene strongly depends on a range of plasma processing parameters. It is essential to optimize the complex plasma-enhanced tar reforming process with multiple input processing parameters and multiple KPIs. In this work, $I_1$ has been used to determine the optimal processing parameters to maximize the effective conversion of naphthalene for process optimization. As shown in Fig. 10d, at a $C_{10}H_8$ concentration of 1.5–1.9 g/Nm$^3$, $I_1$ can be higher than 0.60 if the discharge power and S/C ratio are in the optimal range of 60–80 W and 1.5–2.0, respectively. The highest $I_1$ of 0.69 can be obtained at a discharge power of 60 W, an inlet concentration of 1.7 g/Nm$^3$ and an S/C ratio of 2.0.

In addition, $I_2$, also known as the global desirability function (Mei et al., 2016), has been introduced to determine the optimal processing parameters to maximize both effective tar conversion and energy efficiency simultaneously in the plasma-enhanced tar reforming process. Fig. 10e shows that a relatively high $I_2$ ($> 0.5$) can be obtained when the inlet naphthalene concentration and S/C ratio are in the optimal range of 1.5–1.9 g/Nm$^3$ and 1.5–3.5, respectively when the discharge power is 30–45 W. Furthermore, when $I_2$ reaches the highest value of 0.65, the optimal discharge power (35 W), input $C_{10}H_8$ concentration (1.7 g/Nm$^3$) and S/C ratio (2.0) are determined to achieve the maximum KPIs - conversion (67.2%), carbon balance (81.7%), and energy efficiency (7.8 g/kWh) simultaneously in the plasma reforming of naphthalene.

## 4. Conclusions

A well-trained hybrid ML model incorporating three different algorithms (ANN, SVR, and DT) has been developed to predict and evaluate the influence of different processing parameters (discharge power, tar concentration, and S/C ratio) on the KPIs (tar conversion, carbon balance, and energy efficiency) of the multi-scale and complex plasma tar reforming process. The hyper-parameters of each algorithm in the hybrid ML model have been optimized with the GA. The predicted results from the model agree very well with the experimental data, as confirmed by the high regression coefficient of 0.997. Both the S/C ratio and the discharge power are identified as the most critical parameters affecting the three KPIs, while the inlet concentration of naphthalene is less critical compared to the other two parameters. The hybrid model shows that the discharge power is the most important processing parameter to determine the energy efficiency of the plasma process with a relative importance of 58%, while the S/C ratio is the most critical parameter for the tar conversion with a relative contribution of 38%. When $I_2$ reaches the highest value of 0.65, the optimal discharge power (35 W), input $C_{10}H_8$ concentration (1.7 g/Nm$^3$), and S/C ratio (2.0) are obtained to maximize the three KPIs - conversion (67.2%), carbon balance (81.7%), and energy efficiency (7.8 g/kWh) simultaneously in the plasma reforming of naphthalene. This work has demonstrated that a well-trained hybrid ML model can provide effective, accurate, and fast predictions as well as optimization of the plasma tar reforming process, and has great potential to be used for a range of plasma-based chemical processes.

## CRediT authorship contribution statement

**Yaolin Wang**: Experiment, Conceptualization, Formal analysis, Data curation, Visualization, Writing - original draft. **Zinan Liao:** Conceptualization, Model development, Formal analysis. **Stephanie Mathieu**: Formal analysis, Writing - review & editing. **Feng Bin**: Conceptualization, Co-supervision, Writing - review & editing. **Xin Tu**: Conceptualization, Supervision, Writing - review & editing, Project administration.

## 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.

## Acknowledgements

This project has received funding from European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 722346. The support from the Royal Society - Newton Advanced Fellowship (Ref. NAF/R1/180230) is greatly appreciated.

## References

Bogaerts, A., Tu, X., Whitehead, J.C., Centi, G., Lefferts, L., Guaitella, O., Azzolina-jury, F., Kim, H., Murphy, A.B., Schneider, W.F., Nozaki, T., Jason, C., Rousseau, A., Thevenet, F., Khacef, A., 2020. The 2020 plasma catalysis roadmap. J. Phys. D Appl. Phys. 53, 443001 https://doi.org/10.1088/1361-6463/ab9048.

George, A., Shen, B., Craven, M., Wang, Y., Kang, D., Wu, C., Tu, X., 2021. A review of non-thermal plasma technology: a novel solution for $CO_2$ conversion and utilization. Renew. Sustain. Energy Rev. 135, 109702 https://doi.org/10.1016/j.rser.2020.109702.

Liu, L., Zhang, Z., Das, S., Kawi, S., 2019. Reforming of tar from biomass gasification in a hybrid catalysis-plasma system: a review. Appl. Catal. B Environ. 250, 250–272. https://doi.org/10.1016/j.apcatb.2019.03.039.

Liu, S., Mei, D., Wang, L., Tu, X., 2017. Steam reforming of toluene as biomass tar model compound in a gliding arc discharge reactor. Chem. Eng. J. 307, 793–802. https://doi.org/10.1016/j.cej.2016.08.005.

Wang, Y., Yang, H., Tu, X., 2019. Plasma reforming of naphthalene as a tar model compound of biomass gasification. Energy Convers. Manag. 187, 593–604. https://doi.org/10.1016/j.enconman.2019.02.075.

Mei, D., Wang, Y., Liu, S., Alliati, M., Yang, H., Tu, X., 2019. Plasma reforming of biomass gasification tars using mixed naphthalene and toluene as model compounds. Energy Convers. Manag. 195, 409–419. https://doi.org/10.1016/j.enconman.2019.05.002.

Craven, M., Wang, Y., Yang, H., Wu, C., Tu, X., 2020. Integrated gasification and non-thermal plasma-catalysis system for cleaner syngas production from cellulose. IOP SciNotes 1, 024001. https://doi.org/10.1088/2633-1357/aba7f6.

Zhang, H., Zhu, F., Li, X., Xu, R., Li, L., Yan, J., Tu, X., 2019. Steam reforming of toluene and naphthalene as tar surrogate in a gliding arc discharge reactor. J. Hazard. Mater. 369, 244–253. https://doi.org/10.1016/j.jhazmat.2019.01.085.

Trushkin, A.N., Grushin, M.E., Kochetov, I.V., Trushkin, N.I., Akishev, Y.S., 2013. Decomposition of toluene in a steady-state atmospheric-pressure glow discharge. Plasma Phys. Rep. 39, 167–182. https://doi.org/10.1134/S1063780×13020025.

Liu, S.Y., Mei, D.H., Nahil, M.A., Gadkari, S., Gu, S., Williams, P.T., Tu, X., 2017. Hybrid plasma-catalytic steam reforming of toluene as a biomass tar model compound over $Ni/Al_2O_3$ catalysts. Fuel Process Technol. 166, 269–275. https://doi.org/10.1016/j.fuproc.2017.06.001.

Saleem, F., Harvey, A., Zhang, K., 2019. Low temperature conversion of toluene to methane using dielectric barrier discharge reactor. Fuel 248, 258–261. https://doi.org/10.1016/j.fuel.2019.02.137.

Saleem, F., Zhang, K., Harvey, A., 2019. Temperature dependence of non-thermal plasma assisted hydrocracking of toluene to lower hydrocarbons in a dielectric barrier discharge reactor. Chem. Eng. J. 356, 1062–1069. https://doi.org/10.1016/j.cej.2018.08.050.

Zhu, F., Zhang, H., Yang, H., Yan, J., Li, X., Tu, X., 2020. Plasma reforming of tar model compound in a rotating gliding arc reactor: understanding the effects of $CO_2$ and $H_2O$ addition. Fuel 259, 1–11. https://doi.org/10.1016/j.fuel.2019.116271.

Chun, Y.N., Kim, C.S., Yoshikawa, K., 2012. Decomposition of benzene as a surrogate tar in a gliding arc plasma. Environ. Prog. Sustain Energy 32, 837–845. https://doi.org/10.1002/ep.11663.

Gao, N., Wang, X., Li, A., Wu, C., Yin, Z., 2016. Hydrogen production from catalytic steam reforming of benzene as tar model compound of biomass gasification. Fuel Process Technol. 148, 380–387. https://doi.org/10.1016/j.fuproc.2016.03.019.

Mao, L., Chen, Z., Wu, X., Tang, X., Yao, S., Zhang, X., Jiang, B., Han, J., Wu, Z., Lu, H., Nozaki, T., 2018. Plasma-catalyst hybrid reactor with $CeO_2/\gamma$-$Al_2O_3$ for benzene decomposition with synergetic effect and nano particle by-product reduction. J. Hazard. Mater. 347, 150–159. https://doi.org/10.1016/j.jhazmat.2017.12.064.

Wu, Z., Wang, J., Han, J., Yao, S., Xu, S., Martin, P., 2017. Naphthalene decomposition by dielectric barrier discharges at atmospheric pressure. IEEE Trans. Plasma Sci. 45, 154–161. https://doi.org/10.1109/TPS.2016.2632154.

Liu, L., Liu, Y., Song, J., Ahmad, S., Liang, J., Sun, Y., 2019. Plasma-enhanced steam reforming of different model tar compounds over Ni-based fusion catalysts. J. Hazard. Mater. 377, 24–33. https://doi.org/10.1016/j.jhazmat.2019.05.019.

Mei, D., Liu, S., Wang, Y., Yang, H., Bo, Z., Tu, X., 2019. Enhanced reforming of mixed biomass tar model compounds using a hybrid gliding arc plasma catalytic process. Catal. Today 337, 225–233. https://doi.org/10.1016/j.cattod.2019.05.046.

Chun, Y.N., Kim, S.C., Yoshikawa, K., 2011. Destruction of anthracene using a gliding arc plasma reformer, Korean. J. Chem. Eng. 28, 1713–1720. https://doi.org/10.1007/s11814-011-0162-x.

Kong, X., Zhang, H., Li, X., Xu, R., Mubeen, I., Li, L., Yan, J., 2019. Destruction of toluene, naphthalene and phenanthrene as model tar compounds in a modified rotating gliding arc discharge reactor. Catalysts 9, 6–9. https://doi.org/10.3390/catal9010019.

Liu, L., Wang, Q., Ahmad, S., Yang, X., Ji, M., Sun, Y., 2018. Steam reforming of toluene as model biomass tar to $H_2$-rich syngas in a DBD plasma-catalytic system. J. Energy Inst. 91, 927–939. https://doi.org/10.1016/j.joei.2017.09.003.

Xu, B., Xie, J., Zhan, H., Yin, X., Wu, C., Liu, H., 2018. Removal of toluene as a biomass tar surrogate in a catalytic nonthermal plasma process. Energy and Fuels 32, 10709–10719. https://doi.org/10.1021/acs.energyfuels.8b02444.

Wu, Z., Zhu, Z., Hao, X., Zhou, W., Han, J., Tang, X., Yao, S., Zhang, X., 2018. Enhanced oxidation of naphthalene using plasma activation of $TiO_2$/diatomite catalyst. J. Hazard. Mater. 347, 48–57. https://doi.org/10.1016/j.jhazmat.2017.12.052.

Yang, Y.C., Chun, Y.N., 2011. Naphthalene destruction performance from tar model compound using a gliding arc plasma reformer, Korean. J. Chem. Eng. 28, 539–543. https://doi.org/10.1007/s11814-010-0393-2.

Du, C.M., Yan, J.H., Cheron, B., 2007. Decomposition of toluene in a gliding arc discharge plasma reactor. Plasma Sources Sci. Technol. 16, 791–797. https://doi.org/10.1088/0963-0252/16/4/014.

Yu, L., Li, X., Tu, X., Wang, Y., Lu, S., Yan, J., 2010. Decomposition of naphthalene by dc gliding arc gas discharge. J. Phys. Chem. A 114, 360–368. https://doi.org/10.1021/jp905082s.

Nunnally, T., Tsangaris, A., Rabinovich, A., Nirenberg, G., Chernets, I., Fridman, A., 2014. Gliding arc plasma oxidative steam reforming of a simulated syngas containing naphthalene and toluene. Int. J. Hydrog. Energy. 39, 11976–11989. https://doi.org/10.1016/j.ijhydene.2014.06.005.

Zhu, F., Li, X., Zhang, H., Wu, A., Yan, J., Ni, M., Zhang, H., Buekens, A., 2016. Destruction of toluene by rotating gliding arc discharge. Fuel 176, 78–85. https://doi.org/10.1016/j.fuel.2016.02.065.

Jamróz, P., Kordylewski, W., Wnukowski, M., 2018. Microwave plasma application in decomposition and steam reforming of model tar compounds. Fuel Process Technol. 169, 1–14. https://doi.org/10.1016/j.fuproc.2017.09.009.

Sun, J., Wang, Q., Wang, W., Wang, K., 2018. Plasma catalytic steam reforming of a model tar compound by microwave-metal discharges. Fuel 234, 1278–1284. https://doi.org/10.1016/j.fuel.2018.07.140.

Młotek, M., Ulejczyk, B., Woroszył, J., Krawczyk, K., 2020. Decomposition of Toluene in Coupled Plasma-Catalytic System. Ind. Eng. Chem. Res. 59, 4239–4244. https://doi.org/10.1021/acs.iecr.9b04330.

Tao, K., Ohta, N., Liu, G., Yoneyama, Y., Wang, T., Tsubaki, N., 2013. Plasma enhanced catalytic reforming of biomass tar model compound to syngas. Fuel 104, 53–57. https://doi.org/10.1016/j.fuel.2010.05.044.

Liu, S.Y., Mei, D.H., Shen, Z., Tu, X., 2014. Nonoxidative conversion of methane in a dielectric barrier discharge reactor: prediction of reaction performance based on neural network model. J. Phys. Chem. C 118, 10686–10693. https://doi.org/10.1021/jp502557s.

Zhu, X., Liu, S., Cai, Y., Gao, X., Zhou, J., Zheng, C., Tu, X., 2016. Post-plasma catalytic removal of methanol over Mn - Ce catalysts in an atmospheric dielectric barrier discharge. Appl. Catal. B Environ. 183, 124–132. https://doi.org/10.1016/j.apcatb.2015.10.013.

Istadi, I., Amin, N.A.S., 2007. Modelling and optimization of catalytic-dielectric barrier discharge plasma reactor for methane and carbon dioxide conversion using hybrid artificial neural network-genetic algorithm technique. Chem. Eng. Sci. 62, 6568–6581. https://doi.org/10.1016/j.ces.2007.07.066.

Chang, T., Lu, J., Shen, Z., Huang, Y., Lu, D., Wang, X., Cao, J., Morent, R., 2019. Simulation and optimization of the post plasma-catalytic system for toluene degradation by a hybrid ANN and NSGA-II method. Appl. Catal. B Environ. 244, 107–119. https://doi.org/10.1016/j.apcatb.2018.11.025.

Ye, Z., Yang, J., Zhong, N., Tu, X., Jia, J., Wang, J., 2020. Tackling environmental challenges in pollution controls using artificial intelligence: a review. Sci. Total Environ. 699, 134279 https://doi.org/10.1016/j.scitotenv.2019.134279.

Sun, Y., Liu, L., Wang, Q., Yang, X., Tu, X., 2016. Pyrolysis products from industrial waste biomass based on a neural network model. J. Anal. Appl. Pyrolysis 120, 94–102. https://doi.org/10.1016/j.jaap.2016.04.013.

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.

Drucker, H., Surges, C.J.C., Kaufman, L., Smola, A., Vapnik, V., 1997. Support vector regression machines. Adv. Neural Inf. Process Syst. 1, 155–161.

Safavian, S.R., Landgrebe, D., 1991. A survey of decision tree classifier methodology. IEEE Trans. Syst. Man Cybern. 21, 660–674. https://doi.org/10.1109/21.97458.

Shao, Y., Lunetta, R.S., 2012. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. ISPRS J. Photogramm. Remote Sens. 70, 78–87. https://doi.org/10.1016/j.isprsjprs.2012.04.001.

Mei, D., He, Y., Liu, S., Yan, J., Tu, X., 2016. Optimization of $CO_2$ conversion in a cylindrical dielectric barrier discharge reactor using design of experiments. Plasma Process Polym. 13, 544–556. https://doi.org/10.1002/ppap.201500159.