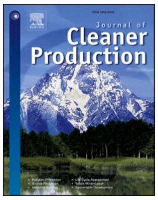
Journal of Cleaner Production 526 (2025) 146662

Contents lists available at ScienceDirect

Journal of Cleaner Production

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



Hybrid pathways for methane production: Merging thermodynamic insights with machine learning

Azita Etminan [a,*], Peter J. Holliman [a], Peyman Karimi [b], Majid Majd [b], Ian Mabbett [a], Mary Larimi [a], Ciaran Martin [c], Anna RL. Carter [d]

[a] Swansea University, Faculty of Science and Engineering, Bay Campus, Swansea, SA1 8EN, UK
[b] Arak University, Faculty of Engineering, Department of Mechanical Engineering, Arak, 38156-88349, Iran
[c] Tata Steel UK, Port Talbot, UK
[d] School of Computer and Information Sciences, Northumbria University, Newcastle Upon Tyne, UK

ARTICLE INFO

Keywords:
Thermodynamic analysis
Catalyst selection
Machine learning
CO_2 Methanation
Energy and Exergy Analysis

ABSTRACT

A comprehensive study was conducted to simultaneously simulate thermodynamic behavior and predict catalyst performance for CH_4 production via CO and CO_2 methanation, using blast furnace gas (BFG) and basic oxygen furnace gas (BOFG) as feedstocks. Thermodynamic equilibrium simulations based on Gibbs free energy minimization identified optimal reaction conditions at moderate temperatures (150–250 °C) and elevated pressures, achieving over 98 % CO_2 conversion with less than 1 wt% carbon formation. In parallel, machine learning models were developed using an augmented dataset of 2777 experimental observations. Atomic-level structural and electronic descriptors were incorporated into the dataset, including unit cell density and formation energy for active metals, promoters, and supports. Feature selection through Pearson correlation and RFECV identified active phase weight, support density, and reduction conditions as the most influential variables. Among all tested algorithms, XGBoost and CatBoost demonstrated the highest accuracy, with R^2 values exceeding 0.93 for predicting CH_4 yield, selectivity, and CO_2 conversion. SHAP and partial dependence analyses showed that catalyst stability and textural properties govern overall performance. This integrated thermodynamic and machine learning approach defines the operating limits for high-efficiency methanation and provides a data-driven framework for catalyst optimization in industrial applications.

Nomenclature

Symbols and Terms	
$\Delta_r G^0$ (T)	Standard Gibbs free energy change of reaction at temperature T (kJ/mol)
$\Delta_r H^0$ (T)	Standard enthalpy change of reaction at temperature T (kJ/mol)
$\Delta_r S^0$ (T)	Standard entropy changes of reaction at temperature T (kJ/mol·K)
$C_{p,m}$	Molar heat capacity at constant pressure (kJ/mol·*K*)
R	Universal gas constant (8.314 J/mol·K)
$\eta_{Total,En,}$	Overall energy efficiency (%)
$H_{,Reactant}$	Total enthalpy of the reactants (kJ)
$H_{,Product}$	Total enthalpy of the products (kJ)
Q_{heat}	Net heat released by the methanation reaction (MJ)
$En_{,out}$	Total useful energy output from the system (MJ/h)
$En_{,in}$	Total energy input supplied to the system (MJ/h)
$LHV_{,CH4}$	Lower heating value of methane (MJ/Kmol)
$LHV_{,H2}$	Lower heating value of hydrogen (MJ/Kmol)
n_i	Molar flow rate of species *i* (Kmol/h)
$\eta_{,Total,Ex}$	Total exergy efficiency (%)
$Ex_{,out}$	Exergy recovered from the process (MJ/h)
$Ex_{,in}$	Exergy supplied to the process(MJ/h)
Ex_{Ch}	Standard chemical exergy of pure substances (MJ/Kmol)
$Ex_{,Ph}$	Standard physical exergy of pure substances (MJ/Kmol)

(*continued on next page*)

---

* Corresponding author.
E-mail address: 2157090@swansea.ac.uk (A. Etminan).

https://doi.org/10.1016/j.jclepro.2025.146662
Received 19 April 2025; Received in revised form 4 September 2025; Accepted 13 September 2025
Available online 17 September 2025

(*continued*)

$E_{x,Q}$	Exergy associated with heat transfer (MJ/h)
T^0	Reference temperature (°C)
T	Process temperature (°C)
$F_{CO2,in}$	Inlet molar flow rate of CO_2 (kmol/h)
$F_{CO2,out}$	Outlet molar flow rate of CO_2 (kmol/h)
$F_{CO,in}$	Inlet molar flow rate of CO (kmol/h)
$F_{CO,out}$	Outlet molar flow rate of CO (kmol/h)
$F_{CH4,out}$	Outlet molar flow rate of CH_4 (kmol/h)
$F_{H2,in}$	Inlet molar flow rate of H_2 (kmol/h)
$F_{C(s)}$	Outlet molar flow rate of solid carbon (kmol/h)
Var (Y)	Total variance of the model output Y
$E_{X-i}[Y\|X_i]$	Conditional expectation of Y given Xi, averaged over all other inputs X−i
$Var_{X-i}(E_{Xi}[Y\|X_{-i}])$	Variance in the conditional expectation of Y due to Xi
Abbreviation	
Active component density	ACD
Active component formation energy	ACFE
Active component content (wt percent)	ACW
Promoter density	PD
Promoter formation energy	PFE
Promoter content (wt percent)	PW
Support a density	SAD
Support a formation energy	SAFE
Support a content (wt percent)	SAW
Support b density	SBD
Support b formation energy	SBFE
Calcination Temperature (C)	CTC
Calcination time (h)	CTH
Reduction Temperature (C)	RTC
Reduction Pressure (bar)	RPB
Reduction time (h)	RTH
Reduced hydrogen content (vol percent)	RHC
Temperature (C)	T
Pressure (bar)	P
Weight hourly space velocity [mgcat/(min·ml)]	WHSV
Content of inert components in raw materials (vol percent)	CIRMW
H_2/CO_2 ratio (mol/mol)	H_2/CO_2

1. Introduction

The combustion of fossil fuels and biomass is a primary source of atmospheric pollutants, posing significant environmental and global health risks (Alamgholiloo et al., 2024). A major contributor to this issue is carbon dioxide (CO_2), a key anthropogenic greenhouse gas that drives global warming (Sedighi et al., 2024). Synthetic methane presents a promising pathway for long-term energy storage, improved grid reliability, and seamless integration with existing natural gas infrastructure (Nemmour et al., 2023). Its scalability and cost-effectiveness position it as a viable solution for reducing emissions across various sectors, including transportation, industry, and residential energy (Hidalgo and Martín-Marroquín, 2020). As such, it plays a vital role in bridging the gap between current energy demands and a more sustainable energy system (Götz et al., 2016). However, the methanation of carbon dioxide and carbon monoxide is a highly exothermic reaction ($\Delta H^\circ = -165$ KJ. mol^{-1}), which introduces significant thermodynamic challenges (Mebrahtu et al., 2019). Excessive heat generation can lead to catalyst deactivation and diminished reaction efficiency (Ewald et al., 2019). Efficient thermal management is therefore essential to maintaining system stability and performance (Coppitters et al., 2023). In particular, avoiding the formation of reactor hot spots is critical, as they can accelerate catalyst degradation and compromise overall process efficiency (Huynh et al., 2022). To enhance CO_2 an CO conversion and CH_4 selectivity, the process typically operates under conditions of 200–400 °C and pressures of 1–30 bar (Rönsch et al., 2016). Addressing these operational complexities requires a fundamental understanding of reaction thermodynamics, which enables the identification of feasible and efficient operating conditions under varying process constraints.

Thermodynamic modelling, based on Gibbs free energy minimization, is a fundamental approach for determining feasible operating conditions in chemical processes (Sahebdelfar and Takht Ravanchi, 2015a). Thermodynamic analysis of CO_2 hydrogenation demonstrates that methanation is the most thermodynamically favorable reaction pathway, rendering it a particularly promising route for sustainable fuel production (Jia et al., 2016). Building on this framework, Hussain et al. thermodynamic modelling with experimental validation to optimize CO_2 methanation for synthetic natural gas production. Simulations using Gibbs free energy minimization identified low temperatures (25–300 °C), moderate pressures (0.1–3 MPa), and a high H_2: CO_2 ratio ($\geq$4:1) for maximizing methane yield and minimizing carbon deposition. These predictions were experimentally confirmed using a metal-free FS@SiO_2-BEA catalyst, which showed high CO_2 conversion and CH_4 selectivity under the same conditions (Hussain et al., 2021a). In a complementary thermodynamic equilibrium study, Jiang et al. examined CO_2 methanation over a CeO_2-promoted Ni/Al_2O_3 catalyst, confirming thermodynamically predicted performance at 260 °C, with ~80 % CO_2 conversion and 99.9 % CH_4 selectivity (Jiang and Lian, 2024). In another equilibrium-based study, CH_4 formation was shown to reach 99.87 % conversion and 99.99 % selectivity at 100 °C, 1 atm, and H_2/CO_2 ratio of 4:1, emphasizing the reaction's viability at low temperatures and low pressures despite kinetic limitations (Yarbaş and Ayas, 2024). Gibbs free energy minimization confirms CH_4 formation is thermodynamically favored under low-temperature, low-pressure conditions, supporting the role of thermodynamic modelling in CO_2 methanation process design (Colelli et al., 2024). Thermodynamic simulations provide reliable, physics-based predictions of equilibrium states, making them useful tools for determining favorable operating conditions in CO_2 conversion processes. However, these models are inherently limited by their focus on equilibrium and do not account for kinetic constraints or catalyst-specific reaction pathways.

CO_2 and CO methanation is a complex process that demands carefully controlled conditions and the use of highly active catalysts to achieve desirable methane yield and selectivity (Shen et al., 2020). Despite being thermodynamically favorable, the reaction proceeds slowly without catalytic assistance due to the inherent chemical stability of CO_2 and CO. Despite ongoing advancements, designing heterogeneous catalysts with high activity for CO_2 methanation under mild conditions remains a significant challenge (Zhang et al., 2025). Conventional catalyst discovery methods rely heavily on chemical intuition and trial-and-error experimentation, which are time-consuming and inefficient given the vast chemical space (Yang et al., 2024a). Similarly, DFT and simulation techniques are limited by computational cost and scale (Sauer, 2024). To address these challenges, machine learning (ML) offers a powerful, data-driven methodology for uncovering complex relationships in material science, enabling the efficient optimization of processes like catalytic conversion (Ma and Liu, 2020). Specifically, artificial intelligence (AI) models, including Artificial Neural Networks (ANN), Support Vector Regression (SVR), and Genetic Algorithms (GA), are increasingly applied to predict complex variables and enhance system performance (Sheikhmohammadi et al., 2024). ML models can also quantify feature importance and predict the activity of untested catalysts, overcoming limitations of traditional methods (Wang et al., 2025). Yang et al. (2025) proposed an ensemble learning approach with active learning to improve catalyst design for CO_2 methanation. They found that recursive feature elimination with cross-validation worked better than autoencoders for feature selection. A random forest model, optimized using active learning, showed relatively high accuracy ($R^2 >$ 0.92). The model helped identify known catalysts like Ni/Al_2O_2 and Ni/CeO_2 with reported high yields. In another study Yılmaz et al. (2023) analyzed a dataset of 4051 entries from 527 experiments on CO_2 methanation using a random forest model to predict CO_2 conversion based on 23 features. The Boruta analysis highlighted temperature, support type, preparation method, and metal loading as key factors. The model performed well (R^2 of 0.95 training, 0.87 testing), and its predictions aligned closely with reported experimental results. In another study by Yang et al. (2024b) a machine learning-assisted reverse design framework proposed to develop CO_2 methanation catalysts with high low-temperature activity. Using a dataset of 2777 entries, they compared multiple models and found the LGBM model to perform best ($R^2 > 0.90$). Feature interpretation showed that adjusting factors like support content and calcination temperature could enhance performance. Coupling the model with a multi-objective optimization algorithm, identified new Ru and Ni based catalysts, with one $Ru–Ba/Cr_2O_2–SrO$ catalyst predicted to achieve over 97 % CO_2 conversion and 72 % CH_4 selectivity at low temperatures. Wang et al. (2025) used an extrapolative machine learning approach to discover over 100 high-performing multi-elemental catalysts for the reverse water-gas shift reaction, surpassing the best known Pt-based catalyst. The top catalyst, $Pt_{(3)}/Rb_{(1)}\text{-}Ba_{(1)}\text{-}Mo_{(0.6)}\text{-}Nb_{(0.2)}/TiO_2$, included Nb, an element not in the training data, highlighting the model's predictive power. The results show ML can uncover unexpected, highly active catalyst compositions with roles tailored by each element. However, machine learning models only use catalyst content and preparation properties. In the heterogeneous catalysis field, the discovery of new materials with desired properties is challenging because of its large parameter space (Shin et al., 2023). Several studies used machine learning on structural atomic properties to evaluate catalyst performance Ulissi et al. (2017) developed a systematic framework to model the diverse active sites of bimetallic catalysts, addressing challenges in predicting their catalytic behavior. Using a neural network surrogate model to reduce the number of DFT calculations, they efficiently screened hundreds of possible sites across low-index facets. Tran and Ulissi (2018) established a link between structural fingerprints and adsorption energies of CO and H_2 to guide catalyst screening for CO_2 reduction and hydrogen evolution. Using target adsorption energies ($E_{CO} = -0.67$ eV and $E_H = -0.27$ eV), they screened alloys containing 31 elements, leading to the identification of 131 CO_2RR and 258 HER candidate surfaces. This approach notably accelerated the discovery of CuAl as a promising CO_2 reduction catalyst. However, the structural and electronic properties of the catalyst, support, and promoter were not considered as input parameters in the experimental outputs. Incorporating these factors could provide valuable insights into the development of novel materials and improved catalytic systems.

In this study, a thermodynamic analysis was conducted to evaluate CH_4 production from industrial off-gases, specifically blast furnace gas (BFG) and basic oxygen furnace gas (BOFG), using Gibbs free energy minimization. Simulations were performed across a wide range of temperatures, pressures, H_2 flow rates, and off-gas compositions to determine the equilibrium distribution of products. The results identify optimal conditions for maximizing methane yield while minimizing carbon deposition, providing a robust thermodynamic baseline for process optimization. However, thermodynamic simulations alone cannot fully capture the complexity of catalytic behavior, which is strongly influenced by the physicochemical properties of catalyst components. Therefore, in parallel, machine learning was applied to an experimental catalyst performance dataset, enriched with atomistic

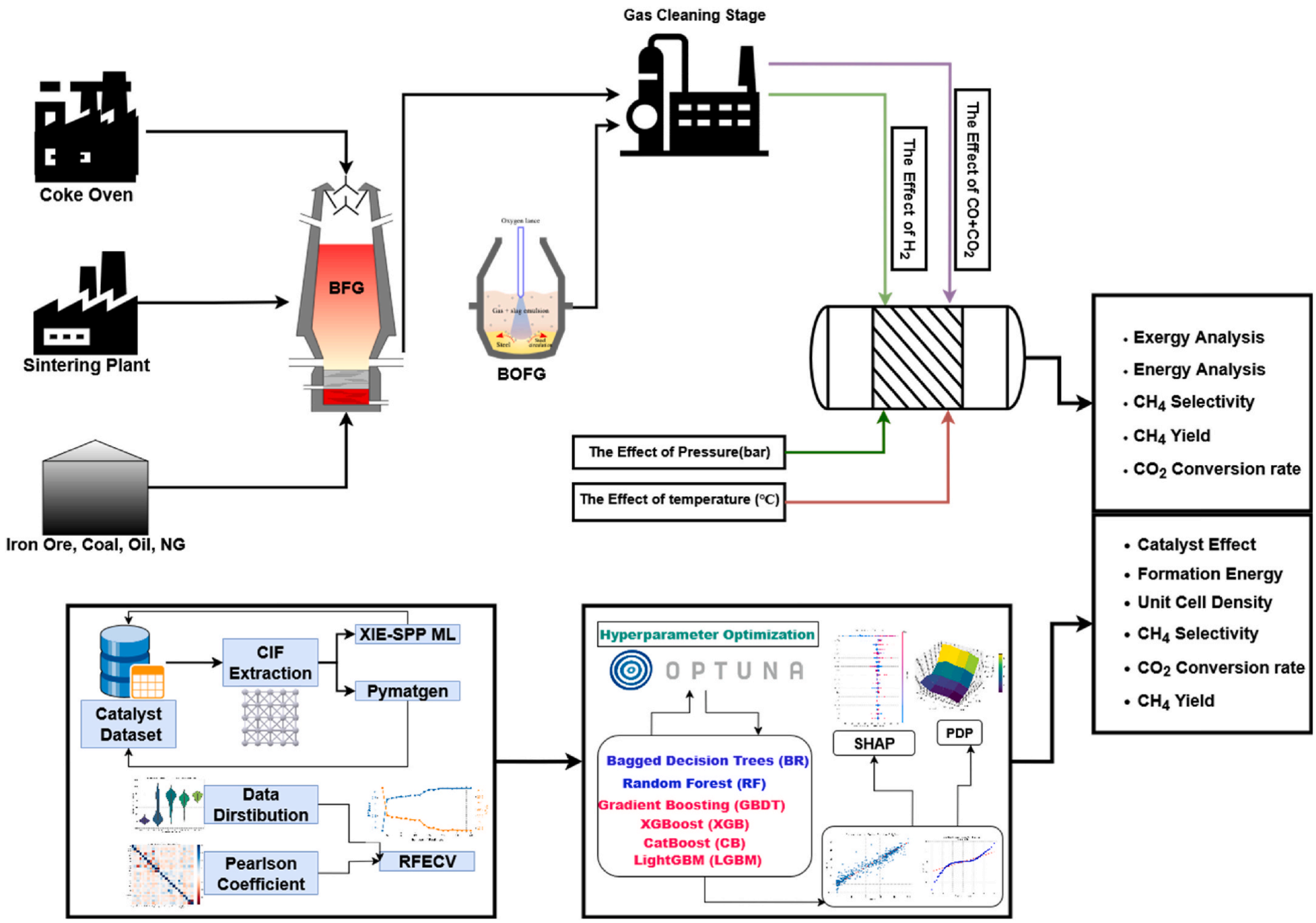


Fig. 1. Thermodynamic pathway for CO_2 methanation with parameter analysis (temperature, H_2/CO_2 ratio, catalyst features) and parallel ML-based catalyst performance prediction (RFECV, OPTUNA, SHAP).

descriptors such as crystal density and formation energy. This parallel strategy enabled the development of predictive models that reveal key factors governing catalyst activity. By integrating thermodynamic insights with data-driven screening, this study bridges process conditions with atomic-level catalyst design, considering the active phase, promoter, and support. The overall workflow is illustrated in Fig. 1.

2. Methods

2.1. *Thermodynamic analysis*

Thermodynamic equilibrium simulations followed by global sensitivity analysis were performed to evaluate COx methanation reactions under varying operating conditions. Calculations were performed using using HSC Chemistry (Version 10, Outotec) based on Gibbs free energy minimization to quantify the effects of temperature, pressure, H_2 flow rate, and off-gas feed rate on COx conversion, CH_4 yield, and energy and exergy efficiencies. The model incorporated key methanation reactions (Schumacher et al., 2023), and carbon formation reactions (Pham et al., 2022) (Table 1) using representative compositions of BFG and BOFG as feedstocks (Table S1). The equilibrium products considered included CH_4, CO, CO_2, H_2O, H_2, and solid carbon, representing the main species formed during CO and CO_2 methanation reactions (Hussain et al., 2021a).

2.2. *Model description*

A computational design of experiments was employed to investigate the influence of key input parameters; temperature, pressure, H_2 inlet flow rate, and off-gas feed rate, on methanation performance, as summarized in Table 2. The investigated parameter ranges were selected to capture the expected variability in operating conditions relevant to the methanation process (Sahebdelfar and Takht Ravanchi, 2015b). Temperature was varied from 50 °C to 500 °C, pressure from 1 to 30 bar, H_2 inlet flow rate from 0.5 to 2.0 $m^3\ h^{-1}$, and off-gas flow rate from 0.5 to 2.0 $m^3\ h^{-1}$. To ensure adequate resolution while maintaining computational feasibility, parameter values were discretized using a Saltelli sampling scheme (Saltelli, 2002), enabling robust variance-based sensitivity analysis with a total of 5000 model evaluations.The total number of simulations was calculated as $N \times (2k + 2)$, where N is the optimized sample size (500) and k is the number of input parameters (4), resulting in 5000 model evaluations. This sampling ensured reliable estimation of Sobol sensitivity indices. Equilibrium compositions for each sampled parameter set were determined by minimizing Gibbs free energy method (Hussain et al., 2021b). Equilibrium compositions were used to calculate COx conversions (Eqs. A.5,6) (Zhang et al., 2024), CH_4 yield (Eq. (A.7) (Song et al., 2024), and carbon formation via carbon balance (Eq. (A.8) (Etminan and Sadrnezhaad, 2022) were calculated. Energy and exergy analyses were conducted to evaluate the thermodynamic performance of COx methanation reactions, focusing on the exergy efficiency of gaseous products relative to reactants while excluding contributions from carbon formation and boundary losses to provide a clear reaction-level assessment (Chen et al., 2012). The total energy efficiency ($\eta_{\text{Total,En}}$) for the exothermic methanation reaction is calculated using Eqs. A9–11 (Martínez-Rodríguez and Abánades, 2020). The Total exergy efficiency ($\eta_{\text{Total, Ex}}$) is determined as the ratio of exergy recovered ($\eta_{\text{Ex,Out}}$) to the exergy input ($\eta_{\text{Ex,in}}$) as defined in Eqs. A.12-16 (Mendoza-Hernandez et al., 2019a). To evaluate the energy and exergy performance of CH_4 as a product, its energy and exergy values were calculated using Eq. (A.16) and Eq. (A.17), respectively (Mendoza-Hernandez et al., 2019b).

Table 2
Summary of input parameters and sampling method used in the thermodynamic equilibrium analysis.

Parameter	Range/ Method	Increment	Description
Temperature (°C)	$50 \le T \le 500$,	$\Delta T = 25$	Captures kinetic and thermodynamic limitations (Sahebdelfar and Takht Ravanchi, 2015c).
Pressure (bar)	$1.00 \le P \le 30.00$,	$\Delta P = 1.00$	Covers industrial operation range (Sahebdelfar and Takht Ravanchi, 2015c).
Inlet H_2 Flow Rate ($m^3.h^{-1}$)	$0.50 \le F \le 2.00$	$\Delta F = 0.25$	Stoichiometric to excess H_2 conditions ensuring the H_2: COx ratio varies between limiting and optimal regimes), which affects conversion and CH_4 yield (Hussain et al., 2021b).
Off-gas Flow Rate ($m^3.h^{-1}$)	$0.50 \le F \le 2.00$	$\Delta F = 0.25$	Feed variation impacts on conversion and yield (Hussain et al., 2021a).
Sampling Method	Saltelli	N = 500, total evaluations = 5000	Generates input samples for Sobol sensitivity analysis (Saltelli, 2002).

The sensitivity analysis was conducted using Sobol's method to quantify the contribution of each input parameter to the variance in outputs (Azzini et al., 2021). First-order Sobol indices (FSi) measured the main effect of a single input parameter is calculated using Eq. (A.18). Total-order Sobol indices (TSi) were also calculated to account for both individual contributions and interactions among parameters (Eq. (A.19). Both indices were computed using the SALib Python library, ensuring rigorous and reproducible implementation. This integrated approach enabled quantification of parameter influences, evaluation of energy and exergy efficiencies, and identification of optimal operating windows for COx methanation processes under industrially relevant conditions.

2.3. *Catalyst data collection and processing*

In this this study we utilized a comprehensive dataset of 2777 experimental observations for CO_2 methanation, curated from the work of Yang et al. (2024c). Each datapoint includes catalyst composition (identities and volumetric proportions of catalyst, support, and promoter), reaction conditions (temperature and pressure), and performance metrics (CO_2 conversion ratio, CH_4 yield, and CH_4 selectivity). The entire dataset was preserved without exclusions to maintain its integrity and representational diversity. To enhance the dataset's utility for predictive modeling and improve generalizability, we augmented it with fundamental physical and electronic descriptors. Recognizing the limitations of atomic weights for characterizing complex porous and crystalline catalysts, we introduced unit cell density as a structural

Table 1
Standard thermodynamic properties of COx methanation, carbon formation reactions at 25 °C and 1 bar.

No.		Reactions	ΔH ($KJ.mol^{-1}$)	ΔS (J.k˙1)	ΔG (KJ)	Log K
1	CO_2 methanation	$CO_{2(g)}+4H_{2(g)}\rightarrow CH_{4(g)}+2H_2O_{(g)}$	−164.755	−172.45	−113.33	19.86
2	CO methanation	$CO_{(g)}+3H_{2(g)}\rightarrow CH_{4(g)} + H_2O_{(g)}$	−205.88	−214.50	−141.93	24.87
3	RWGS	$CO_{(g)} + H_2O_{(g)}\rightarrow CO_{2(g)} + H_{2(g)}$	−41.14	−42.05	−28.60	4.45
4	Boudouard reaction	$2CO_{(g)}\rightarrow CO_{2(g)} + C_{(s)}$	−172.42	−175.81	−120.00	21.02
5	CO reduction	$CO_{(g)} +3H_{2(g)}\rightarrow H_2O_{(g)} + C_{(s)}$	−131.29	−133.77	−91.40	16.05
6	CO_2 reduction	$CO_{2(g)}+2H_{2(g)} \rightarrow H_2O_{(g)} + C_{(s)}$	−90.14	−91.72	−62.80	11.03

descriptor. Density (ρ) was calculated according to Eq. (1), where Σm_i is the total atomic mass within the unit cell, V_{cell} is the unit cell volume, and N_A is Avogadro's constant. These calculations were performed using the pymatgen library (Ong et al., 2013), following the parsing of crystallographic information files (CIFs) with the Atomic Simulation Environment (ASE) toolkit (Hjorth Larsen et al., 2017).

$$\rho = \frac{\sum m_i}{V_{cell} \times N_A} \tag{1}$$

Formation energy (Zhang et al., 2018), a key electronic descriptor indicative of thermodynamic stability and catalyst longevity, was systematically integrated using a prioritized sourcing strategy. Formation energies for crystalline materials were first retrieved directly from the Materials Project database via its API where available (Jain et al., 2013). For compounds absent from this database, we prioritized sourcing both formation energies and densities from ab initio calculations reported in peer-reviewed literature. Subsequently, for materials lacking literature-derived density values, density was computed using the established ASE and pymatgen methodology. Finally, formation energies for the remaining materials without literature or database values were predicted using the XIE-SPP deep learning framework (Davariashtiyani et al., 2021), which derives this property directly from crystal structures. Structural inputs for XIE-SPP predictions were sourced from the Crystallography Open Database (COD) (Vaitkus et al., 2023).

2.4. Machine learning catalyst performance prediction

Following data preprocessing, we implemented a systematic three-phase methodology to construct robust ensemble learning models for predicting CO_2 methanation performance, as illustrated in Fig. 2. (I) Feature selection was conducted using Pearson correlation and Recursive Feature Elimination with Cross-Validation (RFECV) to eliminate redundant variables, ensuring a concise and relevant feature set. (II) Diverse ensemble learning models, encompassing bagging and boosting paradigms, were developed to capture varied predictive patterns. (III) Iterative hyperparameter optimization was performed using Bayesian methods to enhance model accuracy and robustness while maintaining computational efficiency. To evaluate model performance, we employed an 80-20 train-test split, where 80 % of the dataset was used for training to fit the models and 20 % was reserved for testing to assess their predictive generalization on unseen data. This structured workflow ensures optimal feature representation, algorithmic diversity, and reliable model performance.

I. Feature Selection Using Pearson Correlation and RFECV

The feature selection process was conducted in two stages to ensure optimal model performance while minimizing redundancy. First, the Pearson correlation coefficient (PCC) was employed to evaluate linear relationships between input features (Adler and Parmryd, 2010). Highly correlated feature pairs ($|r| \approx 1$) were identified, and one feature from each pair was removed to reduce dimensionality and multicollinearity. This step enhances computational efficiency and model interpretability by eliminating redundant information.

Subsequently, recursive feature elimination with cross-validation (RFECV) was applied to further refine the feature set. Unlike standard feature selection methods, RFECV iteratively removes the least important features while assessing model performance through cross-validation (Yao et al., 2024). This approach ensures robustness against

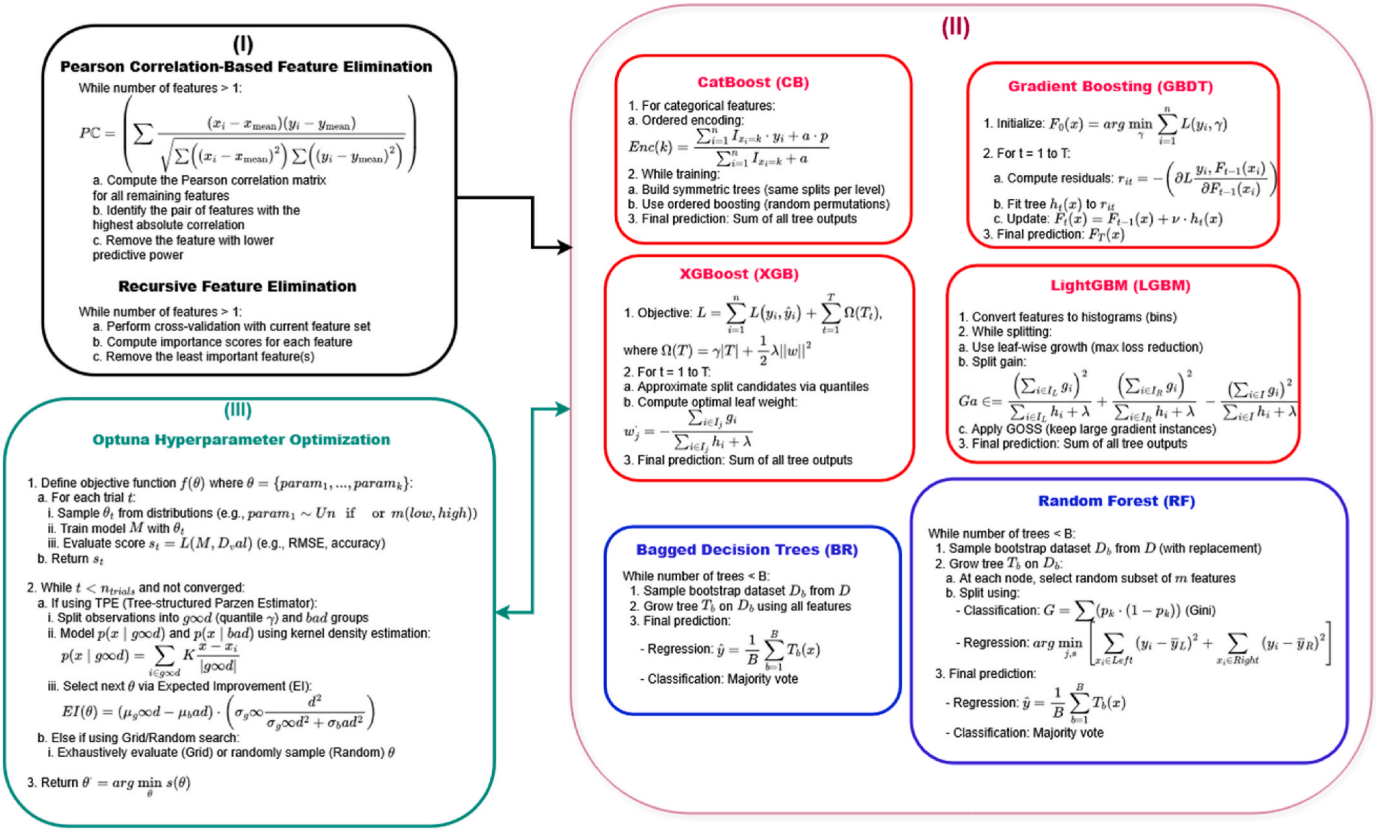


Fig. 2. Flowchart illustrating the three-stage machine learning pipeline: (I) Feature selection using Pearson correlation and RFECV to eliminate redundancy, (II) Ensemble model development comparing bagging (RF, BR) and boosting (GBDT, LGBM, XGB, CB) algorithms, and (III) Hyperparameter optimization via Bayesian tuning with Optuna, forming an iterative refinement loop between stages II and III. The process yields an optimized predictive model with minimized multicollinearity and maximized generalization capability.

overfitting and automatically determines the optimal number of features without manual intervention. By combining PCC and RFECV, the final feature subset maximizes predictive power while maintaining generalization capability.

II. Ensemble Learning Model Development

Six ensemble learning models; Random Forest (RF) (Breiman, 2001), Bagged Decision Tree Regressor (BR) (Breiman, 1996), Gradient Boosting Decision Tree (GBDT), LightGBM (LGBM), XGBoost (XGB), and CatBoost (CB) (Hastie et al., 2009), were implemented to evaluate their suitability for the CO_x methanation process. These models were selected to cover both bagging (RF, BR) and boosting (GBDT, LGBM, XGB, CB) frameworks, ensuring a comprehensive comparison of ensemble techniques. Bagging methods improve stability through bootstrap aggregation, while boosting algorithms iteratively optimize prediction accuracy by correcting errors from previous iterations. The diversity of these models allows for a rigorous assessment of their adaptability to the dataset's unique characteristics.

III. Hyperparameter Optimization via Bayesian Tuning

To maximize model performance, hyperparameter optimization was conducted using Optuna (Akiba et al., 2019), a Bayesian optimization framework. Optuna advances hyperparameter optimization by dynamically refining the search space based on intermediate model outcomes, outperforming traditional grid or random search methods in efficiency. Integrated within an iterative model development framework (Section II), initial training precedes systematic tuning of critical hyperparameters, including learning rate, tree depth, and regularization terms, to balance bias and variance. Models are evaluated using 500 iterations until convergence, ensuring optimized performance tailored to the problem domain. This approach ensures that the final models achieve high predictive accuracy while maintaining computational feasibility. The workflow between model development (II) and hyperparameter tuning (III) forms a closed-loop system, where each iteration refines model performance. After hyperparameter optimization, the best-performing configuration is selected, and the process repeats if further improvements are warranted. This iterative refinement ensures robustness and generalizability, ultimately yielding the most suitable model for the given application.

3. Results and discussion

3.1. Thermodynamic equilibrium characterization of CO_x methanation

The fundamental objective of thermodynamic analysis is to predict the equilibrium composition of a reaction system based on the thermodynamic properties of its components (Hussain et al., 2021a). As illustrated in Fig. S1, the Gibbs free energy and equilibrium constants for CO and CO_2 methanation confirm that these reactions particularly CO methanation, have the high corresponding K values below 400 °C, along with a high negative ΔG values. In contrast, the water-gas shift and Boudouard reactions are less thermodynamically favorable under these conditions, underscoring the importance of low-temperature operation with catalysts active for CO and CO_2 methanation reactions to maximize methane production, which agrees well with the stand points from catalyst systems reported before (Díez-Ramírez et al., 2017). This thermodynamic foundation enables a detailed assessment of the equilibrium behavior under varying operating conditions. The effects of the reactor temperature, pressure, H_2 inlet, and the off-gas ratio were studied, and the results are shown in Fig. S2, and Fig. S3. Equilibrium results presented in Fig. S2(a, b) indicate that increasing the H_2 inlet flow rate (~1.5–1.7 $m^3\ h^{-1}$) significantly enhances CH_4 production while suppressing carbon formation. The maximum CH_4 concentration is observed at a moderate off-gas flow rate (~1.0 $m^3\ h^{-1}$), suggesting optimal H_2 to carbon feed ratios under these conditions. At low H_2 flow rates (<1.0 $m^3\ h^{-1}$) and high off-gas feed rates (>1.5 $m^3\ h^{-1}$), carbon deposition exceeds 24 %, indicating unfavorable conversion conditions. This behavior is consistent with the Boudouard and carbon reduction reactions listed in Table 1, where excess off-gas and limited H_2 supply favor carbon formation via hydrogen-assisted reduction pathways. CO_2 residuals in product gases increase under off-gas dilution at low H_2: CO_2 ratios (Fig. S2c), while CO concentrations remain negligible, indicating effective CO methanation (Fig. S2d). As shown in Fig. S3(c and d), lower reactor temperatures further suppress CO_2 residuals without affecting CO levels. CH_4 production is maximized at temperatures below 250 °C and pressures above 20 bars (Fig. S3a), whereas carbon formation peaks at 250–350 °C and 25–30 bar (Fig. S3b). At low off-gas flow rates (~0.5 $m^3\ h^{-1}$) and high H_2 inlet flows (~1.5–1.7 $m^3\ h^{-1}$), unreacted hydrogen increases significantly, indicating that excess H_2 enhances residual H_2 levels. These results collectively suggest that optimal CH_4 production with minimal carbon formation occurs at H_2 inlet flows around 1.5 $m^3\ h^{-1}$ and off-gas flows between 0.5 and 0.8 $m^3\ h^{-1}$. Few studies have explored thermodynamic analysis in CO and CO_2 methanation; therefore, Sobol sensitivity analysis was conducted to quantify the influence of key thermodynamic parameters on methanation performance. Results showed that methane production exhibited the highest sensitivity to H_2 inlet flow (TSi: 0.85), highlighting hydrogen availability as the primary determinant of CH_4 yield (Fig. 3a). Fig. 3b presents that carbon formation was predominantly influenced by off-gas flow (TSi: 0.85), indicating that higher dilution enhances carbon deposition tendencies. CO_2 flow was governed mainly by H_2 inlet flow (TSi: 0.65) and off-gas flow (0.55), reflecting their roles in determining CO_2 conversion efficiency (Fig. 3c). As shown in Fig. 3d, CO residuals were most sensitive to reactor temperature (TSi: 1.0), underscoring temperature's critical importance for CO conversion. The Sobol sensitivity indexes in Fig. 3e and f, demonstrate that excess H_2 in the product was primarily affected by off-gas flow (TSi: 0.66) and H_2 inlet flow (TSi: 0.59), while H_2O production was governed by off-gas flow (TSi: 0.75), consistent with its dilution-driven behavior.

3.2. Parametric effects on CO_x conversion and carbon formation

The effects of reaction conditions on CO_2 and CO conversion, along with their corresponding Sobol sensitivity indices, are presented in Fig. 4. CO_2 conversion decreased from approximately 92 % at 50 °C to 74 % at 450 °C, exhibited negligible dependence on pressure, declined notably with increasing off-gas flow (from 98 % at 0.5 $m^3\ h^{-1}$ to 68 % at 2.0 $m^3\ h^{-1}$), and increased significantly with higher H_2 inlet flow, reaching nearly complete conversion at 1.7 $m^3\ h^{-1}$. This behavior is corroborated by the sensitivity analysis, which identified H_2 inlet flow as the most influential parameter (FSi ~0.55; TSi~0.75), followed by off-gas flow (FSi ~0.25; TSi ~0.36), while temperature and pressure had minimal effects. As it can been seen in Fig. S4a high H_2 inlet flow combined with low off-gas flow yields near-complete conversion, while Fig. S5a confirms that CO_2 conversion is more sensitive to temperature than pressure, with moderate temperatures (150–250 °C) generally promoting higher conversion across the full pressure range. This trend aligns with previous thermodynamic analysis of CO_2 and CO methanation reactions by Gao et al. (2012). who reported that high H_2/CO_2 ratios ($H_2/CO_2 \geq 4$) significantly enhance CO_2 conversion. CO conversion remained consistently high (>99.5 %) under all conditions, with a slight decrease to ~99.1 % at temperatures above 400 °C, attributable to thermodynamic limitations. Sensitivity analysis confirmed reactor temperature as the dominant factor (FSi ~0.45; TSi ~1.0), while pressure and flow rates had minimal influence. This agrees with Gao et al. (2012) reporting that limited sensitivity to pressure and H_2/CO ratio below 700 °C. Moreover, the reliability of the results is confirmed by Fig. S4b, and Fig. S5b, showing consistently high CO conversion (>99.5 %) across temperature ≤400 °C, and H_2–off-gas flow combinations ranges, with only slight reductions observed at high temperatures and

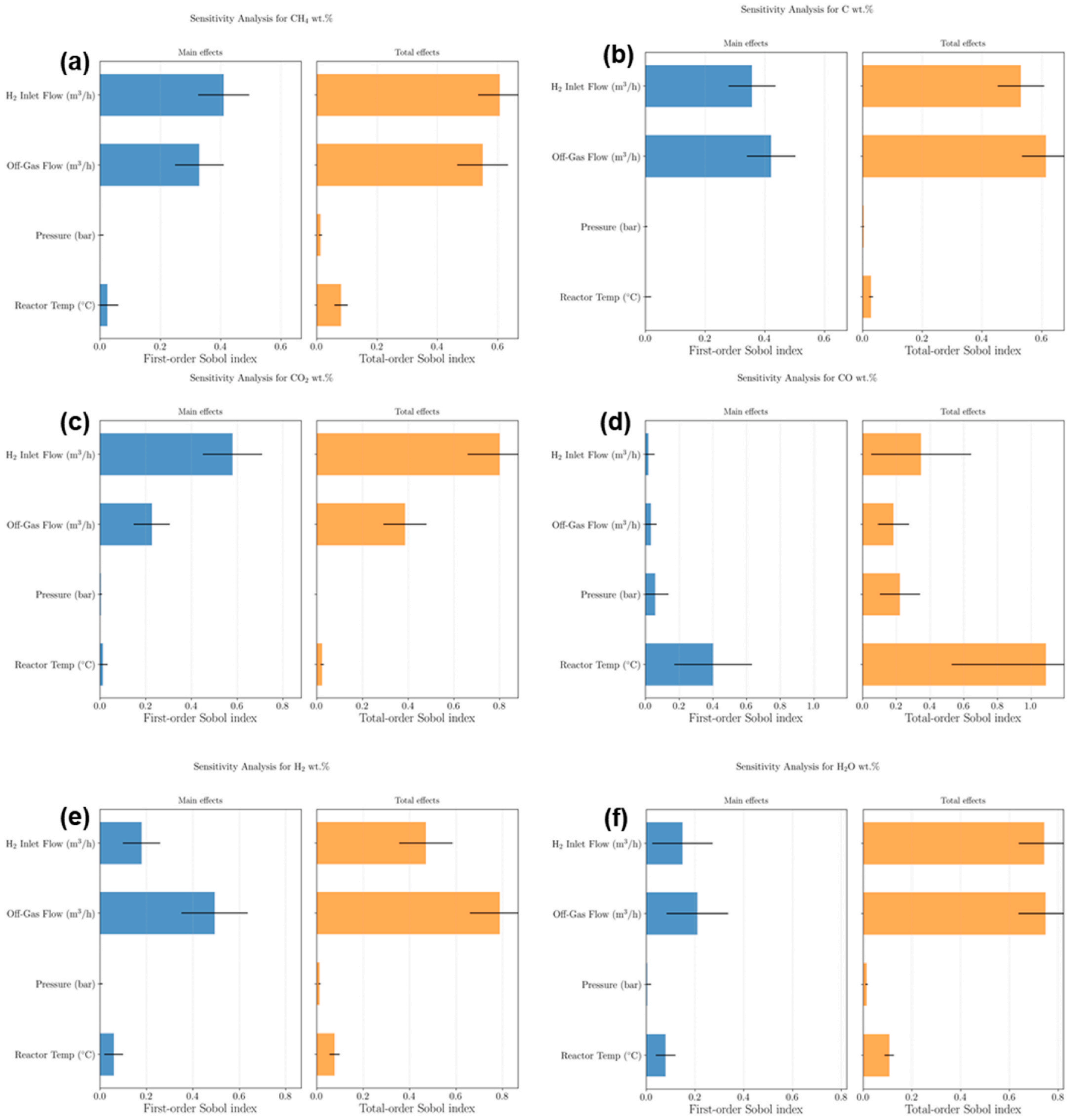


Fig. 3. Sobol sensitivity indices illustrating the influence of temperature (T), pressure (P), H_2 inlet flow rate ($H_{2,in}$), and off-gas flow rate on the equilibrium weight fractions of (a) CH_4, (b) solid carbon (C), (c) CO_2, (d) CO, (e) H_2, and (f) H_2O.

low H_2 flow, supporting the dominant influence of temperature.

Evaluation of CH_4 yield revealed relative stability (~40–48 %) up to 350 °C, followed by a decrease to ~30 % at 450 °C, as shown in Fig. 5a. Pressure had negligible effect, maintaining yields around ~40–45 %. Increasing off-gas flow from 0.5 to 2.0 m^3 h^{-1} caused a substantial decline in CH_4 yield from ~85 % to ~10 %, whereas increasing H_2 inlet flow from 0.5 to 1.7 m^3 h^{-1} markedly increased yield from ~5 % to ~75 %. Sensitivity analysis (Fig. 5c) identified H_2 inlet flow and off-gas flow as the dominant factors influencing CH_4 yield (FSi ~0.45; TSi ~0.55), while temperature and pressure had minimal effects. These results support the finding that CH_4 yield is maximized at H_2 inlet flows of ~1.5–1.7 m^3 h^{-1} combined with low off-gas flows (~0.5 m^3 h^{-1}) (Fig. S4c). The findings obtained are quite compatible with previous thermodynamic analyses, which demonstrate that increasing pressure below 500 °C has little effect on CH_4 yield (Yarbaş and Ayas, 2024). Considering that carbon deposition is a critical cause of catalyst deactivation, its equilibrium behavior is evaluated, with results presented in Fig. 5b–d. Carbon balance remained stable at ~0.5 % across temperature and pressure ranges but decreased substantially from ~0.85 % to ~0.3 % with increasing off-gas flow (0.5–2.0 m^3 h^{-1}). In contrast,

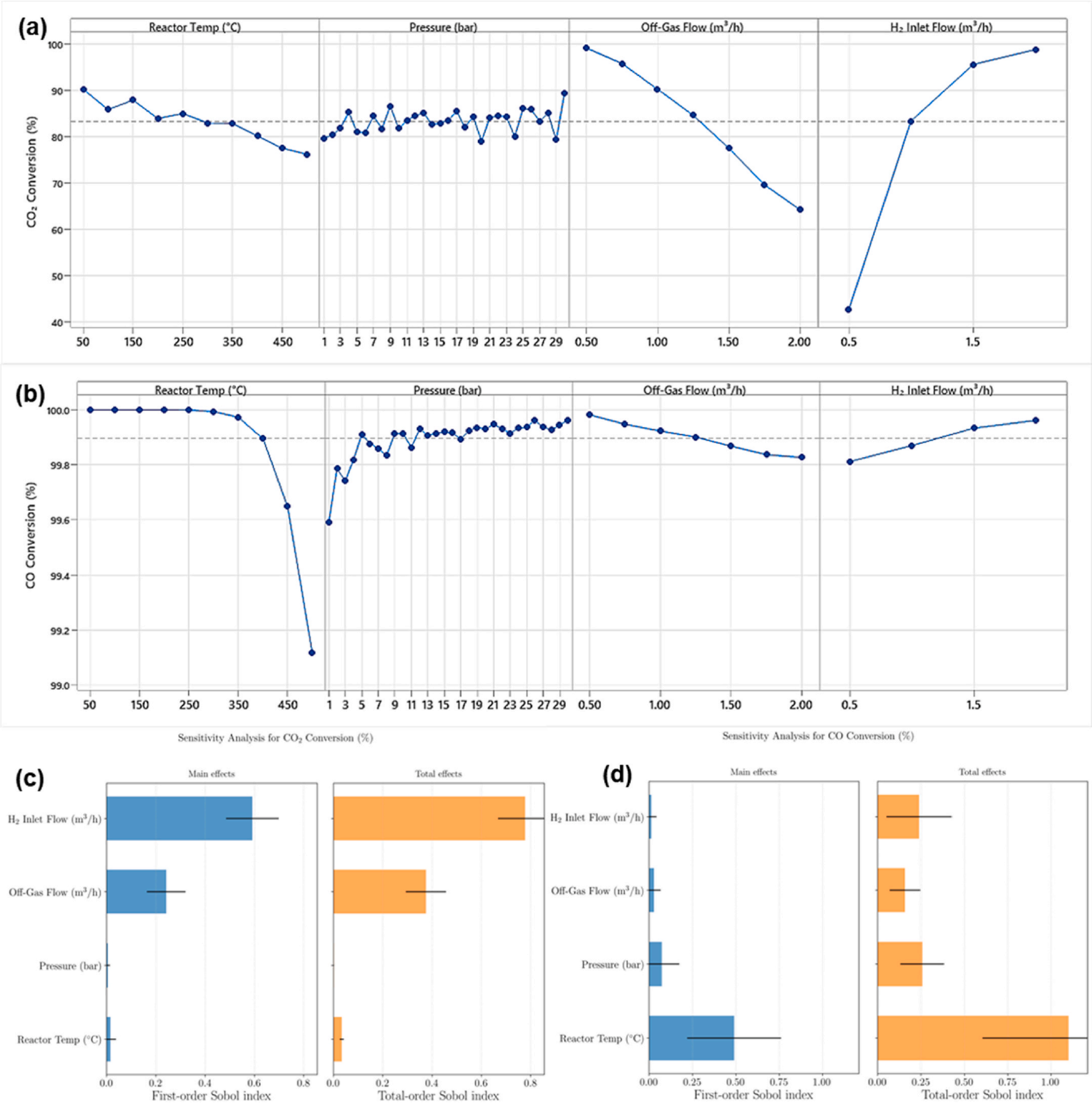


Fig. 4. (a) CO_2, and (b) CO conversion as functions of temperature, pressure, H_2 inlet flow, and off-gas flow, corresponding Sobol sensitivity indices of (c) CO_2, (d) CO conversion.

increasing H_2 inlet flow (0.5–1.7 $m^3\ h^{-1}$) raised carbon balance from ~0.3 % to ~0.8 %. Sensitivity analysis confirmed off-gas flow (TSi ~0.6) as the most influential parameter, followed by H_2 inlet flow (TSi~0.55). Fig. S4d and S5d further support these trends by demonstrating increased carbon balance and yield at high H_2 inlet flows and low off-gas flows. This trend aligns with Ling et al. (2024), who reported that carbon deposition in methanation systems is predominantly influenced by feed composition and increases with lower H_2/CO_x ratios, highlighting the critical role of feed composition in suppressing carbon formation and maintaining catalyst activity.

3.3. Parametric effects of energy and exergy efficiencies

To evaluate process performance of the CO_x methanation reaction, total energy and total exergy efficiencies were assessed under varying operational conditions. As presented in Fig. 6a the total energy efficiency increased with temperature, rising from ~39 % at 50 °C to ~50 % at 450 °C, while pressure effects remained minimal (~42–44 %). Increasing off-gas flow from 0.5 to 2.0 $m^3\ h^{-1}$ reduced total energy efficiency markedly from ~50 % to ~35 %, whereas higher H_2 inlet flow (0.5–1.7 $m^3\ h^{-1}$) improved efficiency from ~37 % to ~45 %. Sensitivity analysis (Fig. 6c) identified off-gas flow (TSi ~0.7) as the most influential parameter, followed by H_2 inlet flow (TSi~0.6), with negligible

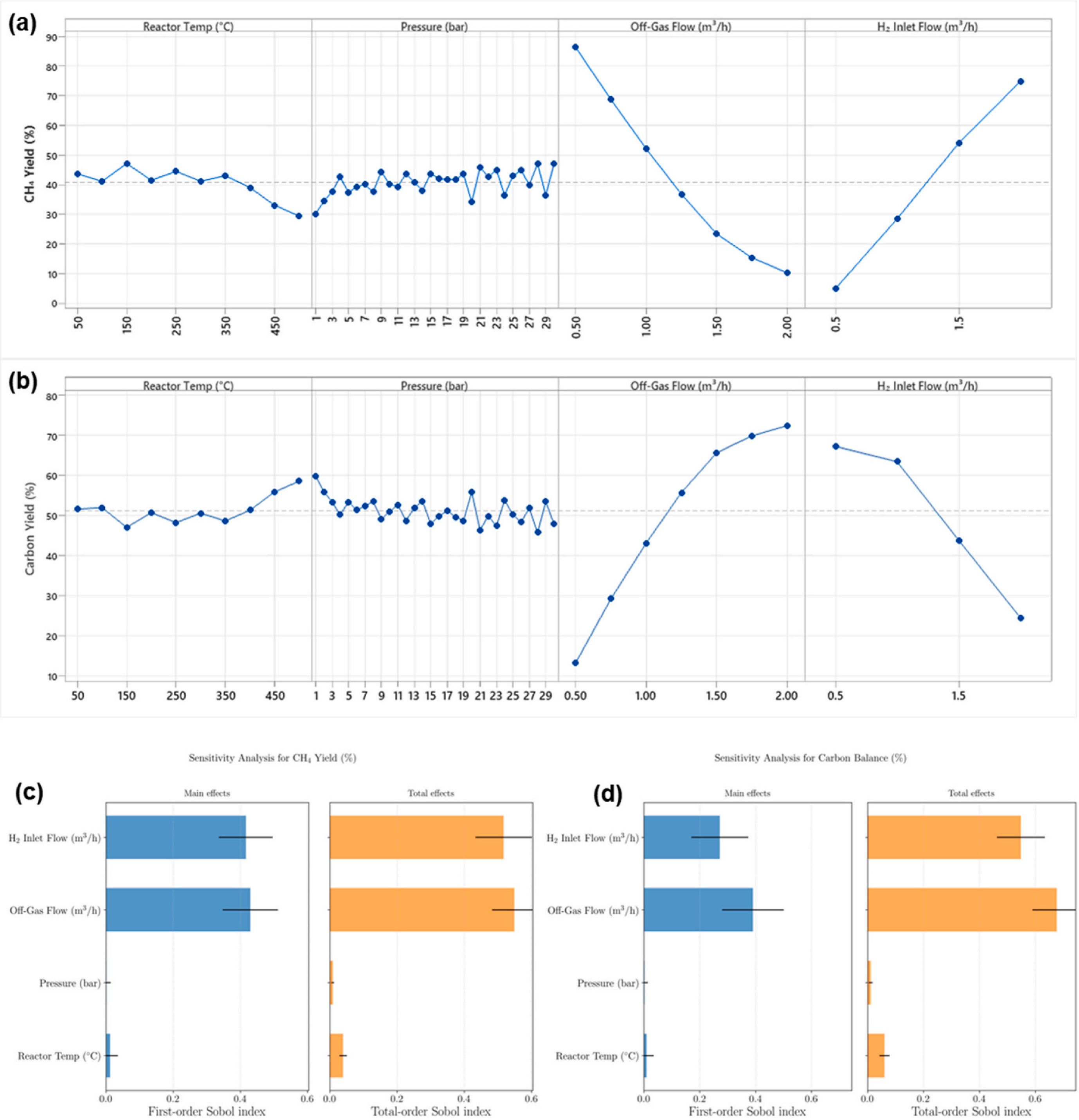


Fig. 5. (a) CH_4 yield, and (b) Carbon balance as functions of temperature, pressure, H_2 inlet flow, and off-gas flow, with corresponding Sobol sensitivity indices of (c) CH_4 yield, (d) Carbon balance.

contributions from temperature and pressure (TSi< 0.2). These trends are further supported by Fig. S6a and Fig. S7a, which show that maximum total energy efficiency is achieved at high temperatures (~450 °C), high H_2 inlet flows (~1.8 m^3 h^{-1}), and low off-gas flows (~0.5 m^3 h^{-1}). Consistent with the trends observed for total energy efficiency, total exergy efficiency also demonstrated strong dependence on feed flow rates. The total exergy efficiency results (Fig. 5b) showed stability at ~38–42 % across temperature and pressure ranges, but efficiency increased significantly with higher H_2 inlet flow (from ~25 % at 0.5 m^3 h^{-1} to ~50 % at 1.7 m^3 h^{-1}) and decreased sharply with increasing off-gas flow (from ~70 % at 0.5 m^3 h^{-1} to ~25 % at 2.0 m^3 h^{-1}). Sensitivity analysis confirmed H_2 inlet flow (TSi ~0.55) and off-gas flow (TSi~0.49) as the dominant factors, while temperature and pressure effects remained negligible (indices <0.05). These findings are further supported by Fig. S6b, and Fig. S7b, which show total exergy efficiency peaking at ~96 % under high H_2 inlet flow (~1.8 m^3 h^{-1}) and low off-gas flow (~0.5 m^3 h^{-1}), highlighting the critical role of feed composition and dilution in maximizing process efficiency.

Assessment of CH_4 energy efficiency (Fig. 7a) under varying operating conditions showed limited sensitivity to temperature (17–22 %)

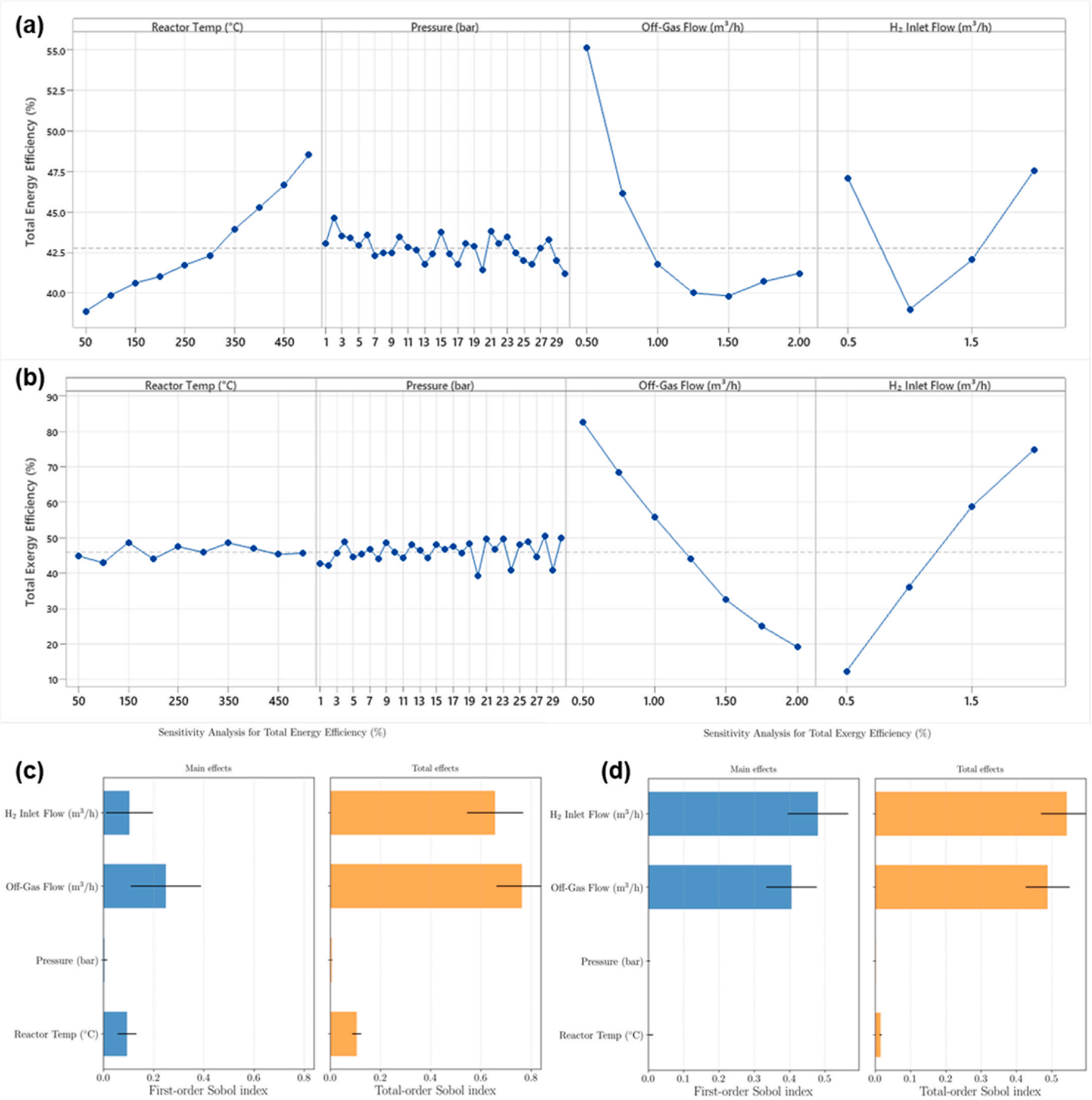


Fig. 6. (a) Total energy efficiency, (b) Total exergy efficiency as a function of temperature, pressure, H_2 inlet flow, and off-gas flow, corresponding Sobol sensitivity indices of (c) Total energy efficiency, (d) Total exergy efficiency.

and pressure (15–22 %). However, efficiency increased with higher H_2 inlet flow (from ~5 % at 0.5 m^3 h^{-1} to ~29 % at 1.7 m^3 h^{-1}) and decreased with increasing off-gas flow (from ~27 % at 0.5 m^3 h^{-1} to ~9 % at 2.0 m^3 h^{-1}). Sensitivity analysis identified H_2 inlet flow (TSi ~0.65) and off-gas flow (TSi~0.5) as the dominant factors. Results in Fig. S6e, and Fig. S7e indicate that maximum CH_4 energy efficiency (~35–37 %) is achieved at H_2 inlet flows of ~1.7–2.0 m^3 h^{-1} with low off-gas flows (~0.5–0.8 m^3 h^{-1}), emphasizing the critical role of feed composition optimization. Building upon the CH_4 energy efficiency analysis, CH_4 exergy efficiency was also assessed to evaluate thermodynamic performance under varying conditions. Results (Fig. 7b) showed that exergy efficiency remained stable at ~45–50 % across the temperature range of 50–450 °C and pressures of 1–30 bar. However, increasing off-gas flow from 0.5 to 2.0 m^3 h^{-1} substantially reduced exergy efficiency from ~83 % to ~20 %, while higher H_2 inlet flow (0.5–1.7 m^3 h^{-1}) markedly enhanced it from ~13 % to ~75 %. Sensitivity analysis (Fig. 7d) identified H_2 inlet flow (TSi ~0.68) as the most influential parameter, followed by off-gas flow (TSi~0.5), with temperature and pressure showing negligible effects. This behavior is further supported by Fig. S6d and Fig. S7d, which indicate that exergy efficiency is maximized at H_2 inlet flows of ~1.5–2.0 m^3 h^{-1} combined with low off-gas flows (~0.5 m^3 h^{-1}), irrespective of temperature and pressure variations. Notably, CH_4 exergy efficiency exhibited similar parametric dependencies to CH_4 energy efficiency but achieved higher absolute

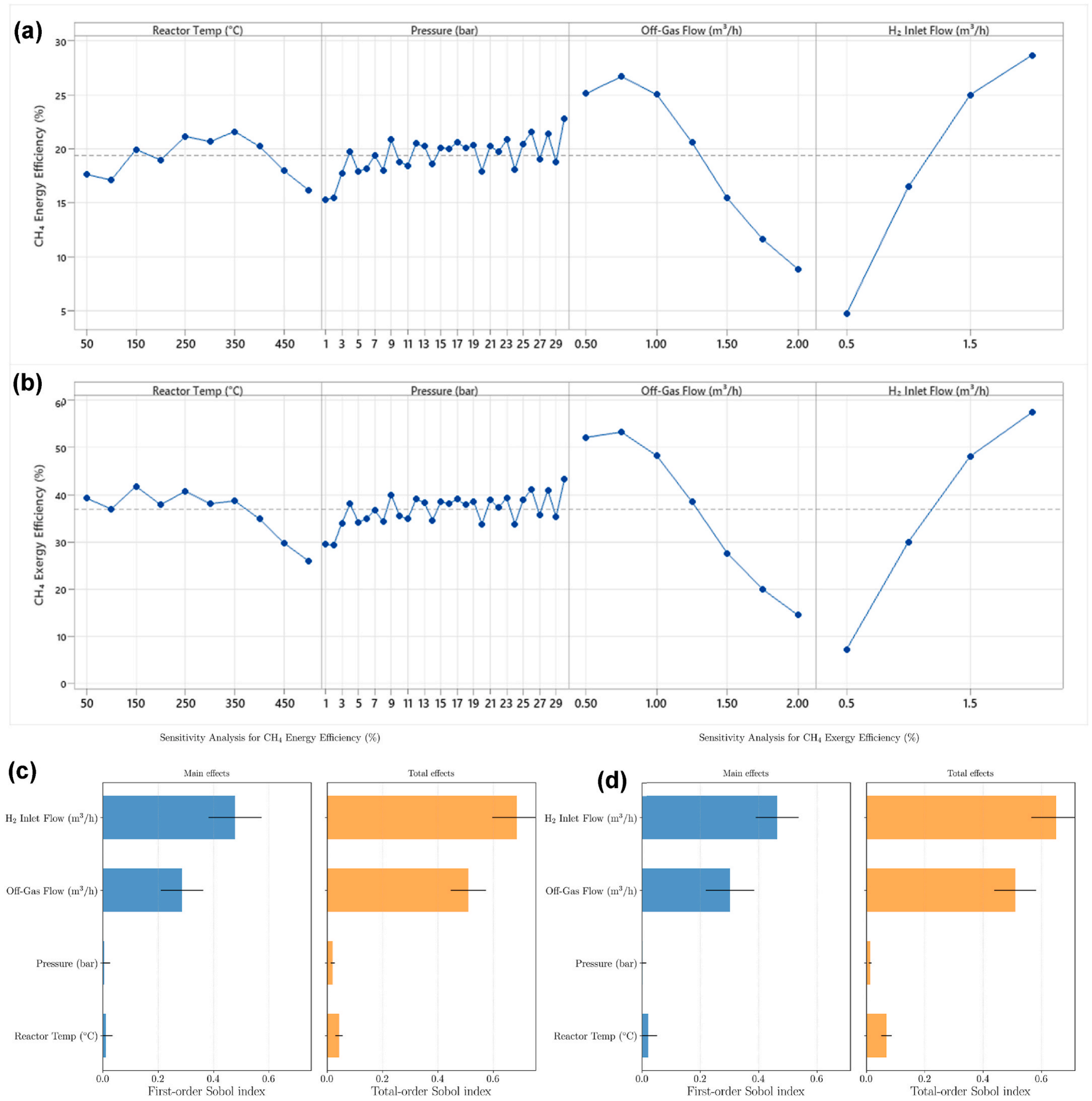


Fig. 7. (a) CH_4 energy efficiency, and (b) CH_4 exergy efficiency as functions of temperature, pressure, H_2 inlet flow, and off-gas flow, corresponding Sobol sensitivity indices of (c) CH_4 energy efficiency, and (d) CH_4 exergy efficiency.

values, highlighting enhanced thermodynamic utilization under optimized feed conditions.

3.4. Machine learning results

3.4.1. Correlation analysis and feature selection

The correlation analysis (Fig. 8 and Fig. S7-8) revealed fundamental relationships between catalyst properties and performance in CO_2 methanation, demonstrating how intrinsic material characteristics govern catalytic behavior. Thermodynamic stability parameters showed consistent negative correlations with activity metrics, as evidenced by the moderate negative correlations of both active component formation energy (ACFE, PCC $= -0.42$) and promoter formation energy (PFE, PCC $= -0.38$) with CO_2 conversion. These results indicate that more thermodynamically stable phases (lower formation energies) yield higher activity, consistent with established principles of catalytic materials where structural stability under reaction conditions maintains active site integrity. Density parameters exhibited distinct positive correlations with selectivity metrics, with active component density (ACD, PCC $=$ 0.35) and support densities (SAD/SBD, PCC $=$ 0.18–0.22) all showing favorable relationships with CH_4 selectivity. This pattern suggests higher density materials promote selective pathways, likely through more uniform surface geometries that favor CH_4 formation transition states and stronger metal-support interactions that stabilize

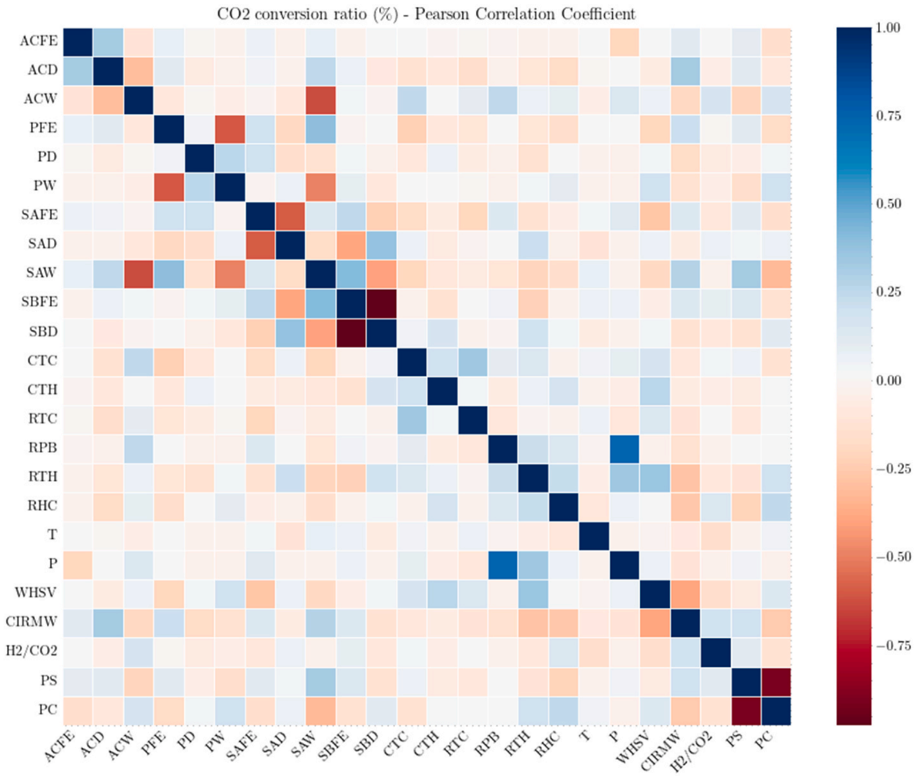


Fig. 8. Pearson correlation matrix (r-values) for process parameters affecting CO_2 conversion.

intermediate species. The more pronounced effects for active components compared to supports indicate the dominant role of active phase microstructure in controlling selectivity. A clear hierarchy emerged regarding parameter importance, with formation energies and densities demonstrating consistently stronger correlations than composition variables (ACW, PW, SAW all PCC <0.25). This hierarchy reflects fundamental materials science principles, where formation energies govern thermodynamic stability and densities reflect atomic packing - both being intrinsic characteristics that directly determine catalytic behavior. In contrast, composition primarily affects the relative abundance of these already-defined phases.

Preparation parameters further reinforced these trends, particularly the reduction conditions which showed strong correlations with performance. The reduction temperature (RTC, PCC $= 0.45$) and hydrogen content (RHC, PCC $= 0.51$) significantly influenced catalyst behavior, consistent with their known roles in determining crystallite size distribution and metal dispersion. These findings align with established catalyst synthesis principles where reduction conditions critically impact final catalyst morphology and oxidation state. The analysis systematically connects fundamental material properties to catalytic performance, demonstrating that formation energies and densities serve as more reliable predictors than compositional factors within the studied ranges.

The recursive feature elimination analysis reveals insights about parameter dependencies in catalytic CO_2 conversion processes. Across all six machine learning models, CO_2 conversion prediction consistently required 21–24 input features, demonstrating the complex multifactorial nature of this transformation. This broad feature dependence persists even in regularized algorithms like XGBoost and LightGBM, suggesting that simplified representations may inadequately capture the intricate network of physicochemical interactions governing CO_2 activation. The high dimensionality contrasts with the more selective feature requirements for CH_4-related predictions, particularly for selectivity where CatBoost achieved optimal performance with only 19 features while maintaining superior predictive accuracy (RMSE $= 0.07$ compared to 0.11–0.12 for other models). This 21 % reduction in parameter space indicates CH_4 selectivity depends more strongly on a constrained set of key variables, likely those directly involved in hydrogenation pathways such as H_2 partial pressure and metal dispersion characteristics.

Fig. 9a–f illustrates distinct algorithmic trajectories in balancing predictive accuracy against feature parsimony. Random Forest and Bagging Regressor achieved competitive performance ($R^2 \approx 0.80$–0.82); however, this required the full feature set (23–24 variables) and was accompanied by indications of overfitting beyond 20 features, as reflected by an 18–22 % increase in Root Mean Square Error (RMSE). In contrast, gradient-boosting architectures demonstrated superior efficacy. XGBoost and LightGBM attained peak R^2 values of 0.85–0.88 using more parsimonious feature subsets (21–23 variables), a result attributable to their inherent regularization techniques which more effectively navigate the accuracy-complexity tradeoff. Notably, CatBoost proved particularly efficient for methane (CH_4) prediction, achieving optimal accuracy with the fewest features. This efficiency stems from its specialized treatment of categorical variables and ordered boosting protocol, which enhances the isolation of critical predictive parameters.

3.4.2. *Evaluation of ML models*

To evaluate the predictive performance of the machine learning models, we utilized four primary evaluation metrics: Root Mean Square Error (RMSE) (Table 3), Coefficient of Determination (R^2) (Table 3), Quantile-Quantile (Q-Q) plots (Fig. S13–S15), and Actual versus Predicted plots (Figs. 10–12). RMSE quantifies the magnitude of prediction errors by calculating the square root of the mean of squared differences between observed and predicted values, with lower values indicating higher predictive accuracy. R^2 measures the proportion of variance in the dependent variable explained by the model, where values approaching 1 reflect stronger explanatory capability. Q-Q plots were employed to assess the normality of model residuals by comparing their distribution to a theoretical normal distribution, with deviations from the diagonal line indicating potential violations of model assumptions.

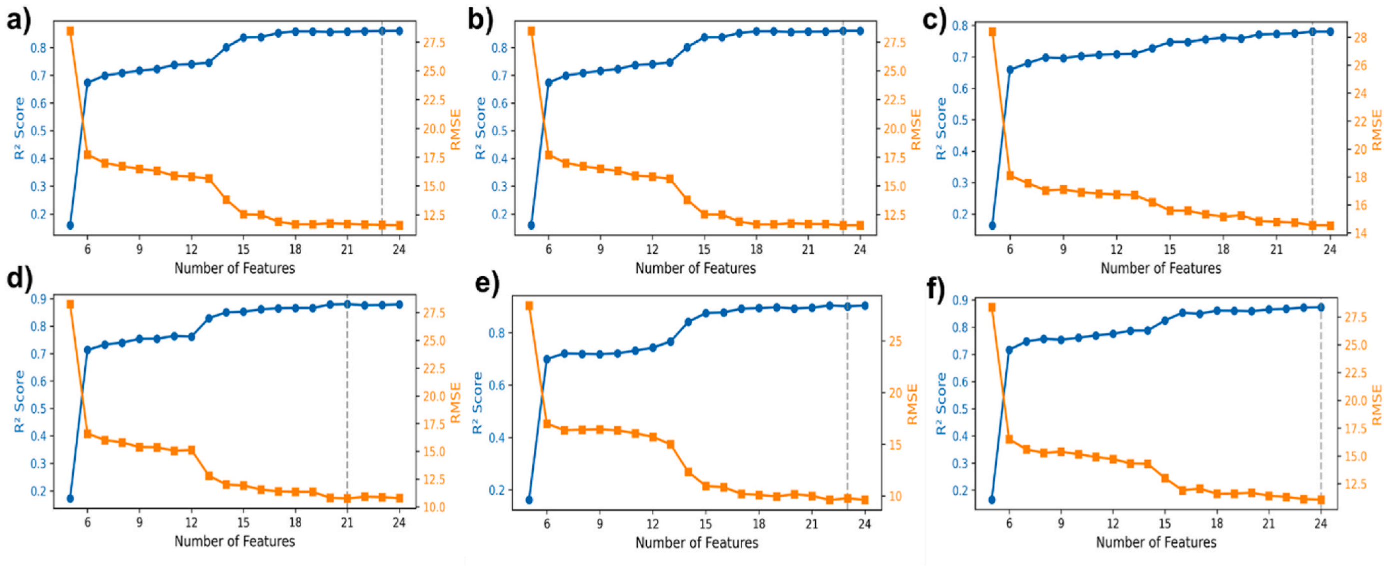


Fig. 9. Recursive Feature Elimination with Cross-Validation (RFECV) performance comparison for CO_2 conversion prediction across six machine learning approaches: (a) Random Forest (RF), (b) Bagging Regressor (BR), (c) Gradient Boosted Decision Trees (GBDT), (d) Light Gradient Boosting Machine (LGBM), (e) Extreme Gradient Boosting (XGBoost), and (f) Categorical Boosting (CatBoost). Results demonstrate the optimal feature subset selection for each algorithm, evaluated using 5-fold cross-validation with mean squared error (MSE) scoring metric.

Table 3
Performance Metrics and Optuna Best R^2 Values for Machine Learning Models Predicting CO_2 Conversion, CH_4 Selectivity, and CH_4 Yield. The table reports R^2 and RMSE for each model across the three target variables, alongside the best R^2 values obtained from Optuna cross-validation.

Metric	Random Forest	Bagging Regressor	Gradient Boosting Decision Tree	LightGBM	XGBoost	CatBoost
CO_2 Conversion R^2	0.8767	0.8755	0.9068	0.9253	0.9384	0.9462
CO_2 Conversion RMSE	10.7986	10.8522	9.3879	8.4035	7.6359	7.1363
CH_4 Selectivity R^2	0.9187	0.9166	0.9017	0.9	0.9374	0.9018
CH_4 Selectivity RMSE	7.1657	7.255	7.8762	7.9468	6.2858	7.8737
CH_4 Yield R^2	0.8893	0.8889	0.9085	0.9325	0.9447	0.9362
CO_2 Conversion Optuna R^2	10.2932	10.3089	9.3568	8.0354	7.2743	7.8111
CH_4 Selectivity Optuna R^2	0.8615	0.8609	0.9998	0.9113	0.9242	0.9262
CH_4 Yield Optuna R^2	0.8486	0.8471	0.9958	0.8552	0.8725	0.8708
CO_2 Conversion Optuna R^2	0.862	0.8617	0.9997	0.9065	0.9191	0.9162

Actual versus Predicted plots provided a visual assessment of model performance, where predictions aligning closely with a 45° line suggest accurate model fit.

The evaluation revealed distinct performance trends across the tested algorithms. For CH_4 yield prediction, XGBoost exhibited the highest R^2 value of 0.9447 and the lowest RMSE of 7.2743. It's Actual versus Predicted plot (Fig. 11) displayed minimal deviation from the ideal line, and the corresponding Q-Q plot (Fig. S14f) indicated residuals consistent with a normal distribution. LightGBM and CatBoost also performed well, though with marginally higher error metrics. In contrast, Random Forest and Bagging Regressor showed larger prediction errors and deviations from normality in their residuals, as evidenced by their Q-Q plots (Fig. S14e and S14a). For CH_4 selectivity, XGBoost again demonstrated the strongest performance, achieving an R^2 of 0.9374 and an RMSE of 6.2858. The residuals exhibited normality in the Q-Q plot (Fig. S13f), and the Actual versus Predicted plot (Fig. 10) confirmed close alignment between predicted and observed values. CatBoost performed competitively but with slightly lower metrics, while Random Forest and Bagging Regressor showed more pronounced deviations in their residual distributions. For CO_2 conversion, CatBoost achieved the highest predictive accuracy, with an R^2 of 0.9462 and an RMSE of 7.1363. Its residuals adhered closely to a normal distribution (Fig. S15b), and the Actual versus Predicted plot (Fig. 12) indicated strong agreement between predicted and actual values. XGBoost performed comparably, while the Gradient Boosting Decision Tree, despite achieving perfect cross-validation scores, exhibited residual skewness in its Q-Q plot (Fig. S15c), suggesting potential overfitting. Comprehensive assessment across these metrics supports the selection of XGBoost for CH_4 yield and CH_4 selectivity predictions and CatBoost for CO_2 conversion predictions. These models consistently outperformed others in terms of predictive accuracy, as indicated by R^2 and RMSE, and demonstrated robust statistical properties through residual normality and alignment in Actual versus Predicted plots. While other models showed varying levels of performance, XGBoost and CatBoost provided the most reliable and accurate predictions, with residual analyses confirming adherence to key statistical assumptions. These characteristics make them well-suited for both predictive and inferential applications in this context.

3.4.3. Analysis of feature contributions of best ML models: XGBoost and CatBoost

The linear perspective provided by Pearson correlation is complemented through the further employment of SHapley Additive exPlanations (SHAP) and Partial Dependence Plots (PDP). The limitations inherent in any single technique are mitigated by this multi-method approach. Feature importance is quantified by SHAP, whereby nonlinear and interactive effects are captured. Concurrently, the functional relationship between features and the predicted outcome is visually illustrated by PDPs. Through this synthesis, the robustness and interpretability of the findings are enhanced. To elucidate the

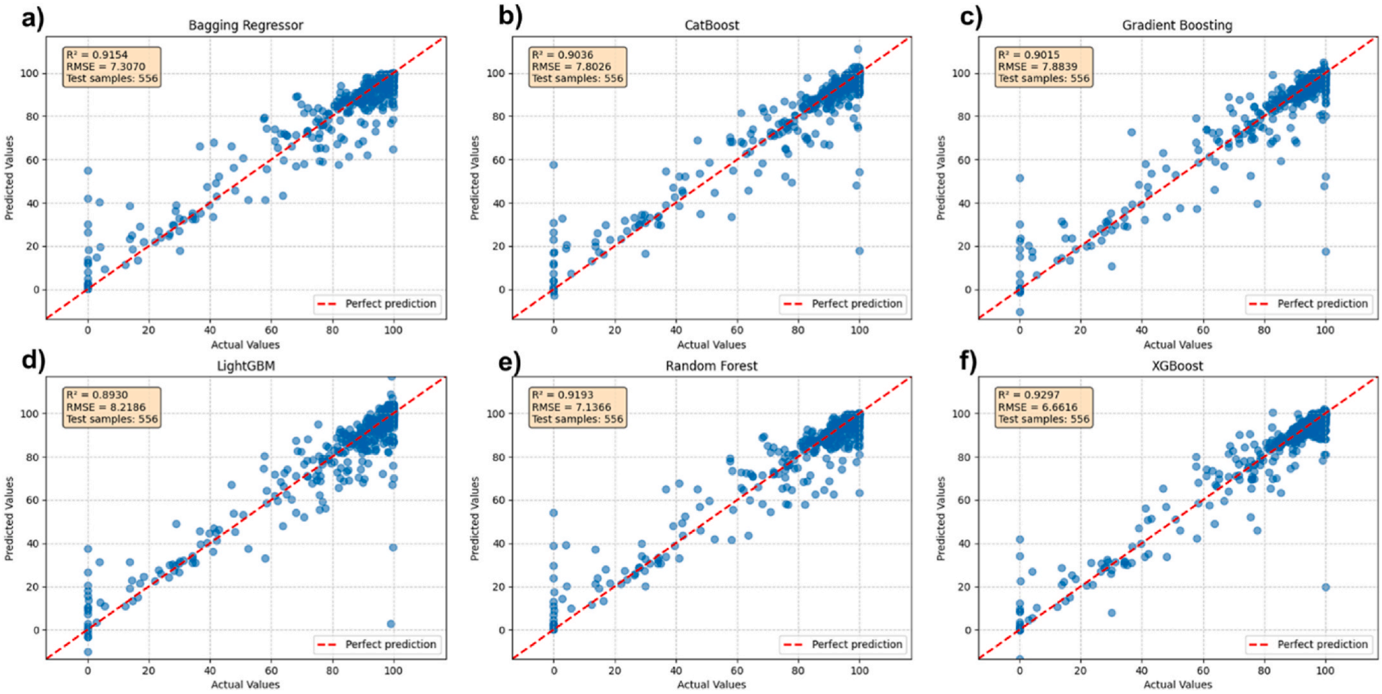


Fig. 10. Actual vs. predicted CH_4 selectivity values for six models: (a) Bagging Regressor (BR), (b) CatBoost (CB), (c) Gradient Boosting Decision Tree (GBDT), (d) LightGBM, (e) Random Forest (RF), and (f) XGBoost (XGB), with the 1:1 line indicating ideal agreement.

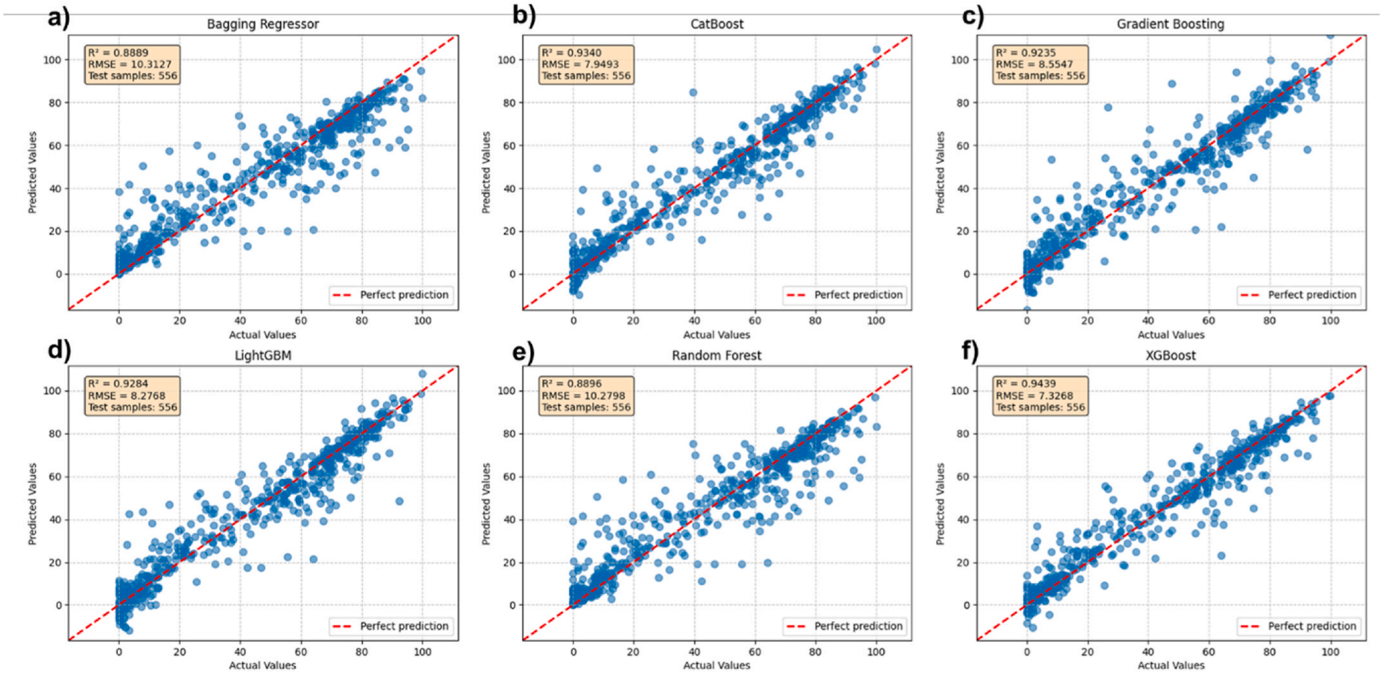


Fig. 11. Actual vs. predicted CH_4 yield values for six models: (a) BR, (b) CB, (c) GBDT, (d) LGBM, (e) RF, and (f) XGB models, with the 1:1 line indicating ideal agreement.

interpretability of the top-performing machine learning models—XGBoost for CH_4 yield (CH_4-Y) and CH_4 selectivity (CH_4-S), and CatBoost for CO_2 conversion (CO_2-CR)—we conducted a comprehensive analysis using SHapley Additive exPlanations (SHAP) and Partial Dependence Plots (PDPs). These interpretative tools provide critical insights into the physicochemical factors driving catalytic performance and reveal key structure-property relationships in CO_2 methanation. This section details the methodologies of SHAP and PDP, presents the results of their application, and discusses their implications for understanding catalyst behavior.

SHAP analysis, grounded in cooperative game theory, quantifies the marginal contribution of each input feature to a model's predictions by calculating Shapley values. These values represent the average contribution of a feature across all possible coalitions of features, ensuring a fair allocation of predictive importance. For a given prediction, the SHAP value for a feature is computed as the difference in the model's

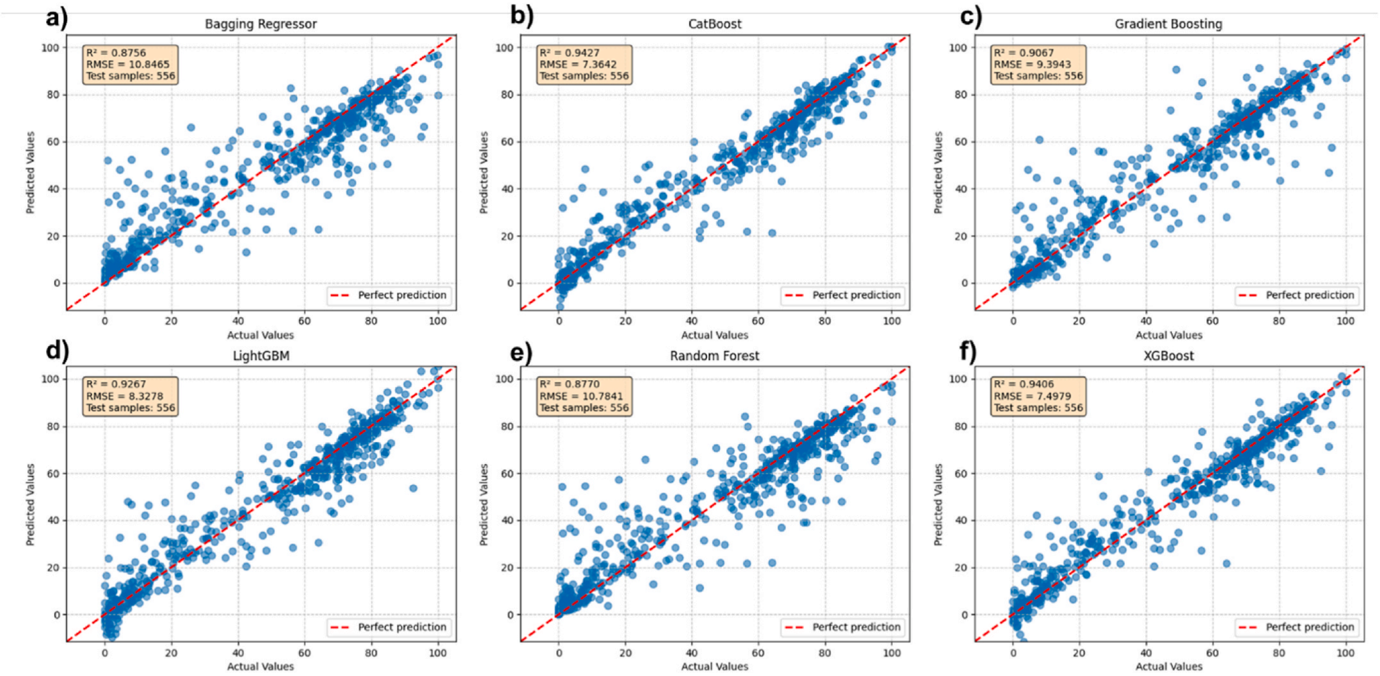


Fig. 12. Actual vs. predicted CO_2 convertion values for six models: (a) BR, (b) CB, (c) GBDT, (d) LGBM, (e) RF, and (f) XGB models, with the 1:1 line indicating ideal agreement.

output when the feature is included versus excluded, averaged over all permutations of feature inclusion. This approach provides a robust measure of feature importance and directionality, revealing how each feature influences the target variable (e.g., CH_4-Y, CH_4-S, CO_2-CR) in terms of both magnitude and sign (positive or negative impact). The results are visualized in SHAP summary plots (Fig. 13, Fig. S16), where features are ranked by their mean absolute SHAP values, and individual data points show the feature's impact on specific predictions. For CH_4 yield (XGBoost, Fig. S16b), the active component weight (ACW) emerged as the dominant feature, followed by support density (SAD), promoter density (PD), support atomic weight (SAW), and weight hourly space velocity (WHSV). The high importance of ACW suggests that the mass fraction of the active component is a primary determinant of methane production efficiency, likely due to its role in providing catalytic sites. SAD and PD indicate that textural properties, such as the spatial distribution of support and promoter materials, significantly influence reaction kinetics by affecting site accessibility and dispersion. SAW and WHSV further modulate yield by governing catalyst stability and reactant-catalyst contact time, respectively. For CH_4 selectivity (XGBoost, Fig. 13b, Fig. S16c), ACW again ranked highest, followed by calcination temperature (CTC), active component bulk modulus-weighted (CIBMW), active component formation energy (ACFE), and active component bulk modulus (IBIC). The prominence of CTC and ACFE highlights the critical role of electronic and thermodynamic properties in directing reaction pathways toward CH_4 formation. CTC likely influences surface electronic structure, facilitating charge transfer during hydrogenation, while ACFE governs the thermodynamic favorability of active site formation. CIBMW and IBIC further underscore the

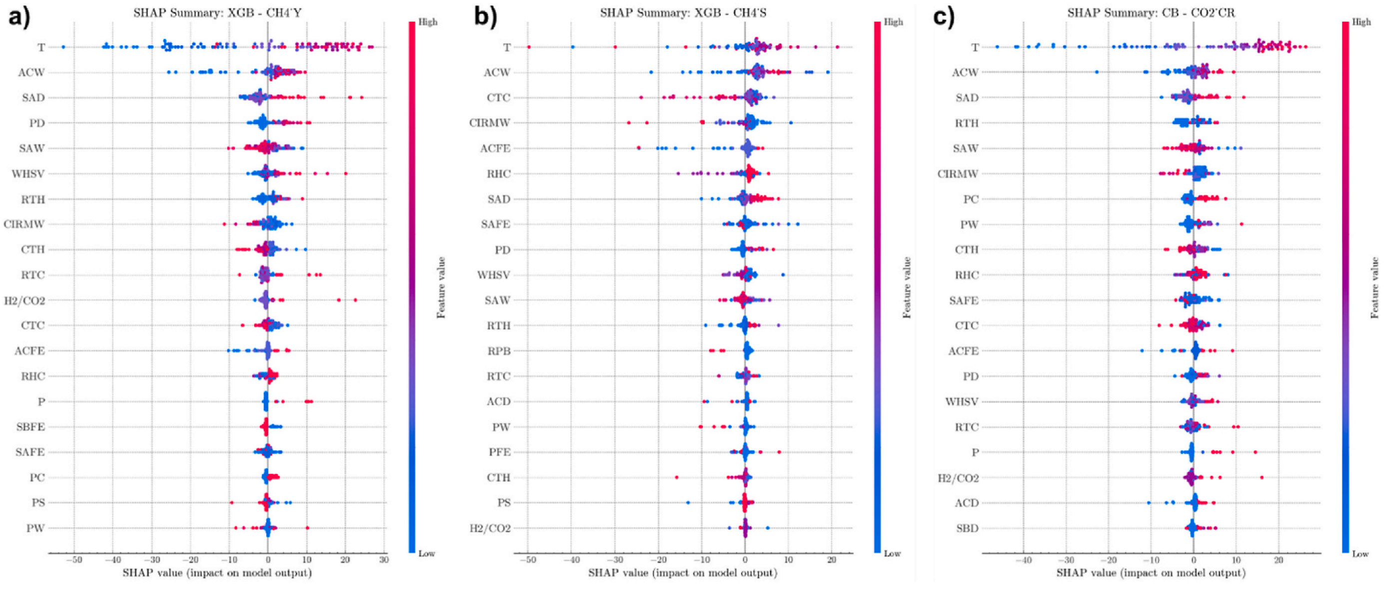


Fig. 13. Shap analysis of model feature importance. (a) XGBoost for CH_4 yield, (b) XGBoost for CH_4 selectivity, and (c) CatBoost for CO_2 conversion.

importance of mechanical and structural properties of the active component in stabilizing selective reaction intermediates. For CO_2 conversion (CatBoost, Fig. 13c, Fig. S16a), ACW, SAD, reduction time (RTH), SAW, and CIBMW were identified as the most influential features. The consistent importance of ACW across all models underscores its universal role in methanation catalysis, likely due to its direct impact on the availability of active sites for CO_2 activation. SAD and RTH suggest that support structure and reduction conditions critically affect CO_2 dissociation kinetics, while SAW and CIBMW indicate that support composition and mechanical properties of the active component further modulate conversion efficiency.

Partial Dependence Plots (PDPs) complement SHAP by illustrating the marginal effect of selected features on the predicted outcome, averaged over the distribution of all other features. For a given feature (or pair of features), the PDP is calculated by fixing the feature(s) at specific values, computing the model's predictions across the dataset, and averaging the results. This process reveals the functional relationship between the feature(s) and the target variable, independent of other inputs. PDPs are particularly useful for identifying linear, nonlinear, or interaction effects, providing mechanistic insights into how feature variations influence catalytic performance. In this study, PDPs were generated for key feature pairs (Figs. 14–16), focusing on formation energies versus densities for promoter, support, and active components. For CH_4 selectivity (XGBoost, Fig. 14), the PDP of promoter formation energy (PFE) versus promoter density (PD) (Fig. 14a) revealed a negative linear dependence, indicating that higher PD reduces energy barriers for intermediate hydrogenation, likely by enhancing promoter dispersion and increasing active site availability. The PDP of support formation energy (SAFE) versus support density (SAD) (Fig. 14b) exhibited an inverse correlation, suggesting that denser supports stabilize metastable active sites, thereby enhancing selectivity. The PDP of active component formation energy (ACFE) versus active component density (ACD) (Fig. 14c) displayed a nonlinear relationship, with optimal CH_4 selectivity achieved at moderate ACD. This suggests a balance between site accessibility and stability, where excessive density may lead to site overcrowding, reducing selectivity. For CH_4 yield (XGBoost, Fig. 15), analogous PDPs (Fig. 15a–c) confirmed that PFE, SAFE, and ACFE interact with their respective de nsities features (PD, SAD, ACD) to tune methane production. Notably, minimum in the PDPs were observed at specific thermodynamic-density equilibria, indicating that optimal catalyst configurations balance formation energies with structural properties to maximize yield. For example, moderate PD and SAD values were associated with higher yields, likely due to optimized dispersion and support stability. For CO_2 conversion (CatBoost, Fig. 16), PDPs (Fig. 16a–c) showed that PFE and SAFE exhibited stronger sensitivity to PD and SAD variations compared to ACFE versus ACD. This suggests that structural modifiers, such as promoters and supports, play a more significant role in enhancing CO_2 dissociation kinetics than the active component's density. Higher PD and SAD values were associated with improved conversion, likely due to enhanced promoter dispersion and support stability facilitating CO_2 activation.

Analysis of the Partial Dependence Plots (PDPs) in Figs. 14–16 reveals specific interaction trends among catalyst parameters in CO_2 methanation. For CH_4 selectivity under XGBoost (Fig. 14), the PDP for promoter formation energy (PFE, ranging from -0.5 to 0.3 eV/atom) versus promoter density (PD, 0 to 1.75 gr/cm^3) exhibits a peak partial dependence of approximately 85 % at PFE values near -0.4 eV and PD around 1.5 atoms/nm (Sedighi et al., 2024), declining to 80 % at higher PFE (0.2 eV/atom) and lower PD (0.5 gr/cm^3), implying that moderately high PD coupled with more negative PFE enhances selectivity (Ullah et al., 2025), yet this may overstate causality given potential multicollinearity in machine learning features. Similarly, the support formation energy (SAFE, -0.5 to 0.2 eV/atom) versus support density (SAD, 0 to 2.0 gr/cm^3) plot shows maximal dependence ($\sim$84 %) at SAFE near -0.4 eV and SAD at 1.0 gr/cm^3, with a drop to 82 % at positive SAFE, suggesting intermediate SAD optimizes support stability but risks extrapolation errors beyond training data (Molinet-Chinaglia et al., 2024). The active component formation energy (ACFE, -0.1 to 0.3 eV/atom) versus active component density (ACD, 0 to 1.75 gr/cm^3) displays a sharp decline from 85 % at low ACFE (-0.1 eV/atom) and low ACD (0.5 gr/cm^3) to 82 % at higher values, indicating sparse ACD favors performance, though model assumptions could amplify noise (Li et al., 2022). These patterns persist in CH_4 yield PDPs (Fig. 15), with peaks at 42 % for similar low formation energies and moderate densities, and in CO_2 conversion under CatBoost (Fig. 16), reaching 50 % at analogous points, but skepticism is warranted as PDPs average marginal effects without isolating confounding variables. Thermodynamically, more negative formation energies correlate with greater catalyst phase stability, potentially lowering Gibbs free energy barriers for CO_2 adsorption and dissociation in the exothermic methanation pathway ($\Delta H < 0$ see Table 1), as stable promoters like Sr or Mg enhance basic sites for CO_2 activation without fully accounting for kinetic limitations at low temperatures (Yarbaş and Ayas, 2024; Cui et al., 2024). However, optimal densities may reflect balanced surface energetics rather than direct equilibrium shifts, and claims of causation demand experimental Gibbs energy calculations to validate against ML-derived trends, lest overfitting obscures endothermic intermediate steps.

4. Conclusions

This study presents an integrated approach that combines thermodynamic simulations with machine learning to evaluate the methanation of CO and CO_2 using industrial off-gases such as blast furnace gas (BFG) and basic oxygen furnace gas (BOFG). Thermodynamic analysis, based

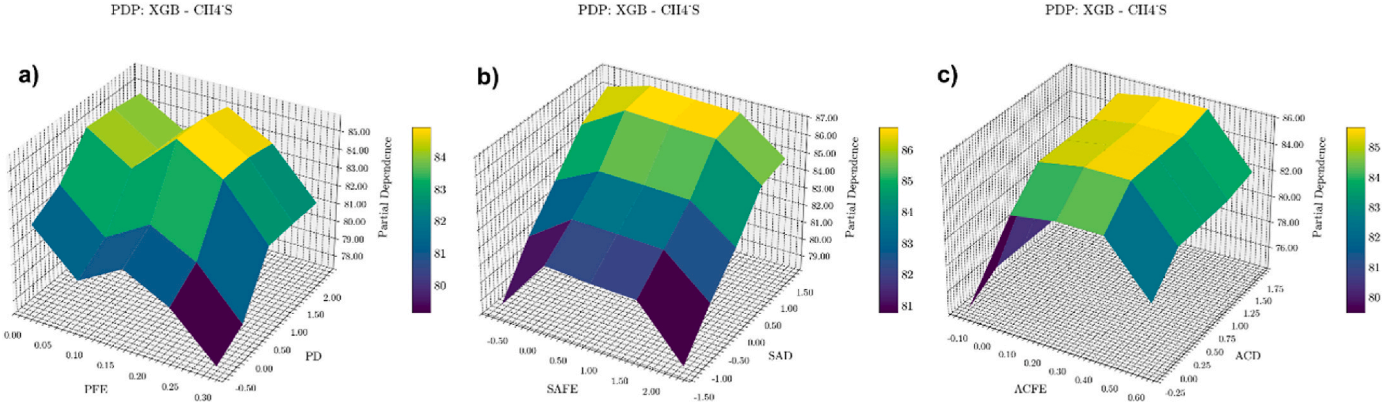


Fig. 14. Partial Dependence Plot (PDP) Analysis for CH_4 Selectivity. (a) XGBoost: Promoter Formation Energy vs. Promoter Density, (b) Support Formation Energy vs. Support Density, and (c) Active Component Formation Energy vs. Active Component Density.

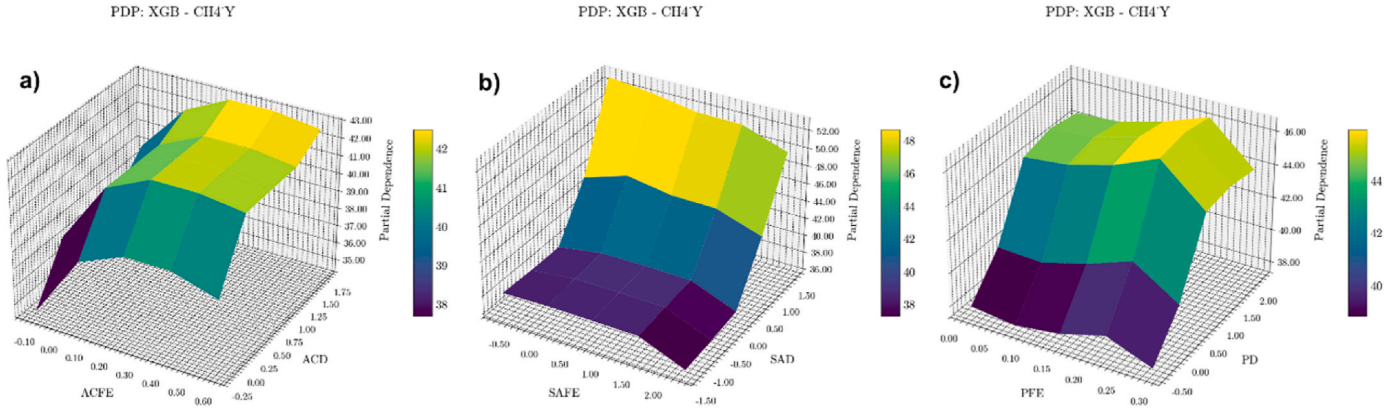


Fig. 15. Partial Dependence Plot (PDP) Analysis for CH_4 Yield. (a) XGBoost: Promoter Formation Energy vs. Promoter Density, (b) Support Formation Energy vs. Support Density, and (c) Active Component Formation Energy vs. Active Component Density.

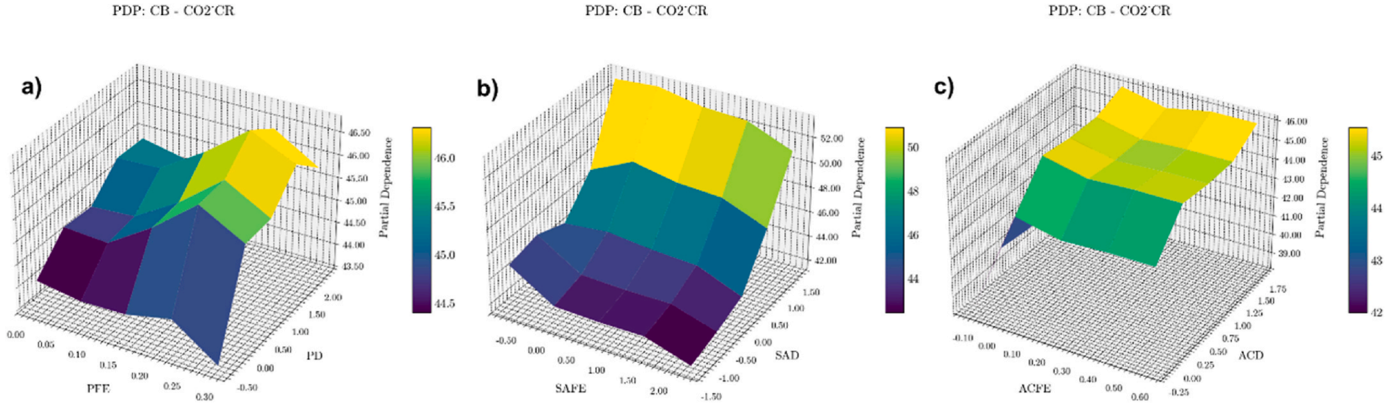


Fig. 16. Partial Dependence Plot (PDP) Analysis for CO_2 Conversion. (a) CatBoost: Promoter Formation Energy vs. Promoter Density, (b) Support Formation Energy vs. Support Density, and (c) Active Component Formation Energy vs. Active Component Density.

on Gibbs free energy minimization, identified the optimal operational conditions for maximizing CH_4 yield: moderate temperatures (150–250 °C) and elevated pressures (20–30 bar). Under these conditions, high CO_2 conversion efficiencies (>98 %) and minimal carbon deposition (<1 wt%) were achieved. Sobol sensitivity analysis highlighted hydrogen availability and off-gas composition as the most influential parameters, exceeding the impact of temperature and pressure, thus offering crucial insights for process optimization and scale-up. In parallel, machine learning models were developed using a curated dataset of 2777 experimental observations (Yang et al., 2024c), augmented with atomistic descriptors including formation energy and unit cell density to elucidate structure–property relationships of catalysts, including both active promoters and supports. Feature selection through Pearson correlation and recursive feature elimination with cross-validation (RFECV) identified key parameters such as active component loading, support density, and conditions reduction as dominant predictors of catalytic performance. Among the models tested, ensemble algorithms, particularly XGBoost and CatBoost, demonstrated superior predictive accuracy ($R^2 > 0.93$) across CH_4 yield, CO_2 conversion, and product selectivity. Model interpretability techniques, including SHAP (SHapley Additive exPlanations) analysis and partial dependence plots, further revealed that thermodynamic stability (as reflected in formation energy) and textural properties (e.g., support density) significantly influence catalytic activity and selectivity. For instance, optimized promoter loading and tailored reduction temperatures were found to enhance methane yield by balancing active site accessibility with structural stability. To guide future catalyst development, the incorporation of multi-objective optimization techniques such as NSGA-III is proposed to address inherent trade-offs among activity, selectivity, and stability, thereby identifying Pareto-optimal catalyst candidates. Additionally, the deployment of hybrid ensemble methods, such as stacked models, may improve generalizability and predictive performance. Ultimately, the integration of thermodynamic insights with machine learning-driven catalyst design and high-throughput screening represents a promising pathway toward the accelerated discovery of efficient and sustainable catalysts for industrial methanation of waste-derived feedstocks.

CRediT authorship contribution statement

Azita Etminan: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Peter J. Holliman:** Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Investigation, Funding acquisition, Data curation. **Peyman Karimi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Majid Majd:** Methodology, Investigation, Formal analysis, Data curation. **Ian Mabbett:** Supervision, Resources, Project administration, Funding acquisition, Software, Writing – review & editing. **Ciaran Martin:** Supervision, Resources, Project administration. **Anna RL. Carter:** Writing – review & editing. **Mary Larimi:** Writing– review & editing.

Funding sources

We gratefully thank EPSRC and Tata Steel for co-sponsoring an iCASE PhD studentship (Voucher number 220106) for AE and EPSRC for funding the Sustain Hub (EP/S018107/1) for PJH and the Centre for Digital Citizens - Next Stage Digital Economy Centre (EP/T022582/1) for ARLC.

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.

Appendix A. Thermodynamic Equations and Sobol Sensitivity Analysis

$$K_p = \exp\left(\frac{-\Delta_r G^\circ(T)}{RT}\right)s \tag{A. 1}$$

$$\Delta_r G^\circ(T) = \Delta_r H^\circ(T) - T\Delta_r S^\circ(T) \tag{A. 2}$$

$$\Delta_r H^\circ(T) = \Delta_r H_m(T) + \int_{298}^{T} \Delta_r C_{p,m} \quad dT \tag{A. 3}$$

$$\Delta_r S^\circ(T) = \Delta_r S_m(T) + \int_{298}^{T} \frac{\Delta_r C_{p,m}}{T} \quad dT \tag{A. 4}$$

$$CO_{2,\text{Conversion}}(\%) = \frac{\text{CO}_{2,\text{in}} - \text{CO}_{2,\text{out}}}{\text{CO}_{2,\text{in}}} \times 100 \tag{A. 5}$$

$$CO_{\text{ Conversion}}(\%) = \frac{\text{CO}_{,\text{in}} - \text{CO}_{,\text{out}}}{\text{CO}_{,\text{in}}} \times 100 \tag{A. 6}$$

$$CH_4\text{Yield}(\%) = \frac{\text{CH}_{4,\text{out}}}{\text{CO}_{2,\text{in}} + \text{CO}_{,\text{in}} + \text{H}_{2,\text{in}}} \times 100 \tag{A. 7}$$

$$C_{Balance} = \frac{CO_{2,ou_t} + CO_{,out}}{CO_{2,in} + CO_{,in}} \times 100 \tag{A. 8}$$

$$Energy_{,in} = = \sum_{i=1}^{n} n_{\text{Reactant}} \times H_{\text{ Reactant}} \tag{A. 9}$$

$$Energy_{,out} = \sum_{i=1}^{n} n_{\text{ Product}} \times H_{\text{ Product}} + Q_{\text{ Heat}} \tag{A. 10}$$

$$Total\ Energy\ Efeciency\ \left(\eta_{,\text{Total En}}\right) = \frac{En_{,\text{out}}}{En_{,\text{in}}} \times 100 \tag{A. 11}$$

$$\text{E}nergy_{,in} = \sum_{i=1}^{n} n_{\text{Reactant}} \times Ex_{\text{ Ch,Reactant}} + \sum_{i=1}^{n} n_{\text{Reactant}} \times Ex_{\text{ Ph,Reactant}} \tag{A. 12}$$

$$\text{E}nergy_{,out} = \sum_{i=1}^{n} n_{\text{Product}} \times Ex_{\text{ Ch,Product}} + \sum_{i=1}^{n} n_{\text{Product}} \times Ex_{\text{ Ph,Product}} + Ex_{\text{ Q}} \tag{A. 13}$$

$$Ex_Q = Q_{\ Heat}\left(1 - \frac{T_0}{T}\right) \tag{A. 14}$$

$$Total\ Exergy\ Efeciency\ \left(\eta_{,\text{Total Ex}}\right) = \frac{Ex_{,\text{out}}}{Ex_{,\text{in}}} \times 100 \tag{A. 15}$$

$$\eta_{\text{ En,CH}_4}(\%) = \frac{n_{CH_4} \times \text{LHV}_{CH_4}}{n_{H_2} \times \text{LHV}_{H_2}} \times 100 \tag{A. 16}$$

$$\eta_{\text{ Ex,CH}_4}(\%) = \frac{n_{CH_4} \times \text{Ex}_{CH_4}}{n_{H_2} \times \text{Ex}_{H_2}} \times 100 \tag{A. 17}$$

$$S_i = \frac{Var_{Xi}(E_{X-i}[Y|X_i])}{Var(Y)} \tag{A. 18}$$

$$S_i = 1 - \frac{Var_{X-i}(E_X[Y|X_{-i}])}{Var(Y)} \tag{A. 19}$$

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclepro.2025.146662.

Data availability

Data will be made available on request.

References

Adler, J., Parmryd, I., 2010. Quantifying colocalization by correlation: the pearson correlation coefficient is superior to the mander's overlap coefficient. Cytometry Pt A 77A (8), 733–742. https://doi.org/10.1002/cyto.a.20896.

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: a next-generation hyperparameter optimization framework. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*; ACM. Anchorage AK USA, pp. 2623–2631. https://doi.org/10.1145/3292500.3330701.

Alamgholiloo, H., Asgari, E., Sheikhmohammadi, A., Ghasemian, N., Hashemzadeh, B., Nourmoradi, H., 2024. Enhancement of the catalytic performance of Co-ZIF/WO3 heterostructures for selective catalytic reduction of NOx. Sci. Rep. 14 (1), 3277. https://doi.org/10.1038/s41598-024-53805-7.

Azzini, I., Mara, T.A., Rosati, R., 2021. Comparison of two sets of monte carlo estimators of sobol' indices. Environ. Model. Software 144, 105167. https://doi.org/10.1016/j.envsoft.2021.105167.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140. https://doi.org/10.1007/bf00058655.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/a:1010933404324.

Chen, B., Liao, Z., Wang, J., Yu, H., Yang, Y., 2012. Exergy analysis and CO2 emission evaluation for steam methane reforming. Int. J. Hydrogen Energy 37 (4), 3191–3200. https://doi.org/10.1016/j.ijhydene.2011.10.102.

Colelli, L., Bassano, C., Verdone, N., Segneri, V., Vilardi, G., 2024. Power-to-Gas: process analysis and control strategies for dynamic catalytic methanation system. Energy Convers. Manag. 305, 118257. https://doi.org/10.1016/j.enconman.2024.118257.

Coppitters, D., Costa, A., Chauvy, R., Dubois, L., De Paepe, W., Thomas, D., De Weireld, G., Contino, F., 2023. Energy, exergy, economic and environmental (4E) analysis of integrated direct air capture and CO 2 methanation under uncertainty. Fuel 344, 127969. https://doi.org/10.1016/j.fuel.2023.127969.

Cui, Y., He, S., Yang, J., Gao, R., Hu, K., Chen, X., Xu, L., Deng, C., Lin, C., Peng, S., Zhang, C., 2024. Research progress of non-noble metal catalysts for carbon dioxide methanation. Molecules 29 (2), 374. https://doi.org/10.3390/molecules29020374.

Davariashtiyani, A., Kadkhodaie, Z., Kadkhodaei, S., 2021. Predicting synthesizability of crystalline materials via deep learning. Commun Mater 2 (1). https://doi.org/10.1038/s43246-021-00219-x.

Díez-Ramírez, J., Sánchez, P., Kyriakou, V., Zafeiratos, S., Marnellos, G.E., Konsolakis, M., Dorado, F., 2017. Effect of support nature on the cobalt-catalyzed CO2 hydrogenation. J. CO2 Util. 21, 562–571. https://doi.org/10.1016/j.jcou.2017.08.019.

Etminan, A., Sadrnezhaad, S.K., 2022. A two step microwave-assisted coke resistant mesoporous Ni-Co catalyst for methane steam reforming. Fuel 317, 122411. https://doi.org/10.1016/j.fuel.2021.122411.

Ewald, S., Kolbeck, M., Kratky, T., Wolf, M., Hinrichsen, O., 2019. On the deactivation of Ni-Al catalysts in CO2 methanation. Appl. Catal. Gen. 570, 376–386. https://doi.org/10.1016/j.apcata.2018.10.033.

Gao, J., Wang, Y., Ping, Y., Hu, D., Xu, G., Gu, F., 2012. Others. A thermodynamic analysis of methanation reactions of carbon oxides for the production of synthetic natural gas. RSC Adv. 2 (6), 2358–2368.

Götz, M., Lefebvre, J., Mörs, F., McDaniel Koch, A., Graf, F., Bajohr, S., Reimert, R., Kolb, T., 2016. Renewable Power-to-Gas: a technological and economic review. Renew. Energy 85, 1371–1390. https://doi.org/10.1016/j.renene.2015.07.066.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning; Springer Series in Statistics. Springer, New York: New York, NY. https://doi.org/10.1007/978-0-387-84858-7.

Hidalgo, D., Martín-Marroquín, J.M., 2020. Power-to-Methane, coupling CO2 capture with fuel production: an overview. Renew. Sustain. Energy Rev. 132, 110057. https://doi.org/10.1016/j.rser.2020.110057.

Hjorth Larsen, A., Jørgen Mortensen, J., Blomqvist, J., Castelli, I.E., Christensen, R., Dułak, M., Friis, J., Groves, M.N., Hammer, B., Hargus, C., Hermes, E.D., Jennings, P. C., Bjerre Jensen, P., Kermode, J., Kitchin, J.R., Leonhard Kolsbjerg, E., Kubal, J., Kaasbjerg, K., Lysgaard, S., Bergmann Maronsson, J., Maxson, T., Olsen, T., Pastewka, L., Peterson, A., Rostgaard, C., Schiøtz, J., Schütt, O., Strange, M., Thygesen, K.S., Vegge, T., Vilhelmsen, L., Walter, M., Zeng, Z., Jacobsen, K.W., 2017. The atomic simulation Environment—A python library for working with atoms. J. Phys. Condens. Matter 29 (27), 273002. https://doi.org/10.1088/1361-648x/aa680e.

Hussain, I., Jalil, A.A., Izan, S.M., Azami, M.S., Kidam, K., Ainirazali, N., Ripin, A., 2021a. Thermodynamic and experimental explorations of CO2 methanation over highly active metal-free fibrous silica-beta zeolite (FS@SiO2-BEA) of innovative morphology. Chem. Eng. Sci. 229, 116015. https://doi.org/10.1016/j.ces.2020.116015.

Hussain, I., Jalil, A.A., Izan, S.M., Azami, M.S., Kidam, K., Ainirazali, N., Ripin, A., 2021b. Thermodynamic and experimental explorations of CO2 methanation over highly active metal-free fibrous silica-beta zeolite (FS@SiO2-BEA) of innovative morphology. Chem. Eng. Sci. 229, 116015. https://doi.org/10.1016/j.ces.2020.116015.

Huynh, H.L., Tucho, W.M., Shen, Q., Yu, Z., 2022. Bed packing configuration and hot-spot utilization for low-temperature CO2 methanation on monolithic reactor. Chem. Eng. J. 428, 131106. https://doi.org/10.1016/j.cej.2021.131106.

Jain, A., Ong, S.P., Hautier, G., Chen, W., Richards, W.D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., Persson, K.A., 2013. Commentary: the materials project: a materials genome approach to accelerating materials innovation. APL Mater. 1 (1). https://doi.org/10.1063/1.4812323.

Jia, C., Gao, J., Dai, Y., Zhang, J., Yang, Y., 2016. The thermodynamics analysis and experimental validation for complicated systems in CO 2 hydrogenation process. J. Energy Chem. 25 (6), 1027–1037. https://doi.org/10.1016/j.jechem.2016.10.003.

Jiang, M., Lian, H., 2024. Enhancing low-temperature CO2 methanation performance by selectively covering Ni with CeO2 to form Ni-O-Ce interface active sites. J. Catal. 438, 115716. https://doi.org/10.1016/j.jcat.2024.115716.

Li, L., Zeng, W., Song, M., Wu, X., Li, G., Hu, C., 2022. Research progress and reaction mechanism of CO2 methanation over Ni-Based catalysts at low temperature: a review. Catalysts 12 (2), 244. https://doi.org/10.3390/catal12020244.

Ling, Q., Li, X., Pei, Q., Lei, Z., Cui, P., Xie, R.L., 2024. Analyzing the methanation thermodynamic feasibility of steel plant byproduct gases. Sci. Rep. 14 (1). https://doi.org/10.1038/s41598-024-62982-4.

Ma, S., Liu, Z.-P., 2020. Machine learning for atomic simulation and activity prediction in heterogeneous catalysis: current status and future. ACS Catal. 10 (22), 13213–13226. https://doi.org/10.1021/acscatal.0c03472.

Martínez-Rodríguez, A., Abánades, A., 2020. Comparative analysis of energy and exergy performance of hydrogen production methods. Entropy 22 (11), 1286. https://doi.org/10.3390/e22111286.

Mebrahtu, C., Krebs, F., Abate, S., Perathoner, S., Centi, G., Palkovits, R., 2019. CO2 methanation: principles and challenges. In: Studies in Surface Science and Catalysis. Elsevier, pp. 85–103. https://doi.org/10.1016/B978-0-444-64127-4.00005-7, 178.

Mendoza-Hernandez, O.S., Shima, A., Matsumoto, H., Inoue, M., Abe, T., Matsuzaki, Y., Sone, Y., 2019a. Exergy valorization of a water electrolyzer and CO2 hydrogenation tandem system for hydrogen and methane production. Sci. Rep. 9 (1). https://doi.org/10.1038/s41598-019-42814-6.

Mendoza-Hernandez, O.S., Shima, A., Matsumoto, H., Inoue, M., Abe, T., Matsuzaki, Y., Sone, Y., 2019b. Exergy valorization of a water electrolyzer and CO2 hydrogenation tandem system for hydrogen and methane production. Sci. Rep. 9 (1). https://doi.org/10.1038/s41598-019-42814-6.

Molinet-Chinaglia, C., Shafiq, S., Serp, P., 2024. Low temperature sabatier CO_2 methanation. ChemCatChem 16 (24), e202401213. https://doi.org/10.1002/cctc.202401213.

Nemmour, A., Inayat, A., Janajreh, I., Ghenai, C., 2023. Green hydrogen-based E-Fuels (E-Methane, E-Methanol, E-Ammonia) to support clean energy transition: a literature review. Int. J. Hydrogen Energy 48 (75), 29011–29033. https://doi.org/10.1016/j.ijhydene.2023.03.240.

Ong, S.P., Richards, W.D., Jain, A., Hautier, G., Kocher, M., Cholia, S., Gunter, D., Chevrier, V.L., Persson, K.A., Ceder, G., 2013. Python materials genomics (pymatgen): a robust, open-source python library for materials analysis. Comput. Mater. Sci. 68, 314–319. https://doi.org/10.1016/j.commatsci.2012.10.028.

Pham, C.Q., Bahari, M.B., Kumar, P.S., Ahmed, S.F., Xiao, L., Kumar, S., Qazaq, A.S., Siang, T.J., Tran, H.-T., Islam, A., Al-Gheethi, A., Vasseghian, Y., Vo, D.-V.N., 2022. Carbon dioxide methanation on heterogeneous catalysts: a review. Environ. Chem. Lett. 20 (6), 3613–3630. https://doi.org/10.1007/s10311-022-01483-0.

Rönsch, S., Schneider, J., Matthischke, S., Schlüter, M., Götz, M., Lefebvre, J., Prabhakaran, P., Bajohr, S., 2016. Review on methanation – from fundamentals to current projects. Fuel 166, 276–296. https://doi.org/10.1016/j.fuel.2015.10.111.

Sahebdelfar, S., Takht Ravanchi, M., 2015a. Carbon dioxide utilization for methane production: a thermodynamic analysis. J. Petrol. Sci. Eng. 134, 14–22. https://doi.org/10.1016/j.petrol.2015.07.015.

Sahebdelfar, S., Takht Ravanchi, M., 2015b. Carbon dioxide utilization for methane production: a thermodynamic analysis. J. Petrol. Sci. Eng. 134, 14–22. https://doi.org/10.1016/j.petrol.2015.07.015.

Sahebdelfar, S., Takht Ravanchi, M., 2015c. Carbon dioxide utilization for methane production: a thermodynamic analysis. J. Petrol. Sci. Eng. 134, 14–22. https://doi.org/10.1016/j.petrol.2015.07.015.

Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. Comput. Phys. Commun. 145 (2), 280–297. https://doi.org/10.1016/s0010-4655(02)00280-1.

Sauer, J., 2024. The future of computational catalysis. J. Catal. 433, 115482. https://doi.org/10.1016/j.jcat.2024.115482.

Schumacher, J., Meyer, D., Friedland, J., Güttel, R., 2023. Methanation of CO/CO2mixtures: evaluation of pellet size effect on methane formation rate and reactant selectivity. Chem. Eng. J. 463, 142451. https://doi.org/10.1016/j.cej.2023.142451.

Sedighi, M., Azarhoosh, M.J., Alamgholiloo, H., Pesyan, N.N., 2024. Engineering CALF-20/Graphene oxide nanocomposites for enhancing CO2/N2 capture performance.

Process Saf. Environ. Prot. 190, 1481–1493. https://doi.org/10.1016/j.psep.2024.08.005.

Sheikhmohammadi, A., Asgari, E., Alamgholiloo, H., Jalilzadeh, Z., Aghanaghad, M., Rahimlu, F., 2024. Unveiling the role of artificial intelligence in tetracycline antibiotics removal using UV/Sulfite/Phenol advanced reduction process. J. Environ. Manag. 370, 122397. https://doi.org/10.1016/j.jenvman.2024.122397.

Shen, L., Xu, J., Zhu, M., Han, Y.-F., 2020. Essential role of the support for nickel-based CO_2 methanation catalysts. ACS Catal. 10 (24), 14581–14591. https://doi.org/10.1021/acscatal.0c03471.

Shin, D., Mok, D.H., Back, S., Na, J., 2023. Explainable formation energy prediction for uncovering the relationship between the electronic structure and stability of the heterogeneous catalyst. In: Computer Aided Chemical Engineering. Elsevier, pp. 691–696. https://doi.org/10.1016/B978-0-443-15274-0.50110-4, 52.

Song, C., Liu, J., Wang, R., Tang, X., Wang, K., Gao, Z., Peng, M., Li, H., Yao, S., Yang, F., Lu, H., Liao, Z., Wen, X.-D., Ma, D., Li, X., Lin, L., 2024. Engineering MOx/Ni inverse catalysts for low-temperature CO2 activation with high methane yields. Nat. Chem. Eng. 1 (10), 638–649. https://doi.org/10.1038/s44286-024-00122-5.

Tran, K., Ulissi, Z.W., 2018. Active learning across intermetallics to guide discovery of electrocatalysts for CO2 reduction and H2 evolution. Nat. Catal. 1 (9), 696–703. https://doi.org/10.1038/s41929-018-0142-1.

Ulissi, Z.W., Tang, M.T., Xiao, J., Liu, X., Torelli, D.A., Karamad, M., Cummins, K., Hahn, C., Lewis, N.S., Jaramillo, T.F., Chan, K., Nørskov, J.K., 2017. Machine-learning methods enable exhaustive searches for active bimetallic facets and reveal active site motifs for CO_2 reduction. ACS Catal. 7 (10), 6600–6608. https://doi.org/10.1021/acscatal.7b01648.

Ullah, S., Yang, X., Khan, Z.U., Aziz, S., Haider, W., Ben, H., 2025. Relationship between catalyst structure and activity in CO_2 methanation of Ru-Based catalysts: recent progress and future prospects. New J. Chem. 49 (17), 7097–7125. https://doi.org/10.1039/D5NJ01108F.

Vaitkus, A., Merkys, A., Sander, T., Quirós, M., Thiessen, P.A., Bolton, E.E., Gražulis, S., 2023. A workflow for deriving chemical entities from crystallographic data and its application to the crystallography open database. J. Cheminf. 15 (1). https://doi.org/10.1186/s13321-023-00780-2.

Wang, C., Wang, B., Wang, C., Li, A., Chang, Z., Wang, R., 2025. A machine learning model with minimize feature parameters for multi-type hydrogen evolution catalyst prediction. npj Comput. Mater. 11 (1), 111. https://doi.org/10.1038/s41524-025-01607-4.

Yang, L., Liu, S., Chang, C., Yang, S., Shen, W., 2024a. An efficient and invertible machine learning-driven multi-objective optimization architecture for light olefins separation system. Chem. Eng. Sci. 285, 119553. https://doi.org/10.1016/j.ces.2023.119553.

Yang, Q., Bao, R., Rong, D., Xiao, J., Zhou, J., Zhao, L., Zhang, D., 2024b. Interpretable machine learning for accelerating reverse design and optimizing CO_2 methanation catalysts with high activity at low temperatures. Ind. Eng. Chem. Res. 63 (33), 14727–14747. https://doi.org/10.1021/acs.iecr.4c01708.

Yang, Q., Bao, R., Rong, D., Xiao, J., Zhou, J., Zhao, L., Zhang, D., 2024c. Interpretable machine learning for accelerating reverse design and optimizing CO_2 methanation catalysts with high activity at low temperatures. Ind. Eng. Chem. Res. 63 (33), 14727–14747. https://doi.org/10.1021/acs.iecr.4c01708.

Yang, Q., Bao, R., Wang, Z., Guo, Q., Pan, Y., Zhou, X., 2025. Unlocking prediction and optimal design of CO2 methanation catalysts via active learning-enhanced interpretable ensemble learning. Chem. Eng. J. 509, 161154. https://doi.org/10.1016/j.cej.2025.161154.

Yao, L., Zhang, Z., Li, Y., Zhuo, J., Chen, Z., Lin, Z., Liu, H., Yao, Z., 2024. Precise prediction of CO2 separation performance of metal–organic framework mixed matrix membranes based on feature selection and machine learning. Separ. Purif. Technol. 349, 127894. https://doi.org/10.1016/j.seppur.2024.127894.

Yarbaş, T., Ayas, N., 2024. A detailed thermodynamic analysis of CO2 hydrogenation to produce methane at low pressure. Int. J. Hydrogen Energy 49, 1134–1144. https://doi.org/10.1016/j.ijhydene.2023.06.223.

Yılmaz, B., Oral, B., Yıldırım, R., 2023. Machine learning analysis of catalytic CO2 methanation. Int. J. Hydrogen Energy 48 (64), 24904–24914. https://doi.org/10.1016/j.ijhydene.2022.12.197.

Zhang, Y., Kitchaev, D.A., Yang, J., Chen, T., Dacek, S.T., Sarmiento-Pérez, R.A., Marques, M.A.L., Peng, H., Ceder, G., Perdew, J.P., Sun, J., 2018. Efficient first-principles prediction of solid stability: towards chemical accuracy. npj Comput. Mater. 4 (1). https://doi.org/10.1038/s41524-018-0065-z.

Zhang, M., Ye, J., Qu, Y., Lu, X., Luo, K., Dong, J., Lu, N., Niu, Q., Zhang, P., Dai, S., 2024. Highly stable and selective Ni/ZrO_2 nanofiber catalysts for efficient CO_2 methanation. ACS Appl. Mater. Interfaces 16 (27), 34936–34946. https://doi.org/10.1021/acsami.4c04124.

Zhang, L., Li, Z., Gao, Y., Li, Q., Deng, Z., Miao, B., He, H., Poh, C.K., Zhang, L., Chan, S. H., 2025. Recent advances in supported metal catalysts for CO2 methanation: mechanisms, materials design, and the promise of perovskite-based supports. Energy Convers. Manag. X 27, 101066. https://doi.org/10.1016/j.ecmx.2025.101066.