16 (2025) 100471

Contents lists available at ScienceDirect

Carbon Capture Science & Technology

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

Perspective

Perspective on artificial intelligence for carbon capture utilization and storage (CCUS) in Petrochemical Industry

Jin Ma [a,1], Yide Han [b,1], Meihong Wang [a,b,*], Weimin Zhong [a], Wenli Du [a], Feng Qian [a,*]

[a] Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, PR China

[b] School of Chemical, Materials and Biological Engineering, The University of Sheffield, Sheffield S1 3JD, UK

ARTICLE INFO

Keywords
Carbon capture, utilisation and storage (CCUS)
Petrochemical industry
Artificial Intelligence
Solvent selection and design
Catalyst Design
Sustainability
Carbon Capture
CO_2 Utilisation

ABSTRACT

The energy-intensive petrochemical industry contributes 14 % of global industrial emissions. In the face of climate change, there is an urgent need for the petrochemical industry transition to low carbon manufacturing. Deployment of carbon capture, utilization and storage (CCUS) technologies can effectively reduce carbon emissions from the petrochemical industry. However, the large-scale deployment of CCUS faces the obstacles of high energy consumption and high cost. Artificial intelligence (AI) has shown great potential to accelerate the large-scale deployment of CCUS in the petrochemical industry. Nevertheless, most AI-based approaches are still largely at the research stage and not yet widely adopted in industrial practice. This paper explores four aspects of AI for petrochemical industry to reduce CO_2 emission, including the solvent selection and design for carbon capture, catalyst design for CO_2 utilisation, hybrid process modelling for optimal design and operation, and life cycle sustainability assessment. We evaluate different promising approaches for AI in each aspect and highlight our key findings, with the goal to accelerate the petrochemical industry transition to carbon neutrality.

1. Introduction

Worldwide, the petrochemical industry contributes to 7 % of the global gross domestic product (GDP) and manufactures approximately 1 billion tonnes of essential commodities such as plastics and synthetic fibres. The petrochemical industry converts feedstocks (e.g. naphtha, liquefied natural gas) to chemical building blocks (e.g. ethylene, ammonia), and then to final products (e.g. plastics, fertilisers) (Cullen et al., 2024). One important product of the petrochemical industry is plastics, which is widely used in packaging, construction and textiles. The annual consumption of plastics is increasing and is considered to be strongly correlated with GDP growth (IEA, 2018). However, significant amount of CO_2 emission is released with the development of petrochemical industry. According to the International Energy Agency (IEA) report, the petrochemical industry worldwide consumes 14 % of global crude oil and 9 % of nature gas. It also generates 1.3 Gt-CO_2/year (roughly 14 % of the global CO_2 emission from industry), making it one of the culprits for the global warming (IEA, 2020). About 60 % of these are energy related CO_2 emission, primarily resulting from fuel combustion for separation and conversion processes (Yan et al., 2024). Under the net-zero scenario, the petrochemical industry is now faced with the major task of reducing carbon emissions and an urgent transition to sustainable petrochemical production.

Currently, *alternative feedstocks* and *recycling* are deployed to reduce CO_2 emissions from the petrochemical industry. However, these measures address only the symptoms rather than the root causes of the problem. For instance, only 9 % of the waste plastic generated globally is successfully recycled (OECD, 2022). Looking ahead, carbon capture, utilisation and storage (CCUS) will play a key role in reducing CO_2 emission from petrochemical industry due to its direct mitigation ability and carbon utilisation potential for producing additional value-added chemicals or fuels. Examples include green ethylene production combining carbon capture with wind power (Nyhus et al., 2024) and blue hydrogen production using natural gas reforming with CCUS (William et al., 2023). The IEA predicts that CCUS will be instrumental in reducing carbon emissions from the petrochemical industry by over 30 % by 2050, serving as the backbone of the petrochemical sector towards net-zero emissions (IEA, 2020). Nevertheless, with the consideration of two factors: (a) the maturation of a chemical process typically takes thirty to fifty years from lab-scale studies to commercial

* Corresponding authors.
E-mail addresses: meihong.wang@sheffield.ac.uk (M. Wang), fqian@ecust.edu.cn (F. Qian).
[1] These two authors are co-first authors.

https://doi.org/10.1016/j.ccst.2025.100471
Received 7 May 2025; Received in revised form 23 July 2025; Accepted 24 July 2025
Available online 25 July 2025

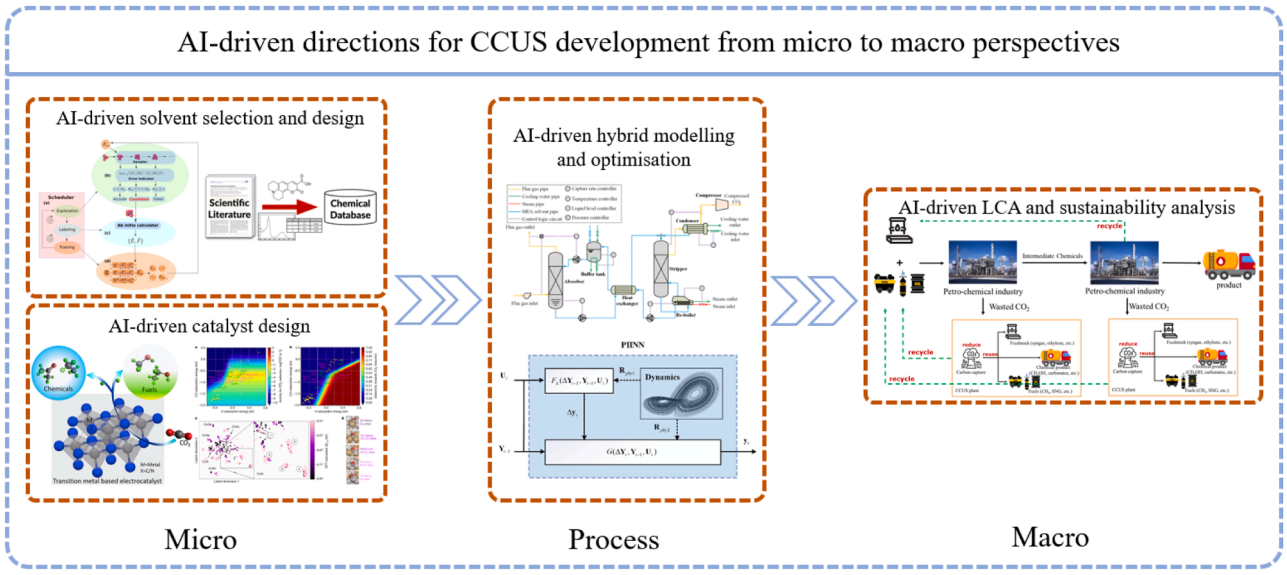

Fig. 1. Four different AI-driven directions for CCUS development in the petrochemical industry from micro to macro perspectives.

deployment; and (b) CCUS faces challenges of high energy consumption and high cost. It is difficult to widely deploy the commercially available CCUS technology for the petrochemical industry to combat the climate change issues in the short term (Gupta et al., 2024; Hong, 2022). For example, China plans to peak its carbon dioxide emissions by 2030, with less than 5 years from now, it is very challenging to develop a mature CCUS technology for the chemical industry using previous paradigm of chemical engineering research.

Artificial intelligence (AI) inspires new ways to accelerate industry process development towards the 4th industrial revolution (Chiang et al., 2022). "AI for science and engineering" has become a time-saving and cost-effective approach to accelerate industrial process development. The AI-driven technologies are already transforming other industries such as AI-driven drug design (Arnold, 2023), natural language processing (NLP) in new materials discovery (Jiang et al., 2025), and machine learning (ML) in molecular design (Lemaoui et al., 2023). With the help of ML, problems in chemical engineering that were difficult to solve with traditional methods due to high computational demand become feasible, such as global optimization of catalytic reaction networks (Margraf et al., 2023), prediction of amine emissions from carbon capture process (Jablonka et al., 2023) and design of gas separation membranes (Li et al., 2025). The momentum is growing for exploring and employing AI in CCUS for the petrochemical industry (Hussin et al., 2023), which will accelerate the research in this field from different perspectives including solvent selection and design, catalyst design, hybrid process modelling, life cycle assessment (LCA) and sustainability.

We suggest four different AI-driven directions (refer to Fig. 1) for CCUS development in the petrochemical industry from micro to macro perspectives as state-of-the-art solutions to accelerate petrochemical industry transition towards sustainability.

2. Potential AI applications, current status and future challenges

2.1. Solvent selection and design for carbon capture

Amine scrubbing is regarded as the most mature carbon capture technology and has been applied for capturing CO_2 from natural gas processing (IEA, 2020) and power plants (Wang et al., 2015). However, solvent-based carbon capture is still very energy-intensive and costly (Borhani et al., 2024). Changing the solvent has great potential to improve the performance of carbon capture. For example, alternative solvents like piperazine (PZ) can reduce energy consumption from 3.7–4 GJ/tonne (using MEA) to less than 3 GJ/tonne for power plants (Otitoju et al., 2021) and for ethylene plants (Ma et al., 2024). This is due to the fact that different substituent groups on the amino nitrogen atom affect the thermodynamic properties of solvents as well as the kinetics of their absorption reactions. Therefore, selecting and designing high-performance solvents is significant in overcoming the main obstacles for the commercial deployment of carbon capture.

The selection and design of solvents in carbon capture requires a balanced consideration of absorption capacity, regeneration energy, solvent selectivity, thermal stability and viscosity. In order to obtain the thermodynamic properties of the solvents and the kinetics of their absorption reactions, experimental discoveries and molecular simulations are commonly used in studies for solvent selection and design. Candidate solvents include not only new solvents developed using quantitative structure-property relationship (QSPR) models, which predict solvent properties using molecular structure, but also mixed solvents composed of multiple solvents blended in different ratios. The large number of potential candidates presents a significant challenge for solvent selection and design in carbon capture (Struebing et al., 2013). In addition, the inherently time-consuming feature of experimental studies and molecular simulations limits the progress of solvent development, which delays the commercial deployment of optimal carbon capture for the petrochemical industry. In the face of this challenge, previous research has attempted to develop prediction models of molecular properties by computer-aided molecular design (CAMD) (Zhou et al., 2020). Integrating computer-aided molecular design with process simulation holds great potential for identifying high-performance solvents. For instance, a recent solvent screening study identified DEA as a promising candidate, achieving a 46.7 % reduction in overall energy consumption for CO_2 capture compared to the benchmark solvent MEA. However, traditional QSPR and CAMD approaches are limited by reliance on manually selected molecular descriptors, low capabiliity to capture complex non-linear structure-property relationships, and poor generalizability due to small datasets. With the emergence of the AI era, NLP-based approach offers a promising solution by data extraction from literature. In addition, emerging machine learning approach can learn directly from DFT without relying on predefined descriptors, thereby enhancing both the predictive accuracy and physical interpretability of molecular property models.

Artificial neural networks, genetic algorithms, Bayesian optimization, and random forests have been used in the past to find new

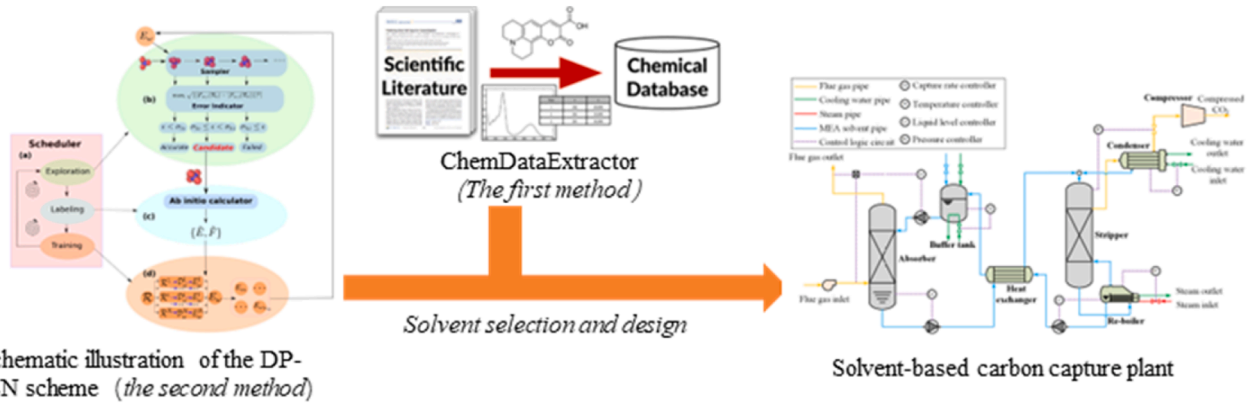

Fig. 2. AI application in solvent selection and design for carbon capture (Swain and Cole, 2016; Zhang et al., 2020; Sha et al., 2025).

structures for amine-based solvents (Hosseinpour et al., 2023). Recent advances in AI have generated two promising methods for accelerating access to experimental data and molecular simulations (refer to Fig. 2). The first method involves the integration of NLP to develop a molecular property database. For example, the open-source software *ChemDataExtractor* originated from the Cavendish Laboratory at the University of Cambridge (Swain and Cole, 2016) is widely used in the world, along with the commercial software *SciFinder* and the open-source software *ChemSpot*. *ChemDataExtractor* is created as an AI tool employing NLP to extract experimental data (i.e. thermodynamics properties and kinetic parameters) about specific chemicals from published literature. The second method is to use deep learning to predict potential energy surface to make molecular simulations faster. This method was initially introduced by Han et al. (2018) and was subsequently improved to develop *DeepMD-kit*—an open-source machine learning tool which trains deep neural networks using density-functional theory (DFT) data to perform molecular dynamics simulation (Wang et al., 2018). Building on this foundation, *DP-GEN* automates Deep Potential model development using an iterative framework, combining configuration space exploration based on large-scale atomic/molecular massively parallel simulator (LAMMPS), performs DFT calculation based on Vienna Ab initio simulation package (VASP), and *DeePMD-kit* training to optimize data representativeness and model accuracy (Zhang et al., 2020).

ChemdataExtractor has been used with large language model (LLM) to implement the collection of experimental metal-organic framework (MOF) data to generate a comprehensive and ready-to-use dataset. Specifically, *ChemdataExtractor* was used to extract texts from over 40,000 journal articles related to MOFs and processed into structured data to feed LLM. This step was crucial for maintaining high data quality and enabling accurate extraction of MOF synthesis conditions and chemical properties (Kang et al., 2025).

DP-GEN has been applied to develop a model to predict the properties of pure copper covering wide temperature (50–2715 K) and pressure ranges (1–50,000 bar) (Zhang et al., 2020). Compared with conventional DFT calculations, the trained Deep Potential model predicts the potential energy surface with <2 % relative error, and accelerates computational efficiency by about 3600 times. The model significantly reduces the computational time required for molecular simulations, thus improving efficiency.

Both methods are expected to be applied to the solvent selection and design for carbon capture. Using *ChemDataExtractor* tool (Swain and Cole, 2016), the properties of candidate solvents documented in the existing literature can be rapidly accessed. Meanwhile, using *DP-GEN* (Zhang et al., 2020) could accelerate the molecular simulation, and the prediction model for candidate solvents properties can reduce computational demand by thousands of times compared to DFT simulations in the foreseeable future. These two tools are open-source. Both *DeepMD-kit* (https://github.com/deepmodeling/deepmd-kit) and *ChemDataExtractor* (https://github.com/CambridgeMolecularEngineering/chemdataextractor2) provide example datasets and benchmark cases for model training and validation. Both methods are time-efficient and significantly accelerate the solvent selection and design process. Subsequent studies could link solvent properties with the carbon capture process to select and design solvents to obtain solvents with the lowest energy consumption for carbon capture process (Lee et al., 2023a). So far, AI-assisted solvent selection and design remains at the proof-of-concept stage, with no real-life application examples. Significant challenges remain, such as accurately modelling the relationship between molecular properties and process performance. Recent studies have begun to address this gap by developing predictive models for carbon capture processes that are applicable to a large number of amine solvents (Lee et al., 2023b). In addition, while NLP tools can provide large-scale data extraction, the inconsistency of various chemical reporting standards remains a major challenge. Manual intervention for post-processing is still required to ensure the reliability of the extracted datasets.

2.2. Catalyst design for CO_2 utilisation

The field of catalysis is undergoing a paradigm shift, ushering in an era where catalyst design is central. The essence of catalyst design is constructing active sites and giving the active site a specific environment to precisely control activity, selectivity and stability. Catalysis is now evolving towards unprecedented precision, as the scale of the catalyst active phase has transitioned from nanoscale and sub-nanoscale dimensions to the single atom dimension. The concept of single-atom catalysts (SACs) was first proposed by Qiao et al. (2011), who employed a co-precipitation method to produce platinum SACs for carbon monoxide oxidation. Currently, SACs have emerged as a frontier in catalysis due to their maximum atom utilization and unique structures. With the rapid progress in controllable synthesis strategies, SACs have been successfully applied across a wide range of fields, including toxic gas handling, fuel cell and CO_2 utilisation (Chen et al., 2018).

The petrochemical industry emits substantial amounts of CO_2. Converting the CO_2 into fuels and chemicals to establish an artificial carbon cycle is important for achieving decarbonization of the petrochemical industry. However, CO_2 utilization process requires the design of efficient and selective catalysts with low energy consumption and low cost, which has always been a great challenge (Leonzio and Shah, 2024). Due to the unique physical-chemical properties and abundant adjustable coordination environments, SACs often exhibit superior performance beyond conventional nanocatalysts in CO_2 utilization. For example, Co-based catalysts for Fischer-Tropsch synthesis (FTS) often suffer from methane overproduction and deactivation due to water-induced oxidation. By stabilizing hexagonal close-packed (HCP) phase of Co nanoparticles, Ir-Co catalyst (a kind of SAC) significantly suppresses methane

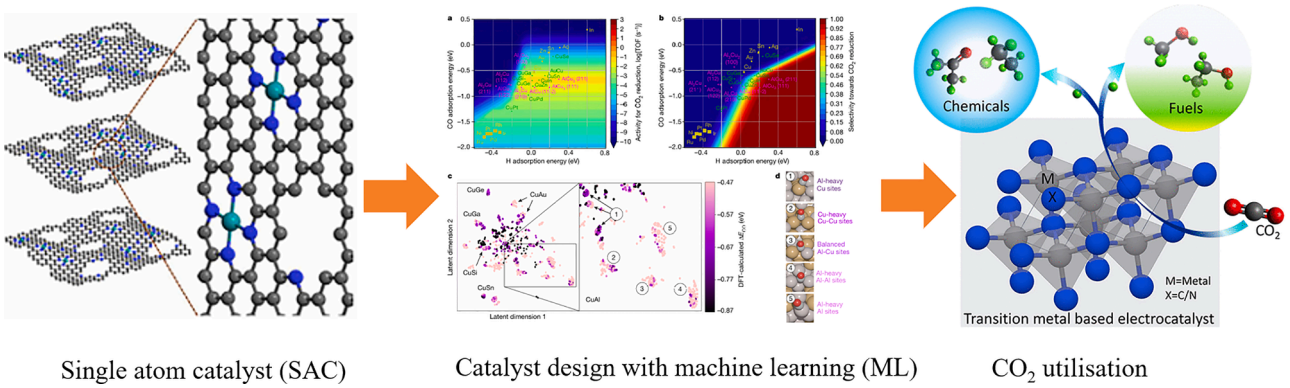

Fig. 3. AI application in catalyst design for CO_2 utilisation (Chen et al., 2018; Zhong et al., 2020).

selectivity—from 10 % with Co catalyst to 2.7 %—and simultaneously enhances durability, maintaining stability to over 1200 hours (better than 200 hours with Co catalyst) (Zhou et al., 2019). In the electrochemical CO_2 reduction reaction (CO_2RR), the high overpotential required to activate the stable $C = O$ bond, coupled with the presence of multiple competing reaction pathways, often results in low conversion efficiency and poor product selectivity. With their maximized atomic utilization efficiency and adjustable coordination environments, SACs exhibit enhanced catalytic activity. Furthermore, the uniform structure of SACs' active sites ensures consistent interactions with substrates, which is beneficial for improving selectivity. Efficient CO2RR processes can be realized using Ni-based SACs (nearly 100 % Faradaic efficiency for CO production) (Chen et al., 2018).

The rational design of SACs remains challenging. Identifying optimal metal-support combinations for catalyst active phase and coordination environments generates thousands of choices. Traditional trial-and-error approaches are limited by the vast combinatorial space of metal-support combinations and coordination configurations. ML models can predict metal coordination numbers, ligand types, and bond distances/angles by learning from extensive DFT databases so as to rapid identification of optimal geometries. In addition, AI accelerates the screening of thousands of potential active site configurations. By combining large DFT datasets with ML algorithms, structure-performance relationships of catalysts can be obtained, which is critical for catalyst design (Xin, 2022). Particularly in CO_2 utilization, the application of AI advances the discovery of catalyst materials (Mazheika et al., 2022). For example, Cu is the dominant electrocatalyst when it is desired to reduce carbon dioxide to ethylene. However, hundreds of copper-containing metal crystals have more than thousands of coordination environments. It is obviously not feasible in terms of time cost to perform DFT simulations for all the possible situations to find the most catalytically active ones. Using DFT simulation results as training data, researchers (Zhong et al., 2020) screened the copper-aluminium alloy using ML regression method and successfully obtained an electrocatalyst with the Faraday efficiency over 80 % compared to 66 % with pure Cu catalyst. In addition to its high selectivity, the AI-designed catalyst demonstrated a power conversion efficiency (PCE) of 55 ± 2 % at 150 mA/cm^2, significantly outperforming conventional Cu-based catalysts (around 35 %). Moreover, volcano plot analysis showed that the AI-designed catalyst approached the optimal CO adsorption energies (ΔE_{CO}) value (~–0.67 eV), indicating a high turnover frequency (TOF). In contrast, the pure copper catalyst exhibits a ΔE_{CO} that deviates significantly from the optimal value, resulting in lower intrinsic performance. Zhong et al. (2020) studied 244 different copper-containing intermetallic compounds, from which 12,229 surface configurations were systematically enumerated and 228,969 potential adsorption sites were identified. To establish a robust AI predictive model, Zhong et al. (2020) conducted systematic DFT simulations across 4000 carefully selected active sites, calculating CO adsorption energies to serve as high-quality training data. A similar approach has been used to design single-atom alloy (SAA, a kind of SAC) catalysts for hydrogen production from methane cracking. Conventional catalysts are prone to deactivation due to carbon deposition, Sun et al. (2024) used DFT-calculated C—H dissociation energy barriers for SAA catalyst surface structures as the training data. A ML model was used to predict and screen the optimal catalysts from over 10,000 SAA catalysts surface structures, and the final screened Ir/Ni surface achieved a methane conversion of 13.87 % at 450 °C and 1 atm, which is much higher than that of the pure Ni catalyst of 3.7 %.

CO_2 utilization involves various catalysis methods (e.g. thermal catalysis, electrocatalysis and photocatalysis) and generates different products such as ethylene, methane and syngas (Kamkeng et al., 2021). Designing catalysts with excellent performance has been a challenge for CO_2 utilization processes (Gao et al., 2017). Currently only the electrocatalytic reduction process for the production of CO and formic acid has the potential for levelized costs below current market prices (Leonzio et al., 2024), which is hardly enough to drive the large-scale deployment of CCUS. The role of SACs in CO_2 utilisation is promising (Yu et al., 2023), especially when atomic-scale insights of the SACs are combined with the predictive power of AI to accelerate catalyst design (Mok et al., 2023). As shown in Fig. 3, SCA can coupled with ML to guide the design of catalysts for CO_2 reduction in to chemical and fuels. However, current progress remains hindered by several critical limitations. Data constraints represent a primary bottleneck, as reliable AI models require training on datasets generated from thousands of DFT calculations, constituting an enormous computational demand. Furthermore, critical questions remain regarding whether training datasets encompass appropriate chemical spaces, data quality, and the inherent gap between DFT predictions and experimental data. In addition, the transferability of AI models across different catalytic systems remains unresolved, while the limited interpretability of AI-derived models poses additional challenges that strongly depend on how physical information is included.

2.3. *Hybrid modelling for process optimisation*

Process modelling and optimization are widely used in petrochemical industry. In terms of modelling, as processes become increasingly nonlinear and multi-scale, the computational demand of first-principles modelling \ simulation continues to increase (Rajulapati et al., 2022). At the same time, the rise of AI advances the development of data-driven modelling that can leverage large datasets to capture complex process behaviour. Hybrid modelling, which combines the advantages of first-principle modelling and data-driven modelling, ensures accurate and efficient model predictions in increasingly complex petrochemical processes, and provides the foundation for optimal design and operation (Sharma and Liu, 2022). In terms of optimisation, a variety of AI-based algorithms have been proposed to enhance the efficiency and effectiveness of process optimization tasks.

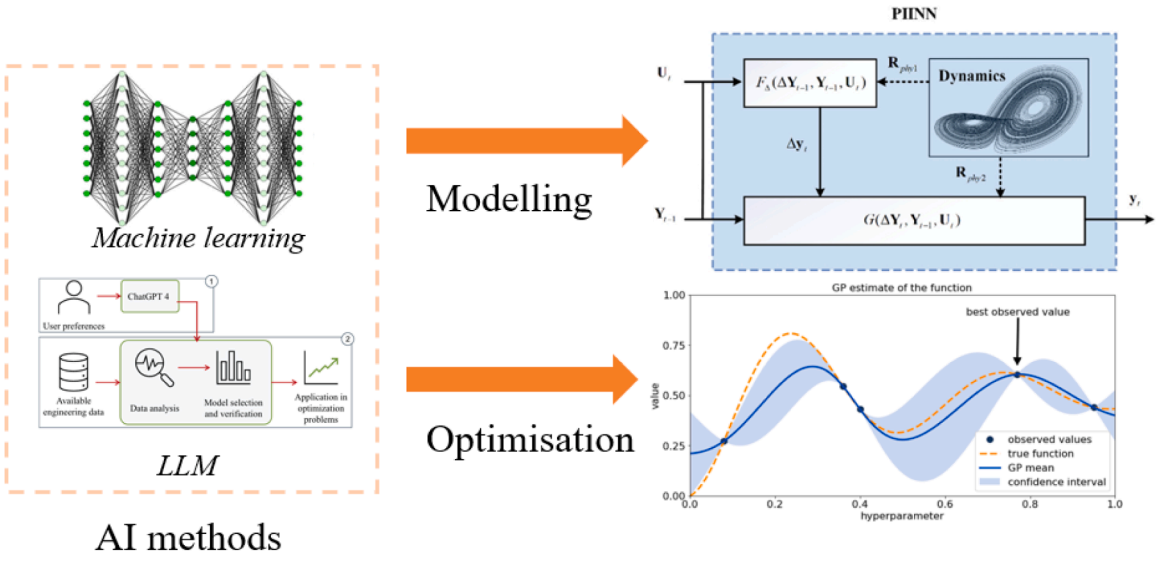

Fig. 4. Hybrid modelling for process optimisation (Liu et al., 2022; Rios et al., 2024; Sha et al., 2025; Chiang et al., 2022).

There are many cases of hybrid modelling for CCUS. For instance, Sha et al. (2025) employed a physically informed neural network (PINN) to combine experimental data with fundamental physical laws to capture the nonlinear dynamics of a carbon capture system. PINN models can effectively capture complex nonlinear process dynamics to cope with variations of the inputs such as lean solvent flowrate, re-boiler steam flowrate, and flue gas flowrate. In contrast, conventional reduced order models based on principal component analysis (PCA), which typically rely on linear projections, are difficult to use for carbon capture systems.

Moreover, design and operation optimization based on AI technology has been applied in the field of decarbonization of industrial processes with CCUS (Zhang et al., 2025). Researchers (Xi et al., 2021) designed a steel mill gas utilization system integrated with CCUS and renewable energy based on a hybrid modelling approach. The system uses particle swarm optimization (PSO) algorithm for multi-objective optimization. To deal with weather and load changes, the carbon capture level and the methanol production rate were flexibly operated, and finally achieved a 62 % reduction in overall system CO_2 emissions.

Recently, with the release of ChatGPT and DeepSeek, the performance of LLM has advanced and is promising for industrial process modelling and optimisation (Javaid et al., 2023). LLM can assist engineers in selecting appropriate hybrid modelling methods as well as optimization strategies. Researchers (Rios et al., 2024) demonstrated how to use LLM to analyse industrial processes cooperating with human users. They used realistic engineering data to interact with ChatGPT-4o to assist the surrogate model selection and process optimisation. LLM-assisted surrogate modelling and optimisation was successfully applied to *gear train* optimization and pressure vessel design, with performance differences in optimization results of <1 % compared to fully manually constructed surrogate models as well as optimisation (Rios et al., 2024).

The AI methods, ML and LLM, shown on the left side of Fig. 4, can be applied to both modelling and optimization. For modelling, we provided an example of carbon capture process modelling. For optimization, we use the example of optimal decarbonization of a steel mill gas utilization system. Even though hybrid process modelling has the potential to significantly advance the development of petrochemical industry decarbonisation, to the best of our knowledge, its application in pilot-scale facilities or industrial digital twins has not yet been documented in the literature. One major limitation is the scarcity of experimental data as operating CCUS pilot plant is both time-consuming and expensive. The limited datasets hinder the development of high-fidelity digital twins and highlight the need for methods that can effectively learn from small datasets. LLM is promising but still facing technical shortcomings such as data not effectively standardized. In addition, while LLM have shown great potential in aiding process modelling and optimization, they are still a distance away from being directly applied in industrial CCUS processes. Current limitations include the need for manual prompt design, the lack of deterministic reproducibility due to inherent randomness, and the limited integration with industrial modelling tools via APIs, indicating that the use of LLMs for modelling and optimization remains at an early stage. Future research should focus on improving small-data learning technologies to fully exploit the potential of "AI for modelling" in CCUS applications. Existing strategies to address data scarcity include the use of ML methods suited for small datasets and data augmentation using generative models. PINN can effectively incorporate domain knowledge into the learning process, enabling accurate predictions even with limited training data. Structured data augmentation can be effectively implemented using generative models such as Generative Adversarial Networks (GAN) and Variant Autoencoders (VAE). They carefully preprocess the data and thoroughly assess the quality of the synthesized data.

2.4. AI for LCA and sustainability analysis

In order to deploy CCUS in the petrochemical industry, a comprehensive assessment of the technology is required (Terlouw et al., 2021). LCA is a standardized methodology for evaluating environmental impacts across a product's lifecycle. Recently it has evolved into life cycle sustainability assessment (LCSA) to encompass social and economic dimensions alongside environmental metrics. LCSA, proposed by Guinée et al. (2011), extends the scope of the traditional LCA by covering the three pillars of sustainability: people, planet and prosperity. However, LCSA faces significant challenges due to the complexity of integrating environmental, social and economic data, often leading to fragmented analyses.

AI facilitates the integration of LLM into LCSA, enhancing data processing and analysis capabilities. AI-assisted LCSA has already proven successful. For instance, the Tiangong Project serves as a general demonstration of how LLMs can automate data extraction and harmonization for LCA (Al-Sakkari et al., 2024; Xu, 2023). Although the software tools developed from the Tiangong Project have not yet been applied to CCUS in the petrochemical industry, they could be easily adapted for use in this sector. In the framework of integrating LLM into LCSA, LLMs can process unstructured text from diverse sources

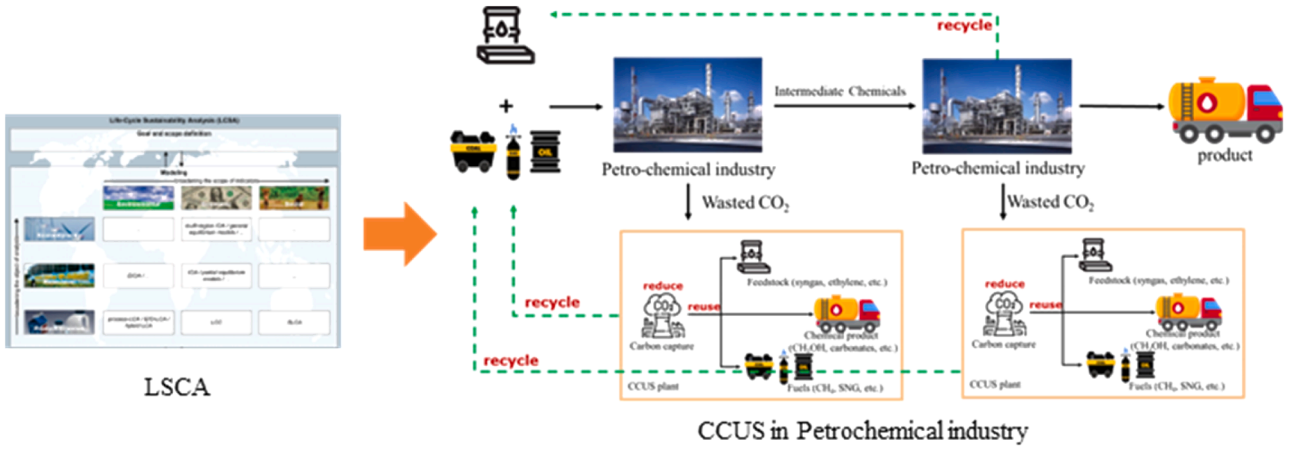

Fig. 5. AI-assisted LCSA for CCUS in petro-chemical industry (Guinée et al., 2011).

(including data sheets, academic publications and regulatory reports) to identify potential life cycle inventory (LCI) data points (Preuss et al., 2024). These data are then validated manually before being entered into an LCI database. This hybrid approach not only accelerates data processing but also enhances accuracy by reducing manual errors. Despite these advances, some open research questions remain. One major gap is the lack of domain-specific LLM trained on high-quality sustainability datasets, which limits the accuracy of the developed LLM.

The application of AI to LCSA holds great potential for enhancing sustainability analysis in the petrochemical industry. However, aligning AI outputs with International Organisation for Standardisation (ISO) standards and ensuring data transparency remain critical challenges. Future research should focus on developing specialized LLMs trained on collated datasets to ensure the reliability and accuracy of the sustainability assessment. By embedding AI into LCSA (refer to Fig. 5), we can address the environmental, social and economic aspects of sustainability in the petrochemical industry, guiding more informed decision-making in CO_2 emission reduction efforts.

3. Conclusion

In this paper, we analysed the challenges faced by petrochemical industry and the role of CCUS in petrochemical industry for decarbonisation. We also explored the opportunities provided by AI for petrochemical industry to reduce CO_2 emission. To support the solvent selection and design for effective carbon capture, we introduce two AI-driven methods using *ChemDataExtractor* (AI as a data mining tool) and *DP-GEN* (AI as a simulation assistant) respectively. Following that, we presented an overview of SACs and their promising potential in CO_2 utilisation and also discussed how AI addresses key challenges in catalyst design for CO_2 utilisation. Subsequently, we introduced the role of hybrid process modelling in petrochemical industry for decarbonisation and provided insights on LLM-assisted surrogate modelling and optimisation. In the end, we used the Tiangong project as an example to demonstrate the potential of LLM for LCSA. Among these four directions, hybrid modelling appears to be most promising for near-term industrial applications due to its relative maturity and extensive academic validation. However, all directions face ongoing challenges. Technically, solvent selection and design, and catalyst design remain at the proof-of-concept stage, with no widespread real-world applications. Hybrid modelling faces data scarcity and fragmentation. In LCSA, the absence of domain-specific LLMs trained on high-quality data limits accuracy and practical value. Economically, access to high-quality industrial data remains costly and restricted. From a regulatory standpoint, there is currently no standardized framework for LCSA tailored specifically to CCUS applications. Continued efforts are needed to establish consistent evaluation methodologies and support future industrial deployment. Through harnessing AI technologies, the petrochemical industry can accelerate its transition to a sustainable low-carbon future.

CRediT authorship contribution statement

Jin Ma: Writing – original draft, Investigation, Conceptualization. **Yide Han:** Writing – original draft. **Meihong Wang:** Writing – review & editing, Writing – original draft, Conceptualization. **Weimin Zhong:** Resources. **Wenli Du:** Validation. **Feng Qian:** Supervision.

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.

Acknowledgement

Professor Meihong Wang is an Advisory Editor of the journal; however, the editor had no access to the peer review or editorial process for this article at any point. The UK authors would like to thank the financial support of the EU RISE project OPTIMAL (Ref: Grant Agreement No:101007963) and CATALYSE (Ref: Grant Agreement No: 10183092).

References

Al-Sakkari, E.G., Ragab, A., Dagdougui, H., Boffito, D.C., Amazouz, M., 2024. Carbon capture, utilization and sequestration systems design and operation optimization: assessment and perspectives of artificial intelligence opportunities. Sci. Total. Environ. 917, 170085. https://doi.org/10.1016/j.scitotenv.2024.170085.

Arnold, C., 2023. Inside the nascent industry of AI-designed drugs. Nat. Med. 29, 1292–1295. https://doi.org/10.1038/s41591-023-02361-0.

Borhani, T.N., Abbasi, M.R., Hosseinpour, M., Salimi, M., Afkhamipour, M., Oko, E., Campbell, K.S., Kahllaghi, N., 2024. CO_2 absorption-desorption cycles: progress, gaps, and future. Carbon Capture Sci. Technol. 13, 100325. https://doi.org/10.1016/j.ccst.2024.100325.

Chen, Y., Ji, S., Chen, C., Peng, Q., Wang, D., Li, Y., 2018. Single-atom catalysts: synthetic strategies and electrochemical applications. Joule 2, 1242–1264. https://doi.org/10.1016/j.joule.2018.06.019.

Chiang, L.H., Braun, B., Wang, Z., Castillo, I., 2022. Towards artificial intelligence at scale in the chemical industry. AIChe J. 68, e17644. https://doi.org/10.1002/aic.17644.

Cullen, L., Meng, F., Lupton, R., Cullen, J.M., 2024. Reducing uncertainties in greenhouse gas emissions from chemical production. Nat. Chem. Eng. 1, 311–322. https://doi.org/10.1038/s44286-024-00047-z.

Gao, P., Li, S., Bu, X., Dang, S., Liu, Z., Wang, H., Zhong, L., Qiu, M., Yang, C., Cai, J., Wei, W., Sun, Y., 2017. Direct conversion of CO2 into liquid fuels with high selectivity over a bifunctional catalyst. Nat. Chem. 9, 1019–1024. https://doi.org/10.1038/nchem.2794.

Guinée, J.B., Heijungs, R., Huppes, G., Zamagni, A., Masoni, P., Buonamici, R., Ekvall, T., Rydberg, T., 2011. Life cycle assessment: past, present, and future. Env. Sci. Technol. 45, 90–96. https://doi.org/10.1021/es101316v.

Gupta, N.C., Tanwar, R., Dipesh, Kaushik, A., Singh, R., Patra, A.K., Sar, P., Khakharia, P., 2024. Perspectives on CCUS deployment on large scale in India: insights for low carbon pathways. Carbon Capture Sci. Technol. 12, 100195. https://doi.org/10.1016/j.ccst.2024.100195.

Han, J., Zhang, L., Car, R., E, W., 2018. Deep Potential: a general representation of a many-body potential energy surface. Commun. Compu. Phys. 23, 629–639. https://doi.org/10.4208/cicp.OA-2017-0213.

Hong, W.Y., 2022. A techno-economic review on carbon capture, utilisation and storage systems for achieving a net-zero CO_2 emissions future. Carbon Capture Sci. Technol. 3, 100044. https://doi.org/10.1016/j.ccst.2022.100044.

Hosseinpour, M., Shojaei, M.J., Salimi, M., Amidpour, M., 2023. Machine learning in absorption-based post-combustion carbon capture systems: a state-of-the-art review. Fuel 353, 129265. https://doi.org/10.1016/j.fuel.2023.129265.

Hussin, F., Md Rahim, S.A.N., Hatta, N.S.M., Aroua, M.K., Mazari, S.A., 2023. A systematic review of machine learning approaches in carbon capture applications. J. CO2. Util. 71, 102474. https://doi.org/10.1016/j.jcou.2023.102474.

IEA, 2018. The Future of Petrochemicals. Paris, https://www.iea.org/reports/the-future-of-petrochemicals.

IEA, 2020. CCUS in clean energy transitions. https://www.iea.org/reports/ccus-in-clean-energy-transitions, Licence: CC BY 4.0.

Jablonka, K.M., Charalambous, C., Sanchez Fernandez, E., Wiechers, G., Monteiro, J., Moser, P., Smit, B., Garcia, S., 2023. Machine learning for industrial processes: forecasting amine emissions from a carbon capture plant. Sci. Adv. 9, eadc9576. https://doi.org/10.1126/sciadv.adc9576.

Javaid, M., Haleem, A., Singh, R.P., 2023. A study on ChatGPT for industry 4.0: background, potentials, challenges, and eventualities. J. Econ. Technol. 1, 127–143. https://doi.org/10.1016/j.ject.2023.08.001.

Jiang, X., Wang, W., Tian, S., Wang, H., Lookman, T., Su, Y., 2025. Applications of natural language processing and large language models in materials discovery. NPJ. Comput. Mater. 11, 79. https://doi.org/10.1038/s41524-025-01554-0.

Kamkeng, A.D.N., Wang, M., Hu, J., Du, W., Qian, F., 2021. Transformation technologies for CO_2 utilisation: current status, challenges and future prospects. Chem. Eng. J. 409, 128138. https://doi.org/10.1016/j.cej.2020.128138.

Kang, Y., Lee, W., Bae, T., Han, S., Jang, H., Kim, J., 2025. Harnessing large language models to collect and analyze metal–Organic framework property data set. J. Am. Chem. Soc. 147, 3943–3958. https://doi.org/10.1021/jacs.4c11085.

Lee, Y.S., Galindo, A., Jackson, G., Adjiman, C.S., 2023a. Enabling the direct solution of challenging computer-aided molecular and process design problems: chemical absorption of carbon dioxide. Comput. Chem. Eng. 174, 108204. https://doi.org/10.1016/j.compchemeng.2023.108204.

Lee, Y.S., Jackson, G., Galindo, A., Adjiman, C.S., 2023b. A predictive model for the techno-economic assessment of CO 2 chemisorption processes applicable to a large number of amine solvents. Chem. Eng. Res. Des. 200, 615–636. https://doi.org/10.1016/j.cherd.2023.11.007.

Lemaoui, T., Boublia, A., Lemaoui, S., Darwish, A.S., Ernst, B., Alam, M., Benguerba, Y., Banat, F., AlNashef, I.M., 2023. Predicting the CO_2 capture capability of deep eutectic solvents and screening over 1000 of their combinations using machine learning. ACS. Sustain. Chem. Eng. 11, 9564–9580. https://doi.org/10.1021/acssuschemeng.3c00415.

Leonzio, G., Hankin, A., Shah, N., 2024. CO_2 electrochemical reduction: a state-of-the-art review with economic and environmental analyses. Chem. Eng. Res. Des. 208, 934–955. https://doi.org/10.1016/j.cherd.2024.07.014.

Leonzio, G., Shah, N., 2024. Recent advancements and challenges in carbon capture, utilization and storage. Curr. Opin. Green. Sustain. Chem. 46, 100895. https://doi.org/10.1016/j.cogsc.2024.100895.

Li, A., Chu, J., Huang, S., Liu, Y., He, M., Liu, X., 2025. Machine learning-assisted development of gas separation membranes: a review. Carbon Capture Sci. Technol. 14, 100374. https://doi.org/10.1016/j.ccst.2025.100374.

Liu, Z., Rolston, N., Flick, A.C., Colburn, T.W., Ren, Z., Dauskardt, R.H., Buonassisi, T., 2022. Machine learning with knowledge constraints for process optimization of open-air perovskite solar cell manufacturing. Joule 6, 834–849. https://doi.org/10.1016/j.joule.2022.03.003.

Ma, J., Dong, Z., Otitoju, O., Wang, M., Du, W., Qian, F., 2024. Heat integration, process design and techno-economic assessment of post-combustion carbon capture using piperazine for large-scale ethylene plant. Chem. Eng. Sci. 284, 119531. https://doi.org/10.1016/j.ces.2023.119531.

Margraf, J.T., Jung, H., Scheurer, C., Reuter, K., 2023. Exploring catalytic reaction networks with machine learning. Nat. Catal. 6, 112–121. https://doi.org/10.1038/s41929-022-00896-y.

Mazheika, A., Wang, Y.G., Valero, R., Viñes, F., Illas, F., Ghiringhelli, L.M., Levchenko, S. V., Scheffler, M., 2022. Artificial-intelligence-driven discovery of catalyst genes with application to CO_2 activation on semiconductor oxides. Nat. Commun. 13, 419. https://doi.org/10.1038/s41467-022-28042-z.

Mok, D.H., Li, H., Zhang, G., Lee, C., Jiang, K., Back, S., 2023. Data-driven discovery of electrocatalysts for CO_2 reduction using active motifs-based machine learning. Nat. Commun. 14, 7303. https://doi.org/10.1038/s41467-023-43118-0.

Nyhus, A.H., Yliruka, M., Shah, N., Chachuat, B., 2024. Green ethylene production in the UK by 2035: a techno-economic assessment. Energy Env. Sci. 17, 1931–1949. https://doi.org/10.1039/D3EE03064D.

OECD, 2022. Plastic pollution is growing relentlessly as waste management and recycling fall short, says OECD.

Otitoju, O., Oko, E., Wang, M., 2021. Technical and economic performance assessment of post-combustion carbon capture using piperazine for large scale natural gas combined cycle power plants through process simulation. Appl. Energy 292, 116893. https://doi.org/10.1016/j.apenergy.2021.116893.

Preuss, N., Alshehri, A.S., You, F., 2024. Large language models for life cycle assessments: opportunities, challenges, and risks. J. Clean. Prod. 466, 142824. https://doi.org/10.1016/j.jclepro.2024.142824.

Qiao, B., Wang, A., Yang, X., Allard, L.F., Jiang, Z., Cui, Y., Liu, J., Li, J., Zhang, T., 2011. Single-atom catalysis of CO oxidation using Pt1/FeOx. Nat. Chem. 3, 634–641. https://doi.org/10.1038/nchem.1095.

Rajulapati, L., Chinta, S., Shyamala, B., Rengaswamy, R., 2022. Integration of machine learning and first principles models. AIChe J. 68, e17715. https://doi.org/10.1002/aic.17715.

Rios, T., Lanfermann, F., Menzel, S., 2024. Large language model-assisted surrogate modelling for engineering optimization, in: 2024 IEEE Conference on Artificial Intelligence (CAI). Presented at the 2024 IEEE Conference on Artificial Intelligence (CAI), IEEE, Singapore, Singapore, pp. 796–803. https://doi.org/10.1109/CAI59869.2024.00151.

Sha, P., Zheng, C., Wu, X., Shen, J., 2025. Physics informed integral neural network for dynamic modelling of solvent-based post-combustion CO_2 capture process. Appl. Energy 377, 124344. https://doi.org/10.1016/j.apenergy.2024.124344.

Sharma, N., Liu, Y.A., 2022. A hybrid science-guided machine learning approach for modeling chemical processes: a review. AIChe J. 68, e17609. https://doi.org/10.1002/aic.17609.

Struebing, H., Ganase, Z., Karamertzanis, P.G., Siougkrou, E., Haycock, P., Piccione, P. M., Armstrong, A., Galindo, A., Adjiman, C.S., 2013. Computer-aided molecular design of solvents for accelerated reaction kinetics. Nat. Chem. 5, 952–957. https://doi.org/10.1038/nchem.1755.

Sun, J., Tu, R., Xu, Y., Yang, H., Yu, T., Zhai, D., Ci, X., Deng, W., 2024. Machine learning aided design of single-atom alloy catalysts for methane cracking. Nat. Commun. 15, 6036. https://doi.org/10.1038/s41467-024-50417-7.

Swain, M.C., Cole, J.M., 2016. ChemDataExtractor: a toolkit for automated extraction of chemical information from the scientific literature. J. Chem. Inf. Model. 56, 1894–1904. https://doi.org/10.1021/acs.jcim.6b00207.

Terlouw, T., Bauer, C., Rosa, L., Mazzotti, M., 2021. Life cycle assessment of carbon dioxide removal technologies: a critical review. Energy Env. Sci. 14, 1701–1721. https://doi.org/10.1039/D0EE03757E.

Wang, H., Zhang, L., Han, J., E, W., 2018. DeePMD-kit: a deep learning package for many-body potential energy representation and molecular dynamics. Comput. Phys. Commun. 228, 178–184. https://doi.org/10.1016/j.cpc.2018.03.016.

Wang, M., Joel, A.S., Ramshaw, C., Eimer, D., Musa, N.M., 2015. Process intensification for post-combustion CO2 capture with chemical absorption: a critical review. Appl. Energy 158, 275–291. https://doi.org/10.1016/j.apenergy.2015.08.083.

William, G.D., Babamohammadi, S., Yang, Y., Masoudi Soltani, S., 2023. The rise of the machines: a state-of-the-art technical review on process modelling and machine learning within hydrogen production with carbon capture. Gas Sci. Eng. 118, 205104. https://doi.org/10.1016/j.jgsce.2023.205104.

Xi, H., Wu, X., Chen, X., Sha, P., 2021. Artificial intelligent based energy scheduling of steel mill gas utilization system towards carbon neutrality. Appl. Energy 295, 117069. https://doi.org/10.1016/j.apenergy.2021.117069.

Xin, H., 2022. Catalyst design with machine learning. Nat. Energy 7, 790–791. https://doi.org/10.1038/s41560-022-01112-8.

Xu, M., 2023. TianGong Initiative - a beacon in sustainable development. https://www.tiangong.earth/.

Yan, Y., Pang, Y.X., Luo, X., Lin, Q., Pang, C.H., Zhang, H., Gao, X., Wu, T., 2024. Carbon dioxide-focused greenhouse gas emissions from petrochemical plants and associated industries: critical overview, recent advances and future prospects of mitigation strategies. Process. Saf. Environ. Prot. 188, 406–421. https://doi.org/10.1016/j.psep.2024.05.136.

Yu, F., Deng, K., Du, M., Wang, W., Liu, F., Liang, D., 2023. Electrochemical CO_2 reduction: from catalysts to reactive thermodynamics and kinetics. Carbon Capture Sci. Technol. 6, 100081. https://doi.org/10.1016/j.ccst.2022.100081.

Zhang, Y., Sha, P., Wang, M., Zheng, C., Huang, S., Wu, X., Cordiner, J., 2025. Toward intelligent and green ethylene manufacturing: An AI-based multi-objective dynamic optimization framework for the steam thermal cracking process. Engineering. https://doi.org/10.1016/j.eng.2025.06.045.

Zhang, Y., Wang, Haidi, Chen, W., Zeng, J., Zhang, L., Wang, Han, E, W., 2020. DP-GEN: a concurrent learning platform for the generation of reliable deep learning based potential energy models. Comput. Phys. Commun. 253, 107206. https://doi.org/10.1016/j.cpc.2020.107206.

Zhong, M., Tran, K., Min, Y., Wang, C., Wang, Z., Dinh, C.T., De Luna, P., Yu, Z., Rasouli, A.S., Brodersen, P., Sun, S., Voznyy, O., Tan, C.S., Askerka, M., Che, F., Liu, M., Seifitokaldani, A., Pang, Y., Lo, S.C., Ip, A., Ulissi, Z., Sargent, E.H., 2020. Accelerated discovery of CO_2 electrocatalysts using active machine learning. Nature 581, 178–183. https://doi.org/10.1038/s41586-020-2242-8.

Zhou, T., McBride, K., Linke, S., Song, Z., Sundmacher, K., 2020. Computer-aided solvent selection and design for efficient chemical processes. Curr. Opin. Chem. Eng. 27, 35–44. https://doi.org/10.1016/j.coche.2019.10.007.

Zhou, W., Cheng, K., Kang, J., Zhou, C., Subramanian, V., Zhang, Q., Wang, Y., 2019. New horizon in C1 chemistry: breaking the selectivity limitation in transformation of syngas and hydrogenation of CO_2 into hydrocarbon chemicals and fuels. Chem. Soc. Rev. 48, 3193–3228. https://doi.org/10.1039/C8CS00502H.