



Original Article

Integration of sustainability assessment into early-stage carbon capture process design with an explainable AI framework

Xin Yee Tai ^a , Oliver Fisher ^b, Lei Xing ^a, Jin Xuan ^{a,*}^a School of Chemistry and Chemical Engineering, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, United Kingdom^b Faculty of Engineering, University of Nottingham, Nottingham, United Kingdom

ARTICLE INFO

Keywords:

Surrogate model
Data-driven model
Life-cycle assessment
Explainable AI
Multi-objective optimisation
CO₂ capture
Sustainability

ABSTRACT

This study introduces a novel framework for reducing environmental impacts by optimising operating conditions using a surrogate modelling approach integrated with Explainable AI (XAI). Two surrogate models were developed: a sequential surrogate model (SSM) with a two-step structure, and a direct surrogate model (DSM) with a single-step architecture. Both were trained on data from a validated physics-based simulation of a monoethanolamine (MEA)-based carbon capture process to predict environmental impacts across human health, ecosystem quality, and resource depletion. SHapley Additive exPlanations (SHAP) were used to enhance transparency by identifying key input variables influencing outcomes. Multi-objective optimisation was conducted using Particle Swarm Optimisation (PSO) and NSGA-II to determine optimal operating conditions. DSM achieved high prediction accuracy (R^2 up to 0.995) and lower errors, while SSM offered better interpretability and broader exploration of Pareto-optimal solutions. This study also shows that our framework identified optimum parameters that reduced environmental impacts by 76–88 % compared with the experiment optimum. This framework supports sustainable process design by combining interpretability, predictive performance, and computational efficiency.

1. Introduction

The chemical industry has evolved through successive industrial revolutions, beginning with mechanisation (Industry 1.0), followed by electrification (Industry 2.0), automation (Industry 3.0), digitalisation (Industry 4.0), and most recently personalisation (Industry 5.0). Each phase has introduced transformative technologies aimed at improving productivity, efficiency, and control. However, across these stages, process optimisation efforts have largely focussed on economic and operational performance, with limited attention paid to the environmental consequences of design and operational decisions. As a result, atmospheric greenhouse gas concentrations have risen from approximately 280 ppm prior to the Industrial Revolution to 422.5 ppm in 2024 (Ahmed et al., 2022; Statista, 2025). In addition, such limitations result in costly redesigns, downtime, or regulatory penalties in the chemical and energy industries. For example, Shell was fined £40,000 for excess propane emissions (The Ferret, 2018), Diageo faced penalties linked to £1.2 million worth of emission (The Ferret, 2022), and Industry Chemical Group Ltd was fined £2.5 million for an uncontrolled acid (Printweek, 2025). In more severe cases, such as the BP Texas refinery

explosion, regulatory and safety failure led to a full site reconstruction lasting over a year, costing the company more than £1.1 billion (Reuters, 2008). In pursuit of responsible and sustainable development, the United Nations urged to decouple economic growth from environmental impacts (UN Environment Programme, 2011). Today, the chemical process industry faces mounting pressure to balance performance with sustainability, driven by tightening environmental regulations and growing societal expectations. Traditional process design approaches, while effective in terms of performance and safety, often fall short in systematically addressing the environmental consequences of chemical manufacturing (European Environmental Bureau, 2021). Optimising environmental impact early in the process design phase is not only essential for sustainable development but also critical for regulatory compliance, resource efficiency, and public trust.

One of the most established methodologies for quantifying the environmental impacts is through life-cycle assessment (LCA), as standardised in ISO 14040-44:2006 (ISO, 2006). It is originally developed in 1976, has steadily evolved in response to growing environmental regulations and sustainability demands within the chemical industry. While LCA provides a comprehensive framework for

* Corresponding author.

E-mail address: j.xuan@surrey.ac.uk (J. Xuan).

quantifying the environmental impacts of chemical processes across their entire life-cycle, its practical application within system-based process design, particularly cradle-to-gate remains limited. This limitation is primarily due to information scarcity during early development stages, when system boundaries, material flows, and process configurations are still uncertain (Haupt et al., 2023; Wowra et al., 2023). LCA can typically only be conducted fully and accurately during the later stages of process design, once detailed and reliable system data are available (Cavalliere et al., 2019). By this point, however, design flexibility is significantly reduced and the cost of implementing changes is prohibitively high. As a consequence, LCA is rarely employed as a proactive decision-making tool during early-stage design, where it could be most effectively influencing sustainable outcomes.

To enable truly sustainable process design, multi-objective optimisation is essential, as it allows simultaneous consideration of economic performance, process efficiency, and environmental impact. Achieving a balance among these often conflicting goals is essential for designing processes that are not only profitable and operationally robust but also environmentally sustainable. Bayesian optimisation, Non-dominated Sorting Genetic Algorithm II (NSGA-II) and Particle Swarm Optimisation (PSO) are widely adopted as optimisation framework. In Bayesian optimisation, Gaussian processes are commonly employed as surrogate models due to their ability to quantify uncertainty, but they often struggle with scalability when applied to large datasets or high-dimensional design spaces (Clayton et al., 2019; Diessner et al., 2022). By contrast, in NSGA-II and PSO framework, deep learning models serve as surrogate models that can efficiently handle large volume of data, offer greater flexibility and scalability while enable rapid evaluation of candidate solutions across the design space. NSGA-II explores a population of potential solutions, while PSO mimics the social behaviour of flocks of birds or schools of fish to find the best solutions. For example, (Ding et al., 2025) demonstrated using NSGA-II to optimise multiple structural parameters with the objective of the CO₂ separation efficiency and pressure-loss ratio. While (Xi et al., 2021) focussed on cost optimisation in carbon capture process using PSO to find the optimal design and operating conditions of the carbon capture process with minimal total annualised cost. However, implementing multi-objective optimisation using traditional first-principle models, also known as physics-based models are often computationally intensive and impractical, especially for complex systems with nonlinear dynamics, uncertain parameters, and high-dimensional design spaces. As a result, AI-based surrogate model has gained increasing attention for their ability to approximate system behaviour with high accuracy and significantly reduced computational cost. These data-driven models enable rapid exploration of trade-offs and facilitate optimisation in cases where mechanistic models are too cumbersome. Recently, deep neural network (DNN) has received substantial attention in chemical sectors due to its high predictive accuracy and low computational cost (Chung and Lee, 2020). For instance, (Oh et al., 2022) employed deep convolutional neural network to predict key performance indicators of the carbon capture process, including CO₂ capture rate, CO₂ concentration in the treated gas, absorber temperature, and stripper pressure under 22 distinct operating conditions. (Ashraf and Dua, 2023) adopted the artificial neural network and support vector machine to predict the CO₂ capture amount from flue gas across a range of operating conditions, including inlet flue gas flow, inlet CO₂ concentration, inlet flue gas pressure, inlet flue gas temperature, lean solvent flow rate, MEA concentration, and lean solvent temperature. The accuracy of data-driven models is highly dependent on both the quantity and quality of the available data. However, many existing studies rely on limited pilot-scale datasets, which can compromise model performance and generalisability. To mitigate this limitation, hybrid modelling approaches are recommended, wherein validated physics-based models are employed to generate high-quality synthetic datasets for training machine learning algorithms (Tai et al., 2022, 2023, 2023). By adopting DNN in PSO and NSGA-II, the framework achieves greater flexibility and

scalability, enabling efficient optimisation of complex, high-dimensional industrial processes where physics-based models and Bayesian optimisation face significant limitations.

These data-driven models can approximate complex process behaviours with high fidelity and dramatically lower computational demands, enabling more efficient exploration of the trade-off space. However, the inherent ‘black-box’ nature of many machine learning (ML) models, such as DNN, poses significant challenges in industrial settings where interpretability, safety, and regulatory compliance are paramount. To address this, Explainable AI (XAI) enables the decision-making process of complex ML models transparent and trackable to human users. Among these intrinsically interpretable models and post-hoc explanation methods are the most common XAI in chemical sectors. For example, Logical Analysis of Data (LAD), an intrinsically interpretable model, was used by (Ragab et al., 2018) to diagnose reduced steam production from black liquor in the boiler, attributed to poor combustion efficiency, heat exchanger scaling, recovery circuit blockages, and abnormal chemical composition affecting reactions. (Wang et al., 2022) used SHapley Additive exPlanations (SHAP) from post-hoc approach to explain relationship between operating and design parameters with current density and N₂ crossover regression in the PEM fuel cell from 9 machine learning algorithms, including extreme gradient boosting (XGBoost), artificial neural network (ANN), and adaptive boosting (AdaBoost), categorical boosting (CatBoost), random forest, extra tree, decision tree, and light gradient boosting machine (LightGBM). In the chemical sector, XAI plays a critical role in building trust among engineers and operators, supporting root cause analysis, enhancing fault detection systems, and enabling regulatory adherence by providing interpretable justifications for model predictions. Moreover, XAI can bridge the gap between data-driven models and domain knowledge, thereby facilitating more robust, human-centred AI integration.

Despite growing interest in sustainable process development, limited attention has been paid to minimising environmental impact through the optimisation of operating conditions. Most existing studies focus on greener process design, material substitution, or life-cycle improvement via structural modifications. While AI has been applied to optimise life-cycle of chemicals based on inventory inputs or process data, its application to operating condition optimisation remains largely unexplored (Kaab et al., 2019; Mayol et al., 2020; Miao et al., 2022). For example, de Jesus et al. used process outputs data such as energy consumption, material usage, heat generation, and waste production as inputs to an adaptive neuro fuzzy inference system (ANFIS) model predict global warming potential (Mayol et al., 2020). This approach can identify optimum inventory inputs, however, there is still a lack of closed-loop framework that directly optimises operating conditions or design parameters. However, even in well-designed systems, suboptimal operating conditions such as temperature, pressure and feed rates can result in excessive energy use, emission and waste generation (Dirza, Krishnamoorthy and Skogestad, 2022). This oversight represents a significant gap, given that operational adjustment is often more flexible and cost-effective than major design changes. To address this challenge, the present study proposes a novel framework to minimise environmental impacts by optimising operating conditions using an AI-driven optimisation framework demonstrated using a case study of monoethanolamine (MEA)-based carbon capture process. This process is selected in this study due to its well-recognised potential for climate change mitigation. However, while its benefit for reducing CO₂ is well established, it is crucial to evaluate its broader environmental impacts to ensure that mitigating one environmental issue does not inadvertently exacerbate others. For example, MEA is synthesised from ethylene oxide and ammonia, with ammonia production alone accounting for approximately 2 % of global CO₂ emissions (IEA, 2021). Additionally, a major limitation of MEA-based carbon capture is the high energy requirement for solvent regeneration, which can substantially offset the net benefits of CO₂ capture, particularly in energy-intensive systems (Salimi et al., 2024). During the capture process, residual CO₂, volatilised MEA, and

wastewater are released into the environment, contributing to a range of environmental impacts, including air pollution, aquatic toxicity, and potential harm to terrestrial ecosystems. MEA-based carbon capture is therefore selected as the case study in this work, as its effectiveness in reducing CO₂ emissions is well established, yet its broader environmental impacts remain underexplored and warrant further investigation.

This study presents a framework of early integration of LCA with XAI framework in the chemical design process to enable sustainability-oriented and transparent decision-making. Two surrogate models with distinct configurations were developed to predict environmental impacts. To improve the reliability of deep data-driven models, SHapley Additive exPlanations (SHAP) were employed to interpret variable influence and reveal patterns within the complex system. Finally, NSGA-II

and PSO algorithms were applied to identify the optimal combinations of operating conditions based on the outputs of the two surrogate models. Given the reduced optimisation performance of algorithms when handling more than three objective functions, this study adopts endpoint-level indicators instead of midpoint-level metrics. The endpoint level summarises overall environmental damage by aggregating midpoint-level metrics across key areas of protection, including human health, ecosystem quality and resource scarcity, which streamlines the optimisation process. This novel framework minimises environmental impacts by optimising operating conditions through an AI-driven surrogate model integrated with XAI, enabling a transparent and informed approach to sustainable process design, where environmental and performance trade-offs can be addressed early, when interventions are more effective and cost-efficient than late-stage

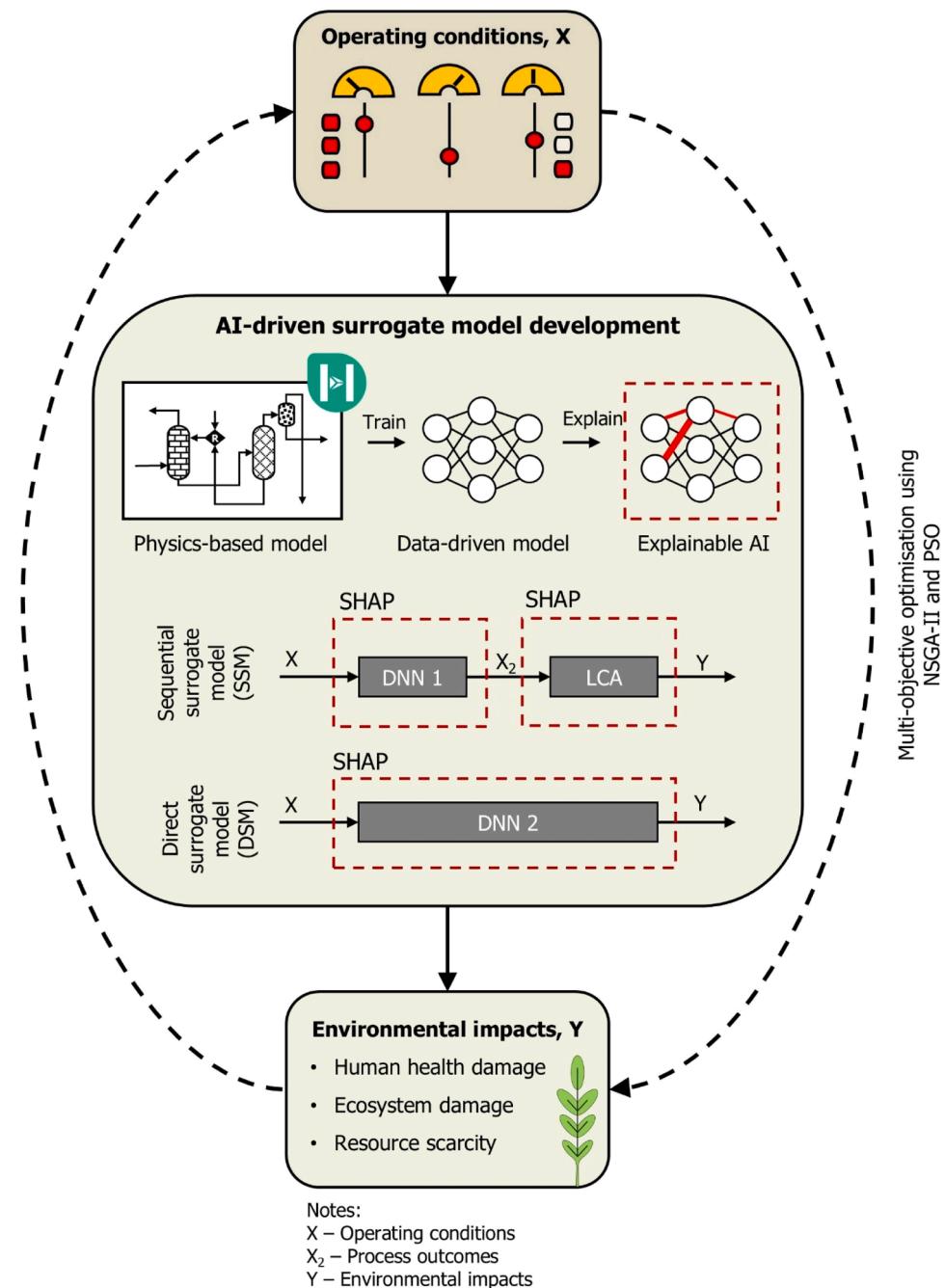


Fig. 1. Schematic diagram of multi-objective optimisation framework using AI-driven surrogate model with XAI, minimising environmental impact of MEA-based carbon capture process through optimising operating conditions.

modifications. The framework shows strong potential in process development by embedding sustainability metrics from LCA into an XAI environment, enabling engineers to visualise environmental impacts early and redesign systems proactively, avoiding costly retrofits. Scaled to the operational stage, it could drive real-time process optimisation that ensures regulatory compliance, maximises economic performance, and mitigates the risk of substantial environmental penalties.

2. Model development

This study focused on minimising the environmental impact of MEA-based carbon capture process through optimising the operating conditions using AI-driven surrogate model with XAI analysis. The carbon capture process separates the CO₂ from the flue gas discharged by a combined cycle power plant. The schematic diagram of this work is demonstrated in Fig. 1. Initially, MEA-based carbon capture process based on pilot data was created in physics-based processing modelling software and generate a database, describing process performance data at various operating condition. Two AI-driven surrogate models, namely sequential and direct approaches, were developed to compare the accuracy and explainability, which both predicted the environmental impacts at various operating conditions. To improve the explainability of the DNN model in this study, SHapley Additive exPlanations (SHAP) from open-source package was adopted (SHAP, 2018). Subsequently, the trained data-driven models were employed as the surrogate model in the multi-objective optimisation problem, minimising total impacts of human health, ecosystem quality and resource scarcity. NSGA-II and

PSO were used in this study to compare the optimisation performance of two AI-driven surrogate model in term of convergence and ability to avoid trap into local minima.

2.1. Physics-based process model

The physics-based process model was developed using Aspen Hysys® based on the pilot plant configuration in Notz et al. (Notz et al., 2012). Aspen Hysys® was selected as physics-based model due to its widespread uses in the other MEA-based carbon capture process studies (Øi, 2012; Qamar et al., 2020) and it handles gas-liquid systems with hydrocarbons and process dynamics well compare to Aspen Plus® (Aspen Tech, 2013). Fig. 2 shows the process flow diagram for capturing CO₂ from flue gas of a combined cycle power plant using aqueous MEA. The simulation composed of (1) absorption column/ absorber in which the CO₂ is captured by amine solvent, (2) desorption column/ desorber in which the solvent is regenerated and recycle and (3) flash column in which to be adopted as an external condenser unit to concentrate the acid gas composition. To ensure consistent performance with the experimental data, the flue gas from a combined cycle power plant was assumed to have been pre-treated to remove H₂S, SO₂, NO₂ and CO. As a result, the flue gas entering the carbon capture process only contains CO₂, N₂, O₂ and H₂O. The flue gas that discharged from pre-treatment process was cooled down before entering at the bottom of absorption column. The flue gas flew upward and contacted with CO₂ absorbing solvent, aqueous MEA in a counter current flow. Vent gas exits the top and rich amine leaves through bottom. Rich amine was pumped to a

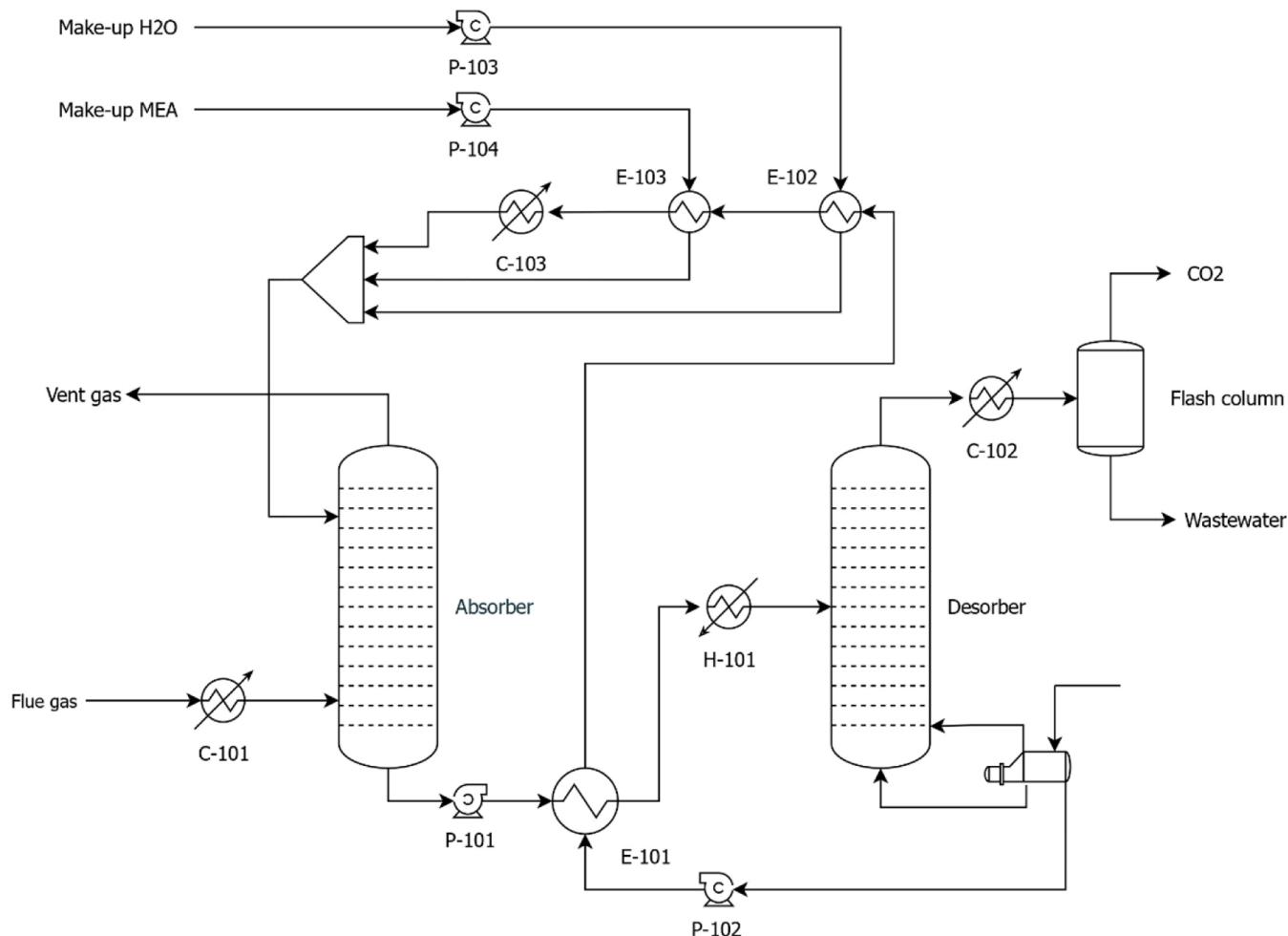


Fig. 2. Process flow diagram of MEA-based carbon capture process.

cross heat exchanger and then to a desorption column for solvent regeneration at high temperature and pressure. During the regeneration process, the reactions between CO₂ and aqueous MEA are reversed using heat, producing CO₂ and water vapour are the top product and regenerated MEA as the bottom products. The top gases were further sent to a condenser in a flash column, improving the purity of CO₂ by removing excess water vapour. The lean amine leaves at the bottom of desorption column which contains regenerated MEA and low concentration of CO₂ is returned to the absorption column. To compensate for the loss of water and MEA during the process, separate make-up streams are added for water and MEA to maintain the efficiency of the CO₂ capture process.

The Acid Gas – Chemical Solvent property package was selected in Aspen HYSYS to model the CO₂ capture process using MEA as a reactive solvent. This package is designed to simulate system involving acid gas removal via chemical solvents, incorporating rate-based models that account for reaction kinetics mass transfer, and thermodynamic equilibrium. In order to reflect the actual pilot carbon capture process, the process model used the same parameters for absorption and desorption columns as the pilot plants, such as packing material, column diameters, and packed heights which shows in SI, Table S1. The physics-based model was validated against experimental data across the full operating range, including solvent flow rate of 75–350 kg/h, flue gas flow rate of 55–100 kg/h, desorber pressure of 100–225 kPa, MEA fraction of 0.1–0.35 kg/kg, MEA temperature of 30–50 °C, flue gas temperature of 23–50 °C, CO₂ mole fraction in flue gas of 0.035–0.059 mol/mol, and reboiler duty of 6–16 kW. It was subsequently employed to generate database of process outputs under varying operating conditions, constrained to the same range as the experimental data.

Latin hypercube sampling was employed to generate 1000 operating conditions sets within the validated range. Of these, 925 data successfully converged in Aspen Hysys. The final dataset therefore comprises 925 observations, each containing 8 operating conditions and 18 process outputs, which are summarised in Table S2 and Table S3, respectively. The operating conditions include solvent flow rate, flue gas flow rate, desorber pressure, MEA fraction, MEA temperature, flue gas temperature, CO₂ mole fraction in the flue gas, and reboiler duty. While the process output composed of material, energy, product, wastewater, and emission data, which would be adopted to evaluate the environmental impacts. Materials include make-up water and make-up MEA, energy inputs comprise the desorber energy, two coolers, two heaters, and four pumps; products include the CO₂ product stream, while waste refers to the wastewater stream, and emissions cover vent gases such as CO₂ and MEA.

2.2. AI-driven surrogate models

In this study, two AI-driven surrogate models were developed in sequential and direct approaches, namely sequential surrogate model (SSM) and direct surrogate model (DSM), respectively to compare the accuracy and explainability, which demonstrates in Fig. 1.

Sequential approach involves two-steps predictions, predicting the process output, X₂ at the first step using DNN model in Eq. (1) and subsequently passing the LCA model to evaluate the environmental impacts in Eq. (2). The physics-based process data, including operating conditions and process outputs were adopted to use as model input and output, respectively to train DNN model. The predicted output from DNN model is fed into the LCA model to evaluate the total damage of human health, ecosystem quality and resource scarcity. Due to large output size involves in the DNN model, dropout technique with a probability of 10 % was used. The SSM is expressed as follow.

$$X_2 = f \left(b + \sum_{i=1}^n X_i \cdot w_i \right) \quad (1)$$

$$[f_1(X), f_2(X), f_3(X)] = \text{LCAmoel}(X, X_2) \quad (2)$$

where X and X₂ are operating conditions and process outputs, respectively. f₁(X), f₂(X) and f₃(X) are human health damage, ecosystem damage and resource scarcity, respectively. f(), b and w are activation function, bias and weights, respectively.

In direct approach, a single-step DNN model was employed to predict the environmental impact at various operating conditions as shown in Eq. (3). Initially, physics-based process data were passed through the LCA model to calculate the corresponding environmental impacts. These operating conditions and their associated environmental impacts were then used as the model's input and output, respectively. The mathematical equation of DSM is stated as follow.

$$[f_1(X), f_2(X), f_3(X)] = f(b + X_i \cdot w_i) \quad (3)$$

All dataset was divided into 70:15:15 ratio for training, validation and testing, respectively. Training and validation samples are randomly selected to ensure unbiased representation and were stored for subsequent hyperparameters optimisation. The training data was normalised to 0 to 1 for effective training efficiency. Grid search method is used to determine optimal batch size, number of epochs, training optimisation algorithm, learning rate, activation function, dropout rate, number of neurons, and number of layers. This identifies the best configuration of DNN model for sequential and direct approach based on the value of mean squared error (MSE). Since normalised values were used to calculate the MSE, the resulting error is unitless. The optimised hyperparameters of DNN models for sequential and direct approaches are shown in Table S4.

2.3. Life-cycle assessment

Life-cycle assessment (LCA) model was developed to evaluate the environmental impacts at various operating condition. In sequential approach, LCA model was used to evaluate the environmental impacts based on the predicted outputs from DNN model. While in direct approach, LCA model was used to evaluate the environmental impacts, and the value was employed as the model output in the DNN model.

The goal of the study is to determine the environmental impacts of MEA-based carbon capture at various operating conditions. The functional unit in this study is capturing 1 ton of CO₂ from pre-treated flue gas at combined cycle power plant using aqueous MEA without its subsequent storage. The assessment method used in this study is Huijbregts et al. (2017). Due to deterioration of algorithm performance in optimising more than three objective function, endpoint level is used, rather than midpoint level to reflect damage at three area of protection, including human health, ecosystem quality, resource scarcity. The area of protection is derived from 17 midpoint impact categories with a constant mid-to-endpoint factor per impact category (Huijbregts et al., 2017). Such methodology is in accordance with the ISO 14040:2006 system standard (ISO, 2006).

Endpoint area of protection includes damages to human health, ecosystems and resources. Human health damage is measured in DALY (disability-adjusted life years), measures the burden of disease resulting from environmental impacts. The inventory data or process outputs are first multiplied by midpoint characterisation factors to derive midpoint indicators. Each midpoint indicator is subsequently multiplied by its corresponding endpoint conversion factor, yielding the total damage in the areas of human health, ecosystems, and resources. One DALY refers to the loss of one year of healthy life due to premature mortality or time lived in less than full health (Huijbregts et al., 2017). This is caused by particulate matter, tropospheric ozone formation, ionising radiation, stratospheric ozone depletion, human toxicity (cancer and non-cancer), global warming and water consumption from midpoint impact category. Ecosystem quality (species.year) quantifies the potential loss of local species over the time (Huijbregts et al., 2017). One species. year represents one species lost in a region for one year due to environmental changes, which includes global warming, water use, freshwater

ecotoxicity, freshwater eutrophication, tropospheric ozone, terrestrial ecotoxicity, terrestrial acidification, land use or transformation, and marine ecotoxicity. While resource scarcity (USD, 2013 equivalent) estimates the increased economic cost of future resource extraction due to depletion of minerals and fossil fuels (Huijbregts et al., 2017). It reflects how much more expensive it will be to extract the next unit of resource because the available reserves are used up, for example, mineral and fossil.

This work simulates a MEA-based carbon capture pilot plant that captures the flue gas from combined cycle power plant at King's Lynn, United Kingdom. The system boundary is cradle-to-gate which illustrates in Fig. 3. Energy for heating, cooling and fluid pumping is supplied by the combined cycle power plant, as detailed in Table S5, where the associated energy impacts are included. The environmental impact of MEA is account for via the make-up stream, due to losses through the vent gas. This impact includes both the manufacturing of MEA and its transport from factory in Lavera, France to King's Lynn, United Kingdom. The environmental assessment also considers CO₂ and MEA emissions through the vent gas, the consumption of deionised water for the make-up water stream, and wastewater discharge.

The midpoint characterisation factors and endpoint conversion factors were taken from Ecoinvent version 3.10, converting inventory data into endpoint impacts to human health, ecosystem damage, and natural resource and integrate them into the Python to evaluate the environmental impact at various operating conditions (ecoinvent, 2023). With this LCA framework, the range of human health damage, ecosystem quality and resource scarcity are 0.62–2.24 DALY, 0.001–0.004 species·year, and 2.4–8.6 × 10⁴ USD, 2013 equivalent which is based on the data framework from Table S2 and Table S3.

2.4. Optimisation of environmental impacts

The objective of this work is to minimise the total impacts of human health, ecosystem quality and resource scarcity and identifying the corresponding optimum operating conditions.

Minimise

$$F(X) = (f_1(X), f_2(X), f_3(X))$$

$$X = [X_1 \ X_2 \ X_3 \ X_4 \ X_5 \ X_6 \ X_7 \ X_8]^T$$

where X₁ to X₈ are solvent flow rate (kg/h), flue gas flow rate (kg/h), desorber pressure (kPa), MEA fraction (kg/kg), MEA temperature (°C), flue gas temperature (°C), CO₂ mole fraction in flue gas (mol/mol) and reboiler duty (kW).

Then the constraints are as follow.

$$\begin{aligned} 75 \leq X_1 &\leq 350, 55 \leq X_2 \leq 100, 100 \leq X_3 \leq 225, 0.10 \leq X_4 \\ &\leq 0.35, 30 \leq X_5 \leq 50, 23 \leq X_6 \leq 50, 0.035 \leq X_7 \leq 0.059, \text{ and } 6 \\ &\leq X_8 \leq 16. \end{aligned}$$

Here, Non-dominated Sorting Genetic Algorithm II (NSGA-II) and Particle Swarm Optimisation (PSO) are used. NSGA-II uses non-dominated sorting method while PSO adopts swarm search to improve the exploitation of the search. NSGA-II offers good convergence while PSO allows to escape local minima. The optimisation parameters for NSGA-II are the same for both the direct and sequential approaches, and similarly, those for PSO are also identical across the two approaches. The NSGA-II algorithm was set with a population size of 200, a crossover probability of 0.7, a mutation probability of 0.5, a mutation strength of 0.05, and a maximum of 100 generations. The PSO was configured with a population and repository size of 200, a maximum of 100 generations, an inertia weight of 0.4, individual and swarm confidence factors of 2, 20 grids per dimension, a maximum velocity of 5 %, and a uniform mutation rate of 0.5 %. Grid search method was applied to identify the optimal parameter settings for NSGA-II and PSO, with evaluation criteria including whether the Pareto front is well-formed, solutions are diverse, and results are consistent across multiple runs. After running the multi-objective optimisation, a Pareto-optimal solution was obtained, representing a set of trade-off solutions where improving one objective would worsen at least one other.

3. Result and discussion

3.1. Model validation

3.1.1. Physics-based process model

The physics-based process model validation of the MEA-based carbon capture process was validated using the pilot plant configuration reported by Notz et al., 2012 (Notz et al., 2012). This validation step is significant for generating high-quality data that accurately reflect the real pilot carbon capture process, which is then used to build an accurate DNN model. The key indicators used to compare the experiment data and physics-based simulation results include specific energy, absorbed CO₂ mass flow, concentration of CO₂ in rich and lean solvent, and temperature profile for absorber and desorber. Fig. 4(a) – (d) shows an excellent agreement between the physics-based process data and experimental pilot data across 33 datasets, covering specific energy consumption, absorbed CO₂ mass flow rate, and absorbed CO₂ in both rich and lean solvents, under a wide range of operating conditions as detailed in Table S2. Fig. 4(e) presents a comparison of the temperature profiles for the desorber and absorber between simulation and experimental results, at a solvent flow rate of 200 kg/h, flue gas flow rate of

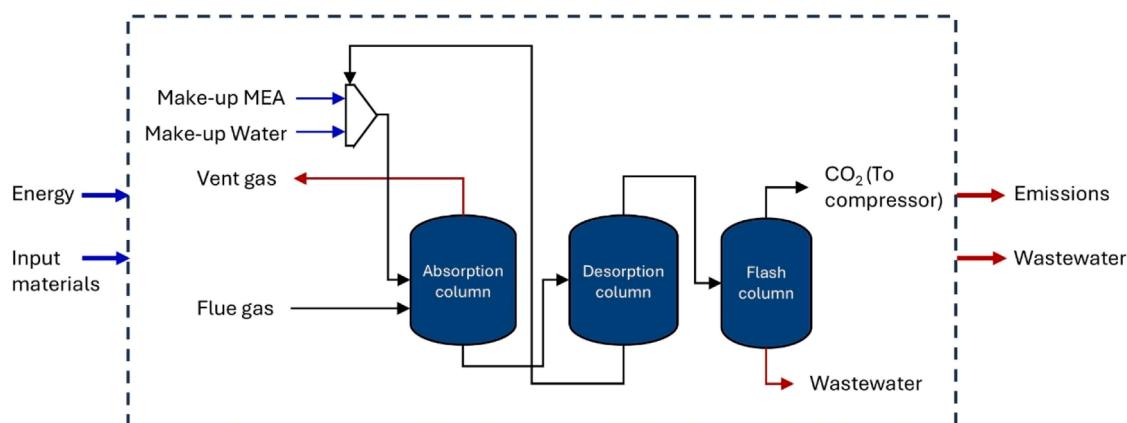


Fig. 3. System boundary of MEA-based carbon capture process.

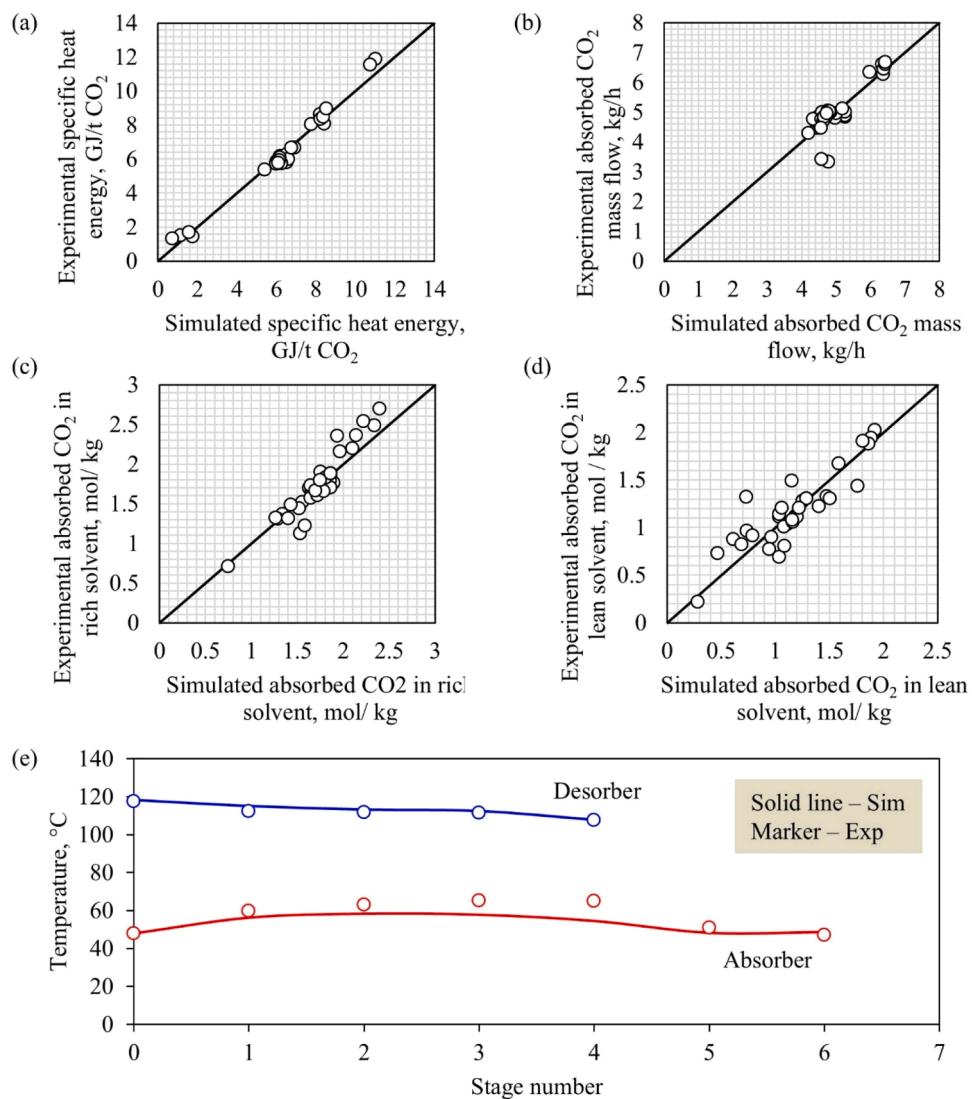


Fig. 4. Results of comparison between physics-based process simulation and experimental pilot data: (a) specific energy, (b) absorbed CO₂ mass flow rate, (c) absorbed CO₂ concentration in rich solvent, (d) absorbed CO₂ concentration in lean solvent, and (e) temperature profile of desorber (blue) and absorber (red). Solid line represents the physics-based process simulation results while the marker refers to experimental data.

71.2 kg/h, flue gas CO₂ concentration of 0.035 mol/mol, desorber pressure of 200 kPa, MEA mass fraction of 0.30 kg/kg, reboiler duty of 7.9 kW, MEA temperature of 40 °C, and flue gas temperature of 48 °C. The largest deviations between experimental and simulated absorber temperatures were observed at stages 3 and 4. These discrepancies may be attributed to limitations of the Kent–Eisenberg model, which is employed by default in the Acid Gas–Chemical Solvent package, as it tends to underestimate CO₂ solubility and the effective absorption driving force within the solvent concentration and temperature ranges considered. The average relative errors for specific energy, absorbed CO₂ mass flow rate, absorbed CO₂ in the rich solvent, absorbed CO₂ in the lean solvent, and the desorber and absorber temperature profiles for 33 runs are 4.54 %, 4.23 %, 4.95 %, 9.73 %, 1.18 %, and 7.08 %, respectively. Given that the comparison between predicted and pilot results shows <10 % uncertainty, the physics-based process model demonstrates a reliable prediction of the carbon capture process using an MEA solvent and serves as a solid foundation for training the DNN model. (Zhang et al., 2009; Li et al., 2016).

3.1.2. Data-driven surrogate model

Adopting data-driven model as the surrogate model significantly

improve the effectiveness of solving multi-objective optimisation problem. In this work, two surrogate models, namely SSM and DSM were developed to predict the environmental impacts at various operating conditions. As SSM involves a two-step prediction, two separate validation steps are required: one for the process outputs and one for the environmental impacts. Test data was employed for process output validation to evaluate the model's generalisation capability on unseen data, thereby ensuring the accuracy of its predictions. Fig. 5 demonstrates 18 process outputs, including material consumption, energy usage and emission rate for carbon capture process, comparing the results between physics-based process data and DNN prediction results. The average R-squared and mean squared error for process output are 0.9845 and 5.382×10^{-4} , respectively.

In SSM, process output was passing through LCA model to evaluate the environmental impacts. To evaluate the model performance, the environmental impacts of SSM are validated by comparison with the target results, derived from original process data passed through the LCA model. Fig. 6(a), (b) and (c) demonstrates human health damage, ecosystem damage and resource scarcity, respectively for SSM between targeted data and prediction results. R-squared and mean square error (MSE) were used to evaluate the model accuracy. The parity plot shows

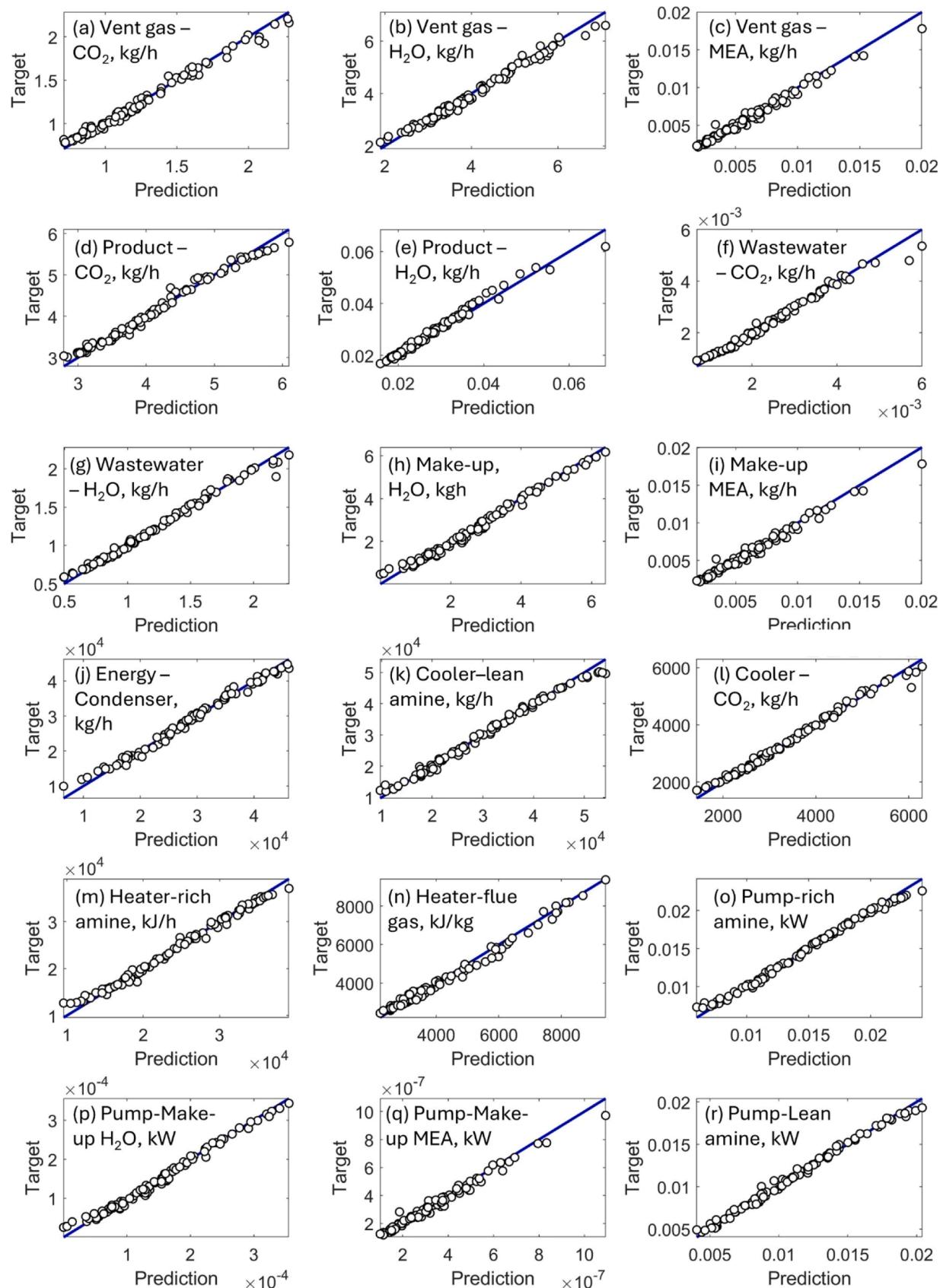


Fig. 5. Comparison of 18 process outputs between physics-based process data and DNN prediction model for SSM. Process outputs include (a) Vent gas - CO₂, kg/h, (b) Vent gas - H₂O, kg/h, (c) Vent gas - MEA, kg/h, (d) Product - CO₂, kg/h, (e) Product - H₂O, kg/h, (f) Wastewater - CO₂, kg/h, (g) Wastewater - H₂O, kg/h, (h) Make-up H₂O, kg/h, (i) Make-up MEA, kg/h, (j) Energy - Condenser, kJ/h, (k) Cooler - Lean amine, kJ/h, (l) Cooler - CO₂ product, kJ/h, (m) Heater - Rich amine, kJ/h, (n) Heater - Flue gas, kJ/kg, (o) Pump - Rich amine, kW, (p) Pump - Make-up H₂O, kW, (q) Pump - Make-up MEA, kW, (r) Pump - Lean amine, kW.

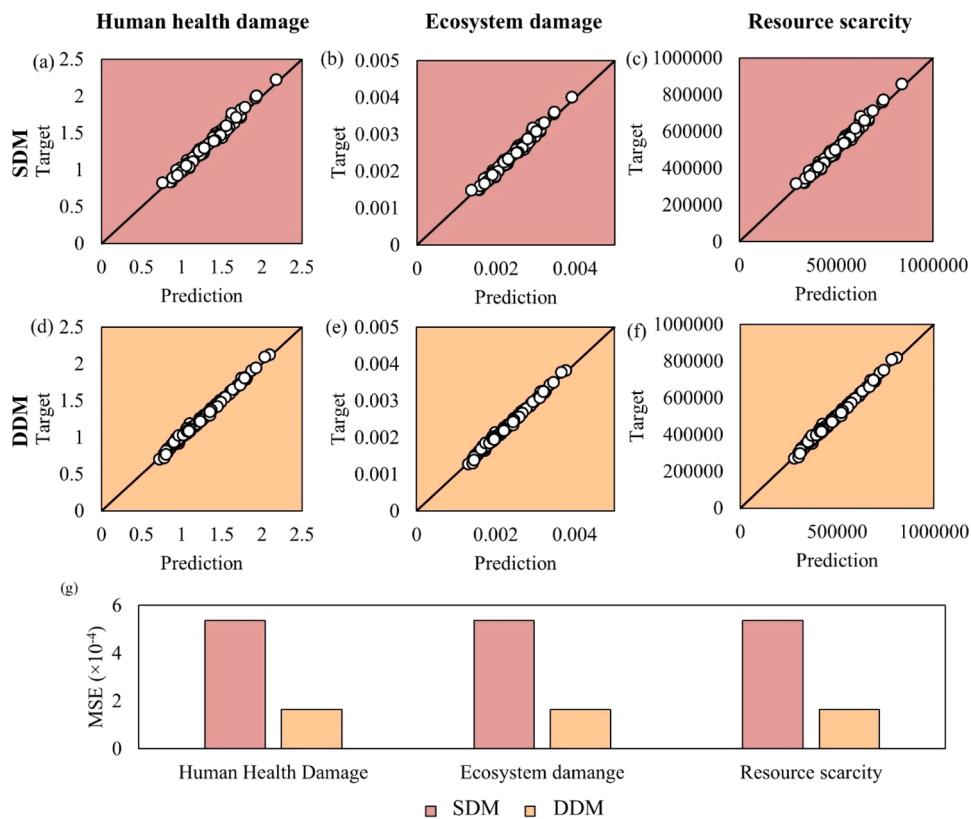


Fig. 6. Model validation results for (a)–(c) SSM and (d)–(f) DSM, across human health damage, ecosystem quality, and resource scarcity. (g) Comparison of MSE results for environmental impacts between SSM and DSM.

the MSE value of 5.356×10^{-4} , 5.369×10^{-4} and 5.355×10^{-4} for human health damage, ecosystem damage and resource scarcity, respectively. The average R-squared value of three total impacts is 0.984.

In DSM, environmental impacts were directly predicted from DNN model. Fig. 6(d) – (f) present the model validation result in a parity plot for DSM. The average R-squared value for three total impacts of area protections is 0.995. The MSE for human health damage, ecosystem quality and resource scarcity for DSM are 1.627×10^{-4} , 1.631×10^{-4} and 1.625×10^{-4} , respectively. Based on the R-squared and MSE values, DSM, which shows lower values for both metrics, demonstrates higher accuracy in predicting environmental impacts compared to SSM. The comparison of SSM and DSM in term of MSE value is shown in Fig. 6(g) and the summary of MSE and R-squared of SSM and DSM are listed in Table 1. Overall, DSM is 80 % quicker than SSM in predicting the environmental impacts and 86 % faster in DNN training. The time taken for training and prediction are listed in Table S6.

3.2. Explainable AI (XAI)

Although machine learning models demonstrated excellent predictive performance, the inherent black-box nature of the model reduces its credibility and reliability, limiting its potential for real-world application. To address this, the open-source package SHAP (SHapley Additive exPlanations) was employed to explain the internal relationships within

the machine learning model, thereby improving its explainability and reliability. As a result, the model's predictive interpretations align closely with existing literature and prior knowledge.

SHAP adopts game theoretic approach to explain the output of machine learning model. As SSM involves a two-step prediction process, two separate SHAP analyses were conducted to identify the importance of operating conditions and intermediate process outputs influencing the environmental impacts.

The SHAP value distributions of process outputs of the SSM for human health damage, ecosystem damage and resource scarcity are demonstrated in Fig. 7(a), (b) and (c), respectively. The x-axis represents the SHAP value (impact on model output), and y-axis shows the most significant 7 studied process outputs. The colour represents the parameter value of each sample. The warmer the colour, the higher the feature values. We can easily identify the most significant process output affecting human health damage, ecosystem damage, and resource scarcity is the CO₂ product yield, followed by the energy profile, which includes lean MEA cooling, desorber reboiler, desorber condenser, rich MEA heating, flue gas cooling, and CO₂ cooling. SSM has 18 process outputs, but only 7 of them have a significant impact on the environmental impacts, and thus Fig. 7(a) – (c) displays only these top 7. CO₂ product yield negatively correlates with human health damage, ecosystem quality and resource scarcity, and the others are positively related. To identify the operating conditions that have the greatest impact on the CO₂ product yield, a SHAP analysis was conducted on the

Table 1
Comparison of MSE and R-squared of environmental impact value in SSM and DSM.

Surrogate model	SSM			DSM		
Environmental impact	Human health damage	Ecosystem damage	Resource scarcity	Human health damage	Ecosystem damage	Resource scarcity
MSE	5.356×10^{-4}	5.369×10^{-4}	5.355×10^{-4}	1.627×10^{-4}	1.631×10^{-4}	1.625×10^{-4}
R-squared	0.984	0.995				

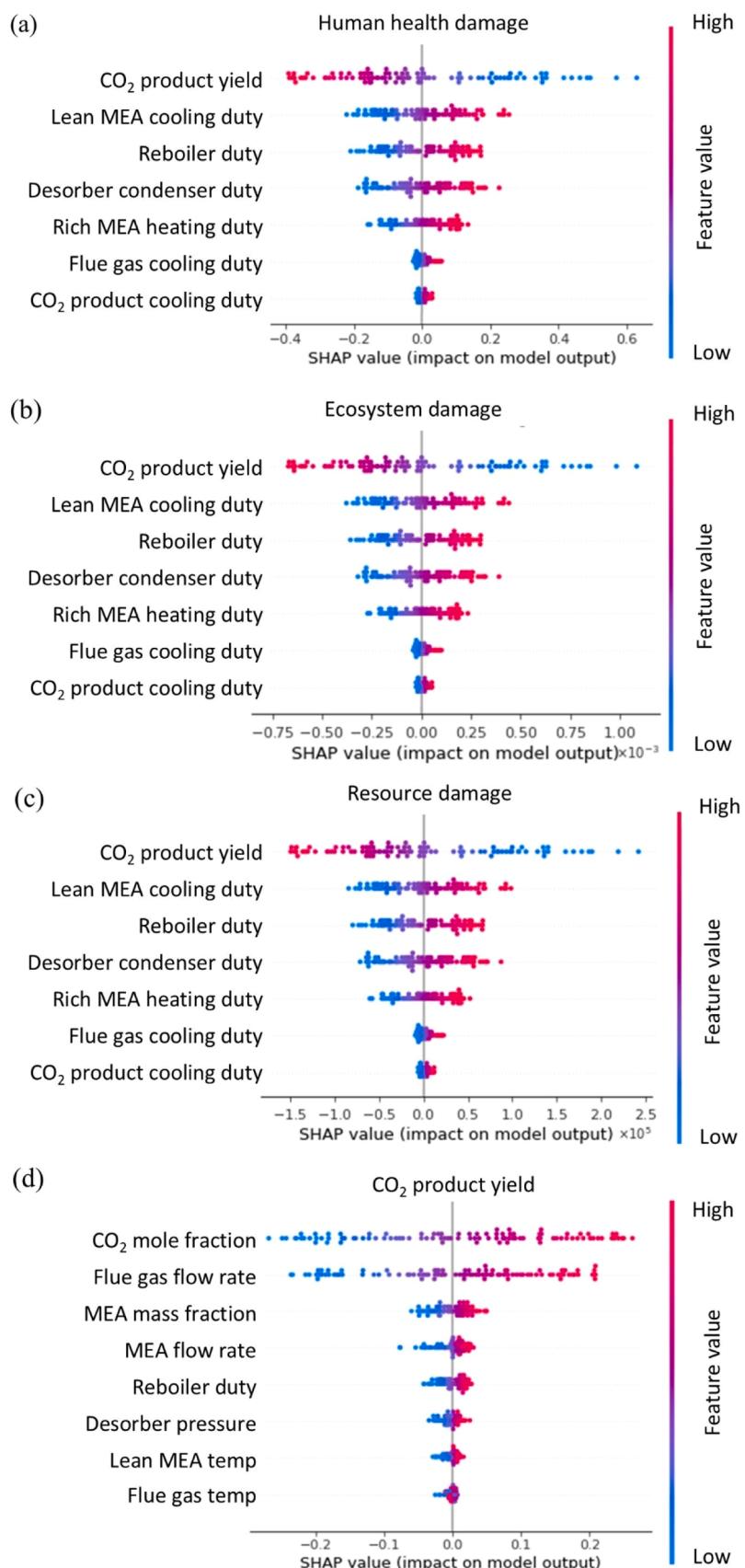


Fig. 7. SHAP analysis for SSM, revealing the impact of process outputs in (a) human health damage, (b) ecosystem quality, and (c) resource damage. (d) Second SHAP analysis was performed to identify the significance of operating conditions to the CO₂ product yield, which is the top process outputs in affecting environmental impacts.

DNN model in SSM. According to Fig. 7(d), CO₂ mole fraction is the most significant factor for CO₂ product yield in SSM. CO₂ mole fraction, flue gas flow rate, MEA mass fraction, MEA flow rate, reboiler duty, desorber pressure, MEA temperature, flue gas temperature is positively correlated with the CO₂ product yield, whereas the others are opposite. We can conclude that high CO₂ mole fraction impacts positively product yield, and negatively affecting human health damage, ecosystem damage and resource scarcity, thereby contributing to improved environmental outcomes.

Fig. 8(a), (b) and (c) shows SHAP result of DSM, representing human health damage, ecosystem damage, resource scarcity, respectively. Since DSM involves one prediction step, therefore, only one SHAP analysis was performed. Similar to SSM, CO₂ mole fraction is the most importance operating conditions for the environmental impacts in DSM. However, the sequence of operating conditions following the CO₂ mole fraction differs from that in SSM, ranking in order as reboiler duty, MEA flow rate, flue gas flow rate, MEA mass fraction, MEA temperature, desorber pressure, and flue gas temperature. This is due to the structure of model in predicting environmental impact from operating conditions.

A parametric study of CO₂ mole fraction was conducted to explain the correlation of operating conditions to the process outputs and environmental impacts. Fig. 9(a) and (b) presents the impact of CO₂ mole fraction on the CO₂ product yield and subsequently affecting the environmental impacts in SSM. When CO₂ mole fraction in the flue gas increase, the CO₂ product yield increase as well, thereby reducing the damages to the human health, ecosystem and resources. Similarly, in DSM, high CO₂ mole fraction reduces the environmental damage of carbon capture CO₂. However, compared to SSM, DSM lacks intermediate information to explain how a high CO₂ mole fraction can lead to low environmental impacts.

In the previous discussion, DSM demonstrated higher prediction accuracy compared to SSM. However, in terms of explainability, SSM performs better than DSM due to the availability of intermediate information on process outputs.

3.3. Optimisation of environmental impacts

NSGA-II and PSO algorithm were employed to minimise the environmental impacts, including human health damage, ecosystem damage and resource scarcity and identifying corresponding optimum operating conditions. The previously trained SSM and DSM were adopted as surrogate model for this multi-objective optimisation. This section will explore how different optimisation algorithms affect the performance of optimisation in both SSM and DSM, each with two distinct surrogate model configurations.

3.3.1. Optimisation with sequential surrogate model (SSM)

Fig. 10 and Fig. 11 illustrate the Pareto front obtained by NSGA-II and PSO in both SSM and DSM, respectively. For SSM, NSGA-II found just one optimum point, which is at human health damage of 0.5324 DALY, ecosystem damage of 9.61×10^{-4} species.year, resource scarcity of 2.051×10^5 USD, 2013 equivalent. The corresponding operating conditions include a solvent flow rate of 95.3 kg/h, flue gas flow rate of 92.8 kg/h, desorber pressure of 205 kPa, MEA mass fraction of 0.32 kg/kg, MEA temperature of 44 °C, flue gas temperature of 40 °C, reboiler duty of 6.6 kW, and a CO₂ mole fraction of 0.059 mol/mol, which demonstrated in Figure S1 (a) – (h).

Whereas PSO produced a range of optimum solutions as shown in Fig. 10(b), including a human health damage of 0.52948 – 0.52952 DALY, ecosystem damage of $9.5684 - 9.5688 \times 10^{-4}$, and resource scarcity of $2.0470 - 2.0471 \times 10^5$ USD, 2013 equivalent. The optimum solvent flow rate and desorber pressure are at the range of 105 – 107 kg/h and 200–204 kPa, respectively, while the other operating conditions including flue gas flow rate, MEA mass fraction, MEA temperature, flue gas temperature, reboiler duty and CO₂ mole fraction have remained at 100 kg/h, 0.33 kg/kg, 30.2 °C, 48.9 °C, 6.175 kW, and 0.059 mol/mol,

respectively. The range of optimum parameters are illustrated in Figure S1.

In SSM, PSO outperformed NSGA-II by avoiding local minima and effectively identifying the optimal parameters that minimise environmental impacts.

3.3.2. Optimisation with direct surrogate model (DSM)

Similarly, NSGA-II and PSO were also adopted to find the minimum environmental impacts in DSM and both Pareto fronts are illustrated in Fig. 11(a). The results show that improving one objective function inevitably worsens at least one other, indicating a Pareto-optimal set. The range of human health damage, ecosystem damage and resource scarcity of NSGA-II includes 0.039 – 0.052 DALY, $5.232 - 7.033 \times 10^{-5}$ species.year and $2.92 - 4.47 \times 10^4$ USD, 2013 equivalent, respectively. IN comparison, PSO identified slightly narrow ranges of 0.039 – 0.051 DALY, $5.242 - 7.185 \times 10^{-5}$ species.year, and $2.93 - 4.33 \times 10^4$ USD, 2013 equivalent, respectively.

NSGA-II identified a broad range of operating parameters for the minimum environmental impacts, including a solvent flow rate of 286–324 kg/h, flue gas flow rate of 94–100 kg/h, and flue gas temperature of 23–29 °C. In comparison, PSO yielded a solvent flow rate of 284–322 kg/h, a flue gas flow rate of 94–100 kg/h, and a flue gas temperature of 23–28 °C. Both optimisation algorithms resulted in the same values for the remaining operating conditions: a desorber pressure of 227.6 kPa, MEA mass fraction of 0.33 kg/kg, MEA temperature of 49.9 °C, reboiler duty of 6.175 kW, and CO₂ mole fraction of 0.059 mol/mol. The comparison of operating condition for both optimisation algorithms are presented in histogram that shows in SI in Figure S2 (a)-(h).

3.3.3. Discussion of NSGA-II and PSO in SSM and DSM

NSGA-II is considered more effective than PSO in this minimisation problem because it consistently converged to slightly lower objective function values in DSM, according to Fig. 11(b), (c) and (d). This indicates better optimisation performance, especially when aiming to minimise environmental impacts. Unlike PSO, which often struggles with premature convergence, NSGA-II is designed to maintain diversity and explore the search space more thoroughly through its non-dominated sorting and crowding distance mechanisms. As a result, it is better at identifying a well-distributed set of Pareto-optimal solutions with minimal trade-offs. Overall, DSM identified the Pareto front approximately 99 % faster than SSM in both NSGA-II and PSO, with a summary of the computational times provided in Table S6.

In contrast to SSM, DSM exhibits a more distinct trade-off between the objective functions and converges to a more diverse set of solutions. In addition, DSM converged to a lower environmental impacts compared to SSM. This difference arises from the structural configuration of the two models. SSM employs a two-step prediction approach, where the operating conditions influence the process output, then subsequently environmental impacts. In other words, the impact of operating conditions on environmental outcomes is indirect in SSM, whereas DSM provides a more direct relationship, enabling a clearer optimisation trade-off.

In this work, reduction in human health damage cannot be achieved without exacerbating either ecosystem damage or resource scarcity. This trade-off is reflected in the optimisation results in both PSO (SSM and DSM) and NSGA-II (DSM) identified multiple solutions illustrating this compromise. In contrast, NSGA-II applied to SSM yielded only a single solution, likely due to convergence to a local minimum. According to Fig. 10(b), the minimum value of human health damage, ecosystem damage and resources scarcity for SSM, optimised using PSO are 0.52949 DALY, 9.5683×10^{-4} species.year and 2.047014×10^5 USD, 2013 equivalent, respectively. Overall, PSO proposed to adopt low solvent flow rate and high desorber pressure in the carbon capture process to minimum the impacts to human health, ecosystem and resource. Slightly change in these two operating conditions can lead to achieve the desired environmental outcomes. For example, using solvent flow rate of

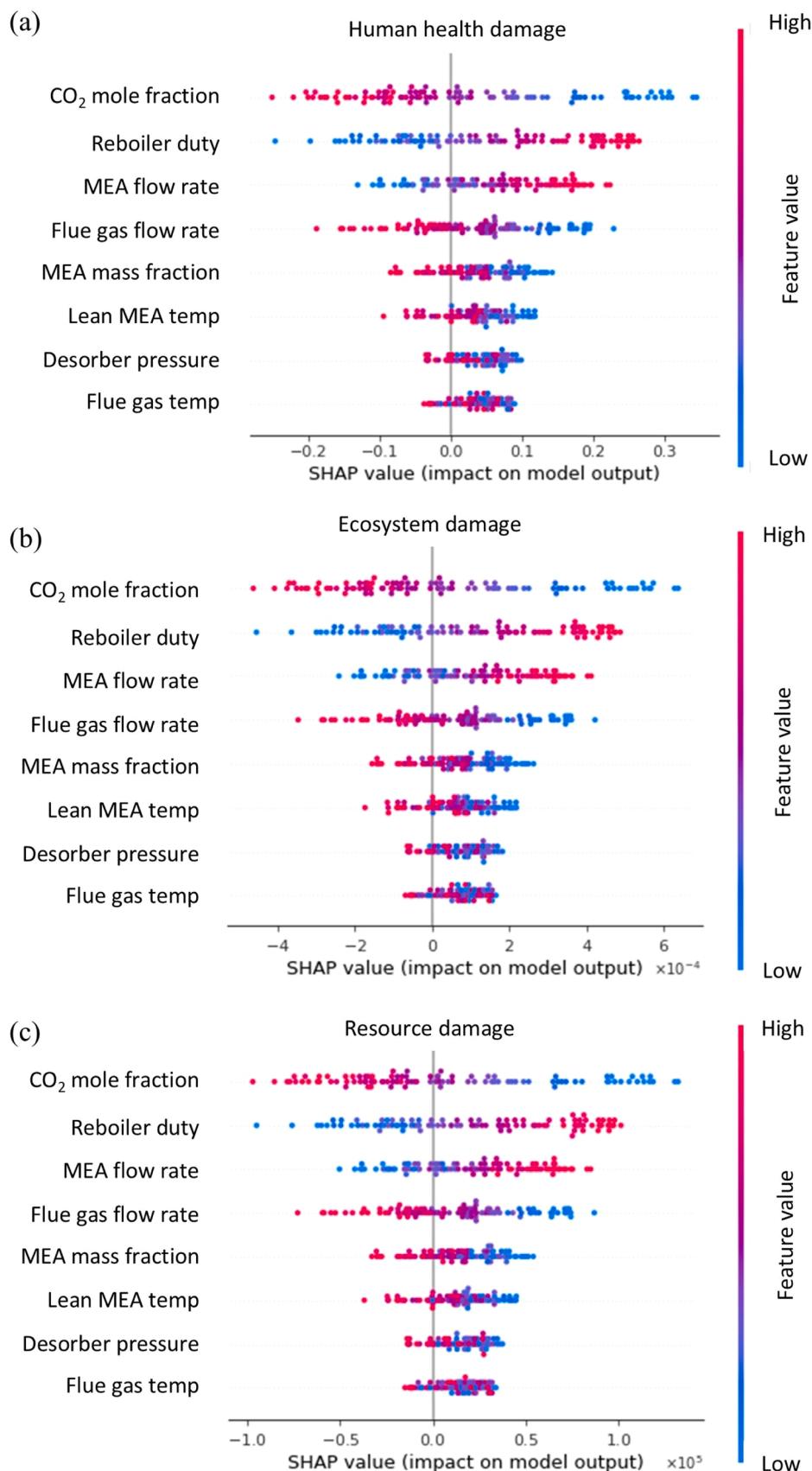


Fig. 8. SHAP summary results for DSM, demonstrating the importance of operating condition to the (a) human health damage, (b) ecosystem damage, and (c) resource scarcity.

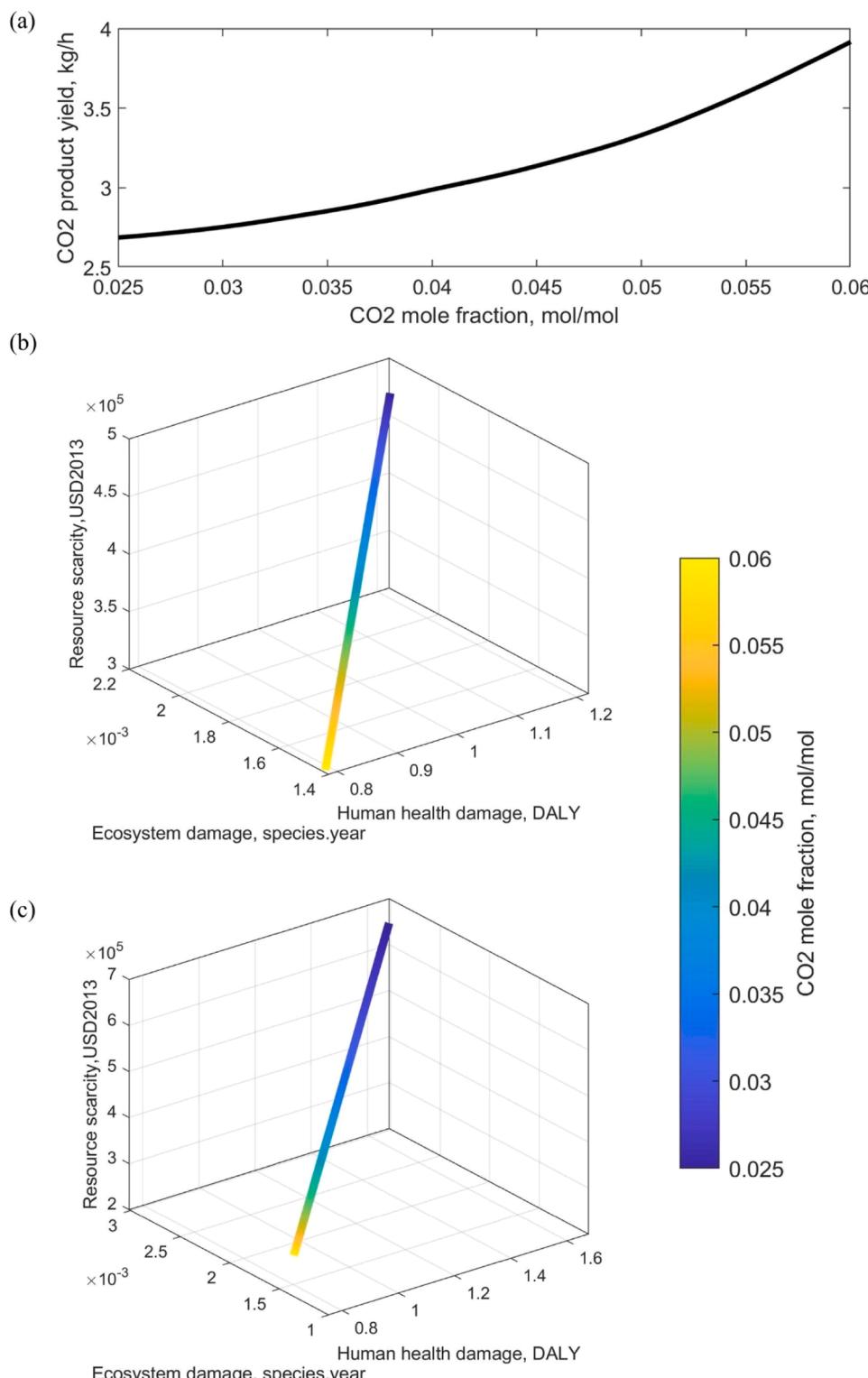


Fig. 9. Parametric study to explain the correlations of (a) – (b) CO₂ mole fraction–process outputs–environmental impacts in SSM, and (c) CO₂ mole fraction–environmental impacts in DSM.

105.7 kg/h and a desorber pressure of 204 kPa in carbon capture process as the benchmark, the minimum human health damage is achieved. Reducing the desorber pressure by 2 kPa, while maintaining the same solvent flow rate, leads to the lowest ecosystem damage. Increasing the solvent flow rate by 2 kg/h at a similar desorber pressure minimises damage to resource availability.

In DSM, the minimum human health damage values obtained were

0.03962 DALY with NSGA-II and 0.03972 DALY with PSO, as shown in Fig. 11(a). The lowest ecosystem damage values were 5.233×10^{-5} species.year for NSGA-II and 5.242×10^{-5} species.year for PSO. For resource scarcity, the minimum values were 2.924×10^4 USD, 2013 equivalent using NSGA-II and 2.930×10^5 USD, 2013 equivalent using PSO. In general, NSGA-II and PSO suggested a high solvent flow rate and flue gas flow rate and low flue gas temperature in carbon capture process

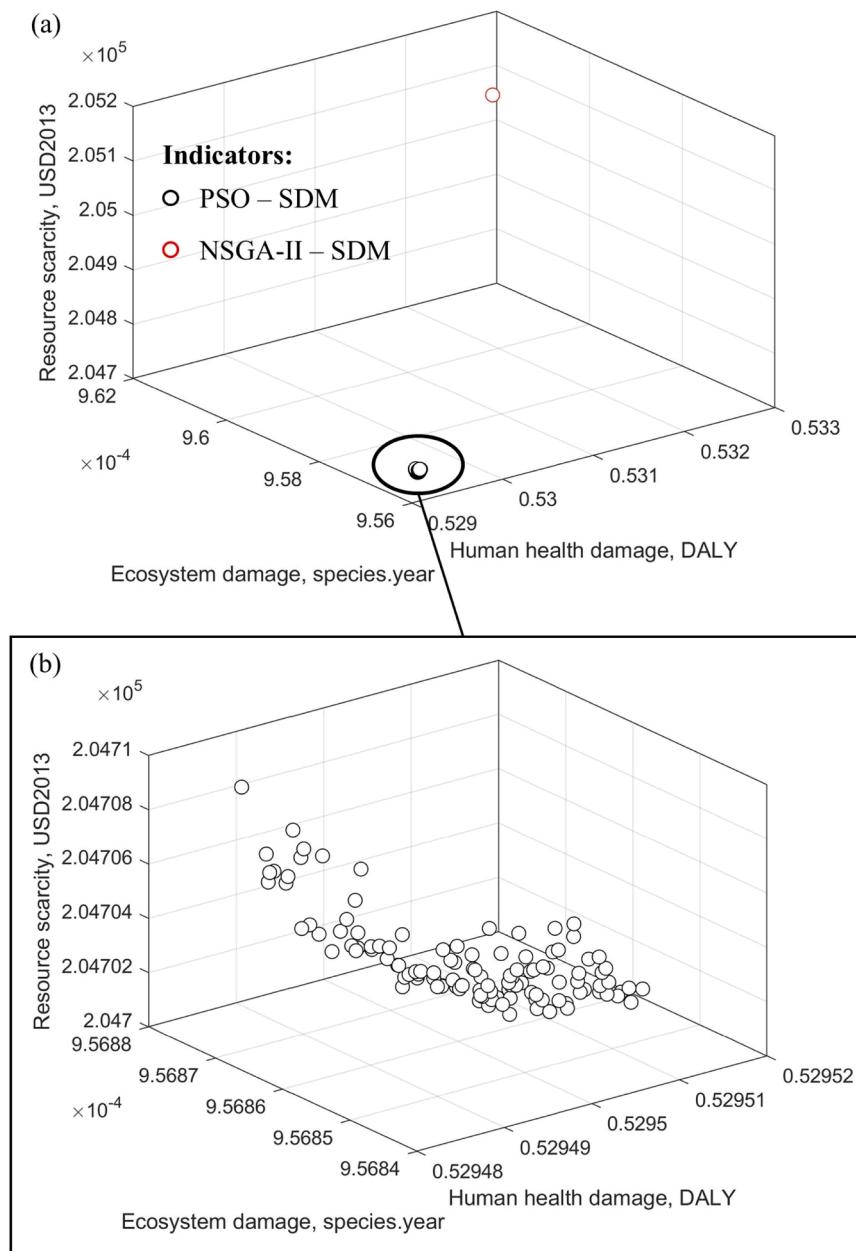


Fig. 10. Pareto front of SSM which converged by (a) NSGA-II and (b) PSO.

for the minimum environmental impacts. To achieve a minimum damage to the human health, NSGA-II and PSO suggested approximate 296 kg/h of solvent flow rate and 96 kg/h of flue gas flow rate and 23 °C of flue gas temperature. Using the operating conditions to achieve the minimum human health damage as the benchmark, increasing solvent flow rate, flue gas flow rate and its temperature by approximate ~26 kg/h, ~4 kg/h and ~5 °C, respectively can help to achieve lowest ecosystem damage. Similarly, reducing the solvent and flue gas flow rate by estimated ~8.5 kg/h and ~1.7 kg/h, respectively can help to minimise the damage to the resource. The operating conditions and the corresponding environmental impacts are summarised in Table S7.

The optimisation results of SSM and DSM for both NSGA-II and PSO in compared with the optimum experimental set are summarised in Table S8. From Notz et al.'s experimental results, the optimum parameters selected for high CO₂ removal efficiency and low energy consumption were 88 % and 5.97 GJ/t CO₂, respectively. The corresponding operating conditions were solvent flow rate of 200 kg/h, flue gas flow rate of 72.3 kg/h, desorber pressure of 200 kPa, MEA

fraction of 0.303 kg/kg, MEA input temperature of 40 °C, flue gas temperature of 47.88 °C, reboiler duty of 6.175 kW, and CO₂ concentration in the flue gas of 0.109 mol/mol. Under these conditions, the environmental impacts were quantified as 0.335 DALY of human health damage, 4.13×10^{-4} species. year of ecosystem damage, and 191,262 USD, 2013 equivalent of resource scarcity. The DSM approach using NSGA-II and PSO identified optimum parameters that reduced environmental impacts by 76–88 % compared with the experimental optimum, whereas the SSM approach yielded higher environmental impacts than the environmental optimum. The accuracy of SSM can be further enhanced by incorporating additional training data, further strengthening its predictive capability. This underscores the critical role of AI-driven multi-objective optimisation frameworks, such as the one developed in this study, in systematically navigating complex trade-offs between performance criteria, enabling the identification of solutions with substantially improved environmental outcomes, and expediting the development of sustainable process designs beyond the capabilities of experimental approaches alone.

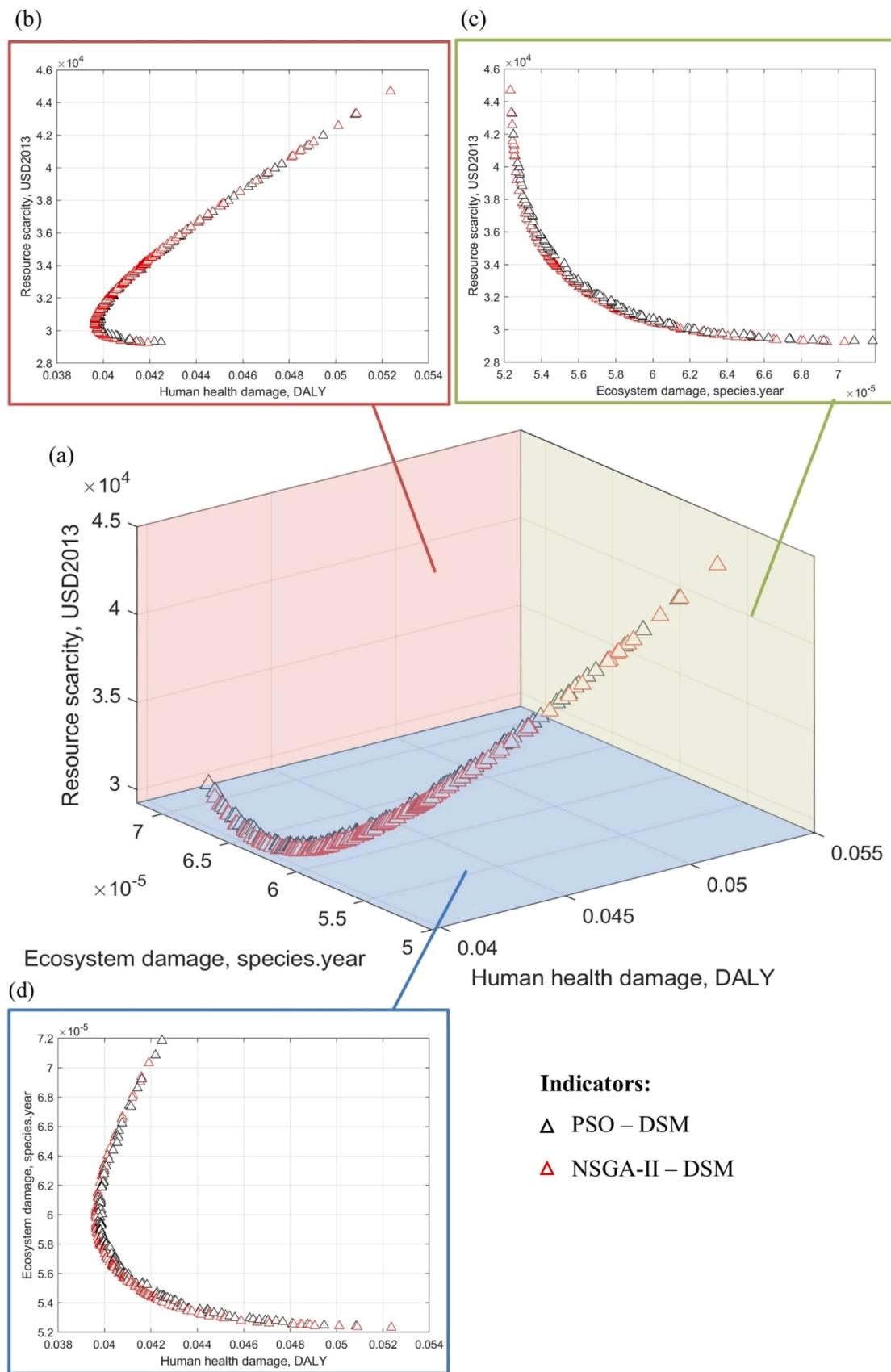


Fig. 11. (a) Pareto front that obtained by NSGA-II and PSO in DSM, view in (b) Human health damage – Resource scarcity, (c) Ecosystem damage – Resource scarcity and (d) Human health damage – ecosystem damage.

4. Conclusion

This work presents a minimisation of environmental impacts through optimising operating conditions by integrating LCA in an XAI optimisation framework in the case study of MEA-based carbon capture process. MEA-based carbon capture process is chosen as the case study due to its effectiveness in reducing CO₂ emissions, yet its broader environmental impacts remain underexplored. Two configurations of data-driven surrogate model with sequential and direct approaches, i.e. SSM and DSM were developed to compare the prediction accuracy and explainability. An experiment-validated physics-based model with <10 % uncertainty was employed to generate the database, addressing the limitation of insufficient experimental data. Based on the results, although the DSM shows a lower MSE than the SSM, both models are considered relatively accurate in their predictions, with the DSM demonstrating better overall accuracy. This is shown by its lower mean squared errors of 1.627×10^{-4} for human health, 1.631×10^{-4} for the ecosystem, and 1.625×10^{-4} for resources, compared to SSM's two-step prediction approach with values of 5.356×10^{-4} , 5.369×10^{-4} , and 5.355×10^{-4} respectively. Additionally, DSM achieves a higher coefficient of determination, with an R-squared value of 0.995 compared to 0.984 for SSM. To address the inherent black-box nature of the data-driven surrogate model, SHAP analysis was conducted to enhance its explainability, thereby supporting its adoption in real-world applications. SSM demonstrated better explainability than DSM. This indicates that CO₂ product yield is the most important process output influencing human health damage, ecosystem damage, and resource scarcity. Further analysis of SSM revealed that CO₂ mole fraction is the most significant operating condition affecting CO₂ product yield. Similarly, in DSM, CO₂ mole fraction emerged as the most significant operating condition influencing environmental impacts. Although DSM demonstrates excellent predictive ability, SSM offers better explainability. This is attributed to the intermediate information provided by SSM, which offers a better understanding of the process performance while accounting for the associated environmental impacts.

In this study, NSGA-II and PSO were adopted to minimise the damage to human health, ecosystem and resources in MEA-based carbon capture process and identifying the corresponding operating conditions. At the endpoint level, human health, ecosystem quality, and resource availability not only provide a consolidated summary of total environmental damage but also align with the current capability of optimisation frameworks, which are generally limited to handling up to three objectives effectively. In SSM, PSO is capable of identifying multiple Pareto front solutions, whereas NSGA-II can only find a single set, due to PSO's advantage in avoiding local minima. In DSM, both NSGA-II and PSO can identify a set of Pareto front solutions and with a clear trade-off between objective functions. NSGA-II converged to solutions with lower environmental impacts than PSO, owing to its use of crowding distance, which enables a more thorough exploration of the search space. This study also shows that our framework, particularly DSM identified optimum parameters that reduced environmental impacts by 76–88 % compared with the experimental optimum.

This study presents a novel framework of integrating LCA with advanced data-driven approaches to enable sustainability-oriented decision-making. This framework offers a streamlined approach to integrating sustainability insights during process design, enabling early-stage identification of environmentally optimal operating conditions without the need for computationally intensive simulations. In future work, additional data can be incorporated to improve the accuracy of the SSM, which is something that can be addressed in real-life applications where abundant process data can be used to enhance model accuracy. Besides, we aim to incorporate real process data to strengthen the validation of the framework and ensure that results can be readily compared against established benchmarks.

CRediT authorship contribution statement

Xin Yee Tai: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Oliver Fisher:** Writing – review & editing, Software, Investigation. **Lei Xing:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Jin Xuan:** Writing – review & editing, Supervision, Resources, Investigation, Funding acquisition, Conceptualization.

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

The authors acknowledge the financial support from EPSRC via grant numbers EP/W018969/2 and EP/W03722X/1 and the Leverhulme Trust via grant number PLP-2022-001.

Data availability

Data will be made available on request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dche.2025.100265.

References

- Ahmed, F., et al., 2022. The environmental impact of industrialization and foreign direct investment: empirical evidence from Asia-Pacific region. Env. Sci. Pollut. Res. Int 29 (20), 29778. <https://doi.org/10.1007/S11356-021-17560-W>. Available at.
- Ashraf, W.M., Dua, V., 2023. Machine learning based modelling and optimization of post-combustion carbon capture process using MEA supporting carbon neutrality. Digit. Chem. Eng. 8, 100115. <https://doi.org/10.1016/J.DCHE.2023.100115>. Available at.
- Aspen Tech, 2013. Aspen. HYSYS. |. Lead. Process. Simul. Softw. Oil. Gas. |Available at <https://www.aspentech.com/en/products/engineering/aspen-hysys>. Accessed: 15 August 2025.
- Cavalliere, C., et al., 2019. Consistent BIM-led LCA during the entire building design process. IOP. Conf. Ser.: Earth. Environ. Sci. 323 (1), 012099. <https://doi.org/10.1088/1755-1315/323/1/012099>. Available at.
- Chung, W., Lee, J.H., 2020. Input-output surrogate models for efficient economic evaluation of amine scrubbing CO2Capture processes. Ind. Eng. Chem. Res. 59 (42), 18951–18964. https://doi.org/10.1021/ACS.IECR.0C02971/ASSET/IMAGES/LARGE/IECR02971_0005.JPG. Available at.
- Clayton, A.D., et al., 2019. Algorithms for the self-optimisation of chemical reactions. React. Chem. Eng. 4 (9), 1545–1554. <https://doi.org/10.1039/C9RE00209J>. Available at.
- Diessner, M., et al., 2022. Investigating bayesian optimization for expensive-to-evaluate black box functions: application in fluid dynamics. Front. Appl. Math. Stat. 8, 1076296. <https://doi.org/10.3389/FAMS.2022.1076296>; BIBTEX. Available at.
- Ding, H., et al., 2025. Multi-objective optimization of supersonic separator for gas removal and carbon capture using three-field two-phase flow model and non-dominated sorting genetic algorithm-II (NSGA-II). Sep. Purif. Technol. 358, 130363. <https://doi.org/10.1016/J.SEPPUR.2024.130363>. Available at.
- Dirza, R., Krishnamoorthy, D., Skogestad, S., 2022. Primal-dual feedback-optimizing control with direct constraint control. Comput. Aided. Chem. Eng. 49, 1153–1158. <https://doi.org/10.1016/B978-0-323-85159-6.50192-5>. Available at.
- ecoinvent, 2023. ecoinvent v3.10 - ecoinvent. Available at <https://ecoinvent.org/ecoinvent-v3-10/>. Accessed: 9 June 2025.
- European Environmental Bureau, 2021. New EU environmental standards for chemical industry fall short of Green Deal ambition. Available at <https://eeb.org/new-eu-environmental-standards-for-chemical-industry-fall-short-of-green-deal-ambition/>. Accessed: 3 June 2025.
- Haupt, J., et al., 2023. Challenges of prospective life cycle assessment of emerging recycling processes: case study of battery materials recovery. Procedia. CIRP 116, 23–28. <https://doi.org/10.1016/J.PROCIR.2023.02.005>. Available at.
- Huijbregts, M.A.J., Steimann, Z.J.N., Elshout, P.M.F., et al., 2017. ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. Int. J. Life. Cycle. Assess. 22, 138–147. <https://doi.org/10.1007/s11367-016-1246-y>.

- IEA, 2021. Executive Summary – Ammonia Technology Roadmap – Analysis - IEA. Available at <https://www.iea.org/reports/ammonia-technology-roadmap/executive-summary>. Accessed: 15 May 2025.
- ISO, 2006. ISO 14040:2006 - environmental management — Life cycle assessment — Principles and framework. Available at <https://www.iso.org/standard/37456.html>. Accessed: 8 June 2025.
- Kaab, A., et al., 2019. Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. *Sci. Total. Environ.* 664, 1005–1019. <https://doi.org/10.1016/J.SCITOTENV.2019.02.004>. Available at.
- Li, K., et al., 2016. Systematic study of aqueous monoethanolamine-based CO₂ capture process: model development and process improvement. *Energy. Sci. Eng.* 4 (1), 23–39. <https://doi.org/10.1002/ESE3.101>. Available at.
- Mayol, A.P., et al., 2020. Environmental impact prediction of microalgae to biofuels chains using artificial intelligence: a life cycle perspective. *IOP. Conf. Ser.: Earth. Environ. Sci.* 463 (1), 012011. <https://doi.org/10.1088/1755-1315/463/1/012011>. Available at.
- Miao, Y., et al., 2022. Integrated artificial intelligence and life cycle assessment to predict environmental and economic impacts for natural gas hybrid electric vehicle with multi-dimension. *J. Phys.: Conf. Ser.* 2198 (1), 012016. <https://doi.org/10.1088/1742-6596/2198/1/012016>. Available at.
- Notz, R., et al., 2012. Post combustion CO₂ capture by reactive absorption: pilot plant description and results of systematic studies with MEA. *Int. J. Greenh. Gas. Control* 6, 84–112. <https://doi.org/10.1016/J.IJGGC.2011.11.004>. Available at.
- Oh, D.H., et al., 2022. Prediction of CO₂ capture capability of 0.5 MW MEA demo plant using three different deep learning pipelines. *Fuel* 315, 123229. <https://doi.org/10.1016/J.FUEL.2022.123229>. Available at.
- Öi, L.E., 2012. Comparison of Aspen HYSYS and Aspen Plus simulation of CO₂ absorption into MEA from atmospheric gas. *Energy. Procedia* 23, 360–369. <https://doi.org/10.1016/J.EGYPRO.2012.06.036>. Available at.
- Printweek, 2025. Available at <https://www.printweek.com/content/news/chemicals-firm-fined-25m-after-serious-breaches/>. Accessed: 10 August 2025.
- Qamar, R.A., et al., 2020. Aspen HYSYS simulation of CO₂ capture for the best amine solvent. *J. Adv. Res. Fluid. Mech. Therm. Sci.* 68 (2), 124–144. <https://doi.org/10.37934/ARFMTS.68.2.124144>. Available at.
- Ragab, A., et al., 2018. Fault diagnosis in industrial chemical processes using interpretable patterns based on logical analysis of data. *Expert. Syst. Appl* 95, 368–383. <https://doi.org/10.1016/J.ESWA.2017.11.045>. Available at.
- Reuters, 2008. Available at <https://www.reuters.com/article/business/bps-texas-city-compensation-bill-tops-2-billion-idUSL04900882/>. Accessed: 10 August 2025.
- Salimi, M., et al., 2024. Techno-economic study of different process configurations of CO₂ capture by MEA absorption and utilization of an MCDM method for their ranking. *Energy. Rep.* 12, 593–607. <https://doi.org/10.1016/J.EGYR.2024.06.043>. Available at.
- SHAP, 2018. Welcome to the SHAP documentation — SHAP latest documentation. Available at <https://shap.readthedocs.io/en/latest/>.
- Statista, 2025. Available at <https://www.statista.com/statistics/1091999/atmospheric-c-concentration-of-co2-historic/>. Accessed: 4 June 2025.
- Tai, X.Y., et al., 2022. Multi-objective optimisation with hybrid machine learning strategy for complex catalytic processes. *Energy. AI* 7, 100134. <https://doi.org/10.1016/J.EGYAI.2021.100134>. Available at.
- Tai, X.Y., Xing, L., Christie, S.D.R., et al., 2023a. Deep learning design of functionally graded porous electrode of proton exchange membrane fuel cells. *Energy*, 128463. <https://doi.org/10.1016/J.ENERGY.2023.128463>. Available at.
- Tai, X.Y., Xing, L., Zhang, Y., et al., 2023b. Dynamic optimisation of CO₂ electrochemical reduction processes driven by intermittent renewable energy: hybrid deep learning approach. *Digit. Chem. Eng.*, 100123 <https://doi.org/10.1016/J.DCHE.2023.100123>. Available at.
- The Ferret, 2018. Shell fined £40,000 for breaching climate pollution rules at Fife plant. Available at <https://theferret.scot/shell-fined-40000-climate-pollution-rules-fife/>. Accessed: 10 August 2025.
- The Ferret, 2022. Drinks giant Diageo fined £1.2m for climate breaches. Available at <https://theferret.scot/diageo-fined-climate-breaches/>. Accessed: 10 August 2025.
- UN Environment Programme, 2011. Decoupling: natural resource use and environmental impacts from economic growth. Available at <https://www.unep.org/resources/report/decoupling-natural-resource-use-and-environmental-impacts-economic-growth>. Accessed: 4 June 2025.
- Wang, J., et al., 2022. Comparison of state-of-the-art machine learning algorithms and data-driven optimization methods for mitigating nitrogen crossover in PEM fuel cells. *Chem. Eng. J.* 442, 136064. <https://doi.org/10.1016/J.CEJ.2022.136064>. Available at.
- Wowra, K., et al., 2023. Estimating environmental impacts of early-stage bioprocesses. *Trends. Biotechnol.* 41 (9), 1199–1212. <https://doi.org/10.1016/J.TIBTECH.2023.03.011>. Available at.
- Xi, H., et al., 2021. Simultaneous parametric optimization for design and operation of solvent-based post-combustion carbon capture using particle swarm optimization. *Appl. Therm. Eng.* 184, 116287. <https://doi.org/10.1016/J.APPLTHERMALENG.2020.116287>. Available at.
- Zhang, Y., et al., 2009. Rate-based process modeling study of CO₂ Capture with aqueous monoethanolamine solution. *Ind. Eng. Chem. Res.* 48 (20), 9233–9246. <https://doi.org/10.1021/IE900068K>; JOURNAL:JOURNAL:IECRED;PAGE:STRING:ARTICLE/CHAPTER. Available at.