Journal of CO2 Utilization 106 (2026) 103402

Contents lists available at ScienceDirect

Journal of CO2 Utilization

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

Integrated design and optimization for a unified carbon capture and storage system using a machine-learning-assisted multi-objective optimization framework

Suin Choi [a,b], Minkyung Chae [a], Soeun Yoon [a,c], Tea-Woo Kim [d,*], Suryeom Jo [e], Byungin Ian Choi [f], Honggeun Jo [g], Baehyun Min [a,*]

[a] Department of Climate and Energy Systems Engineering, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul, 03760, Republic of Korea
[b] Hildebrand Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, 200 E Dean Keeton St, Austin, TX, 78712, the United States of America
[c] POSCO International, 165, Convensia-daero, Yeonsu-gu, Incheon, 21998, Republic of Korea
[d] CO2 Geological Storage Research Center, Climate Change Response Division, Korea Institute of Geoscience and Mineral Resources, 124, Gwahak-ro, Yuseong-gu, Daejeon, 34132, Republic of Korea
[e] Mineral Exploration and Mining Research Center, Mineral Resources Division, Korea Institute of Geoscience and Mineral Resources, 124, Gwahak-ro, Yuseong-gu, Daejeon, 34132, Republic of Korea
[f] Computer Modelling Group Ltd. (CMG), 3710 33 St NW, Calgary, Alberta, T2L 2M1 Canada
[g] Department of Energy Resources Engineering, Inha University, Incheon, 22212, Republic of Korea

ARTICLE INFO

Keywords:
Unified CCS system
Dynamic nodal analysis
Multi-objective optimization
Machine-learning
Proxy

ABSTRACT

Efficient deployment of carbon capture and storage (CCS) systems requires the integrated design of surface facilities and subsurface storage reservoirs. Herein, we propose a unified CCS system design framework in which dynamic nodal analysis, multi-objective optimization, and machine-learning are combined to support efficient and scalable decision-making. Optimization is focused on four key design variables: the CO_2 discharge pressure at the hub terminal, the inner diameters of the pipeline and injection tubing, and the injection temperature, with storage capacity, injection safety, and economic feasibility as the performance objectives. Dynamic nodal analysis couples surface and subsurface flow physics—specifically, compressible pipe flow dynamics for surface transport and Darcy-based porous media flow for subsurface storage—to simulate transient system behavior. A multi-objective particle swarm optimization algorithm linked to a full-physics facility-reservoir model was used to derive an initial Pareto-optimal front covering approximately 3.8% of the feasible design space. This was refined and expanded by using machine-learning-based proxies, including neural network, random forest, and boosting techniques trained on the optimization results. The framework was applied to a potential CCS assessment project in the Republic of Korea, an onshore CO_2 terminal, a 175-km subsea pipeline, and a 1-km vertical injection well. The machine-learning models demonstrated high predictive performance ($R^2 > 0.96$; Mean Absolute Percent Error ≦ 5%) and significantly reduced computational time compared to a full-physics simulation. This hybrid framework offers a practical and accurate approach for early-stage CCS system design and lays the foundation for real-world deployment under complex multi-objective constraints.

1. Introduction

Carbon capture and storage (CCS) is a pivotal technology for mitigating industrial CO_2 emissions and achieving global climate change targets [1,2]. It encompasses the integrated operation of CO_2 capture, transport, and geological storage processes that collectively form a unified CCS system [3,4]. As global decarbonization efforts accelerate, CCS is becoming increasingly regarded as a viable and scalable solution. However, successful deployment requires system-level design strategies that not only maximize CO_2 injection capacity but also ensure operational safety and economic viability. Although prior research has been conducted to address individual segments of the CCS chain,

* Corresponding authors.
E-mail addresses: tw_kim@kigam.re.kr (T.-W. Kim), bhmin01@ewha.ac.kr (B. Min).

https://doi.org/10.1016/j.jcou.2026.103402
Received 14 October 2025; Received in revised form 20 February 2026; Accepted 18 March 2026
Available online 24 March 2026
2212-9820/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

interdependencies among the components—such as the influence of capture conditions on the pipeline pressure or the impact of storage reservoir constraints on the injection strategy—highlight the need for integrated optimization approaches [5]. Despite growing interest, comprehensive design methodologies that account for these cross-process interactions remain underdeveloped [6–10].

Implementing a unified CCS system design is inherently challenging due to the distinct physical principles and the governing equations applied to the various system components. CO_2 transport through pipelines and wellbores is primarily governed by compressible fluid dynamics, whereas CO_2 migration within the pore network of deep saline formations can be described using Darcy's law for porous media flow. Although the fluid properties and flow behaviors remain continuous across the surface and subsurface domains, the fundamental contrast in their governing physics principles makes their integration within a single simulation framework complicated.

Nodal analysis offers a viable solution by connecting surface and subsurface flow regimes through Inflow Performance Relationships (IPR) and Outflow Performance Relationships (OPR), which assists in consistently identifying system-wide operating conditions at interface points such as the wellhead and bottomhole [11–13]. Traditionally, nodal analysis has been employed in a static form using segmented modeling tools, which often requires extensive external data exchange between simulation time steps [14,15]. In contrast, dynamic nodal analysis facilitates the transient evaluation of coupled pressure and temperature evolution by integrating surface facility performance with subsurface flow behavior [16,17]. Although dynamic nodal analysis has recently gained attention in geoenergy research as a promising method for integrated system evaluation, its practical application in the CCS context remains limited and underexplored [18–22].

Effective CCS design must account for non-linear thermophysical behavior of CO_2 in the vicinity of its critical point (7.4 MPa, 31 °C). In this near-critical regime, small perturbations in pressure and temperature would induce abrupt variations in density and viscosity. These might lead to flow instabilities, phase segregation, and pronounced pressure transients, thereby undermining injectivity control and system safety [23,24]. A practical risk-mitigation strategy is to maintain CO_2 in a stable dense phase along the entire transport-injection pathway. Consequently, operating at pressures sufficiently above the critical pressure reduces exposure to near-critical conditions and supports predictable, robust hydraulic performance.

Designing a CCS system inherently involves reconciling multiple conflicting objectives—such as maximizing injectivity, maintaining safe pressure limits, and ensuring economic feasibility—thus making it a multi-objective optimization (MOO) problem. In practice, the operational aim is to maximize CO_2 storage capacity to support emission reduction goals, minimize pressure buildup to maintain operational safety, and improve economic indicators to ensure project viability. These competing objectives require unavoidable trade-offs, thereby necessitating systematic optimization. Based on the principle of non-dominance, MOO provides a structured methodology to identify the Pareto-optimal front (POF), where no objective can be improved without sacrificing another [25,26]. However, evaluating each design candidate solution typically requires computationally intensive full-physics simulations, in which the simulation load increases rapidly with the number of objectives and design variables [27,28]. As the number of objectives is increased, the resulting computational burden creates a bottleneck that impedes the search performance and limits the resolution of POF exploration, especially in complex systems with nonlinear behavior. While classical MOO studies often rely on standard non-dominated sorting, its computational advantage is limited when the number of solutions is relatively small (typically a few hundred). In this study, the integration of machine learning (ML) proxies with MOO expands the design space to more than hundreds of thousands of solutions, where standard non-dominated sorting becomes computationally prohibitive. To overcome this bottleneck, a divide-and-conquer-based non-dominated sorting algorithm is adopted, thereby enabling efficient POF identification in the ML refinement stage. Such advanced algorithms such as fast non-dominated-sorting and divide-and-conquer strategies facilitate the exploration of high-dimensional design spaces and delineation of the POF [29–32]. Incorporating decision-makers' preferences or constraints on the objectives further enhances efficiency by narrowing the design space to relevant trade-off regions [27,28,33,34,35]. Although MOO has been successfully applied to various components of the CCS chain—including CO_2 capture [8], transport planning [9], supply chain design [6], and injection/storage optimization [10,36,37]—its full potential has so far remained underutilized in unified CCS system design, particularly when tightly coupled with integrated flow simulation models. To complement the computational burden of full-physics simulations in MOO, ML has emerged as a promising and effective tool in geoenergy system design. ML-based surrogate models—or proxies—can approximate complex system behavior with high accuracy while drastically reducing the computation time, thereby enabling rapid evaluation of a significantly larger number of design alternatives [38–41]. In CCS applications, ML techniques have been successfully employed for site screening, injection optimization, and storage capacity estimation, often achieving prediction errors below 5% and reducing the computational cost by more than 80% [42–45]. Beyond acceleration, ML enables high-resolution refinement of the POF by supporting rapid exploration of refined, potentially continuous design variable spaces derived from initial MOO results. Moreover, ensemble learning approaches—in which neural networks, random forests, and boosting algorithms are combined—can further enhance predictive performance and model robustness by reducing variance and mitigating bias [46–50]. When these ML-based proxies are used as surrogates inside the optimization loop, the resulting framework is surrogate-based MOO, which reduces the number of expensive high-fidelity evaluations and expedites the search [51–54]. Nevertheless, ML-based proxies must be carefully integrated into the optimization workflow, and all final solutions should be validated against full-physics models to ensure physical reliability and operational feasibility [55].

To overcome the limitations of conventional CCS system optimization approaches, we herein propose a hybrid framework that integrates dynamic nodal analysis, MOO, and ML to support the design of a unified CCS system. Dynamic nodal analysis, serving as the foundation of the modeling framework, captures the transiently coupled behavior of surface facilities and subsurface reservoirs. MOO is used to construct an initial POF that balances CO_2 storage capacity, injection safety, and economic feasibility based on full-physics simulations. To alleviate the computational burden and enhance solution diversity, ML-based proxies were trained on the MOO outcomes and subsequently used to rapidly explore and refine the POF across a broader design space. This hybrid framework enables scalable, comprehensive optimization while significantly reducing the computational cost without compromising accuracy.

2. Methodology

We developed a hybrid framework in which dynamic nodal analysis, MOO, and ML are integrated to support the design of a unified CCS system. The methodology consists of three main components: (1) modeling of the unified CCS system as a multi-objective problem, (2) Pareto-optimal solution searching using MOO, and (3) refinement of the solution space using ML-based proxies.

2.1. Modeling of the unified CCS system as a multi-objective problem

The design of a unified CCS system is formulated as an MOO problem defined over a set of N_{var} design variables and N_{obj} objective functions as follows:

$$f(\mathbf{x}) = f(x_1, x_2, \cdots, x_{N_{var}}) = \left\{ f_1(\mathbf{x}), f_2(\mathbf{x}), \cdots, f_{N_{obj}}(\mathbf{x}) \right\} \quad (1)$$

where $\mathbf{x}$ represents the vector of design variables and $f(\mathbf{x})$ denotes the corresponding set of objective functions to be optimized.

The unified CCS model comprises the full CCS value chain from source to sink—CO_2 transport, injection, and geological storage—with corresponding design variables and performance objectives. As illustrated in Fig. 1, the model in this study is defined using four design variables and three objectives (i.e., $N_{var} = 4$ and $N_{obj} = 3$). This MOO problem can be formulated as a minimization problem as follows:

$$f(\mathbf{x}) = f(P_{hub}, d_{pipe}, d_{tubing}, T_{wh}) = \left\{\frac{1}{M_{co_2}}, \text{BHP}, \frac{1}{\text{BCR}}\right\} \quad (2)$$

under constraints that $M_{co_2} \geq M_{CO_2}^{min}$,

$\text{BHP} \leq \text{BHP}^{max}$, and $\text{BCR} \geq \text{BCR}^{min}$,where P_{hub} is the CO_2 discharge pressure at the hub terminal (i.e., the source), d_{pipe} is the inner diameter of a subsea CO_2 transport pipeline, d_{tubing} is the inner diameter of the injection well tubing, and T_{wh} is the injection temperature at the wellhead.

Although NPV is widely regarded as a primary economic indicator for investment decisions in geoenergy projects [56–58], it focuses on absolute monetary gain. In this study, BCR is employed as an economic objective because it highlights the relative return per unit cost, which is a critical metric in CCS projects where public funding and policy incentives play important roles [59–62]. BCR also offers the advantage of being readily normalized, similar to the other two objectives, thereby facilitating balanced MOO. We clarify that the NPV values for all scenarios are calculated but are not disclosed due to data confidentiality.

The selected design variables each play a critical role in system behavior and performance: P_{hub} affects CO_2 transport efficiency and associated energy consumption; d_{pipe} and d_{tubing} influence flow assurance by controlling flow rate and pressure losses, and also impact infrastructure cost; and T_{wh} governs both injectivity and the thermodynamic behavior of CO_2 in the geological formation (i.e., the sink). These four design variables can be used collectively to determine the system's performance, which is quantified by using three objective functions: the cumulative mass of injected CO_2 (M_{CO_2}), the bottomhole pressure (BHP), and the benefit-cost ratio (BCR). These objectives capture the core operational aspects of a CCS system: injectivity, safety, and economic feasibility. In this formulation, $M_{CO_2}^{min}$ represents the minimum total CO_2 mass required over the entire injection period to meet the emission reduction target, while the total injection mass over a year is assumed to be uniformly distributed. The BHP limit (BHP^{max}) represents the maximum allowable BHP required for safe CCS operation relative to the estimated fracture pressure. As the estimated fracture pressure cannot be disclosed due to data confidentiality, BHP serves as an indirect yet practical indicator of injection safety: the higher the BHP, the sooner the allowable limit is reached. For this reason, BHP was treated as an objective function rather than as a design parameter, complementing storage performance and economic feasibility in the optimization framework. This choice reflects that the CO_2 discharge pressure at the hub terminal (P_{hub}) is already defined as a design variable, and wellhead and bottomhole pressures are automatically determined from P_{hub} without independent control. Meanwhile, the minimum BCR threshold (BCR^{min}) represents the lowest value required to ensure the economic viability of the CCS project for attaining sufficient CO_2 injection and storage.

The values of objective functions M_{co_2}, BHP, and BCR are computed based on the time-dependent evolution of flowrate Q and pressure P throughout the injection period. Specifically, $M_{co_2} = \int Q \rho_{co_2} dt$, where ρ_{co_2} denotes the CO_2 density. In this study, dynamic nodal analysis is explicitly coupled with MOO, such that the unified CCS model calculates the values of Q and P discretely. This extends time-independent nodal analysis by incorporating temporal updates, thereby enabling a more comprehensive evaluation of injection well performance over time. This approach captures spatiotemporal variations in both reservoir properties and surface facility conditions [63,64].

Fig. 2 illustrates the application of dynamic nodal analysis to compute Q and P for a single-well CO_2 injection system while using the bottomhole as the analysis node. In Fig. 2(a), "inflow" refers to the movement of CO_2 from the wellhead to the bottomhole and is governed

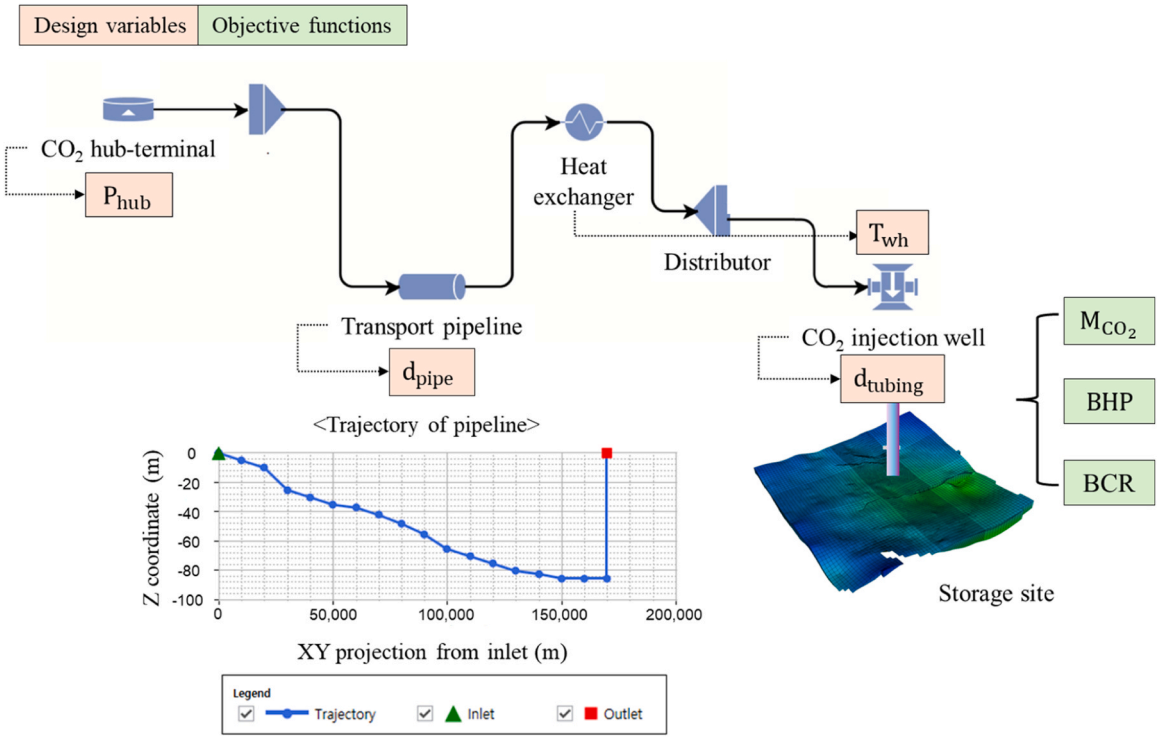


Fig. 1. A schematic of a unified carbon capture and storage system from the CO_2 hub terminal to the geological storage site via a CO_2 transport pipeline incorporating the design variables and objective functions considered in this study.

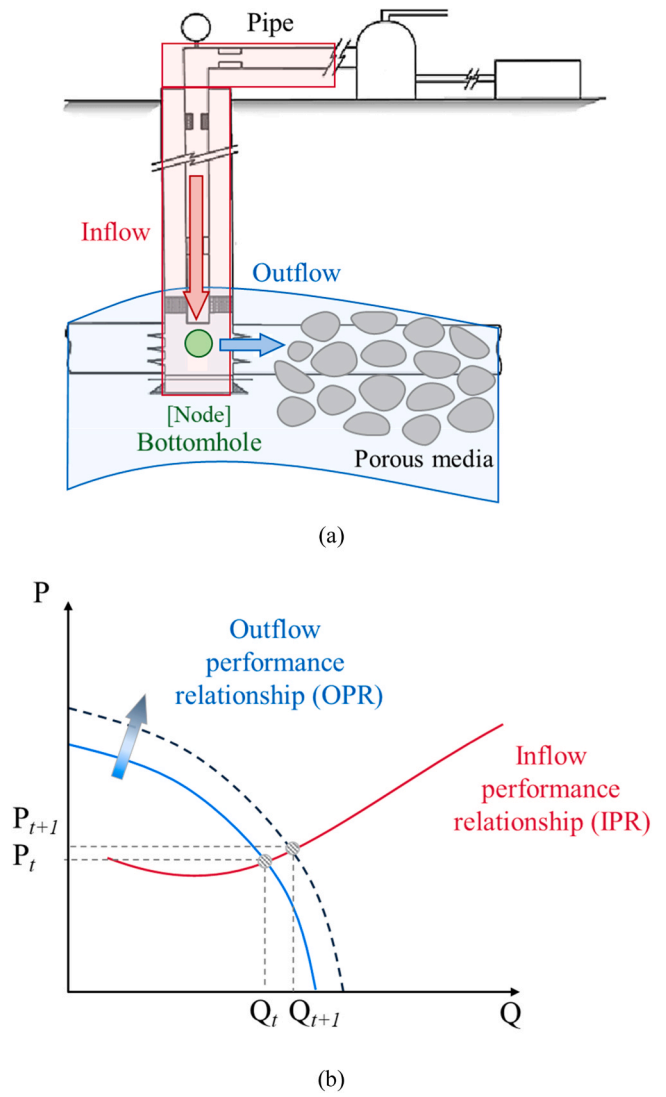


Fig. 2. Dynamic nodal analysis of the unified carbon capture and storage system: (a) the inflow and outflow of the injection system and (b) the dynamic inflow and outflow performance relationship. Inflow is primarily affected by surface facilities such as the transport pipeline and injection well, whereas outflow is governed by reservoir conditions. The intersection between the inflow performance relationship (IPR) and the outflow performance relationship (OPR) represents the stable flow condition at the node (bottomhole) and differs from time to time.

primarily by the compressible pipe flow mechanics. Meanwhile, "outflow" describes the flow of CO_2 from the bottomhole into the geological storage formation, governed by Darcy's law for porous media flow. The pipe flow and subsurface Darcy flow are implicitly coupled to calculate a consistent inflow-outflow balance [16]. The inflow characteristics are largely governed by surface facility parameters such as pipeline and tubing specifications, whereas the outflow behavior is influenced by reservoir properties, including formation pressure, porosity, and permeability. Fig. 2(b) demonstrates that the intersection between the IPR and OPR defines a unique operating point (Q, P) at each time step where both governing equations are simultaneously satisfied. As injection continues and storage reservoir pressure increases, the OPR curve progressively shifts outward. Given a fixed facility configuration that results in a fixed IPR curve, the IPR-OPR intersection varies over time. This dynamic nodal analysis serves as the foundation for the unified CCS model by enabling stepwise calculations of Q and P under evolving system conditions.

2.2. MOO

Conventional MOO algorithms operate iteratively by applying non-dominated sorting and crowding-distance sorting to rank the solution candidates. In each generation, the algorithm propagates the top-ranked solutions toward the POF, thereby evolving the solution set across successive iterations. In this study, the optimization was terminated after a predefined number of evaluations ($N_{gen}xN_{pop}$), where N_{gen} is the number of generations and N_{pop} is the population size per generation. While early stopping can be adopted if optimization stagnates, we employed the maximum number of evaluations as the termination criterion to ensure robust identification of the POF.

2.2.1. Non-dominated sorting and crowding-distance sorting

For each generation, non-dominated sorting is applied to classify N_{pop} solutions into distinct non-dominated ranks (also known as non-dominated fronts) based on the principle of Pareto dominance [65]. Eq. (3) represents the formal condition under which solution $\mathbf{x}^a$ is said to dominate solution $\mathbf{x}^b$ in a minimization problem with N_{obj} objectives:

$$\forall\, i \in \{1,\ \cdots,\ N_{obj}\}\ :\ f_i(\mathbf{x}^a) \le f_i(\mathbf{x}^b) \text{and} \exists\ j \in \{1,\ \cdots,\ N_{obj}\}\ :\ f_j(\mathbf{x}^a) < f_j(\mathbf{x}^b) \tag{3}$$

where f_i and f_j represent the i^{th} and j^{th} objective functions, respectively.

Fig. 3(a) provides a visual example of the dominance relationships among six candidate solutions ($\mathbf{x}^1$ to $\mathbf{x}^6$) in two-objective design space Ω. In this example, $\mathbf{x}^1$ dominates $\mathbf{x}^2$, which in turn dominates $\mathbf{x}^3$, and similarly, $\mathbf{x}^4$ dominates $\mathbf{x}^5$, which dominates $\mathbf{x}^6$. However, $\mathbf{x}^1$and $\mathbf{x}^4$ do not dominate one another because $f_i(\mathbf{x}^1) < f_i(\mathbf{x}^4)$ while $f_j(\mathbf{x}^1) > f_j(\mathbf{x}^4)$, and vice versa. As a result of iteratively applying Eq. (3), the solutions are grouped into three non-dominated ranks: Rank 1 includes $\mathbf{x}^1$and $\mathbf{x}^4$, Rank 2 includes $\mathbf{x}^2$and $\mathbf{x}^5$, and Rank 3 includes $\mathbf{x}^3$and $\mathbf{x}^6$. Solutions within the same non-dominated front are considered trade-offs of equal value in the absence of information required to set a preference.

Crowding-distance sorting is applied to preserve the diversity of trade-off solutions across generations. When the number of solutions in a non-dominated front exceeds population capacity N_{pop}, crowding-distance sorting is used to select a representative subset that preserves solution diversity in the next generation. The crowding-distance metric quantifies the local density of solutions in the objective space and is used to prioritize those in sparsely populated regions. As illustrated in Fig. 3 (b), boundary solutions (e.g., those at the extreme ends of the objective space) are assigned crowding distances of infinity to ensure their selection in the subsequent generation. For interior solutions, the crowding distance is calculated based on the spacing between adjacent solutions as follows:

$$d_k = \sum_{i=1}^{N_{obj}} \frac{f_i^{k+1} - f_i^{k-1}}{f_i^{max} - f_i^{min}} \tag{4}$$

where $f_i^{k+1}$and f_i^{k-1} denote the i^{th} objective function values of the two immediate neighbors surrounding the k^{th} solution and terms $f_i^{max}$and f_i^{min} represent the maximum and minimum values of the i^{th} objective across the population, respectively. This process assigns lower crowding distances to solutions in densely clustered areas while favoring the selection of solutions in sparsely populated regions to maintain diversity.

2.2.2. Divide-and-conquer-based non-dominated sorting

One limitation of conventional non-dominated sorting is its computational inefficiency when dealing with a large N_{pop} value due to it exhibiting a time complexity of $O(N_{obj}N_{pop}{}^2)$. To address this bottleneck, [32] developed a divide-and-conquer-based non--dominated-sorting algorithm with improved time complexity of $O(N_{pop}\log N_{pop} + N_{obj}N_{pop})$. This enhanced method adopts a divide-and-conquer strategy that decomposes the problem into smaller

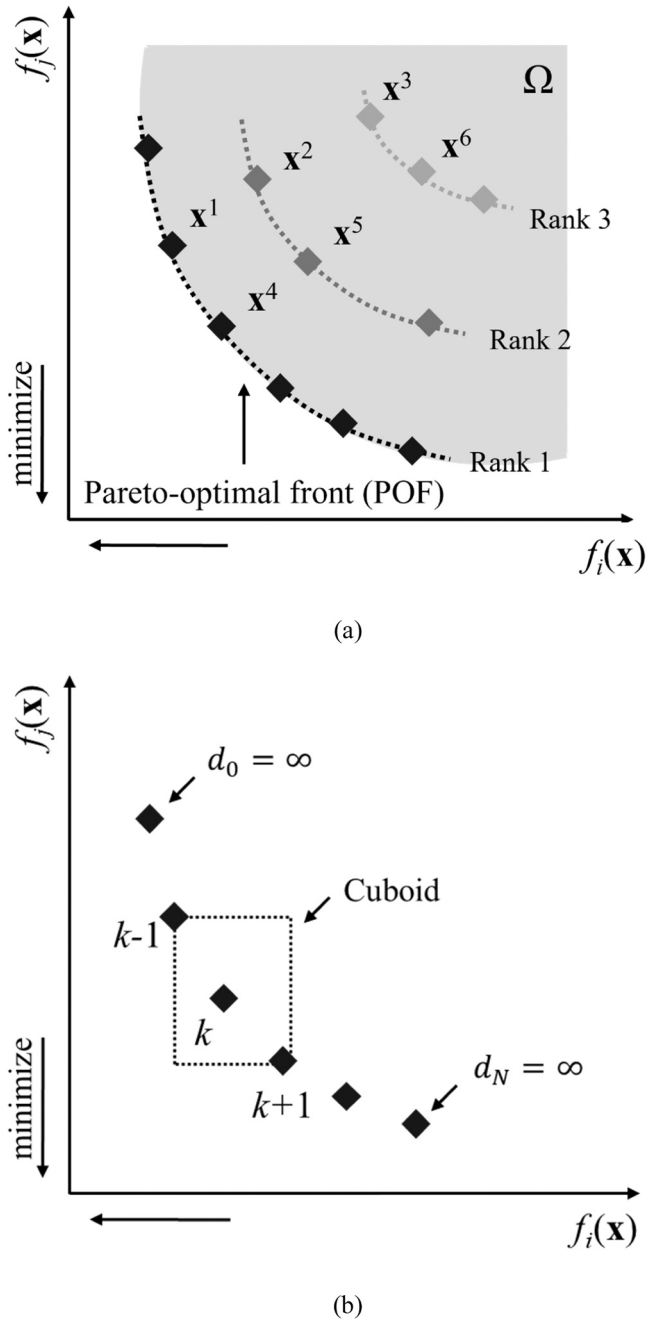


Fig. 3. A schematic showing the search for the Pareto-optimal front (POF) for a bi-objective minimization problem based on six candidate solutions: (a) non-dominated sorting and (b) crowding-distance sorting.

subproblems, solves each independently, and then merges the results to reconstruct the full non-dominated-sorting structure. By combining pre-sorting, hierarchical merging, and efficient insertion operations, the algorithm significantly reduces redundant dominance comparisons and enhances scalability for large-scale MOO tasks. Details of the pseudo-code for this algorithm are provided in [32].

2.2.3. Performance metrics for MOO

The performance of MOO can be evaluated using two metrics: the averaged Hausdorff distance ($\Delta_{\mathscr{P}^*}$) to assess the convergence of the approximated POF toward the true POF and the hypervolume per solution (HV^n) to quantify solution diversity along the front [66]. Prior to computing $\Delta_{\mathscr{P}^*}$ and HV^n, all objective function values are normalized to within the [0, 1] range to ensure comparability using the following relationship:

$$f_i(\mathbf{x})^n = \frac{f_i(\mathbf{x}) - f_i^{\min}}{f_i^{\max} - f_i^{\min}}.\# \tag{5}$$

Convergence metric $\Delta_{\mathscr{P}^*}$ is defined as the maximum of two distance-based measures: the generational distance ($GD_{\mathscr{P}^*}$) and the inverse generational distance ($IGD_{\mathscr{P}^*}$), as defined in Eq. (6) [67]. This combined metric quantifies the bidirectional distance between the approximated POF (O) and the true POF ($\mathscr{P}^*$).

$$\Delta_{\mathscr{P}^*} = \max\{GD_{\mathscr{P}^*}(O, \mathscr{P}^*), IGD_{\mathscr{P}^*}(O, \mathscr{P}^*)\} \tag{6}$$

The individual components of $\Delta_{\mathscr{P}^*}$ are defined as follows:

$$GD_{\mathscr{P}^*}(O, \mathscr{P}^*) = \left(\frac{1}{N_O}\sum_{i=1}^{N_O} \mathrm{dist}(o_i, \mathscr{P}^*)^2\right)^{\frac{1}{2}} \tag{7}$$

$$IGD_{\mathscr{P}^*}(O, \mathscr{P}^*) = \left(\frac{1}{N_{\mathscr{P}^*}}\sum_{i=1}^{N_{\mathscr{P}^*}} \mathrm{dist}(O, p_i)^2\right)^{\frac{1}{2}} \tag{8}$$

where O denotes the set of obtained non-dominated-sorting solutions (i.e., the approximated POF) and $\mathscr{P}^*$ represents the true POF ($o_i \in O$ and $p_i \in \mathscr{P}^*$); N_O and $N_{\mathscr{P}^*}$ denote the number of solutions in O and $\mathscr{P}^*$, respectively; $\mathrm{dist}(o_i, \mathscr{P}^*)$ refers to the shortest Euclidean distance from o_i to any solution in $\mathscr{P}^*$ and $\mathrm{dist}(p_i, O)$ is defined analogously. Since $\Delta_{\mathscr{P}^*}$ is a non-negative scalar, a lower value indicates better convergence. Since in practice, the true POF $\mathscr{P}^*$ is typically unknown in real-world problems, we assume that $\mathscr{P}^*$ is the origin of the objective space (i.e., $\mathscr{P}^* = (0, 0, 0)$)—the ideal solution in a three objective minimization context.

The diversity of O is measured using the hypervolume per solution (HV^n), which quantifies the volume of the objective space dominated by the solution set with respect to a predefined reference point r and is computed as follows:

$$HV^n = \frac{1}{N_O} HV = \frac{1}{N_O}\sum_{i=1}^{N_O}\prod_{j=1}^{N_{\mathrm{obj}}}(r_j - o_{i,j}) \tag{9}$$

where r_j is the reference value for the j^{th} objective, and $o_{i,j}$ is the j^{th} objective value for the i^{th} solution. When O well-converges toward $\mathscr{P}^*$, a higher HV^n indicates greater solution diversity and a more accurate representation of the trade-offs along $\mathscr{P}^*$. In this study, the reference point was set to $r = (1, 1, 1)$, following the recommendation of [68].

2.3. ML modeling

2.3.1. ML-based proxies

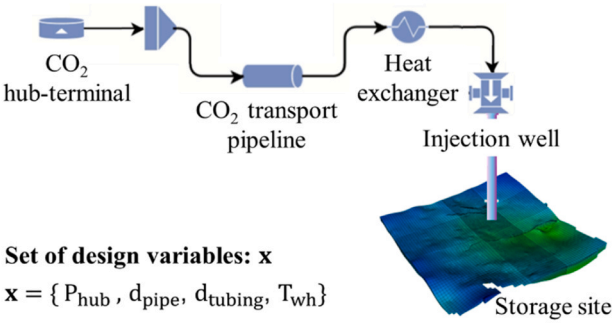
ML-based proxies are constructed using the design variables and corresponding objective function values obtained from the MOO results. They are trained to emulate the full-physics simulation outputs, thereby enabling rapid evaluation of a large number of feasible design solutions across the objective space. Once the surrogate models are trained, divide-and-conquer-based non-dominated-sorting and crowding-distance sorting algorithms are applied to the proxy-evaluated results to identify the optimal set of non-dominated-sorting solutions that approximate $\mathscr{P}^*$. Thereby, implementation and comparison of the predictive performances of five ML algorithms: artificial neural network (ANN), random forest (RF), gradient boosting (GB), extreme gradient boosting (XGB), and adaptive boosting (Ada) were accomplished. These methods are widely used in engineering optimization because they can capture complex nonlinear relationships in high-dimensional design spaces [43,46,48,69,70].

2.3.2. Performance metrics for the ML algorithms

The predictive performances of the ML-based proxies were evaluated using three commonly used metrics: the coefficient of determination (R^2), the root mean squared error (RMSE), and the mean absolute per-

1. Modeling of a unified CCS model

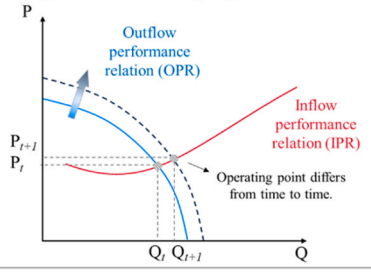
Objective: Model the CCS system integrating surface facilities and subsurface storage site.



Set of design variables: x

$\mathbf{x} = \{P_{hub}, d_{pipe}, d_{tubing}, T_{wh}\}$

[Dynamic nodal analysis]
- captures time-varying interactions



2. Initial Pareto search using MOO

Objective: Identify trade-offs among injectivity, safety, and economics

Set of objective functions: $f(\mathbf{x})$

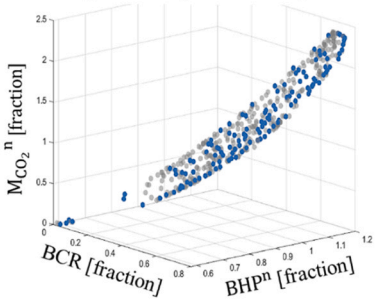
$f(\mathbf{x}) = \{M_{CO_2}, BHP, BCR\}$ { maximize M_{CO_2}; minimize BHP; maximize BCR

Preference-based constraints:

$M_{co_2} \geq M_{CO_2}^{min}, BHP \leq BHP^{max}, BCR \geq BCR^{min}$

Output: Simulation results under various combinations of design variables

[Pareto-optimal front]



3. Refined Pareto search using an ML-based proxy

Objective: Predict multi-objective values for refined design variables

Dataset: Results of multi-objective optimization

Input: Refined design variables (P_{hub} , d_{pipe}, d_{tubing}, T_{wh})

Output: Approximated objective function values (M_{CO_2}, BHP, BCR)

(1) Train and test ML proxy
(2) Evaluate all feasible solutions
(3) Filter within preference
(4) Apply divide-and-conquer-based non-dominated sorting and crowding distance sorting
(5) Derive refined Pareto-optimal front

Nomenclature & Superscript

- P_{hub}: discharge pressure at CO_2 hub-terminal
- d_{pipe}: pipeline inner diameter
- d_{tubing}: injection well tubing inner diameter
- T_{wh}: CO_2 temperature at injection wellhead
- ^{n}: normalized value
- M_{CO_2}: cumulative CO_2 injection
- BHP : bottomhole pressure
- BCR: benefit-cost ratio
- ^{min}: minimum value
- ^{max}: maximum value

Fig. 4. The hybrid framework including multi-objective optimization (MOO) used to analyze the unified carbon capture and storage (CCS) system.

centage error (MAPE) respectively defined as

$$R_i^2 = 1 - \frac{\sum_{k=1}^{N}\left(f_i^k - \widehat{f_i^k}\right)^2}{\sum_{k=1}^{N}\left(f_i^k - \bar{f}_i\right)^2} \tag{10}$$

$$\mathrm{RMSE}_i = \sqrt{\sum_{k=1}^{N}\frac{\left(f_i^k - \widehat{f_i^k}\right)^2}{N}} \tag{11}$$

$$\mathrm{andMAPE}_i = \frac{1}{N}\sum_{k=1}^{N}\left|\frac{f_i^k - \widehat{f_i^k}}{f_i^k}\right| \times 100 \quad (\%).\# \tag{12}$$

Here, subscript i denotes the i^{th} objective function, f_i^k denotes the full-physics-simulation-derived value of the i^{th} objective for the k^{th} design point, $\widehat{f_i^k}$ represents the corresponding value predicted by the ML proxy, $\bar{f}_i$ is the mean of simulation-derived values, and N is the number of samples used for model evaluation.

The R^2 metric can be used to evaluate the explanatory power of the ML proxy across the training, validation, and testing phases. A value close to 1 indicates that the proxy accurately captures the variance in simulation-derived objective values. RMSE can be applied as a primary loss function during model training and quantifies the standard deviation of prediction errors: a lower RMSE value corresponds to improved prediction fidelity. MAPE can be used to measure the average percentage deviation between the predicted and actual values, thereby providing an intuitive interpretation of prediction error. Ideally, a high-performing ML proxy will achieve an R^2 value near 1, along with low RMSE and MAPE values, indicating its strong generalization capability and high predictive accuracy.

2.4. Development of the hybrid MOO-ML framework for the pareto-optimal design of the unified CCS model

This study develops a hybrid framework for identifying the Pareto-optimal operating conditions of a unified CCS system. As shown in Fig. 4, the framework consists of three sequential stages: (1) modeling of the unified CCS system, (2) an initial Pareto search using MOO, and (3) a refined Pareto search using the ML-based proxies.

(1) Modeling of a unified CCS model: The first stage formulates a unified CCS model that encompasses the key physical components spanning the surface-to-subsurface continuum: the onshore CO_2 hub terminal, subsea transport pipeline, heat exchanger, injection well, and deep saline storage reservoir. System performance was evaluated using dynamic nodal analysis, where operating points are sequentially updated to reflect reservoir pressure evolution under continuous injection, while assuming steady-state flow within each time step. The design space was defined by using four key variables $\mathbf{x} = \{P_{hub}, d_{pipe}, d_{tubing}, T_{wh}\}$, which govern transport performance, injection behavior, and thermal efficiency. These variables influence the system's physical response and define the feasible region for the MOO.

(2) Initial Pareto search using MOO: The second stage applies an MOO algorithm was applied to evaluate trade-offs among injectivity, operational safety, and economic feasibility. The three objective functions—M_{CO_2}, BHP, and BCR—were evaluated under varying design configurations, with optimization in terms of operator-defined constraints on minimum injection mass ($M_{CO_2}^{min}$), maximum allowable pressure (BHP^{max}), and minimum economic viability (BCR^{min}) being evaluated for each one. The MOO algorithm produced an initial POF by simulating the full-physics CO_2 transport-injection scenario across the design space.

(3) Refined Pareto search using an ML-based proxy: In the final stage, ML-based surrogate models were developed to approximate objective function values across refined combinations of design variables. These proxies were trained on the MOO results and used to rapidly evaluate new design alternatives. This stage enhances resolution in the solution space while reducing computational demand. Refined design solutions were filtered based on preference constraints and ranked using divide-and-conquer-based non-dominated sorting and crowding-distance sorting. The final outcome was a refined POF that delineates the optimal trade-offs between the objectives.

3. Results and discussion

The proposed hybrid framework was applied to design a unified CO_2 transport, injection, and storage model for a medium-scale offshore CCS project in the Republic of Korea and to evaluate its geological CO_2 storage potential. A saline aquifer in Basin A was identified as a promising CO_2 storage reservoir due to its capacity potential and geological stability, as supported by multichannel seismic profiles and borehole data [71,72]. As the project was in the early planning phase without existing infrastructure, the aim of the present study was to derive Pareto-optimal combinations of front-end engineering design and operational parameters to support early-stage field development planning. Full-physics modeling and simulation of the unified CCS system were performed using CoFlow™, a geoenergy system management program with integrated surface facility and reservoir modeling capabilities [73]. MOO was performed using CMG's CMOST™, and ML algorithms were implemented using Python.

3.1. Unified CCS system design and evaluation: a case study for basin A, Republic of Korea

3.1.1. Model description

The unified CCS model consists of five main components: an onshore CO_2 hub terminal, a subsea transport pipeline, an offshore heat exchanger, an injection well, and a deep saline storage aquifer (Fig. 1). The onshore CO_2 terminal is connected to the Basin A aquifer via a 175-km subsea pipeline. The pipeline design accounts for seabed topography and flow assurance requirements using bathymetric data from the General Bathymetric Chart of the Oceans [74]. CO_2 is transported through the subsea pipeline and riser system to the injection wellhead, where a single vertical well located at the center of the saline aquifer enables continuous CO_2 injection over a 29-year operational period. The offshore platform is positioned on the continental shelf, and the total well length was determined based on the average water depth at the site. The onshore surface facilities were designed to operate without pump stations, so system performance evaluation was conducted under the condition of natural pressure buildup during injection. As a result, wellhead and bottomhole pressures were not directly controlled but were automatically calculated from P_{hub} through the transport and injection system. Although long-term CCS operations with multiple injection wells may require pump stations to control formation pressure [75], such infrastructure considerations were excluded from the scope of the present study given the early planning stage of the project. This

Table 1
Properties of the Basin A aquifer model.

Property	Value
Reservoir depth [m]	1150
Reservoir thickness [m]	60
Porosity [%]	20
Permeability [mD]	60
Initial pressure [MPa]	12
Reservoir temperature [°C]	50

modeling choice explains why BHP was formulated as an objective function in this study, rather than treated as a controllable design parameter.

Table 1 provides the key reservoir properties used to construct a homogeneous 3D model of the saline aquifer based on the geological and petrophysical data collated by [72,76]. The model was developed primarily from exploration-phase geological and seismic datasets, in the absence of well-based measurements (e.g., cores and wireline logs). Given these data limitations and the associated uncertainty, we adopted a homogeneous representation for the current modeling stage. Because reservoir heterogeneity can materially affect injectivity and pressure propagation, the optimal designs reported here should be interpreted as proof-of-concept results under a homogeneous baseline. Future work will extend the framework to geostatistical realizations and uncertainty-aware optimization to quantify design robustness. The storage site at approximately 1150 m below sea level has a thickness of 60 m, uniform porosity of 20%, and permeability of 60 mD. The initial reservoir conditions were a formation pressure of 12 MPa and a temperature of 50 °C, both of which are above the critical pressure (7.4 MPa) and temperature (31 °C) of CO_2 [77]. Accordingly, minimizing BHP within the allowable range of 15–20 MPa does not risk falling below the supercritical threshold. As further confirmed in the flow assurance analysis, the injected CO_2 remained in the dense supercritical phase under all evaluated conditions.

3.1.2. Experimental setup

As per Eq. (2), the unified CCS model comprises four key design variables: P_{hub}, d_{pipe}, d_{tubing}, and T_{wh}. Captured CO_2 is discharged directly from the onshore hub terminal and transported to the offshore aquifer through a fixed-length pipeline and injection well. Within this fixed infrastructure layout, d_{pipe} and d_{tubing} function as facility design parameters, while adjustments to T_{wh} are made via a heat exchanger embedded within the unified CCS model.

The three objective functions introduced in Eq. (1) were normalized and reformulated as follows:

$$f(\mathbf{x}) = f(P_{hub}, d_{pipe}, d_{tubing}, T_{wh}) = \left\{ \frac{1}{M_{CO_2}{}^{n}}, \ BHP^{n}, \frac{1}{BCR} \right\} \tag{13}$$

under the constraints that $M_{co_2} \geq M_{CO_2}^{min}$,

$BHP \leq BHP^{max}$, and $BCR \geq BCR^{min}$.

The first objective, $M_{CO_2}{}^{n}$, represents the total injected CO_2 mass over the 29-year injection period normalized to the minimum required cumulative injection mass, $M_{CO_2}^{min}$:

$$M_{CO_2}{}^{n} = \frac{M_{co_2}}{M_{CO_2}^{min}} \tag{14}$$

Here, $M_{CO_2}{}^{n} \geq 1$ indicates that the storage capacity preference is satisfied. The value of $M_{CO_2}^{min}$ was determined based on representative annual injection rates for medium-scale CCS projects (100,000–500,000 tCO_2/year) and the project duration. For confidentiality reasons, the exact value has not been disclosed.

The second objective, BHP^{n}, is used to evaluate the storage formation's pressure stability. It is defined as the maximum BHP observed during the injection period normalized to a predefined threshold, BHP^{max}:

$$BHP^{n} = \frac{\max\left\{ BHP_{t_0}, BHP_{t_1}, \cdots, BHP_{t_k}, \cdots, BHP_{t_{N_t^m}} \right\}}{BHP^{max}}, \tag{15}$$

where t_m is the time index in months ($t_m = 0, 1, 2, \cdots, N_t^m$) and N_t^m is the total number of months in the 29-year period. BHP^{max} reflects the safety criteria suggested by the US Electronic Code of Federal Regulations [78], which recommends maintaining a pressure below 90% of the estimated fracture pressure. In the present study, this threshold was estimated based on analogous sedimentary formations in the Republic of Korea. A value of $BHP^{n} \leq 1$ satisfies the safety requirement.

The third objective, BCR, represents the system's economic feasibility and is defined as the ratio of discounted benefits to costs:

$$BCR = \frac{\sum_{t=0}^{N_t^m} \frac{C_{benefit,t}}{\left(\frac{1+r_d}{12}\right)^t}}{\sum_{t=0}^{N_t^m} \frac{C_{cost,t}}{\left(\frac{1+r_d}{12}\right)^t}} \tag{16}$$

$$C_{benefit,t} = C_{CO_2,credit} \times M_{CO_2,t} \tag{17}$$

$$C_{cost,t} = C_{CO_2,inj} \times M_{CO_2,t} + C_{pipe} + C_{tubing} \tag{18}$$

where $C_{benefit,t}$ and $C_{cost,t}$ represent the monetary benefit and cost at month t, respectively, and r_d is the annual discount rate. $C_{benefit,t}$ incorporates the CO_2 credit price ($C_{CO_2,credit}$) and monthly injected CO_2 mass ($M_{CO_2,t}$) while $C_{cost,t}$ includes the CO_2 injection cost ($C_{CO_2,inj}$), as well as the infrastructure costs for the pipeline (C_{pipe}) and tubing (C_{tubing}).

Table 2 summarizes the economic parameters used to calculate the BCR: r_d was set at 0.1/year, reflecting typical rates used in energy infrastructure planning; $C_{CO_2,credit}$ was set at 87.5 USD/tCO_2, representing the average carbon price for 2030 (ranging from 75 to 100 USD/tCO_2) by the International Energy Agency [79]; and $C_{CO_2,inj}$ derived from [80] provides a comprehensive cost profile that includes CO_2 capture, transport, and storage tailored for offshore CCS project settings. The capital costs for pipeline and injection well tubing—C_{pipe} and C_{tubing}, respectively—were calculated using [81]'s empirical cost model (Eq. (19)), which accounts for materials, labor, miscellaneous, and right-of-way costs:

$$C_{pipe} = C_{tubing} = \left(674 \times d^2 + 11,754 \times d + 234,085\right) \times (1.60934 \times L) + 405,000, \tag{19}$$

where d denotes the inner diameter (inches) of the pipeline or tubing, L represents the length (km), and the resulting cost (USD). The values of C_{pipe} and C_{tubing} vary according to their respective inner diameters (d_{pipe}

Table 2
Economic parameters used to calculate the benefit-cost ratio.

Parameter	Value	Ref.
r_d [fraction]	0.1	-
$C_{CO_2,credit}$ [USD/tCO_2]	87.5	[65]
$C_{CO_2,inj}$ [USD /tCO_2]	76.8	[66]
C_{pipe} [million USD]	158.73–319.79 (varying by pipeline diameter)	[67,68]
C_{tubing} [million USD]	1.56–1.93 (varying by pipeline diameter)	[67,68]

and d_{tubing}) and lengths. Material costs are scaled linearly with pipeline or tubing length, while labor typically represents 40–50% of the total construction cost. Miscellaneous costs include design, supervision, and contingencies, while right-of-way costs vary by location. These values were adjusted to 2022 USD using the average inflation rate provided by the Organization for Economic Co-operation and Development (OECD) [82] to ensure consistency across references.

The economic feasibility preference, denoted as BCR^{min}, was set to 0.5 based on the outcomes of feasibility studies over the past five years for publicly funded infrastructure projects in the Republic of Korea. While a BCR above 1 is generally considered the threshold for commercial viability, public investment decisions are often influenced by broader criteria, including policy relevance, environmental impact, and social benefits, many of which are difficult to quantify or monetize directly [83]. Broader assessments involving these dimensions—such as regional development or climate policy alignment—are beyond the scope of this technical analysis. Instead, we referenced empirical BCR data from publicly funded infrastructure and energy-related projects between 2020 and 2024. During this period, projects approved for implementation reported an average BCR of 0.69, while rejected proposals averaged 0.30 (Korea Institute of S&T Evaluation and Planning (KISTEP) [84]). Based on these reference values, a threshold value of $BCR^{min} = 0.5$ was selected to indicate that an economic preference criterion reasonably aligned with public sector decision-making standards. Table 3 summarizes the preference criteria adopted for the three objective functions—CO_2 storage capacity, operational safety, and economic feasibility—reflecting priorities relevant to both government agencies and operating entities.

In the MOO stage of the hybrid framework, a multi-objective particle swarm optimization (MOPSO) algorithm, which has been widely applied to a variety of engineering problems [85–87], was employed to explore the design space of the unified CCS model. The search space was defined based on standard engineering design specifications (Table 4). The ranges for each of the four design variables—P_{hub}, d_{pipe}, d_{tubing}, and T_{wh}—were set after considering guidelines from the American Petroleum Institute (API) [88] and the literature on CCS system design [75]. To explicitly verify phase stability within these design bounds, Fig. 5 presents a CO_2 phase diagram with the simulated P–T trajectories overlaid along the path from the source to the reservoir. Within the prescribed operating range, the minimum source conditions were 7 MPa and 10 °C. The operating P–T trajectories consistently remain outside the vapor–liquid envelope, thereby avoiding crossing the saturation (boiling) line across all evaluated cases. These operating constraints keep CO_2 in a condensed state at the source and enable a direct transition to a supercritical state in the reservoir, ensuring robust single-phase flow operation. As a result, the risk of phase separation is effectively removed, supporting stable hydraulics and reliable injection performance throughout the system.

A uniform grid search with a step size of 1 was applied within each variable's range, resulting in 13,377 design combinations. The search space was explicitly defined by the bounds of the decision variables, and candidate solutions were initialized within these bounds. Constraint-handling ensured that infeasible solutions were penalized, keeping the search within the feasible region [89]. The algorithm retained the best solutions identified during the search and updated an archive of non-dominated solutions, from which the POF was identified after the final generation using non-dominated sorting and crowding-distance sorting [90,91]. To efficiently sample this space, the MOPSO algorithm evaluated 503 candidate solutions. Each simulation based on full-physics modeling required approximately 30 min of computational time. The MOPSO hyperparameters—inertia weight, cognitive component, and social component—were set according to the tuning recommendations of [92] (Table 5). In this study, the number of evaluations was used as the termination criterion. These settings ensured balanced search behavior between global exploration and local refinement within the multi-dimensional objective space.

Table 3
Decision maker's preference criteria for the three objective functions.

Objective Function	Preference	Reason for Preference
$M_{CO_2}{}^{n}$	$M_{co_2} \geq M_{CO_2}^{min}$	Annual injection rate (0.1–0.5 Mtpa) for medium-scale CCS projects
BHP^{n}	$BHP \leq BHP^{max}$	U.S. Code of Federal Regulations on fluid injection into geological formations
BCR	$BCR \geq BCR^{min} = 0.5$	BCR for public project feasibility studies in South Korea from 2020 to 2024

Table 4
Ranges of design variables for multi-objective optimization.

Design Variable	Base-Case	Lower Limit	Upper Limit	Step Size	Number of Steps
P_{hub} [MPa]	10	7	13	1	7
d_{pipe} [inch]	12	6	18	1	13
d_{tubing} [inch]	5	2	8	1	7
T_{wh} [°C]	20	10	30	1	21
Number of feasible combinations of design variables					13,377

In the ML modeling stage of the hybrid framework, five ML proxies—ANN, RF, GB, XGB, and Ada—were trained, validated, and tested to learn the nonlinear relationships between the four design variables and the three objective functions. A total of 504 data points, comprising 503 solutions from the MOO stage and one base-case scenario, were used as the learning data. These data were partitioned into training, validation, and test sets with a ratio of 60:20:20. Prior to model training, all features and targets were standardized to mitigate scale-related bias and improve numerical stability. Robustness was ensured through five-fold cross-validation, and a representative fold was selected for each ML proxy based on the mean R^2 value across the three objectives. The selected model was then used to evaluate a refined design space. To further ensure the consistency of the optimal solutions, we aimed to verify that the derived Pareto-optimal results were reproducible and stable across runs. Previous studies have shown that MOO algorithms such as MOPSO tend to converge toward a similar POF across independent runs due to their population-based search mechanisms, selection strategies, and diversity-preservation methods [93,94]. In addition, five independent ML models were trained during cross-validation to confirm the robustness and reproducibility of the ML-based POF derivation.

In contrast to the relatively coarser discretization used in the MOO stage (Table 4), the step sizes in the ML stage were refined for key facility design variables while maintaining the original bounds. Specifically, the step size was reduced from 1 to 0.5 in. for d_{pipe} and from 1 to 0.125 in. for d_{tubing}. These refined step sizes are consistent with standard API specifications that support common fractional dimensions such as 1½ and $4^{3}/_{8}$ inches [88]. As a result, the total number of feasible design combinations expanded from 13,377 to 180,075—an approximately 13.5-fold increase—thereby enabling significantly higher resolution in the design space (Table 6).

This expansion allowed the trained ML proxies to explore a far denser solution space at negligible computational cost compared to a full-physics simulation. To identify Pareto-optimal solutions, divide-and-conquer-based non-dominated sorting was used for the ML-evaluated solutions, while traditional non-dominated sorting was applied to the MOO set. The hyperparameters for each ML proxy were systematically optimized using GridSearch, and model selection was based on validation performance considering R^2, RMSE, and MAPE. Robustness was further ensured through five-fold cross-validation. The train/validation/test split (60/20/20) and feature/target standardization were applied consistently across all models to ensure a fair comparison. For the ANN, for example, the selected topology comprised Dense(64, ReLU), Dense (32, ReLU), and Dense(3, Linear) layers, trained with Adam and mean-

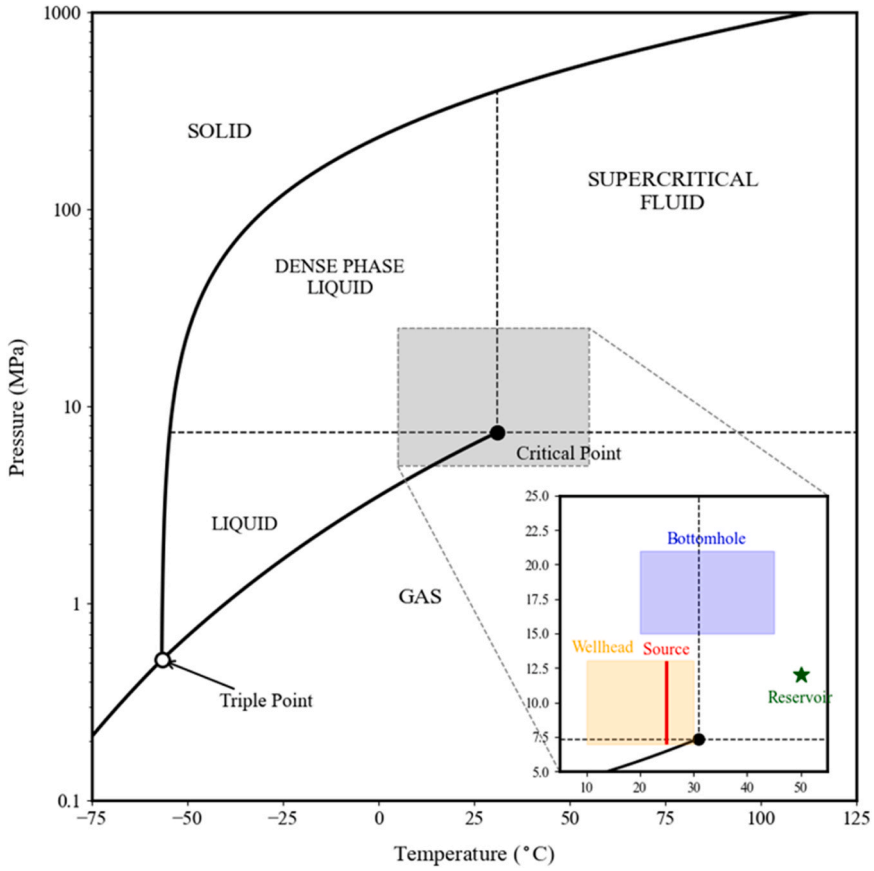


Fig. 5. CO_2 Phase diagram with simulated operating P–T trajectories from source to reservoir.

Table 5
Experimental settings for the multi-objective particle swarm optimization algorithm.

Parameter	Value
Number of evaluations	503
Inertia weight	0.7298
Cognitive component	1.49618
Social component	1.49618

Table 6
Ranges of the design variables in the refined search space.

Design Variable	Base-Case	Lower Limit	Upper Limit	Step Size	Number of Steps
P_{hub} [MPa]	10	7	13	1	7
d_{pipe} [inch]	12	6	18	0.5	25
d_{tubing} [inch]	5	2	8	0.125	49
T_{wh} [°C]	20	10	30	1	21
Number of feasible combinations of design variables					180,075

squared-error loss (see Table 7 for the hyperparameter specifications of ANN, RF, GB, XGB, and Ada). Note that the objective of this comparison is to verify that well-tuned ML proxies provide robust and reproducible performance within the MOO–ML framework, rather than to prioritize a particular algorithm.

3.2. Identification of the POF

3.2.1. MOO results

The convergence behavior of the MOO process is summarized in Fig. 6, which reports the cumulative HV and the relative HV improvement over generations. As shown in Fig. 6(a), HV increases sharply during the early generations, reflecting efficient exploration and rapid expansion of the dominated objective-space region. After approximately 20 generations, the HV curve approaches a plateau, indicating that further evaluations provide only marginal gains and that the solution set has converged toward a stable POF. This interpretation is corroborated by Fig. 6(b), where the relative HV improvement becomes negligible from the 15th generation onward. The near-simultaneous stagnation of cumulative HV and its improvement rate demonstrates that the optimization was effectively completed within the predefined evaluation budget, achieving convergence and computational efficiency.

To establish a performance baseline prior to performing MOO using the MOPSO algorithm, a base-case simulation was conducted using median expert-prescribed values for the design variables: P_{hub}

Table 7
Hyperparameters used for the five machine-learning proxies.

Proxy Parameter	ANN	RF	GB	XGB	Ada
Number of estimators	–	100	100	400	200
Maximum depth	–	None	3	3	3
Learning rate	–	None	0.1	0.03	0.05
Minimum sample leaf	–	1	None	None	4
Minimum sample split	–	2	None	None	5
Layers	Dense (64,ReLU), Dense (32,ReLU), Dense (3, Linear)	–	–	–	–
Optimizer	Adam	–	–	–	–
Loss function	Mean squared error	–	–	–	–

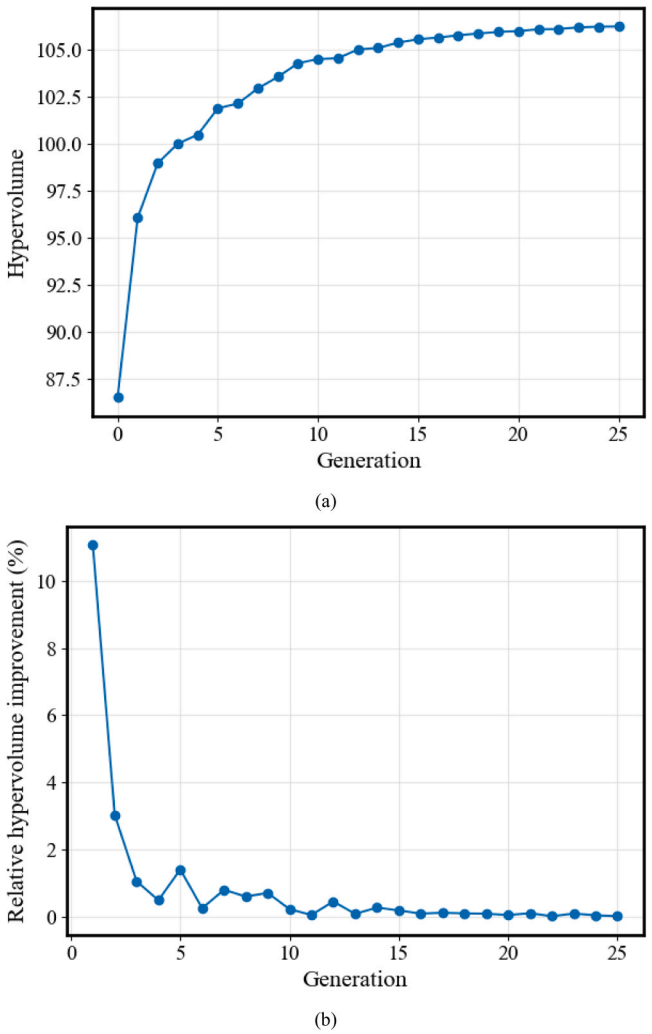


Fig. 6. Convergence analysis of the multi-objective optimization: (a) cumulative hypervolume and (b) relative hypervolume improvement over generations.

= 10 MPa, d_{pipe} = 12 in., d_{tubing} = 5 in., and T_{wh} = 20 °C; the resulting objective values were $M_{CO_2}{}^n = 1.5$, $BHP^n = 0.99$, and BCR = 0.63. These results indicate that the base-case attains 150% of the minimum required CO_2 injection mass, maintains the injection pressure at 99% of the allowable threshold, and exceeds the economic preference threshold. Because this baseline performance already satisfies all of the preference criteria, the purpose of the MOO process is to explore trade-offs and identify marginally improved or more diverse solutions across the objective space.

Fig. 7 illustrates the distribution of 503 candidate solutions (gray dots) evaluated using the MOPSO algorithm, representing approximately 3.8% of the total 13,377 design combinations defined in the MOO design space. Among them, 128 solutions (highlighted in blue in Fig. 7(a)) were identified as the best non-dominated solutions to form the initial POF. The total evaluated set and the best non-dominated solutions subset were both uniformly distributed in the 3D objective space, thus clearly revealing the trade-offs between CO_2 storage capacity ($M_{CO_2}{}^n$), injection safety (BHP^n), and economic feasibility (BCR). Notably, solutions achieving higher $M_{CO_2}{}^n$ values tended to exhibit correspondingly higher BHP^n values, thereby reflecting physical coupling between the storage mass and the storage reservoir pressure. Meanwhile, solutions with higher BCR values were typically concentrated near the median of the $M_{CO_2}{}^n$ distribution rather than at its extremities, indicating that maximum economic performance does not necessarily coincide with either the minimal or maximal CO_2 injection mass. Fig. 7(b) provides a filtered subset of the 46 best non-dominated solutions (red dots) from the original 128 that satisfy all three objective preferences: $M_{CO_2}^{min}$, BHP^{max}, and BCR^{min}. Meanwhile, the remaining 82 solutions in the best non-dominated set violated one or more of these preference thresholds. This filtering process could enable decision-makers to focus on a region of interest (ROI) within the operational, regulatory, and economic constraints. By narrowing the solution space to only those that meet decision-maker-defined criteria, the framework facilitates targeted evaluation and selection of optimal designs that are both technically and economically viable.

To assess the quality of the optimization results, both convergence and diversity were evaluated using convergence metric $\Delta_{\mathscr{P}^*}$ and diversity metric HV^n, respectively. For the complete set of 503 MOPSO-evaluated solutions, $\Delta_{\mathscr{P}^*}$ was 0.640, and HV^n was 0.393. Among the 128 best non-dominated solutions, these metrics improved slightly to $\Delta_{\mathscr{P}^*} = 0.638$ and $HV^n = 0.391$, indicating marginally better convergence toward the true POF accompanied by a slight reduction in diversity. This modest loss in diversity can be attributed to the elimination of extreme solutions that had previously contributed to broader coverage of the objective space. In contrast, the subset of 46 preferred best non-dominated solutions exhibited a markedly improved convergence score of $\Delta_{\mathscr{P}^*} = 0.540$ and a higher diversity value of $HV^n = 0.457$. These improvements suggest that preference-based filtering not only sharpens proximity to the ideal POF but also enhances the evenness of solution distribution within the practical ROI. As such, incorporating decision-makers' preferences into the selection process yields a refined, interpretable solution set that better aligns with real-world deployment criteria.

Fig. 8(a–c) presents correlation matrixes for three groups of solutions: all 503 MOPSO-evaluated solutions, the 128 best non-dominated-sorting solutions, and the 46 best non-dominated-sorting solutions satisfying all preference criteria, respectively. Because preference filtering was applied after the full set of simulations had been evaluated, the correlation patterns in Fig. 8(a) primarily reveal the intrinsic physical relationships among the design variables and objective functions that are independent of the decision-maker constraints. Comparing the correlation patterns in Fig. 8(a) with those in Fig. 8(b) reveals stronger correlations using the non-dominated-sorting solution set, thereby indicating that MOO clearly exposed trade-offs and interdependencies. Notably, the correlations between d_{tubing} and the objective functions increased from 0.43 to 0.59 for $M_{CO_2}{}^n$, 0.44–0.60 for BHP^n, and 0.37–0.42 for BCR. Similarly, the influence of T_{wh} became more pronounced, with its correlations shifting from −0.32 to −0.48 for $M_{CO_2}{}^n$, −0.27 to −0.46 for BHP^n, and −0.37 to −0.60 for BCR. Across the objective functions, the perfect correlation between $M_{CO_2}{}^n$ and BHP^n remained unchanged at 1.00 while the correlations between $M_{CO_2}{}^n$ and BCR, and between BHP^n and BCR, increased modestly from 0.80 to 0.83. These heightened correlation coefficients confirm that non-dominated solutions not only capture stronger functional dependencies but also explicitly represent the trade-off relationships that characterize the POF, where improvement of one objective is generally accompanied by the sacrifice of another. Fig. 8(c) shows the correlation matrix for the 46 best non-dominated solutions that meet all of the predefined objective preferences. Compared to Figs. 8(a) and 8(b), this filtered subset exhibited more distinct and interpretable correlation patterns that point toward its alignment with the decision-maker priorities. One of the most notable features is the strong negative correlation (–0.76) between d_{pipe} and BCR, which reflects the economic impact of pipeline sizing. Since the pipeline accounts for nearly half of the total system cost, increasing d_{pipe} significantly raises the capital expenditure, thereby reducing BCR.

For the preference-filtered subset, the perfect positive correlation between $M_{CO_2}{}^n$ and BHP^n was consistent for all three objective functions (Figs. 8(a) and 8(b)) whereas the correlations between BCR and the other two objectives weakened; the latter indicates that maximizing the economic return is not strictly tied to maximizing or minimizing the

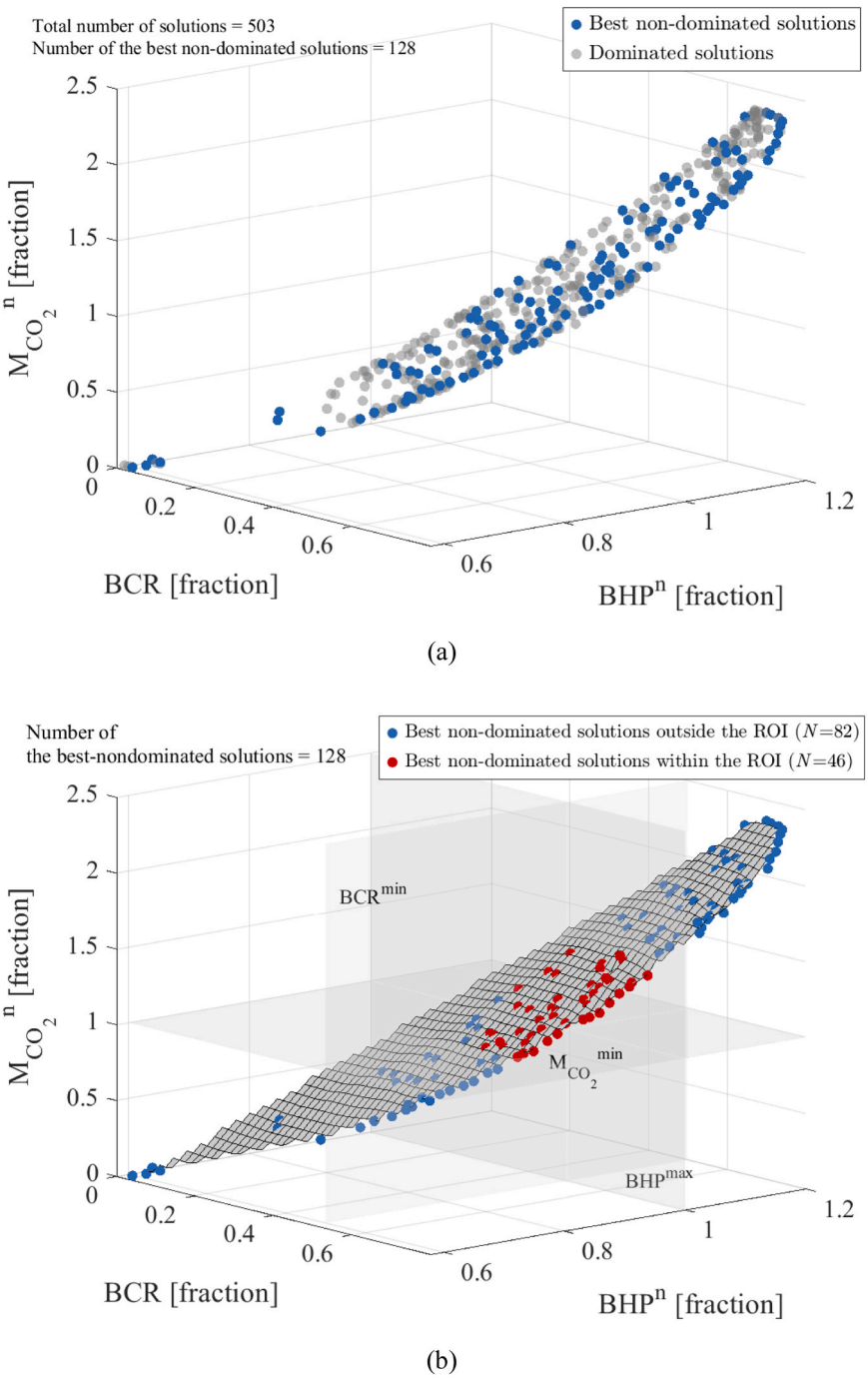


Fig. 7. Distribution of the evaluated solutions in the 3D objective space: (a) all 503 evaluated solutions and (b) the 128 best non-dominated solutions classified via preference satisfaction (46 and 82 solutions satisfying and not satisfying the preferences, respectively).

injection mass or storage reservoir pressure. As illustrated by the red dots of Fig. 7(b), solutions with the highest BCR values tended to cluster near the median of the $M_{CO_2}^{n}$ distribution rather than at its extremities. This suggests that the most cost-effective designs are not necessarily those that maximize CO_2 injection or minimize pressure but rather those that strike a balanced trade-off among the competing goals. Overall, preference-based filtering highlights these trade-offs in an interpretable and relevant manner.

3.2.2. *ML proxy-based approximation results*

Five ML proxies (ANN, RF, GB, XGB, and Ada) were developed to approximate the three objective functions efficiently using data generated from the MOO stage. These proxies enabled high-resolution exploration of the POF without the need for additional full-physics simulations. Fig. 9(a–c) presents Taylor diagrams for the three objective functions ($M_{CO_2}^{n}$, BHP^{n}, and BCR, respectively) to compare the predictive performance of the five ML proxies against simulation results using three key metrics: correlation coefficient (R), normalized standard deviation (σ^{n}), and RMSE. The models trained on the 504 MOO-labeled solutions demonstrated robust predictive accuracy across all of the evaluated metrics using their representative cross-validation folds.

As an example of using the ML proxies, Fig. 10 illustrates the sequential process for deriving the POF using the RF model. First, the trained ML proxy was used to predict the objective function values for all

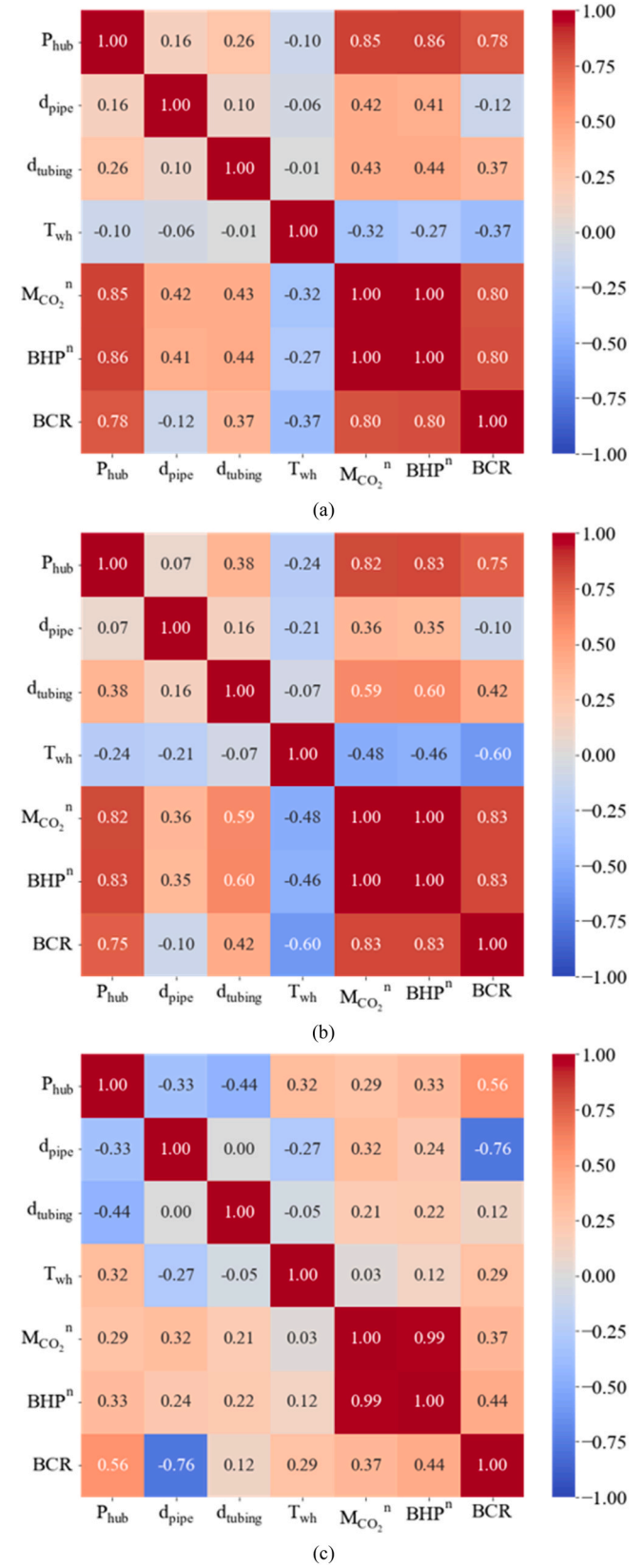


Fig. 8. Correlation matrixes of solutions obtained from the multi-objective optimization process: (a) all 503 evaluated solutions, (b) the 128 best non-dominated solutions, and (c) the 46 best non-dominated solutions satisfying the objective preferences.

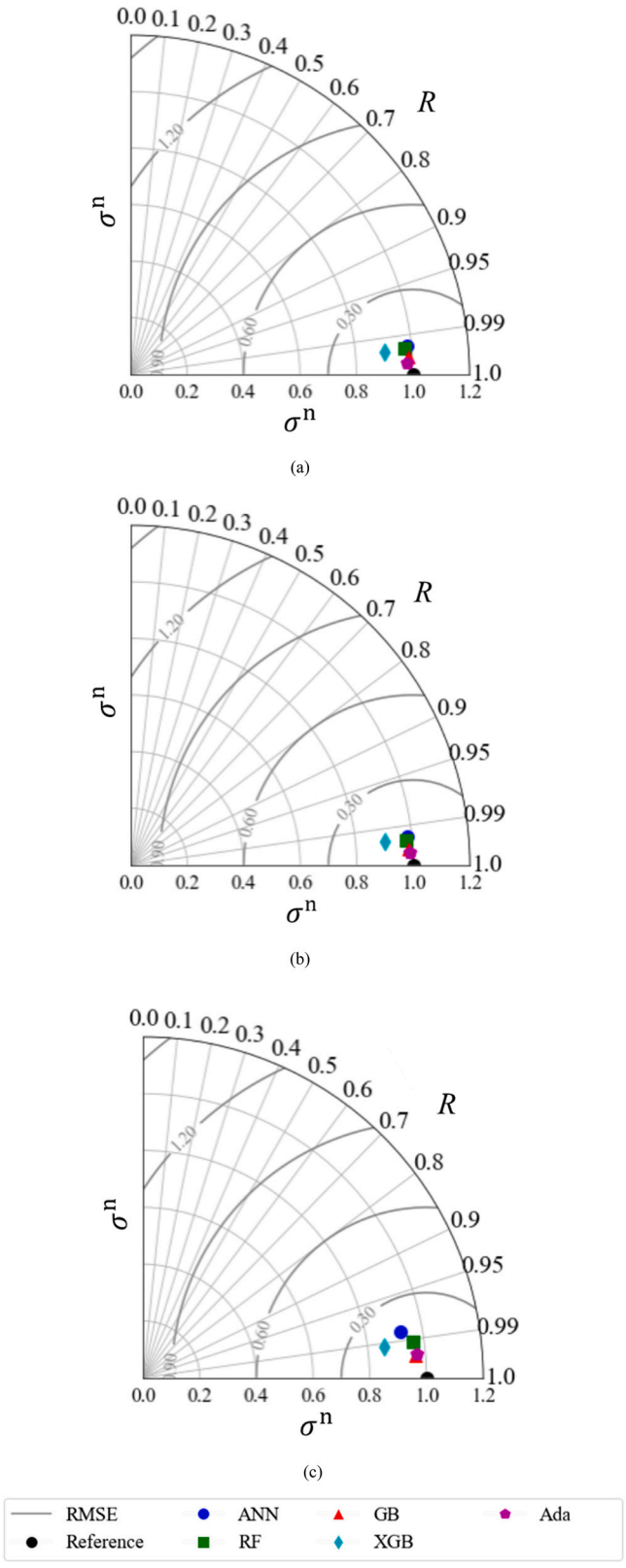


Fig. 9. Taylor diagrams illustrating the predictive performances of the machine-learning proxies in terms of (a) $M_{CO_2}{}^n$, (b) BHP^n, and (c) BCR.

180,075 refined combinations of design variables and yield a comprehensive solution set (Fig. 10(a)). Next, solutions were filtered based on the defined ROI, resulting in 66,517 preferred designs that satisfied all of the objective preference thresholds (Fig. 10(b)). The other proxies

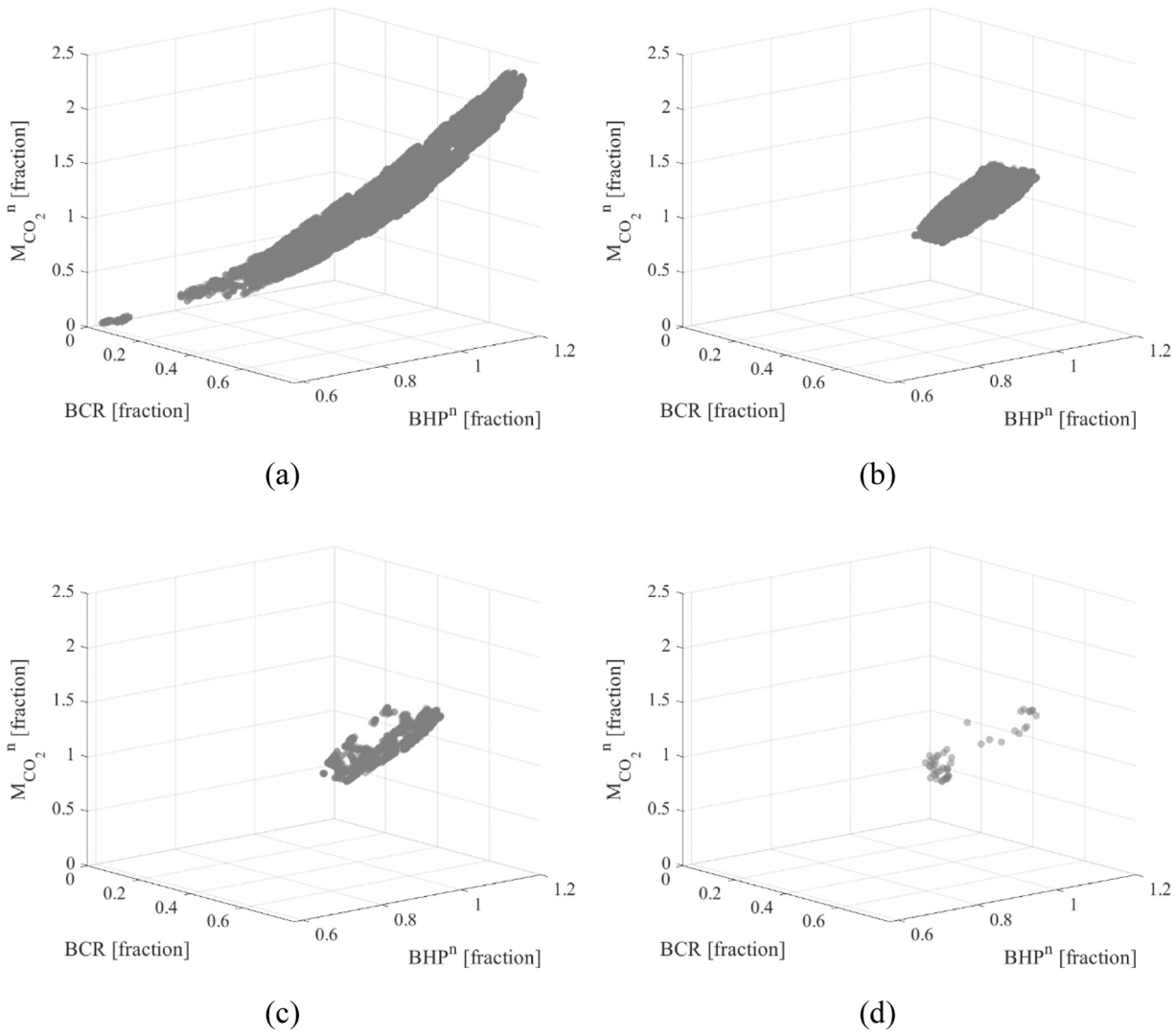


Fig. 10. Distributions of the solutions in the refined 3D objective space evaluated using the random forest (RF) proxy trained on the multi-objective optimization results: (a) all of the evaluated solutions (180,075), (b) solutions filtered within the preference range (66,517), (c) Pareto-optimal solutions after divide-and-conquer non-dominated sorting (7542), and (d) the final 50 selected Pareto-optimal solutions after additional crowding-distance sorting.

produced similar filtered subsets ranging from 64,000 to 72,000 solutions. Based on the filtering results, the divide-and-conquer-based non-dominated-sorting algorithm was applied to identify the optimal non-dominated front, which further narrowed the pool to 7542 solutions in the RF case (Fig. 10(c)). Across all of the proxies, the number of non-dominated solutions varied from 1776 to 7542 solutions. Finally, crowding-distance sorting was applied to select a diverse subset of and high-performing solutions that balance convergence and an even spread

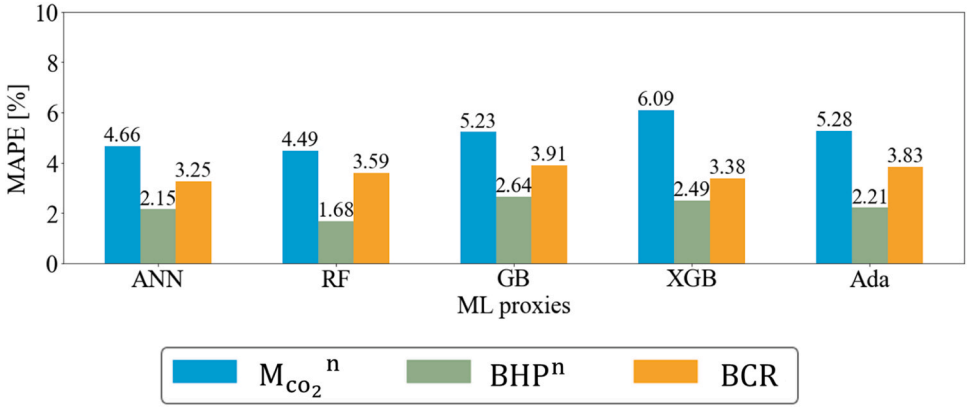


Fig. 11. Mean average percentage error (MAPE) results for the three objective function predictions made using the five machine-learning proxies evaluated on the final 50 selected Pareto-optimal solutions.

along the POF (Fig. 10(d)). Specifically, 50 representative solutions were retained to approximate the 3D convex POF while preserving adequate coverage of the trade-off structure and minimizing redundancy among neighboring designs [95]. This user-defined limit of 50 ensured computational efficiency for the subsequent dynamic nodal and flow-assurance evaluations, while still providing decision-makers with a manageable set of optimal design alternatives.

For validation and accurate objective function evaluation, full-physics simulations were rerun for all 50 selected solutions and the simulation-derived values were used in place of ML-based predictions. Fig. 11 presents the MAPE results for the predictions using the five ML proxies across the three objective functions that were calculated based on the full-physics simulation results for the 50 final design scenarios selected through the Pareto-optimal filtering process. The results demonstrate strong predictive performance, with average MAPE values of 5.15% for $M_{CO_2}{}^n$, 2.23% for BHP^n, and 3.59% for BCR. For all of the

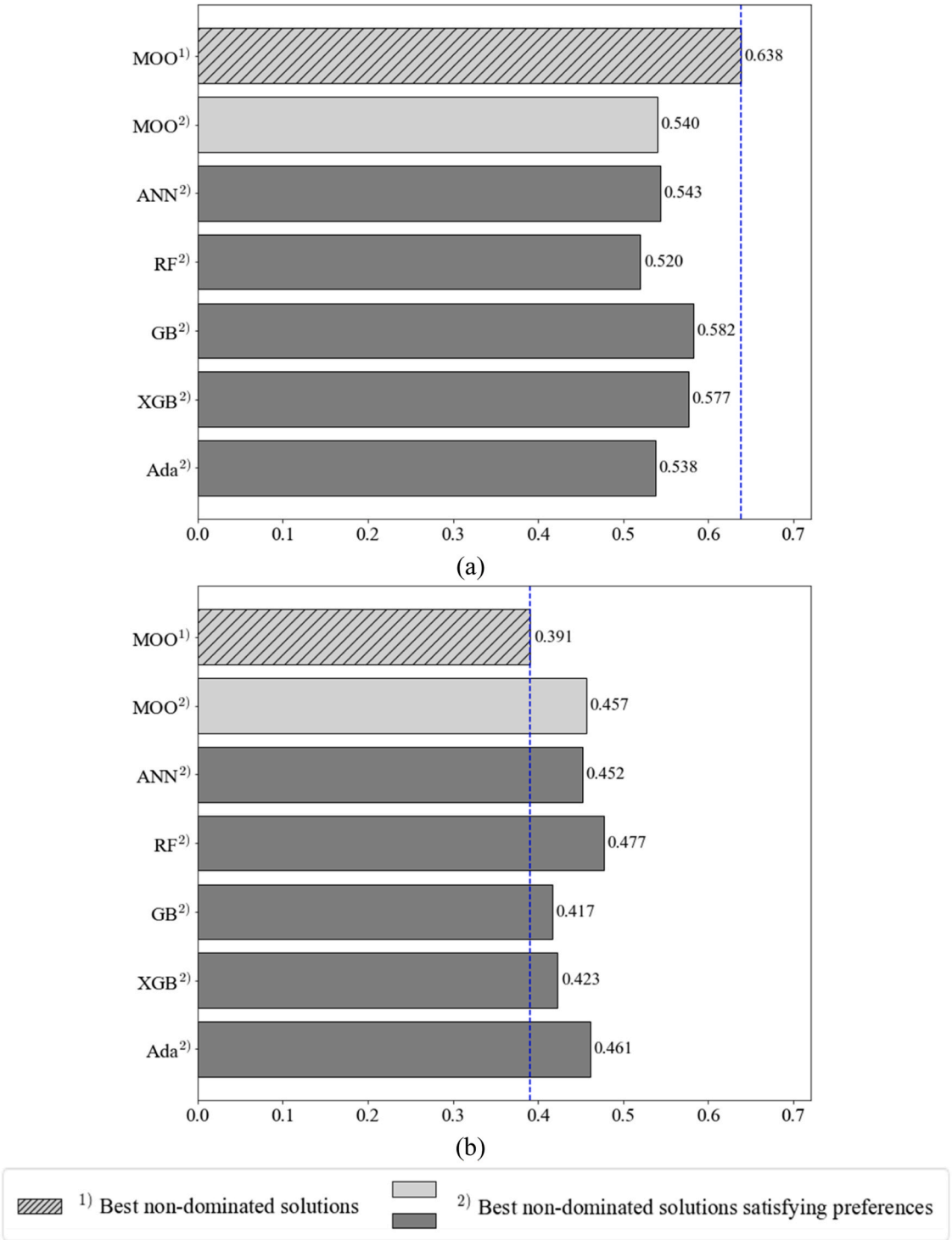


Fig. 12. Performance metrics when using multi-objective optimization (MOO)-only or MOO with one of the five machine-learning (ML) proxies applied to analyze the unified carbon capture and storage system in terms of (a) convergence ($\Delta_{\mathscr{P}}$) and (b) diversity (HV^n). The striped bars and vertical dashed lines highlight the best non-dominated solutions using MOO only, while the solid bars indicate the best non-dominated solutions satisfying the preferences after applying MOO.

ML proxies, the prediction error remained well within 5% on average, thus confirming the reliability of the trained models for approximating the objective functions in a high-resolution design space. These findings validate the use of ML proxies as a computationally efficient yet accurate alternative for exploring and refining optimal CCS system designs.

Fig. 12 shows the performances of the MOO-only and the five ML proxy-derived solutions in terms of convergence ($\Delta_{\mathscr{P}^*}$) (Fig. 12(a)) and diversity (HV^n) (Fig. 12(b)). For MOO, both the full set of 128 non-dominated solutions (striped bar) and the preference-satisfying subset of 46 solutions (light gray bar) are displayed. For the ML proxies, the final sets of 50 preference-satisfying solutions are shown (solid gray bars). As previously reported in Section 3.2.1, the 128 MOO-derived non-dominated solutions exhibited $\Delta_{\mathscr{P}^*} = 0.638$ and $HV^n = 0.391$, while the refined 46-solution set showed improvements to $\Delta_{\mathscr{P}^*} = 0.540$ and $HV^n = 0.457$, respectively. The ML proxies achieved $\Delta_{\mathscr{P}^*}$ values ranging from 0.520 to 0.582 (mean = 0.552, standard deviation = 0.02) and HV^n values from 0.417 to 0.477 (mean = 0.446, standard deviation = 0.02). These results demonstrate that the ML proxies trained on the MOO-derived simulation data can efficiently extend the design space search to a much finer resolution, thereby identifying diverse, high-quality solutions that align with the decision-maker preferences. Despite the broader scope, the ML-derived solutions achieved convergence and diversity metrics comparable to those from simulation-based MOO but at a fraction of the computational cost, thus highlighting the practical effectiveness of ML-assisted optimization for early-stage CCS system design.

Table 8 provides selected Pareto-optimal solutions obtained from the MOO and ML stages that highlight the design achieving the best performance for each objective. The MOO-only solutions exhibited notable improvements over the base-case scenario, specifically a 6.00% increase in $M_{CO_2}{}^n$, a 15.15% reduction in BHP^n, and a 14.29% increase in BCR. In comparison, the ML-driven solutions on average achieved a 6.87% in $M_{CO_2}{}^n$, a 14.75% reduction in BHP^n, and a 14.92% increase in BCR relative to the base-case scenario. Thus, the ML proxies extended the optimization to a refined design space and uncovered additional configurations with comparable or enhanced performance. These outcomes reflect that applying an ML proxy not only preserved the trade-off structures identified through MOO but also expanded them into high-resolution regions of the design space.

Fig. 13 visualizes the distributions and statistics of the three objectives ($M_{CO_2}{}^n$, BHP^n, and BCR) for the 50 Pareto-optimal solutions derived using the RF model. The diagonal panels show density histograms, where colored vertical lines denote the five quantiles including minimum (Q_0), first quartile (Q_1), median (Q_2), third quartile (Q_3), maximum (Q_4), and the black line marks the mean with its standard deviation (σ) reported next to it. The off-diagonal panels provide pairwise scatterplots that reveal dependence and trade-off structure among the objectives. This figure provides a quantitative overview of the range, variability, and extreme solutions (minimum and maximum) of the Pareto-optimal results, complementing Table 8 by highlighting the medians and percentage differences across the solution set. While MOO studies typically emphasize extreme solutions or diversity metrics due to the quasi-uniform distribution of Pareto-optimal results [25,90,96], Fig. 13 complements the case-specific performance comparisons in Table 8 by quantitatively summarizing the overall distribution of optimal designs.

Fig. 14 illustrates the computational advantage of using ML proxies for refined scenario analysis. All computations were performed on a workstation equipped with an Intel® Xeon® Silver 4214 R CPU (2.40 GHz), 64 GB RAM, and an NVIDIA GeForce RTX 3090 GPU. While the MOO stage required approximately 251.5 h to complete 503 full-physics simulations, the trained ML proxies evaluated all 180,075 refined design combinations in just 19.24 s. In particular, the divide-and-conquer-based algorithm enabled non-dominated sorting of 180,075 solutions to be completed within 20.2 min. As a result, the entire process of deriving the refined POF—filtering, sorting, and selection—was completed in an average of 22 min and 11 s. This efficiency is notable, as performing non-dominated sorting for such a large solution set was computationally challenging with our available resources. To ensure accuracy, 50 additional simulations were performed to validate the ML-derived results, bringing the total number of full-physics runs to 554 (one base-case, 503 MOO simulations, and 50 validations)—which accounts for only 0.3% of the refined design space. These outcomes underscore the scalability and practicality of the hybrid framework, making it possible to perform high-resolution design optimization with negligible additional simulation cost.

3.3. Operational feasibility and flow assurance of the pareto-optimal solutions

The operational feasibility of the refined Pareto-optimal solutions was assessed using IPR-OPR curves from the dynamic nodal analysis. As a representative case, the POF generated by the RF model was used in the endeavor. Fig. 15 displays the IPR-OPR curves and their intersection points (i.e., operating conditions) for all 50 Pareto-optimal solutions. In this analysis, we used the average normalized flow rate (Q^n) and the average normalized BHP (P^n) over the injection period to evaluate the overall system behavior. Both the horizontal and vertical axes have been

Table 8
Pareto-optimal solutions with the highest $M_{CO_2}{}^n$, lowest BHP^n, or highest BCR.

Parameter Case	Method	P_{hub} [MPa]	d_{pipe} [inch]	d_{tubing} [inch]	T_{wh} [°C]	$M_{CO_2}{}^n$ [fraction]	BHP^n [fraction]	BCR[fraction]
Highest $M_{CO_2}{}^n$	Base-case	10	12	5	20	1.50	0.99	0.63
	MOO	10	12	3	10	1.59	0.99	0.65
	ANN	10	15	3	10	1.61	1	0.58
	RF	11	8.5	3	11	1.61	1	0.71
	GB	10	16	3	10	1.61	1	0.59
	XGB	12	7	5.875	15	1.57	1	0.73
	Ada	10	11.5	3.5	12	1.61	1	0.66
Lowest BHP^n	MOO	9	11	2	10	1.01	0.84	0.54
	ANN	10	6	2.75	15	1.03	0.85	0.63
	RF	9	6	3	10	1.00	0.84	0.62
	GB	9	6	3.5	10	1.02	0.85	0.62
	XGB	9	6	4.875	12	1.00	0.84	0.62
	Ada	9	6	3	10	1.00	0.84	0.62
Highest BCR	MOO	12	7	4	19	1.50	0.99	0.72
	ANN	13	7	2.625	10	1.60	1	0.73
	RF	13	6	4.75	10	1.45	0.96	0.72
	GB	13	7	2.5	10	1.56	0.99	0.73
	XGB	12	7	5.875	15	1.57	1	0.73
	Ada	13	6	3.75	15	1.39	0.95	0.71

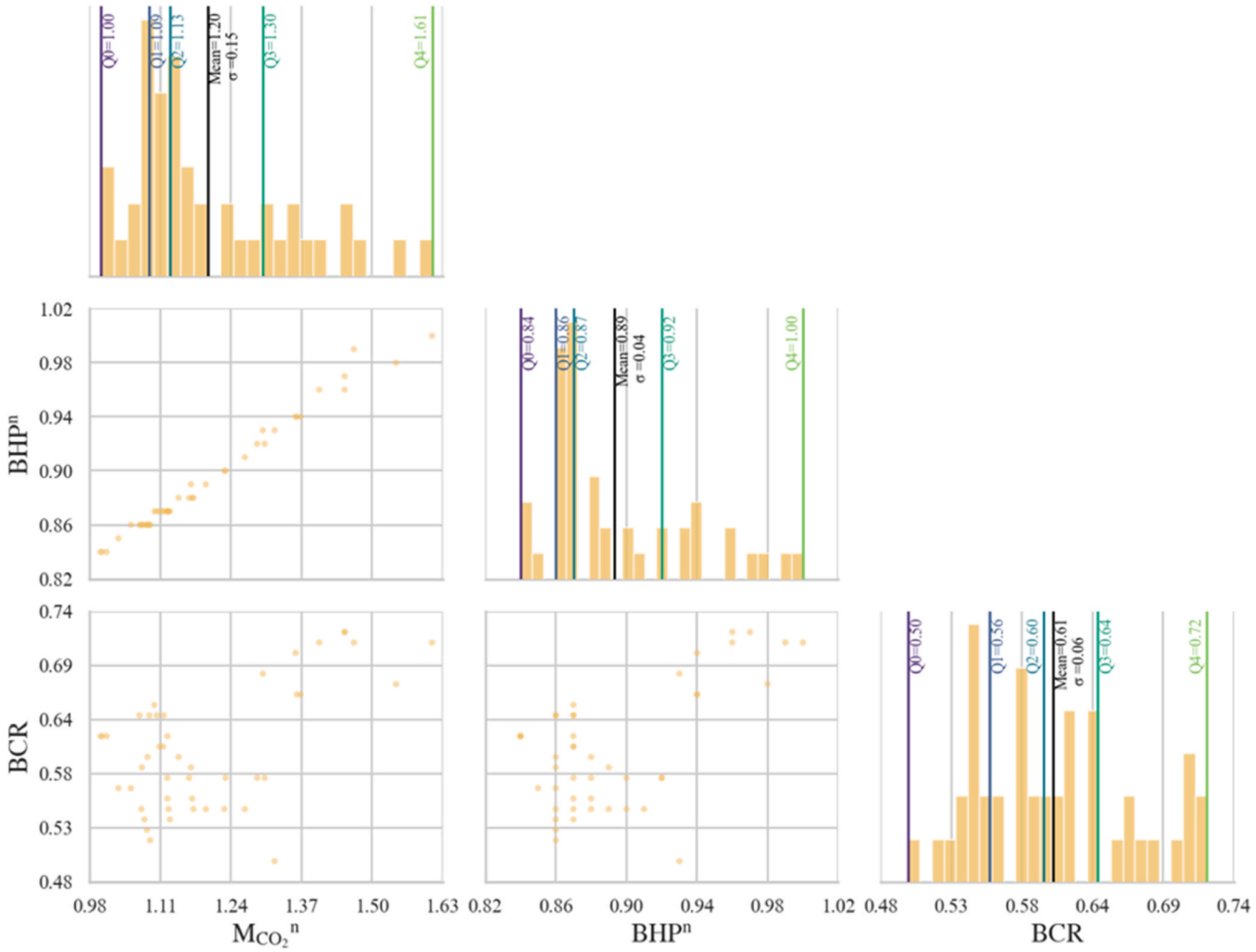


Fig. 13. Distributions and pairwise scatterplots of objective functions ($M_{CO_2}{}^n$, BHP^n, and BCR) with quartiles, mean and standard deviation, across 50 Pareto-optimal solutions derived using the random forest proxy.

normalized relative to $Q_{CO_2}^{min}$ (i.e., the annual CO_2 injection rate derived by uniformly distributing $M_{CO_2}^{min}$ over the injection period) and BHP^{max}, respectively. All 50 operating points are located within the ROI, thus indicating that all of the solutions satisfied the specified performance thresholds. On average, the solutions exhibited $Q^n = 1.18$ and $P^n = 0.88$, with ranges from 1.00 to 1.63 for Q^n and 0.81–1.00 for P^n. The most conservative solution (located in the bottom-left of the ROI) operated at 81% of BHP^{max} with $Q_{CO_2}^{min}$. Conversely, the most aggressive configuration (located at the top-right of the ROI) reached 100% of BHP^{max} with 1.63-times $Q_{CO_2}^{min}$. These results demonstrate that all of the selected solutions can maintain operational stability while meeting the constraints of $P^n \leq 1$ and $Q^n \geq 1$, thus satisfying both safety and injectivity preferences.

Fig. 16 presents flow assurance results for the CO_2 transport pipeline based on 10 representative Pareto-optimal solutions. These cases share the most frequently occurring design configuration among the 50 Pareto-optimal solutions, characterized by $P_{hub} = 9$ MPa and $d_{tubing} = 2.25$ in.. As can be seen in Fig. 16(a), the pipeline pressure profile begins at 9 MPa at the hub terminal and gradually increased to between 9.4 and 9.7 MPa at the riser inlet. This increase is primarily driven by hydrostatic head buildup along the upward-sloping section of the pipeline, which extends to approximately 150 km. Beyond this point, the pressure stabilized in the horizontal segment (150–175 km). At the 175 km mark—coinciding with the beginning of the vertical riser—there was a noticeable pressure drop attributable to the loss of hydrostatic support and increased frictional resistance from upward CO_2 flow. A larger pipeline diameter resulted in a higher pressure increase along the route. For instance, 9-inch and 12.5-inch pipelines yielded terminal pressures of approximately 9.4 MPa and 9.7 MPa, respectively. Fig. 16(b) shows that superficial CO_2 flow velocities remained in the range of 0.14–0.24 m/s across all cases, well below the erosion risk threshold of 4–5 m/s. The highest and lowest velocities (0.24 and 0.14 m/s, respectively) occurred in the smallest (9 in.) and largest (12.5 in.) diameter pipelines, respectively, thereby confirming that all of the examined configurations pose minimal erosion risk under the assessed operating conditions.

Fig. 17 provides detailed flow assurance profiles for the same representative design configuration used in the pipeline analysis in terms of pressure, temperature, fluid density, fluid viscosity, and liquid holdup, respectively, thereby offering a comprehensive assessment of flow conditions over the injection depth. Fig. 17(a) shows a linear pressure increase from 9 MPa at the wellhead to approximately 17 MPa at the bottomhole, indicating a stable pressure gradient with no abrupt changes that could induce phase transitions. Fig. 17(b) exhibits a steady increase in temperature along the wellbore, with bottomhole temperatures ranging from 20 to 30 °C, thus confirming thermal stability therein. Fig. 17(c) illustrates that the fluid density remained consistently between 850 and 920 kg/m^3, so this configuration prevented large velocity fluctuations. Similarly, fluid viscosity was confined to a narrow

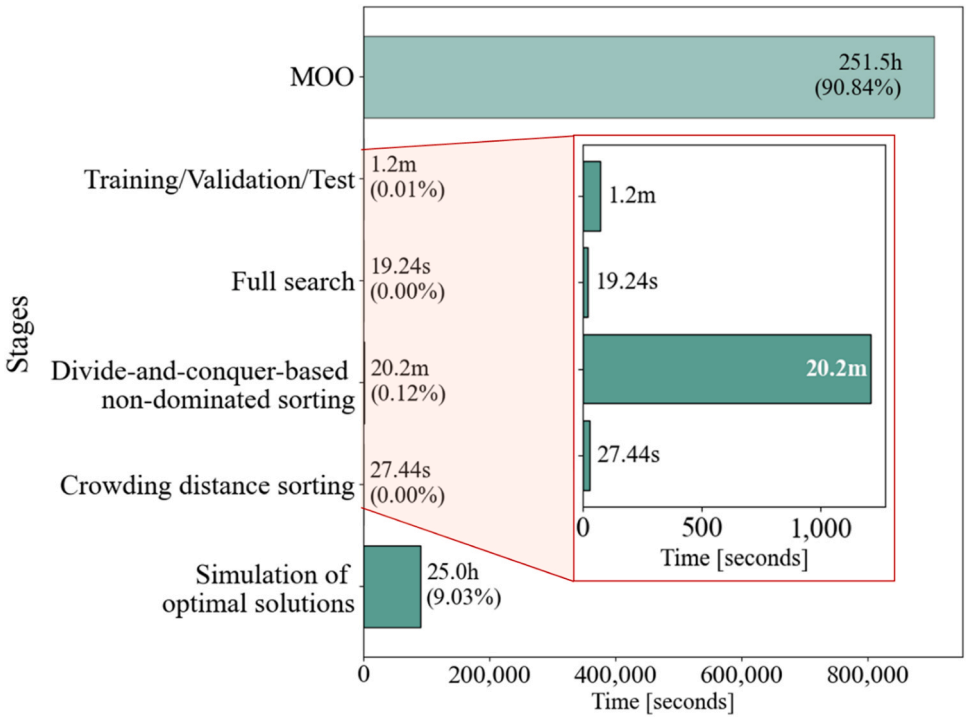


Fig. 14. Average computational times for the initial Pareto search using multi-objective optimization (MOO; light green) and the refined Pareto search using the five machine-learning proxies (green) broken down by stages. Inset (red box): an enlarged view highlighting the computing time required for each stage, with particular emphasis on the 20.2-minute refined approximation phase.

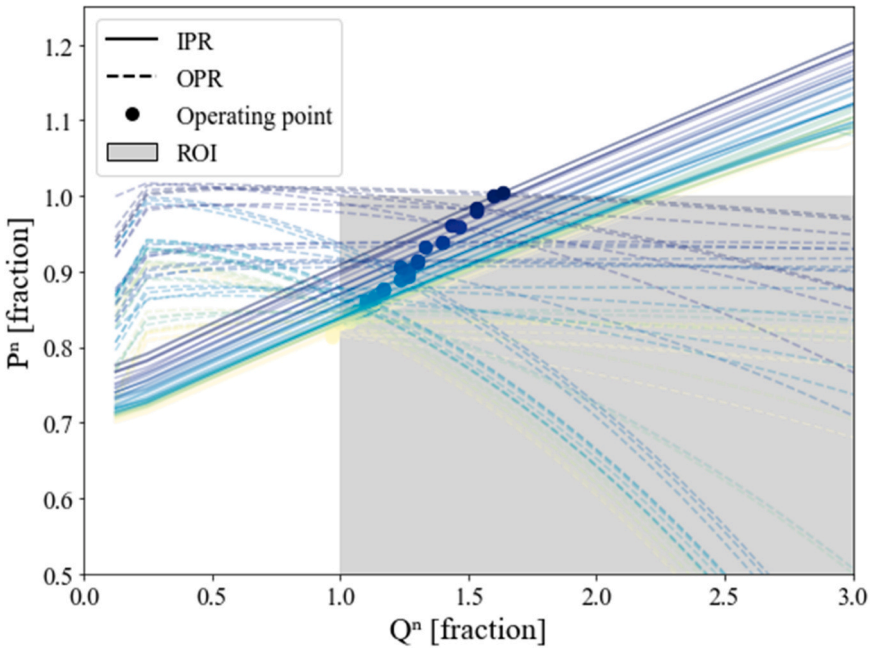


Fig. 15. Inflow Performance Relationship (IPR)-Outflow Performance Relationship (OPR) curves for 50 Pareto-optimal solutions derived using the random forest proxy. P^n, normalized pressure; Q^n, normalized flow rate; ROI, region of interest.

range of 0.07–0.08 mPa·s ((Fig. 17(d)), thus confirming that the configuration supported smooth flow behavior and minimal frictional pressure loss. Finally, Fig. 17(e) shows that liquid holdup remained fixed at 1.0 throughout the wellbore, thereby indicating single dense-phase CO_2 flow with no entrained gas or multiphase instability. Taken together, these results validate that the selected Pareto-optimal solutions can ensure operational robustness by maintaining stable single-phase CO_2 injection conditions, with smooth gradients in pressure, temperature, and flow parameters throughout the injection column.

Reservoir responses were further evaluated by tracking the spatio-temporal evolution of pressure and CO_2 saturation, as shown in Fig. 18. Two representative Pareto-optimal solutions were selected to exemplify characteristic dynamics during injection. Case 1 was defined by $P_{hub} = 9$ MPa, $d_{pipe} = 12.5$ in., $d_{tubing} = 2.25$ in., and $T_{wh} = 13$ °C, whereas Case 2 adopted identical source and tubing conditions but with a smaller pipeline diameter (12 in.) and lower injection temperature (10

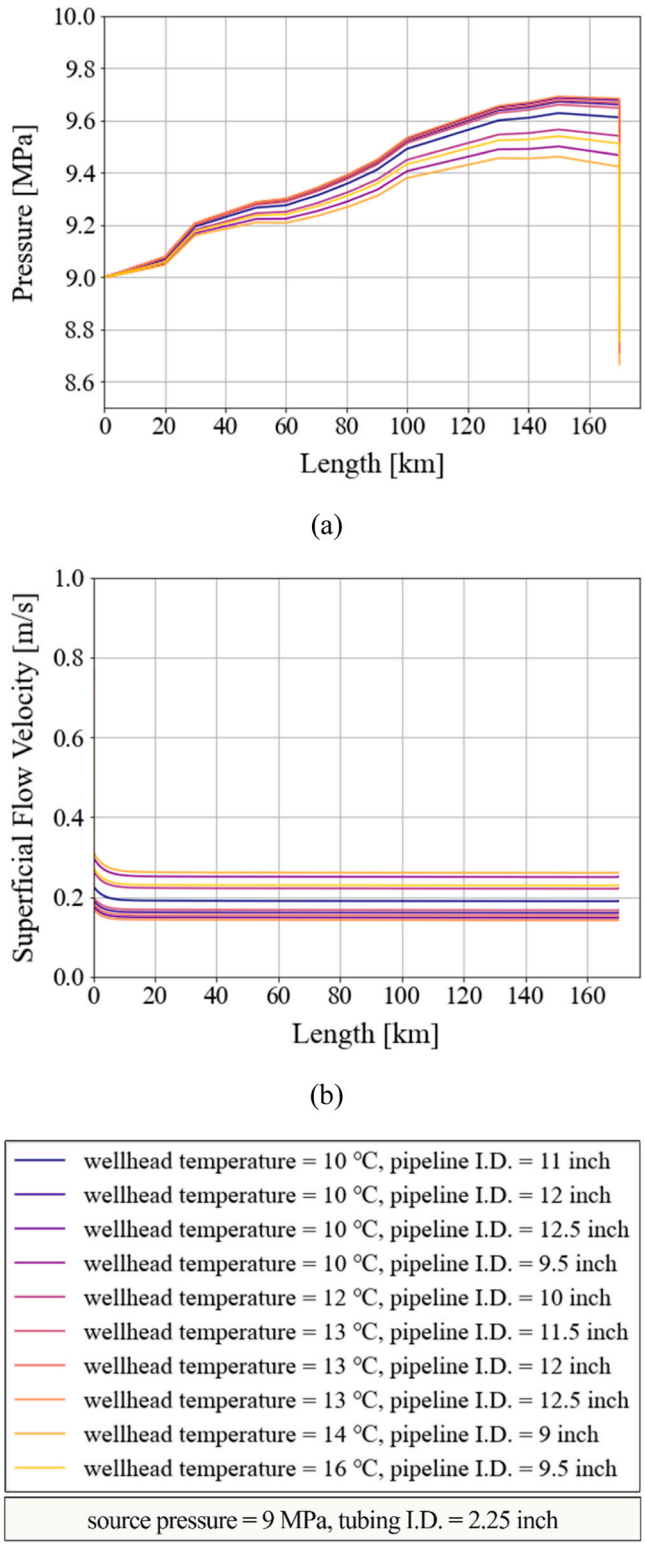


Fig. 16. Profiles for the CO_2 transport pipeline employing 10 Pareto-optimal solutions derived using the random forest proxy: (a) pressure and (b) superficial flow velocity.

°C). In both cases, pressure buildup remained below the allowable BHP^{max}, and plume migration exhibited upward and lateral spreading consistent with the reservoir structure. These results demonstrate that the proposed framework captures coupled transport–reservoir interactions, providing a quantitative basis for evaluating injection performance and storage safety.

3.4. Discussion

The proposed hybrid framework produced by integrating full-physics simulations, MOO, and ML-based refinement successfully optimized the design of a unified CCS system by balancing injectivity, pressure safety, and cost-effectiveness through Pareto-optimal exploration. The framework achieved high predictive accuracy with substantially reduced computational effort compared to full-physics simulations. Operational assessments confirmed the feasibility and stability of the derived solutions, including single-phase flow, acceptable erosion risk, and compliance with design constraints. While these results demonstrate the framework's robustness and practical value for early-stage CCS system design, several areas require further development to improve its applicability under more realistic and complex conditions.

In the current implementation, we assumed that the aquifer properties are homogeneous which simplifies modeling and lowers the computational demand. However, this assumption limits the ability to capture the spatial variability found in the geological conditions of real-world reservoirs. Advances in site characterization will provide additional well logs, core measurements, and geophysical surveys that can be leveraged to build geologically heterogeneous reservoir model that more realistically reflects stratigraphy and subsurface structure. Such models will enable the simulation of localized injectivity contrasts, preferential flow pathways, and spatially variable pressure propagation. Incorporating heterogeneity into the optimization process stage is expected to improve physical fidelity and expose nonlinear interactions that may be obscured under homogeneous assumptions. In particular, coupling stochastic geological modeling (e.g., ensembles of equiprobable realizations) with the proposed ML-assisted MOO framework allows uncertainty in petrophysical properties to be propagated into the design space, thereby supporting more robust, uncertainty-aware decisions [97–99].

Our correlation analysis results reveal a strong linear relationship between normalized CO_2 injection mass and BHP, thereby indicating that injectivity and pressure containment are tightly coupled under homogeneous conditions. This raises the possibility of reducing the number of objectives in similarly uniform systems to simplify optimization. However, in more complex reservoirs—such as those with heterogeneity, anisotropy, or multi-well interactions—these variables are likely to behave independently [100,101]. To provide generalizability and support robust field-scale decision-making, it is essential to retain all three objectives: injection performance, pressure safety, and economic viability in the modeling process.

Currently, the framework is limited to a single-well injection scenario, which is suitable for preliminary evaluations but insufficient for full-scale deployment [102]. Expanding the model to include multiple injection wells and strategic well placement would significantly improve its operational relevance. Multi-well configurations enhance reservoir sweep efficiency, accommodate geologic heterogeneity, and increase total storage capacity. Co-optimizing well locations in coordination with reservoir properties and infrastructure constraints would enable more adaptive and cost-effective system designs, which is particularly important for large-scale CCS projects. Moreover, increasing design dimensionality and the inclusion of integrated safety constraints will require more advanced optimization strategies to efficiently navigate high-dimensional trade-off spaces. The framework can also be extended to address transient operational scenarios, including injection interruption and restart, hydrate formation, and salt precipitation.

While the framework accounted for economic feasibility via BCR, the cost model assumed a deterministic carbon price of 87.5 USD/tCO_2. The resulting BCR is sensitive to carbon-price uncertainty: decreasing the carbon price to 75 USD/tCO_2 reduces BCR by 17% on average, whereas increasing it to 100 USD/tCO_2 improves BCR by 13%. Given this sensitivity, carbon-price uncertainty could be incorporated through stochastic optimization, for example, by maximizing the expected BCR over probabilistic price scenarios or through robust formulations that

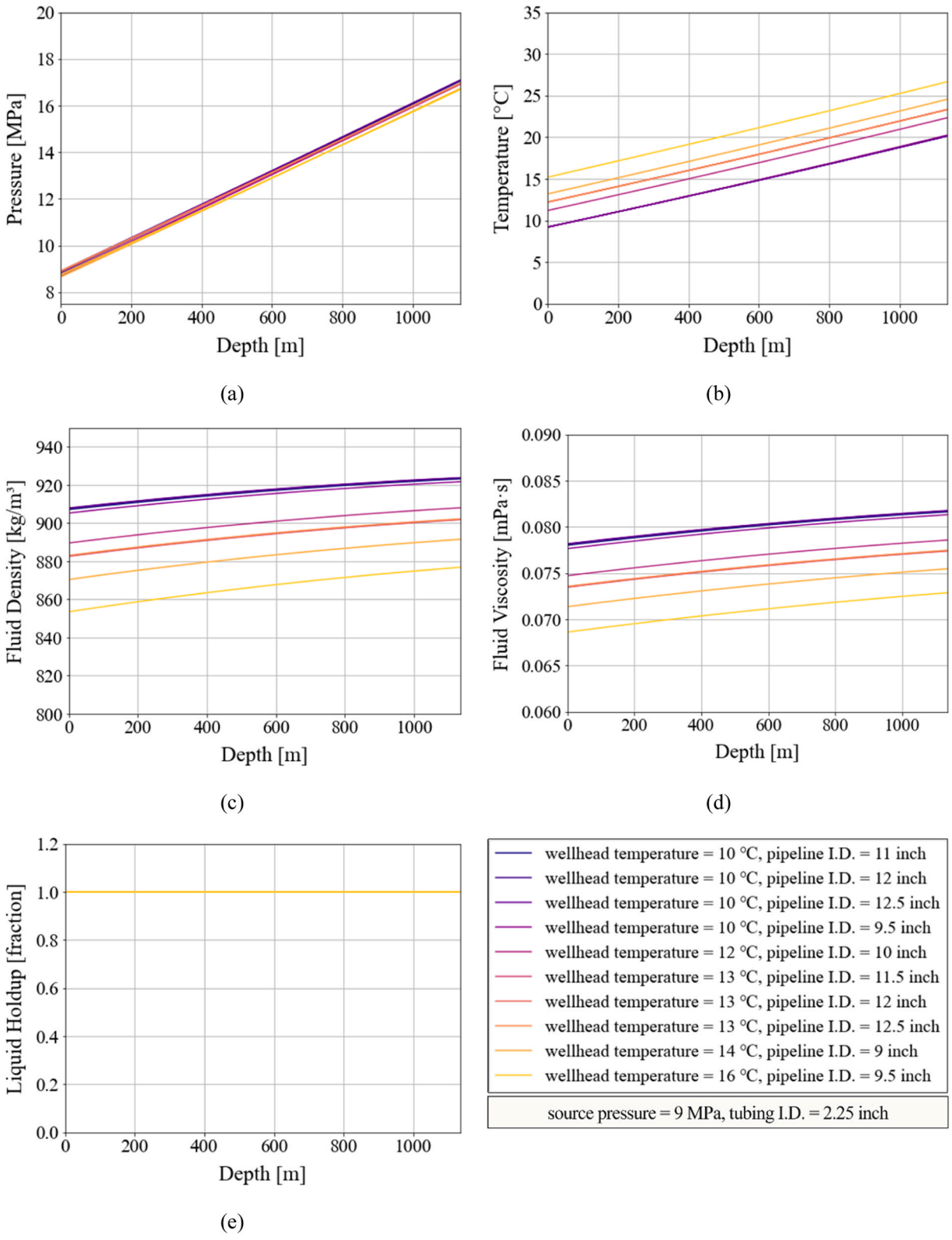


Fig. 17. Profiles for CO_2 injection well tubing utilizing 10 Pareto-optimal solutions derived using the random forest proxy: (a) pressure, (b) temperature, (c) fluid density, (d) fluid viscosity, and (e) liquid holdup.

preserve feasibility under adverse price conditions. Further extensions might also adopt a more detailed techno-economic representation, including well-specific drilling and completion costs, compression and energy requirements, long-term operating expenditure, and project-specific carbon credit pricing [103,104]. These additions would improve the accuracy and policy relevance of system-level financial assessments and enable comparisons across different project configurations and regulatory environments.

The ML proxies employed in the framework were trained on full-physics simulation data to approximate system performance over a

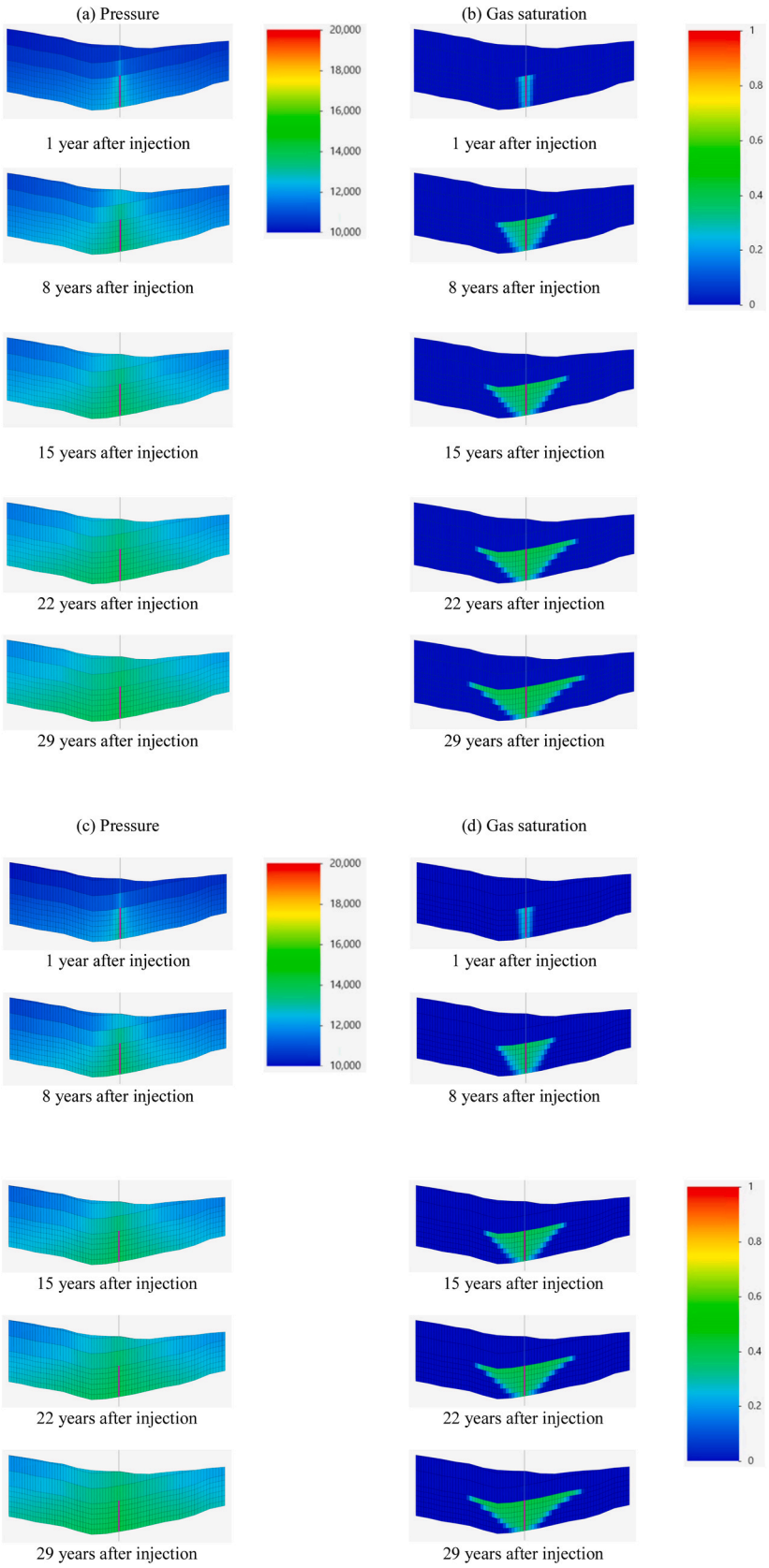


Fig. 18. Reservoir responses under two Pareto-optimal solutions; (a, b) Case 1–P_{hub}= 9 MPa, d_{pipe}= 12.5 in, d_{tubing}= 2.25 in, T_{wh}= 13℃; (c, d) Case 2–P_{hub}= 9 MPa, d_{pipe}= 12 in, d_{tubing}= 2.25 in, T_{wh}= 10℃. Subplots show pressure distribution (a, c) and gas saturation (b, d).

high-resolution design space. Although this approach proved effective, we intend to explore alternative surrogate modeling strategies to further enhance scalability and adaptability. These may include reduced-order models derived from physics-based simplifications [105], physics-informed neural networks that embed governing equations into learning architectures [106], or ensemble models that combine multiple surrogate types to improve robustness [107]. Such techniques would reduce dependence on costly simulations and enable more efficient optimization of complex high-dimensional CCS systems.

In the present study, decision-maker preferences were incorporated after the initial design space exploration rather than embedded directly in the MOO algorithm. This posterior filtering approach enabled broad unbiased exploration, followed by targeted refinement to align with the decision-maker goals. While this structure offers flexibility, future studies could be conducted to investigate hybrid workflows that integrate preferences earlier in the process. Embedding preference information through soft constraints [108], utility functions [109], or objective weighting [110] could reduce the number of unnecessary evaluations and improve convergence, particularly in large-scale or real-time design environments.

4. Conclusions

We built an ML-assisted MOO framework for the integrated design of a unified CCS system by combining dynamic nodal analysis, full-physics simulation, and data-driven surrogate modeling. The framework was developed to optimize key design variables—including CO_2 discharge pressure, pipeline inner diameter, injection tubing diameter, and injection temperature—while balancing storage capacity, injection safety, and economic feasibility. By integrating physically based simulation with data-driven acceleration, the framework could provide scalable and efficient system design under complex, competing objectives. In particular, by incorporating BCR together with injection performance and pressure safety, the proposed framework offers a robust and integrative basis for guiding CCS system design and operation. The framework was applied to a potential CCS assessment project targeting a deep saline aquifer in Basin A, Republic of Korea, which includes an onshore CO_2 terminal, a 175-km subsea pipeline, and a 1-km vertical injection well. The initial MOPSO process was used to evaluate 3.8% of 13,377 feasible designs, from which 46 Pareto-optimal solutions that exhibited strong inter-objective correlations (0.83–1.0) were identified within the ROI. To further expand the design space, ML-based proxies—ANN, RF, GB, XGB, and Ada —were trained on the simulation results and used to explore 180,075 refined solutions. These surrogate models achieved high predictive accuracy ($R^2 > 0.96$; MAPE $\leq$ 5%) and showed strong agreement with full-physics simulation outputs. Computation of the ML-assisted optimization process was completed in approximately 22 min, thus demonstrating efficient evaluation of high-resolution design alternatives within a reasonable computation time. Taken together, these results indicate that the divide-and-conquer-based non-dominated sorting algorithm, while offering limited benefit in conventional MOO problems with only hundreds of solutions, was essential in the ML refinement stage where the number of solutions increased by several orders of magnitude. Its adoption not only removed the computational bottleneck but also enhanced the practicability of the proposed MOO–ML hybrid framework for large-scale CCS system design. The refined POF preserved the key optimization characteristics, with convergence metric values comparable to those derived from the full-physics-based MOO ($\Delta_{\mathscr{P}^*}$ = 0.552 vs. 0.540) and diversity (HV^n = 0.446 vs. 0.457). The results of dynamic nodal and flow assurance analyses confirm the operational feasibility of 50 representative solutions that can maintain stable dense-phase CO_2 transport and injection without phase transitions or erosional risks. In summary, the proposed hybrid framework provides a computationally efficient, accurate, and scalable approach for early-stage CCS system design. Its successful application to the Basin A case study demonstrates its practicability and highlights its potential for broader applicability to future CCS deployment and other geoenergy projects requiring integrated system optimization.

Nomenclatures

Variables
C Monetary value
d Inner diameter
d_k Crowding distance
$\Delta_{\mathscr{P}^*}$ Averaged Hausdorff distance
f_i Objective function
HV^n Hypervolume per solution
L Length
M_{co_2} Cumulative CO_2 injection mass
N Number
O Set of obtained non-dominated solutions
P Pressure
$\mathscr{P}^*$ True Pareto-optimal front
Q Flow rate
R Correlation coefficient
R^2 Coefficient of determination
r Reference point
r_d Annual discount rate
T_{wh} CO_2 temperature at the injection wellhead
t_m Monthly time index
$\mathbf{x}$ Set of design variables
ρ Density
σ Standard deviation
Superscripts
max Maximum
min Minimum
n Normalized
Subscripts
gen Generation
hub CO_2 hub terminal
obj Objective function
pipe Subsea CO_2 transport pipeline
pop Population
tubing Injection well tubing
var Variable
Abbreviations
Ada Adaptive boost
ANN Artificial neural network
API American Petroleum Institute
BCR Benefit-cost ratio
BHP Bottomhole pressure
CCS Carbon capture and storage
GB Gradient boosting
IPR Inflow performance relationship
MAPE Mean absolute percentage error
ML Machine-learning
MOO Multi-objective optimization
OPR Outflow performance relationship
POF Pareto-optimal front
RF Random forest
ROI Region of interest
RMSE Root mean squared error
XGB Extreme gradient boosting

CRediT authorship contribution statement

Baehyun Min: Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis. **Tea-Woo Kim:** Writing – review & editing, Supervision, Project

administration, Methodology, Conceptualization. **Soeun Yoon:** Validation. **Minkyung Chae:** Writing – review & editing, Validation, Software. **Suin Choi:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Honggeun Jo:** Validation. **Byungin Ian Choi:** Writing – review & editing, Validation, Resources. **Suryeom Jo:** Writing – review & editing, Validation.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (No. RS-2023–00208297) and the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Ministry of Climate, Energy and Environment (MCEE) of the Republic of Korea (No. RS-2025–16063236). Tea-Woo Kim and Suryeom Jo were partially supported by the Basic Research Project (GP2025–021) at the Korea Institute of Geoscience and Mineral Resources (KIGAM), funded by the Ministry of Science and ICT.

Declaration of Competing Interest

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

Acknowledgments

Suin Choi, Soeun Yoon, and Baehyun Min are grateful to the Weather and Climate Data Convergence Analysis Specialized Graduate School supported by the Korea Meteorological Institute. We would like to thank the Computer Modelling Group Ltd. (CMG) for providing its reservoir simulation software package.

Data availability

The data that has been used is confidential.

References

[1] International Energy Agency. (2023). World energy outlook 2023. ⟨https://www.iea.org/reports/world-energy-outlook-2023⟩.
[2] Global CCS Institute. (2024). Global status of CCS 2024. ⟨https://www.globalccsinstitute.com/resources/global-status-report/⟩.
[3] J. You, W. Ampomah, Q. Sun, Development and application of a machine learning based multi-objective optimization workflow for CO_2-EOR projects, Fuel 264 (2020) 116758, https://doi.org/10.1016/j.fuel.2019.116758.
[4] B. Li, H. Yu, W. Xu, H. Huang, M. Huang, S. Meng, et al., A multi-physics coupled multi-scale transport model for CO_2 sequestration and enhanced recovery in shale formation with fractal fracture networks, Energy 284 (2023) 129285, https://doi.org/10.1016/j.energy.2023.129285.
[5] R.S. Middleton, J.M. Bielicki, A comprehensive carbon capture and storage infrastructure model, Energy Procedia 1 (1) (2009) 1611–1616, https://doi.org/10.1016/j.egypro.2009.01.211.
[6] S. Zhang, Y. Zhuang, R. Tao, L. Liu, L. Zhang, J. Du, Multi-objective optimization for the deployment of carbon capture utilization and storage supply chain considering economic and environmental performance, J. Clean. Prod. 270 (2020) 122481. ⟨https://10.1016/j.jclepro.2020.122481⟩.
[7] S. Ding, F. Wen, N. Wang, Y. Zhang, R. Lu, Y. Gao, et al., Multi-objective optimization of CO_2 enhanced oil recovery and storage processes in low permeability reservoirs, Int. J. Greenh. Gas. Control 121 (2022) 103802, https://doi.org/10.1016/j.ijggc.2022.103802.
[8] N. Van Nguyen, V. Pirouzfar, H. Soheilinezhad, C.H. Su, Optimizing the CO_2 capture and removal process to recover energy and CO_2 from the flue gas of boilers and gas turbines outlet with considering the techno-economic analysis, Energy 291 (2024) 130281, https://doi.org/10.1016/j.energy.2024.130281.
[9] T. Hennequin, D. Roberts, R. van Zelm, S. Brown, R.T. Porter, H. Mahgerefteh, et al., Tiered multi-objective optimization of CO_2 transport for CCS. Proceedings of the 17th Greenhouse Gas Control Technologies Conference (GHGT-17), SSRN. 5015647, 2024, https://doi.org/10.2139/ssrn.5015647.
[10] J. Liu, J. Liu, Y. Zhu, W. Sun, D. Zhang, H. Pan, Multi-objective optimization for efficient CO_2 storage under pressure buildup constraint in saline aquifer, Appl. Energy 382 (2025) 125175, https://doi.org/10.1016/j.apenergy.2024.125175.
[11] K.E. Brown, J.F. Lea, Nodal systems analysis of oil and gas wells, J. Pet. Technol. 37 (10) (1985) 1751–1763, https://doi.org/10.2118/14714-PA.
[12] I. Salaudeen, D. Bopbekov, A. Abdulkarim, Optimization of petroleum production system using nodal analysis program, Niger. J. Technol. Dev. 19 (1) (2022) 1–8, https://doi.org/10.4314/njtd.v19i1.1.
[13] A.U. Yadua, K.A. Lawal, A new method to evaluate integrated production system capacity in oil fields, Pet. Explor. Dev. 50 (3) (2023) 665–674, https://doi.org/10.1016/S1876-3804(23)60418-5.
[14] J.M. Vissers, A. Alkindi, A. Hassan, The integration of miscible gas injection in oil reservoirs with the development of a sour gas condensate field (Paper No. SPE-154992-MS), SPE EOR Conf. Oil Gas. West Asia (2012, March), https://doi.org/10.2118/154992-MS.
[15] F. Sierra, A. Monge, S. Leon, F. Vasquez, J. Garcia, G. Llerena, et al., A real case study: field development plan evaluations for Oso field (Paper No. SPE-177071-MS), SPE Lat. Am. Caribb. Pet. Eng. Conf. (2015, November), https://doi.org/10.2118/177071-MS.
[16] V. Pathak, C. Palaschak, O. Martinez, J. Hudson, J. Ita, F. Saaf, et al., Solving the challenges of short- and long-term production forecasting and uncertainty using a fully coupled implicit integrated production modelling system (Paper No. SPE-181427-MS). SPE Annual Technical Conference and Exhibition, 2016, September, https://doi.org/10.2118/181427-MS.
[17] V. Pathak, C. Palaschak, A. Kumar, R. Eastick, R. Schulz, A. Al-Hadhrami, Coupled multi-reservoir multi-network production optimization of an onshore Middle-Eastern reservoir with sour gas reinjection for miscible EOR (Paper No. SPE-188377-MS). Abu Dhabi International Petroleum Exhibition and Conference, 2017, November, https://doi.org/10.2118/188377-MS.
[18] V. Pathak, Y. Hamedi, O. Martinez, S. Vermeulen, A. Kumar, Quantifying geological uncertainty for complex integrated production systems with multiple reservoirs and production networks (Paper No. SPE-195477-MS). SPE Europec featured at the 81st EAGE Conference and Exhibition, 2019, June, https://doi.org/10.2118/195477-MS.
[19] K. Kumar, V. Pathak, P. Agrawal, Z. Alias, T. Narwal, A. Al-Hadhrami, et al., Integrated production optimization using implicitly coupled multi-reservoirs and facility modeling with miscible gas injection EOR in South Oman (Paper No. SPE-203151-MS). SPE Abu Dhabi International Petroleum Exhibition and Conference, 2020, November, https://doi.org/10.2118/203151-MS.
[20] Y. Hamedi Shokrlu, V. Pathak, A. Kumar, V. Salazar, Integrated production system modelling and optimization for advanced EOR application in pre-salt offshore carbonate reservoir (Paper No. OTC-30713-MS). Offshore Technology Conference, 2020, May, https://doi.org/10.4043/30713-MS.
[21] M. Abdellatif, M. Hashemi, S. Azizmohammadi, Large-scale underground hydrogen storage: integrated modeling of reservoir-wellbore system, Int. J. Hydrog. Energy 48 (50) (2023) 19160–19171, https://doi.org/10.1016/j.ijhydene.2023.01.227.
[22] N. Surin, K.C. Kok, N. Hutamin, S.K. Rofieq, S.M. MousaviMirkalaei, G. Syah, et al., Application of integrated production system modelling (IPSM) for long-term production forecasting and optimization: A case study in deepwater assets, Malaysia (Paper No. SPE-221169-MS). APOGCE 2024: Asia Pacific Oil and Gas Conference and Exhibition, 2024, October, https://doi.org/10.2118/221169-MS.
[23] F. Kazemifar, D.C. Kyritsis, Near-critical CO_2 flow measurement and visualization, J. Energy Res. Technol. 137 (1) (2015) 012002, https://doi.org/10.1115/1.4027961.
[24] F. Bai, W. Lou, Y. Lu, A robust drift-flux model for two-phase CO_2 pipe flow, Gas. Sci. Eng. 119 (2023) 205140, https://doi.org/10.1016/j.jgsce.2023.205140.
[25] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2) (2002, April) 182–197, https://doi.org/10.1109/4235.996017.
[26] K. Deb, H. Gupta, Searching for robust Pareto-optimal solutions in multi-objective optimization, in: C.A. Coello Coello, A. Hernández Aguirre, E. Zitzler (Eds.), Evolutionary multi-criterion optimization. EMO 2005. Lecture notes in computer science, 3410, Springer, 2005, pp. 150–164, https://doi.org/10.1007/978-3-540-31880-4_11.
[27] B. Min, J.M. Kang, S. Chung, C. Park, I. Jang, Pareto-based multi-objective history matching with respect to individual production performance in a heterogeneous reservoir, J. Pet. Sci. Eng. 122 (2014) 551–566, https://doi.org/10.1016/j.petrol.2014.08.023.
[28] B. Min, C. Park, I. Jang, J.M. Kang, S. Chung, Development of Pareto-based evolutionary model integrated with dynamic goal programming and successive linear objective reduction, Appl. Soft Comput. 35 (2015) 75–112, https://doi.org/10.1016/j.asoc.2015.06.007.
[29] S. Tang, Z. Cai, J. Zheng, A fast method of constructing the non-dominated set: Arena's principle. Proceedings of the 2008 Fourth International Conference on Natural Computation, IEEE, 2008, October, pp. 391–395, https://doi.org/10.1109/ICNC.2008.823.
[30] H. Fang, Q. Wang, Y.C. Tu, M.F. Horstemeyer, An efficient non-dominated sorting method for evolutionary algorithms, Evol. Comput. 16 (3) (2008) 355–384, https://doi.org/10.1162/evco.2008.16.3.355.
[31] S. Mishra, S. Saha, S. Mondal, Divide and conquer based non-dominated sorting for parallel environment. 2016 IEEE Congress on Evol. Comput, IEEE, 2016, July, pp. 4297–4304, https://doi.org/10.1109/CEC.2016.7744336.
[32] S. Mishra, S. Saha, S. Mondal, C.A.C. Coello, A divide-and-conquer based efficient non-dominated sorting approach, Swarm Evol. Comput. 44 (2019) 748–773, https://doi.org/10.1016/j.swevo.2018.08.011.
[33] H. Wang, M. Olhofer, Y. Jin, A mini-review on preference modeling and articulation in multi-objective optimization: current status and challenges,

Complex Intell. Syst. 3 (4) (2017) 233–245, https://doi.org/10.1007/s40747-017-0053-9.
[34] A.R. Adekoya, M. Helbig, Decision-maker's preference-driven dynamic multi-objective optimization, Algorithms 16 (11) (2023) 504, https://doi.org/10.3390/a16110504.
[35] R. Ozaki, K. Ishikawa, Y. Kanzaki, S. Takeno, I. Takeuchi, M. Karasuyama, Multi-objective Bayesian optimization with active preference learning, in: Proceedings of the AAAI Conference on Artificial Intelligence, 38, 2024, pp. 14490–14498, https://doi.org/10.1609/aaai.v38i13.29364.
[36] B. Min, A.Y. Sun, M.F. Wheeler, H. Jeong, Utilization of multiobjective optimization for pulse testing dataset from a CO_2-EOR/sequestration field, J. Pet. Sci. Eng. 170 (2018) 244–266, https://doi.org/10.1016/j.petrol.2018.06.035.
[37] M. Azizi, S.M. Hasheminezhad, S. Moeinpour, M. Kanaani, B. Sedaee, Multi-objective optimization of CO_2 injection process into oil reservoirs using machine learning algorithms: incorporating carbon sequestration mechanisms, Energy Fuels 38 (19) (2024) 18814–18830, https://doi.org/10.1021/acs.energyfuels.4c02859.
[38] A. Sircar, K. Yadav, K. Rayavarapu, N. Bist, H. Oza, Application of machine learning and artificial intelligence in oil and gas industry, Pet. Res. 6 (4) (2021) 379–391, https://doi.org/10.1016/j.ptlrs.2021.05.009.
[39] Z. Tariq, M.S. Aljawad, A. Hasan, M. Murtaza, E. Mohammed, A. El-Husseiny, et al., A systematic review of data science and machine learning applications to the oil and gas industry, J. Pet. Explor. Prod. Technol. (2021) 1–36, https://doi.org/10.1007/s13202-021-01302-2.
[40] M.V.B. Machado, A. Khanal, M. Delshad, Unveiling the essential parameters driving mineral reactions during CO_2 storage in carbonate aquifers through proxy models, Appl. Sci. 14 (4) (2024) 1465, https://doi.org/10.3390/app14041465.
[41] S. Kim, T.W. Kim, S. Jo, Artificial intelligence in geoenergy: bridging petroleum engineering and future-oriented applications, J. Pet. Explor. Prod. Technol. 15 (2) (2025) 35, https://doi.org/10.1007/s13202-025-01939-3.
[42] I. Pan, M. Babaei, A. Korre, S. Durucan, Artificial neural network based surrogate modelling for multi-objective optimisation of geological CO_2 storage operations, Energy Procedia 63 (2014) 3483–3491, https://doi.org/10.1016/j.egypro.2014.11.377.
[43] J. You, W. Ampomah, Q. Sun, Development and application of a machine learning based multi-objective optimization workflow for CO_2-EOR projects, Fuel 264 (2020) 116758, https://doi.org/10.1016/j.fuel.2019.116758.
[44] C. Park, J. Oh, S. Jo, I. Jang, K.S. Lee, Multi-objective optimization of CO_2 sequestration in heterogeneous saline aquifers under geological uncertainty, Appl. Sci. 11 (20) (2021) 9759, https://doi.org/10.3390/app11209759.
[45] M. Nassabeh, Z. You, A. Keshavarz, S. Iglauer, Sub-surface geospatial intelligence in carbon capture, utilization and storage: a machine learning approach for offshore storage site selection, Energy 305 (2024) 132086, https://doi.org/10.1016/j.energy.2024.132086.
[46] F. Fathalian, S. Aarabi, A. Ghaemi, A. Hemmati, Intelligent prediction models based on machine learning for CO_2 capture performance by graphene oxide-based adsorbents, Sci. Rep. 12 (1) (2022) 21507, https://doi.org/10.1038/s41598-022-26138-6.
[47] T.H. Vo, D.S. Dashtgoli, H. Zhang, B. Min, Machine-learning-based prediction of oil recovery factor for experimental CO_2-foam chemical EOR: implications for carbon utilization projects, Energy 278 (2023) 127860, https://doi.org/10.1016/j.energy.2023.127860.
[48] T.H. Vo, H. Zhang, Z. Dai, T. Zhang, S. Tangparitkul, B. Min, Data-driven machine learning models for the prediction of hydrogen solubility in aqueous systems of varying salinity: implications for underground hydrogen storage, Int. J. Hydrog. Energy 55 (2024) 1422–1433, https://doi.org/10.1016/j.ijhydene.2023.12.131.
[49] Z. Tariq, M. Ali, A. Hassanpouryouzband, B. Yan, S. Sun, H. Hoteit, Predicting wettability of mineral/CO_2/brine systems via data-driven machine learning modeling: implications for carbon geo-sequestration, Chemosphere 345 (2023) 140469, https://doi.org/10.1016/j.chemosphere.2023.140469.
[50] C. Wu, L. Jin, J. Zhao, X. Wan, T. Jiang, K. Ling, Determination of gas–oil minimum miscibility pressure for impure CO_2 through optimized machine learning models, Geoenergy Sci. Eng. 242 (2024) 213216, https://doi.org/10.1016/j.geoen.2024.213216.
[51] Younis, A.A.H. (2010). Space exploration and region elimination global optimization algorithms for multidisciplinary design optimization (Doctoral dissertation).
[52] A. Younis, Z. Dong, Global optimization using mixed surrogates and space elimination in computationally intensive engineering designs, Int. J. Comput. Methods Eng. Sci. Mech. 13 (4) (2012) 272–289, https://doi.org/10.1080/15502287.2012.682196.
[53] A. Younis, Z. Dong, High-fidelity surrogate based multi-objective optimization algorithm, Algorithms 15 (8) (2022) 279, https://doi.org/10.3390/a15080279.
[54] A. Younis, Z. Dong, Adaptive surrogate assisted multi-objective optimization approach for highly nonlinear and complex engineering design problems, Appl. Soft Comput. 150 (2024) 111065, https://doi.org/10.1016/j.asoc.2023.111065.
[55] S. Kwon, M. Ji, M. Kim, J.Y. Leung, B. Min, Determination of sequential well placements using a multi-modal convolutional neural network integrated with evolutionary optimization, Math 13 (1) (2024) 36, https://doi.org/10.3390/math13010036.
[56] J.E. Smith, K.F. McCardle, Options in the real world: lessons learned in evaluating oil and gas investments, Oper. Res. 47 (1) (1999) 1–15, https://doi.org/10.1287/opre.47.1.1.
[57] X. Zhang, X. Wang, J. Chen, X. Xie, K. Wang, Y. Wei, A novel modeling based real option approach for CCS investment evaluation under multiple uncertainties, Appl. Energy 113 (2014) 1059–1067, https://doi.org/10.1016/j.apenergy.2013.08.047.
[58] M.M. Rajabi, M. Chen, M.R.H. Javaran, A. Al-Maktoumi, A. Izady, Y. Dong, Probabilistic net present value analysis for designing techno-economically optimal sequential CO_2 sequestration and geothermal energy extraction, J. Hydrol. 612 (2022) 128237, https://doi.org/10.1016/j.jhydrol.2022.128237.
[59] D.C. Major, Benefit–cost ratios for projects in multiple objective investment programs, Water Resour. Res. 5 (6) (1969) 1174–1178, https://doi.org/10.1029/WR005i006p01174.
[60] B. Wang, X. Xia, J. Zhang, A multi-objective optimization model for the life-cycle cost analysis and retrofitting planning of buildings, Energy Build. 77 (2014) 227–235, https://doi.org/10.1016/j.enbuild.2014.03.025.
[61] C. Zhou, B. Zhang, C. Mu, Z. Chu, L. Sun, Multi-objective optimization considering cost–benefit ratio for the placement of gas detectors in oil refinery installations, J. Loss Prev. Process Ind. 62 (2019) 103956, https://doi.org/10.1016/j.jlp.2019.103956.
[62] R. Rathore, B.R. Bakshi, H. Kodamana, M. Ramteke, Multi-objective optimization of crop land allocation for sustainability within the FEWLF nexus, Comput. Chem. Eng. 201 (2025) 10, https://doi.org/10.1016/j.compchemeng.2025.109207.
[63] A.B. Bitsindou, M.G. Kelkar, Gas well production optimization using dynamic nodal analysis (Paper No. SPE-52170-MS), SPE Cont. Oper. Symp. (1999, March), https://doi.org/10.2118/52170-MS.
[64] M.H. Schindler, Dynamic nodal analysis in well testing interpretation (Paper No. SPE-107239-MS). SPE Latin American & Caribbean Petroleum Engineering Conference, 2007, April, https://doi.org/10.2118/107239-MS.
[65] N. Srinivas, K. Deb, Muiltiobjective optimization using nondominated sorting in genetic algorithms, Evol. Comput. 2 (3) (1994, September) 221–248, https://doi.org/10.1162/evco.1994.2.3.221.
[66] N. Riquelme, C. Von Lücken, B. Baran, Performance metrics in multi-objective optimization. 2015 Latin American Computing Conference, IEEE, 2015, October, pp. 1–11, https://doi.org/10.1109/CLEI.2015.7360024.
[67] O. Schutze, X. Esquivel, A. Lara, C.A.C. Coello, Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization, IEEE Trans. Evol. Comput. 16 (4) (2012) 504–522, https://doi.org/10.1109/TEVC.2011.2161872.
[68] E.J. Hughes, Evolutionary many-objective optimisation: many once or one many?. 2005 IEEE Congress on Evol. Comput IEEE, 2005, September, pp. 222–227, https://doi.org/10.1109/CEC.2005.1554688.
[69] F. Ascione, N. Bianco, G.M. Mauro, G.P. Vanoli, A new comprehensive framework for the multi-objective optimization of building energy design: Harlequin, Appl. Energy 241 (2019) 331–361, https://doi.org/10.1016/j.apenergy.2019.03.028.
[70] Z. Liu, Z. Cui, M. Wang, B. Liu, W. Tian, A machine learning proxy based multi-objective optimization method for low-carbon hydrogen production, J. Clean. Prod. 445 (2024) 141377, https://doi.org/10.1016/j.jclepro.2024.141377.
[71] Y.K. Kwon, Y.J. Shinn, Suggestion for technology development and commercialization strategy of CO_2 capture and storage in Korea, Econ. Environ. Geol. 51 (4) (2018) 381–392, https://doi.org/10.9719/EEG.2018.51.4.381.
[72] K.J. Kim, M. Kang, S. Cheong, K.Y. Kim, I. Song, 2024, Comprehensive site screening and selection for large-scale carbon storage in the early Miocene formations of the Eastern Gunsan Basin, Yellow Sea, offshore Korea.
[73] Computer Modelling Group Ltd. (2024). CoFlow (Version 2024.1) [Computer software]. ⟨https://www.cmgl.ca⟩.
[74] GEBCO. (2024, July 30). Gridded bathymetry data. ⟨https://www.gebco.net/data_and_products/gridded_bathymetry_data/⟩.
[75] M.M.J. Knoope, W. Guijt, A. Ramírez, A.P.C. Faaij, Improved cost models for optimizing CO_2 pipeline configuration for point-to-point pipelines and simple networks, Int. J. Greenh. Gas. Control 22 (2014) 25–46, https://doi.org/10.1016/j.ijggc.2013.12.016.
[76] T.W. Kim, K.J. Kim, Y.K. Lee, S. Jo, S. Choi, B. Min, B.I. Choi, Optimizing transportation and storage design for CO_2 geological sequestration using multiobjective optimization and nodal analysis: a case study from the Gunsan basin, South Korea, Int. J. Energy Res. 2025 (1) (2025) 6686996, https://doi.org/10.1155/er/6686996.
[77] Z. Chen, Y. Zhou, H. Li, A review of phase behavior mechanisms of CO_2 EOR and storage in subsurface formations, Ind. Eng. Chem. Res. 61 (29) (2022) 10298–10318, https://doi.org/10.1021/acs.iecr.2c00204.
[78] US Code of Federal Regulations. (n.d.). Injection well operating requirements. ⟨https://www.ecfr.gov/current/title-40/chapter-I/subchapter-D/part-146/subpart-H/section-146.88⟩.
[79] International Energy Agency. (2018, July 9). The importance of real-world policy packages to drive energy transitions. ⟨https://www.iea.org/commentaries/the-importance-of-real-world-policy-packages-to-drive-energy-transitions⟩.
[80] W.J. Schmelz, G. Hochman, K.G. Miller, Total cost of carbon capture and storage implemented at a regional scale: northeastern and midwestern United States, Interface Focus 10 (5) (2020) 20190065, https://doi.org/10.1098/rsfs.2019.0065.
[81] Parker, N. (2004). Using natural gas transmission pipeline costs to estimate hydrogen pipeline costs. (Report No. UCD-ITS-RR-04-35). Institute of Transportation Studies, University of California, Davis. ⟨https://escholarship.org/uc/item/9c06n11x⟩.
[82] Organisation for Economic Co-operation and Development. (n.d.). Consumer price indices – Data explorer. ⟨https://data-explorer.oecd.org/vis?df[ds]=DisseminateFinalDMZanddf[id]=DSD_PRICES%40DF_PRICES_ALLanddf[ag]=OECD.SDD.TPSanddq=.M.N.CPI.PA._T.N.GYandlom=LASTNPERIODSandlo=13andto[TIME_PERIOD]=false⟩.

[83] Kim, H. (2015, October). Understanding and application of AHP analysis in preliminary feasibility studies. 11th Public Investment Research Forum, Chungnam Institute, Korea.
[84] Korea Institute of S&T Evaluation and Planning. (n.d.). Publication – Preliminary Feasibility Study. ⟨https://www.kistep.re.kr/reportList.es?mid=a10305070000&rpt_tp=831-003⟩.
[85] S. Lalwani, S. Singhal, R. Kumar, N. Gupta, A comprehensive survey: applications of multi-objective particle swarm optimization (MOPSO) algorithm, Trans. Comb. 2 (1) (2013) 39–101, https://doi.org/10.22108/toc.2013.2834.
[86] W. Zhao, Z. Luan, C. Wang, Parameter optimization design of vehicle E-HHPS system based on an improved MOPSO algorithm, Adv. Eng. Softw. 123 (2018) 51–61, https://doi.org/10.1016/j.advengsoft.2018.05.011.
[87] Z. Tian, Z. Zhang, K. Zhang, X. Tang, S. Huang, Statistical modeling and multi-objective optimization of road geopolymer grouting material via RSM and MOPSO, Constr. Build. Mater. 271 (2021) 121534, https://doi.org/10.1016/j.conbuildmat.2020.121534.
[88] Trident Steel. (n.d.). API 5L Pipe: What is it?, ⟨https://www.tridentsteel.co.in/api-5l-pipe.html#what⟩.
[89] Y.G. Woldesenbet, G.G. Yen, B.G. Tessema, Constraint handling in multiobjective evolutionary optimization, IEEE Trans. Evol. Comput. 13 (3) (2009) 514–525, https://doi.org/10.1109/TEVC.2008.2009032.
[90] C.C. Coello, M.S. Lechuga, MOPSO: A proposal for multiple objective particle swarm optimization, in: Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600), 2, IEEE, 2002, pp. 1051–1056, https://doi.org/10.1109/CEC.2002.1004388.
[91] M. Reyes-Sierra, C.C. Coello, Multi-objective particle swarm optimizers: a survey of the state-of-the-art, Int. J. Comput. Intell. Res. 2 (3) (2006) 287–308, https://doi.org/10.5019/J.IJCIR.2006.68.
[92] M. Clerc, J. Kennedy, The particle swarm – Explosion, stability, and convergence in a multidimensional complex space, IEEE Trans. Evol. Comput. 6 (1) (2002, February) 58–73, https://doi.org/10.1109/4235.985692.
[93] K. Deb, K. Sindhya, J. Hakanen, Multi-objective optimization. Decision Sciences, CRC Press, 2016, pp. 161–200.
[94] N. Gunantara, A review of multi-objective optimization: methods and its applications, Cogent Eng. 5 (1) (2018) 1502242, https://doi.org/10.1080/23311916.2018.1502242.
[95] J. Branke, Multiobjective optimization: Interactive and evolutionary approaches, 5252, Springer Science & Business Media, 2008, https://doi.org/10.1007/978-3-540-88908-3.
[96] E. Zitzler, K. Deb, L. Thiele, Comparison of multiobjective evolutionary algorithms: empirical results, Evol. Comput. 8 (2) (2000) 173–195, https://doi.org/10.1162/106365600568202.
[97] M.J. Pyrcz, C.V. Deutsch, Geostatistical reservoir modeling, Oxford university press, 2014.
[98] H. Jo, M.J. Pyrcz, F. Laugier, M. Sullivan, Sensitivity analysis of geological rule-based subsurface model parameters on fluid flow, AAPG Bull. 107 (6) (2023) 887–906, https://doi.org/10.1306/10102221083.
[99] W.J. Al-Mudhafar, T.H. Vo, D.A. Wood, B. Min, Stochastic lithofacies and petrophysical property modeling for fast history matching in heterogeneous clastic reservoir applications, Sci. Rep. 14 (2024) 22, https://doi.org/10.1038/s41598-023-50853-3.
[100] V. Honari, X. Wang, J. Underschultz, A. Garnett, X. Wang, R. Johnson, et al., Evaluation of CO_2 injectivity for CO_2-EOR using a two-stage well test approach: Case study of a Chinese tight oil reservoir (Paper No. SPE-202374-MS). SPE Asia Pacific Oil and Gas Conference and Exhibition, 2020, November, https://doi.org/10.2118/202374-MS.
[101] A.R. Shokri, M. Kamali, Z. Movahedzadeh, R. Narayanasamy, E. Nickel, R. Chalaturnyk, Significance of multiple CO_2 plume management and pressure interactions during geological CO_2 storage operation based on the Canadian Aquistore CCS Project (Paper No. SPE-222195-MS). Abu Dhabi International Petroleum Exhibition and Conference, 2024, November, https://doi.org/10.2118/222195-MS.
[102] A. Afanasyev, A. Penigin, M. Dymochkina, E. Vedeneeva, S. Grechko, Y. Tsvetkova, et al., Reservoir simulation of the CO_2 storage potential for the depositional environments of West Siberia, Gas. Sci. Eng. 114 (2023) 204980, https://doi.org/10.1016/j.jgsce.2023.204980.
[103] M. Jang, S. Yoon, S. Jung, B. Min, Simulating and assessing carbon markets: application to the Korean and the EU ETSs, Renew. Sustain. Energy Rev. 195 (2024) 114346, https://doi.org/10.1016/j.rser.2024.114346.
[104] M. Jang, B. Min, Stochastic forecasting of long-term greenhouse gas emissions and energy transitions: a comparative analysis of the US, EU, China, and Korea, Energy 327 (2025) 136376, https://doi.org/10.1016/j.energy.2025.136376.
[105] G.S.H. Pau, Y. Zhang, S. Finsterle, H. Wainwright, J. Birkholzer, Reduced order modeling in iTOUGH2, Comput. Geosci. 65 (2014) 118–126, https://doi.org/10.1016/j.cageo.2013.08.008.
[106] M. Zhao, Y. Wang, M. Gerritsma, H. Hajibeygi, A physics-constraint neural network for CO_2 storage in deep saline aquifers during injection and pst-injection periods, Adv. Water Resour. 193 (2024) 104837, https://doi.org/10.1016/j.advwatres.2024.104837.
[107] T. Kadeethum, V.L.S. Silva, P. Salinas, C.C. Pain, H. Yoon, Boosting Barlow Twins reduced order modeling for machine learning-based surrogate models in multiphase flow problems, Water Resour. Res. 60 (10) (2024) e2023WR035778, https://doi.org/10.1029/2023WR035778.
[108] U. Kaymak, J.M. Sousa, Weighting of constraints in fuzzy optimization, in: Proceedings of the 10th IEEE International Conference on Fuzzy Systems, 2, IEEE, 2001, December, pp. 1131–1134, https://doi.org/10.1109/FUZZ.2001.1008854.
[109] Y. Yongping, Z. Rongrong, D. Liqiang, O. Masek, J. Oakey, Study on multi-objective optimization of load dispatch including renewable energy and CCS technologies, Int. J. Energy Res. 34 (8) (2010) 702–715, https://doi.org/10.1002/er.1590.
[110] D. Asadi, E.M. Atkins, Multi-objective weight optimization for trajectory planning of an airplane with structural damage, J. Intell. Robot. Syst. 91 (4) (2018) 667–690, https://doi.org/10.1007/s10846-017-0753-9.