

Article

Not peer-reviewed version

A Fully Automated Design of Experiments-Based Method for Rapidly Screening Near-Optimal CO₂ Injection Strategies

[Epameinondas Theofanis Diplas](#) , [Sofianos Panagiotis Fotias](#) , Ismail Ismail , [Spyridon Bellas](#) , [Vassilis Gaganis](#) *

Posted Date: 24 December 2025

doi: 10.20944/preprints202512.2077.v1

Keywords: CO₂ storage; design of experiments; CCS optimization; reservoir simulation; radial basis function; D-optimal



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Fully Automated Design of Experiments-Based Method for Rapidly Screening Near-Optimal CO₂ Injection Strategies

Epameinondas Theofanis Diplas ¹, Sofianos Panagiotis Fotias ¹, Ismail Ismail ¹, Spyridon Bellas ² and Vassilis Gaganis ^{1,*}

¹ School of Mining and Metallurgical Engineering, National Technical University of Athens, Athens 73100, Greece

² Institute of Geoenergy, Foundation for Research and Technology Hellas, M1 Building, University Campus, GR-73100 Chania, Greece

* Correspondence: vgaganis@metal.ntua.gr

Abstract

Injection well placement and rate allocation are among the most decisive factors in determining the efficiency and bankability of CCS projects. However, optimizing these parameters is notoriously complex: even a small number of injection wells leads to a virtually infinite set of injection scenarios, while traditional optimization techniques typically require thousands of high-fidelity reservoir simulations. For project developers, this computational burden can stall critical Final Investment Decisions (FID). The proposed approach here addresses this bottleneck by using a Design of Experiments (DoE) framework combined with nonlinear surrogate modeling, which efficiently maps the relationship between injection rates and storage performance, to identify near-optimal solutions with a minimal number of simulations. We show that our method achieves up to 97% of the initially targeted CO₂ sequestration with as few as 15 simulations, demonstrating a step-change reduction in time and cost. From a business standpoint, CCS operators can de-risk projects earlier, accelerate FID timelines, and evaluate multiple site configurations in parallel while minimizing computational overhead. Rather than waiting weeks or months for exhaustive optimization, decision-makers can gain timely, reliable insights that directly support capacity commitments, regulatory submissions, and ultimately revenue realization.

Keywords: CO₂ storage; design of experiments; CCS optimization; reservoir simulation; radial basis function; D-optimal

1. Introduction

The twenty-first century is witnessing an unprecedented acceleration of climatic change, driven predominantly by the persistent accumulation of anthropogenic carbon dioxide and other greenhouse gases in the atmosphere. Global mean surface temperatures have already risen by approximately 1.1 °C above pre-industrial levels, and this warming is accompanied by profound disruptions: polar ice sheets are retreating at record rates, sea levels are rising faster than at any time in the past two millennia, and extreme weather events, from record-breaking heatwaves to catastrophic floods, have become more frequent and intense [1]. The Intergovernmental Panel on Climate Change (IPCC) projects that, without decisive intervention, global warming could exceed 1.5 °C as early as the 2030s, with cascading impacts on ecosystems, agriculture, human health, and socio-economic stability [2]. These trends underscore that climate change is not a distant threat but an immediate crisis demanding integrated scientific, technological, and societal responses.

Even with aggressive deployment, certain industrial processes will continue to emit substantial CO₂, necessitating complementary technologies capable of immediate, scalable mitigation to sustain

economic growth. Among these, Carbon Capture and Storage (CCS) stands out as a transformative solution—bridging the gap between current emission realities and long-term climate goals, while ensuring that industrial growth, economic resilience, and energy security are preserved in the transition toward a sustainable economy [3]. Carbon capture represents the first step in the CCS chain, extracting CO₂ from industrial or power-generation emissions using chemical, physical, or biological processes. Capture technologies are broadly categorized into pre-combustion, post-combustion, and oxy-fuel approaches, each tailored to specific industrial configurations and emission sources [4]. Following capture, CO₂ is transported to storage sites through dedicated pipelines, shipping, or other engineered transport solutions, with robust safety and monitoring protocols to prevent leaks or operational failures [5]. While utilization remains an optional pathway converting CO₂ into chemicals, fuels, or construction materials [6], the storage pillar guarantees permanent removal, forming the cornerstone of climate mitigation in hard-to-abate sectors [3].

Storage involves a coordinated chain of surface and subsurface operations. Surface infrastructure includes wells, pumps, compressors, and monitoring stations that regulate CO₂ injection and ensure operational safety. Subsurface storage relies on carefully selected geological formations where CO₂ is immobilized through a combination of natural forces and mechanisms. Structural and stratigraphic trapping confines CO₂ beneath impermeable rock layers, residual trapping immobilizes it within pore spaces, solubility trapping allows CO₂ to dissolve in formation water, and mineral trapping converts CO₂ into stable carbonate minerals over longer timescales. Effective storage depends on precise placement and control of wells injection rates [7,8], continuous monitoring of well integrity, and rigorous observation of subsurface conditions to prevent leakage and ensure permanent containment, [9,10].

The design and operation of CO₂ storage systems represent one of the most complex engineering challenges within the broader CCS value chain. Unlike capture or transport, the subsurface storage phase is governed by strongly coupled physical, chemical, and operational processes that evolve across multiple spatial and temporal scales. Decisions related to injection design such as placement, sequencing, rate control, and pressure management directly influence the efficiency, safety, and longevity of the storage operation [3].

At the core of any credible and technically sound Storage Development Plan (SDP) lies the injection strategy, or more precisely, the injection schedule a plan that defines how, where, and when CO₂ is injected into the reservoir [11]. The objective is to maximize storage efficiency while satisfying a complex set of multidimensional constraints. These constraints arise from the interplay between subsurface technical factors, operational limitations, and broader economic, commercial, and regulatory conditions. From a technical perspective, pressure buildup within the reservoir must remain below a defined fraction of the fracture pressure to preserve caprock integrity and ensure containment [12]. Simultaneously, plume migration must be carefully managed to prevent CO₂ from escaping the designated storage complex. The geological heterogeneity of the formation variations in porosity, permeability, and capillary entry pressures further complicates this task, rendering deterministic optimization approaches insufficient. Moreover, operational strategies such as brine production or reinjection, which are employed to manage pressure and enhance storage capacity through dissolution mechanisms, add additional degrees of freedom and complexity to the control problem [13]. From an economic standpoint, CO₂ storage optimization involves continuous trade-offs between cost and performance. Higher injection rates or additional wells may accelerate project timelines and increase total storage capacity, but they also raise capital and operating expenditures [14]. Conversely, conservative injection strategies reduce risk yet may underutilize available capacity and delay commercial returns. Balancing these competing objectives requires an integrated optimization framework capable of coupling technical performance metrics with financial, regulatory, and risk-based considerations [11].

Traditional work involves manual heuristic design of the injection schedule however, optimization frameworks those relying on explicit analytical formulations are inadequate for addressing such systems. The objective functions governing CCS operations are inherently non-

linear, non-convex, and path-dependent, reflecting the physics of multiphase flow, geomechanics, and chemical interactions in the subsurface [11]. Furthermore, the operational constraints are time-dependent and spatially distributed, rendering closed-form solutions infeasible. As a result, numerical simulation-based optimization becomes indispensable. This approach leverages reservoir simulators to evaluate candidate injection policies, iteratively adjusting control parameters through computational algorithms to converge on optimal or near-optimal solutions.

Meanwhile, simulation-based optimization introduces its own challenges. Each simulation run is computationally expensive [15], and the solution space defined by the combination of well configurations, rates, pressures, and time schedules is vast. Exhaustive search methods are therefore impractical. This motivates this work, where the use of approaches that identify near-optimal strategies—achieving performance close to the true optimum at a fraction of the computational cost—offer a pragmatic alternative. In this study, we propose a method capable of effectively identifying a near-optimal CO₂ storage capacity solution by optimizing wells injection rates with minimal computational effort. The approach is based on two key principles: first, an experimental design of initial points used to train a surrogate model targeted for maximization; second, an iterative refinement of the surrogate model by incorporating the optimal points along with their true simulation responses, continuing until the maximum allowed simulation budget is reached.

In laboratory science, experiments require materials, apparatus, and time, whereas, in computational reservoir engineering, experiments are conducted through numerical simulations, where the primary cost lies in CPU time rather than physical resources [16,17]. Traditional passive data collection in simulations can lead to biased or redundant datasets, reducing the efficiency of model calibration. By contrast, in the first key principle, Design of Experiments (DoE) provides a systematic methodology for varying input factors in an informative and structured manner, ensuring that maximum information is extracted from a limited number of runs [18]. DoE methods can thus be effectively applied to the selection of computational simulations, enabling efficient exploration of the simulation space while minimizing computational cost. In the present framework, DoE is used to select the minimum number of test injection scenarios, still carrying the maximum amount of information on the behavior of the CO₂ storage performance function to be optimized. By revealing the nature of the performance function using the DoE sampled injection scenarios, the user can continue with its optimization to identify the optimal one while honoring all physical, mathematical or operational constraints.

Among various DoE methods, D-optimal designs are particularly suitable when simulation costs are high and the design space is large. The D-optimality criterion seeks to maximize the information captured from the design space for a given amount of computational effort, thereby minimizing parameter uncertainty in the surrogate model to follow. For example, [19] developed a D-optimal design framework to build rapid assessment models for CO₂ flooding in high water cut oil reservoirs, using five key reservoir variables and a quadratic regression model to predict recovery factor and net present value. Their study demonstrated that D-optimal experimental design can effectively reduce the number of required simulations while maintaining prediction accuracy, making it a valuable screening tool for CO₂-EOR applications. Similarly, [20] developed a framework for optimizing of reservoir simulation studies that incorporated D-optimal design to efficiently determine optimal well locations under geological uncertainty. Their results showed that D-optimal design helps focus simulation runs on the most informative well configurations, yielding reliable optimization outcomes with far fewer simulations. This approach ensures that each selected simulation contributes maximally to the surrogate model calibration, making D-optimal designs especially effective for screening CO₂ injection strategies in saline aquifers.

Even with carefully chosen DoE originating injection scenarios, the number of simulations may remain insufficient to fully characterize reservoir responses. To address this, surrogate models—simplified mathematical approximations of the true simulation response—are widely employed. Once these surrogates are constructed, optimization techniques can be applied directly to identify promising strategies. For instance, [21] employed a Design of Experiments (DoE) approach combined

with Radial Basis Function neural networks (RBF-ANNs) to optimize the Gas–Downhole Water Sink–Assisted Gravity Drainage (GDWS-AGD) process. Their study demonstrated that RBF-based surrogate models can efficiently replace full-physics simulations, accurately capturing the effect of key operational parameters—particularly bottom-hole injection pressure—while markedly reducing computational effort. This highlights the potential of combining DoE with data-driven surrogates to explore complex reservoir behaviors at a fraction of the computational cost.

This work proposes and evaluates a hybrid optimization workflow for CO₂ injection in saline aquifers, designed to rapidly identify near-optimal strategies while minimizing computational demand. The approach is built on five interconnected objectives. First, a D-optimal design of experiments (DoE) is employed to select a minimal but highly informative set of injection scenarios. Second, these scenarios are executed through full-physics reservoir simulations, ensuring that realistic operational constraints such as injection limits, bottom-hole pressure thresholds, and caprock integrity are honored [22]. Third, surrogate models based on polynomial regression or Radial Basis Functions (RBFs) are constructed to approximate the relationship between injection rates and CO₂ storage capacity. Fourth, the surrogate models are optimized using a gradient based optimization algorithm or a Particle Swarm Optimization (PSO) algorithm to efficiently identify near-optimal injection strategies. Finally, the optimized solutions are validated against additional full-physics simulations using the Open Porous Media (OPM) software[23]. Three case studies are presented, demonstrating the efficiency of the proposed method.

By integrating experimental design, simulation, surrogate modeling, and tailored optimization techniques, the proposed workflow enables efficient exploration of the injection strategy design space, balancing computational feasibility with optimization accuracy.

The remainder of this paper is structured as follows. Section 2 reviews the theoretical principles that form the basis of the proposed methodology. Section 3 presents the proposed screening methodology. Section 4 reports the case study, including the static model and operational constraints. Section 5 presents and discusses the results of the three cases. Finally, Section 6 concludes the work.

2. Theoretical Background

This section presents the mathematical foundations that are essential for comprehending the methodology of the present research. It focuses on the principles of the D-optimal experimental design and on the construction of the information matrix. Subsequently, it discusses Polynomial Regression, RBF surrogate modeling and the Particle Swarm Optimization (PSO) algorithm. All those components are involved in various steps of the proposed methodology.

2.1. D-Optimal Design of Experiments Method

For a static process governed by a model of the form $y = f(\xi)$, where y is the process response and ξ is the vector containing all factors contributing to the response, Design of Experiments (DoE) provides a systematic framework for planning and conducting experiments to efficiently explore the effects of multiple design variables on one or more system responses. Instead of changing one factor at a time, DoE methods vary several factors simultaneously according to a predefined design matrix, allowing the model to estimate main effects, interactions, and nonlinearities with a minimum number of runs [24]. In computational studies, DoE is crucial for selecting the most informative simulation points within the feasible domain, reducing the total computational effort while preserving model accuracy.

Let Ξ_N , of size $N \times p$ denotes the full set of all N possible candidate experiments, where each row corresponds to one combination of factors and each column to one of the p design variables. From this set, a subset of n experiments ξ_i , $i \in [1, 2, \dots, n]$, is chosen to build the model matrix $\mathbf{X}$ of size $(n \times p)$, $\mathbf{X} = [\xi_1, \dots, \xi_n]^T$. The information matrix $\mathbf{M}$ summarizes how much information the selected experiments provide about the model parameters, defined as:

$$\mathbf{M}_{(p \times p)} = \frac{1}{n} \mathbf{X}^T \mathbf{X} \quad (1)$$

A higher determinant value of $\mathbf{M}$ maximizes the volume spanned by the design vectors in p -dimensional factor space. Geometrically, this means the selected experiments cover the factor space as fully and evenly as possible, providing the maximum amount of independent information about the system [25]. Among all possible subsets ξ_i drawn from Ξ_N , the D-optimal design is the one that maximizes the determinant of the information matrix, described in Equation (2) for a fixed number of runs, n [26].

$$\{\xi_1, \dots, \xi_n\} = \operatorname{argmax}(|\mathbf{M}|) \quad (2)$$

D-optimal designs are therefore preferred when the total number of feasible experiments is limited or when each model evaluation is computationally expensive, as they provide the maximum information from the smallest possible set of runs [27]. For the purposes of this research, the BoFire (Bayesian Optimization Framework for Real Experiments) library [28] was employed to determine the D-optimal injection scenarios.

2.2. Polynomial Regression

Regression aims to build function forms $y = f(\mathbf{x})$ upon a dataset $\{\mathbf{x}_i, y_i\}$, where more than one or more independent variables appear in the input $\mathbf{x}$ and one or more output ones may appear in y [29]. In its simplest form, the polynomial regression, we regress one dependent variable on monomials of a single independent variable. In that case, the model can be expressed as:

$$y = \theta_0 + \theta_1 x + \theta_2 x^2 + \dots + \theta_k x^k + \varepsilon \quad (3)$$

where k is the degree of the polynomial, θ_l , with $l = 0, 1, \dots, k$, are the coefficients of the polynomial and ε is the fitting error. Equation (3) can be generalized for multiple variables $x_1, x_2, \dots, x_p$ whereas interaction or other non-linear terms between them can also be included, for example $x_i \cdot x_j$. This allows the model to capture not only the individual effect of each variable on the response but also the combined influence of variable pairs or higher-order interactions [30].

In multiple linear regression, the model of one dependent (response) variable y on a set of p independent (predictor) variables $\mathbf{x} = [x_1, \dots, x_p]^T$ can be expressed in matrix algebra term by:

$$y = \boldsymbol{\theta}^T \mathbf{x} + \varepsilon \quad (4)$$

The regression surrogate in this study is constructed using the scikit-learn library [31], which fits the model to the training data by minimizing the least-squares error, thereby ensuring an accurate approximation of the response surface.

Once a linear or quadratic polynomial surrogate model (first- and second-order respectively) has been developed, thus the assumption that the response surface is either planar (linear) or convex/concave (quadratic) respectively, the optimal point of the model is then sought [32]. In the case of quadratic model regression, there is a single optimum within the feasible space. This property simplifies the identification of the optimum point to the use of classic gradient-based methods. However, for the problem at hand, there are also equality and inequality constraints, analyzed in the following sections. Those constraints now allow for the optimization of linear surrogate models, or even quadratic ones with the optimum living outside the feasible space. In such cases, the most suitable approach for solving the constrained optimization problem is the Sequential Least Squares Quadratic Programming (SLSQP) algorithm, implemented using the SciPy library [33].

2.3. Radial Basis Function

Radial basis functions (RBFs) are a class of real-valued functions whose values depend on the distance of the operating $\mathbf{x}$ from a center point, also known as the radial distance. These basis

functions have long been explored in function approximation, machine learning, and pattern recognition tasks [34].

The basic idea in RBFs is to model a function by linearly combining several radially symmetric functions, each centered at a different point in the input space. The output of an RBF network is a linear combination of radial basis functions, weighted by adjustable coefficients. Mathematically, an RBF network with M centers can be represented as:

$$f(\mathbf{x}) = \sum_{i=1}^M w_i \varphi(\|\mathbf{x} - \mathbf{c}_i\|) \quad (5)$$

where w_i are the adjustable weights or coefficients and $\varphi(\cdot)$ is the radial basis function, which depends on the distance between the input $\mathbf{x}$ and the center of each RBF, $\mathbf{c}_i$ [35]. The Gaussian function is one of the most widely used RBFs, defined as:

$$\varphi(r) = \exp(-(\gamma \cdot r)^2) \quad (6)$$

$$r = \|\mathbf{x} - \mathbf{c}_i\| \quad (7)$$

where r is the radial distance, and γ is a parameter value that controls the width or spread of the function [36]. A smaller shape parameter γ value corresponds to a “flatter” or “wider” basis function. The limit $\gamma \rightarrow 0$ is often referred as the “flat” limit, because $\gamma = 0$ corresponds to a constant basis function. In this study, the SciPy library [33] was employed to perform RBF interpolation for constructing the surrogate model.

2.4. Particle Swarm Optimization Algorithm

The Particle Swarm Optimization (PSO) algorithm is a population-based optimization technique inspired by the social behavior of bird flocking and fish schooling. Introduced in 1995, PSO has been extensively applied to a wide range of complex optimization problems in fields such as engineering, machine learning, and computational intelligence.

PSO is initialized with a population of randomly distributed particles, and the optimal solution is sought through iterative updates. Each particle is characterized by two parameters: its velocity u^k and its position x^k at iteration k . In each iteration, the particle updates itself by tracking two 'extremes'. One extremum is the optimal solution found by the particle itself, which is called the individual extremum P_{best} and the other extremum is the optimal solution currently found by the entire population, which is called the global extremum G_{best} [37]. After identifying these two optima, the particle's velocity and position are updated according to the standard PSO update Equations (A1) and (A2), provided in Appendix A.1.

Through iterative updates of particle velocities and positions, each particle moves toward the global best position that minimizes the objective function, thereby improving the accuracy of the predicted optimum. The optimization procedure terminates when convergence is achieved, defined as a change in the objective function smaller than a specified tolerance [38]. In this study, the PySwarms Python library [39] was employed to implement the PSO framework and perform the optimization of the surrogate model.

3. Screening Methodology

3.1. Problem Definition and DoE Sampling

The injection schedule optimization problem considers n injection wells each with its own constant injection rate I_i (Mtpa), $i = 1, 2, \dots, n$, over a period of y years. The objective is to identify the I_i value for each well which achieve a cumulative CO₂ storage target of T (Mtpa), where T corresponds to the desired total annual injection rate of the field which is screened. Each injection well operates within predefined bounds, expressed by the inequality constraint (8).

$$I_{min} \leq I_i \leq I_{max}, i = 1, 2, \dots, n \quad (8)$$

The lower limit of I_{min} (Mtpa) is imposed to avoid the unrealistic scenario of equating the operational power of some well to zero, while the upper limit I_{max} (Mtpa) corresponds to the typically maximum feasible injection capacity of a single well, mostly depending on the tubing diameter and the CO₂ injection temperature [40].

The total injection rate is fixed at the desired T value, to serve the emitters need for a guaranteed performance of the field:

$$\sum_{i=1}^n I_i = T \quad (9)$$

Given the equality constraint (9), the problem dimensionality can be reduced to $n - 1$ independent variables, with I_n been expressed by:

$$I_n = T - \sum_{i=1}^{n-1} I_i \quad (10)$$

Applying the bound condition (8) to I_n and substituting Equation (10), results in the inequality constraint Equation (11), which can be rearranged to Equation (12). Finally, isolating I_{n-1} leads to in a single inequality constraint, Equation (13), defining its feasible region.

$$I_{min} \leq T - \sum_{i=1}^{n-1} I_i \leq I_{max} \rightarrow \quad (11)$$

$$T - I_{max} \leq \sum_{i=1}^{n-1} I_i \leq T - I_{min} \quad (12)$$

$$T - I_{max} - \sum_{i=1}^{n-2} I_i \leq I_{n-1} \leq T - I_{min} - \sum_{i=1}^{n-2} I_i \quad (13)$$

Understanding the feasible design space is essential to correctly interpret the optimization results presented later. In particular, it is important to examine the geometric expression of inequality constraint (13), together with the bound constraints (8), within the $(n - 1)$ - dimensional decision space.

Regarding constraint (8), each inequality restricts the feasible region through two parallel hyperplanes defined by the constants I_{min} , I_{max} . Since these bounds do not depend on any other variable I_i , they generate $(n - 1)$ axis- aligned, zero-slope constraints that simply limit each I_i to lie within a fixed half-space.

In contrast, inequality (13) introduces a coupling between the variables through the summation term. Geometrically, Equation (13) defines two linear, parallel hyperplanes in the $(n - 1)$ - dimensional space, not aligned with the axes, each involving all variables $I_1, I_2, \dots, I_{n-1}$. More specifically, the left-hand and right-hand sides of Equation (13) impose lower and upper bounds on I_n that vary as functions of the remaining $n - 2$ variables. This creates two oblique hyperplanes with non-zero slope, thereby constraining I_{n-1} relative to the combined values of the other injection rates.

As a result, the feasible region is no longer a simple hyper-rectangle defined solely by the independent bounds, Equation (8), but a truncated region whose shape depends on how the total injection constraint (9) is distributed among the wells.

To make the formulation of the feasible design space more interpretable, consider a representative example with $n = 3$ injection wells, a total field injection rate of $T = 2.0$ Mtpa, and operational bounds $I_{min} = 0.15$ Mtpa and $I_{max} = 1.0$ Mtpa. From constraint (8), each well rate is

bounded between 0.15 and 1.0 Mtpa, whereas Equation (10) reduces the decision space to a two-variable one. The governing relations (8), (10), (13) for this example are therefore:

$$0.15 \leq I_i \leq 1.0, i = 1, 2 \quad (14)$$

$$I_3 = 2.0 - (I_1 + I_2) \quad (15)$$

$$1.0 - I_1 \leq I_2 \leq 1.85 - I_1 \quad (16)$$

As discussed earlier, inequality (14) alone restricts the feasible region to a rectangular design space, as illustrated in Figure 1a. Inequality (16) introduces two additional restrictions. To visualize this for the present three-well example, we note that $I_1 \in [0.15, 1.0]$ from Equation (14). The two extreme values of I_1 therefore define the two boundary lines of constraint (16), which correspond to the limits of the feasible region in the (I_1, I_2) plane. On the one hand, when $I_1 = 0.15$ Mtpa, the bounds for I_2 are $[0.85, 1.70]$, but according to Equation (14) it is truncated to $[0.85, 1.0]$. On the other hand, when $I_1 = 1.0$ Mtpa, the bounds for I_2 are $[0, 0.85]$, but according to Equation (14) it is truncated to $[0.15, 0.85]$. The resulting restriction is illustrated in Figure 1b.

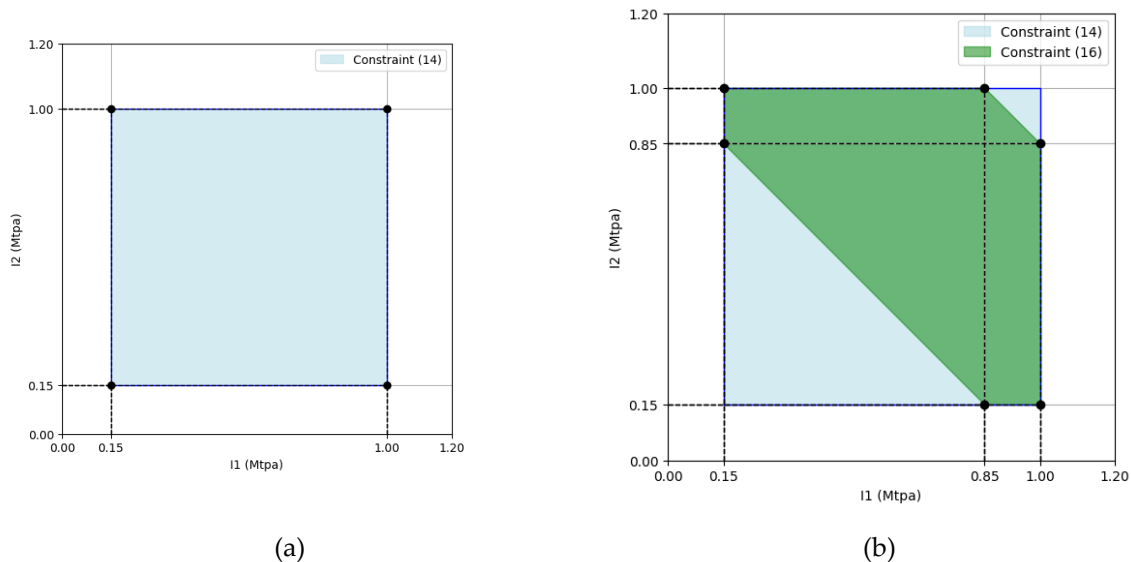


Figure 1. Illustration of the feasible design space formatting constraints (14), (a), and (16), (b), in the two-dimensional space.

After the feasible design space definition, a D-optimal design is generated, constrained to produce only injection schedules within the feasible space, described according to inequalities (8), (12), resulting in a design matrix $\mathbf{X}$ where each column corresponds to a decision variable and each row represents a sample point, i.e. an injection scenario that needs to be simulated to evaluate the system response, hence the objective function.

The number of simulation points generated, n , does not follow a general closed-form formula based on the number of variables and constraints, but rather depends on the available time and computational resources. Indicatively, for the case studies in this work incorporating three injection wells (i.e., two variables) four D-optimal points were generated, whereas for four injection wells (i.e., three variables) nine D-optimal points were produced. If the surrogate model type considered is polynomial (e.g., a quadratic model), then a D-optimal design needs to be generated with a number of points at least equal to the number of coefficients to ensure unique identification of the model parameters, while adding more points can improve precision [29]. In cases where an RBF model is deemed more suitable for capturing the underlying relationship, the goal is to select a number of simulations that adequately cover the feasible space without exceeding the available computational resources. The objective in all cases is to cover the feasible space with as few simulations as possible while still extracting sufficient information. Increasing the number of the D-optimal produced points

allows for surrogate models to be fitted in higher accuracy in subsequent steps, while preserving and utilizing the information collected in the first run.

The next step is to perform reservoir simulations at the designed sampling points. The numerical reservoir flow simulations at each design point, are performed using the Open Porous Media Flow reservoir simulator [23], thereby generating the initial dataset of CO₂ sequestration outcomes by recording the total CO₂ stored at the end of each simulation, the CO₂ recycled, the brine produced and the pressure and saturation distribution over the whole reservoir.

Subsequently, the outcomes are evaluated from a reservoir engineering perspective to assess whether the surrogate modeling and optimization techniques that will follow can further improve the maximum point identified, or whether alternative measures might be required to enhance efficiency. For example, by adding an additional injection well to the system.

3.2. Optimization Method

In low-complexity problems, such as systems with one to three injection wells, a simple polynomial surrogate, such as a linear model extended to include interaction effects or a full quadratic model, may be sufficient to accurately represent the D-optimal dataset. In the latter case, the surrogate model is expressed, subject to Equations (8) and (13), as:

$$CO_2^{seq}(Mtn) = a_0 + \sum_{i=1}^{n-1} b_i \cdot I_i + \sum_{i=1}^{n-2} \sum_{j=1}^{n-2} c_{ij} \cdot I_i \cdot I_j \quad (17)$$

where CO_2^{seq} denotes the CO₂ total sequestered surrogate's prediction.

In such cases, a straightforward gradient-based optimization algorithm, such as SLQP can be used to identify the optimal operating point $\xi^* = [I_1, \dots, I_n]^*$ at which CO₂ sequestration is maximized within the feasible space and n is the number of the injection wells.

In contrast, for more complex cases involving four or more injection wells, a Radial Basis Function (RBF) surrogate model with a Gaussian kernel is preferred, as it can capture complex nonlinear interactions among injection rates without incorporating a huge number of the surrogate model parameters. The surrogate model is expressed according to Equations (5), (6) and (7) as:

$$CO_2^{seq}(Mtn) = \sum_{i=1}^m w_i \varphi(\|\xi - \mathbf{c}_i\|) \quad (18)$$

$$\varphi(r) = \exp[-(\gamma \cdot r)^2] \quad (19)$$

where CO_2^{seq} is the surrogate prediction of the total CO₂ sequestered in Mtn, m is the number of simulation points, ξ is the variable operating point, $\mathbf{c}_i$ are the RBF centers pointing at the D-optimal solution (training points) and $r = \|\xi - \mathbf{c}_i\|$. Conceptually, γ value determines the range of influence of each Gaussian center: larger values produce narrower Gaussians with more localized effects, while smaller values produce broader Gaussians affecting a wider region.

For the initial RBF surrogate, the γ value is selected randomly between 0.1 and 1.0. This approach aims to create a smoothed surrogate in which the influence of individual points is more evenly distributed, rather than being dominated by the maximum point. The choice of the initial random γ value does not yet affect the optimization result, as it will be adjusted in the subsequent optimization step.

After constructing the RBF surrogate from the D-optimal design points, a primary Particle Swarm Optimization (PSO) algorithm is employed to identify the operating point ξ^* that maximizes CO₂ sequestration, described in Equation (20) subject to the constraint (8), by minimizing L , Equation (21). Note that L can be extended to consider any additional effects, such as plume expansion, contribution of different storage mechanisms, leakage penalty or cash-flow related terms.

$$\xi^* = \operatorname{argmin}(-L) \quad (20)$$

$$L = -CO_2^{seq} \quad (21)$$

Followingly, a reservoir simulation of the resulted ξ^* is performed. The actual CO₂ sequestered value ($CO_2^{seq}_{sim}$), as obtained by another simulation run, is recorded and its absolute prediction error is calculated by the Equation (22) and also documented.

$$|\varepsilon_{\xi^*}| = |CO_2^{seq}_{sim} - CO_2^{seq}| \quad (22)$$

The next step is to employ a secondary PSO algorithm to determine the γ^* parameter value that minimizes the surrogate's absolute prediction error at ξ^* by minimizing $|\varepsilon_{\xi^*}|$, as described in Equation (23). This γ parameter value optimization step ensures that the surrogate is optimally fitted in the region where the best solution is expected to lie, thereby improving local accuracy and the reliability of the subsequent optimization.

$$\gamma^* = \operatorname{argmin}(|\varepsilon_{\xi^*}|) \quad (23)$$

The procedure subsequently continues by incorporating the identified ξ^* into the surrogate's training set and refitting the RBF surrogate model with the updated dataset and the optimized γ^* value. Followingly, the primary PSO algorithm is re-executed to search for the new optimal solution that minimizes L based on the updated surrogate. The entire procedure is repeated iteratively by running one additional simulation per iteration, until either three consecutive optimal ξ^* solutions are obtained and found equal up to some threshold, or until a total of 15 sampling points has been reached. All supporting details regarding the PSOs specifications, are provided in Tables A1 and A2, Appendix A.2.

4. Case Studies

For this research, a detailed aquifer case study is developed to evaluate the proposed CO₂ storage optimization strategies. The reservoir static model has been employed in previous studies, [7,41], focusing on optimal well placement and flow rates allocation, leveraging Bayesian optimization frameworks. By employing this established static model, where the well placement has already been optimized, the influence of well location on the results is eliminated. This ensures that the present study focuses solely on evaluating the impact of operational and economic factors, such as injection strategies and CO₂ storage performance, without any interference from well placement variability.

The corner-point geometry reservoir grid comprises of $120 \times 163 \times 4$ cells (X, Y, Z), totaling 78,240 cells, of which 9,212 are active. Inactive cells define non-flow zones and are excluded from flow computations. The aquifer exhibits an inclined structural profile with a steep anticlinal crest, which acts as a trap for buoyant CO₂, illustrated in Figure 2.

The depth of the reservoir, shown in Figure 2, ranges from 9,311 ft at the crest to 10,850 ft at the deepest point. A hydrostatic pressure–depth gradient of 0.465 psi/ft [42] was used in the model, resulting in a normal initial reservoir pressure between approximately 4,330 psi at the crest and 5,045 psi at the deepest point, with an average pressure of about 4,700 psi.

For the temperature–depth relationship, a geothermal gradient of 0.0142 °F/ft [43] was applied. The temperature at depth D (ft) is calculated using Equation (24):

$$T_D = T_s + G \cdot D \quad (24)$$

where $T_s = 77$ °F is the surface temperature and G is the geothermal gradient. Based on this, the crest temperature is approximately 209 °F, while at the deepest point the temperature is about 231 °F, giving an average reservoir temperature of roughly 219 °F.

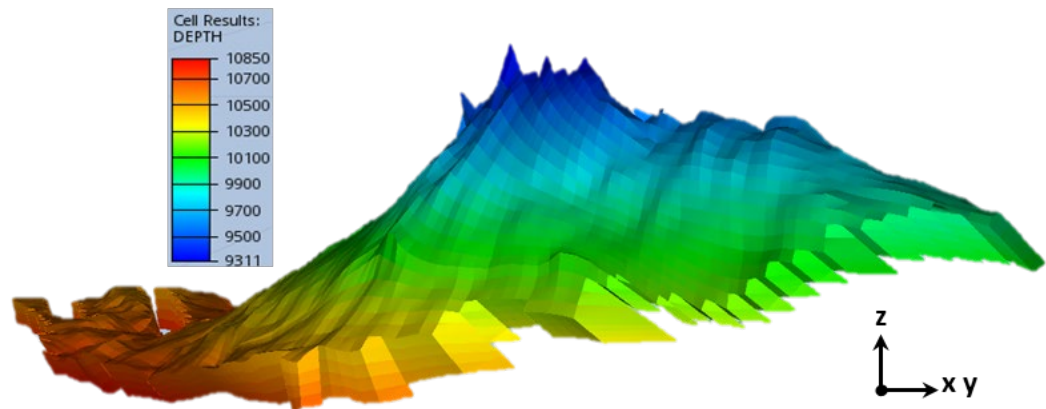


Figure 2. Case studies aquifer depth.

The reservoir is structurally simple and sealed by impermeable shale, with a bulk volume of 240 MMMft³, initial water in place of 1.7 MMMSTB, salinity of ~65,000 ppm. Permeability is anisotropic and heterogeneous: horizontal permeability ($k_x = k_y$) has P90 = 91 mD, P50 = 272 mD, P10 = 572 mD, while vertical permeability (k_z) is one order of magnitude lower, favoring lateral flow. The spatial distribution of permeability X is provided in Figure 3. Porosity is uniform at 0.25, indicating stable storage capacity. The CO₂-brine (water) relative permeabilities are also given in Table 1, Appendix B.1 [15], and their plot is presented in Figure 4.

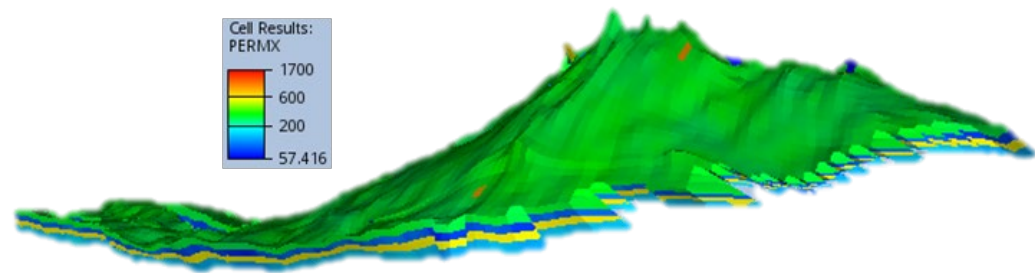


Figure 3. Aquifer grid block permeability in X direction.

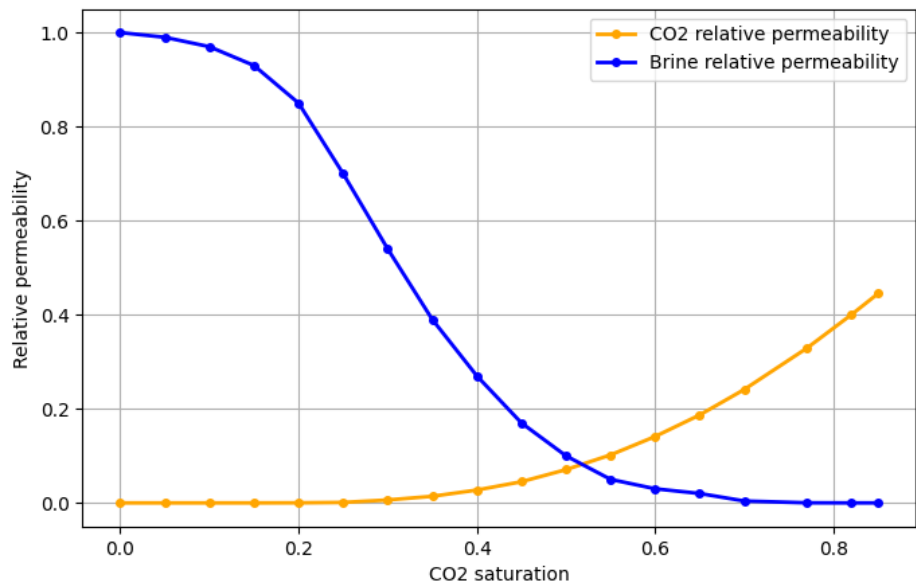


Figure 4. CO₂- brine relative permeabilities plot.

Three vertical injectors I_1 , I_2 , I_3 are placed in the downdip reservoir region to promote controlled upward CO_2 migration, whereas eight brine producers P_1 to P_8 are positioned to manage reservoir pressure and direct the plume movement, Figure 5. Notably, P_6 acts as a pressure sink along the plume path, and P_4 ensures uniform migration despite buoyancy-driven asymmetry. The specific placement of each well is also provided in Table B2, Appendix B.1, and the operational specifications of all wells are summarized in Table 1 where, to facilitate notation, the wells are distinguished according to their operational role into producers and injectors.

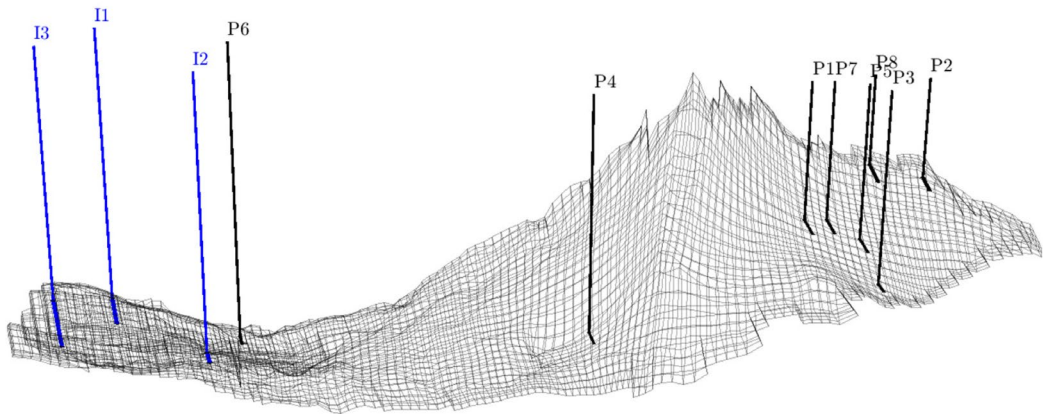


Figure 5. Injection and production optimal well placement in the modeled aquifer (Fotias, Ismail, & Gaganis, 2024).

The storage process is scheduled for 40 years. Field-level constraints limit brine production below 80 Mstb/day based on the brine treatment facilities capacity, while individual production wells are allowed a brine production of 60 Mstb/day as dictated by the wellbore capacity. CO_2 field recycling of 6 MMscf/day at maximum, CO_2 recycling of 4 MMscf/day at maximum per well, bottom-hole pressure (BHP) floored at 3,000 psi for producers and capped at 9,000 psi for injectors, are also constraining the problem solution.

Bottom-hole pressure constraints have different purposes for producers and injectors. On the one hand, for the injectors, the BHP constraint is applied to ensure that the pressure does not exceed the upper fracture limit of the caprock, thereby preventing the activation of potential CO_2 leakage pathways [3]. On the other hand, the BHP constraint for producers is imposed to ensure that reservoir depressurization occurs in a controlled and gradual manner [44]. In this study, a value of 3,000 psi is adopted for producers, following previous work [7] based on the BHP required to lift the produced brine at surface, without explicitly assessing its influence on the overall problem solution.

It is noted that when a constraint is activated at some well, the simulator reduces the CO_2 injection rate or the brine production rate to honor the constraint [45]. For instance, when an injection well reaches its BHP limit of 9,000 psi, the simulator reduces the CO_2 injection rate to maintain the well pressure at its upper bound (9,000 psi).

Initially, the field CO_2 injection target is set at 2 Mtpa, which corresponds to approximately 107.4 MMscf/day, aiming at a total cumulative injection of 1.57 Tscf (equivalent to 80 Mtn) over the project’s lifetime. The term Mtn denotes million (10^6) tonnes and Mtpa million tonnes per annum, while Mscf stands for thousand (10^3) surface cubic feet, MMscf = 10^6 scf. All simulations were run using the CO_2 storage module in Open Porous Media Flow [22].

Table 1. Operational specifications of field wells and well groups.

Group	Maximum brine production (Mstb/day)	Maximum CO_2 production (MMscf/day)	BHP limits (psi)
-------	-------------------------------------	--	------------------

Field	80	6.0	-
Producers	60	4.0	3,000
Injectors	-	-	9,000

It is important to highlight that the only way to predict, identify and guide the system’s behavior during injection stages is by imposing operational constraints on the operating wells, as these wells provide the only direct measurements available under real field conditions [46]. For this reason, it is essential to carefully define the operating constraints of the wells.

5. Results & Discussion

5.1. Case 1: Total Field CO₂ Sequestration Target 80 Mtn with Three Injection Wells

The injection schedule optimization problem involves three wells with constant injection rates I_1, I_2, I_3 (Mtpa) over a period of 40 years. With an operations management decision of achieving a design total target of 2.0 Mtpa × 40 years= 80 Mtn, the total injection rate T is fixed at 2.0 Mtpa. The problem dimensionality reduces to two independent variables I_1, I_2 as described by Equation (26) and subject to the condition (25), in accordance with Section 3.1.

$$0.15 \leq I_i \leq 1.0, i = 1, 2, 3$$

(25)

$$1.0 \leq I_1 + I_2 \leq 1.85$$

(26)

The feasible space for the two variables under the constraint is illustrated in Figure 6. A D-optimal experimental design was conducted to cover this space as effectively as possible by only using a user specified low number of points. When using three sample points, the algorithm selects points 1, 2, and 3, illustrated in Figure 6a, simply capturing the extremes of the feasible space. To gain additional insight into the system’s behavior in other regions of the design space, the D-optimal design algorithm was further asked to generate a fourth distinct combination of the well injection rates I_1, I_2 , indicated by point 4 in Figure 6a. The feasible space restricted by the problem’s constraints is demonstrated by the light green area

After generating the DoE design schedules, simulations were performed using the Open Porous Media (OPM) Flow reservoir simulator to collect the reservoir response, i.e. pressure build up, CO₂ sequestered, CO₂ recycled and brine production. The desired CO₂ sequestration capacity was verified for all sampling points, yielding 80 Mtn in 40 years of operation as seen in Table 2.

Table 2. D-optimal design observations of I_1, I_2 (Mtpa), and simulated $CO_2^{seq}_{sim}$ (Mtn).

I_1 (Mtpa)	I_2 (Mtpa)	$CO_2^{seq}_{sim}$ (Mtn)
1.00	1.00	80
1.00	0.50	80
0.50	1.00	80
0.75	0.75	80

The fact that the desired target storage of 80 Mtn has been directly achieved, even before any optimization takes place, demonstrates that additional sequestration capacity can be realized; for

example, targeting an injection rate of 2.5 Mtpa over 40 years, which would result in a total sequestration of 100 Mtn, might be feasible.

5.2. Case 2: Total Field CO₂ Sequestration Target 100 Mtn with Three Injection Wells

In this case, the objective is to achieve a maximum total target of 2.5 Mtpa × 40 years = 100 Mtn, hence the field injection rate T is set at 2.5 Mtpa. Consistent with the procedure applied in Section 3.1, the feasible design space is now constrained by Equation (27), subject to condition (25).

$$1.50 \leq I_1 + I_2 \leq 2.35 \quad (27)$$

The feasible space for the two variables, along with the four D-optimal points generated, is illustrated in **Figure 6b**. Note that the feasible space has been remarkably reduced due to the increased storage request.

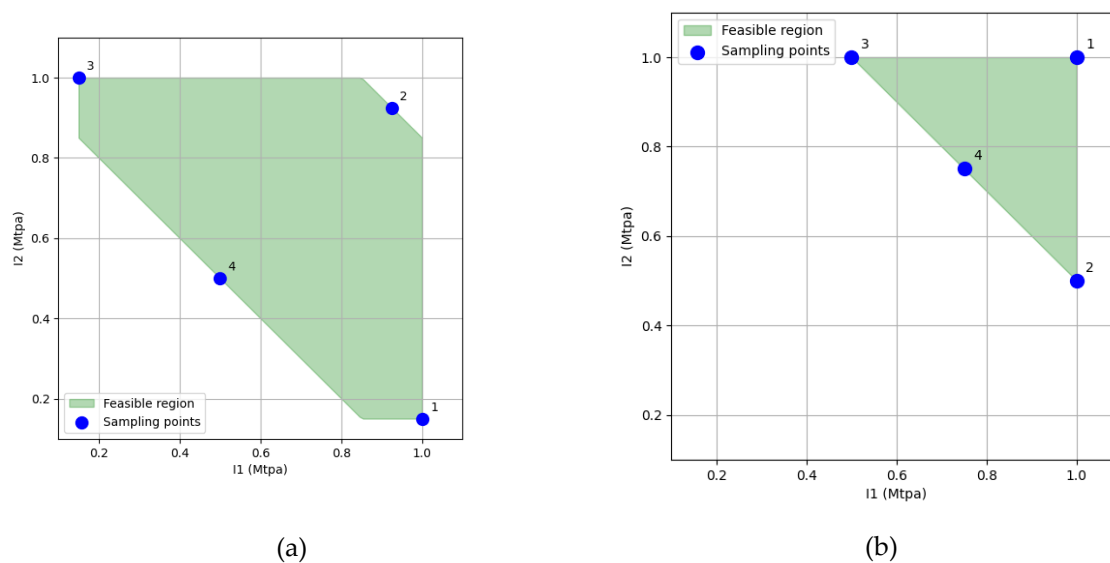


Figure 6. The feasible design areas together with the placement of the design points for the first (a) and the second (b) case respectively.

The simulation results recorded at the end of the reservoir simulation of each D-optimal design point, where $CO_{2, sim}^{seq}$ is CO₂ sequestration in Mtn, are shown in Table 3. The highest sequestration capacity was obtained at point $\xi = (1.0, 0.5, 1.0)$, which lies to the boundary of the feasible space, yielding 93.3 Mtn of CO₂, corresponding to 93.3% of the initial 100 Mtn target. It should be noted that the achieved sequestration capacity is influenced by the constraints imposed on the system. Specifically, under the bottom-hole pressure (BHP) constraints, the injectors were unable to inject as much CO₂ as desired, since the system automatically adjusts injection rates to comply with the prescribed pressure limits. In other words, the simulator is instructed to maintain a constant rate all over the full injection period, however the rate achieved subject to the operational constraints proved to be lower than the desired one.

Table 3. D-optimal design observations of I_1 , I_2 (Mtpa), and simulated $CO_{2, sim}^{seq}$ (Mtn).

I_1 (Mtpa)	I_2 (Mtpa)	$CO_{2, sim}^{seq}$ (Mtn)
0.75	0.75	90.7
1.00	0.50	93.3
0.50	1.00	89.5
1.00	1.00	86.7

Although the D-optimal design alone proved sufficient to identify a high-quality solution, fitting a surrogate model based on these points and the optimization process using a gradient-based algorithm was still pursued. This would validate the surrogate model prediction accuracy and investigate whether an alternative configuration within the feasible space might yield a superior outcome. By fitting a second-order surrogate model with only interaction effects to the D-optimal points, Equation (28) is obtained, where $I_i \in [0, 0.15]$.

$$CO_2^{seq} (Mtn) = 94.30 + 5.60 \cdot I_1 - 2.00 \cdot I_2 - 11.20 \cdot I_1 \cdot I_2 \quad (28)$$

Figures 7, 8 effectively illustrate the relationship between the prediction CO_2^{seq} values and the problem's input variables (I_1, I_2). It also shows that, outside the feasible space, even higher predicted solutions could be obtained by the surrogate model. However, these solutions are not considered not only for lying outside the feasible space, that is they violate the per-well operating bounds of 0.15 Mtpa (lower) and 1 Mtpa (upper), subject to the condition (25). From a fitting point of view, such values are produced by pure extrapolation, hence they are fully uncontrollable.

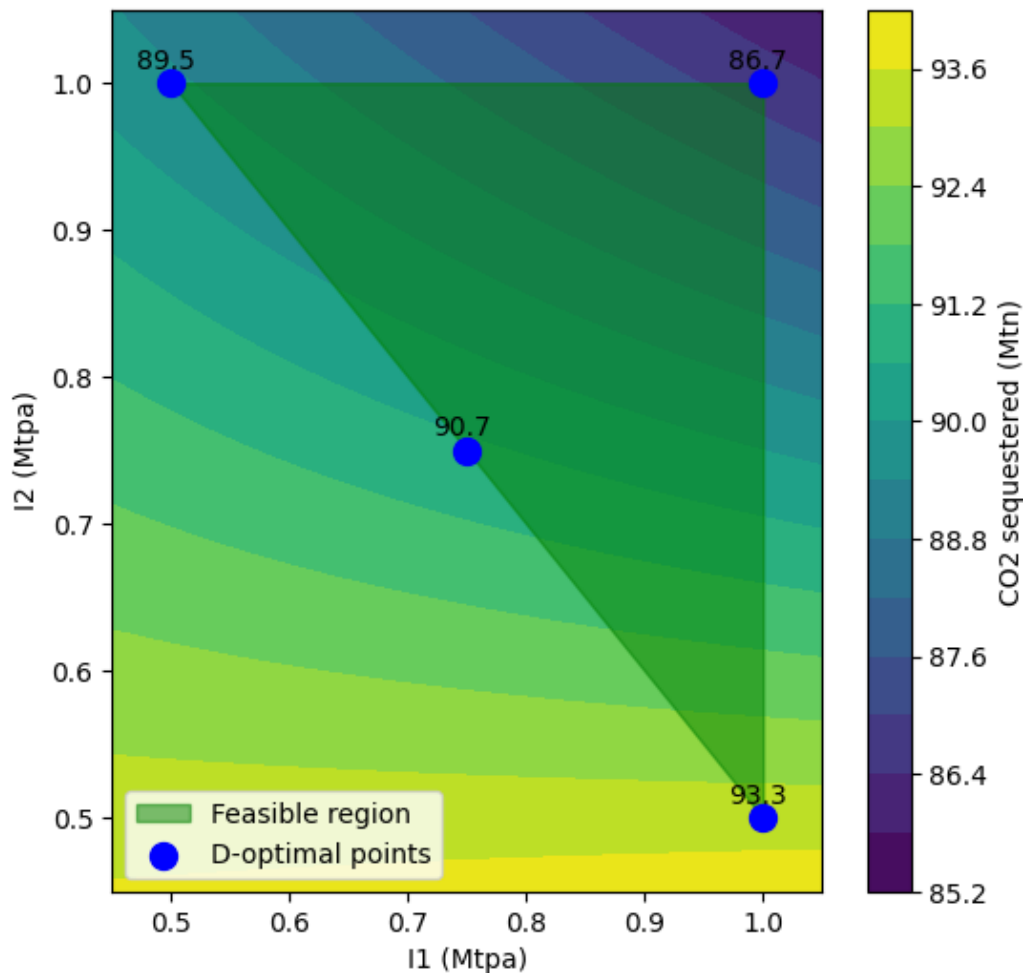


Figure 7. Response surface of the system within the feasible region (light green) and outside the feasible space, modeled using a second-order surrogate with interaction effects.

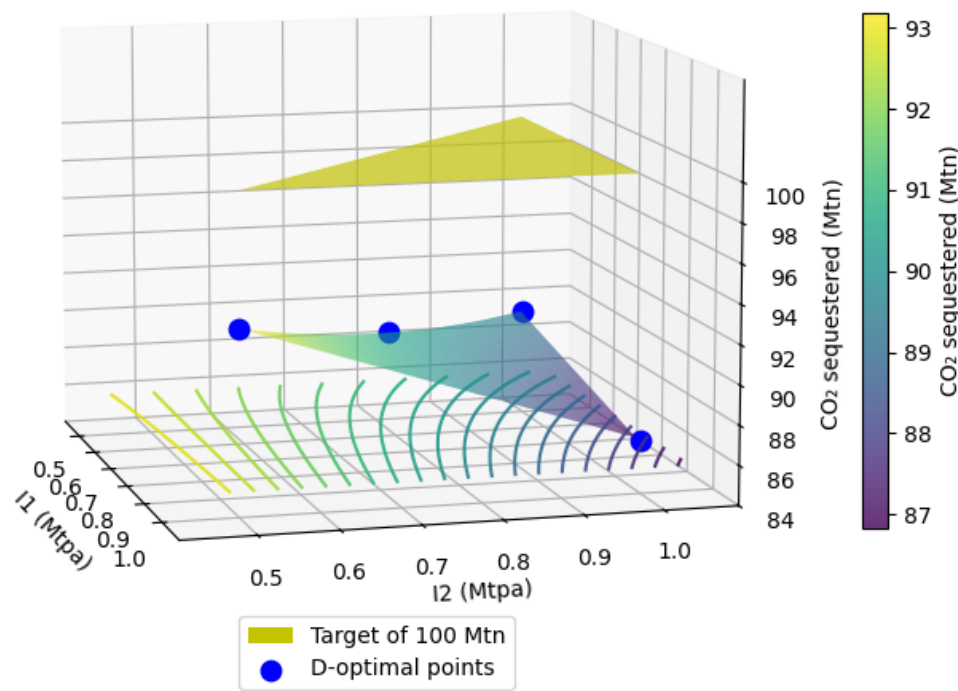


Figure 8. Surrogate model of CO₂ sequestration constructed from the D-optimal design points, overlaid with the targeted 100 Mtn CO₂ plane. The figure illustrates the predicted sequestration surface within the feasible design space, highlighting the alignment of the surrogate model with the D-optimal points and the location of the target plane.

The optimization of the surrogate model resulted in a ξ^* optimal point with $I_1 = 1.0$ Mtpa and $I_2 = 0.5$ Mtpa (corresponding to $I_3 = 1.0$ Mtpa), that is the best already obtained D-optimal result (93.3 Mtn) indicating that the surrogate model optimization did not achieve any better than the DoE sampled points. This is attributed to the fact that the surrogate model proved to be monotonic within the feasible space with no internal optimal. Nevertheless, the predicted CO₂ sequestration at three additional random points within the feasible space was compared against validation simulations. The results (Table D1, Appendix D.1) confirm the high accuracy of the selected surrogate and support the conclusion that no scenario surpasses the initial D-optimal maximum.

Figure 9 demonstrates the CO₂ plume expansion at four distinct time steps. The simulation starts with all cells exhibiting 100% water saturation (not shown). At $t = 10$ years, injection wells I_1 and I_3 have the most pronounced effect on CO₂ migration, as initially planned. At $t = 20$ years, the three CO₂ plumes merge into a single plume, which at $t = 30$ years begins to pass over the anticline. Finally, in $t = 40$ years, the plume has fully traversed the anticline, with a total CO₂ sequestration of 93.3 Mtn.

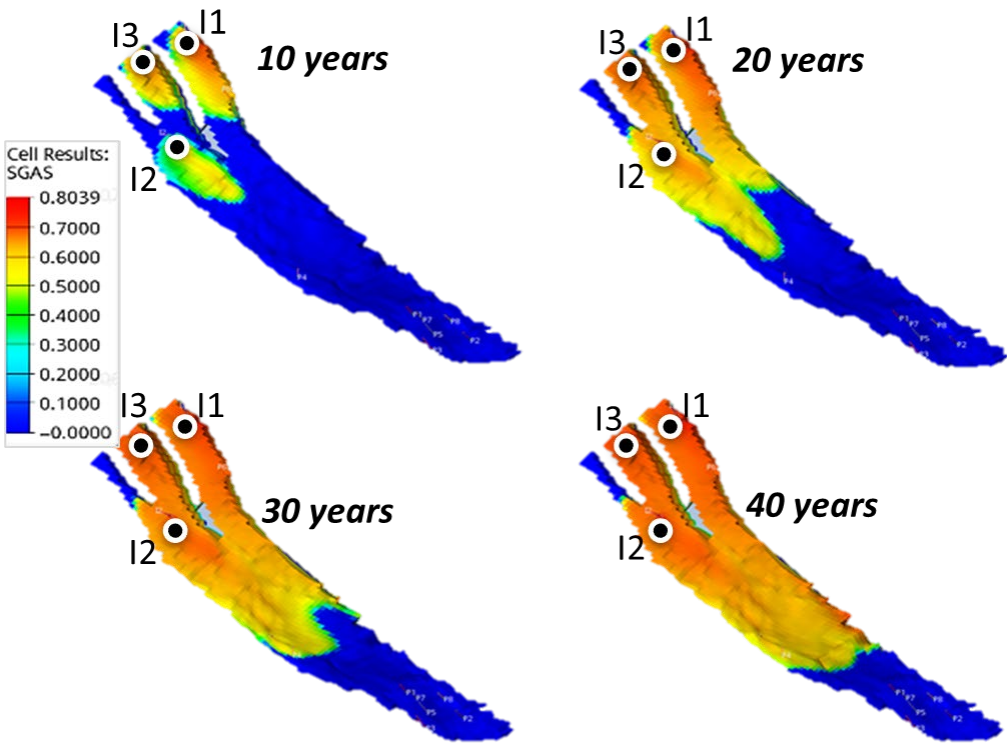


Figure 9. Distribution of CO₂ saturation at four distinct time steps: $t=10$ years, $t=20$ years, $t=30$ years, and $t=40$ years. The figure illustrates the spatial evolution of the CO₂ plume over time.

Figure 10 illustrates the spatial distribution of reservoir pressure at the end of the injection period. The most critical zone is the caprock, where maintaining pressures below the fracture threshold is essential to ensure mechanical integrity during CO₂ injection. The spatial distribution indicates that the maximum reservoir pressure reaches approximately 9,000 psi, remaining safely below the caprock’s fracture-pressure limit and thereby preventing overpressurization. Figure 11 also shows that the initial average reservoir pressure of 4,700 psi decreases during the first five years and then begins to rise, reaching approximately 8,100 psi by the end of the project.

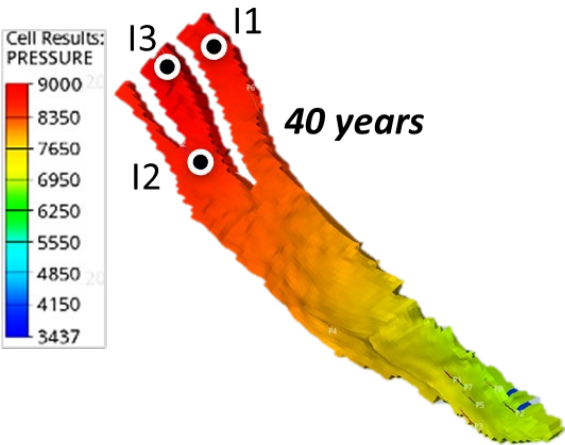


Figure 10. Distribution of pressure build up at the end of the project: $t=40$ years.

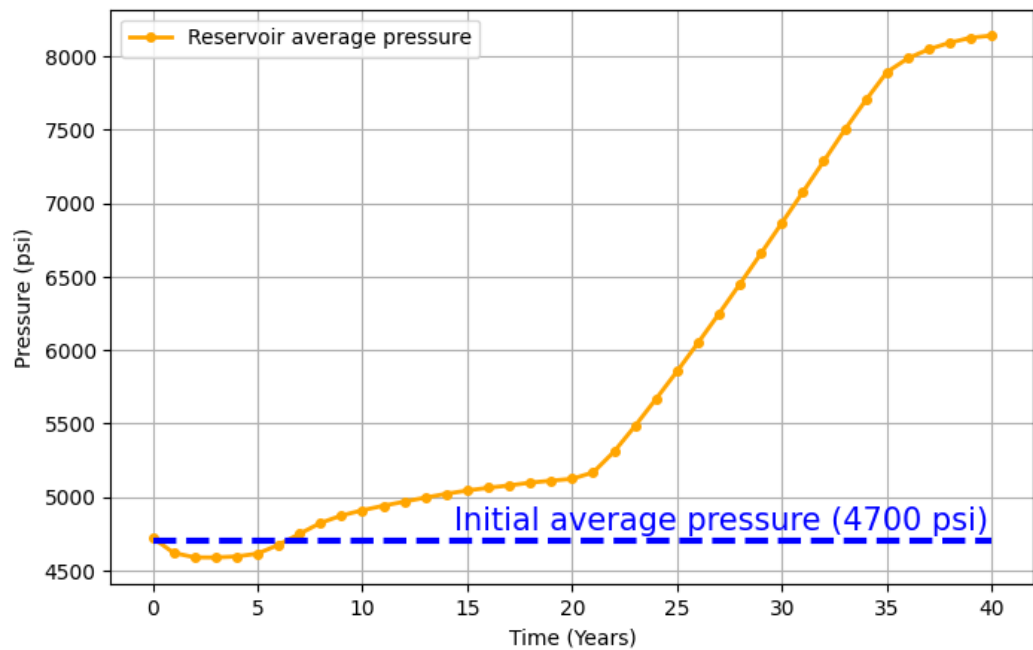


Figure 12. Field performance profiles: Average reservoir pressure.

Note that ξ^* corresponds to the intended injection schedule identified by the optimizer. When the simulator implements that, it is possible the amount of CO₂ actually arriving bottom hole might be less than the optimal due to potential activation of the injectors BHP constraint (9,000 psi). In this case, the injection process begins at a maximum rate of 134.5 MMscf/day (2.5 Mtpa), which is sustained for 34 years. After this period, a decline occurs for the remaining 6 years of the injection period, eventually arriving at 83 MMscf/day (1.54 Mtpa). On the other hand, the field’s maximum water production rate of 77,000 stb/day is never reached. Undesired CO₂ production begins after approximately 5 years of operation and gradually increasing until a peak of 6.0 MMscf/day (0.11 Mtpa), reached after 22 years. The CO₂ field injection and production rates over time are demonstrated in Figure 12.

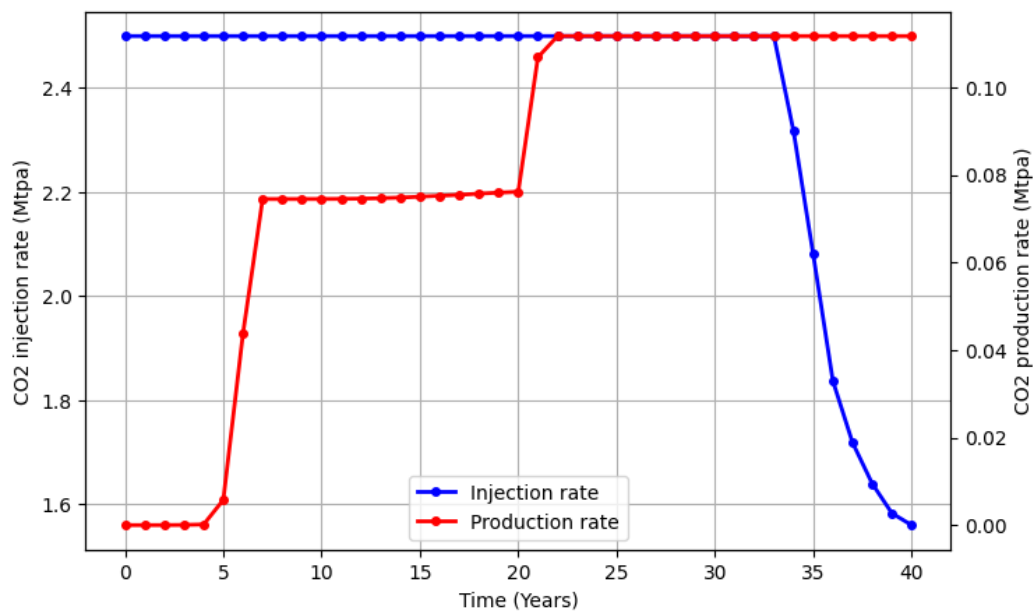


Figure 12. Field performance profiles: CO₂ injection rate (blue) and CO₂ field recycled (red).

The injection wells performance is further illustrated in Figures 13, 14, which show the bottom-hole pressure of each well and their injection rates over time respectively. The initial wells' bottomhole pressure (BHP) appear significantly higher than the pressure at the deepest point of the reservoir. This is due to the localized pressure buildup that develops at the very beginning of CO₂ injection. Moreover, each well shows a different starting BHP, which is explained by two factors: the differences in well depth and the fact that each well begins injecting at a different initial injection rate. I₃ (green) initially maintains a constant injection rate of 53.7 MMscf/day (1.0 Mtpa) for the first 34 years. Subsequently, bottom-hole pressure limit is reached, resulting in a declining injection rate for the remaining 6 years. I₁ (blue) maintains a steady injection rate of 53.7 MMscf/day (1.0 Mtpa) for 35 years until the bottom-hole pressure constraint is activated. After that its rate declines over the final 5 years. I₂ (purple) consistently injects 27.1 MMscf/day (0.50 Mtpa) during the whole operation since no bottom-hole pressure reached.

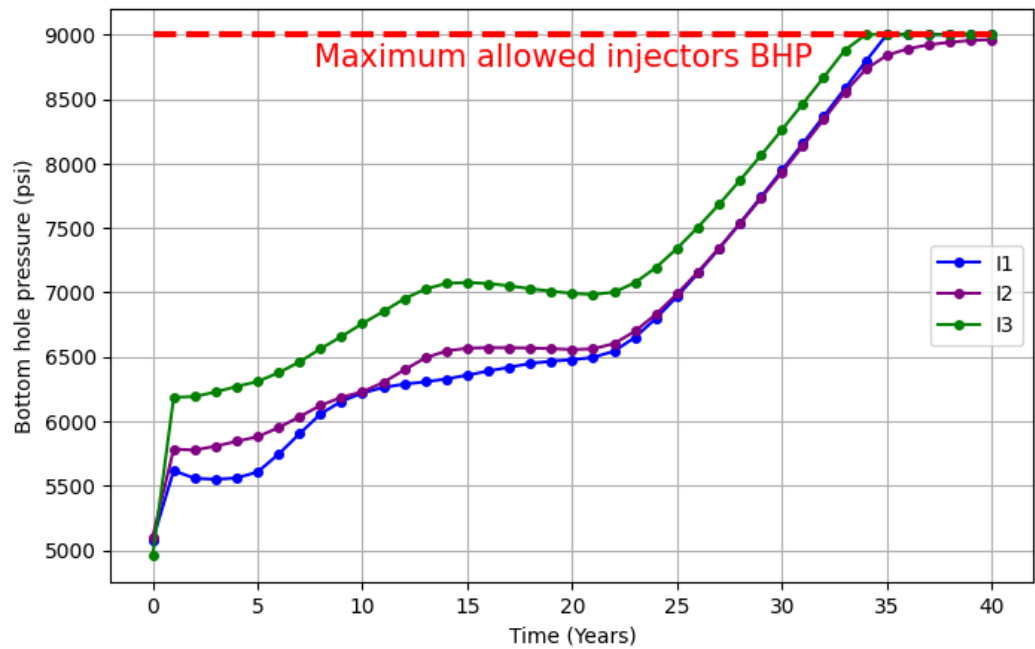


Figure 13. Performance profiles by well: Bottom-hole pressures of wells I₁, I₂, and I₃ over the 40-year injection period.

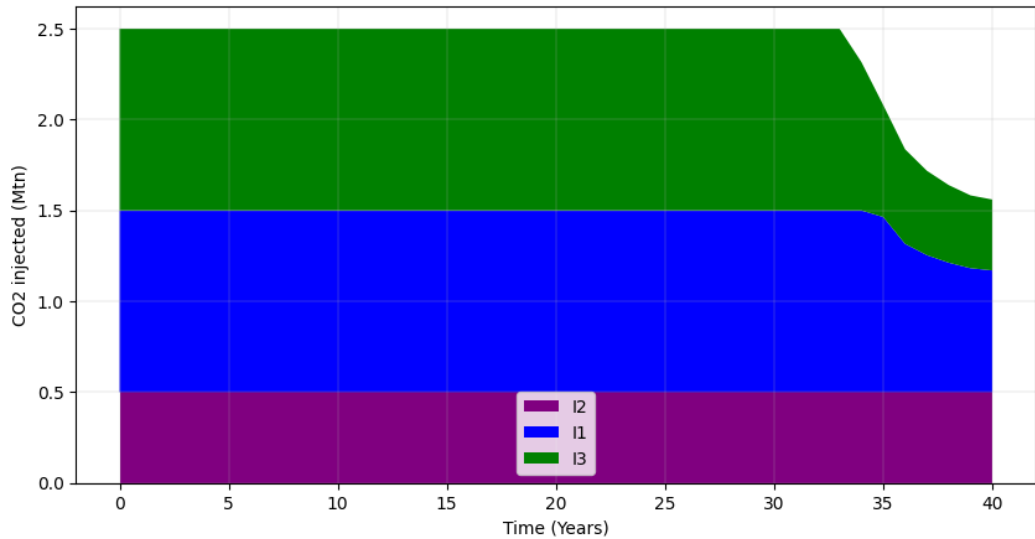


Figure 14. Performance profiles by well: Injection rates of wells I₁, I₂, and I₃ over the 40-year injection period.

5.3. Case 3: Total Field CO₂ Sequestration Target 100 Mtn with Four Injection Wells

Although, the procedure in Case 2 proved sufficient to identify a near-optimal solution, achieving 93.3% of the targeted sequestration capacity, the full goal was not reached, due to the injector’s BHP constraints. To address this shortfall, the potential benefits of adding an additional injection well could be investigated, as this is expected to improve efficiency by enhancing the feasible space dimensionality, thus allowing greater sequestration capacity.

By keeping the injection target at 100 Mtn while introducing one more injector, *I*₄, bounded by the same minimum and maximum rate constraints as the other wells, the design space is expanded. The location of the new introduced well *I*₄ is illustrated in Figure 18. Nine D-optimal sample points were selected within this enhanced 3D feasible space. The sampling points and their distribution are illustrated in Table 4 and in Figure 15 respectively. The operational constraint for the three variables problem in this four-well configuration was defined with Equation (29) subject to the condition (25).

$$1.50 \leq I_1 + I_2 + I_3 \leq 2.35$$

(29)

Simulations were performed using the Open Porous Media (OPM) Flow reservoir simulator, producing a dataset of observations. Table 4 illustrates the simulation results for these D-optimal points, where the amount of CO₂ sequestered is expressed in Mtn. The highest sequestration capacity was obtained at the sixth D-optimal simulation point (*I*₁ = 0.15, *I*₂ = 1.0, *I*₃ = 0.9, *I*₄ = 0.45), yielding a total of 96.1 Mtn of CO₂, hence demonstrating that the additional injection well provides a better, still not optimal solution.

Table 4. D-optimal design observations with corresponding values of *I*₁, *I*₂, *I*₃, *I*₄ (Mtpa), and simulated CO₂ sequestration (Mtn).

Point	ξ=(<i>I</i> ₁ , <i>I</i> ₂ , <i>I</i> ₃ , <i>I</i> ₄) (Mtpa)	<i>CO</i> ₂ ^{seq} _{sim} (Mtn)
1	(0.35, 1.00, 1.00, 0.15)	95.3
2	(0.35, 0.15, 1.00, 1.00)	89.0
3	(0.78, 0.15, 0.57, 1.00)	83.6
4	(0.15, 0.47, 1.00, 0.88)	94.2
5	(1.00, 0.15, 1.00, 0.35)	94.9
6	(0.15, 1.00, 0.90, 0.45)	96.1
7	(0.65, 0.61, 0.62, 0.62)	87.6
8	(1.00, 1.00, 0.35, 0.15)	87.3
9	(0.78, 0.57, 1.00, 0.15)	93.0

By fitting a second-order surrogate model keeping only the interaction effects to the D-optimal points the resulting fitted model is described in Equation (30), where *I*_{*i*} are in Mtpa, *I*_{*i*} ∈ [0, 0.15].

$$CO_2^{seq} (Mtn) = 32.42 + 41.43 \cdot I_1 + 100.09 \cdot I_2 + 48.48 \cdot I_3$$

$$- 72.55 \cdot I_1 \cdot I_2 - 21.47 \cdot I_1 \cdot I_3 - 67.29 \cdot I_2 \cdot I_3$$

(30)

The gradient-based optimization predicted a maximum of approximately 120 Mtn at the point *I*₁ = 0.15, *I*₂ = 1.0, *I*₃ = 0.35 (corresponding to *I*₄ = 1.0), which is physically unrealistic given that the true maximum possible CO₂ sequestration is 100 Mtn (40 years injection of 2.5 Mtpa). This overestimation arises because the second-order model limited to the interaction effects only, is a simplistic approximation of the true response surface, highlighting the limitations of the surrogate model form when extrapolating beyond the sampled domain.

To improve the surrogate model, a full quadratic model was fitted after incorporating the 10th D-optimal design point *I*₁ = 0.78, *I*₂ = 0.57, *I*₃ = 1.0, *I*₄ = 0.15 and running the corresponding simulation, which yielded 91.3 Mtn of sequestered CO₂. The resulting quadratic model is given by:

$$\begin{aligned} CO_2^{seq}(Mtn) = & 17.92 + 126.01 \cdot I_1 + 121.56 \cdot I_2 - 19.74 \cdot I_3 \\ & - 32.88 \cdot I_1^2 - 31.93 \cdot I_2^2 + 66.27 \cdot I_3^2 - 75.27 \\ & \cdot I_1 \cdot I_2 - 61.24 \cdot I_1 \cdot I_3 - 51.07 \cdot I_2 \cdot I_3 \end{aligned} \tag{31}$$

Optimizing this quadratic model with a simple gradient-based algorithm thanks to its constant curvature, the predicted optimum occurs at $I_1 = 0.63, I_2 = 1.00, I_3 = 0.15$ with a maximum predicted CO₂ of 111.55 Mtn, which again is physically unrealistic, though less exaggerated than in the previous attempt. This demonstrates that, while quadratic models improve the fit within the sampled space, they still may overestimate values outside the convex hull of the design points, emphasizing the need for caution required when the built surrogate models are extrapolated.

Considering that a third-order model would require 20 design points, and in order to avoid exceeding computational costs—since the method aims to screen near-optimal solutions with as few simulations as possible—we adopt an alternative approach, described in **Section 3.2**. Specifically, the D-optimal points are first used to fit an RBF surrogate model equipped with gaussian kernel, which is then optimized using PSO. This strategy is designed to limit the total number of simulations to a maximum of 15, while still effectively exploring the feasible space to identify high-quality solutions.

When an RBF model is involved additional simulation runs can also be used to optimize the surrogate model’s hyperparameter γ . Indeed, the RBF surrogate based on the nine D-optimal sampling points dataset was fitted, using a randomly selected shape parameter γ value of 0.85. The PSO algorithm was then ran and identified the new ξ^* solution, which was subsequently simulated using OPM Flow, yielding a CO₂ sequestration of 96.8 Mtn. Next, PSO algorithm was employed again, this time to determine the γ value that minimizes the absolute prediction error of $CO_{2\ sim}^{seq}$ (Mtn) at the ξ^* solution, resulting in γ value equal to 0.73, point 10 in Table 5. The ξ^* solution, along with its $CO_{2\ sim}^{seq}$ response, was then added to the surrogate’s training set, and the RBF was refitted. The results at each iteration of this process are presented in Table 5. Finally, after three consecutive iterations, it was concluded that the identified point (0.21, 0.93, 0.96, 0.40) maximizes the sequestration capacity at 97.4 Mtn, its placement is plotted in Figure 15. This dual approach kept producing new sampling points, on top of the first D-optimal based ones, to improve the local accuracy of the surrogate model, hence its optimum well rates point.

Table 5. Step-by-step optimization with corresponding values of I_1, I_2, I_3, I_4 (Mtpa), γ parameter optimization and simulated CO₂ sequestration (Mtn).

Point	$\xi = (I_1, I_2, I_3, I_4)$ (Mtpa)	γ	$CO_{2\ sim}^{seq}$ (Mtn)
10	(0.15, 0.92, 1.0, 0.43)	0.73	96.8
11	(0.15, 0.95, 1.0, 0.40)	0.77	96.7
12	(0.15, 0.85, 1.0, 0.50)	4.20	96
13	(0.21, 0.93, 0.96, 0.40)	4.23	97.4
14	(0.21, 0.93, 0.96, 0.40)	4.18	97.4
15	(0.21, 0.93, 0.96, 0.40)	4.19	97.4

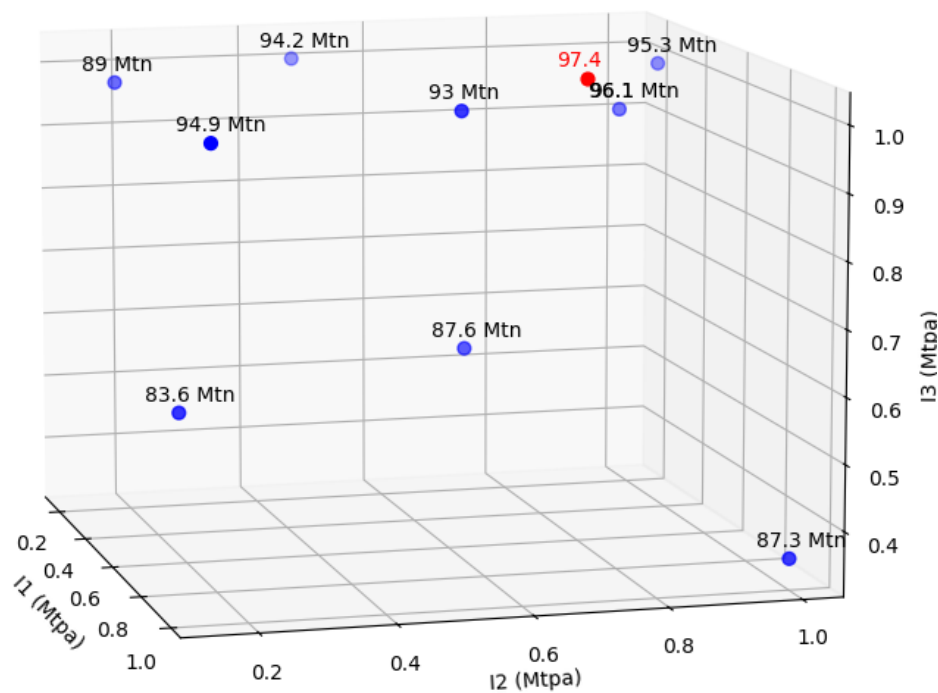


Figure 15. The plot shows the placement of the nine D-optimal design points (blue) and the relative position of the final near-optimal solution (red) obtained from the optimization methodology.

To assess the predictive reliability and generalization capability of the final optimized surrogate, an independent validation was performed. The performance of the final RBF surrogate, with its optimized γ parameter, γ^* , was validated using five additional random testing samples within the feasible design space. The analytical results (Table C2, Appendix C.2) indicate that the absolute total error across the five testing points is 3%, with a standard deviation of 0.313, confirming the high predictive accuracy of the RBF surrogate with gaussian kernel. This validation demonstrates the effectiveness of the proposed screening methodology and its applicability to optimization problems of similar dimensionality.

Furthermore, the effect of the optimized γ parameter was assessed to evaluate how the final model would perform without its optimization, using a randomly chosen value instead. As described in Section 3.2, the γ parameter was optimized to ensure that the RBF surrogate performs optimally, particularly in the region surrounding the maximum solution point. In the proposed methodology, the γ parameter was randomly initialized at 0.85, and the optimizer converged to $\gamma^* = 4.2$, following the path shown in Figure 16.

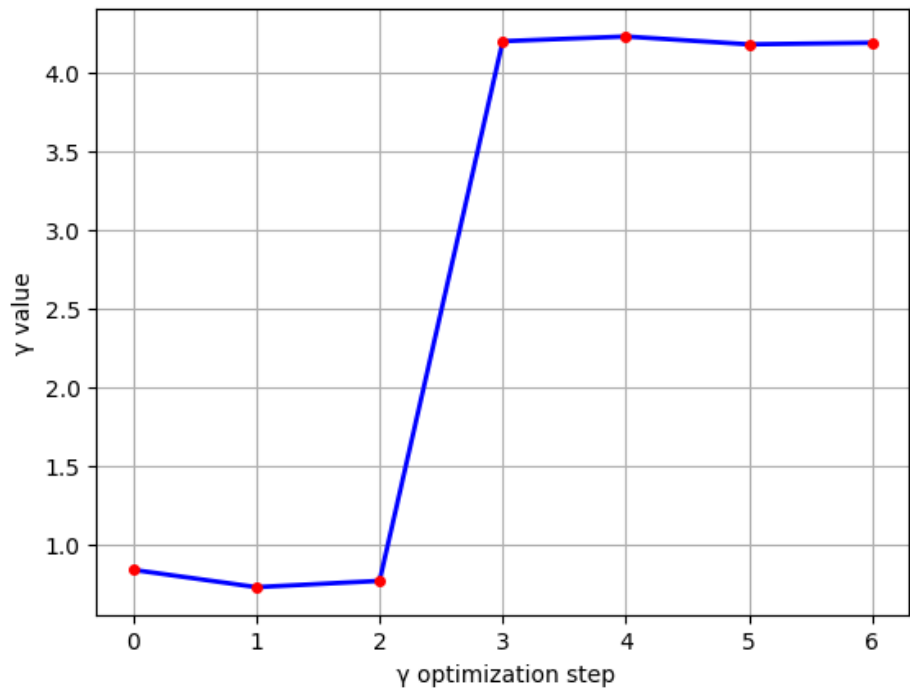


Figure 16. Plot illustrating the optimization path of the γ parameter. Step 0 represents the RBF surrogate constructed from the nine D-optimal points. Points 1 through 6 indicate each γ optimization step, with the procedure concluding at the 15th point (the 6th optimization step).

To highlight the impact of γ optimization, the fifteen-point RBF surrogate was evaluated at the same five testing points used for overall surrogate accuracy, considering eight γ values: 0.77, 2.4, 3.3, 4.2, 5, 10, 50, and 100. The results (Table C3, Appendix C.2) show that low γ values (<2) lead to a high total absolute error (~8%), intermediate values around 4 minimize the error (~3.2%), and high values (>10) cause a slight increase. Figure 17 illustrates these trends, including a zoomed-in view around the minimum absolute error. These results validate that the proposed algorithm reliably determines γ^* , achieving optimal surrogate performance.

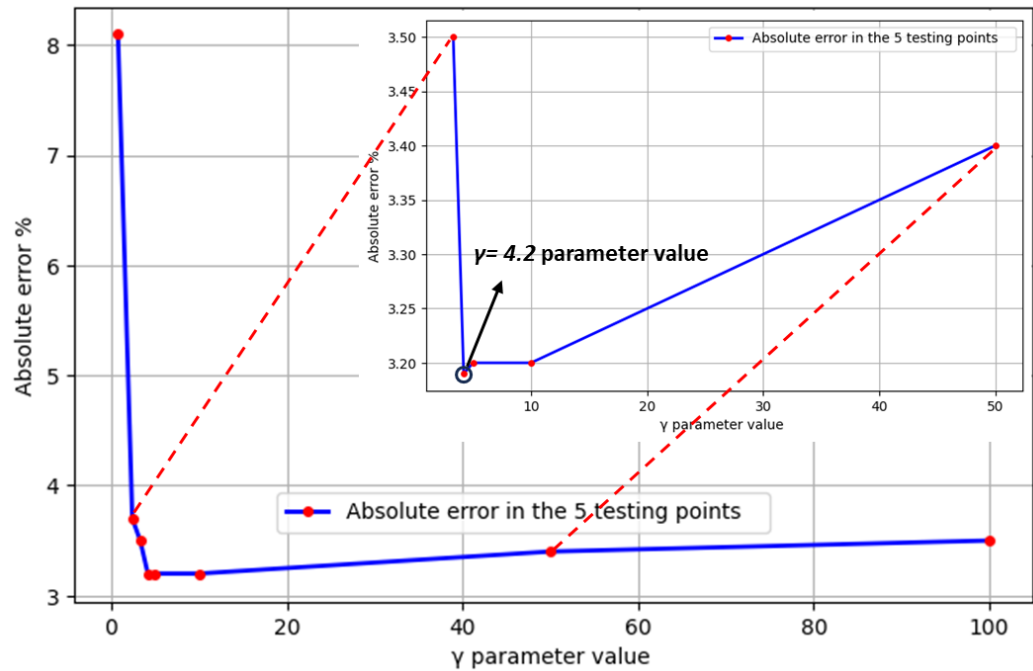


Figure 17. Plot demonstrating the influence of the γ parameter (eight different cases) on the absolute total error of the five testing points (expressed as relative error, %).

The predicted CO₂ sequestration at the optimum corresponds to 97.4 Mtn, representing an additional 4.1 Mtn compared to the three-well model and bridging the performance of the optimized injection schedule further close to the desired 100 Mtn target. This result demonstrates that adding an additional operating well can efficiently achieve a better outcome, since the capacity increased. Furthermore, it indicates that the system may not be able to exceed the upper bound of 100 Mtn under the current operational constraints. This proposed approach not only identifies a physically consistent solution—since, in contrast to previous polynomial surrogates (second order), it yielded a value below the 100 Mtn—but also provides an improved performance relative to the original three-well configuration.

The detailed simulation results corresponding to the maximum response identified from the optimizer are also presented. Similarly to the results in Case 2, Figure 18 illustrates the CO₂ plume at four distinct time steps. At $t=10$ years, injection wells I_2 and I_3 exert the strongest influence on CO₂ migration, consistent with the initial design. At $t=20$ years, the I_2 , I_3 , I_4 CO₂ plumes have merged into a single stream that has started to migrate across the anticline. From $t=30$ years to $t=40$ years, the plume has fully traversed the anticline, with a total CO₂ sequestration of 1.91 Tscf, approximately 97.4 Mtn.

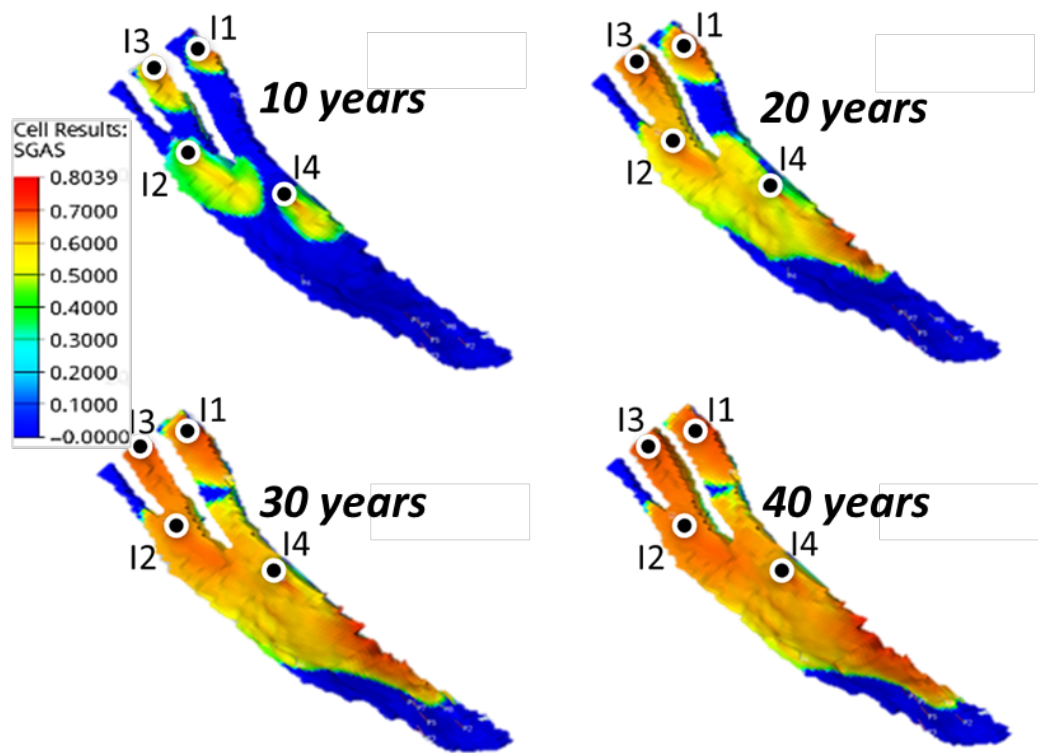


Figure 18. 3D distribution of CO₂ saturation at $t=10$ years, $t=20$ years, $t=30$ years, and $t=40$ years.

Figure 19 demonstrates the spatial distribution of reservoir pressure at the end of the injection period. It is noted the criticality to ensure that the pressure remains below the caprock fracture threshold while the reservoir is being pressurized by CO₂ injection. Furthermore Figure 20, indicate that the maximum reservoir pressure reaches 9,000 psi, remaining safely below the caprock fracture pressure limit and thus preventing overpressurization.

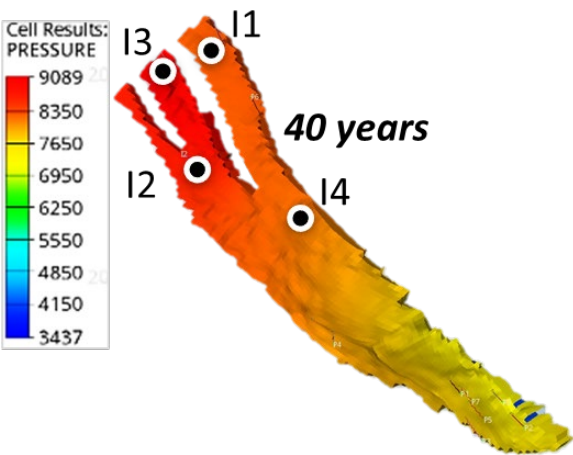


Figure 19. Distribution of pressure build up at the end of the project at $t=40$ years.

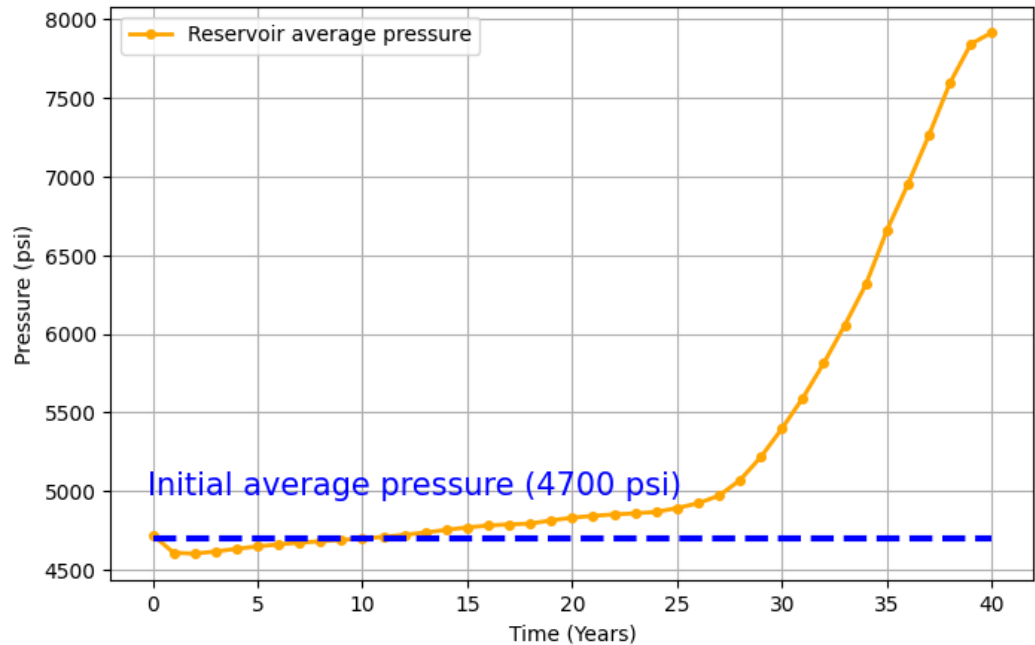


Figure 20. Field performance profiles: Average reservoir pressure.

The optimal scenario field CO₂ injection and recycling rates are demonstrated in Figure 21. The injection process begins at a maximum rate of 134.5 MMscf/day (2.5 Mtpa), which is sustained for 38 years, that is almost all the duration of the project. A decline is observed only during the last 2 years of the injection period, eventually arriving at 113.8 MMscf/day (2.12 Mtpa), which still corresponds to 85% of the design rate on the very last day. On the other hand, the field’s maximum water production rate of 77 Mstb/day is never encountered. Undesired CO₂ production due to recycling through the brine producers begins after approximately 18 years of operation gradually increasing to a peak of 6.0 MMscf/day (0.11 Mtpa), reached at $t=27$ years, and remains constant until the end of the project.

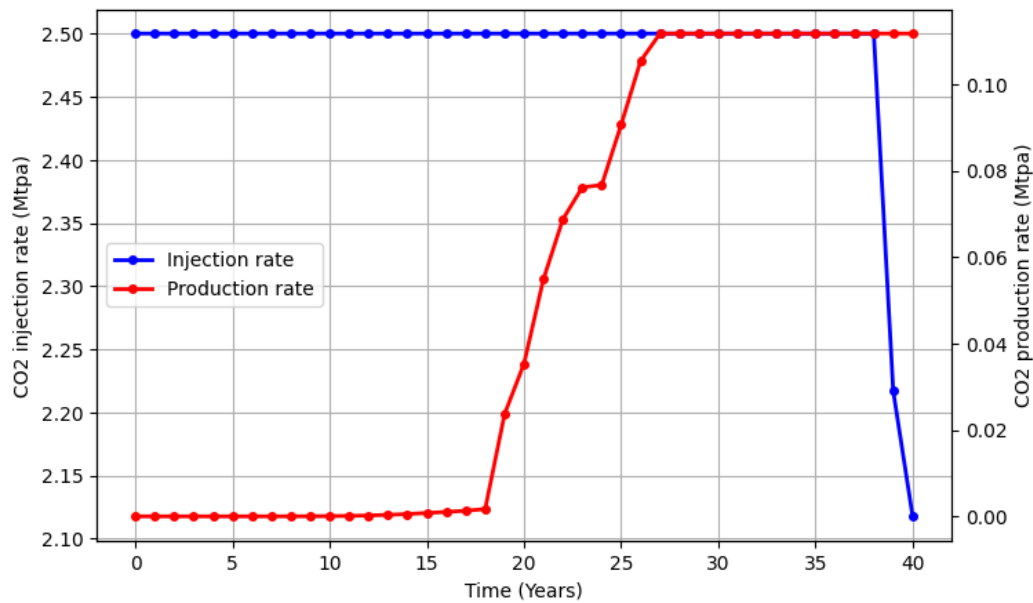


Figure 21. Field performance profiles: CO₂ injection rate (blue) and CO₂ field recycled (red).

The individual well performance at the optimized schedule is further illustrated in Figures 22, 23, which show the bottom-hole pressures of each well with and their injection rates over time. I₄ shows a lower starting BHP compared to the other wells. This is explained by the shallower depth of its perforations—being located at a higher structural position.

I₃ (green) maintains a steady injection rate of 51.5 MMscf/day (0.96 Mtpa) for 38 years. Subsequently, the bottom-hole pressure limit is reached, resulting in a declining injection rate for the last 2 year. I₂ (purple area) maintains a steady injection rate of 49.9 MMscf/day (0.93 Mtpa) for 39 years until the bottom-hole pressure constraint is reached. I₄ (orange) and I₁ (blue) consistently inject 21.5 MMscf/day (0.40 Mtpa) and 11.2 MMscf/day (0.21 Mtpa) respectively, throughout the entire 40-year period with no decline. The well rate profiles are in alignment to the pressure profiles which confirm that I₃ gets BHP constrained at $t=38$ years, while I₄ does so at $t=39$ years. I₁ and I₂ are not limited, hence they maintain their design rate, as dictated by ξ^* , until the completion of the project.

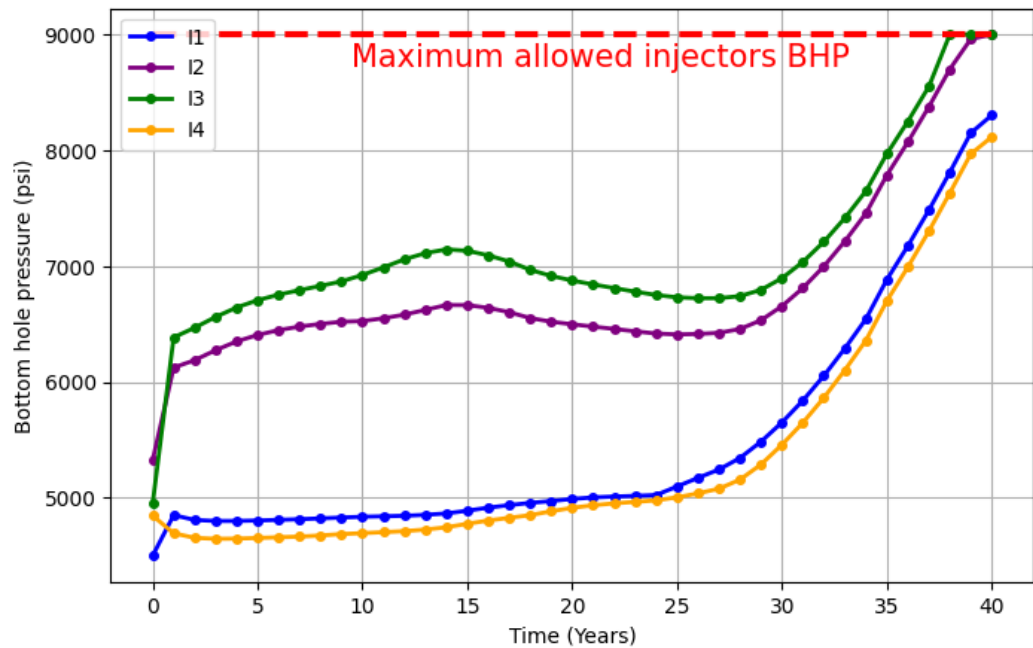


Figure 22. Performance profiles by well: Bottom-hole pressures of wells I₁, I₂, I₃ and I₄ over the 40-year injection period.

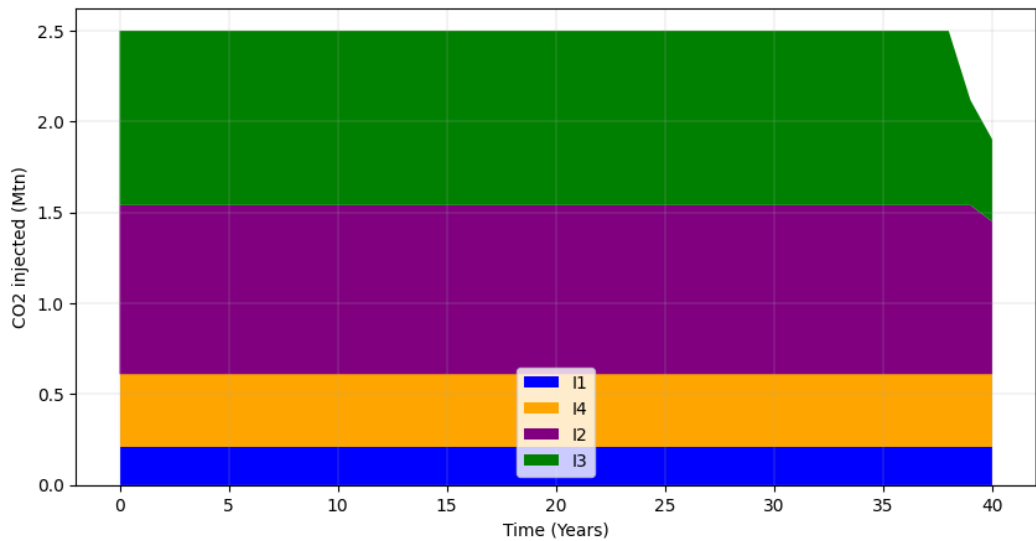


Figure 23. Performance profiles by well: Injection rates of wells I₁, I₂, I₃ and I₄ over the 40-year injection period.

5.4. Discussion

Table 6 summarizes the outcomes of the various CO₂ injection scenarios examined. Each case differs in the number of injection wells and the targeted sequestration amount. The achieved CO₂ sequestration closely matches the targets, demonstrating the effectiveness of the designed injection strategies. Table 6 also lists the number of D-optimal design points and the total simulations required for each scenario, highlighting the efficiency of the experimental design approach.

Table 6. Cases comparison table.

Case	Number of injectors	Targeted CO ₂ sequestration (Mtn)	Achieved CO ₂ sequestration (Mtn)	D-optimal design points	Simulations required
Case 1	3	80	80	4	4
Case 2	3	100	93.3	4	5
Case 3	4	100	97.4	9	15

The results indicate that, under the specified operational constraints, the system captures significantly more CO₂ than Case 1’s target of 80 Mtn in Case 2, despite having the same number of wells. In Case 3, although the proposed methodology successfully captures system relationships and provides a physically consistent solution, the addition of a fourth well yields only a marginal increase in sequestration, 4.1 Mtn, compared to Case 2 with three wells. From a reservoir engineering perspective, this solution is less attractive economically, as the additional well increases both CAPEX and OPEX disproportionately to the small gain in storage.

Despite the diminishing returns observed when adding more wells, the methodology itself proves robust and effective in capturing the system’s behavior and identifying optimal injection configurations making it a reliable tool for decision-making. Beyond maximizing CO₂ storage, the framework can be readily extended to more complex, multi-objective optimization problems, such as maximizing net present value (NPV) or balancing economic efficiency with storage performance. This

flexibility demonstrates that the proposed approach is not only effective for single-objective optimization but also capable of supporting more comprehensive reservoir management strategies.

For future extensions, the methodology can be adapted to incorporate time-dependent injection strategies by splitting the project period into multiple time steps and optimizing injection controls for each step [11]. This approach allows more dynamic management of CO₂ injection, potentially improving performance or balancing economic and operational objectives over time. The same surrogate-based framework is likely suitable for this extension; however, increasing the number of time steps also increases the problem's dimensionality, creating a trade-off between computational cost and surrogate accuracy. Careful experimental design and potentially adaptive surrogate updating would be needed to maintain efficiency while preserving reliable predictions.

Beyond CO₂ storage, the methodology provides a practical and scalable tool for real-world dynamic subsurface management [47] and can be extended to other applications, including open reservoir CO₂ storage systems [48], geothermal production [49] and hydrogen storage [50]. The effect of CO₂-stream impurities on storage mechanisms can also be investigated, further broadening the applicability and robustness of the approach [51].

6. Conclusions

The proposed methodology addresses this need by integrating experimental design, surrogate modeling, and intelligent optimization to efficiently identify near-optimal injection strategies under realistic physical, operational, and commercial constraints. Rather than relying on computationally expensive exhaustive reservoir simulations, the approach minimizes simulation cost while preserving decision quality, making it particularly attractive for portfolio screening, feasibility studies, and scenario evaluation under contractual injection commitments.

Specifically, the methodology implements an integrated optimization framework that combines D-optimal experimental design, radial basis function (RBF) surrogate modeling, and particle swarm optimization (PSO). This framework is specifically tailored to support early-stage decision making and operational planning, where rapid assessment of feasible injection strategies is critical for both technical and commercial viability.

From a commercial and strategic perspective, large-scale CO₂ storage (CCS) projects must balance three competing objectives: maximizing contracted CO₂ intake, minimizing operational and computational costs, and reducing technical and financial risk over multi-decadal project lifetimes. Injection strategies directly affect a project's ability to honor long-term CO₂ offtake agreements with emitters, ensure stable revenue streams, and protect capital investments in wells and surface infrastructure. Consequently, decision-makers require optimization tools that are not only technically robust, but also fast, interpretable, and suitable for early-stage screening and operational planning.

Author Contributions: Conceptualization, E.T.D and V.G.; methodology, E.T.D; software, E.T.D. and S.P.F.; validation, E.T.D., V.G. and S.P.F.; formal analysis, E.T.D. and V.G.; investigation, E.T.D.; data curation E.T.D.; writing- original draft preparation, E.T.D. and I.I.; writing- review and editing, V.G. and S.B.; visualization, E.T.D.; supervision V.G.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Appendix A

Appendix A.1.

$$u_i^{k+1} = \omega u_i^k + c_1 r_1 (P_{best} - x_i^k) + c_2 r_2 (G_{best} - x_i^k), \quad (A1)$$

$$x_i^{k+1} = x_i^k + u_i^{k+1} \quad (A2)$$

In these expressions, ω denotes the inertia weight, which balances global and local search abilities. A larger value of ω enhances global exploration, while a smaller value strengthens local exploitation. The parameters c_1 and c_2 , are the acceleration coefficients, typically both are assigned values between 0 and 2. Moreover, the random variables r_1 and r_2 are uniformly distributed in the interval $[0,1]$ [38].

Appendix A.2.

Table A1. Subordinator PSO algorithm specifications, for identifying optimal ε parameters.

Acceleration coefficients c_1 & c_2	γ value searching bounds	Particle swarm size	Maximum iterations
0.5	(0.1, 100)	100	100

Table A2. Primary PSO algorithm specifications, for identifying optimal ξ^* .

Acceleration Coefficients c_1 & c_2	I_i well rates searching bounds	Particle swarm size	Maximum iterations
0.5	(0.15, 1.0)	100	100

Appendix B

Appendix B.1.

Brine relative permeability is 1.0 at zero CO₂ saturation and decreases sharply as CO₂ saturation rises, while CO₂ relative permeability starts near zero and gradually increases, reflecting typical two-phase flow in water-wet reservoirs. The connate water saturation ($Sw_c \approx 0.23$) leaves only 77% of the pore space available for CO₂ storage, and the residual gas saturation ($Sgr \approx 0.20$) indicates limited CO₂ residual trapping.

Table A3. Relative permeabilities table.

CO ₂ saturation	CO ₂ relative permeability	Brine relative permeability
0.00	0.0000	1.0000
0.05	0.0000	0.9900
0.10	0.0000	0.9700
0.15	0.0000	0.9300
0.20	0.0002	0.8500
0.25	0.0010	0.7000
0.30	0.0062	0.5400
0.35	0.0140	0.3900
0.40	0.0273	0.2700
0.45	0.0450	0.1700
0.50	0.0707	0.1000
0.55	0.1020	0.0500
0.60	0.1412	0.0300
0.65	0.1870	0.0200
0.70	0.2412	0.0040

0.77	0.3288	0.0002
0.82	0.4000	0.0000
0.85	0.4450	0.0000

Table A4. Wells’ locations, connections and skin factors.

Well name	Group	Wellhead location grid cells (x, y)	Connection grid cells (upper, lower)	Skin factor
I ₁	Injectors	84, 15	1, 4	0.708
I ₂	Injectors	60, 26	1, 4	0.708
I ₃	Injectors	74, 9	1, 4	0.708
P ₁	Producers	62, 100	1, 3	0.708
P ₂	Producers	70, 117	4, 4	0.708
P ₃	Producers	57, 108	1, 3	0.708
P ₄	Producers	53, 72	4, 4	0.708
P ₅	Producers	62, 107	1, 3	0.708
P ₆	Producers	76, 29	4, 4	0.708
P ₇	Producers	63, 103	1, 3	0.708
P ₈	Producers	70, 110	4, 4	0.708

Appendix C

Appendix C.1.

Table A5. Additional simulations for error estimation of the Case 2 surrogate.

I_1 (Mtpa)	I_2 (Mtpa)	CO_2^{seq} (Mtn)	$CO_2^{seq}_{sim}$ (Mtn)	Absolute Error (%)
1	0.75	90	90.75	0.8
0.8	0.9	88.9	89.2	0.3
0.9	0.7	90.9	91.2	0.32

Appendix C.2.

Table A6. Additional simulations for error estimation of the Case 3 surrogate.

(I_1, I_2, I_3, I_4) (Mtpa)	$CO_2^{seq}_{sim}$ (Mtn)	CO_2^{seq} (Mtn) ($\gamma^*=4.2$)	Absolute Error (%)
(0.34, 0.78, 0.73, 0.65)	92.7	90.4	2.54
(0.91, 0.22, 0.51, 0.86)	88.7	85.9	3.26
(0.61, 0.34, 0.65, 0.90)	88.7	86.2	2.9

(0.84, 0.16, 0.83, 0.67)	93.8	91.1	2.96
(0.96, 0.44, 0.23, 0.87)	87	84.2	3.32

Table A7. Parameter’s γ influence estimation in RBF surrogate performance.

γ parameter value	Absolute Error (%)
$\gamma_1= 0.77$	8.10
$\gamma_2= 2.4$	3.70
$\gamma_3= 3.3$	3.50
$\gamma_4= 4.2$	3.20
$\gamma_5= 5$	3.20
$\gamma_6= 10$	3.20
$\gamma_7= 50$	3.40
$\gamma_8= 100$	3.50

References

1. H. Lee and J. Romero, “AR6 Synthesis Report,” Geneva, 2023.

2. I. P. on C. C. (IPCC), Ed., “Annex I: Glossary,” in *Global Warming of 1.5°C: IPCC Special Report on Impacts of Global Warming of 1.5°C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*, Cambridge: Cambridge University Press, 2022, pp. 541–562. doi: DOI: 10.1017/9781009157940.008.

3. I. Ismail and V. Gaganis, “Carbon Capture, Utilization, and Storage in Saline Aquifers: Subsurface Policies, Development Plans, Well Control Strategies and Optimization Approaches—A Review,” Jun. 01, 2023, MDPI. <https://doi.org/10.3390/cleantech5020031>.

4. L. Fu *et al.*, “Research progress on CO2 capture and utilization technology,” *Journal of CO2 Utilization*, vol. 66, p. 102260, Dec. 2022. <https://doi.org/10.1016/J.JCOU.2022.102260>.

5. H. Wang, J. Chen, and Q. Li, “A Review of Pipeline Transportation Technology of Carbon Dioxide,” *IOP Conf Ser Earth Environ Sci*, vol. 310, no. 3, p. 032033, 2019. <https://doi.org/10.1088/1755-1315/310/3/032033>.

6. S. Fotias, V. Gaganis, and S. Bellas, “Integrated Raw Material Approach to Sustainable Geothermal Energy Production: Harnessing CO2 for Enhanced Resource Utilization,” *Technical Annals*, vol. 1, pp. 1–20, Jun. 2024. <https://doi.org/10.12681/ta.37117>.

7. S. P. Fotias, I. Ismail, and V. Gaganis, “Optimization of Well Placement in Carbon Capture and Storage (CCS): Bayesian Optimization Framework under Permutation Invariance,” *Applied Sciences (Switzerland)*, vol. 14, no. 8, Apr. 2024. <https://doi.org/10.3390/app14083528>.

8. I. Ismail and V. Gaganis, "Well Control Strategies for Effective CO₂ Subsurface Storage: Optimization and Policies," MDPI AG, Jan. 2024, p. 74. <https://doi.org/10.3390/materproc2023015074>.
9. G. Qi, W. Zhang, and W. Huang, "Urban Near-surface Characterization Using Passive Fiber-optic Accelerometer Array Records: A Case Study," vol. 2025, no. 1, pp. 1–3, 2025. <https://doi.org/10.3997/2214-4609.202572170>.
10. M. Amvrazis, I. Ismail, D. Ktenas, V. Gaganis, E. Tartaras, and A. Stefatos, "Prinos CO₂ Storage Permit Application: An Integrated Technical and Regulatory Workflow," vol. 2025, no. 1, pp. 1–3, 2025. <https://doi.org/10.3997/2214-4609.202536047>.
11. I. Ismail, S. P. Fotias, and V. Gaganis, "A Holistic Framework for Optimizing CO₂ Storage: Reviewing Multidimensional Constraints and Application of Automated Hierarchical Spatiotemporal Discretization Algorithm," *Energies (Basel)*, vol. 18, no. 22, 2025. <https://doi.org/10.3390/en18225926>.
12. "citations-20251214T204338".
13. J. A. White *et al.*, "Geomechanical behavior of the reservoir and caprock system at the In Salah CO₂ storage project," *Proceedings of the National Academy of Sciences*, vol. 111, no. 24, pp. 8747–8752, 2014. <https://doi.org/10.1073/pnas.1316465111>.
14. K. Zhang, H. C. Lau, and Z. Chen, "Extension of CO₂ storage life in the Sleipner CCS project by reservoir pressure management," *J Nat Gas Sci Eng*, vol. 108, p. 104814, Dec. 2022. <https://doi.org/10.1016/J.JNGSE.2022.104814>.
15. I. Ismail, S. P. Fotias, D. Avgoulas, and V. Gaganis, "Integrated Black Oil Modeling for Efficient Simulation and Optimization of Carbon Storage in Saline Aquifers," *Energies (Basel)*, vol. 17, no. 8, Apr. 2024. <https://doi.org/10.3390/en17081914>.
16. J. Chen, E. Gildin, and G. Kompantsev, "Optimization of pressure management strategies for geological CO₂ storage using surrogate model-based reinforcement learning," *International Journal of Greenhouse Gas Control*, vol. 138, p. 104262, 2024. <https://doi.org/10.1016/j.ijggc.2024.104262>.
17. E. M. Kanakaki, S. P. Fotias, and V. Gaganis, "Application of an Automated Machine Learning-Driven Grid Block Classification Framework to a Realistic Deep Saline Aquifer Model for Accelerating Numerical Simulations of CO₂ Geological Storage," *Processes*, vol. 13, no. 8, 2025. <https://doi.org/10.3390/pr13082658>.
18. R. C. St. John and N. R. Draper, "D-Optimality for Regression Designs: A Review," in *Technometrics*, vol. 17, 1975.
19. Song and Zhaojie, "D-optimal design for Rapid Assessment Model of CO₂ flooding in high water cut oil reservoirs," *J Nat Gas Sci Eng*, vol. 21, 2014.
20. J. Zhang, M. Delshad, and K. Sepehrnoori, "Development of a framework for optimization of reservoir simulation studies," *J Pet Sci Eng*, vol. 59, no. 1–2, pp. 135–146, Oct. 2007. <https://doi.org/10.1016/J.PETROL.2007.03.013>.
21. Al-Mudhafar, Watheq J., Wojtanowicz, and Andrew K., "Leveraging Designed Simulations and Machine Learning to Develop a Surrogate Model for Optimizing the Gas–Downhole Water Sink–Assisted Gravity Drainage (GDWS-AGD) Process to Improve Clean Oil Production," *Processes*, vol. 12, 2024.
22. D. Baxendale, "OPEN POROUS MEDIA OPM Flow Reference Manual," 2024.
23. A. F. Rasmussen *et al.*, "The Open Porous Media Flow reservoir simulator," *Computers & Mathematics with Applications*, vol. 81, pp. 159–185, Jan. 2021. <https://doi.org/10.1016/J.CAMWA.2020.05.014>.
24. L. Kang, X. Deng, and R. Jin, "Bayesian D-Optimal Design of Experiments with Quantitative and Qualitative Responses," Apr. 2023, [Online]. Available: <http://arxiv.org/abs/2304.08701>
25. William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, *Numerical Recipes in C The Art of Scientific Computing*, Second. Cambridge: Press Syndicate of the University of Cambridge, 1992.

26. Z. Song *et al.*, "D-optimal design for Rapid Assessment Model of CO₂ flooding in high water cut oil reservoirs," *J Nat Gas Sci Eng*, vol. 21, pp. 764–771, Nov. 2014. <https://doi.org/10.1016/J.JNGSE.2014.10.005>.
27. P. F. de Aguiar, B. Bourguignon, M. S. Khots, D. L. Massart, and R. Phan-Than-Luu, "D-optimal designs," *Chemometrics and Intelligent Laboratory Systems*, vol. 30, no. 2, pp. 199–210, Dec. 1995. [https://doi.org/10.1016/0169-7439\(94\)00076-X](https://doi.org/10.1016/0169-7439(94)00076-X).
28. J. P. D. K, Thomas S. Asche, and Johanna Kleinekorte, "BoFire: Bayesian Optimization Framework Intended for Real Experiments," *ArXiv*, 2024.
29. Jorge Nocedal and Stephen J. Wright, *Numerical Optimization*, 2nd ed. California: New York: Springer, ©2006, 2006.
30. E. Ostertagová, "Modelling using Polynomial Regression," *Procedia Eng*, vol. 48, pp. 500–506, Jan. 2012. <https://doi.org/10.1016/J.PROENG.2012.09.545>.
31. F. Pedregosa, G. Varoquaux, A. Gramfort, and V. Michel, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, 2011.
32. Christopher M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer Science + Business Media, 2006.
33. Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, and Matt Haberland, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nat Methods*, vol. 17, 2020.
34. Y. Zhang, J. Cao, B. Zhang, X. Zheng, and W. Chen, "A comparative study of different radial basis function interpolation algorithms in the reconstruction and path planning of γ radiation fields," *Nuclear Engineering and Technology*, vol. 56, no. 7, pp. 2806–2820, Jul. 2024. <https://doi.org/10.1016/J.NET.2024.02.043>.
35. Adrijan Rogan, Andrej Kolar-Požun, and Gregor Kosec, "A Numerical Study of Combining RBF Interpolation and Finite Differences to Approximate Differential Operators," *ArXiv*, 2025.
36. Z. Li, "Kolmogorov-Arnold Networks are Radial Basis Function Networks," 2024.
37. L. Abualigah, "Particle Swarm Optimization: Advances, Applications, and Experimental Insights," *Computers, Materials and Continua*, vol. 82, no. 2, pp. 1539–1592, Feb. 2025. <https://doi.org/10.32604/CMC.2025.060765>.
38. X. Z. Huangfu, "A hybrid global maximum power point tracking control method based on particle swarm optimization (PSO) and perturbation and observation (P&O)," *Electric Power Systems Research*, vol. 248.
39. L. J. V. Miranda, "PySwarms: a research toolkit for Particle Swarm Optimization in Python," *J Open Source Softw*, 2018.
40. Suat Bagci, "CO₂ Injection Well Design Considerations: Completion Design to Injectivity Analysis," in *SPE Annual Technical Conference and Exhibition*, 2024.
41. S. P. Fotias, I. Ismail, E. Tartaras, A. Stefatos, and V. Gaganis, "Optimizing CO₂ Injection Rates in CCS: a Bayesian Approach for Realistic Business Efficiency," vol. 2024, no. 1, pp. 1–5, 2024. <https://doi.org/10.3997/2214-4609.202421204>.
42. F. F. Edward A. Beaumont, "Exploring for Oil and Gas Traps."
43. Geology In, "Geothermal Gradient."
44. B. Chen, D. R. Harp, Y. Lin, E. H. Keating, and R. J. Pawar, "Geologic CO₂ sequestration monitoring design: A machine learning and uncertainty quantification based approach," *Appl Energy*, vol. 225, pp. 332–345, Sep. 2018. <https://doi.org/10.1016/J.APENERGY.2018.05.044>.
45. E. Kanin *et al.*, "Geomechanical risk assessment for CO₂ storage in deep saline aquifers," *Journal of Rock Mechanics and Geotechnical Engineering*, 2024. <https://doi.org/10.1016/j.jrmge.2024.04.016>.

46. I. Ismail, V. Gaganis, E. Tartaras, and A. Stefatos, "Spatiotemporal Hierarchical Optimization of CO₂ Injection Schedules: An Industry-Ready Framework," in *Fifth EAGE Eastern Mediterranean Workshop*, Cairo, Egypt: European Association of Geoscientists & Engineers, 2025.
47. M. Amvrazis, I. Ismail, D. Ktenas, V. Gaganis, E. Tartaras, and A. Stefatos, "Prinos CO₂ Storage Permit Application: An Integrated Technical and Regulatory Workflow," in *Fifth EAGE Eastern Mediterranean Workshop*, Cairo: European Association of Geoscientists & Engineers, 2025.
48. A. K. Furre, O. Eiken, H. Alnes, J. N. Vevatne, and A. F. Kiær, "20 Years of Monitoring CO₂-injection at Sleipner," *Energy Procedia*, vol. 114, pp. 3916–3926, Jul. 2017. <https://doi.org/10.1016/J.EGYPRO.2017.03.1523>.
49. S. P. Fotias and V. Gaganis, "Sensitivity of CO₂ Flow in Production/Injection Wells in CPG (CO₂ Plume Geothermal) Systems.," *Materials Proceedings*, 2025.
50. P.-M. Trimi *et al.*, "A Review of Caprock Integrity in Underground Hydrogen Storage Sites: Implication of Wettability, Interfacial Tension, and Diffusion.," *Hydrogen*, 2025.
51. R. Mokhtari, K. Ghahramani, S. Khojamli, D. Mihrin, and K. L. Feilberg, "Experimental Investigation of the Effect of Major Impurities in the CO₂ Stream on Carbon Storage in Chalk Reservoirs.," in *SPE Europe Energy Conference and Exhibition*, Turin, Italy, 2024.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.