

A NOVEL OPTIMIZATION FRAMEWORK BASED ON SURROGATE MODELING FOR UNDERGROUND HYDROGEN STORAGE IN DEPLETED NATURAL GAS RESERVOIRS

Zhilei Han , Zeeshan Tariq , Bicheng Yan 

*Physical Science and Engineering (PSE) Division, King Abdullah University of Science and Technology (KAUST),
Thuwal, Saudi Arabia*

Correspondence to:

Bicheng Yan,
bicheng.yan@kaust.edu.sa

How to Cite:

Han, Z., Tariq, Z., & Yan, B. (2025). A novel robust optimization framework based on surrogate modeling for underground hydrogen storage in depleted natural gas reservoirs. *InterPore Journal*, 2(3), IPJ250825-6.
<https://doi.org/10.69631/ipj.v2i3nr69>

RECEIVED: 7 Dec. 2024

ACCEPTED: 18 Jun. 2025

PUBLISHED: 25 Aug. 2025

ABSTRACT

Underground hydrogen storage (UHS) plays a vital role in global net-zero energy systems, enabling the storage of excess renewable energy for future use. However, physical reservoir model-based optimization for UHS system design and operation is computationally expensive due to complex geological properties and well-operational controls. This study developed a novel, efficient framework for UHS stochastic optimization to address this challenge, integrating advanced compositional reservoir simulation, accurate surrogate modeling, and stochastic optimization techniques. First, a base reservoir simulation model was developed to capture compositional fluid flow, hydrogen methanation reactions, gravity segregation, hysteresis, and capillary effects. To rapidly evaluate various well controls and reservoir configurations, convolutional neural network (CNN)-bi-directional long short-term memory (BiLSTM)-Attention models were trained as surrogate models using a comprehensive dataset generated from reservoir simulations. The CNN transforms three-dimensional (3D) geological fields into one-dimensional (1D) vectors, effectively capturing spatial features. The BiLSTM network learns the temporal evolution of the input features over time by processing them in both forward and backward directions. Subsequently, the attention mechanism enhances prediction accuracy by identifying and emphasizing the most significant features at critical time steps. The well-trained surrogate models were seamlessly integrated into the stochastic optimization framework based on the genetic algorithm, aiming to maximize the net present value (NPV) from UHS projects. The results demonstrate that the surrogate model exhibits satisfactory performance in the context of prediction accuracy, computational efficiency, and scalability. Notably, the newly developed framework based on surrogate models achieves an

approximate 4878-fold speedup compared to an approach relying solely on reservoir simulation, while maintaining comparable accuracy. Overall, the proposed framework offers a promising solution for UHS optimization, providing valuable insights for the design and management of sustainable energy infrastructure.

KEYWORDS

Underground hydrogen storage, Robust optimization, Surrogate modeling, Deep learning



@2025 The Authors

This is an open access article published by InterPore under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0) (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. INTRODUCTION

The vision of a low-carbon economy has driven the rapid development of renewable energy technologies, primarily dominated by wind and solar energy. However, their short-term and rapid fluctuations challenge the power grid's ability to respond effectively, thereby accelerating the development of various energy storage technologies, as shown in [Figure 1](#). Considering its long discharge time and high storage capacity, Power-to-Gas (Hydrogen or Methane) becomes a viable solution for the long-term, large-scale storage of electrical energy. In recent years, the growing global demand for renewable energy has sparked significant interest in hydrogen as a clean and sustainable energy carrier ([15](#), [39](#), [48](#)).

Geological formations, such as depleted hydrocarbon reservoirs (DHR), salt caverns, and saline aquifers, are promising options for seasonal and large-scale underground hydrogen storage ([15](#)). Among these, depleted natural gas reservoirs (DGRs) present a viable solution for managing renewable power intermittency and over-generation, ensuring a stable and scalable energy supply ([23](#)). The formation of DGRs has illustrated their geological trapping ability on natural gas, sealed by impermeable overburden mudrock or salt layers

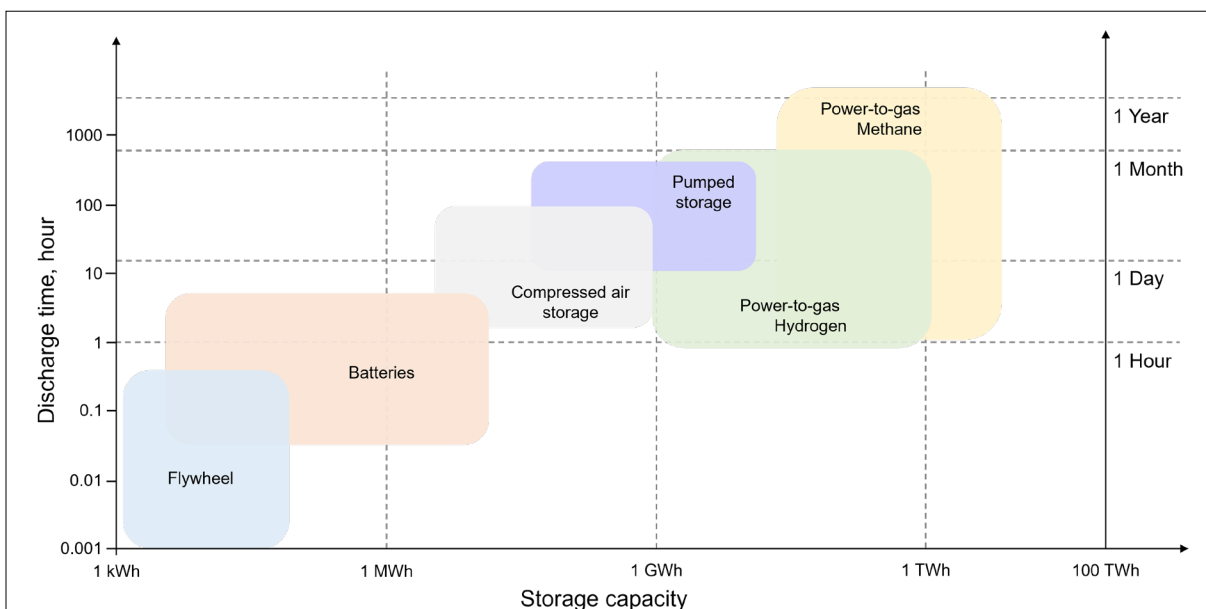


Figure 1: Various energy storage technologies (see [44](#)).

and surrounded by bottom and edge water layers. Moreover, extensive operational expertise and long-term production experience make these reservoirs well-characterized (42). This includes insights into reservoir connectivity, caprock integrity, and related factors. Such understanding allows for accurate estimation of reservoir storage capacity, reduces uncertainty, and mitigates leakage risks, thereby leading to significant cost savings. As a result, DGRs offer lower capital and operating expenditures by reusing existing equipment and wells after evaluation, ensuring the economic feasibility of underground hydrogen storage (UHS) operations (3). While extensive experience exists in underground natural gas storage (UGS) (29), the significant differences in the properties of hydrogen and methane present unique challenges when attempting to directly apply UGS experience to UHS (46). Underground hydrogen storage operations in DGRs are still in the research and exploration period. To gain deeper insights into hydrogen recovery in UHS projects, reservoir simulation and optimization are essential to optimize decision parameters during operations. In recent years, researchers have extensively explored UHS simulations using both commercial and open-source simulators, such as CMG GEM (55), ECLIPS (22), *DuMux*^x (22), COMSOL (47), and TOUGH2 (36). Key physics of focus include black-oil type simulations (24), compositional simulations with cushion gas types (16, 27), cycling schedules (6, 38), caprock integrity (26, 37), microbial activity (14, 53), geochemical reactions (18, 59), and rock-fluid behaviors (2, 19, 58). Although these studies have provided valuable insights into various factors influencing UHS operations, they are often limited to specific scenarios and operational conditions. To address this limitation, it is essential to develop a more comprehensive approach to UHS optimization, exploring a broader and more generalized search space to identify optimal well controls. Such an approach facilitates a deeper understanding of operational performance across a wider range of conditions.

However, conducting physical reservoir model-based optimizations often requires extensive computational resources, making the process time-consuming and resource-intensive. This limits their practicality for real-time decision-making and UHS optimization. To overcome this challenge, researchers started to develop surrogate models as efficient alternatives for various optimization problems (4, 45). These models significantly accelerate the optimization process, mitigating the computational time constraints associated with traditional methods (57). Despite this potential, the integration of deep learning (DL) based surrogate models for UHS optimization remains underexplored, with only a few preliminary efforts reported in the literature. For instance, Kanaani et al. (28) introduced a multi-objective co-optimization framework for UHS and carbon dioxide storage using machine learning (ML) algorithms. They employed a multi-layer neural network (MLNN) to predict multiple outputs simultaneously. Similarly, Sun et al. (51) proposed a framework for UHS in saline aquifers, utilizing two Gaussian support vector machine (SVM) models as proxies to address computational challenges in the optimization process. Further advancements include reduced-order models (ROMs) based on deep neural networks (DNNs) developed to predict UHS performance in DGRs (40), and to select optimal subsurface hydrogen storage sites in saline aquifers (7). Despite the growing interest in surrogate model-based UHS optimization, significant research gaps remain. For instance, these optimization studies ignore some critical physical processes in UHS, such as the impact of biochemically mediated hydrogen loss (33, 50), which can significantly influence storage performance and recovery efficiency. Additionally, stochastic optimization approaches, which quantify the effects of uncertainties in geological parameters on project design and performance, have not yet been reported in the context of UHS operations. Addressing these gaps is essential for developing more reliable and comprehensive stochastic optimization frameworks for UHS applications.

This study proposes a novel stochastic optimization framework for UHS in DGRs, leveraging accurate and efficient surrogate models that capture the dynamic behavior of UHS systems under various operating conditions. We develop an advanced reservoir simulation model to account for compositional fluid flow, gas component diffusion, bio-reaction (i.e., methanation), gravity segregation, relative permeability (drainage and hysteresis), and capillary effects. Based on the comprehensive dataset generated from these simulations,

we build the surrogate models for time-series sequence-to-sequence prediction problems based on a convolutional neural network (CNN), a bi-directional long short-term memory (BiLSTM) network, and an attention mechanism. In particular, BiLSTM is designed to capture dependencies in time-sequential data by processing it in both forward and backward directions. This capability makes BiLSTMs particularly effective for tasks where understanding both past (preceding) and future (succeeding) contexts is crucial. The CNN-BiLSTM-Attention model is rigorously validated for accuracy, efficiency, and reliability. Subsequently, the validated surrogate model is seamlessly integrated into a genetic algorithm (GA) based stochastic optimization workflow, aiming at maximizing net present value (NPV).

The rest of the paper is structured as follows: [Section 2](#) details the governing equations for two-phase compositional fluid flow, the development of the physical reservoir simulation model, the construction and performance evaluation of the surrogate models, and the stochastic optimization workflow. [Section 3](#) provides an in-depth analysis of the reservoir simulation results and the stochastic optimization outcomes. [Section 4](#) summarizes the key findings of this study.

2. METHODOLOGY

This section describes the key steps involved in developing the stochastic optimization framework based on the proposed surrogate model for the UHS system. First, the governing equations underlying the physical processes are introduced, followed by a detailed description of the numerical simulation model and the generation of a high-fidelity dataset. Next, the architecture and components of the proposed CNN-BiLSTM-Attention model are presented. Finally, the formulation of the objective function and the optimization algorithm are discussed, enabling the quantitative evaluation of NPV.

2.1. Governing equations

To capture the dynamic behavior of hydrogen injection into depleted natural gas reservoirs, as well as the influence of cushion gas type, we consider a two-phase (gaseous and aqueous) and multi-component (H_2, N_2, CH_4, CO_2) fluid flow in subsurface porous media. By incorporating the advection and molecular diffusion mechanisms, the governing equation of mass conservation for each component is given as ([Eq. 1](#)):

$$\frac{\partial}{\partial t} \left(\sum_{\alpha} \phi S_{\alpha} \rho_{\alpha} X_{\alpha}^i \right) + \nabla \cdot \sum_{\alpha} \left(\rho_{\alpha} u_{\alpha} X_{\alpha}^i - \phi S_{\alpha} D_{\alpha}^i \nabla (\rho_{\alpha} X_{\alpha}^i) \right) = q^i \quad (1)$$

where α and i denote phase and component, respectively; ϕ and S_{α} are porosity and saturation, respectively; ρ_{α} is mole density; X_{α}^i is mole fraction of component i in phase α ; u_{α} is Darcy's velocity; D_{α}^i is diffusion coefficient; q^i is source or sink term. Darcy's velocity is expressed as ([Eq. 2](#)):

$$u_{\alpha} = - \frac{\mathbf{K} k_{r\alpha}}{\mu_{\alpha}} (\nabla p_{\alpha} - \rho_{\alpha} g \nabla z) \quad (2)$$

where $\mathbf{K}$ is the absolute permeability; $k_{r\alpha}$ is the relative permeability of phase α ; u_{α} is phase viscosity; p_{α} is phase pressure; g is gravitational acceleration; and z is depth.

Moreover, the capillary effect is integrated into the pressure term through the inclusion of capillary pressure, as expressed by ([Eq. 3](#)):

$$p_c = p_{nw} - p_w \quad (3)$$

where p_w , p_{nw} , and p_c are the pressures of the wetting phase (water), non-wetting phase (gas), and the capillary pressure, respectively.

The flash splitting calculation is utilized to calculate the mole fraction and density with the assumption of local thermodynamic equilibrium. The equation governing phase equilibrium is the equivalence of the fugacities of the component in both the gaseous and aqueous phases (Eq. 4):

$$f_{ig} = f_{iw} \quad (4)$$

where f_{ig} and f_{iw} are fugacities of component i in the gaseous and aqueous phase, respectively.

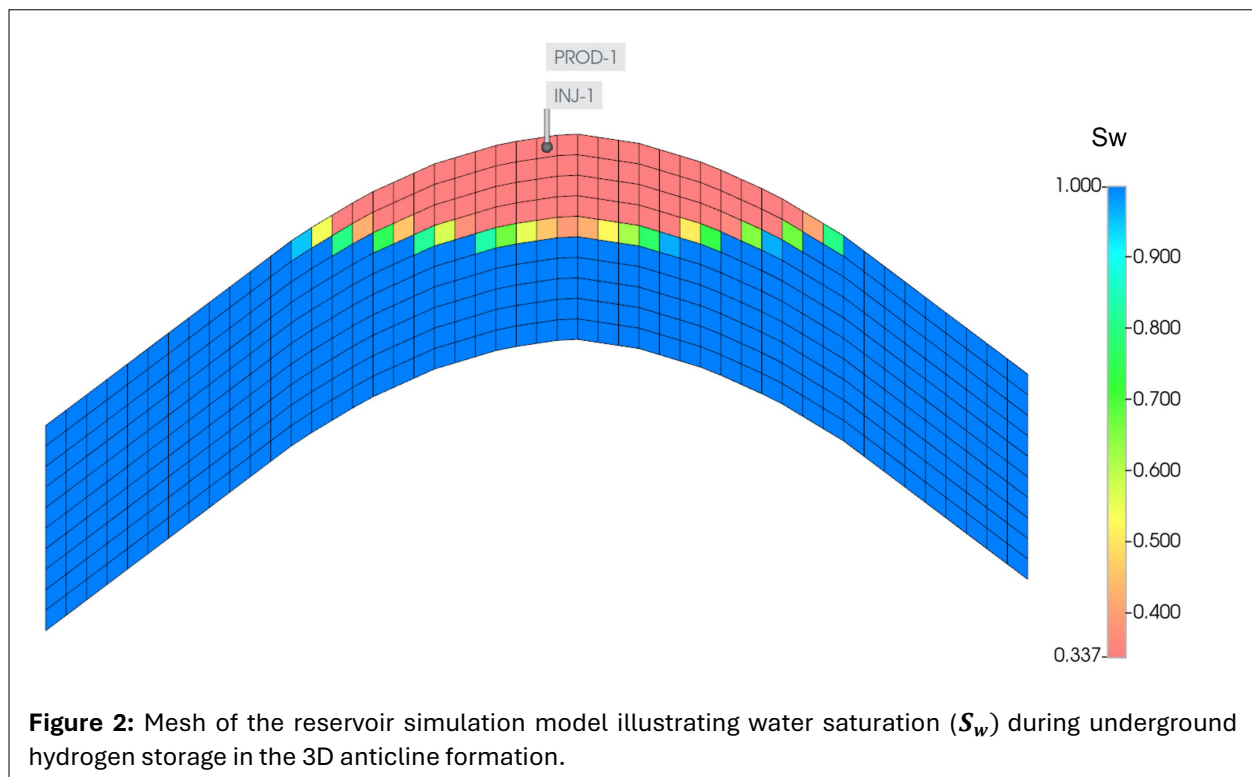
In particular, this study incorporates the Soave-Redlich-Kwong (SRK) Equation of State (EOS) (49) to effectively model the phase behavior of H_2 and other gas components (23), while Henry's law is applied to estimate the solubility of these gas components in the aqueous phase, where f_{iw} is given as (32) (Eq. 5):

$$f_{iw} = x_{iw} * H_i \quad (5)$$

where x_{iw} is mole fraction of component i in aqueous phase; H_i is Henry's law constant.

2.2. Reservoir simulation model

A comprehensive reservoir model is developed to facilitate multi-phase compositional simulations for UHS operations in DGR by using the commercial reservoir simulator GEM^a from the Computer Modeling Group (CMG) (8). This model utilizes the finite difference method for spatial discretization, along with an adaptive implicit scheme for simulating multi-phase multi-component flow with phase and bio-geochemical equilibrium using a fully coupled approach. By discretizing and solving Equation 1 through the Newton-Raphson method, the primary variables, such as pressure and number of moles for each component, can be determined for each cell at each time step. Figure 2 depicts a three-dimensional (3D) anticline structure employed as the representative mesh model for this research, characterized by dimensions of 1500 m in length, 500 m in width, and 100 m in vertical thickness. The model is discretized using the corner-point grid consisting of 12,500 cells ($50 \times 25 \times 10$) and spans a depth range from 800 to 1040 m. The reservoir is



^a <https://www.cmgl.ca/solutions/software/gem/>

assumed to be heterogeneous and anisotropic, with its side boundaries connecting to an infinite edge water aquifer. In particular, reservoir anisotropy is quantitatively characterized by permeability, where the permeabilities in the two horizontal directions are equal, and the ratio of vertical to horizontal permeability is set to 0.1. A central vertical well serves as both an injection and production well. To prevent water coning, only the first top layer is perforated.

Regarding the rock-fluid interactions, the drainage and imbibition relative permeability curves of H_2 - H_2O (Fig. 3a) (9), as well as capillary pressure curve (Fig. 3b) (20), are accounted for. The maximum residual H_2 saturation is set at 0.337.

In this study, the hysteresis is incorporated in the relative permeability-saturation curves, but not in the capillary pressure-saturation relationship. This modeling choice is based on the available experimental data for H_2 - H_2O systems, which exhibit hysteresis in relative permeability but not in capillary pressure. Similar modeling methods have been adopted in previous UHS studies (e.g. 9). In particular, the hysteresis behavior in this work is modeled using Land's equation (30), which has been integrated in GEM. The detailed formulations are presented in Equation 6 to Equation 9.

The gas relative permeability along the drainage-to-imbibition scanning curve for a given gas saturation, S_g , is expressed as (Eq. 6):

$$K_{rg}(S_g) = K_{rg}^{drain}(S_{gf}) \quad (6)$$

where S_{gf} is the free gas saturation, calculated as (Eq. 7):

$$S_{gf} = \frac{1}{2}(S_g - S_{grh}) + \frac{1}{2}\sqrt{(S_g - S_{grh})^2 + \frac{4}{C}(S_g - S_{grh})} \quad (7)$$

where C is Land's parameter, S_{grh} is the residual gas saturation of imbibition process. These parameters are calculated as follows (Eq. 8, Eq. 9):

$$C = \frac{1}{S_{grmax}} - \frac{1}{S_{gmax}} \quad (8)$$

$$S_{grh} = \frac{S_{gh}}{1 + C * S_{gh}} \quad (9)$$

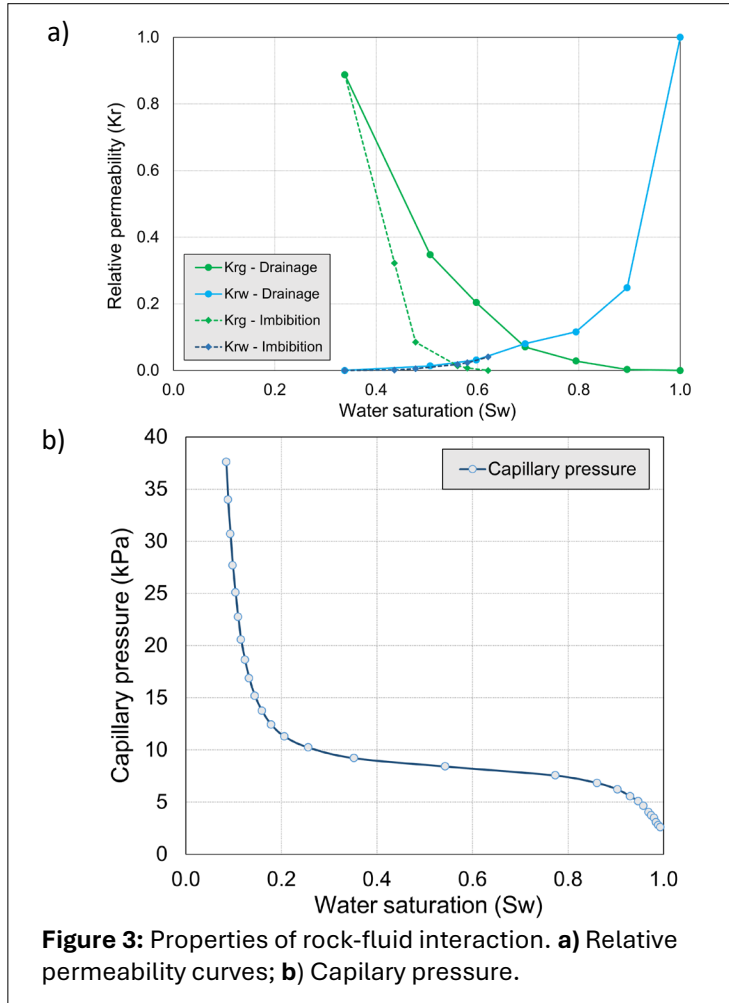


Figure 3: Properties of rock-fluid interaction. **a)** Relative permeability curves; **b)** Capillary pressure.

where S_{gh} is the maximum gas saturation in drainage process; S_{grmax} is the maximum residual gas saturation; $S_{grmax} = 1 - S_{wcon} + S_{oirg}$; S_{wcon} and S_{oirg} are connate water saturation and irreducible oil saturation.

Table 1: Other major properties of the reservoir simulation model.

Parameter	Value	Unit
Pressure at $z = 900$ m	8920	kPa
Temperature at $z = 900$ m	55	°C
Gas-water contact (GWC)	850	m
Initial global mole fraction of CH_4	0.690	Dimensionless
Initial global mole fraction of C_2H_6	0.300	Dimensionless
Initial global mole fraction of CO_2	0.005	Dimensionless
Initial global mole fraction of N_2	0.005	Dimensionless
Initial global mole fraction of H_2	0.000	Dimensionless

During reservoir shut-in periods, it is assumed that water recharge from the infinite edge aquifer brings the reservoir to equilibrium prior to UHS operations. As illustrated in **Figure 2**, the water saturation (S_w) across the vertical layers ranges from 0.337 to 1.0, governed primarily by the equilibrium between gravitational and capillary forces. A transition zone is assumed near the gas-water

contact (GWC). Under this equilibrium assumption, the initial reservoir pressure is modeled as hydrostatically balanced, while the initial reservoir temperature is determined using the reservoir depth and the geothermal gradient. Other major properties of the reservoir model used in this study can be found in **Table 1**.

The initial reservoir gas components consist of CH_4 , C_2H_6 , CO_2 and N_2 , with C_2H_6 being a trace component that exhibits relatively low solubility. The initial global mole fraction of the reservoir gas components are detailed in **Table 1**. The ions present in the aqueous phase include Na^+ , H^+ , OH^- , HCO_3^- and CO_3^{2-} , with initial concentrations (in molarity) of 10^{-1} , 10^{-7} , 10^{-7} , 10^{-7} , and 10^{-16} , respectively. Additionally, molecular diffusion of H_2 , N_2 , CH_4 and CO_2 is considered in this study, whereas the diffusion of the chemical ions is neglected. The geochemical reactions involved in this study are (**Eq. 10**, **Eq. 11**, **Eq. 12**):



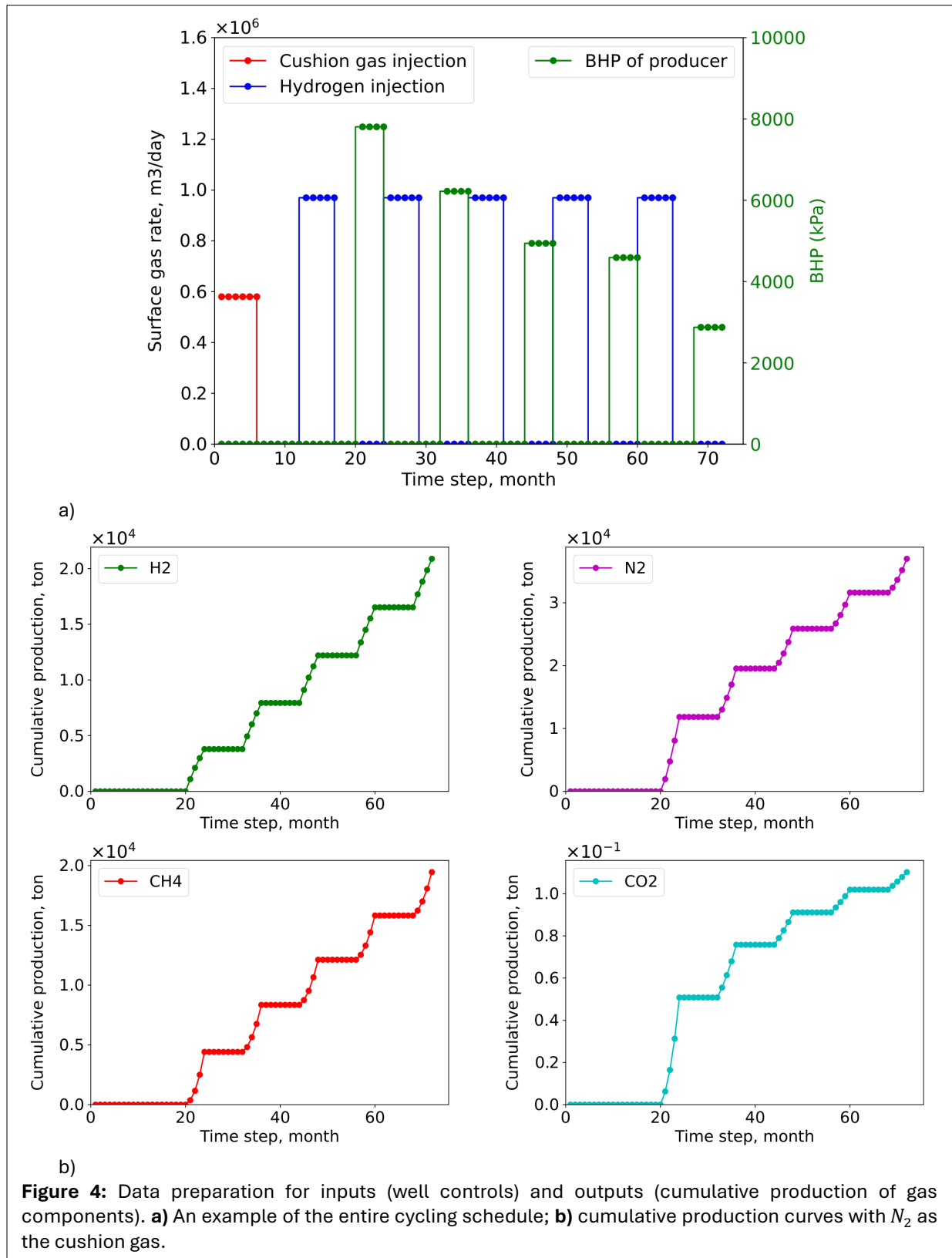
To address microbial-induced H_2 loss (**13**), an Arrhenius-type reaction is incorporated to model methanation (**36**) resulting from micro-bio reactions (**Eq. 13**):



2.3. Surrogate modeling workflow

The stochastic optimization of the UHS system through numerical modeling is often computationally intensive due to complex geological properties and dynamic well controls. To address this challenge, a deep learning model can be developed as an efficient surrogate to significantly accelerate the optimization process. These models are designed to map the high-dimensional decision parameters to the desired output space, leveraging the universal approximation capability of neural networks. This study introduces a novel surrogate model to predict the spatial evolution of the cumulative production of various gas components during UHS operations. The workflow for constructing the surrogate model is as follows:

- Generating decision parameters using the Latin Hypercube Sampling (LHS) method;
- Performing the numerical simulations in parallel and building a high-fidelity simulation dataset;
- Training and testing the surrogate models, including hyper-parameter optimization;
- Evaluating performance metrics to assess model accuracy and robustness.



2.3.1. Problem setting

To recover the hydrogen from the depleted natural gas reservoir and account for the effect of cushion gas type on hydrogen recovery, the cycling schedule in this study starts with six months of cushion gas injection and six months of idle period, followed by five cycles of H_2 injection, idle and production spanning five consecutive years, as shown in [Figure 4a](#). Each month represents a single time step, resulting in a total of 72 time steps over six years. It is assumed that the injection rate of H_2 remains constant to simulate a steady injection of H_2 . To ensure a higher recovery efficiency of H_2 , decreasing BHPs are assumed throughout the entire cycle. In particular, a discretized, monotonically decreasing exponential decline curve ([Eq. 14](#)) is introduced to generate five decreasing BHP values corresponding to the five operational cycles. To meet the BHP constraint, the last value generated from the decline curve is set to be greater than the lower bound of BHP, as shown in [Equation 10](#).

$$y = coeff \cdot e^{(-k \cdot x)} \quad (14)$$

This approach reduces the five BHP values to two parameters: *coeff* and *k*, where *coeff* follows the upper bound of BHP, and *k* ranges between 0 and 1.

Consequently, there are a total of six well controls: the injection rate of cushion gas, the injection rate of H_2 , *coeff*, *k*, the duration of well opening, and the duration of well shut-in.

The optimization in this study aims to maximize the expected NPV, where geological uncertainty is introduced through variations in permeability across multiple realizations. A formal mathematical description of the optimization problem is presented in [Section 2.4.1](#).

2.3.2. Data preparation

Latin Hypercube Sampling (LHS) ([17](#)) is used to generate a comprehensive dataset of decision variables and geological parameters, which follows a uniform distribution and covers a wide range of realistic scenarios for practical UHS projects. The decision parameters include the injection rate of cushion gas and H_2 , bottom hole pressure (BHP) of the production well, as well as the duration of well open-up and shut-in periods. Since porosity can be estimated from permeability, the geological parameter subject to uncertainty is permeability. The ranges for these parameters are shown in [Table 2](#), with all parameters uniformly distributed within these specified ranges.

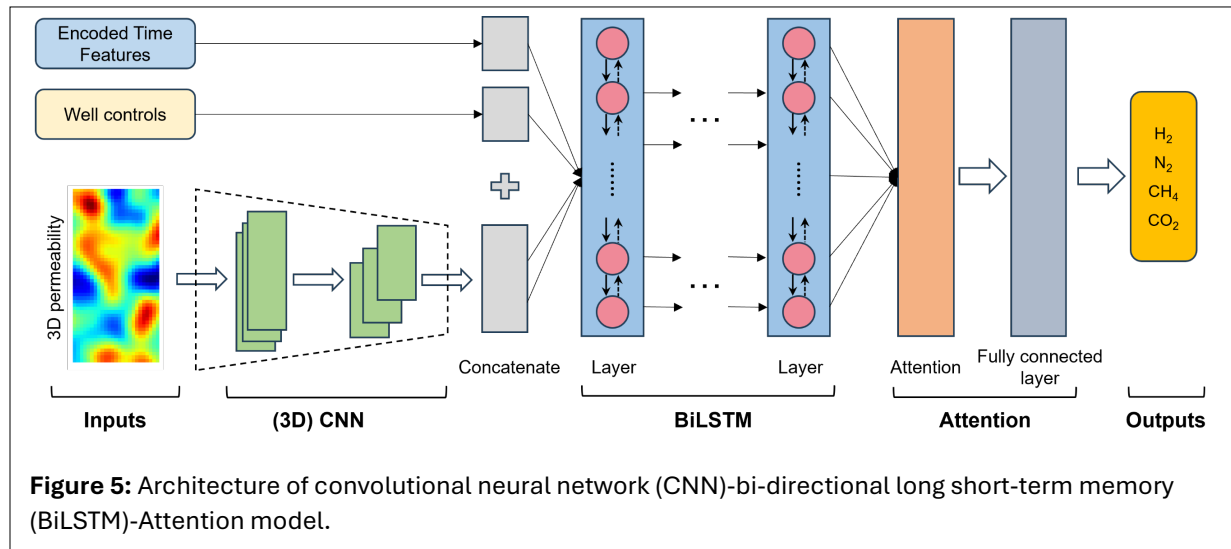
Specifically, the 3D permeability field is constructed utilizing the Dykstra-Parsons method ([25](#)), with the Dykstra-Parsons coefficient of 0.6. The porosity field is then derived from the generated permeability using the empirical equation outlined in [Equation 15](#) ([60](#)), where ϕ is porosity:

$$\phi = 0.05 * \log_{10}(K) - 0.03 \quad (15)$$

After sampling and parameterizing these inputs shown in [Figure 4a](#), we generate a comprehensive database of approximately 1000 simulation cases by executing CMG in parallel. The outcomes of these simulations include the cumulative production of four components: H_2 , N_2 , CH_4 , and CO_2 , as illustrated in [Figure 4b](#). These production curves are used as outputs for training the surrogate models.

Table 2: Ranges of geological parameters and well controls.

Parameter	Lower bound	Upper bound	Unit
Permeability	100	5000	mD
Injection rate of cushion gas	1.0×10^3	1.0×10^6	m^3/day
Injection rate of H_2	1.0×10^3	1.0×10^6	m^3/day
BHP of the production well	2.0×10^3	8.0×10^3	kPa
Duration of injection	1	6	month
Duration of production	1	6	month



2.3.3. Surrogate model development

Based on the Pytorch library (43), we develop four surrogate models with identical architectures to separately predict the cumulative production of H_2 , N_2 , CH_4 , and CO_2 . As shown in Figure 5, the CNN-BiLSTM-Attention model is selected as the deep learning architecture of the surrogate model in this research, which combines the strengths of both CNN (31) and LSTM (21) to capture spatial-temporal dependencies and patterns in UHS systems. As an encoder, CNN (a dashed-line box with light green blocks in Fig. 5) transforms a 3D heterogeneous permeability field into one-dimensional (1D) vectors. These 1D vectors, later along with other 1D well controls, serve as inputs for the BiLSTM-Attention network to predict the cumulative production of each gas component. A BiLSTM network (light blue blocks in Fig. 5) is a type of recurrent neural network (RNN) designed to capture the order and time dependencies of the input data by processing it in both forward and backward directions. The attention mechanism (52) (a light orange block in Fig. 5) is integrated to emphasize key time steps, thereby improving the prediction accuracy. Finally, the output of the attention mechanism is then passed through a fully connected layer (dark gray block in Fig. 5) with a Rectified Linear Unit (ReLU) activation function to produce the final feature. Once trained, the model serves as an efficient surrogate for computationally expensive reservoir simulations, enabling rapid evaluation of various operational strategies and reservoir configurations. The architecture of the surrogate model is illustrated in Figure 5.

Figure 6 depicts the dimensionality reduction process for geological permeability using 3D CNN. As described in Section 2.2, the initial 3D permeability field has dimensions of $50 \times 25 \times 10$ corresponding to the depth, width, and thickness of the reservoir model. The dimensionality reduction is achieved through two sequential layers, each comprising a 3D CNN followed by a MAXPooling operation. The resulting tensor is then

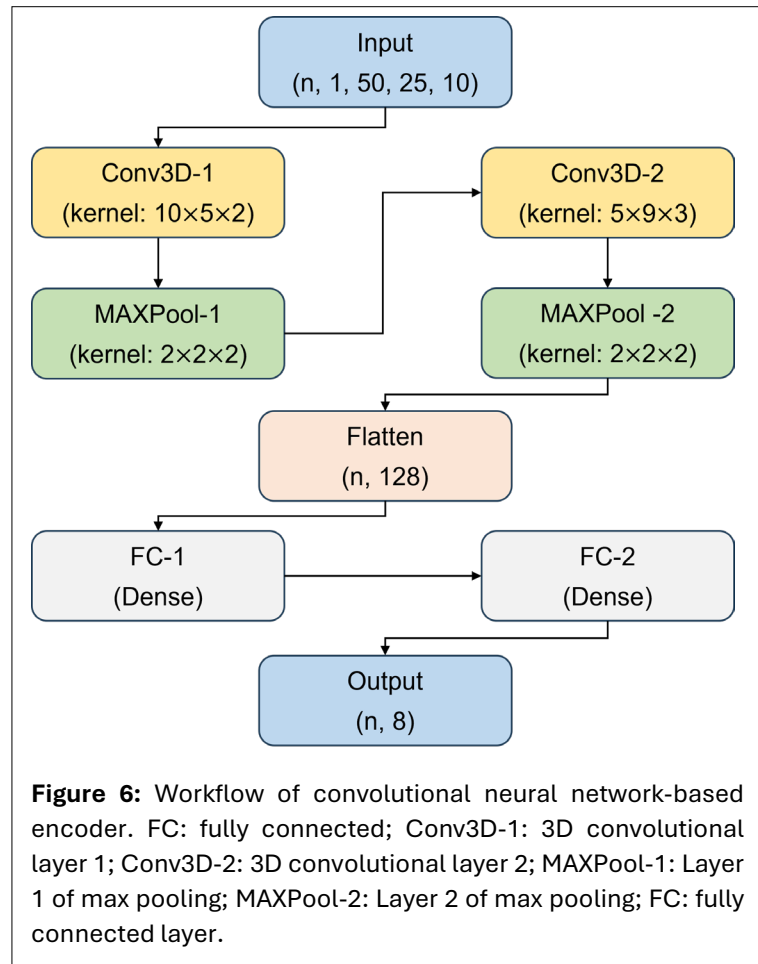
Table 3: Summary of the 3D convolutional neural network block.

Layer	Kernel size	Output size
Input	/	(n, 1, 50, 25, 10)
Conv3D-1, 32 filters of size, stride 1	$10 \times 5 \times 2$	(n, 32, 41, 21, 9)
MAXPool-1	$2 \times 2 \times 2$	(n, 32, 20, 10, 4)
Conv3D-2, 16 filters of size, stride 1	$5 \times 9 \times 3$	(n, 16, 16, 2, 2)
MAXPool-2	$2 \times 2 \times 2$	(n, 16, 8, 1, 1)
Flatten	/	(n, 128)
FC-1	/	(n, 64)
FC-2	/	(n, 8)

Conv3D-1: 3D convolutional layer 1; Conv3D-2: 3D convolutional layer 2; MAXPool-1: Layer 1 of max pooling; MAXPool-2: Layer 2 of max pooling; FC: fully connected layer.

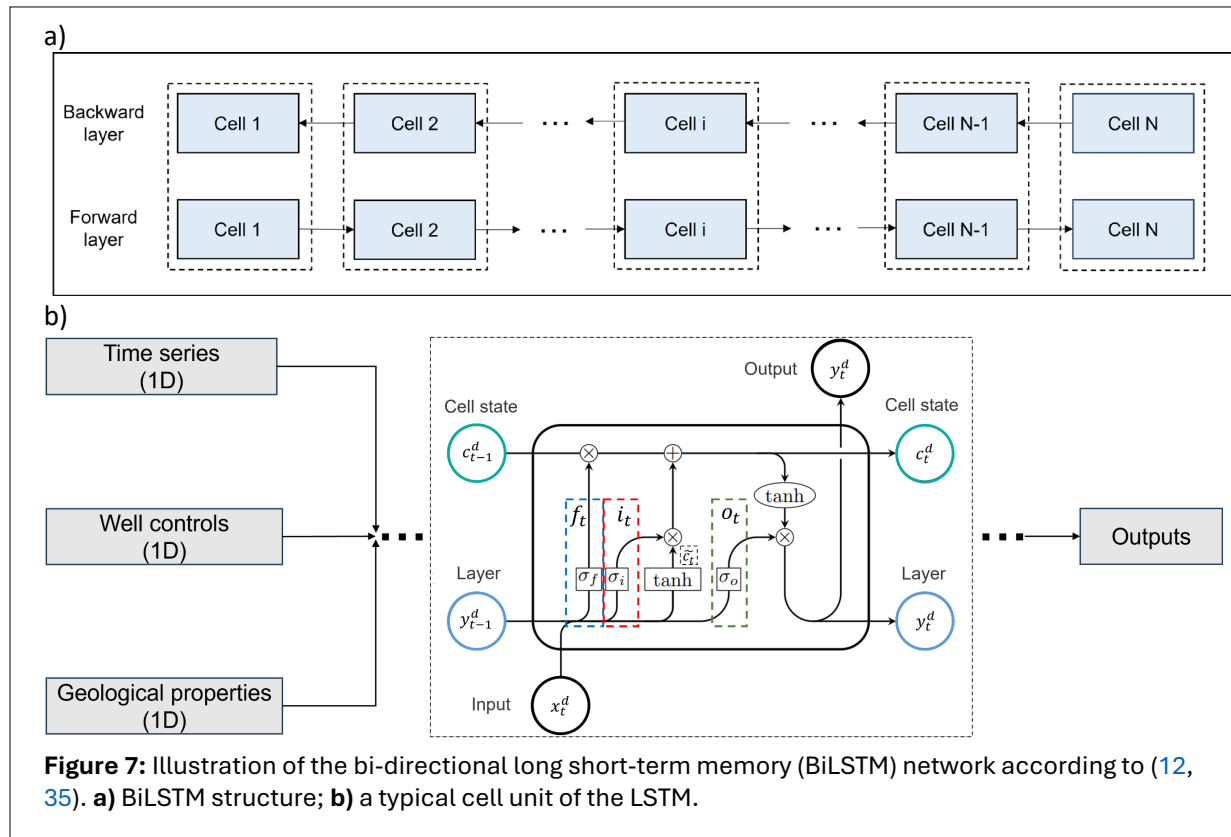
passed through a flattening layer and two fully connected (FC) layers, reducing its dimensionality from $[n, 1, 50, 25, 10]$ to $[n, 8]$. Here, n denotes the batch size, and 1 represents the channel size. The architecture details of the 3D CNN are presented in [Table 3](#).

Although sequence-to-sequence modeling with recurrence theoretically does not require explicit encoding of the time series, incorporating positional embeddings (11) can enhance model performance. Specifically, embedding the absolute positions of input features along the time axis equips the model with a good understanding of the order. This is particularly beneficial for capturing the strong periodic patterns (e.g., seasonal behavior) inherent in the UHS operations. Using a simple sequential index (e.g., $t = 1, 2, \dots, 72$) to represent the position of a time step, however, poses a risk that the model may misinterpret them as linear features, thereby failing to capture the cyclic or seasonal nature of time. To address this, the time series are encoded using learned embeddings in this study, which can more effectively capture their periodicity and complex patterns.



The well controls are preprocessed according to the cycling schedule, as illustrated in [Figure 4a](#). To enhance the weights of well controls within the overall input parameters during the training of the surrogate model, these scalar values are converted into 1D vectors by repeating their values at each time step, matching the dimension of the 1D latent representation of permeability after the 3D CNN block. The processed 1D well controls are then concatenated with the 1D permeability latent and encoded time series along the feature dimension, forming the inputs for the subsequent BiLSTM layers.

As illustrated in [Figure 7a](#), the BiLSTM model consists of forward and backward layers. The forward LSTM processes the input sequence from the beginning (i.e., cell 1 or $t=1$) to the end (i.e., cell N or $t=T$), capturing how the UHS system evolves over time based on preceding conditions. Conversely, the backward LSTM processes the input sequence from the end (i.e., cell N or $t=T$) to the beginning (i.e., cell 1 or $t=1$), incorporating information from later states, such as cumulative production, to improve prediction accuracy. At each timestep, the outputs from both directions are concatenated to provide a more comprehensive prediction. By processing the time sequence in both directions, the BiLSTM effectively captures the dependencies in the data that span across timesteps, whether they occur earlier or later in the sequence. During prediction, the model proceeds in a step-by-step (autoregressive) manner, where only past information is available. Thus, the BiLSTM is applied to observed sequences during training to extract informative representations, while the final predictive model uses these learned features to generate forecasts in a strictly unidirectional manner. In other words, the backward LSTM becomes naturally inactive



during prediction due to the absence of future production data as input. This dual perspective enhances the ability of the surrogate model to predict dynamic system behaviors. **Figure 7b** depicts the structure of a typical LSTM cell unit, including an input gate (i_t), forget gate (f_t), output gate (o_t), and cell state (c_t). The BiLSTM utilizes these gates to selectively memorize or forget past information while passing relevant details through hidden states. The mathematical equations of these gate mechanisms and the cell state are provided in **Equations 16 to 19**.

$$i_t = \sigma(w^i[y_{t-1}^d, x_t^d] + b_i) \quad (16)$$

$$f_t = \sigma(w^f[y_{t-1}^d, x_t^d] + b_f) \quad (17)$$

$$o_t = \sigma(w^o[y_{t-1}^d, x_t^d] + b_o) \quad (18)$$

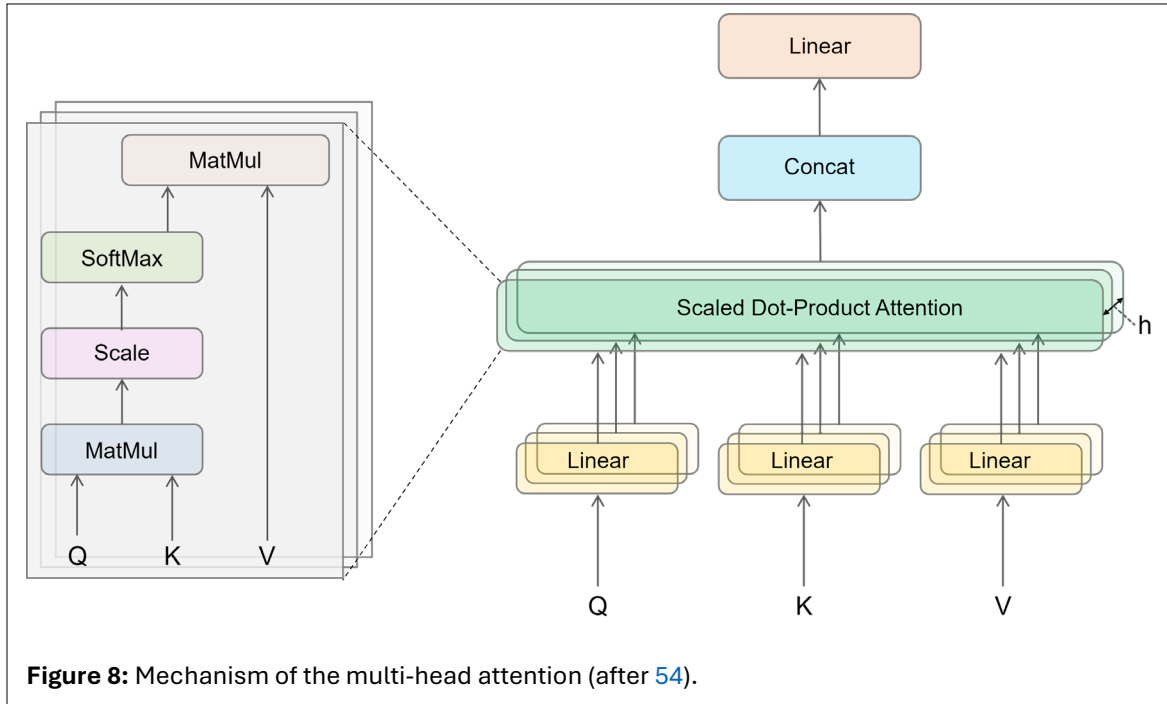
$$c_t^d = f_t * c_{t-1}^d + i_t * \tanh(w^c[y_{t-1}^d, x_t^d] + b_c) \quad (19)$$

The output of LSTM (y_t^d) can then be calculated by (**Eq. 20**):

$$y_t^d = o_t * \tanh(c_t^d) \quad (20)$$

where t and d represent time step and direction of LSTM, respectively; x_t represents the inputs; i_t , f_t , and o_t are the outputs of the input, forget, and output gates, respectively; y_{t-1}^d and c_{t-1}^d denote the layer (or hidden) state and cell state at timestep $t - 1$; y_t^d and c_t^d represent the layer (or hidden) state and cell state at time step t ; the weights of the input, forget, and output gates, as well as the hidden state are denoted by w_i , w_f , w_o , and w_c , respectively; b_i , b_f , b_o , and b_c are the corresponding biases. The activation functions are represented by σ and $\tanh$, while $*$ is the convolution operator.

A multi-head self-attention mechanism (10, 54) is then integrated after the BiLSTM layers to enhance the capability of the model to capture critical time steps and spatial relationships. It enables the model to



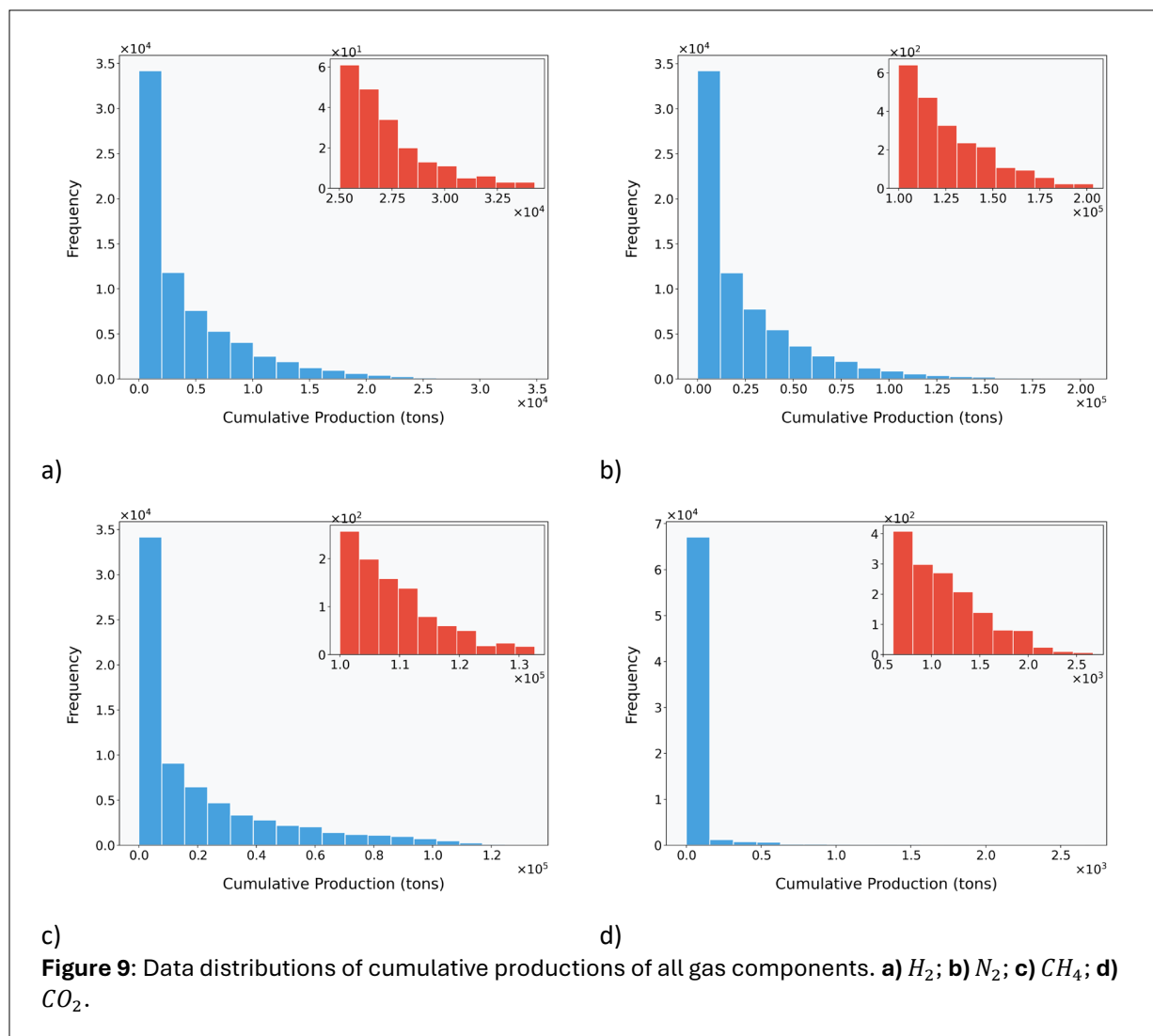
dynamically weigh different parts of the input sequence when making a prediction, rather than treating all past information equally.

Figure 8 illustrates the mechanism of multi-head attention. The output states from the BiLSTM layer are first linearly projected into multiple sets of Query (Q), Key (K), and Value (V) vectors, corresponding to each attention head. Each head independently computes attention scores by evaluating the similarity between queries and keys, allowing it to capture different types of temporal dependencies or spatial correlations. The resulting attention weights are then multiplied by the value vectors (see Eq. 21), highlighting the most relevant features in the input sequence. By using multiple attention heads, the model can jointly attend to information from different representation subspaces at different positions, which significantly enhances its ability to learn complex and multi-scale interactions. The outputs of all attention heads are then concatenated and passed through a linear transformation layer to form the final attended representation. These attention-enhanced features are subsequently used for prediction, allowing the model to selectively emphasize influential time steps (e.g., sharp changes in well controls or responses) and suppress less informative ones. This integration improves both the interpretability and accuracy of the production forecasts, particularly under highly dynamic well control scenarios.

$$Attention = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (21)$$

where Q denotes the query vector, representing the information we are seeking within the sequence; K stands for the key vector, which is used to compare against the query to assess the relevance of each part of the sequence; and V is the value vector, which carries the actual information to be passed on after the query and key are compared. The symbol T is the transpose operation, and $\sqrt{d_k}$ is the scaling factor, where d_k is the dimensionality of the key vectors. The dot product of the query and key produces the raw scores, and the *softmax* function converts these scores into probabilities.

To increase the stability of the model training process, we implement a strategy where the monthly production rate is first predicted, and the cumulative production is then calculated by accumulating the



predicted monthly values. This approach is particularly effective since the production can be zero during well shut-in periods. Given that the output data spans several orders of magnitude (as shown in Fig. 9), a log transformation is applied to all production curves of the four gas components. Subsequently, both the input and output parameters are normalized using the min-max method to streamline the training and validation process of the surrogate model. In particular, the comprehensive dataset with around 1000 cases is divided into training, validation, and testing subsets in an 8:1:1 ratio. The CNN-BiLSTM-Attention model is trained on an NVIDIA RTX A6000 using the AdamW optimizer (34), mean square error (MSE) loss, and ReLU activation function.

The Python-based, open-source OPTUNA library (1) is used to tune the hyperparameters, as outlined in Table 4. The optimized hidden size and number of attention heads for all models are 128 and 2, respectively. Specifically, the objective function minimized by OPTUNA is the MSE of the predicted production rate on the validation dataset during hyperparameter tuning.

Table 4: Optimized hyperparameters for the four surrogate models.					
Surrogate model	Batch size	Learning rate	No. of BiLSTM layers	Dropout size	Weight decay
CH_4	32	7.21×10^{-4}	3	0.1	7.32×10^{-4}
CO_2	8	9.83×10^{-5}	2	0.3	1.17×10^{-4}
H_2	8	4.30×10^{-4}	2	0.4	1.45×10^{-4}
N_2	8	2.90×10^{-4}	2	0.3	1.44×10^{-4}
BiLSTM: bi-directional long short-term memory.					

2.3.4. Performance metrics

Two metrics are used to quantify the performance of the CNN-BiLSTM-Attention surrogate model. Mean Square Error is a commonly used evaluation metric. The MSE loss is calculated by Equation 22:

$$MSE = \frac{1}{N} \sum_{i=1}^N y_i - \hat{y}_i \quad (22)$$

where N is the number of samples; y_i and $\hat{y}_i$ are the ground truth and the surrogate model predictions, respectively.

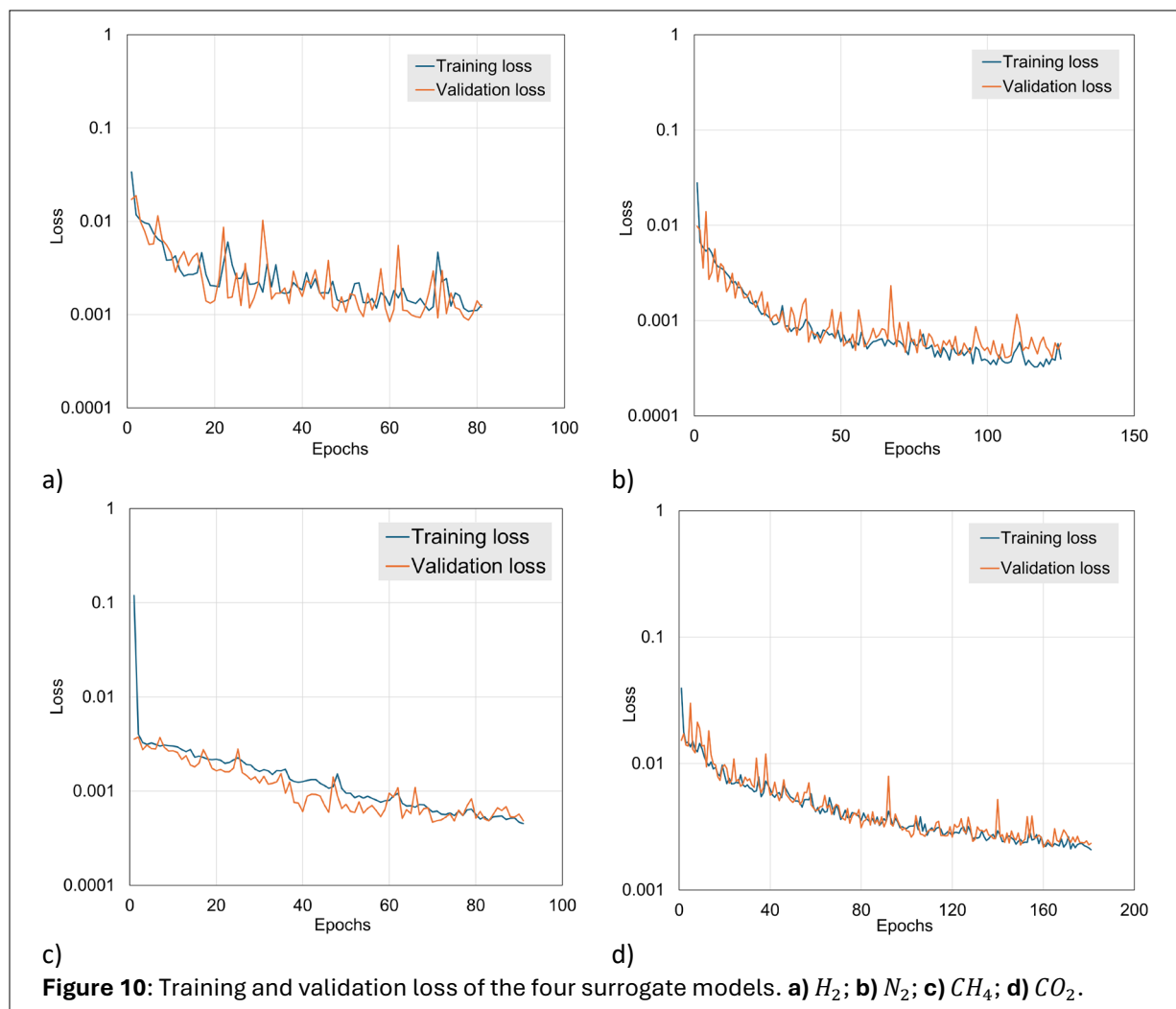
Additionally, R^2 is another commonly used metric and is defined by Equation 23:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (23)$$

2.3.5. Performance of surrogate model

Based on the comprehensive simulation dataset, we developed four surrogate models using CNN-BiLSTM-Attention network to separately predict the production of CH_4 , N_2 , H_2 , and CO_2 . The temporal evolution of training and validation losses for four surrogate models is illustrated in Figure 10. All models demonstrate consistent convergence in terms of MSE loss, initiating from approximately 10^{-1} and descending to 10^{-3} – 10^{-4} . The H_2 model exhibits mild oscillations during training while maintaining convergence, reaching stability at approximately 80 epochs. The N_2 model displays a smoother convergence trajectory, requiring an extended training duration of 130 epochs to achieve optimal performance. Notably, the CH_4 model demonstrates the most stable convergence, with minimal divergence between training and validation losses, completing training within 90 epochs. The CO_2 model takes the longest training period (180 epochs) and shows moderate fluctuations in validation loss, though ultimately achieving stable convergence. The close alignment between training and validation losses across all models indicates effective generalization without significant overfitting, suggesting robust model architectures suitable for their respective prediction tasks.

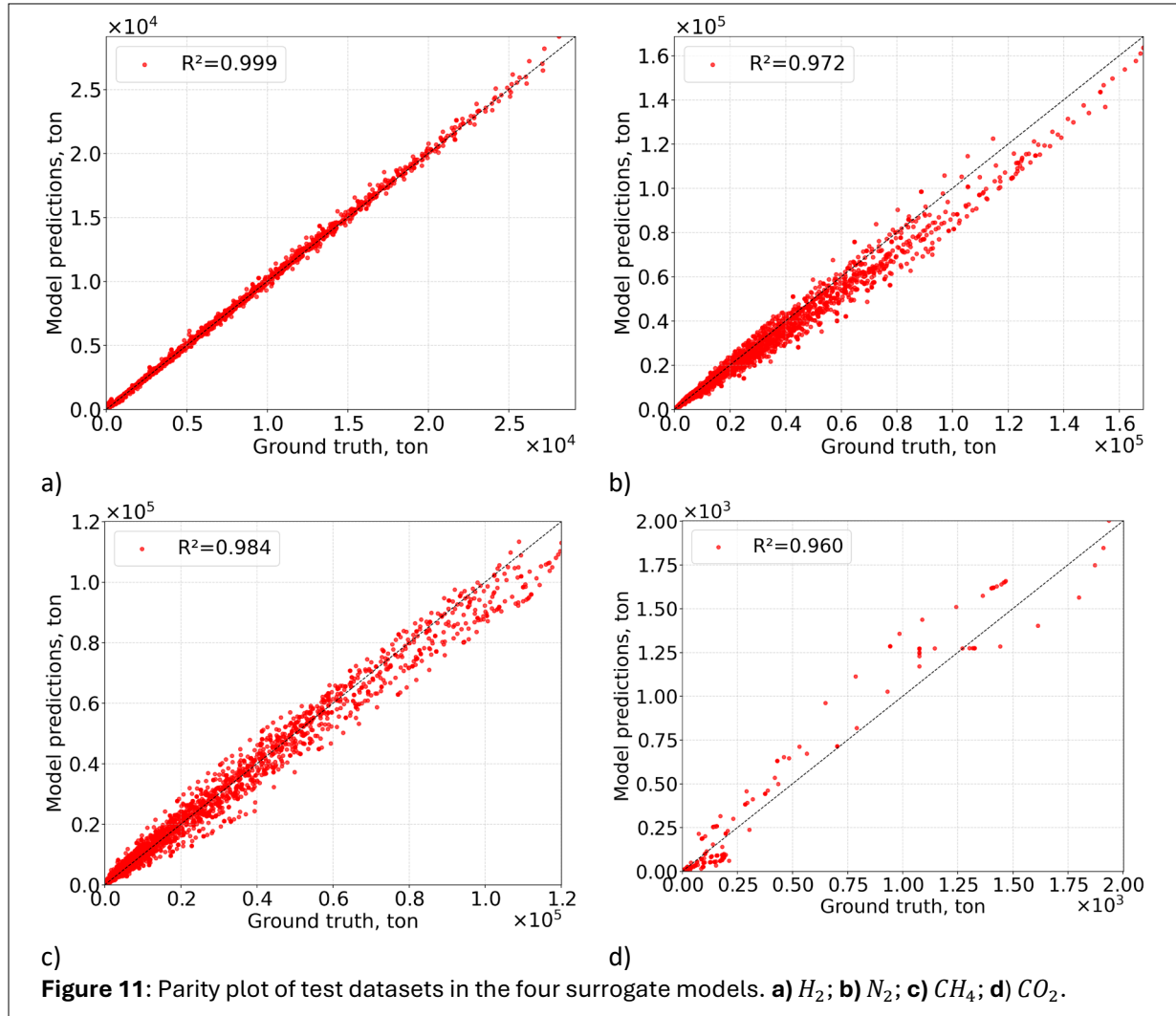
Figure 11 presents parity plots comparing model predictions against ground truth values for the four surrogate models. The H_2 model achieves exceptional performance with an R^2 value of 0.999, demonstrating nearly perfect alignment along the parity line and minimal scatter in predictions. The N_2 model yields an R^2 value of 0.972, exhibiting good correlation despite showing increased scatter at higher values. The CH_4 model maintains strong predictive capability with an R^2 of 0.984, displaying consistent performance across the prediction range with moderate scatter at elevated values. The CO_2 model, while achieving a respectable R^2 of 0.960, shows the most pronounced deviation from the parity line, particularly in the higher value regions, and exhibits distinct clustering patterns in its predictions. All models maintain R^2 values above 0.960, indicating strong overall predictive capabilities, though with varying degrees of accuracy and consistency.



The performances demonstrate that the CNN-BiLSTM-Attention architecture successfully captures the underlying patterns in all four different produced gas components, with particularly outstanding performance in predicting H_2 behavior. The varying levels of prediction accuracy across different fluid components may inform future model refinements and applications in similar chemical systems.

2.4. Stochastic optimization

Optimization is commonly used to explore the search space of the simulated model, providing a deeper understanding of the problem and identifying the optimal values for decision parameters. This process involves addressing challenges, such as local optima, computational cost, and uncertainties (41). Stochastic optimization (56) aims to find optimal solutions that exhibit minimal variability in response to probable uncertainties in reservoir model parameters. Given the unavoidable uncertainties in geological properties and varying operational variables, developing stochastic optimization is essential to quantitatively assess these uncertainties and provide reliable estimates for their impact on UHS operations. In this context, the aforementioned well-designed surrogate models are seamlessly integrated into the stochastic optimization workflow, providing the effectiveness of the optimized well controls in a probabilistic form with variations in model inputs. Genetic Algorithm (GA) is used for the UHS optimization problem due to its key advantages, including its non-gradient descent formulation and its ability to effectively handle the complexities of real-



world problems. The primary objective of this optimization process is to maximize the Net Present Value (NPV), ensuring economically viable and reliable operational strategies under uncertain conditions.

2.4.1. Optimization problem of UHS

In this work, the optimization problem focuses on estimating the well controls that maximize the NPV of the UHS in a depleted natural gas reservoir. The NPV is defined as (Eq. 24):

$$J(m_i, u) = \sum_{t=1}^{N_t} \left[\frac{1}{(1+R)^2} \left(\sum_{j=1}^{N_{prod}} (q_{H_2,j}^t \times P_{H_2,sell} + q_{cg,j}^t \times P_{cg} - q_{w,j}^t \times P_w) - \sum_{k=1}^{N_{inj}} (q_{H_2,k}^t \times P_{H_2,buy}) \right) \right] - Q_{cg,inj} \times P_{cg} + (Q_{CO_2,inj} - Q_{CO_2,prod}) \times B_{CO_2} - CapEx - OpEx \quad (24)$$

where m_i is the i -th realization of the reservoir model; u is an N_u -dimensional column vector that contains all the well controls at all wells during the UHS operations; t denotes the t -th timestep of the reservoir simulation; N_t is the total number of timesteps; R is the annual discount rate; N_{inj} and N_{prod} represent the total number of injection and production wells, respectively; the variables $q_{H_2,prod}^t$, $q_{cg,prod}^t$, $q_{w,prod}^t$ denote

the production rate of H_2 , cushion gas and water in CO_2/day , respectively; $q_{H_2, inj}^t$ denotes the injection rate of H_2 in m^3/day ; $Q_{cg, inj}^t$ is the total injection of cushion gas in m^3 ; $Q_{CO_2, inj}^t$ and $Q_{CO_2, prod}^t$ are the total injection (as cushion gas) and production amount of CO_2 in m^3 respectively; the prices $P_{H_2, sell}$ and $P_{H_2, buy}$ represent the selling and buying price of H_2 in $$/kg$; P_{cg} is the cost of cushion gas in $$/kg$; P_w is the disposal cost associated with water in $$/kg$; B_{CO_2} is the bonus received for storing CO_2 in $$/kg$. The bonus is calculated based on the effective accumulated storage of CO_2 in the subsurface reservoir, rather than being accounted for on a daily basis. Additionally, $CapEx$ and $OpEx$ denote the capital and operating expenditure, respectively, both measured in $€$. For the examples considered in this work, we neglect the disposal costs associated with the separation of produced gas mixtures. However, we account for fluctuations in the market price of H_2 while assuming a fixed price for the cushion gas. The cost of purchasing cushion gas is separated from the daily calculations as the cushion gas is injected only once at the beginning of each cycling schedule with a fixed amount, and no additional cushion gas is injected during the subsequent cycles. Moreover, the water-related terms are set to zero, as no water is produced.

To determine the optimal well controls u for UHS operations, stochastic optimization is employed to maximize the reservoir performance metric, while accounting for uncertainties in the geological parameters, denoted by m_i . Due to these uncertainties, an ensemble of stochastic forward models is evaluated, resulting in a corresponding ensemble of J values. Consequently, the expected value of J is used as the objective function in the stochastic optimization and is defined as (Eq. 25):

$$J_E(u) = \frac{1}{N_e} \sum_{i=1}^{N_e} J(m_i, u) \quad (25)$$

where $J_E(u)$ denotes the approximated expectation of NPV across all geological realizations under well control strategy u ; N_e is the number of geological model realizations used to characterize reservoir uncertainty. In this study, 100 equally probable permeability fields were used to assess the impact of geological uncertainty on NPV during the optimization process.

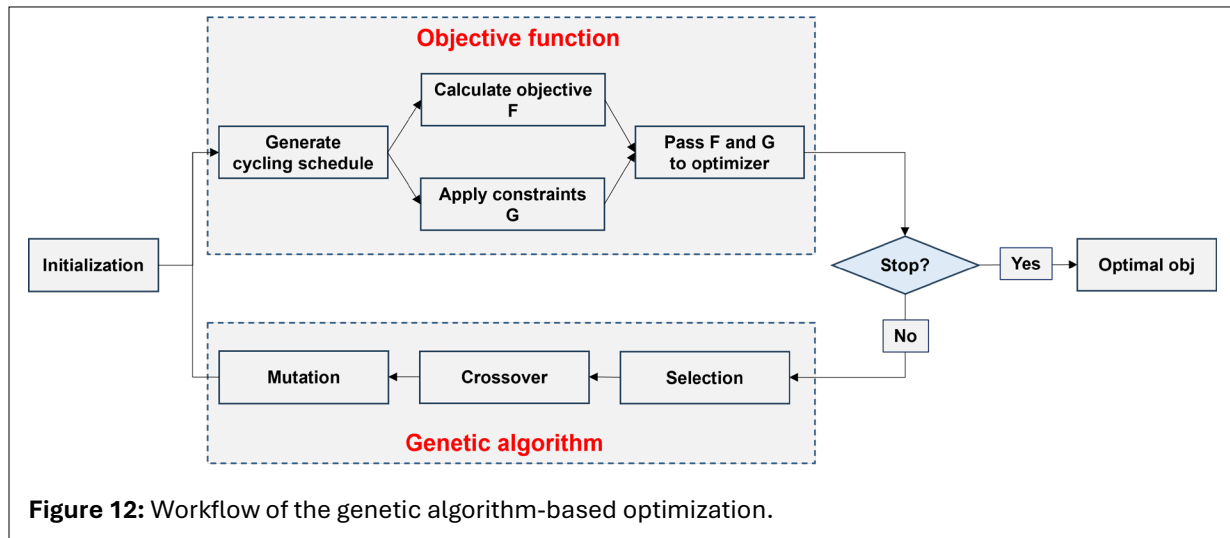
As introduced in Section 2.3.1, the well control strategy is parameterized by $N_u = 6$ decision variables, including the coefficient $coeff$ and exponential index k of the decline curve used to constrain the BHPs. These variables are subject to simple bound constraints, with their respective lower and upper bounds denoted as u_i^{low} and u_i^{up} . The stochastic optimization problem is therefore formulated as follows (Eq. 26):

$$\begin{aligned} & \text{minimize} && -J_E(u)_{u \in R^{N_u}} \\ & \text{subject to} && g_j(u) = BHP_{cal}^{5th} - BHP^{low} \geq 0, \quad j = 1, 2, 3, \dots, N_e \\ & && u_i^{low} \leq u_i \leq u_i^{up}, \quad i = 1, 2, 3, \dots, N_u \end{aligned} \quad (26)$$

where BHP_{cal}^{5th} is the BHP during the final well-open period, as determined by the exponential decline of BHP; BHP^{low} is the lower bound of BHP, set to 2000 psia as shown in Table 2; g_j denotes the j -th inequality constraint, which is explicitly enforced during optimization to ensure that the BHP during the final well-open period remains above 2000 psia.

2.4.2. Genetic algorithm

The optimization process integrates the objective function with a genetic algorithm (GA) (41), as illustrated in Figure 12. The workflow begins with generating an initial population by sampling from the search space. Each generation consists of 60 individuals, where each individual is a vector representing well controls across 72 time steps. The CNN-BiLSTM-Attention surrogate models are then utilized to compute the objective function (F), serving as a computationally efficient surrogate for the numerical simulator. In parallel, a hard constraint (G) is defined to ensure that the BHP during the final well-open period remains above its specified lower bound. Both the objective function (F) and the constraint (G) are passed to the optimizer, which



internally handles constraint feasibility during the selection process. After evaluating the current generation, the GA checks the termination criteria. If these criteria are not satisfied, the algorithm applies a series of evolutionary operators to generate the next generation. First, individuals are selected using the binary tournament method combined with crowding distance sorting, ensuring diversity among the selected solutions. Next, crossover is applied to generate offspring by recombining two parent solutions, typically using single-point or double-point crossover methods. This operator primarily exploits the search space by combining promising solutions. Mutation introduces small alterations to one or more genes in the offspring to maintain genetic diversity and enhance global exploration. These steps produce a new generation, and the process repeats until the termination condition is met, such as reaching the maximum number of generations. The maximum generation size in this study is 40. This structured approach balances global and local search, leveraging crossover for exploitation and mutation for exploration, ultimately converging toward optimal solutions.

As discussed in Section 2.3.1, to satisfy the constraints on BHP and enhance hydrogen recovery efficiency, a monotonically decreasing discrete exponential decline curve is introduced to generate five decreasing BHP values corresponding to the five cycles. To meet the BHP constraint, the last value of the decline curve is set to be greater than the lower bound of BHP, serving as a constraint within the optimization algorithm. This approach reduces the five BHP values to two parameters: *coeff* and *k*.

The optimization algorithm is implemented based on the open-source library pymoo (5). The hyperparameters of the optimization algorithm are detailed in Table 5.

Table 5: Hyperparameters of the genetic algorithm

Parameter	Value
Generation size	40
Population size	60
Offspring size	60
Crossover method	Stimulated binary
Crossover probability	0.9
Mutation method	Polynomial
Mutation probability	0.1

3. RESULTS AND ANALYSIS

3.1. Reservoir simulation results

3.1.1. Sensitivity analysis

In this section, we conduct a sensitivity analysis to evaluate the effects of microbial reaction and cushion gas type on H_2 production during UHS operations. The six scenarios considered are summarized in Table 6.

Table 6: Designed cases for sensitivity analysis	
Case	Description
Case 1	Includes heterogeneity, anisotropy, relative permeability and cycling schedule, without cushion gas injection or micro-bio reactions
Case 2	Built on case 1 by adding microbial reaction (methanation)
Case 3	Built on case 1 by adding CO_2 as cushion gas
Case 4	Built on case 3 by adding microbial reaction (methanation)
Case 5	Built on case 1 by adding N_2 as cushion gas
Case 6	Built on case 5 by adding microbial reaction (methanation)

Figure 13 presents the results, illustrating cumulative H_2 production over time for the six cases. The figure highlights the interplay of microbial activity and cushion gas selection. Case 1, serving as the base model, exhibits a consistent stepwise increase in H_2 production and acts as the benchmark. Introducing microbial reaction (Case 2) results in a slight reduction of H_2 production, approximately 1.7% at the final timestep. This decrease is attributed to the methanation reaction, where H_2 is consumed by the microbes to form CH_4 . Cases 3 and 5 explore the use of CO_2 and N_2 as cushion gases, respectively, without considering microbial reactions. Initially, H_2 production in Case 3 is lower than in Case 5, but this trend reverses after 2005. This behavior is likely due to the higher solubility of CO_2 than that of N_2 , as CO_2 dissolves in the aqueous phase initially and remains in the gaseous state once saturation is reached. Another possible reason is that N_2 is much more compressible than CO_2 . Cases 4 and 6 incorporate microbial reactions into Cases 3 and 5, respectively. The combination of CO_2 and microbial reactions (Case 4) results in a substantial decrease in H_2 production, as CO_2 reacts with H_2 to form CH_4 . This exacerbates the effect of microbial activity, yielding the lowest performance among all cases. In contrast, adding microbial reactions to N_2 as a cushion gas (Case 6) leads to slightly lower H_2 production than Case 5, reflecting losses due to microbial activity, but the impact is much less severe compared to CO_2 -related cases. The comparison between Case 5, 6 and Case 1 further highlights the suitability of N_2 as a cushion gas. The inset

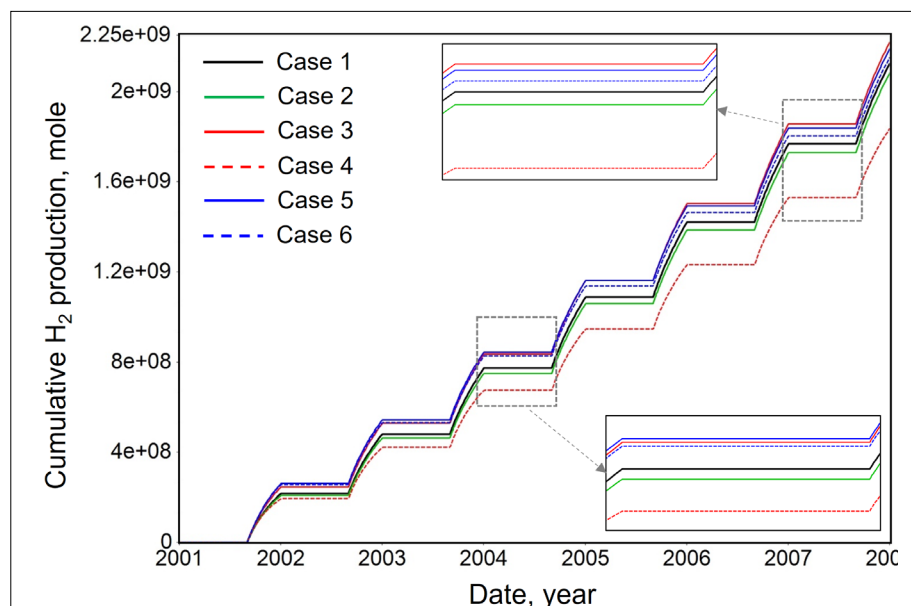
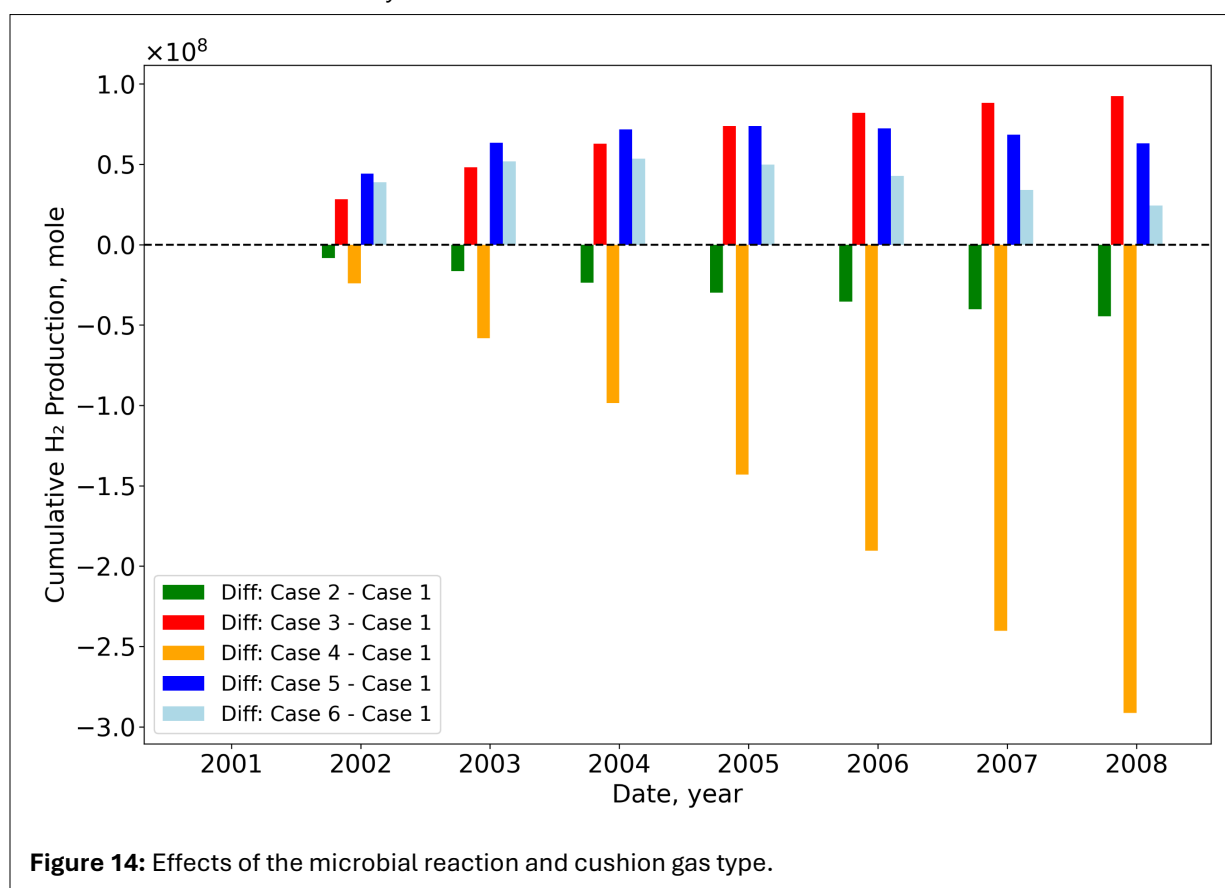


Figure 13: Results of the sensitivity analysis.

boxes in the figure emphasize critical intervals, magnifying the differences between cases, particularly highlighting the distinctive variation features of CO_2 -related cases. Overall, the analysis demonstrates that N_2 is the optimal cushion gas for maximizing H_2 recovery, while the combination of CO_2 and microbial activity poses significant challenges that must be addressed.

To highlight the effects of microbial reaction and cushion gas types on H_2 production, we replot Figure 13 as Figure 14. This figure shows the differences in cumulative H_2 production between Cases 2–6 and the base case (Case 1), evaluated at the beginning of each year from 2001 to 2008. The values represent the positive or negative impacts on H_2 production. The green bars confirm that H_2 production decreases when the microbial reaction is considered and exhibits an overall increasing trend. The orange bars further emphasize this finding, showing that adding CO_2 as a cushion gas intensifies the negative effect. On the other hand, the red and blue bars demonstrate that adding cushion gases can enhance H_2 production. Specifically, when N_2 is used as a cushion gas, H_2 production still increases even in the presence of microbial reaction (as depicted by the light blue bars). However, the magnitude of this increase is reduced compared to the scenario without microbial activity.



3.1.2. A comprehensive case study

Using N_2 as an example of cushion gas, Figure 15 illustrates the cyclical evolution of the N_2 mole fraction in the gas phase. Given that the density of N_2 is higher than that of CH_4 and less than that of H_2O , the injected N_2 descends toward the interface of natural gas and water, as shown in Figure 15a-c. Due to the combined effects of buoyancy and diffusion, the N_2 plume transitions from a near-symmetric shape around the well to an irregular shape, as observed in Figure 15a,b. During the following hydrogen injection and production, a significant amount of N_2 is also produced by comparing Figure 15b and d. By comparing the Figure 15a

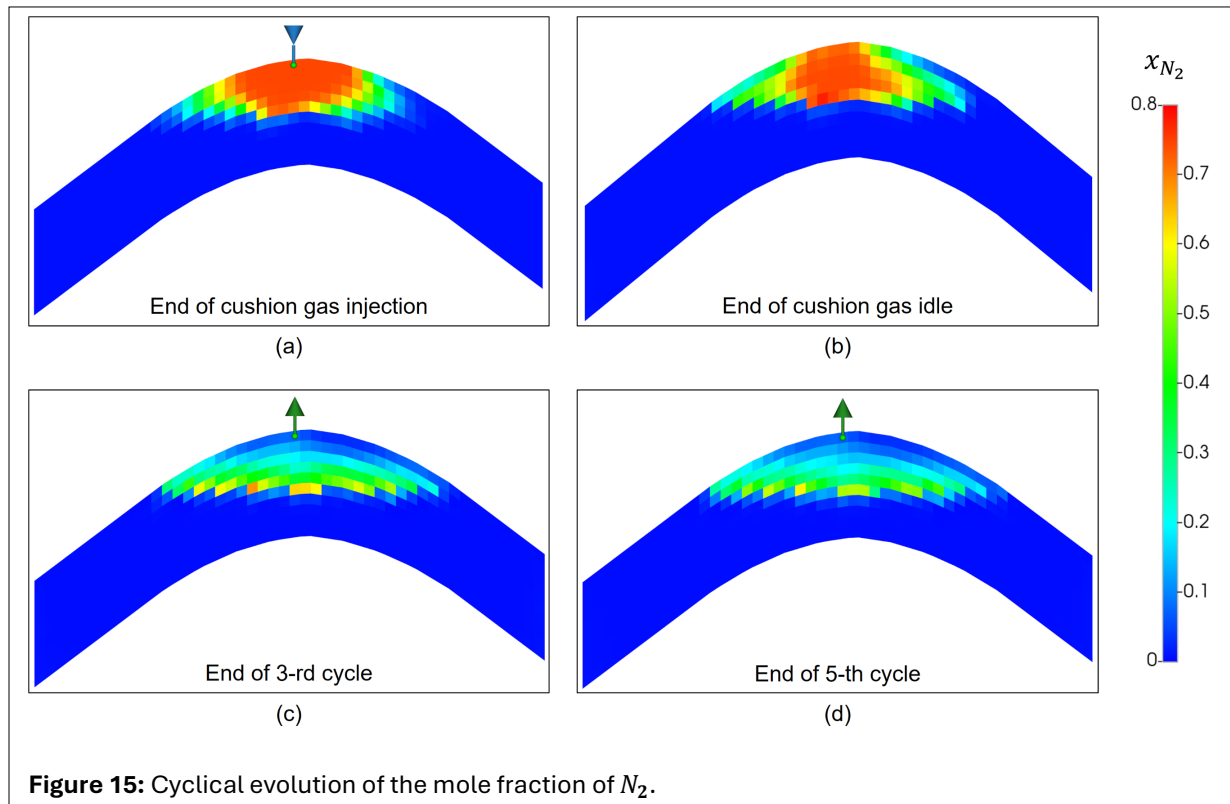


Figure 15: Cyclical evolution of the mole fraction of N_2 .

and **d**, the affected area of the injected N_2 progressively expands as it is displayed by the advancing H_2 plume.

The cyclical evolution of gaseous H_2 mole fraction in transverse and longitudinal directions is depicted in **Figure 16a** and **b**, respectively. This simulation provides insights into the transport behavior of gaseous H_2 during UHS operations. As mentioned above, five hydrogen injection, idle, and production cycles are simulated in this study. **Figure 16a** presents the results for the first, third, and fifth cycles of case 504 in the top reservoir layer. Each cycle in this figure includes the contours of gaseous H_2 at key stages: injection start, injection end, idle end, and production end. Taking the first cycle as an example, the gaseous H_2 plume expands in an irregular shape during the injection period. After a four-month idle, the plume area increases further. In the production period, the plume near the well takes on distinct shapes, leaving behind residual gaseous H_2 . The evolution of the H_2 plume is mainly due to the heterogeneity of the reservoir. During the well idle period, the H_2 plume undergoes significant evolution due to potential forces, such as diffusion, capillary pressure, and other mechanisms. By comparing the three different cycles, the footprint of the H_2 plume progressively enlarges with successive cycles, indicating the accumulation of trapped H_2 in the reservoir. Notably, a significant amount of residual H_2 remains after each cycle for this specific case. **Figure 16b** illustrates the vertical evolution of gaseous H_2 plume after the injection period. Due to gravity, the H_2 plume remains concentrated near the top of the reservoir. The mole fraction of gaseous H_2 peaks near the wellbore and decreases horizontally with the distance away from the well. The affected area of the injected H_2 continues to expand with repeated injections, underscoring the cumulative nature of the storage process.

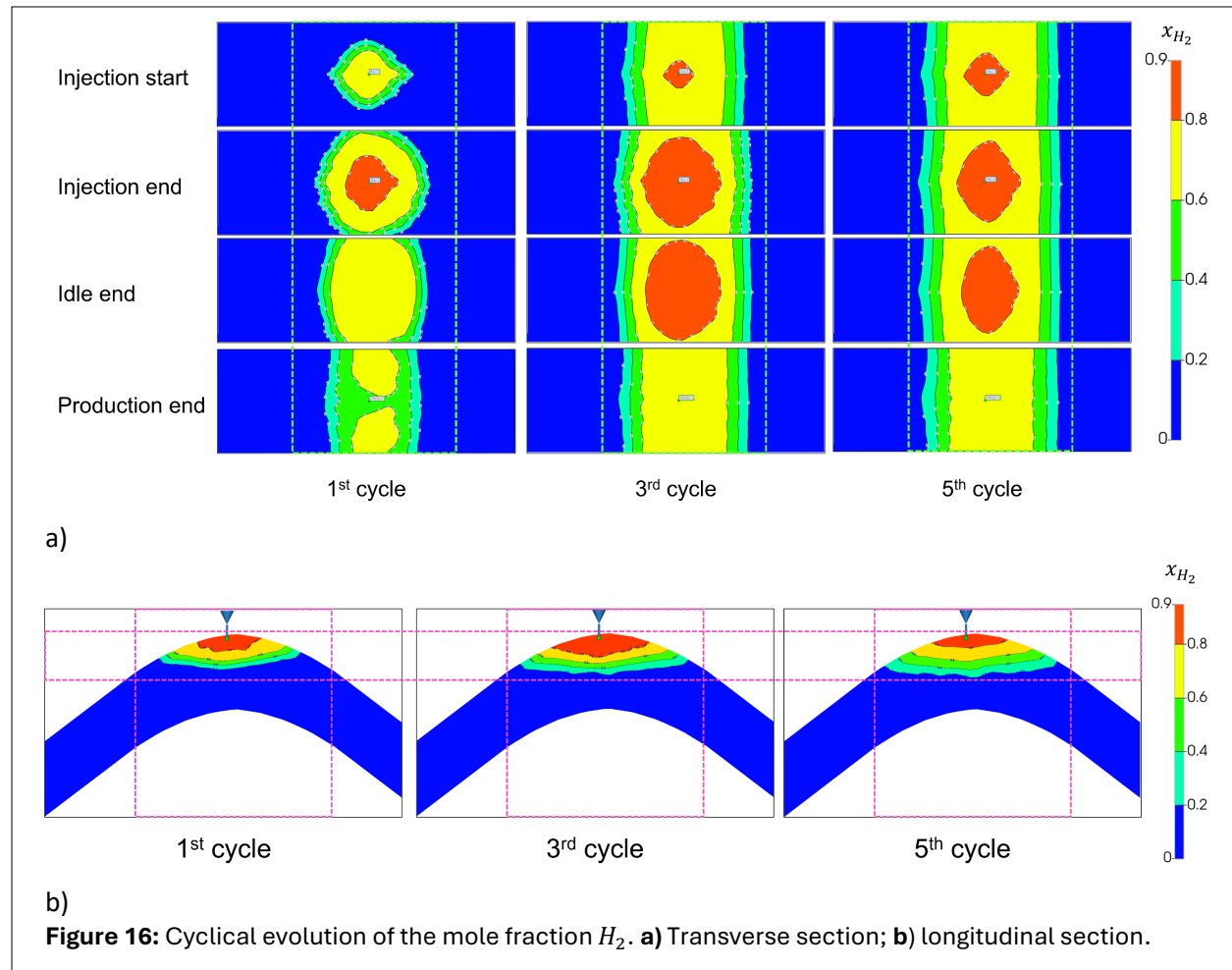
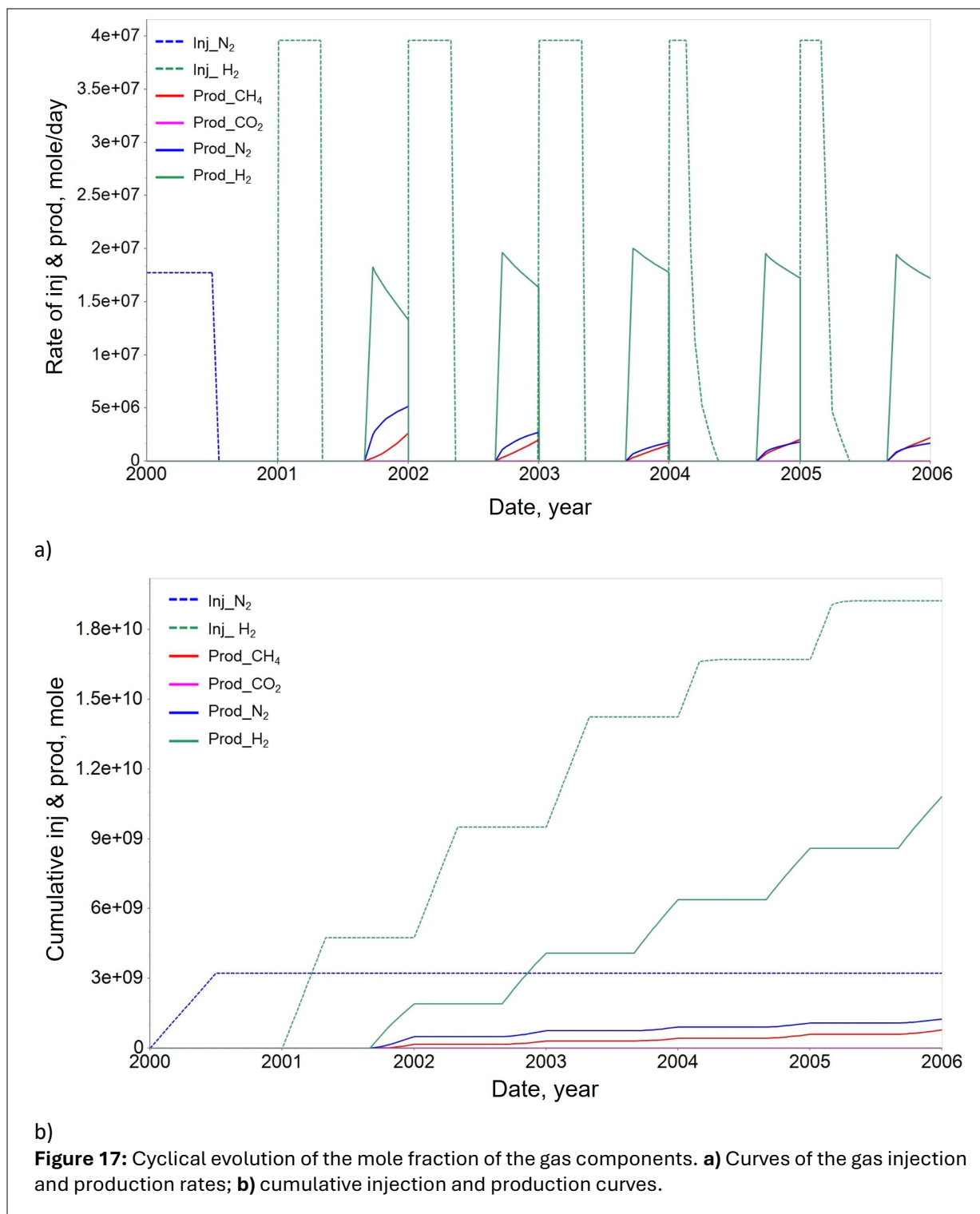


Figure 17 shows the cyclic evolution of gas components through two complementary plots: instantaneous rates (a) and cumulative quantities (b) from January 1, 2000 (i.e., the first timestep) to January 1, 2006 (i.e., the last timestep). **Figure 17a** presents the injection and production rates for various gases throughout the study period. The injection of H_2 (green dashed line) demonstrates a distinct cyclic pattern, characterized by sharp rectangular pulses reaching approximately 4.0×10^7 moles/day. Similarly, N_2 (blue dashed line) shows injection rates with less pronounced variations. The production rates of H_2 (solid green line) exhibit a strong response to the injection cycles, with peak rates reaching approximately 2.0×10^7 moles/day following each injection event. With an increasing number of cycles, the H_2 production rate rises due to reservoir pressure buildup. In contrast, the production rates of N_2 , CH_4 and CO_2 are lower and consistently decrease throughout the cycles. **Figure 17b** illustrates the cumulative quantities of injected and produced gases. After 5 years (2006), the cumulative injected H_2 reaches approximately 1.9×10^{10} moles, following a stepwise increase aligned with the cyclic injection strategy. In comparison, the cumulative injection of N_2 plateaus early at around 3.2×10^9 moles. The cumulative production profiles reveal that H_2 (solid green line) shows the most substantial growth among the produced gases, reflecting its significant recovery over time. N_2 and other produced gases exhibit more modest cumulative increases. Together, the figures highlight the effectiveness of a cyclic injection strategy, particularly for H_2 , in achieving distinct operational periods and demonstrate the system's dynamic response in both short-term production rates and long-term material balance.



3.2. Optimization results

3.2.1. Deterministic optimization for a specific permeability field

To validate the optimization results obtained using surrogate models, we compare them with those derived from the simulator, both initialized with the same well controls for a specific permeability field. **Figure 18** shows the NPV values throughout the optimization process. The optimization result from the surrogate model-based method is labeled as NPV_optimal_DL (solid red line with triangles), while the optimization result based on the physical simulation-based method is denoted as NPV_optimal_sim (dotted blue line with diamonds). Additionally, NPVs are computed using the physical simulator by directly inputting the well control decision vectors obtained at each generation of the DL-based optimization. This result is referred to as NPV_calc_sim and illustrated as a dotted blue line with stars. Initially, NPV_optimal_sim demonstrates a steeper improvement in NPV, rapidly approaching near-optimal solutions within fewer generations. In contrast, NPV_optimal_DL exhibits a slower rate of improvement at early generations but ultimately converges to a slightly higher NPV. The trend of NPV calc sim initially follows NPV_optimal_DL, and subsequently aligns with NPV_optimal_sim. Eventually, the NPVs from all three methods converge to a similar level of objective function values, with final optimal NPVs of 1.43×10^8 \$, 1.32×10^8 \$ and 1.28×10^8 \$, respectively. This comparison supports the validity and reliability of the surrogate model-based optimization strategy.

Although the simulator-based optimization converges faster, it is also essential to compare the CPU time required for the optimization process. To search for the global optimum of NPV, the optimizer requires a large number of simulation evaluations using forward models. As described in [Section 2.4.2](#), the population size is set to 60, meaning that each generation (or iteration) involves 60 simulation runs. With a maximum of 40 generations, this results in a total of 2400 simulation runs for the GA optimization process. The computational cost of forward simulations based on the physical reservoir model is quite high due to the complex physics involved. From our numerical experiments, the average CPU time for a single CMG simulation for one specific permeability field is approximately 210 *seconds*. Assuming an optimistic scenario where all 60 CMG simulations in each GA generation can run in parallel, the total time required per generation would still be 210

seconds. However, this means a total of 140 *minutes*, considering 40 generations. Furthermore, this estimate is optimistic, as parallel computing for coupled simulation models is often constrained by practical memory limitations and license availability. In contrast, the surrogate model, which can be evaluated efficiently in a graphics processing unit (GPU) environment using batch processing, offers significant computational advantages. The total average CPU time for a single prediction of the four

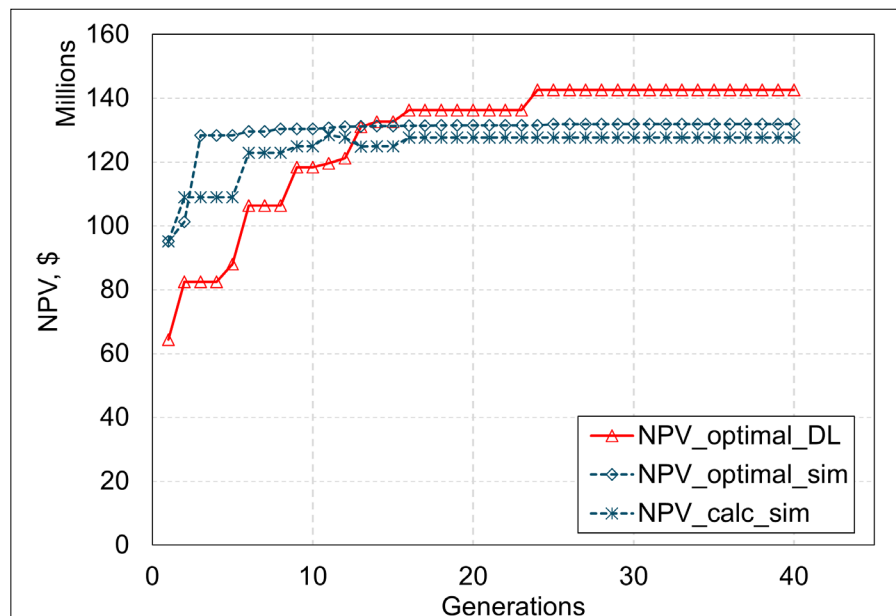
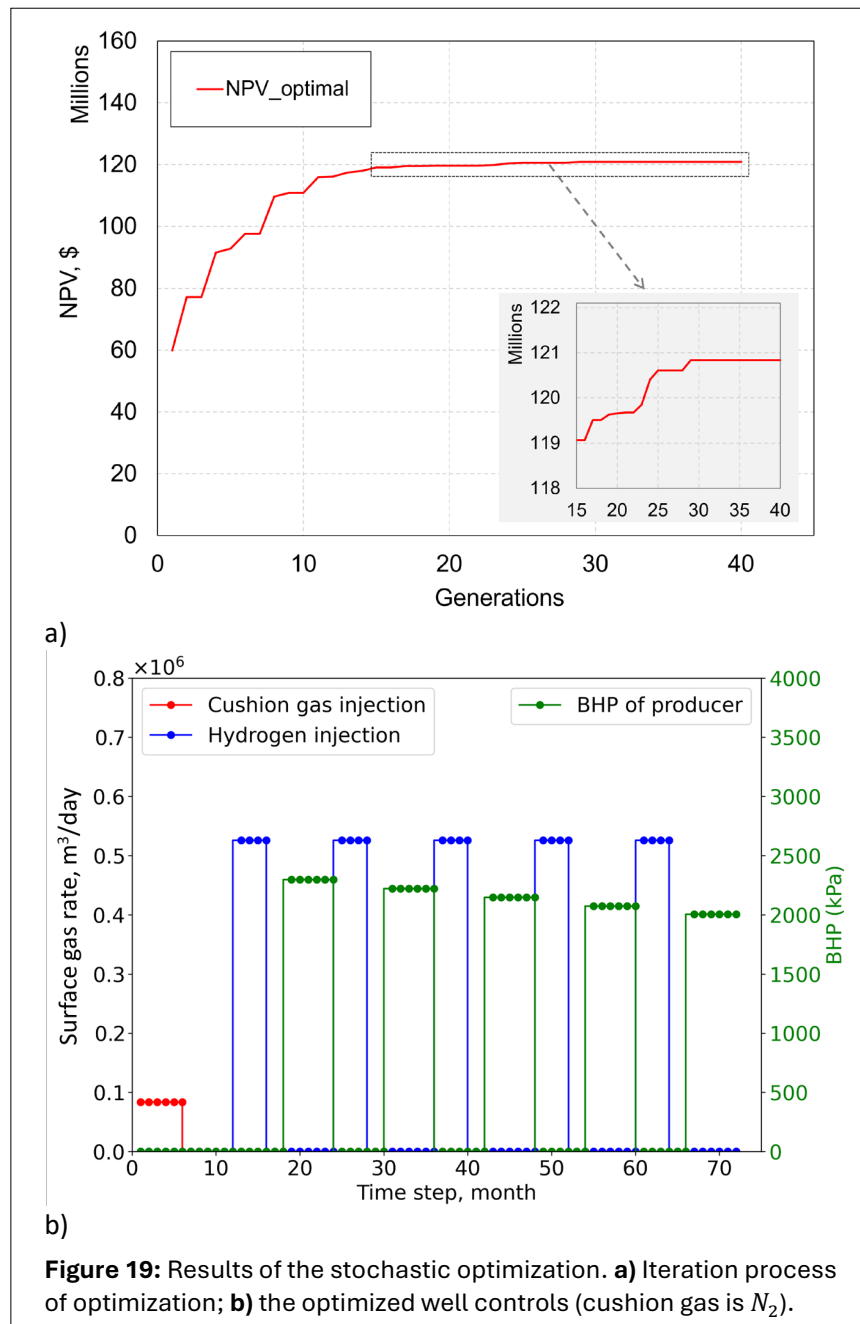


Figure 18: Iteration process of optimization. NPV: net present value; DL: deep learning; calc: calculated; sim = simulation.

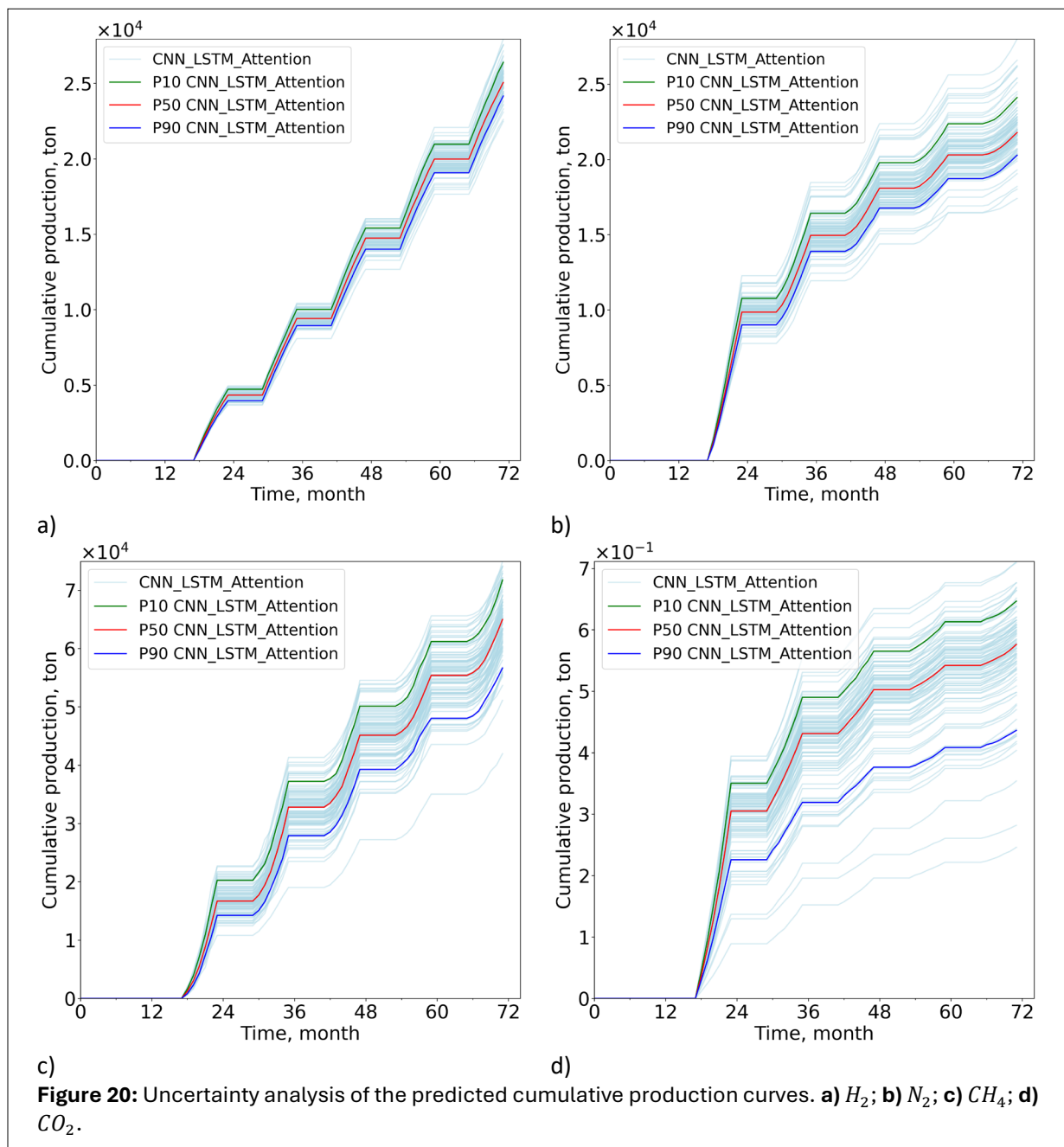
components with the surrogate model is only 4.3×10^{-2} seconds. Considering its low memory occupation and efficient batch processing, the surrogate model requires just 2.87×10^{-2} minutes (1.722 seconds) for each optimization task (40 iterations). This makes the GA optimization using the surrogate model approximately 4878 times faster than the physics-based reservoir simulator. Consequently, the developed surrogate model provides a highly efficient and viable alternative for intensive optimization tasks, significantly reducing computational costs while maintaining accuracy.

3.2.2. Stochastic optimization considering uncertainties in permeability fields

In this section, the GA-based optimization framework is coupled with the surrogate models based on the CNN-BiLSTM-Attention architecture and generalized to account for the uncertainty in permeability fields for UHS operations in depleted natural gas reservoirs. We select 100 equally probable permeability fields to quantify the uncertainty of the geological model during the stochastic optimization. **Figure 19a** illustrates the stochastic optimization process for these permeability fields. In this figure, the solid red line represents the evolution of optimal NPV over one generation, while the dashed blue line denotes the evolution of the average NPV across the same generation. As the number of generations increases, both the optimal and average values of the NPV converge and gradually stabilize. The zoomed-in inset indicates that the optimal value remains unchanged, stopping optimization after reaching the maximum generations (i.e., 40). The final optimal NPV value is 1.208×10^8 \$. **Figure 19b** depicts the well controls corresponding to the optimal NPV. The injection rates of cushion gas and hydrogen are 8.368×10^4 m³/day and 5.259×10^5 m³/day, respectively. The BHPs decrease from 2297.85 kPa to 2004.64 kPa. The well injection and production periods last for 4 months and 6 months, respectively.



Using the 100 equal-probable uncertain permeability fields, the optimized well controls shown in **Figure 19b** are applied to generate 100 cases. These cases are then evaluated using the surrogate models to predict the cumulative production of the four components, as illustrated in **Figure 20**. The light blue lines represent individual predictions from the ensemble model, illustrating variability across 100 realizations of permeability fields. The red dashed lines denote the P50 (median) predictions, while the green and blue dashed lines correspond to the P10 (10th percentile) and P90 (90th percentile), respectively. Specifically, the problem of maximizing NPV is reformulated as minimizing-NPV for our actual optimization implementation. The results show the uncertainty range of predictions, with the spread between P10 and P90 providing a measure of confidence.



The prediction of H_2 exhibits minimal variability, with tightly clustered percentile bands indicating high prediction consistency. In contrast, N_2 and CH_4 demonstrate wider uncertainty bands, suggesting greater variability across the ensemble. The production of CO_2 is the lowest among the four fluid components, with percentile bands reflecting its unique production dynamics. The step-like increments in production curves correspond to the designed well-cycling schedule. This visualization highlights the predictive uncertainty inherent in the ensemble model for different gases under identical production scenarios.

4. CONCLUSION

We develop an efficient approach for robust UHS optimization by integrating a CNN-BiLSTM- Attention surrogate model into a genetic algorithm optimizer tailored for depleted natural gas reservoirs. These surrogate models, trained on comprehensive reservoir simulation datasets, deliver fast and accurate predictions of UHS performance across various geological and operational conditions. The workflow introduced in this study enables rapid and precise evaluation of UHS performance, thereby facilitating the development of effective UHS projects in depleted natural gas reservoirs. The key conclusions are as follows:

1. An advanced reservoir simulation model is developed to capture the complexities of a depleted natural gas reservoir, including heterogeneous permeability fields, relative permeability (drainage and hysteresis), capillary effects, compositional fluid flow, gas diffusion, bio-reactions (e.g., methanation), and gravity segregation. This model provided a rich dataset for training surrogate models.
2. The CNN-BiLSTM-Attention surrogate model demonstrates excellent performance in time-series prediction tasks for UHS applications. The BiLSTM effectively captures temporal dependencies in both forward and backward directions, providing a comprehensive understanding of time-sequential data. The model shows high prediction accuracy overall, with an R_2 score of at least 0.960 for all four components. However, the performance for CO_2 predictions is relatively lower, as evidenced by a lower R_2 value. This is mainly due to the higher variability in CO_2 production across different permeability realizations, which presents a greater challenge for accurate prediction.

Despite this limitation, the model offers a significant advantage in computational efficiency, with an average CPU time per prediction of just 4.3×10^{-2} approximately 4878 times faster than traditional physics-based reservoir simulation models. This work also highlights the value of the developed optimization framework, including uncertainty quantification and the use of the genetic algorithm, which remains highly relevant for future applications. Further improvements to surrogate models, such as exploring alternative architectures, may enhance performance, particularly for gases like CO_2 .

3. A novel stochastic optimization workflow is developed by integrating the surrogate model into a genetic algorithm-based optimizer. The optimizer refines operational parameters to maximize the NPV, ensuring optimal performance of UHS operations.

STATEMENTS AND DECLARATIONS

Acknowledgements

The authors thank the King Abdullah University of Science and Technology (KAUST) for the research funding through grant BAS/1/1423-01-01.

Author Contributions

Zhilei Han: Conceptualization, Methodology, Programming and Analysis, Writing. Zeeshan Tariq: Methodology, Programming, Reviewing. Bicheng Yan: Supervision, Funding Acquisition, Conceptualization, Methodology, Reviewing.

Conflicts of Interest

There are no conflicts of interest to declare.

Data, Code & Protocol Availability

Data and codes can be provided upon request.

Funding Received

Funding was received from KAUST under grant BAS/1/1423-01-01.

ORCID IDs

Zhilei Han

 <https://orcid.org/0000-0002-9042-6034>

Zeeshan Tariq

 <https://orcid.org/0000-0001-5456-7115>

Bicheng Yan

 <https://orcid.org/0000-0002-3356-7594>

REFERENCES

1. Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623–2631. <https://doi.org/10.1145/3292500.3330701>
2. Al-Yaseri, A., Wolff-Boenisch, D., Fauziah, C. A., & Iglaier, S. (2021). Hydrogen wettability of clays: Implications for underground hydrogen storage. *International Journal of Hydrogen Energy*, 46(69), 34356–34361. <https://doi.org/10.1016/j.ijhydene.2021.07.226>
3. Bauer, J. F., Amro, M. M., Nassan, T., & Alkan, H. (2024). Reservoir engineering aspects of geologic hydrogen storage. *International Petroleum Technology Conference*, D031S097R001. <https://doi.org/10.2523/IPTC-23943-MS>
4. Bhosekar, A., & Ierapetritou, M. (2018). Advances in surrogate based modeling, feasibility analysis, and optimization: A review. *Computers & Chemical Engineering*, 108, 250–267. <https://doi.org/10.1016/j.compchemeng.2017.09.017>
5. Blank, J., & Deb, K. (2020). Pymoo: Multi-objective optimization in python. *IEEE Access*, 8, 89497–89509. <https://doi.org/10.1109/ACCESS.2020.2990567>
6. Chai, M., Chen, Z., Nourozieh, H., & Yang, M. (2023). Numerical simulation of large-scale seasonal hydrogen storage in an anticline aquifer: A case study capturing hydrogen interactions and cushion gas injection. *Applied Energy*, 334, 120655. <https://doi.org/10.1016/j.apenergy.2023.120655>
7. Chen, F., Chen, B., Mao, S., Malki, M., & Mehana, M. (2024). Integrating capacity and efficiency for optimal hydrogen storage site selection in saline aquifers. *Energy & Fuels*, 38(5), 4733–4742. <https://doi.org/10.1021/acs.energyfuels.3c04852>
8. Computer Modeling Group. (2022). *GEM User Manual*. Computer Modeling Group (CGM).
9. Delshad, M., Alhotan, M. M., Fernandes, B. R. B., Umurzakov, Y., & Sepehrnoori, K. (2023). Modeling flow and transport in saline aquifers and depleted hydrocarbon reservoirs for hydrogen energy storage. *SPE Journal*, 28(05), 2547–2565. <https://doi.org/10.2118/210351-PA>
10. Feng, Z., Tariq, Z., Shen, X., Yan, B., Tang, X., & Zhang, F. (2024). An encoder-decoder ConvLSTM surrogate model for simulating geological CO₂ sequestration with dynamic well controls. *Gas Science and Engineering*, 125, 205314. <https://doi.org/10.1016/j.jgsce.2024.205314>
11. Gehring, J., Auli, M., Grangier, D., Yarats, D., & Dauphin, Y. N. (2017). Convolutional sequence to sequence learning. *Proceedings of the 34th International Conference on Machine Learning*, 1243–1252. <https://proceedings.mlr.press/v70/gehring17a.html>

12. Graves, A., & Schmidhuber, J. (2005). Framework phoneme classification with bidirectional LSTM and other neural network architectures. *Neural networks*, 18(5-6), 602-610. <https://doi.org/10.1016/j.neunet.2005.06.042>
13. Hagemann, B. (2017). *Numerical and analytical modeling of gas mixing and bio-reactive transport during underground hydrogen storage* [PhD Thesis, Université de Lorraine]. <https://theses.hal.science/tel-01735019>
14. Hagemann, B., Ganzer, L., & Panfilov, M. (2018). *Field scale modeling of bio-reactions during underground hydrogen storage*. 2018, 1–12. <https://doi.org/10.3997/2214-4609.201802116>
15. Heinemann, N., Alcalde, J., Miocic, J. M., Hangx, S. J. T., Kallmeyer, J., et al. (2021). Enabling large-scale hydrogen storage in porous media – the scientific challenges. *Energy & Environmental Science*, 14(2), 853–864. <https://doi.org/10.1039/D0EE03536J>
16. Heinemann, N., Scafidi, J., Pickup, G., Thaysen, E. M., Hassanpouryouzband, A., et al. (2021). Hydrogen storage in saline aquifers: The role of cushion gas for injection and production. *International Journal of Hydrogen Energy*, 46(79), 39284–39296. <https://doi.org/10.1016/j.ijhydene.2021.09.174>
17. Helton, J. C., & Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81(1), 23–69. [https://doi.org/10.1016/S0951-8320\(03\)00058-9](https://doi.org/10.1016/S0951-8320(03)00058-9)
18. Hemme, C., & Van Berk, W. (2018). Hydrogeochemical modeling to identify potential risks of underground hydrogen storage in depleted gas fields. *Applied Sciences*, 8(11), 2282. <https://doi.org/10.3390/app8112282>
19. Higgs, S., Da Wang, Y., Sun, C., Ennis-King, J., Jackson, S. J., et al. (2022). In-situ hydrogen wettability characterisation for underground hydrogen storage. *International Journal of Hydrogen Energy*, 47(26), 13062–13075. <https://doi.org/10.1016/j.ijhydene.2022.02.022>
20. Higgs, S., Wang, Y. D., Sun, C., Ennis-King, J., Jackson, S. J., et al. (2024). Direct measurement of hydrogen relative permeability hysteresis for underground hydrogen storage. *International Journal of Hydrogen Energy*, 50, 524–541. <https://doi.org/10.1016/j.ijhydene.2023.07.270>
21. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
22. Hogeweg, S., Strobel, G., & Hagemann, B. (2022). Benchmark study for the simulation of Underground Hydrogen Storage operations. *Computational Geosciences*, 26(6), 1367–1378. <https://doi.org/10.1007/s10596-022-10163-5>
23. Huang, T., Moridis, G. J., Blasingame, T. A., Abdulkader, A. M., & Yan, B. (2023). Compositional reservoir simulation of underground hydrogen storage in depleted gas reservoirs. *International Journal of Hydrogen Energy*, 48(92), 36035–36050. <https://doi.org/10.1016/j.ijhydene.2023.05.355>
24. Jafari Raad, S. M., Ranjbar, E., Hassanzadeh, H., & Leonenko, Y. (2023). Hydrogen-brine mixture PVT data for reservoir simulation of hydrogen storage in deep saline aquifers. *International Journal of Hydrogen Energy*, 48(2), 696–708. <https://doi.org/10.1016/j.ijhydene.2022.09.222>
25. Jensen, J. L. (Ed.). (2000). *Statistics for petroleum engineers and geoscientists* (2nd ed). Elsevier.
26. Kanaani, M., & Sedaee, B. (2022). Impact of dilation and irreversible compaction on underground hydrogen storage in depleted hydrocarbon reservoirs. *Energy & Fuels*, 36(22), 13506–13517. <https://doi.org/10.1021/acs.energyfuels.2c02150>
27. Kanaani, M., Sedaee, B., & Asadian-Pakfar, M. (2022). Role of cushion gas on underground hydrogen storage in depleted oil reservoirs. *Journal of Energy Storage*, 45, 103783. <https://doi.org/10.1016/j.est.2021.103783>
28. Kanaani, M., Sedaee, B., Asadian-Pakfar, M., Gilavand, M., & Almahmoudi, Z. (2023). Development of multi-objective co-optimization framework for underground hydrogen storage and carbon dioxide storage using machine learning algorithms. *Journal of Cleaner Production*, 386, 135785. <https://doi.org/10.1016/j.jclepro.2022.135785>
29. Katz, D. L., & Tek, M. R. (1981). Overview on underground storage of natural gas. *Journal of Petroleum Technology*, 33(06), 943–951. <https://doi.org/10.2118/9390-PA>
30. Land, C. S. (1968). Calculation of imbibition relative permeability for two- and three-phase flow from rock properties. *Society of Petroleum Engineers Journal*, 8(02), 149–156. <https://doi.org/10.2118/1942-PA>
31. Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>

32. Li, Y., & Nghiem, L. X. (1986). Phase equilibria of oil, gas and water/brine mixtures from a cubic equation of state and Henry's law. *The Canadian Journal of Chemical Engineering*, 64(3), 486–496.
<https://doi.org/10.1002/cjce.5450640319>
33. Liu, N., Kovscek, A. R., Fernø, M. A., & Dopffel, N. (2023). Pore-scale study of microbial hydrogen consumption and wettability alteration during underground hydrogen storage. *Frontiers in Energy Research*, 11, 1124621.
<https://doi.org/10.3389/fenrg.2023.1124621>
34. Loshchilov, I., & Hutter, F. (2017). *Decoupled weight decay regularization*. arXiv.
<https://doi.org/10.48550/ARXIV.1711.05101>
35. Lu, W., Li, J., Wang, J., & Qin, L. (2021). A CNN-BiLSTM-AM method for stock price prediction. *Neural Computing and Applications*, 33(10), 4741–4753. <https://doi.org/10.1007/s00521-020-05532-z>
36. Luboń, K., & Tarkowski, R. (2020). Numerical simulation of hydrogen injection and withdrawal to and from a deep aquifer in NW Poland. *International Journal of Hydrogen Energy*, 45(3), 2068–2083.
<https://doi.org/10.1016/j.ijhydene.2019.11.055>
37. Luboń, K., & Tarkowski, R. (2021). Influence of capillary threshold pressure and injection well location on the dynamic CO₂ and H₂ storage capacity for the deep geological structure. *International Journal of Hydrogen Energy*, 46(58), 30048–30060. <https://doi.org/10.1016/j.ijhydene.2021.06.119>
38. Lysy, M., Fernø, M., & Ersland, G. (2021). Seasonal hydrogen storage in a depleted oil and gas field. *International Journal of Hydrogen Energy*, 46(49), 25160–25174. <https://doi.org/10.1016/j.ijhydene.2021.05.030>
39. Malki, M. L., Chellal, H., Mao, S., Rasouli, V., & Mehana, M. (2024). A critical review of underground hydrogen storage: From fundamentals to applications, unveiling future frontiers in energy storage. *International Journal of Hydrogen Energy*, 79, 1365–1394. <https://doi.org/10.1016/j.ijhydene.2024.07.076>
40. Mao, S., Chen, B., Malki, M., Chen, F., Morales, M., et al. (2024). Efficient prediction of hydrogen storage performance in depleted gas reservoirs using machine learning. *Applied Energy*, 361, 122914.
<https://doi.org/10.1016/j.apenergy.2024.122914>
41. Mirjalili, S. (2019). Genetic algorithm. In S. Mirjalili, *Evolutionary Algorithms and Neural Networks* (Vol. 780, pp. 43–55). Springer International Publishing. https://doi.org/10.1007/978-3-319-93025-1_4
42. Muhammed, N. S., Haq, M. B., Al Shehri, D. A., Al-Ahmed, A., Rahman, M. M., et al. (2023). Hydrogen storage in depleted gas reservoirs: A comprehensive review. *Fuel*, 337, 127032. <https://doi.org/10.1016/j.fuel.2022.127032>
43. Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., et al. (2017). *Automatic differentiation in PyTorch*.
<https://openreview.net/forum?id=BJJsrnfCZ>
44. *Power-to-Gas: The Case for Hydrogen White Paper*. (2015, October 18). California Hydrogen Business Council.
<https://static1.squarespace.com/static/58e8f58d20099ea6eb9ab918/t/5e62a5d3f6a1d1586bc39be5/1583523286672/CHBC+Hydrogen+Energy+Storage+White+Paper+FINAL.pdf>
45. Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., & Kevin Tucker, P. (2005). Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*, 41(1), 1–28. <https://doi.org/10.1016/j.paerosci.2005.02.001>
46. Reitenbach, V., Ganzer, L., Albrecht, D., & Hagemann, B. (2015). Influence of added hydrogen on underground gas storage: A review of key issues. *Environmental Earth Sciences*, 73(11), 6927–6937. <https://doi.org/10.1007/s12665-015-4176-2>
47. Sainz-Garcia, A., Abarca, E., Rubi, V., & Grandia, F. (2017). Assessment of feasible strategies for seasonal underground hydrogen storage in a saline aquifer. *International Journal of Hydrogen Energy*, 42(26), 16657–16666. <https://doi.org/10.1016/j.ijhydene.2017.05.076>
48. Sekar, L. K., Kiran, R., Okoroafor, E. R., & Wood, D. A. (2023). Review of reservoir challenges associated with subsurface hydrogen storage and recovery in depleted oil and gas reservoirs. *Journal of Energy Storage*, 72, 108605. <https://doi.org/10.1016/j.est.2023.108605>
49. Soave, G. (1972). Equilibrium constants from a modified Redlich-Kwong equation of state. *Chemical Engineering Science*, 27(6), 1197–1203. [https://doi.org/10.1016/0009-2509\(72\)80096-4](https://doi.org/10.1016/0009-2509(72)80096-4)
50. Stolten, D., & Emonts, B. (2016). *Hydrogen Science and Engineering: Materials, Processes, Systems, and Technology, 2 volume set / Wiley* (Vols. 1 & 2). Wiley. <https://www.wiley.com/en-us/Hydrogen+Science+and+Engineering%3A+Materials%2C+Processes%2C+Systems%2C+and+Technology%2C+2+Volume+Set-p-9783527332380>

51. Sun, Q., Zhang, M., & Ertekin, T. (2024). The design of hydrogen saline aquifer storage processes using a machine-learning assisted multiobjective optimization protocol. *SPE Journal*, 29(04), 2086–2105. <https://doi.org/10.2118/218405-PA>
52. Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136. [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5)
53. Tsao, D.T. (2003). Overview of Phytotechnologies. In: Tsao, D.T. (eds) Phytoremediation. Advances in Biochemical Engineering/Biotechnology, vol 78. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45991-X_1
54. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., et al. (2017). *Attention is all you need*. arXiv. <https://doi.org/10.48550/ARXIV.1706.03762>
55. Wang, G., Pickup, G., Sorbie, K., & Mackay, E. (2021). *Compositional simulation of hydrogen storage with carbon dioxide cushion gas in subsurface heterogeneous porous media*. 2021, 1–5. <https://doi.org/10.3997/2214-4609.202121032>
56. Yan, B., Gudala, M., & Sun, S. (2023). Robust optimization of geothermal recovery based on a generalized thermal decline model and deep learning. *Energy Conversion and Management*, 286, 117033. <https://doi.org/10.1016/j.enconman.2023.117033>
57. Yan, B., Harp, D. R., Chen, B., Hoteit, H., & Pawar, R. J. (2022). A gradient-based deep neural network model for simulating multiphase flow in porous media. *Journal of Computational Physics*, 463, 111277. <https://doi.org/10.1016/j.jcp.2022.111277>
58. Yekta, A. E., Manceau, J.-C., Gaboreau, S., Pichavant, M., & Audigane, P. (2018). Determination of hydrogen–water relative permeability and capillary pressure in sandstone: Application to underground hydrogen injection in sedimentary formations. *Transport in Porous Media*, 122(2), 333–356. <https://doi.org/10.1007/s11242-018-1004-7>
59. Zeng, L., Vialle, S., Ennis-King, J., Esteban, L., Sarmadivaleh, M., et al. (2023). Role of geochemical reactions on caprock integrity during underground hydrogen storage. *Journal of Energy Storage*, 65, 107414. <https://doi.org/10.1016/j.est.2023.107414>
60. Zhong, Z., Sun, A. Y., & Jeong, H. (2019). Predicting CO₂ plume migration in heterogeneous formations using conditional deep convolutional generative adversarial network. *Water Resources Research*, 55(7), 5830–5851. <https://doi.org/10.1029/2018WR024592>