Citation (APA)
Loyola, A., Voskov, D., Farajzadeh, R., de Borst, K., & Geiger, S. (2026). Evaluation of operational strategies for underground hydrogen storage in depleted gas fields under diverse geological scenarios: guidelines for site screening and development planning. *Environmental Science: Advances*. https://doi.org/10.1039/d6va00162a

Environmental Science Advances

PAPER

Published on 25 June 2026

Cite this: DOI: 10.1039/d6va00162a

Evaluation of operational strategies for underground hydrogen storage in depleted gas fields under diverse geological scenarios: guidelines for site screening and development planning

Ana Loyola, [*ac] Denis Voskov,[a] Rouhi Farajzadeh,[ab] Karin de Borst[b] and Sebastian Geiger[a]

Underground hydrogen storage in depleted gas fields is a potential solution for large-scale, seasonal storage of hydrogen, in support of the decarbonization of energy systems and other industrial activities. Its viability depends on the performance of the storage operations, which is influenced by the interaction between reservoir geology and operational strategies. However, general guidelines for development planning that account for geological uncertainty are still lacking. In addition, existing site screening criteria remain limited in that they do not account for how operational decisions can alter the suitability of a reservoir geology for hydrogen storage. Here, we employ a numerical model of flow and transport to evaluate a set of operational strategies in varying geological scenarios for depleted methane gas reservoirs of the Bunter Sandstone, an important formation in the North Sea. We investigate the following strategies for their impact on performance and interaction with geological features that are common in the Bunter sandstone: depletion level, injected hydrogen mass, cushion gas, well perforation, idle period, production rates, and methane reinjection. We found that depletion level, injected mass, and well perforation interact strongly with geology and are critical for site selection. The methane reinjection strategy provides pressure support that increases hydrogen production, though at the cost of purity in the long-term. Furthermore, cushion gas strategies show significant optimization potential but limited interaction with geology, whereas the duration of the idle period and target rates have low optimization potential. Based on these findings, we propose a site selection and development planning framework for underground hydrogen storage in depleted gas fields. The site selection phase introduces a novel screening criterion, the gravity–purity number, which integrates geological and operational considerations. The development phase provides criteria and guidelines for planning operational strategies, and establishes a hierarchy based on their optimization potential.

Received 8th April 2026
Accepted 18th June 2026

DOI: 10.1039/d6va00162a

rsc.li/esadvances

1 Introduction

Underground hydrogen storage (UHS) is an alternative for large-scale hydrogen storage, enabling seasonal balancing of supply and demand for applications such as energy systems and industrial feedstock. Numerical models of UHS [*e.g.* ref. 37 and 40] and recent pilot projects[3,11] indicate that hydrogen gas could be effectively stored in salt caverns, aquifers, or depleted gas fields to be later recovered in periods of high energy demand.[42]

The North Sea hosts hundreds of depleted gas fields, several of which are planned to be decommissioned and repurposed to support greenhouse gas reduction. Current initiatives in the region include emerging projects for the permanent capture and storage of CO_2 [*e.g.* ref. 1 and 2] and the expansion of offshore wind energy capacity.[7,9] The geological storage of hydrogen is also being assessed as an alternative to enhance energy security and provide seasonal flexibility for the expanding wind farms.[16,21]

The successful performance of a UHS operation depends on the effective containment of hydrogen within geological traps and on the recovery of the stored gas with minimal loss. Moreover, the produced gas must have a level of purity that meets the requirements of its intended applications.[6] In part, meeting these performance requirements depends on the selection of an appropriate geological formation, which must

[a]*Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands*
[b]*Shell Global Solutions International B.V., The Hague, The Netherlands*
[c]*Navier, ENPC, Institut Polytechnique de Paris, Univ Gustave Eiffel, CNRS, Marne-la-Vallée, France. E-mail: ana.loyola@univ-eiffel.fr*

ensure sufficient storage capacity and deliverability,[38,49] and provide an environment that is not prone to hydrogen loss through mechanisms such as (bio)geochemical reactions, capillary trapping, and gravity override.[18]

To support the procedure for site screening and selection, we have recently investigated the geological controls on the performance of UHS in depleted gas fields.[28] We targeted, as a case study, the Bunter Sandstone, a formation that hosts tens of depleted gas fields in the North Sea. We performed several numerical simulations of UHS in the ensemble of geological models designed by Alshakri *et al.*,[4] which were extended to account for reservoir thickness, depth and dip. This model ensemble presents 32 plausible realizations (geological scenarios) for key geological features of the Bunter, considering their uncertainty. For this ensemble, the main geological controls on UHS performance were the thickness of the reservoir and of the sedimentological layers, and reservoir dip, which control the strength of buoyancy forces. The lateral continuity of low-permeability mudstones and of aeolian sandstones were also key controlling factors, because the former can act as baffles that compartmentalize flow and the latter promote easier spreading of the hydrogen at the top of the reservoir. All of these controlling geological features can either boost or mitigate gravity override, the main source of hydrogen loss observed in these models. Based on these insights, we proposed, as a site-screening criterion, a modified gravity number that incorporates geological heterogeneities.

These conclusions were drawn for a specific set of operational conditions, including a fully perforated well, a given depletion level, and operation at full storage and recovery capacity. However, the performance of an UHS operation is not dictated by geology alone, but also by the development strategy itself. The effect of operational parameters on UHS performance in aquifers and depleted oil and gas fields has been subject to several studies. The most commonly tested strategies include well placement and perforation;[16,29,34,40] the duration and target rates of cyclic operations;[5,12,22,34,48] the amount and composition of the cushion gas;[12,22,29,48] and the depletion level.[22,48]

However, these studies evaluate operational strategies in a single reservoir, whether synthetic or real, which makes their conclusions case-specific. Consequently, they do not provide generalizable engineering guidelines to optimize UHS while accounting for geological uncertainty that is captured by considering different geological scenarios. This generalization remains difficult when considering all studies collectively, partly because they often report divergent effects for certain operational parameters, such as the best cushion gas composition or the impact of higher flow rates.[22,48] These discrepancies exist because each study considers different residual fluids, operations and geology.

On the other hand, there is agreement regarding the impact of some strategies on the performance of UHS. For example, there are multiple reports on the improvement of UHS performance when the well is placed updip[34,40] and with shorter idle periods.[22,48] Different studies also report higher purity when using hydrogen as cushion gas.[29,48] Still, even these cases require a more systematic analysis to translate these findings into practical engineering guidelines. Such an analysis would benefit from identifying which strategies have the greatest potential impact on performance and should therefore be prioritized in development planning.

Furthermore, reservoir-specific findings provide a limited understanding of the interactions between geology and operational strategy. For optimization of development, it is useful to determine which operational choices are best suited to a specific geological scenario and geological heterogeneity. For site screening, it is valuable to understand if optimal operational choices can compensate for geological scenarios that are more susceptible to hydrogen losses; or, conversely, if poor operational decisions can undermine the potential of a geological formation with good reservoir properties that should lead to high performance.

This work aims to address these gaps. To do so, we use numerical models of flow and transport to simulate different UHS operational strategies for the ensemble of geological models of the Bunter sandstone designed by Alshakri *et al.*[4] and extended by ref. 28. We consider eight distinct operational tests, each modifying an operational strategy that we adopted previously.[28] These operational strategies include the depletion level, the amount of stored hydrogen, the amount and composition of the cushion gas, the duration of the idle period, the target injection and withdrawal rates, the length of the well perforation, and a two-well strategy for reinjecting the co-produced methane. To the best of our knowledge, this combination of parameters is investigated here for the first time.

Using this approach, this study pursues the following objectives, each providing a novel contribution to the topic of development planning and site screening for UHS in depleted gas fields: (i) evaluate the impact on UHS performance of different operational strategies in varying geological scenarios; (ii) assess the sensitivity of these impacts to geology; (iii) update our earlier screening criteria[28] to account for operational strategies; and (iv) provide general guidelines for the development of UHS that integrate geological aspects. These guidelines will provide criteria for development planning and establish a hierarchy of operational strategies based on their potential influence on performance.

This paper is structured as follows. Section 2 presents the computational models, including the geological model ensemble and the numerical model. Section 3 recapitulates the main results and findings of the tests performed by Loyola *et al.*[28] to assess the geological controls on UHS, here referred to as the reference simulations. Section 4 presents and discusses the results of the operational tests. Section 5 uses these results to propose updated screening criteria and a framework for development planning of UHS in depleted gas fields. Finally, Section 6 presents the conclusions and perspectives for future work.

2 Computational models

To analyze the impact of UHS operational strategies on different geological scenarios of the Bunter sandstone, we employ the open-access geological models developed by Alshakri *et al.*[4] and later modified by Loyola *et al.*[28] These models are used to simulate UHS operations with the Open Delft Advanced Terra Simulator (open-DARTS).[47]

Section 2.1 presents the ensemble of geological models, and Section 2.2 describes the open-DARTS numerical model.

2.1 Geological models

The Bunter Sandstone is a fluvial–aeolian formation of Triassic age that extends across several regions of Europe, notably the North Sea, where it hosts several gas fields. Alshakri *et al.*[4] identified six main stratigraphic bodies within the Bunter: multistorey, multilateral, and cemented fluvial sandstones; aeolian sandstones; channelized fluvial sandstones; floodplain and sabkha (silty) mudstones; and clay-rich lacustrine mudstones (Fig. 1).

These sedimentological bodies vary considerably in their volumetric proportions, thicknesses and petrophysical properties, as interpreted from seismic, well logs and core-sample data [*e.g.* ref. 11, 32, 44 and 45]. Based on an extensive review of such data, Alshakri *et al.*[4] used the Rapid Reservoir Modeling (RRM) software[20,36] to design the ensemble of 32 geological models that we use in our simulations. This ensemble was built using a factorial Design of Experiment (DOE) approach, in which selected factors are varied between low and high settings to quantify their impact on a given response. Eight key heterogeneities of the Bunter were included, mostly associated with the thickness, lateral continuity, and permeability of its main stratigraphic bodies. Each combination of these factors results in a unique geological scenario.

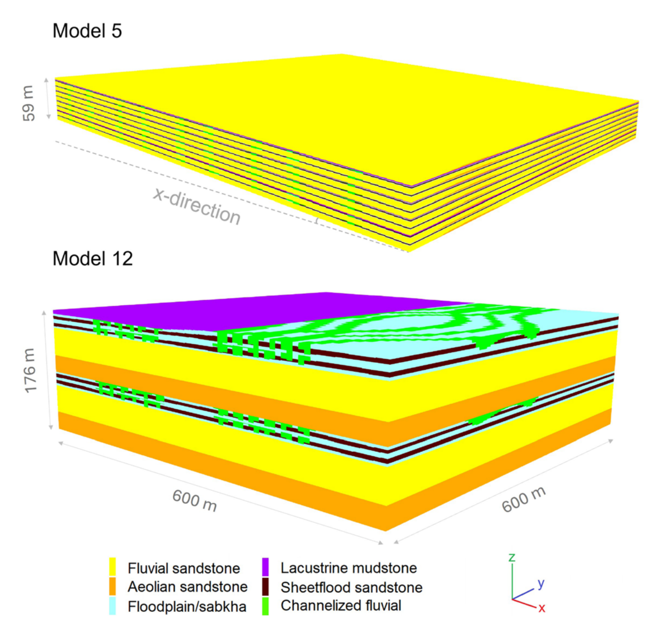

Fig. 1 Models 5 and 12 of the geological model ensemble designed by Alshakri *et al.*[4] and modified by Loyola *et al.*[28] Model 5 has low layer and reservoir thickness and a high dip, whereas Model 12 has high layer and reservoir thickness and no dip. See Table 1 for the full description of the geological factors and Table 7 for their settings.

In addition to these sedimentological properties, the Bunter also presents significant spatial variation in burial depths and formation thickness.[11] Loyola *et al.*[28] extended the DOE of Alshakri *et al.*[4] to include other factors such as reservoir thickness and structural dip, establishing low and high values for these values based on the well data of Dutch Triassic formations compiled by Korevaar *et al.*[25] Table 1 lists all the factors varied in the ensemble, while Table 7 in Appendix A presents the extended DOE. The rationale for the low and high settings of these factors is detailed in Alshakri *et al.*[4] and in the SI of Loyola *et al.*[28]

The resulting models have dimensions of 600 m × 600 m in the *x*–*y* directions and a thickness of 59 m or 176 m in the *z*-direction, depending on it assuming a high or low setting in the DOE. These lateral dimensions of 600 m × 600 m are compatible with the dimensions of small gas fields in the North Sea that are considered for UHS.

Fig. 1 shows Models 5 and 12 as two representative examples from the model ensemble. The reservoir depth is fixed at 1500 m in all cases. We have previously simulated the model ensemble considering reservoir depths of 1500 m and 3000, and concluded that higher depths slightly decrease UHS performance metrics due to lower gas compressibility and higher density contrasts between H_2 and CH_4. For further details on this comparison, the reader is referred to 28.

2.2 Numerical model

This section describes the numerical model implemented in the open-DARTS framework and the set of simulations performed. Section 2.2.1 presents the main assumptions and the mathematical formulation of the model, and Sections 2.2.2 to 2.2.4 describe the setup of the model, including the initial conditions, fluid and rock properties, mesh, boundary conditions, and the well model.

The ensemble of geological models is used to perform several operational tests, each corresponding to a set of simulations that assess the impact of a different operational strategy on the UHS performance metrics for the model ensemble. These operational tests are compared to a set of simulations referred to as the reference simulations, which correspond to the simulations performed by Loyola *et al.*[28] Section 2.2.5 presents the operational parameters used in the reference simulations, Section 2.2.6 describes the operational tests, and Section 2.2.7 introduces the performance metrics used for comparison.

2.2.1 Mathematical description. The reservoirs are assumed to be depleted methane (CH_4) gas reservoirs with no aquifer connection. We solve the problem of injection and production of hydrogen (H_2) in these reservoirs with a one-phase (gas), two-component isothermal flow and transport model. The mass conservation equation is described as

$$\frac{\partial(\phi x_c \rho_p)}{\partial t} + \nabla\left(x_c \rho_g \boldsymbol{u}_g + \phi \rho_p \boldsymbol{J}_c\right) - q_c = 0. \quad (1)$$

Here, the subscript c denotes a component (H_2 and CH_4), the subscript g refers to the gas phase, ϕ is the effective rock

Table 1 Factors analyzed in the Design of Experiments (DOE) and their low and high settings. Factors A to F were used by Alshakri *et al.*[4] to build the ensemble of 32 geological models. The remaining factors were added by Loyola *et al.*[28] to extend their DOE. In factor *G*, k_v and k_h denote the vertical and horizontal permeability, respectively. Modified from Loyola *et al.*[28]

Factor id	Factor		Low[a]	High[a]
A	Thickness of layers	Fluvial	3.4/10 m	13.3/40 m
		Floodplain/sabkha	1.7/5 m	6.7/20 m
		Aeolian	1.7/5 m	8.3/25 m
		Lacustrine	0.7/2 m	1.7/5 m
B	Lateral continuity of aeolian bodies		Discontinuous lens	Continuous sheet
C	Lateral continuity of lacustrine layers		Discontinuous sheet	Continuous sheet
D	Proportion of channelized fluvial sandbodies in floodplain/sabkha		17%	33%
E	Lateral connectivity of channelized fluvial sandbodies in floodplain/sabkha		Isolated clusters	Connected network
F	Lateral continuity of sheetflood sandbodies in floodplain/sabkha		Discontinuous lens	Continuous sheet
G	k_v/k_h ratio of fluvial sandbodies		0.05	0.5
H	Permeability of lacustrine layers		0.001 mD	0.01 mD
I	Reservoir thickness		59 m	176 m
J	Reservoir dip		0°	3°

[a] The two values separated by bars correspond to the thickness of layers when the reservoir thickness is low and thigh, respectively.

porosity, x_c is the molar concentration of component c, ρ_g is the molar density of the gas phase, $\boldsymbol{u}_g$ is the velocity of the gas phase, $\boldsymbol{J}_c$ is diffusive flux of component c and q_c denotes source of the component c.

The phase velocity $\boldsymbol{u}_g$ is given by Darcy's law as

$$\boldsymbol{u}_{\rm g} = -\frac{\boldsymbol{K}}{\mu_{\rm g}}\left(\nabla p_{\rm g} - \rho_{\rm g}\boldsymbol{g}\right), \tag{2}$$

where $\boldsymbol{K}$ is the permeability tensor, μ_g the viscosity of the gas phase and $\boldsymbol{g}$ is the gravity vector. The diffusive flux $\mathbf{J}_c$ is given by Fick's law as

$$\boldsymbol{J}_{\rm c} = -D_{\rm c}\nabla x_{\rm c}, \tag{3}$$

where D_c is the molecular diffusion coefficient of component c.

The open-DARTS model employs a cell-centered Finite Volume Two Point Flux approximation scheme for the spatial discretization of both fluxes and a fully implicit scheme for the time discretization.[30] All simulations were performed on the DelftBlue supercomputer.[8]

Note that depleted gas fields with no significant aquifer contribution are still expected to contain a residual amount of water. Although this water is immobile, it can give rise to capillary and solubility trapping of H_2 and influence its mobility through interfacial phenomena. These effects are neglected in our single-phase flow formulation.

2.2.2 Initial reservoir conditions. The reservoir temperature is set to 55 °C, computed for the depth of 1500 m using the temperature-depth relationship derived for the borehole data from the Netherlands Oil and Gas Portal (NLOG)[33] (see the SI of ref. 28). For a depth of 1500 m, we estimate a hydrostatic pressure of 150 bar and assume it to be the original reservoir pressure before depletion. This hydrostatic pressure is used to define an initial pressure at the start of the UHS operation, based on a specified depletion level, to be defined in Sections 3 and 4.

2.2.3 Rock and fluid properties. Table 2 presents the porosity and the intrinsic permeability of the sedimentological layers as defined by Alshakri *et al.*[4] and Loyola *et al.*[28] The permeability is considered to be isotropic for all sedimentological layers except the fluvial sandstones. For these latter, Table 2 gives the horizontal permeability, and the vertical permeability is computed according to the given permeability ratio, which is a factor in the DOE (Table 1). The permeability values in Table 2 are consistent with the lower-end of the permeability ranges reported by Foote *et al.*,[13] Fig. 9 for the considered sedimentological layers.

The gas density is computed using the Peng–Robinson equation of state. The gas viscosity is estimated using the datasets from Hassanpouryouzband *et al.*[17] and the NIST Chemistry WebBook,[27] which provide viscosity values for various combinations of pressure, temperature, and gas composition. We use the multilinear interpolation scheme implemented in open-DARTS[23,31] to interpolate viscosity at arbitrary states. For further details, see Fig. S3 in ref. 28.

The molecular diffusion coefficient is assumed constant and equal to 0.078 m^2 per day, as obtained from the correlation proposed by Kobeissi *et al.*[24] The pressure input for this

Table 2 Intrinsic permeability and porosity of the sedimentological bodies present in the Bunter sandstone models

Geobody	Porosity (%)	Permeability (mD)
Lacustrine mudstone	5	0.001 or 0.01[a]
Floodplain and sabkha mudstone	7	1
Sheetflood fluvial sandstone	7	1
Fluvial sandstone[b]	17	500
Aeolian sandstone	24	3000

[a] Lacustrine mudstone permeability is a factor in the DOE. [b] Permeability in horizontal directions.

Table 3 Operational parameters used for the reference simulations. The target rates and well perforation differ between the low- and high-reservoir-thickness models. The well perforation corresponds to the reservoir thickness in each case, and the target rates are scaled accordingly

Initial pressure (bar)	50	
Minimum/maximum BHP (bar)	75/150	
	Low res. thick	High res. thick
Target rate – initialization (Mmol per day)	35	105
Target rate – injection (Mmol per day)	52	156
Target rate – production (Mmol per day)	104	312
Well perforation (m)	59	176

correlation corresponds to the average of the minimum and maximum borehole pressures (Table 3).

2.2.4 Mesh, boundary conditions, and wells. The finite volume grid is fully structured, containing 71 × 71 × 132 cells for models with low reservoir thickness and 71 × 71 × 396 cells for those with high reservoir thickness. The grid cell size is the same across all models of the ensemble. This resolution was defined by Loyola *et al.*[28] based on mesh convergence studies. In models with high dip, the dip is implemented by translating grid columns in a stepwise manner along the x-direction, so that the highest elevation is reached at the minimum x-coordinate (Fig. 2 in ref. 28).

All reservoir boundaries are closed (no-flow). These boundary conditions imply negligible leakage of H_2 through the caprock and other sealing layers. The validity of this assumption should be assessed for target reservoirs based on an evaluation of the sealing properties of the cap rock and overburden, including the role of potential faults.

UHS operations are simulated in a single vertical well with a diameter of 0.16 m. The well is positioned at the coordinates (13 m, 300 m) in the x–y plane, so that it perforates cells located at center of the y-axis and close to the minimum x-coordinate. In models with high dip, this positioning ensures that the well is located structurally updip.

H_2 injection and production are modeled by defining target flow rates at the well (source terms). The flow transfer between the well and the reservoir is modeled following the formulation by Peaceman,[35] with no skin factor considered. When the bottom-hole pressure (BHP) reaches its upper or lower limit, the flow rate is adjusted to ensure that the respective pressure constraints are not violated.

2.2.5 Operational parameters in the reference simulations. The UHS operations of the reference simulations consist of two phases. The first phase is an initialization period during which pure H_2 is injected as cushion gas. This period ends when the minimum BHP is reached in the well. Following the initialization, the cyclic storage operation starts. In the reference simulations, each cycle lasts one year and comprises a six-months injection period followed by three-months idle and production periods. The idle periods correspond to a shut-in period where zero flow rates are assigned to the well. The minimum and maximum BHPs are 75 bar and 150 bar, respectively.

Table 3 summarizes the operational parameters in the reference simulations, including well perforation, initial reservoir pressure, BHP limits, and target rates. The well is fully perforated throughout the thickness of the reservoir, and the depletion level is set to two-thirds of the hydrostatic pressure, resulting in an initial pressure of 50 bar.

The target rates in Table 3 were defined to ensure that the reservoirs operate at their full storage and recovery capacities, constrained by the minimum and maximum BHP. The target rates for the models with high reservoir thickness were three times those for the models with low reservoir thickness. This scaling ensures that all models operate at comparable flow rates, H_2 pore volume occupation, and pressure ranges, allowing for a reliable comparison across the different geological scenarios.[28]

2.2.6 Operational tests. Table 4 presents the eight operational tests performed in this study. These tests modify the operational parameters of the reference simulations. If any change in an operational parameter is not explicitly stated, that parameter remains identical to the reference simulations.

Test 1 (depletion level) assesses the impact of increasing the depletion level of the reservoirs to 90% of their original hydrostatic pressure. This results in an initial reservoir pressure (at the start of the operation) of 15 bar. The minimum BHP is reduced to 40 bar. Since, in the initialization phase, the H_2 cushion gas is injected until the minimum bottomhole pressure is reached, this ensures additional pressure support of approximately 25 bar provided by the cushion gas, similarly to the reference simulations. This change in the minimum BHP is based on the assumption that lower working pressures do not compromise reservoir containment or the deliverability of the

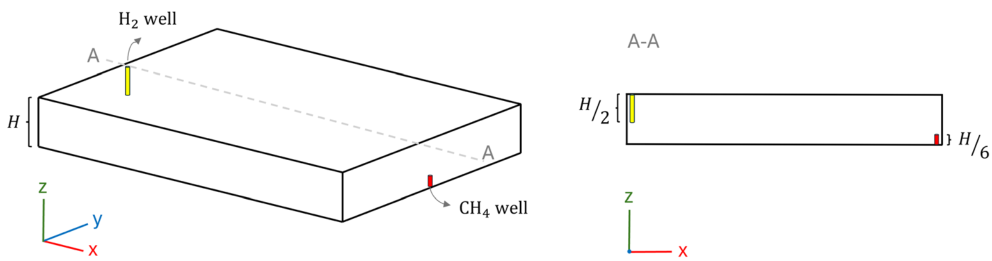

Fig. 2 Configuration of the wells in Test 6 (reinjection of CH_4).

Table 4 The performed operational tests

Test id	Test name (%)	Description
1	Depletion level	Initial reservoir pressure is set to 10% of the hydrostatic pressure Initial pressure = 15 bar, minimum BHP = 40 bar
2	Injected amount	Cyclic injection of H_2 is reduced by 50% Same target rates, injection and production duration are decreased by 50%
3	No cushion gas	No initialization phase Minimum BHP = 50 bar
4	Cushion gas composition	The cushion gas is a mixture of 50% CH_4 and 50% H_2
5	Well perforation	Well perforation is half of the reservoir thickness
6	Reinjection of CH_4	Co-produced CH_4 is reinjected in the reservoir
7	No idle period	No idle period in the cyclic operation
8	Target rate	Cycle flow rates increased by a factor of 2 Injection and production periods are decreased by 50%

well system, since such pressures were already achieved during CH_4 production.

Test 2 (injected amount) assesses the impact of reducing the amount of H_2 injected during the cyclic operation, *i.e.* the working volume of H_2. In the reference simulations, the reservoirs operate at full storage and recovery capacity, *i.e.* the well pressure varies between the minimum and maximum BHPs during the cyclic operation. However, UHS may involve reservoirs whose storage capacity exceeds the projected storage demand. To investigate an operation below storage capacity, Test 2 reduces the target mass of H_2 injected during the cyclic operation by half. The target mass rates are kept unchanged, while the duration of the injection and production periods are reduced by half. The amount of injected H_2 during the initialization phase remains unchanged.

Test 3 (no H_2 cushion gas) omits the initialization phase, starting the storage operation directly with the injection–production cycles. We assume that the decision to not use cushion gas implies that no pressure support is required to ensure well deliverability or safe containment. Consequently, the minimum bottomhole pressure is set equal to the initial reservoir pressure of 50 bar.

Test 4 changes the composition of the cushion gas injected during initialization, using a mixture of 50% H_2 and 50% CH_4, in terms of molar fraction.

Test 5 (well perforation) reduces the length of the perforated interval to half of the reservoir thickness. This reduced perforation is connected to all sedimentological layers present on the top half of the reservoir (see the H_2 well in Fig. 2).

Test 6 (reinjection of CH_4) implements a strategy for reinjecting the co-produced CH_4. This reinjection may be required due to the lack of an alternative to handle the co-produced gas, for environmental compliance, or for increasing reservoir pressure. To promote the separation of H_2 and CH_4, the well configuration shown in Fig. 2 is used. The H_2 well remains at the same *x*–*y* coordinates and is perforated over half of the reservoir thickness, as in Test 5 (well perforation). A second well (CH_4 well) is placed on the opposite side of the reservoir and perforated until its highest depth. This second well is used for reinjecting the CH_4 produced from the H_2 well. The length of the perforation of the CH_4 well is equal to $\frac{1}{6}$ of the reservoir thickness. This length was defined to avoid pressure interference with the H_2 well. During the production period, CH_4 co-produced at the H_2 well is accumulated for 10 days and reinjected in the CH_4 well over periods of 1 to 2 days, depending on the amount of CH_4 and the target rates. The period of 10 days was set to ensure that the target rates for reinjection of CH_4 are high-enough to avoid numerical convergence issues in the simulation.

Test 7 (no idle period) does not have an idle period. While the reference simulations have a three-months idle period between the injection and recovery phases, the cyclic operation in Test 7 alternates between a six-months injection and three-months production.

Test 8 (target rates) doubles the target rates of the injection and production periods, while reducing their durations by half so that the total amount of injected H_2 remains equal to that of the reference case.

Although the target rates and the duration of the idle period are tested here as operational strategies, they are not design parameters *per se*, as they cannot be flexibly planned in UHS operations. Rather, they are determined by the schedules of H_2 production and demand arising from the use case that justifies the need for a H_2 storage solution. Thus, they are tested here to assess whether external variables that determine H_2 production and demand may affect the performance of UHS operations.

2.2.7 Performance metrics. Four performance metrics are computed in the simulations to quantify the efficiency of an operational strategy: cycle recovery factor (RF), ultimate recovery factor (ultimate RF), cycle mass purity (MP) and flow purity (FP). The recovery factor (RF) of cycle i is given by

$$\mathrm{RF}^i = \frac{M_{\mathrm{prodH_2}}{}^i}{M_{\mathrm{injH_2}}{}^i}, \tag{4}$$

where $M_{\mathrm{prodH_2}}{}^i$ and $M_{\mathrm{injH_2}}i$ are the produced and injected mass of hydrogen, respectively, in cycle i, respectively. The ultimate RF is computed as

$$\mathrm{RF} = \frac{\sum_i M_{\mathrm{prodH_2}}{}^i}{\sum_i M_{\mathrm{injH_2}}{}^i}, \tag{5}$$

Mass purity (MP) of cycle i is given by

$$\mathrm{MP}^i = \frac{M_{\mathrm{prodH_2}}{}^i}{M_{\mathrm{rec}}{}^i}, \tag{6}$$

where $M_{\mathrm{prodH_2}}{}^i$ and $M_{\mathrm{rec}}{}^i$ are the produced mass of H_2 and total produced mass of fluids in cycle i, respectively. The flow purity FP at time t of the withdrawal phase of cycle i is given by

$$\mathrm{FP}^i(t) = \frac{Q_{\mathrm{H_2}}{}^i(t)}{Q_t{}^i(t)}, \tag{7}$$

where $Q_{\mathrm{H_2}}{}^i$ and $Q_t{}^i$ are outflow rates of hydrogen and total outflow rates in time t of the production phase of cycle i, respectively.

Throughout this text, when we refer to UHS performance we refer specifically to these metrics that measure H_2 recovery and purity.

3 Summary of the reference simulations

We present here a summary of the results of the reference simulations[28] that are most relevant to the present study, as well as their interpretation.

The reference simulations resulted in a wide range of performance metrics for the model ensemble. The variability of these metrics is highest in cycle 1 and decreases as the operation progresses (Fig. 3). This variability is due to the geological heterogeneities captured by the ensemble. As the injection–production cycles increase, these heterogeneities become less influential because CH_4 initially in place is progressively replaced by H_2 and there is sufficient time for H_2 to enter the lower-permeability layers.

We frequently refer to the performance in cycle 1 because this is the most critical stage of the operation—where the performance metrics of each model are at their minimum, and their variability is highest. From here on, we denote the recovery factor and methane purity of cycle 1 as RF_1 and MP_1, respectively.

In the reference simulations, RF_1 range from 70% to 96%, MP_1 from 71% to 96%, and the ultimate RF from 89% to 96%. Models 5 and 12 (Fig. 1) exhibited the highest and lowest values for these metrics, respectively. These two models are representative of the geological factors that control UHS performance. Loyola *et al.*[28] identified the thickness of the sedimentological layers, reservoir thickness, and dip as the three most influential factors affecting the performance metrics for cycle 1. Model 5 has low layer thickness, low reservoir thickness, and high dip, whereas Model 12 has high layer and reservoir thickness and no dip.

The geological controls identified by Loyola *et al.*[28] influence UHS performance through their effect on gravity-driven mechanisms. The displacement of CH_4 by H_2 naturally carries a risk of unstable displacement due to the significantly lower viscosity and density of H_2. Here, the terms unstable displacement or unstable fronts refer specifically to gravity override, in which H_2, driven by buoyancy forces, migrates upward and spreads laterally along the reservoir top. This instability is directly related to H_2 loss and CH_4 production, since the CH_4 that is not efficiently displaced can reach the well and be produced.

The degree of instability, however, is strongly controlled by the geological scenario. This is illustrated by comparing Fig. 4 and 5, which show cross-sectional views of the H_2 plume at different stages during cycle 1 for Models 5 and 12, respectively.

In Model 5, the H_2 plume develops a relatively stable front with the resident CH_4, approaching a piston-like displacement during the initialization phase. During the cyclic operation, some gravity override occurs away from the well, but viscous forces driving longitudinal migration outweigh buoyancy forces through most of the reservoir. The buoyancy forces are weak in this model because the sedimentological layers are thin. In addition, the continuous lacustrine mudstones in Model 5 further mitigate buoyancy forces by compartmentalizing flow,

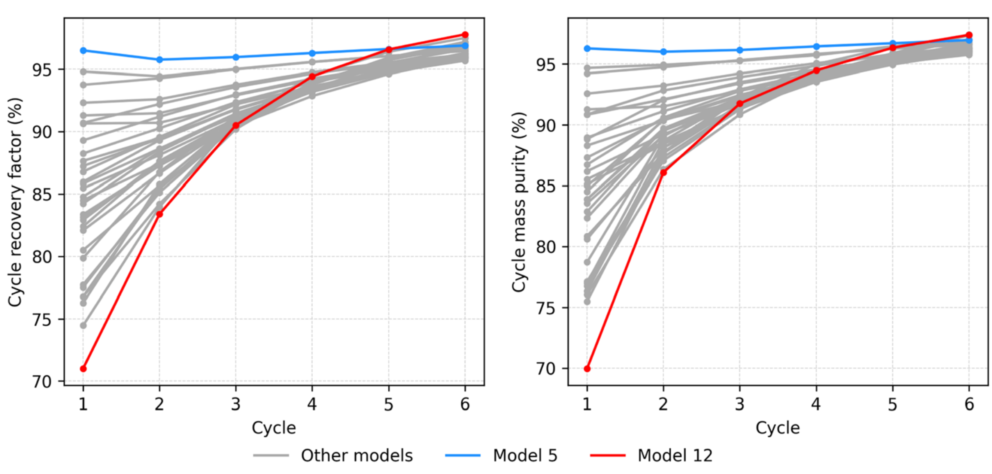

Fig. 3 Cycle recovery factor and mass purity observed in the reference tests. Models 5 and 12 (Fig. 1) are highlighted with distinct colors, as they presented the highest and lowest performance metrics in cycle 1, respectively. Results for the other models are shown in grey to represent the ensemble as a whole.

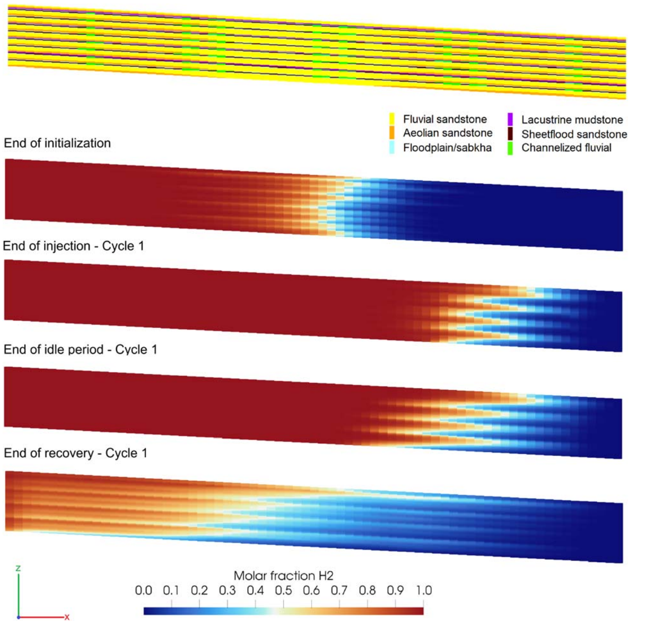

Fig. 4 Cross-sectional view of the sedimentological layers and the H_2 plume in the reference simulation of Model 5 during different stages of cycle 1.

and the high dip helps to keep CH_4 segregated toward the bottom of the reservoir.

In contrast, Model 12 exhibits pronounced gravity override already during the initialization phase, and this instability develops much closer to the well. This behavior results from stronger buoyancy forces, which are proportional to the thickness of the depositional cycles (*i.e.*, the repeating pattern of sandstone and mudstone couplets). The occurrence of gravity override is also favored by the zero dip in this model. As H_2 spreads along the reservoir top, CH_4 rapidly approaches the well and is produced, leading to the low performance metrics observed in this model.

In the absence of significant viscous forces, the interface between fluids of contrasting density tends toward a hydrostatic equilibrium configuration, leading to a horizontally segregated front. Model 12 approaches this limiting condition by the end of the idle period, rendering the well, which is perforated across the entire reservoir thickness, particularly exposed to CH_4 during production.

3.1 Modified gravity number

Based on the gravity-controlled behavior of the reference tests, Loyola *et al.*[28] adapted the gravity number derived by Shook *et al.*[41] to consider the key geological heterogeneities of the Bunter sandstone for UHS. This modified gravity number is given by

$$N_g = \frac{k_x \lambda \Delta\rho g \cos\alpha}{u}\frac{H}{L}, \tag{8}$$

where k_x is the upscaled horizontal permeability of the reservoir (here computed with flow based upscaling in RRM), λ is the mobility of the injected phase, $\Delta\rho$ is the density difference between the two fluids, g is the gravitational acceleration, α is the dip angle, u is the velocity of the injected fluid, H is the thickness of the depositional cycle bounded by mudstone layers, and L is the length of the reservoir. The modifications proposed by Loyola *et al.*[28] include the use of the upscaled permeability and the replacement of the thickness of the reservoir with the thickness of the depositional cycles.

The modified N_g correlates strongly with the performance metrics computed for the Bunter sandstone model ensemble, especially when clustering the reservoirs according to their

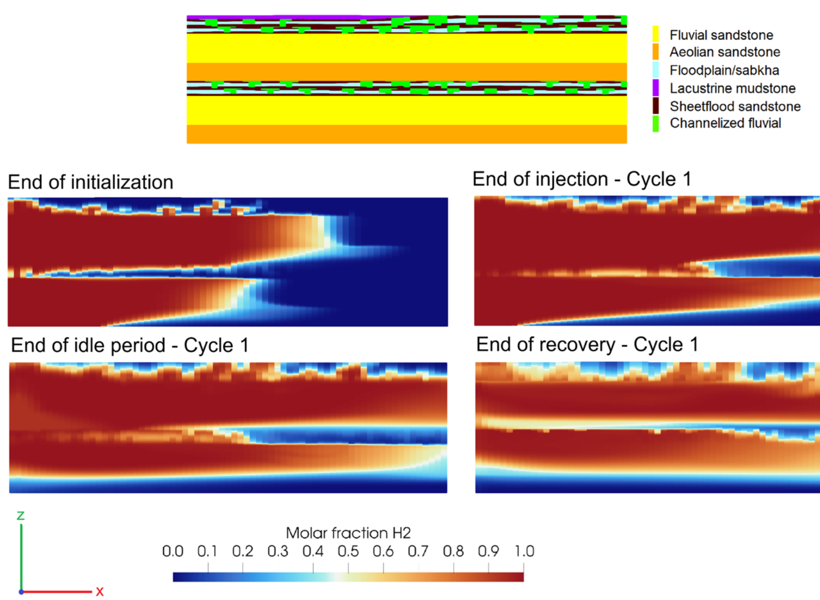

Fig. 5 Cross-sectional view of the sedimentological layers and the H_2 plume in the reference simulation of Model 12 during different stages of cycle 1.

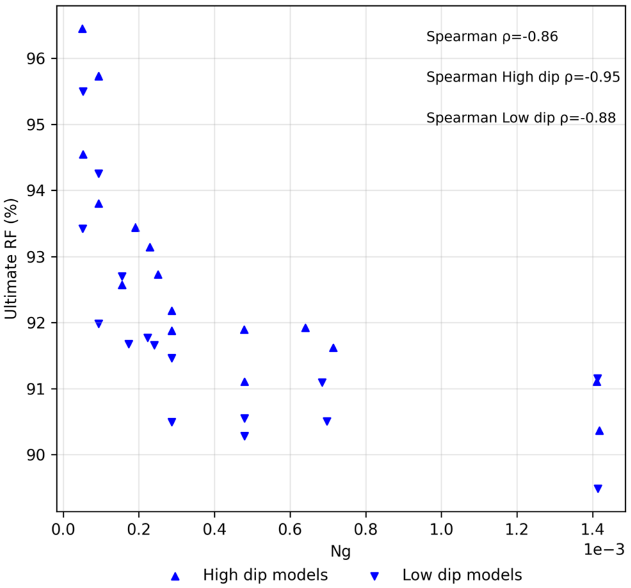

Fig. 6 Spearman correlation between the ultimate recovery factor (RF) of the reference tests and the modified gravity number (N_g) proposed by Loyola *et al.*[28] The Spearman coefficient is presented for all the models, low-dip models and high-dip models.

dip.[28] In addition to the results presented by Loyola *et al.*[28] for RF_1, we present here the correlation between N_g and the ultimate RF of the reference test, measured by the Spearman coefficient (Fig. 6).

4 Results of the operational tests

Fig. 7 shows the range of RF_1, MP_1, and ultimate RF for the reference models and the operational tests. Further detail on the evolution of performance in the operational tests can be found in Fig. S1 to S7 of the SI. Fig. S8 presents the evolution of average reservoir pressures during the UHS operations for the operational tests.

To evaluate whether the modified N_g remains correlated to the performance metrics in each operational test, Table 5 shows the Spearman coefficients for the ultimate RF and N_g. In the Supplementary Material, Fig. S9 to S11 present the correlation of N_g with RF_1, MP_1 and the ultimate RF. Except for Test 5 (well perforation) and Test 6 (reinjection of CH_4), the correlation between N_g and the performance metrics is negative and strong in all tests, with absolute values greater than 0.8. Therefore, in all tests except those two, higher performance metrics generally correspond to lower N_g values and *vice versa*.

Table 5 Spearman coefficients between ultimate recovery factor and the modified gravity number for the reference and operational tests

Test	Low dip	High dip
Reference	−0.88	−0.95
1 – depletion level	−0.84	−0.84
2 – injected amount	−0.87	−0.88
3 – No cushion gas	−0.83	−0.93
4 – Cushion gas comp	−0.95	−0.86
5 – Well perforation	0.40	0.41
6 – Reinjection of CH_4	−0.47	−0.12
7 – No idle	−0.96	−0.87
8 – Target rate	−0.95	−0.87

To analyze the impact of operational strategies under different geological conditions, the model ensemble is divided into two groups based on N_g. The lowest N_g group includes the eight models with the lowest N_g values (bottom quartile), while the highest N_g group includes the eight models with the highest N_g values (top quartile). Table 6 presents the average values for RF_1, MP_1 and ultimate RF for each group. It also presents the average value for $G_g = k_x H \cos(\alpha)$, which is the part of N_g that includes reservoir characteristics, namely the upscaled permeability k_x, the dip angle α and the thickness of the depositional cycle H. Moreover, Fig. 8 shows the average differences in the performance metrics between the operational tests and the reference cases, calculated for the highest N_g, lowest N_g, and the entire ensemble of 32 models. All highest N_g models have high thickness of the reservoir and of the sedimentological layers, while all lowest N_g models have low setting for both factors in the DOE.

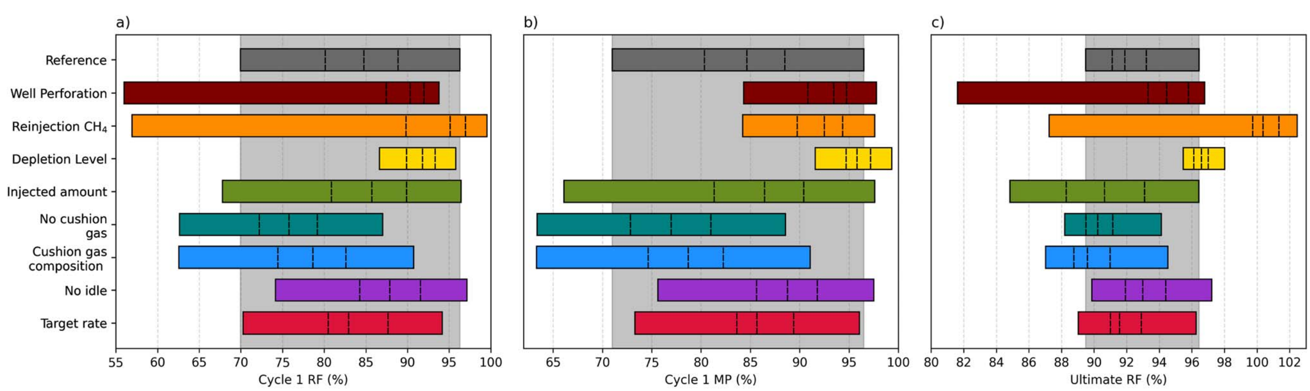

Fig. 7 Range of variation of (a) ultimate recovery factor (RF), (b) RF of cycle 1 and (c) the mass purity (MP) of cycle 1 in the reference simulations and operational tests. The left, middle, and right dashed lines crossing each bar correspond to the first, second and third quartiles of the data, respectively.

Table 6 Average geometrical component of the gravity number (G_g), recovery factor of cycle 1 (RF_1), mass purity of cycle 1 (MP_1), and ultimate recovery factor (ultimate RF) for the highest and lowest N_g groups, which include the eight models with the highest and lowest N_g values, respectively

Group	Models	G_g (m^3)	RF_1 (%)	MP_1	Ultimate RF
Highest N_g	1, 3, 5, 7, 25, 27, 29, 31	71 312	78	78	90
Lowest N_g	10, 12, 14, 16, 18, 20, 22, 24	4925	92	92	94

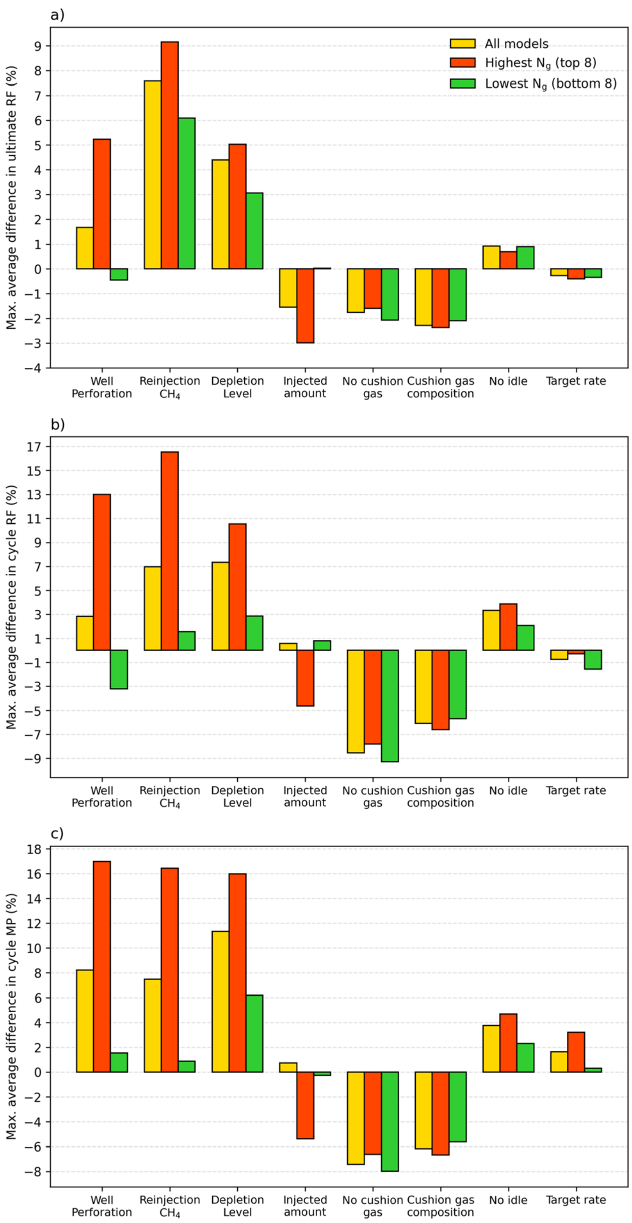

Fig. 8 Average differences in the (a) ultimate recovery factor, (b) cycle recovery factor (RF) and (c) cycle mass purity (MP) between each operational test and the reference case as computed for the whole model ensemble, the lowest N_g models and the highest N_g models. The differences in cycle RF and cycle MP correspond to the cycle that presented the maximum average differences, which is cycle 2 for Test 2 (injected amount of H_2) test and cycle 1 for the other operational tests.

The next sections analyze and discuss these results for each operational test, with focus on: (i) the variations in the range of performance metrics induced by different operational strategies; (ii) the main flow mechanisms driving these variations; (iii) the interaction between operational strategies and geology; and (iv) comparisons with findings from previous studies when pertinent.

4.1 Test 1—depletion level

Compared to the reference simulation, the higher depletion level of 90% significantly improves the performance across the model ensemble. This improvement is most evident in the minimum values of the performance metrics (Fig. 7), which generally correspond to the highest N_g values (Table 5). The minimum RF_1 increases from 70% to 87%, the minimum MP_1 from 71% to 92%, and the minimum ultimate RF from 89% to 96%. The maximum RF_1 remains at 96%, the maximum MP_1 increases from 97% to 99%, and the maximum ultimate RF increases from 96% to 98%. The high minimum values observed for a higher depletion level indicate that less favorable geological scenarios, as classified by high N_g, can still achieve good performance for UHS.

The benefit of higher depletion levels is explained by two key factors. First, a lower initial amount of CH_4 reduces the amount of impurity that can be produced and increases the volumetric content of H_2 in the reservoir. By the end of the first cycle injection, on average, 89% of the reservoir pore volume (PV) is occupied by H_2, compared to 69% in the reference simulations. Second, lower reservoir pressures (Fig. S8) result in smaller density contrasts between H_2 and CH_4, which mitigates buoyancy forces and thereby reduces gravity override.

The lowest N_g models are less affected by the depletion level than the highest N_g models (Fig. 8); they have high average performance metrics, above 90%, for both the depletion levels of 67% and 90%. These models are less sensitive to the initial CH_4 content because their geological scenarios prevent CH_4 from reaching the well, since their thin sedimentological layers, and sometimes high dip, mitigate gravity override.

The impact of the depletion level on the performance of UHS operations has also been investigated by Kanaani *et al.*[22] and Zamehrian and Sedaee[48] for a depleted oil reservoir and a depleted gas condensate reservoir, respectively. Both studies report two competing effects of higher depletion levels: first, an increased recoverable amount of the formation fluid, which tends to decrease H_2 recovery; and second, higher initial reservoir pressures, which are beneficial to H_2 recovery in their case studies.

These competing effects led to complex dependencies of performance on the depletion level. For example, Zamehrian and Sedaee[48] observed an increase in RF when the depletion level increases from 30% to 60%; however, the highest tested depletion level of 70% resulted in the lowest RF because it did not yield sufficient reservoir pressure to produce gas at the target flow rates. This behavior probably occurs because the minimum BHP was the same in all scenarios; therefore, at higher depletion levels, part of the injected H_2 in early cycles is trapped for pressure support, ultimately working as a cushion gas instead of recoverable H_2.[29]

In contrast, we reduce the minimum BHP and inject sufficient cushion gas to maintain this minimum working pressure before the cyclic operation. This comes from the assumption that lower depletion levels allow for lower minimum BHPs, since these pressures have already been safely reached before the UHS operations during production of CH_4. Because of this assumption, our model ensemble benefits from the lower amounts of CH_4 in the reservoir without having the H_2 deliverability reduced by shorter working pressure ranges.

4.2 Test 2—injected amount of H_2

Similarly to the depletion level, decreasing the amount of injected working gas affects mostly the minimum values of the performance metrics (Fig. 7). Compared to the reference simulations, the minimum RF_1 decreases from 70% to 68%, the minimum cycle MP_1 from 71% to 66% and the minimum ultimate RF from 89% to 85%. The maximum values for RF_1 and the ultimate RF remains largely constant, while the maximum value of MP_1 only increases by 1%.

The decrease in the injected mass of H_2 reduces the proportional amount of H_2 in the reservoir, which tends to decrease RF and purity. By the end of the first cycle injection, on average, 58% of the reservoir PV is occupied by H_2, against 69% for the reference simulations. The highest N_g models are more sensitive to this effect. They show an average decrease of 5% in RF_1, 6% in MP_1 and 3% in the ultimate RF. In contrast, the lowest N_g models show a maximum average decrease of 1% in their performance metrics (Fig. 8). Therefore, in geologies that intensify gravity override, operating at high storage capacity is more important. In contrast, geological scenarios where gravity override is less prevalent can achieve good performance with lower H_2 injection.

4.3 Test 3—no H_2 cushion gas

The absence of the H_2 initialization phase leads to a decrease in performance metrics for the entire model ensemble (Fig. 7). Both the minimum and maximum values of the performance metrics decrease similarly compared to the reference test. The minimum RF_1 decreases from 70% to 63%, the minimum MP_1 from 71% to 63%, and the minimum ultimate RF from 90% to 88%. The maximum RF_1 decreases from 96% to 87%, the maximum MP_1 from 97% to 89%, and the maximum ultimate RF from 96% to 94%.

The injection of H_2 cushion gas enhances the purity of the produced gas stream through two main mechanisms. The first mechanism is the increase in the volumetric content of H_2 before the cyclic operations, which ensures a minimum purity during production. H_2 occupies, on average, 64% of the reservoir PV by the end of the first cycle injection, compared to 69% in the reference cases (with cushion gas).

The second mechanism through which H_2 cushion gas enhances purity is by displacing CH_4 initially in place farther away from the well, which delays its production. The H_2 cushion gas is more effective in displacing and segregating CH_4 in low-N_g models, where it creates stable fronts. Fig. 4 and 5 show, for the reference simulations, the H_2 plumes after initialization for Model 5 (lowest N_g) and Model 12 (highest N_g), respectively. The more effective displacement of CH_4 in Model 5 results in a later decline in purity and a higher minimum FP of 89%, compared to 62% for Model 12 (Fig. 9). In the absence of cushion gas, the minimum FP in Cycle 1 for both models is similar, around 53%, illustrating how the presence of cushion gas impacts the minimum purity of the produced gas.

The absence of cushion gas causes similar differences in the performance metrics for the lowest and highest N_g (Fig. 8), suggesting that the impact of the amount of cushion gas on UHS performance is largely insensitive to the geological scenario.

In addition to improving purity, the cushion gas can provide pressure support, maintaining the working pressure above the minimum required BHP. Higher minimum pressures can also reduce pressure gradients, which limits the lateral spreading of H_2. However, in our models, pressure support plays a negligible role because pressure gradients are small[28] and the minimum BHP is equal to the initial reservoir pressure. The studies that report improvement in UHS performance due to cushion gas pressure support use models where the minimum BHP is higher than the initial reservoir pressure, or the reservoirs are large and cannot be uniformly pressurized [*e.g.* ref. 22 and 29].

4.4 Test 4—cushion gas composition

Similar to Test 3, the change in cushion gas composition causes a reduction in the performance metrics for the entire model ensemble. The minimum RF_1 decreases from 70% to 63%, the minimum MP_1 from 71% to 63% and the minimum ultimate RF from to 89% to 87%. The maximum RF_1 decreases from 96% to 91%, the maximum MP_1 from 97% to 91% and the maximum ultimate RF from 96% to 95%.

The range of performance is similar to that of Test 3 (no cushion gas), indicating that using a mixture of H_2 and other component as cushion gas may yield no advantage compared to an operational strategy where there is no cushion gas (Fig. 7 and 9). This result is because in our test with no cushion gas (Test 3), the well deliverability was not affected, due to the assumption of a lower required minimum BHP. Moreover, the CH_4–H_2 mixture as cushion gas does not significantly increase the proportion of H_2 in the reservoir. By the end of the first cycle injection, H_2 occupies, on average, 64% of the PV when no cushion gas is used (Test 3) and 62% of the PV when a mixture of H_2 and CH_4 is used as cushion gas.

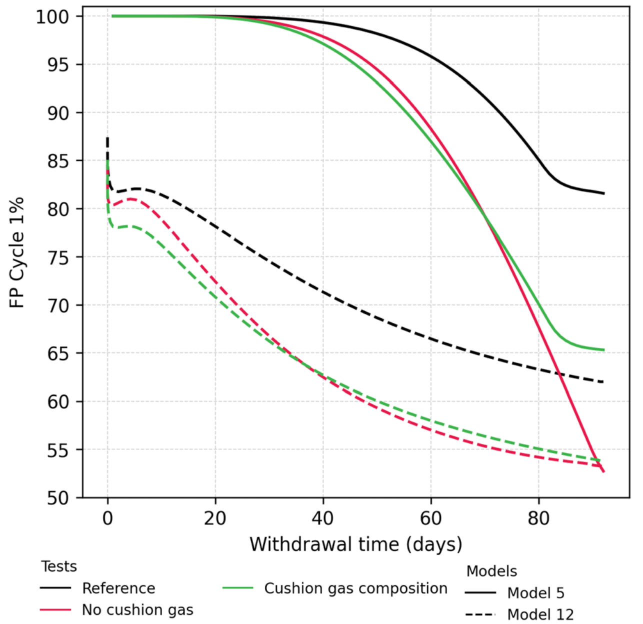

Fig. 9 Evolution of the flow purity (FP) during the production of cycle 1 in Models 5 and 12 in the reference simulations, Test 3 (no H_2 cushion has) and Test 4 (cushion gas composition).

Using a mixture of H_2 and CH_4 as cushion gas can be effective when it results in a significant increase of the H_2 fraction in the reservoir. It is also useful when pressure support is needed. This pressure support can also be achieved with cushion gases that do not contain H_2, at the expense of a decrease in purity.[22]

In the case of UHS in aquifers or in gas fields with significant water content, the cushion gas also plays an important role in displacing water. In this context, the composition of the cushion gas can further affect the efficiency of this displacement by influencing interfacial tension and wettability,[19] ultimately affecting hydrogen mobility and the strength of capillary trapping.

4.5 Test 5—well perforation

Decreasing the well perforation results in a sharp decrease in the minimum RF. The minimum RF_1 decreases from 70% to 56%, and the minimum ultimate RF from 90% to 82%. Conversely, the minimum MP_1 increases from 71% to 84% (Fig. 7).

The reduction of the well perforation is one of the only tested strategies that result in poor correlation between the performance metrics and the N_g 5. The highest N_g models experience significant performance improvement—on average, increases of 13% in RF_1, 17% in MP_1, and 5% in the ultimate RF are observed. Meanwhile, the lowest N_g models experience average decreases of 3% and 0.5% in RF_1 and the ultimate RF, respectively, and an average increase of 1% in MP_1. As a result, the highest-N_g models show better performance than most of the lowest-N_g models.

Nearly all models where RF decreases have low thickness of layers and high continuity of lacustrine mudstones in the DOE (Fig. 10). Half of the lowest-N_g models share these characteristics (see Models 5, 7, 29, and 31 in Table 7), explaining their average reduction in RF. In the model ensemble by Alshakri *et al.*,[4] models with a high setting for the thickness of layers contain only one lacustrine mudstone layer at the top of the reservoir, whereas models with low thickness of layers include several of these mudstones (see Models 5 and 12 in Fig. 1).

These frequent lacustrine mudstone layers act as baffles that vertically compartmentalize the reservoir because their permeability can be 4 to 5 orders of magnitude lower than that of the fluvial sandstones (Table 2). When reducing the well perforation, this compartmentalization causes pressure discontinuities across the lacustrine mudstones. This results in low average reservoir pressures and reduction in the performance of UHS.

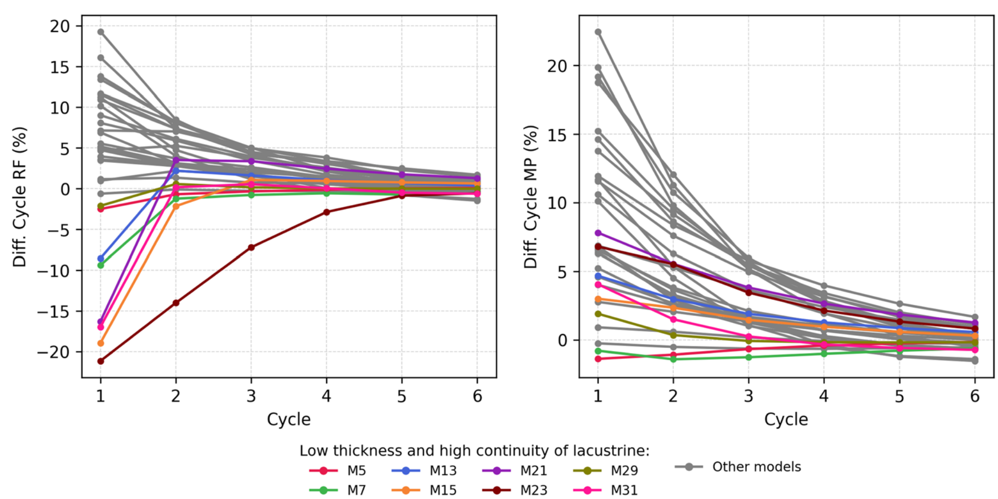

Fig. 10 Differences in the cycle recovery factor (left) and mass purity (right) between the reference tests and Test 5 (well perforation) for the model ensemble. The models with low thickness of the layers and high continuity of the lacustrine mudstone in the DOE are colored. They form a group of models where the initial recovery factor is negatively affected by the reduction in well perforation.

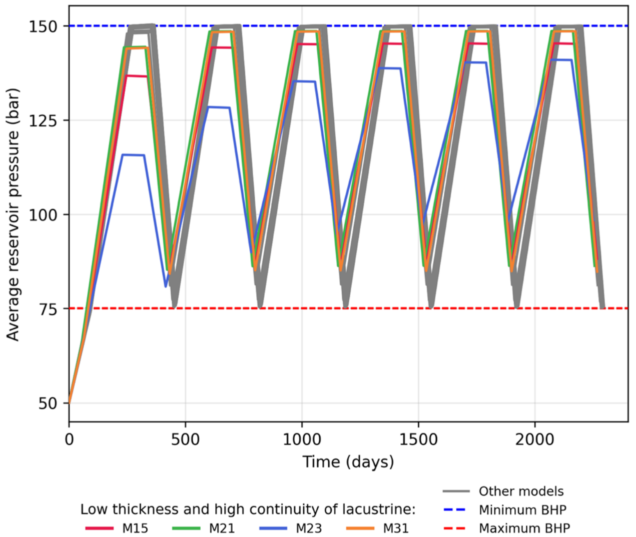

Fig. 11 Evolution of the average reservoir pressure during UHS for the entire model ensemble in Test 5 (well perforation). The models that exhibit the largest decrease in cycle RF are colored.

Fig. 11 shows that Models 15, 21, 23, and 31, which exhibit the strongest decreases in cycle RF (Fig. 10), reach lower maximum pressures and higher minimum pressures compared to the other models.

The mechanism of H_2 loss behind the distinct pressure range of these models is illustrated in Fig. 12 for model 31. The vertical reservoir compartments connected to the well experience higher pressure buildup than the deeper compartments that are not connected to the well. Due to this vertical pressure gradient, H_2 can reach the deeper compartments, especially through the channelized sandstones that allow for vertical pressure communication. During the withdrawal of H_2, the top compartment (connected to the well) reaches the minimum BHP constraint, ceasing production even though the pressure in the deeper compartments remain above this limit. As a result of this complex depressurization, some H_2 remains trapped beneath the lower lacustrine mudstones, leading to low RF.

While this early cessation of production causes a decrease in RF for the models with low layer thickness and continuous lacustrine mudstones, it also causes an increase in MP for most of these models (Fig. 10). This increase occurs because the purity of the produced gas decreases with time during a cycle production. A sharp decrease of FP tends to occurs in the late stages of production due to CH_4 breakthrough (see, for example, the evolution of FP for Model 5 in Fig. 9. Thus, if the production is stopped early, the minimum FP increases because CH_4 breakthrough is prevented.

A reduced well perforation is beneficial for the models where the thickness of the reservoir and the sedimentological layers have a high in the DOE, that is, for the highest-N_g models. In these models, the well intersects thick sandstone layers and has no lacustrine mudstone below the perforated interval. For example, Fig. 13 shows the evolution of the H_2 plume for Model 12 (which has high N_g) during cycle 1. The shorter perforation interval promotes a vertical downwards filling that prevents the occurrence of gravity override. The resulting H_2 plume has an almost horizontal interface with the *in situ* CH_4, forming a H_2 column thicker than the perforated interval. This configuration keeps most of the CH_4 segregated at the bottom of the reservoir, reducing the influx of CH_4 into the well. Note that this behavior is only possible because the injected H_2 volume is large enough to completely fill reservoir along the perforation interval, so the well is mostly hydraulically connected to a region of the reservoir that is filled with H_2.

The performance metrics for the model ensemble are highly sensitive to the perforation of the well. The reference simulations showed that continuous lacustrine mudstones, *i.e.* the reservoir baffles, enhance UHS performance by mitigating

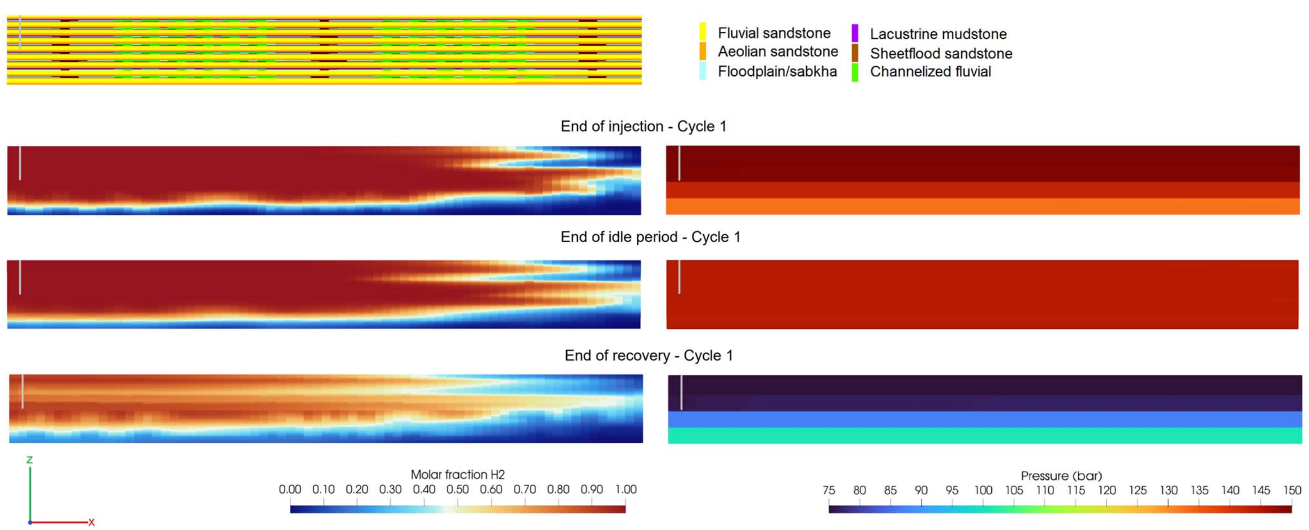

Fig. 12 Cross-sectional view of the sedimentological layers, the H_2 plume and the fluid pressures in Model 31 in Test 5 (well perforation) at different stages of cycle 1. The gray solid line in the models indicates the position of the perforated well.

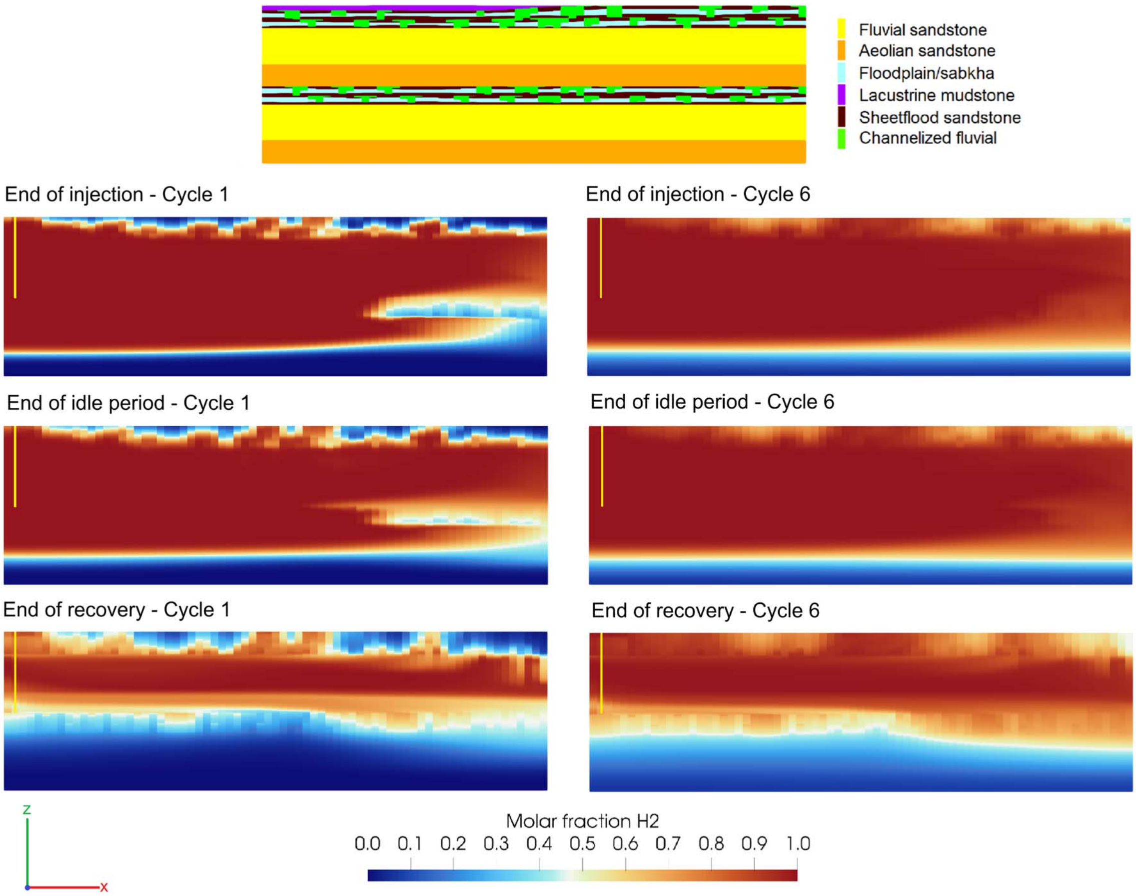

Fig. 13 Cross-sectional view of the sedimentological layers and the H_2 plume in Model 12 in Test 5 (well perforation) at different stages of cycles 1 and 6. The yellow solid line in the models indicates the position of the well perforation.

buoyancy forces through compartmentalization.[28] However, if the well perforation does not reach zones below these baffles, they can trap H_2 and reduce RF. In contrast, the results for models with high with N_g, *i.e.* a presumably less favorable geology, show good performance with shorter well perforations. Among the eight highest values of ultimate RF, seven correspond to models in the highest-N_g set.

4.6 Test 6—reinjection of co-produced CH_4

The proposed strategy for the reinjection of co-produced CH_4 (Fig. 2) results in significant changes in the performance metrics for the model ensemble. These changes are caused by the combined action of the reinjection of CH_4 in a well perforated at the bottom of the reservoir (CH_4 well) and the reduced perforation of the well used for the injection of H_2 (H_2 well).

Since the impact of the reduced well perforation on UHS performance is already assessed in Test 5 (well perforation), we take this previous test as a reference for comparison, to isolate the influence of the reinjection of CH_4 on the results. Compared to Test 5, the minimum and maximum MP and the minimum RF_1 remain approximately the same, but there are significant changes in the other performance metrics (Fig. 7). The maximum RF_1 increases from 94% to 100%, the minimum ultimate RF from 82% to 87% and the maximum ultimate RF from 96% to 102%. The maximum values of RF_1 and ultimate RF above 100% indicate that some of the H_2 initially injected as cushion gas is recovered.

The ultimate RF for the lowest and highest N_g models shows average increases (Fig. 8) that are largely due to the reinjection of co-produced CH_4. Compared to the reference simulations, the ultimate RF for the lowest N_g models shows an average increase of 6%, against an average decrease of 0.5% when the perforation of the H_2 well is reduced without reinjection of CH_4 (Test 5). The highest N_g models show an average increase of 9% in the ultimate RF, against 5% in Test 5.

The strategy for the reinjection of co-produced CH_4 enhances RF because it repressurizes the reservoir during production, which helps to maintain the average reservoir pressure above the minimum BHP (Fig. S8) and prolongs H_2 deliverability. Moreover, the reinjected CH_4 replaces part of the H_2 cushion gas in its function of maintaining a minimum working pressure. As a consequence, part of the cushion gas can be recovered, leading to a cycle RF above 100%. In addition to the pressure support provided by the reinjected CH_4, the high values of RF also result from the placement of the CH_4 well at the bottom of the reservoir, which promotes the vertical segregation of H_2 and CH_4 and reduces CH_4 influx into the H_2 well.

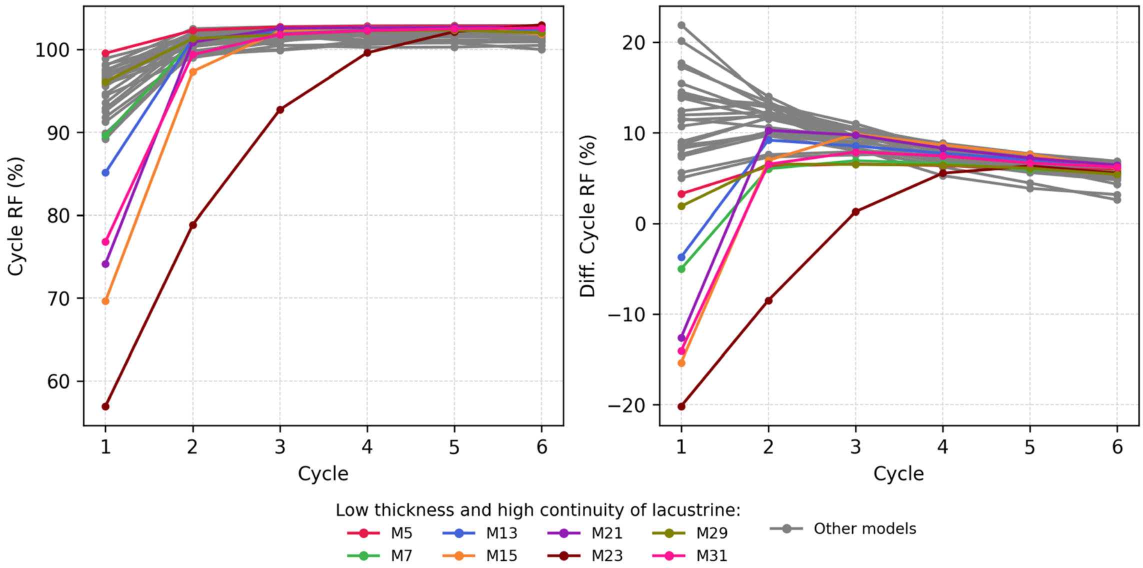

Fig. 14 Cycle recovery factor (left) and its difference from the reference tests (right) for Test 6 (reinjection of co-produced CH_4). The models with low thickness of the layers and high continuity of the lacustrine mudstones in the DOE are colored. They form a set of models whose initial recovery factor is negatively affected by reduced well perforation.

Fig. 14 presents the cycle RF and the difference in cycle RF, compared to the reference simulations. For most models, the cycle RF is always higher than in the reference simulations, and reaches a value of 100% or higher by cycle 3. As an exception, some models with high continuity of the lacustrine mudstones and low thickness of layers in the DOE face a pronounced decrease in RF_1, when compared to the reference simulations (Fig. 14). These models require more time for the pressure support provided by the CH_4 to become effective, because of the uneven pressurization caused by the baffles (*i.e.* the lacustrine mudstones) and the reduced well perforation (Fig. 11). Eventually, these models surpass the cycle RF of the reference simulations and reach values greater than 100% for this metric.

Fig. 15 presents the cycle MP and the difference in cycle MP compared to the reference simulations. For nearly all models, MP shows a trend that is significantly different from the reference simulations and the other operational tests (Fig. S1 to S7). In these other simulations, the cycle MP increases with increasing cycles due to the replacement of CH_4 by H_2 in the reservoir.[28] But when reinjecting the produced CH_4, the cycle MP remains approximately constant for each model, because the amount of CH_4 in the reservoir does not decrease.

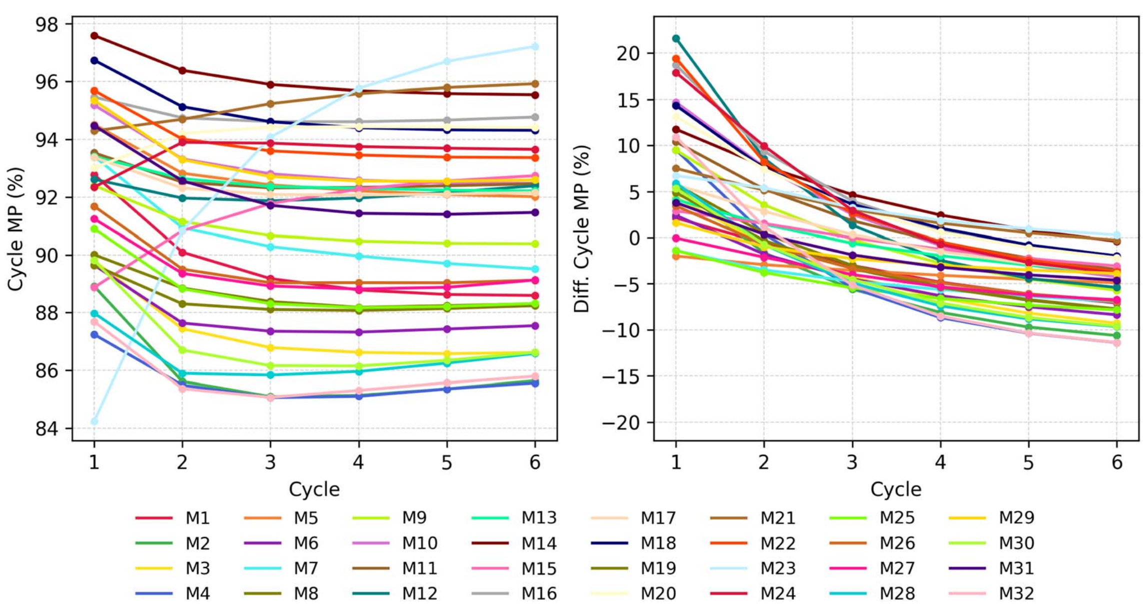

Fig. 15 Cycle mass purity (left) and its difference from the reference simulations (right) computed for Test 6 (reinjection of CH_4).

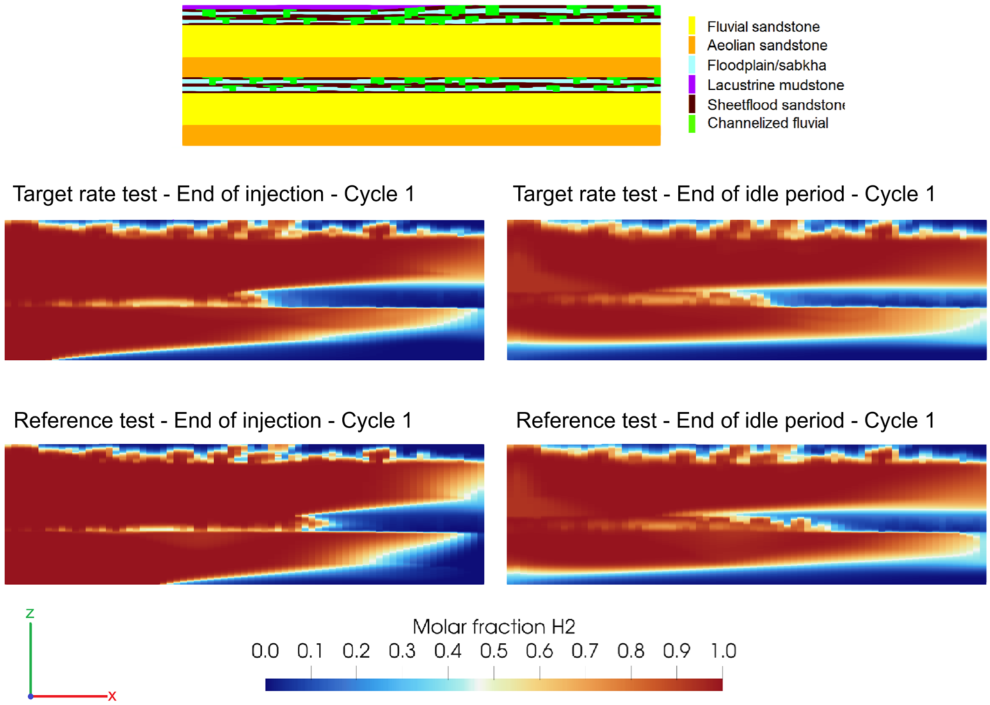

Fig. 16 Cross-sectional view of the sedimentological layers and the H_2 plume in Model 12 at the end of the injection and the idle period of cycle 1 for Test 7 (target rate) and the reference simulations.

For most of the models in the ensemble, MP_1 is higher when performing the reinjection of co-produced CH_4 than in the reference simulations. However, the increase in the cycle MP tends to be lost after a few cycles. By cycle 6, the cycle MP for the entire model ensemble is equal to or lower than the cycle MP for the reference simulations, with a difference that can reach −12% (Fig. 15). Thus, the gains in RF obtained from the reinjection of co-produced CH_4 come at the cost of purity in the long-term.

Our models do not prevent the production of the H_2 cushion gas, part of which becomes unnecessary for pressure support once CH_4 starts to be reinjected. However, the H_2 cushion gas can also serve to maintain a minimum purity, as observed when reducing the amount of cushion gas (Test 3) and when changing its composition (Test 4). Thus, alternatively, production can be stopped at each cycle when RF reaches 100%.

4.7 Test 7—no idle period

The performance metrics for the model ensemble are improved when there is no idle period in the cyclic operations. Compared to the reference simulations, where the cycles comprise a three-months idle period, the minimum RF_1 increases by 4%, the minimum MP_1 by 6% and the minimum ultimate RF remains stable around 90%; the maximum RF_1, maximum MP_1 and maximum ultimate RF increase by approximately 1% (Fig. 7).

During the idle period, mixing by molecular diffusion and buoyancy forces develop without being counteracted by viscous forces. Although these phenomena contribute to decrease RF and purity, the overall influence of the idle period on the performance metrics is considerably smaller than that of the other tested operational strategies (Fig. 8).

In our simulations, the dominating physical mechanism during the idle period is gravity segregation promoted by buoyancy forces (Fig. 5). Although this phenomenon is intensified during the idle period, the buoyancy forces are already significant during H_2 injection, as evidenced by the unstable fronts observed in high N_g models (Fig. 5). Thus, most of the H_2 loss due to gravity occurs already during the injection, the idle period being a secondary source of loss.

Furthermore, the action of buoyancy forces is mostly restricted to initial cycles, when more CH_4 is present in the reservoir. As CH_4 is produced, the importance of these forces decreases. As a result, the absence of idle periods leads to an increase in the ultimate RF of only 1%. These results are consistent with Kanaani *et al.*[22] and Zamehrian and Sedaee,[48] who reported small increases of approximately 1% in the ultimate RF after suppressing a two-month idle period.

4.8 Test 8—target rate

Among all tested operational strategies, doubling the target rates produced the smallest impact on the performance metrics (Fig. 8). The change in ultimate RF is small and slightly negative, with an average decrease of 0.3%. This variation is marginal and likely falls within the numerical error associated with time stepping and numerical diffusion.

Higher injection rates mitigate buoyancy effects by promoting a more viscous-dominated and less gravity-dominated flow regime. This effect is illustrated in Fig. 16, which shows the H_2 plume in Model 12 at the end of the injection of cycle 1. Compared to the reference simulation, gravity override develops farther from the well. Despite this change in the H_2 plume, doubling the target rates does not impact RF for Model 12 significantly, which is representative of the average behavior of the entire model ensemble.

In early cycles, the low impact of the target rates on the performance metrics can be explained in part by a counter effect that occurs during the idle period. Once the injection is stopped, the H_2 plume rearranges towards a static vertical equilibrium, offsetting the mitigation of gravity override promoted by higher rates during H_2 injection. For example, the H_2 plume in Model 12 at the end of the idle period of cycle 1 is similar for the reference simulation and for the simulation where the target rates are doubled.[15] In later cycles, the impact of target rates tends to become even less significant because the potential loss of H_2 due to gravity override is smaller, since the amount of CH_4 in the reservoir decreases.

A second mechanism that counteracts the expected improvement in UHS performance at higher target rates is the development of stronger pressure gradients along the well. The well intervals where pressures are lower cease production early, before higher-pressure intervals reach their full recovery potential. This early cessation of production tends to decrease RF and also explains the average increase in MP_1, as discussed for Test 5 (well perforation) in Section 4.5.

It is important to note that, apart from mitigating gravity override, higher flow rates can also boost mixing through hydrodynamic dispersion. This latter is not considered in our models, and could potentially change the impact of higher flow rates on UHS performance.

5 Guidelines for site screening and field development

We observed that the interplay between reservoir geology and operational constrains can impact the performance of UHS in different ways. We use these observations to establish a general framework that provides guidelines to identify depleted gas fields that are more suitable for UHS and to plan operational strategies that optimize UHS operations in these fields.

Fig. 17 presents this general framework, which can be divided in guidelines for site screening and for development planning. The screening procedure integrates the geological features previously identified as key controls, quantified through N_g in eqn (8), together with two operational parameters: depletion level and injected amount of H_2. The screening is also informed by, and can be updated based on, well planning, which therefore sits at the interface between screening and development planning.

Depletion level, injected amount of H_2, and well perforation are included in the screening criteria because of how they interact with geology. The impact of these strategies on UHS performance significantly depends on the reservoir N_g (Fig. 8). Moreover, variations in UHS performance due to different well perforations are impacted by the presence of continuous lacustrine mudstones. This has important implications because by changing the operational strategies that interact with geology, reservoirs with less favorable geology can yield good UHS performance and *vice versa*. Higher depletion levels and a short-enough well perforation yield good UHS performance for geological scenarios that have high N_g, *i.e.* for reservoirs with less favorable geology. Moreover, the UHS performance for reservoirs with high N_g is hindered when the

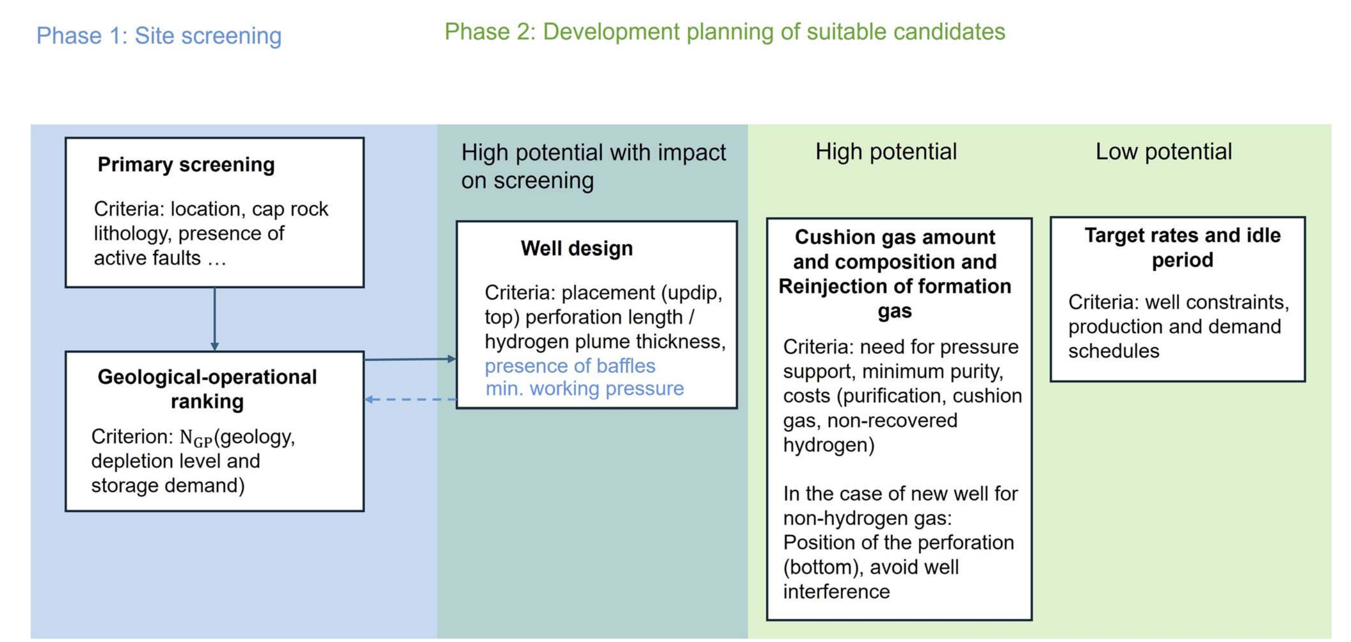

Fig. 17 Framework for the site screening (step 1) and development planning (step 2) for UHS in depleted gas fields. The planning of well perforation phase integrates both the screening and the development planning steps.

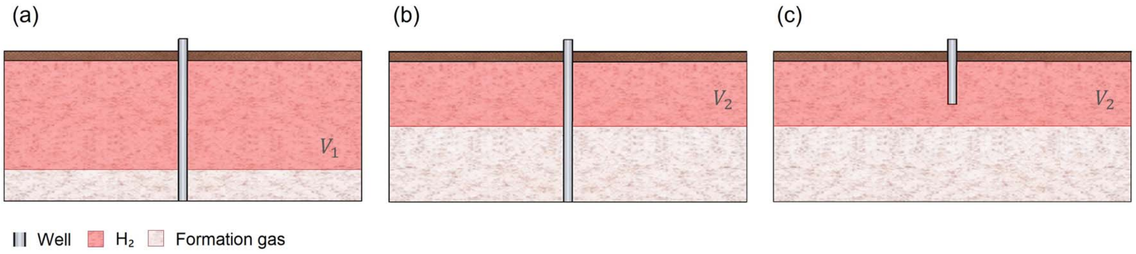

Fig. 18 Conceptual models for the configuration of H_2 and a denser formation gas after the full development of buoyancy forces in identical reservoirs, considering two injected volumes of H_2 (V_1 and V_2, with $V_1 > V_2$) and two well perforations. The reservoirs in (a) and (b) have the same well perforation but different injected amounts of H_2. As a result, the minimum purity of the produced H_2 is lower in reservoir (b). The reservoir in (c) has the lowest injected amount of H_2 but a shorter well perforation than the others; as a result, it recovers H_2 at a higher purity than reservoir (b), despite having the same injected H_2 volume.

stored amount of H_2 is significantly below full storage capacity, while reservoirs that have low N_g can perform well for a wider range of depletion levels injected amounts of H_2. For these reasons, the three operational strategies (depletion level, injected amount of H_2 and well perforation) integrate the screening step.

The framework for development planning includes guidelines for well perforation, which informs and is informed by the screening step, and four operational strategies that do not influence the site selection process: cushion gas amount and composition, reinjection of co-produced formation gas, target rates and the duration of idle periods. These four operational strategies are divided in two categories—low potential and high potential—according to the potential they show in our operational tests to influence UHS performance. The strategies with low potential include the duration of the idle period and the target rates. The strategies with high potential include the amount and composition of the cushion gas and the reinjection of co-produced formation gas (CH_4). Despite of their significant influence on UHS performance, these latter strategies can be planned after site screening because they are largely independent from geology. The amount and composition of cushion gas improves UHS performance similarly across all tested geological scenarios (Fig. 8). As for the reinjection of co-produced CH_4, although its impact on UHS performance is highly sensitive to N_g and to the presence of baffles, this sensitivity arises from the well perforation rather than from the CH_4 reinjection itself.

5.1 Criteria and guidelines for site screening

The site-screening step consists of a primary screening followed by a geological–operational screening. The primary screening considers social-technical-economical criteria and characteristics of the reservoir, with the objective of disqualifying reservoirs that can be rapidly judged as unsuitable for UHS. The proposal of these high-level criteria is not the focus of this work. However, several studies suggest approaches to screen candidate reservoirs for UHS worldwide, commonly considering criteria such as the proximity of the site to renewable-energy installations, cities, and pipeline networks,[16,39] as well as the lithology of the caprock[10,21,26,34] and the presence of faults.[10,26,34]

The main novelty proposed here is the geological-operational approach to further screen and rank candidates that are qualified in the primary screening. This approach provides a method to rank and select the depleted gas fields building on the analysis presented here and our earlier work.[28] This geological-operational screening should be viewed as a step within a broader screening procedure, and does not preclude the addition of further screening steps based on physical phenomena that are not accounted for in our models, such as H_2 loss due to biogeochemical reactions;[10,43] multiphase displacement effects governed by wettability, interfacial tension, and residual trapping;[14,46] and leakage through the caprock, bounding faults, and abandoned wells.[15,46]

Candidate reservoirs for UHS have different geological scenarios, depletion levels, and storage capacities. The interaction that was observed between geology (characterized by the reservoir N_g), depletion level and stored amount of H_2 implies that they must be considered together for site screening. Depletion level and stored mass of H_2 ultimately modify the relative proportions of H_2 and CH_4 in the reservoir. The strength of buoyancy forces, given by N_g, determines the susceptibility of the reservoir to producing CH_4 due to gravity override; but the amount of CH_4 that actually is produced depends on the fluid composition in the reservoir *i.e.* the proportional amount of CH_4 and H_2.

Fig. 18a and b illustrate this concept. They depict scenarios in which different volumes of H_2 are injected into two reservoirs with identical geology and identical rock and fluid properties. Because buoyancy forces depend only on these properties, their magnitude is the same in both reservoirs. The limiting state for the action of buoyancy forces is the hydrostatic equilibrium shown in the figures, where a column of H_2 develops over a column of denser formation gas, which here is CH_4. The volumes occupied by these columns in the reservoir depend on the proportional mass of injected H_2 and CH_4, and have implications for the composition of the gas produced at the well. The reservoir with the lower injected volume of H_2 contains a higher proportional amount of CH_4; as a result, it leads to a lower minimum purity under the same buoyancy forces than in the other reservoir.

Based on this concept, we propose scaling the buoyancy-force component of the modified gravity number (eqn (8)) with the mass ratio of CH_4 and H_2 to define the gravity–purity number, given by:

$$N_{\mathrm{GP}} = F_{\mathrm{b}}\frac{M_{\mathrm{CH_4}}}{M_{\mathrm{H_2}}} = k_x H g \Delta\rho \cos\alpha \frac{M_{\mathrm{CH_4}}}{M_{\mathrm{H_2}}}, \tag{9}$$

where F_b is buoyancy force, k_x is the reservoir upscaled longitudinal permeability, H is the thickness of the depositional cycle (repeating sandstone–mudstone patterns), α is the reservoir dip angle, g is gravitational acceleration, $\Delta\rho$ is the density difference between H_2 and CH_4, M_{CH4} is the mass of CH_4 initially in place, and M_{H_2} is the maximum stored mass of H_2. The density difference $\Delta\rho$ should be computed at a consistent reference pressure and temperature for all reservoirs. Here, we computed the fluid densities at the minimum BHP, following Loyola *et al.*,[28] and at the constant reservoir temperature of 55 °C. The mass M_{H_2} is the maximum stored mass of H_2 in one cycle. The mass M_{CH_4} is computed as:

$$M_{\mathrm{CH_4}} = V_{\mathrm{r}}\bar{\phi} p_{\mathrm{CH_4}}(T_0, p_0), \tag{10}$$

where V_r and $\bar{\phi}$ are the reservoir volume and average porosity, respectively, and $p_{CH_4}(T,p_0)$ is the density of CH_4 as computed for the initial reservoir pressure (p_0) and temperature (T_0).

Note that N_{GP} is not dimensionless, unlike N_g. It has units of force, as it scales the magnitude of buoyancy forces with the H_2–CH_4 mass ratio to express their potential impact on purity and recovery factor. We omit the viscous-force component of the gravity number, because flow rates were observed to have limited impact on UHS performance; therefore, they are considered only in later stages of the development planning step (Fig. 17).

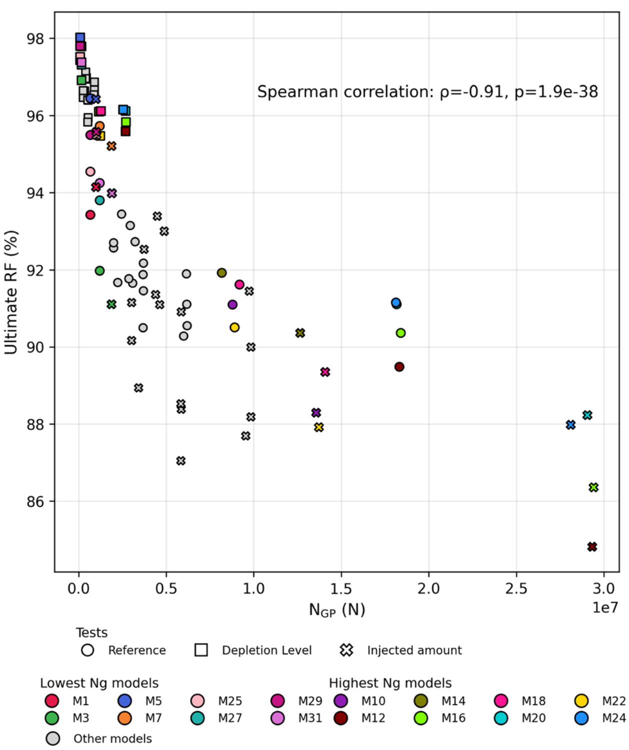

Fig. 19 Correlation between the gravity–purity number (N_{GP}) and the ultimate recovery factor (RF) for the entire model ensemble in the reference simulations, Test 1 (depletion level), and Test 2 (injected amount of H_2). The lowest-N_g and highest-N_g models are highlighted. Different operational tests are represented by different symbols.

Fig. 19 shows that, when considering the reference tests together with Test 1 (depletion level) and Test 2 (injected amount), N_{GP} correlates strongly with the ultimate RF. For this reason, we propose N_{GP} as the main criterion in the geological-operational screening. N_{GP} allows the ranking of storage units using geological features (*e.g.* reservoir thickness, dip, and upscaled permeability), depletion level, and store mass of H_2. By incorporating operational strategies, N_{GP} improves the modified N_g with the consideration of operational strategies that control the mass proportion of H_2 in the reservoir.

Fig. 19 illustrates that N_{GP} captures the high improvement in UHS performance for highest N_g models when the depletion level is high. It also shows that the lowest N_g models achieve ultimate RF values above 91% regardless of the tested depletion level or injected H_2. Thus, models with low N_g imply good UHS performance in situations where depletion level and storage volumes are uncertain.

N_{GP} assigns a higher ranking to a candidate field as the depletion level increases. This is a direct consequence of the assumption made in our operational tests, according to which the deliverability of H_2 at the well is not compromised by lower reservoir pressures and BHPs. It is possible, nonetheless, that target production rates require a minimum BHP that is well above the pressure of the depleted gas field. In that case, the proposed site screening can be further enhanced by disqualifying reservoirs with pressures significantly lower than the desired minimum working pressure. The estimation of a minimum working pressure for the well system should be feasible during site selection, because target rates are dictated by a previously identified H_2 demand that justifies the UHS operation. This disqualification represents one way in which well planning can influence the site selection process. Nonetheless, providing criteria for this disqualification is out of the scope of the current work, since it would require consideration of the design of wellbore hydraulics and surface facilities.

5.2 Criteria and guidelines for development planning

The framework for development planning provides guidelines to design different operational strategies for UHS. These strategies are categorized according to their low or high potential to optimize UHS performance.

We propose the types of criteria that can be used to design operational strategies, without defining precise quantitative values or detailing specific evaluation methods. The purpose of this framework is to provide guidelines that serve as starting points for more detailed, interdisciplinary and site-specific assessments of UHS operational strategies.

5.2.1 Well perforation. A sufficiently short well perforation can mitigate gravity override in geological scenarios with high N_g. Fig. 17 presents the criteria to optimize well perforation. If a depleted reservoir contains multiple wells, these criteria can also be used to identify the most suitable wells to be repurposed for injecting H_2.

The well should be placed updip, and be perforated the top of the reservoir, since the H_2 tends to migrate upwards due to gravity segregation from the formation gas. In addition, perforation length is a key factor controlling UHS performance (see Section 4.5). If the injected volume of H_2 is large enough to fill the reservoir to a thickness greater than the well perforation, the formation gas is displaced downward away from the well (Fig. 13 and 18c). Hence, we propose that the well perforation should be short enough to be hydraulically connected only to reservoir zones that are filled with H_2. This recommendation is particularly relevant for for reservoirs with higher N_g, which are more susceptible to buoyancy forces.

Fig. 18b and c provide conceptual models illustrating the interplay between the well perforation and the injected amount of H_2. If the well perforation is short relative to the thickness occupied by the H_2 column, inflow to the well will consist predominantly of H_2 (Fig. 18c). Otherwise, the well will also be hydraulically connected to deeper zones occupied by formation gas (Fig. 18b).

The presence of laterally continuous baffles is also a critical factor to consider when planning well perforation. In our model ensemble, the baffles are laterally continuous mudstones that have maximum permeability of 0.01 mD, which is 4 to 5 orders of magnitude smaller than the permeability of the fluvial sandstones. These baffles compartmentalize the reservoir and reduce the effective thickness over which buoyancy forces act, thereby mitigating gravity override and potentially improving UHS performance.[28] However, they also pose a risk of trapping H_2 in geobodies below them that are not connected to the well. Therefore, if laterally extensive baffles are present in the reservoir, the well perforation should be long enough to reach zones below the baffles to which H_2 can migrate.

In practice, there is often significant uncertainty regarding the lateral extent of baffles intersecting wells. If a reservoir contains frequent baffles of uncertain extent, the risk of H_2 trapping beneath them exists. This risk can be mitigated by perforating the well across the entire reservoir thickness. When this strategy is not feasible, the most conservative approach do deal with this uncertainty is to disqualify such reservoirs for UHS. In this way, well-perforation planning informs and updates the screening stage.

While the criteria for well perforation planning are based on numerical models involving a single well for H_2 injection and production, these criteria can be applied individually to each well when multiple wells are used, taking into account their respective zones of influence within the reservoir.

We restrict ourselves to providing criteria for well perforation, which directly interacts with reservoir geology and is considered in our tests. However, as stated in Section 5.1, other aspects of well design, notably working pressure constraints, could also inform the site screening process.

5.2.2 Cushion gas and reinjection of formation gas strategy. The cushion gas amount and composition have high potential to improve UHS performance. These strategies should be defined based on at least three criteria: the need for pressure support, the required minimum purity, and total costs.

Pressure support is required when the reservoir pressure is lower than the minimum working pressure. It can also be used to improve H_2 production when the reservoir is expected to develop high pressure gradients during storage operations, which favor the migration of hydrogen away from the well. The required minimum purity depends on the intended application for the H_2. The use of cushion gas can improve the purity of the H_2 produced at the well, potentially reducing purification costs.

The cushion gas composition is particularly relevant for meeting purity requirements. While pressure support can be provided by any type of gas, improvements in the minimum purity of the produced H_2 are achieved mainly through the use of H_2-containing cushion gas.

All these considerations need to be evaluated together with a cost assessment. Ultimately, a cushion gas strategy is only effective if the costs associated with the cushion gas are compensated by gains in H_2 production or by reductions in purification costs.

The reinjection of co-produced formation gas is a waste management solution that can also provide pressure support to the reservoir. Similar to the planning of the cushion gas strategy, this strategy requires consideration of the need for pressure support, as well as careful evaluation of minimum purity requirements and costs. The reinjection of co-produced formation gas tends to maintain the purity of the produced H_2 at approximately constant levels and to increase purification costs in the long term (see Section 4.6). Moreover, this strategy requires an additional injection well. To limit the co-production of formation gas after reinjection, this well should be perforated near the bottom of the reservoir and should not have pressure interference with the H_2 injection–production well. The same considerations apply if a separate well is planned for injecting any cushion gas other than H_2.

5.2.3 Idle period duration and target rates. The duration of the idle period and the target rates should be adjusted to factors external to the reservoir, such as expected schedules for H_2 production and energy demand. The planning of target rates should also consider the constraints of the injection system and the well. Since these parameters are influenced by external factors, they are not flexibly adjustable. Moreover, since they depend on H_2 production demand, they are expected to vary. The limited flexibility in their adjustment and their variability are not detrimental to the operation and should not interfere with the site selection process, as both operational strategies showed low impact on UHS performance metrics and limited interaction with geology in our operational tests.

6 Conclusions

We performed several operational tests, corresponding to simulations of UHS with different operational strategies, using a model ensemble that contains various geological scenarios for the Bunter sandstone. The tested operational strategies include depletion level, injected amount of hydrogen, cushion gas amount and composition, well perforation, reinjection of co-produced gas, duration of the idle period and target rates.

The operational tests are based on a set of reference simulations, based on which we identified the geological controls on UHS performance and proposed a modified gravity number as a geological screening criterion. We identified the potential impact of the tested strategies on UHS performance by comparing the results of the operational tests with those of the reference simulations. We also assessed the interaction between operational strategies and reservoir geology by comparing how the strategies affected UHS performance metrics across different geological scenarios, mainly relying on comparisons between reservoir models with low and high gravity numbers. The main findings of the operational tests are summarized below.

• UHS performance in reservoirs with high gravity numbers is sensitive to the depletion level and can be adequate if the depletion level is high. On the other side, reservoirs with low gravity number can achieve strong UHS performance at a higher range of depletion levels because they are less susceptible to gravity override. This conclusion is based on the assumption that lower minimum working pressures can be used for higher depletion levels without affecting well deliverability.

• UHS performance in reservoirs with high gravity numbers is also sensitive to the injected amount of hydrogen. In these reservoirs, operating below storage capacity compromises performance. In contrast, reservoirs with lower gravity numbers perform well across the tested injected masses of hydrogen.

• Changes in the amount and composition of cushion gas cause significant differences on UHS performance metrics that are largely independent of geology.

• UHS performance in reservoirs with high gravity numbers can be significantly improved if the well perforation is sufficiently short and placed at the top of the reservoir. The perforation length should be short enough to be hydraulically connected to reservoir zones filled mostly with hydrogen.

• If the reservoir has frequent and laterally continuous baffles, there is a risk of hydrogen trapping. In these cases, the well perforation should be long enough to reach the geobodies below the baffles, where hydrogen may be trapped.

• The reinjection of produced formation gas is a waste management solution that also serves to provide pressure support to the reservoir. This strategy can improve the recovery factor at the cost of a long-term reduction in purity.

• The duration of the idle period and the target rates have low potential to affect UHS performance.

In summary, reservoirs with high gravity number can be engineered to yield good UHS performance if critical operational strategies such as well perforation and depletion level are carefully planned. Reservoirs with low gravity numbers offer more flexibility for different operational strategies and are therefore better-suited to handle operational uncertainties.

Based on these findings, we proposed a framework for site screening and development planning for UHS in depleted gas fields. The site screening framework includes a screening criterion called the gravity–purity number, an extension of our modified gravity number that integrates geological and operational aspects. This criterion can be used to select candidate gas reservoirs that minimize the loss of hydrogen through buoyancy effects, considering their geology and the proportional amount of hydrogen and resident gas in the reservoir. This latter depends on operational parameters such as depletion level and the stored volume of hydrogen. The screening step is also informed by the identification of laterally extensive baffles in the reservoir, which pose the risk of trapping of hydrogen.

The gas reservoirs identified as suitable for UHS by the screening procedure should undergo more detailed studies as part of their development planning, which includes the design of the well system and the choice of cushion gas. Our proposed framework indicates the types of criteria that should guide the optimisation of these strategies. These criteria are often external to the reservoir—typically related to cost analysis and well-system constraints—so development planning is inevitably interdisciplinary. We do not provide objective criteria for these external factors, which should be addressed in complementary studies to the reservoir geology and engineering analyses presented here. The selection of optimal UHS strategies for target fields requires an integrated approach combining techno-economic analyses with coupled reservoir-wellbore flow simulations.

The physical phenomena considered in the numerical models are isothermal gas flow and molecular diffusion. The main source of hydrogen loss in this setting is gravity override; thus, the proposed guidelines for screening and development planning target the mitigation of undesired gravity effects. However, other physical processes that may cause hydrogen loss, such as hydrogen-consuming reactions, dispersion, water flow due to aquifer support, capillary effects and other interfacial phenomena, were not included. Future studies should incorporate these physical phenomena and evaluate their impact on hydrogen loss to refine the screening and development planning framework.

Conflicts of interest

There are no conflicts of interest to declare.

Data availability

The numerical simulations performed in this study use a completely open-source frame work, based on open-DARTS, Rapid Reservoir Modeling (RRM), and the ensemble of open-access geological models for the Bunter sandstone designed by Alshakri *et al.* (2022). Open-DARTS is available at: https://

gitlab.com/open-darts/open-darts; RRM is available at: https://bitbucket.org/rapidreservoirmodelling/rrm; and the Bunter geological models are available at: https://figshare.com/articles/dataset/RRM_models_of_Sherwood_and_Bunter_Sandstones/19210002. Moreover, the open-DARTS models for the reference simulations of this study are also available at: https://zenodo.org/records/19442719. Note that these models are compatible with version 1.2.1 of open DARTS.

Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d6va00162a.

Appendix

A Design of Experiment

Table 7 DOE composed of 32 different combinations of the factors A to J in Table 1, where L stands for low setting and H for high setting

Models 1–16											Models 17–32										
	A	B	C	D	E	F	G	H	I	J		A	B	C	D	E	F	G	H	I	J
1	L	L	L	L	H	L	L	L	L	L	17	L	L	L	L	L	H	L	L	H	H
2	H	L	L	L	H	L	H	H	L	L	18	H	L	L	L	L	H	H	H	H	H
3	L	H	L	L	L	L	H	H	L	L	19	L	H	L	L	H	H	H	H	H	H
4	H	H	L	L	L	L	L	L	L	L	20	H	H	L	L	H	H	L	L	H	H
5	L	L	H	L	L	L	L	H	L	H	21	L	L	H	L	H	H	L	H	H	L
6	H	L	H	L	L	L	H	L	L	H	22	H	L	H	L	H	H	H	L	H	L
7	L	H	H	L	H	L	H	L	L	H	23	L	H	H	L	L	H	H	L	H	L
8	H	H	H	L	H	L	L	H	L	H	24	H	H	H	L	L	H	L	H	H	L
9	L	L	L	H	L	L	H	L	H	L	25	L	L	L	H	H	H	H	L	L	H
10	H	L	L	H	L	L	L	H	H	L	26	H	L	L	H	H	H	L	H	L	H
11	L	H	L	H	H	L	L	H	H	L	27	L	H	L	H	L	H	L	H	L	H
12	H	H	L	H	H	L	H	L	H	L	28	H	H	L	H	L	H	H	L	L	H
13	L	L	H	H	H	L	H	H	H	H	29	L	L	H	H	L	H	H	H	L	L
14	H	L	H	H	H	L	L	L	H	H	30	H	L	H	H	L	H	L	L	L	L
15	L	H	H	H	L	L	L	L	H	H	31	L	H	H	H	H	H	L	L	L	L
16	H	H	H	H	L	L	H	H	H	H	32	H	H	H	H	H	H	H	H	L	L

Acknowledgements

The authors thanks Shell Global Solutions International B.V. for sponsoring this work through ASSET program and granting permission to publish this work. The authors also thank the Shell hydrogen team for discussions and feedback during the project. Sebastian Geiger thanks Energi Simulation for supporting his Chair in Sustainable Geoenergy.

References

1 *Northern lights project*, https://norlights.com/, accessed in November 2025.
2 *Porthos project*, https://www.porthosco2.nl/en/project/, accessed in November 2025.
3 International Energy Agency, *Global Hydrogen Review 2025*, 2025, pp. 139–142.
4 J. Alshakri, G. J. Hampson, C. Jacquemyn, M. D. Jackson, D. Petrovskyy, S. Geiger, J. D. Machado Silva, S. Judice, F. Rahman, and M. C. Sousa, *A Screening Assessment of the Impact of Sedimentological Heterogeneity on CO_2 Migration and Stratigraphic-Baffling Potential: Sherwood and Bunter Sandstones, UK*, Geological Society Special Publication, 2023, vol. 528, pp. 245–266, DOI: 10.1144/SP528-2022-34.
5 Z. Bo, M. Boon, H. Hajibeygi and S. Hurter, Impact of experimentally measured relative permeability hysteresis on reservoir-scale performance of underground hydrogen storage (UHS), *Int. J. Hydrogen Energy*, 2023, **48**(4), 13527–13542, DOI: 10.1016/j.ijhydene.2022.12.270.
6 T. Busch, J. Derichs, T. Klütz, J. Linßen and D. Stolten, The hydrogen supply chain—a comprehensive literature review incorporating purity analysis, *Int. J. Hydrogen Energy*, 2025, **164**, 149367, DOI: 10.1016/j.ijhydene.2025.04.532.
7 H. Cleijne, M. de Ronde, M. Duvoort, W. de Kleuver and J. Raadschelders, *North Sea Energy Outlook*, 2020.
8 Delft High Performance Computing Centre (DHPC), *DelftBlue Supercomputer (Phase 2)*, 2024, https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase2.

9 Department of Energy and Climate Change, *UK Renewable Energy Roadmap*, 2011.
10 N. Diamantakis, A. Peecock, O. Shahrokhi, S. Pitchaimuthu and J. M. Andresen, A review of analogue case studies relevant to large-scale underground hydrogen storage, *Energy Rep.*, 2024, **11**(6), 2374–2400, DOI: 10.1016/j.egyr.2024.02.002.
11 E. Cecchetti, A. W. Martinius, P.-O. Bruna, A. Bender and H. A. Abels, Structural controls on the triassic main buntsandstein sediment distribution in the roer valley graben, the netherlands, *Neth. J. Geosci.*, 2024, **103**, e23, DOI: 10.1017/njg.2024.17.
12 R. Ershadnia, M. Singh, S. Mahmoodpour, A. Meyal, F. Moeini, S. Abolfazl Hosseini, D. Murray Sturmer, M. Rasoulzadeh, Z. Dai and M. R. Soltanian, Impact of geological and operational conditions on underground hydrogen storage, *Int. J. Hydrogen Energy*, 2023, **48**(1), 1450–1471, DOI: 10.1016/j.ijhydene.2022.09.208.
13 N. Foote, S. J. Hatter, E. Gray and D. Charlaftis, Reservoir characterization of the triassic bunter sandstone in the 43/28-nepbh1 borehole: implications for CO_2 storage in the neighbouring endurance structure, southern north sea, UK, *Geoenergy*, 2025, **3**(1), DOI: 10.1144/geoenergy2025-012.
14 I. Gomez Mendez, W. M. M. El-Sayed, A. H. Menefee and Z. T. Karpyn, Insights into underground hydrogen storage challenges: A review on hydrodynamic and biogeochemical experiments in porous media, *Energy Fuels*, 2024, **38**(21), 20015–20032, DOI: 10.1021/acs.energyfuels.4c03142.
15 R. Guo, H. Wang, J. Leng and S. A. Hosseini, Evaluation of hydrogen leakage through abandoned wells to overlaying saline aquifers during underground hydrogen storage in depleted natural gas reservoirs, *Gas Sci. Eng.*, 2025, **140**, 205659, DOI: 10.1016/j.jgsce.2025.205659.
16 S. Harati, S. Rezaei Gomari, M. Ramegowda and T. Pak, Multi-criteria site selection workflow for geological storage of hydrogen in depleted gas fields: A case for the uk, *Int. J. Hydrogen Energy*, 2024, **51**, 143–157, DOI: 10.1016/j.ijhydene.2023.10.345.
17 A. Hassanpouryouzband, E. Joonaki, K. Edlmann, N. Heinemann and J. Yang, Thermodynamic and transport properties of hydrogen containing streams, *Sci. Data*, 2020, **7**, 222, DOI: 10.1038/s41597-020-0568-6, DOI: 10.1038/s41597-020-0568-6.
18 A. Hassanpouryouzband, E. Joonaki, K. Edlmann and R. Stuart Haszeldine, *Offshore Geological Storage of Hydrogen: Is This Our Best Option to Achieve Net-Zero?*, 2021, ISSN 23808195.
19 Z. Dalal Isfehani, A. Jafari, J. Fahimpour, M. Hosseini, S. Iglauer and A. Keshavarz, Sandstone wettability and mixed gas composition: unraveling the impact of CO_2 in hydrogen geo-storage, *Int. J. Hydrogen Energy*, 2024, **59**, 1352–1366, DOI: 10.1016/j.ijhydene.2024.02.120.
20 C. Jacquemyn, M. E. H. Pataki, G. J. Hampson, M. D. Jackson, D. Petrovskyy, S. Geiger, C. C. Marques, J. D. Machado Silva, S. Judice, F. Rahman and M. C. Sousa, Sketch-based interface and modelling of stratigraphy and structure in three dimensions, *J. Geol. Soc.*, 2021, **178**(4), jgs2020–jgs2187, DOI: 10.1144/jgs2020-187.
21 J. Juez-Larré, S. van Gessel, R. Dalman, G. Remmelts and R. Groenenberg, Assessment of underground energy storage potential to support the energy transition in the netherlands, *First Break*, 2019, **37**(7), 57–66, DOI: 10.3997/1365-2397.n0039.
22 M. Kanaani, B. Sedaee and M. Asadian-Pakfar, Role of cushion gas on underground hydrogen storage in depleted oil reservoirs, *J. Energy Storage*, 2022, **45**(1), 103783, DOI: 10.1016/j.est.2021.103783.
23 M. Khait and D. Voskov, Adaptive parameterization for solving of thermal/compositional nonlinear flow and transport with buoyancy, *SPE J.*, 2018, **23**(2), 522–534, DOI: 10.2118/182685-pa.
24 S. Kobeissi, N. N. A. Ling, K. Yang, E. F. May and M. L. Johns, Dispersion of hydrogen in different potential cushion gases, *Int. J. Hydrogen Energy*, 2024, **60**(3), 940–948, DOI: 10.1016/j.ijhydene.2024.02.151.
25 S. Korevaar, R. Dalman, S. Nelskamp, S. Atkins, E. Boter, E. Wiarda, M. Nolten and K. Beintema, *Play 5 triassic*, https://www.geodeatlas.nl/pages/play-5-triassic, 2023, accessed: 04 07 2025.
26 L. Lankof, K. Luboń, Y. L. Gallo and R. Tarkowski, The ranking of geological structures in deep aquifers of the polish lowlands for underground hydrogen storage, *Int. J. Hydrogen Energy*, 2024, **62**(4), 1089–1102, DOI: 10.1016/j.ijhydene.2024.03.106.
27 E. W. Lemmon, I. H. Bell, M. L. Huber and M. O. McLinden, Thermophysical properties of fluid systems, *NIST Chemistry WebBook*, NIST Standard Reference Database Number, vol. 69, 2025, retrieved April 10, 2025.
28 A. Loyola, D. Voskov, R. Farajzadeh, K. Bisdom, K. d. Borst and S. Geiger, Geological controls on underground hydrogen storage in depleted gas fields, *EarthArXiv*, 2026, preprint, DOI: 10.31223/X5V18J.
29 M. Lysyy, M. Fernø and G. Ersland, Seasonal hydrogen storage in a depleted oil and gas field, *Int. J. Hydrogen Energy*, 2021, **46**, 25160–25174, DOI: 10.1016/j.ijhydene.2021.05.030.
30 X. Lyu and D. Voskov, Advanced modeling of enhanced co2 dissolution trapping in saline aquifers, *Int. J. Greenhouse Gas Control*, 2023, **127**, 103907, DOI: 10.1016/j.ijggc.2023.103907.
31 X. Lyu, M. Khait and D. Voskov, Operator-based linearization approach for modeling of multiphase flow with buoyancy and capillarity, *SPE J.*, 2021, **26**(4), 1858–1878, DOI: 10.2118/205378-PA.
32 G. Medici, L. West and N. Mountney, Sedimentary flow heterogeneities in the Triassic U.K. Sherwood sandstone group: Insights for hydrocarbon exploration, *Geol. J.*, 2018, **54**, 1361–1378, DOI: 10.1002/gj.3233.
33 Netherlands Oil and Gas Portal (NLOG), *Temperature data*, https://www.nlog.nl/en/temperature-data, accessed in August 2025.
34 E. Rita Okoroafor, S. D. Saltzer and A. R. Kovscek, Toward underground hydrogen storage in porous media: Reservoir

engineering insights, *Int. J. Hydrogen Energy*, 2022, **47**(9), 33781–33802, DOI: 10.1016/j.ijhydene.2022.07.239.

35 D. W. Peaceman, Interpretation of well-block pressures in numerical reservoir simulation, *Soc. Pet. Eng. J.*, 1978, **18**(3), 183–194, DOI: 10.2118/6893-PA.

36 D. Petrovskyy, C. Jacquemyn, S. Geiger, M. D. Jackson, G. J. Hampson, J. D. Machado Silva, S. Judice, F. Rahman and M. Costa Sousa, Rapid flow diagnostics for prototyping of reservoir concepts and models for subsurface CO_2 storage, *Int. J. Greenhouse Gas Control*, **124**(3), 2023, DOI: 10.1016/j.ijggc.2023.103855.

37 W. T. Pfeiffer, S. A. al Hagrey, D. Köhn, W. Rabbel and S. Bauer, Porous media hydrogen storage at a synthetic, heterogeneous field site: numerical simulation of storage operation and geophysical monitoring, *Environ. Earth Sci.*, 2016, **75**, 8, DOI: 10.1007/s12665-016-5958-x.

38 R. A. Sadkhan and W. J. Al-Mudhafar, Key aspects of underground hydrogen storage in depleted hydrocarbon reservoirs and saline aquifers: a review and understanding, *Energy Geosci.*, 2024, **5**(4), 100339, DOI: 10.1016/j.engeos.2024.100339.

39 A. Safari, Y. Sugai, M. Sarmadivaleh and M. Imai, Screening and ranking japanese gas fields for underground H_2 storage potential: impact of the reservoir drive mechanism, *J. Energy Storage*, 2023, **70**, 107679, DOI: 10.1016/j.est.2023.107679.

40 A. Sainz-Garcia, E. Abarca, V. Rubi and F. Grandia, Assessment of feasible strategies for seasonal underground hydrogen storage in a saline aquifer, *Int. J. Hydrogen Energy*, 2017, **42**(6), 16657–16666, DOI: 10.1016/j.ijhydene.2017.05.076.

41 M. Shook, D. Li and L. Lake, Scaling immiscible flow through permeable media by inspectional analysis, *In Situ*, 1992, **16**(4), 311–349.

42 R. Tarkowski, Underground hydrogen storage: characteristics and prospects, *Renewable Sustainable Energy Rev.*, 2019, **105**(5), 86–94, DOI: 10.1016/j.rser.2019.01.051.

43 E. M. Thaysen, S. McMahon, G. J. Strobel, I. B. Butler, B. T. Ngwenya, N. Heinemann, M. Wilkinson, A. Hassanpouryouzband, C. I. McDermott and K. Edlmann, Estimating microbial growth and hydrogen consumption in hydrogen storage in porous media, *Renewable Sustainable Energy Rev.*, 2021, **151**, 111481, DOI: 10.1016/j.rser.2021.111481.

44 M. v. Eijk, *Volpriehausen Prospectivity Review*, MSc thesis, VU University, Amsterdam, 2014.

45 B. M. M. van Kempen, H. F. Mijnlieff and J. van der Molen, *Data Mining in the Dutch Oil and Gas Portal: a Case Study on the Reservoir Properties of the Volpriehausen Sandstone Interval*, Geological Society, London, Special Publications, 2018, vol. 469, pp. 253–267, DOI: 10.1144/SP469.1.

46 W. A. van Rooijen and H. Hajibeygi, Site selection for underground hydrogen storage in porous media: critical review and outlook, *Energy Fuels*, 2025, **39**(41), 19600–19613, DOI: 10.1021/acs.energyfuels.5c03665.

47 D. Voskov, I. Saifullin, A. Novikov, M. Wapperom, L. Orozco, G. S. Seabra, Y. Chen, M. Khait, X. Lyu, X. Tian, S. De Hoop and A. Palha, open Delft Advanced Research Terra Simulator(open-DARTS), *J. Open Source Softw.*, 2024, **9**(99), 6737, DOI: 10.21105/joss.06737.

48 M. Zamehrian and B. Sedaee, Underground hydrogen storage in a partially depleted gas condensate reservoir: Influence of cushion gas, *J. Pet. Sci. Eng.*, 2022, **212**(5), 110304, DOI: 10.1016/j.petrol.2022.110304.

49 L. Zeng, R. Sander, Y. Chen and Q. Xie, Hydrogen storage performance during underground hydrogen storage in depleted gas reservoirs: a review, *Engineering*, 2024, **40**, 211–225, DOI: 10.1016/j.eng.2024.03.011.

