



Local Hybrid Newton method for the acceleration of well event handling in the simulation of CO₂ storage using supervised learning

Antoine Lechevallier, Sylvain Desroziers, Thibault Faney, Eric Flauraud, Frédéric Nataf, Harald Sandve

► To cite this version:

Antoine Lechevallier, Sylvain Desroziers, Thibault Faney, Eric Flauraud, Frédéric Nataf, et al.. Local Hybrid Newton method for the acceleration of well event handling in the simulation of CO₂ storage using supervised learning. 2024. <hal-04774181>

HAL Id: hal-04774181

<https://hal.science/hal-04774181v1>

Preprint submitted on 8 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

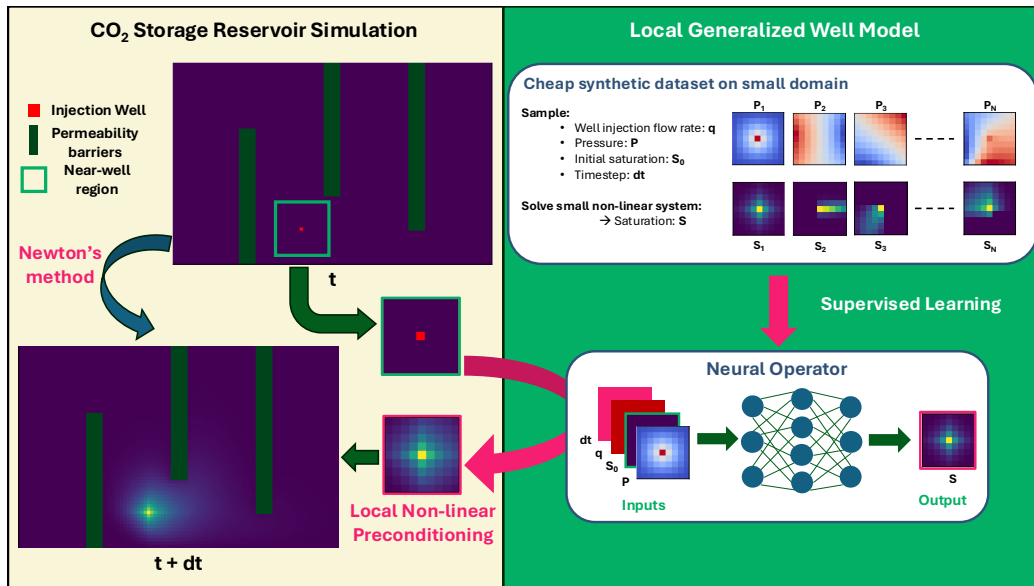


HAL Authorization

Graphical Abstract

Local Hybrid Newton method for the acceleration of well event handling in the simulation of CO₂ storage using supervised learning

Antoine Lechevallier, Sylvain Desroziers, Thibault Faney, Eric Flauraud, Frédéric Nataf, Tor Harald Sandve



Highlights

Local Hybrid Newton method for the acceleration of well event handling in the simulation of CO₂ storage using supervised learning

Antoine Lechevallier, Sylvain Desroziers, Thibault Faney, Eric Flauraud, Frédéric Nataf, Tor Harald Sandve

- **Innovative Approach:** We develop a local machine-learning based Newton preconditioning method, showcasing its potential through a proof of concept.
- **Local Hybrid Initialization:** Our approach integrates Newton’s method with a local initialization strategy, enabling faster convergence while preserving numerical guarantees.
- **Application in CO₂ Storage Simulation:** Our method is applied to accelerate the numerical simulation of CO₂ storage, using an open-source Python framework.
- **Initialization Strategy:** We initialize the pressure using a linear approximation and the saturation using a Fourier Neural Operator neural network trained in a supervised way.
- **Performance Boost:** Our method significantly reduces the number of Newton iterations across various well events and well locations, showcasing its efficiency.

Local Hybrid Newton method for the acceleration of well event handling in the simulation of CO₂ storage using supervised learning

Antoine Lechevallier^a, Sylvain Desroziers^d, Thibault Faney^b, Eric Flauraud^b, Frédéric Nataf^c, Tor Harald Sandve^a

^a*NORCE, Norway*

^b*IFP Energies Nouvelles, France*

^c*Laboratoire Jacques-Louis Lions, France*

^d*Michelin SA, France*

Abstract

Reservoir simulation is crucial in understanding and predicting the behavior of subsurface reservoirs, aiding in efficient resource management. By modeling fluid flow, pressure changes, and other dynamic processes, it provides valuable insights for optimizing development strategies. This helps in making informed decisions, minimizing risks, and maximizing the sustainable use of natural resources such as water, geothermal energy, or carbon storage. However, simulating fluid flow in porous media is computationally intensive; accurately modeling a single injection scenario for a large CO₂ reservoir can take several hours on a high-performance computing cluster. This becomes a bottleneck when running numerous simulations, such as in the 'Uncertainty Quantification' process. For any context involving reservoir simulation (such as CO₂ storage, Hydrogen storage, or oil extraction), well events, including their opening and closure, introduce significant numerical challenges due to their immediate effects on pressure and saturation variables. This often necessitates a drastic reduction in time step size to solve the non-linear equations derived from the discretization of the continuous mathematical model. Despite this, the impact of these specific well events tends to be relatively similar over space and time. We propose a local preconditioning strategy in the near-well region to mitigate the impact of well events. We complement the standard fully implicit solver by predicting an initialization in the near-well region of Newton's method using supervised learning. More specifically, we replace the initialization in pressure by a linear approximation obtained

through an implicit solver and we use a Fourier Neural Operator (FNO) to predict a local saturation initialization. We apply our methodology to a test case derived from a CO_2 storage in saline aquifer benchmark (SHPCO₂). We test out two different datasets for the training of the Fourier Neural Operator, the first "physical" one is generated by running simulations at a single location using the SHPCO₂ case and the second "synthetic" one is generated using a cheap synthetic dataset. We then apply both models to either the single well location dataset used to train the first model or to multiple well locations to assess the generalization capabilities of the models. Overall, we reduce the required number of Newton iterations, i.e the required number of linear system to solve; to handle a well opening by at least 27% up to 45% depending on the test case. Finally, we discuss a way to estimate the quality of a "synthetic" dataset used for training in order to make predictions on a "physical" case.

Keywords: Newton's method, Non-linear preconditioning, Neural Operator, CO_2 storage

1. Introduction

Reservoir simulation is crucial in understanding and predicting the behavior of subsurface reservoirs, aiding in efficient resource management. By modeling fluid flow, pressure changes, and other dynamic processes, it provides valuable insights for optimizing development strategies. This helps in making informed decisions, minimizing risks, and maximizing the sustainable use of natural resources such as water, geothermal energy, carbon storage or hydrogen storage.

Numerical simulations of fluid flow in porous media require substantial computational resources. Simulating a single injection scenario in a large CO_2 reservoir with precision can take hours on an HPC cluster due to the complexity and the size of the physical processes. This becomes a bottleneck when numerous simulations are needed.

Industrial solvers frequently employ fully implicit methods (FIM) (Dawson et al., 1997; Palomino Monteagudo and Firoozabadi, 2007; Liu et al., 2013; Eymard et al., 2000; LeVeque, 2002), known for their unconditional stability, which allows them to handle problems over a wide range of time scales. However, this stability requires solving a large system of equations with mixed elliptic and hyperbolic characteristics. The resulting nonlinear

20 system is typically addressed using Newton’s method, which is not guaran-
21 teed to converge (Cai et al., 2002; Skogestad et al., 2013). Achieving a so-
22 lution within a specified tolerance for a given time step involves performing
23 multiple Newton iterations until the nonlinear algebraic equations converge.
24 Yet, when faced with poor initial estimates or large time steps, the standard
25 Newton method may struggle to converge. In such cases, a common approach
26 is to retry with a smaller time step, leading to increased computational costs.

27 In recent decades, significant efforts have been made to improve the con-
28 vergence of Newton’s method for multiphase flow problems. This has in-
29 cluded modifications to the time-step size and the use of nonlinear precon-
30 ditioners (Yang et al., 2016). These preconditioners can either enhance the
31 initial guess by implementing either explicit methods (IMPES) (Chen et al.,
32 2004, 2006) or implicit sequential methods (IMPIMS) Quandalle and Savary
33 (1989); Ouaki (2013), as well as domain decomposition techniques like the
34 Additive Schwarz Preconditioned Inexact Newton method (ASPIN) (Cai and
35 Keyes, 2001; Hwang and Cai, 2005; Liu and Keyes, 2015). Another approach
36 at the Newton iteration level involves damping, where a cell-wise damping
37 factor is applied to ensure that the absolute variations in saturations stay
38 below a specified threshold. This threshold can be either heuristic Schlum-
39 berger (2013) or based on physics, such as in Trust regions (Pour et al., 2023;
40 Klemetsdal et al., 2019), which define safe updates for saturation regions us-
41 ing inflection points.

42 Well events, such as openings and closures, introduce significant numer-
43 ical challenges due to their substantial effects on pressure and saturation
44 variables. This often necessitates a considerable reduction in time step size
45 to effectively solve the nonlinear equations generated by the discretization of
46 the continuous mathematical model.

47 However, these well events tend to exhibit similar characteristics across
48 different spatial and temporal contexts. The degree of similarity between two
49 well events is influenced by several factors, including the injection conditions,
50 the reservoir state at the time of the event, boundary conditions, and the
51 porous media properties (like permeability and porosity) surrounding each
52 well. Therefore, exploring the potential of machine learning algorithms to
53 mitigate the numerical difficulties associated with well events is a promising
54 avenue for investigation.

55 Recent interest in applying machine learning to predict physical processes
56 has led to the emergence of ”Physics-Informed Deep Learning.” This ap-
57 proach involves using machine learning models to either replace or augment

58 traditional numerical algorithms while maintaining the essential constraints
59 of the physical model.

60 A common strategy to learn the behavior of physical processes is super-
61 vised learning, where a significant amount of numerical simulations are run
62 to generate a database, and the Deep Learning model is trained to match the
63 labeled data from this often expensive database. Specifically, the supervised
64 learning of mappings between infinite-dimensional function spaces using neu-
65 ral operators (Kovachki et al., 2021; Lu et al., 2019) has produced significant
66 results in learning solutions to PDEs (Takamoto et al., 2022). In particular,
67 the Fourier Neural Operator (FNO) Li et al. (2020) has shown impressive
68 performance, resulting in the development of numerous recent architectures
69 built upon it (Wen et al., 2022; Li et al., 2021; Raonić et al., 2023).

70 Physics-Informed Deep Learning has been applied to the geological se-
71 questration of CO_2 using various machine learning techniques (Seabra et al.,
72 2024; Sasal et al., 2024; Diab and Kobaisi, 2024; Kompantsev et al., 2024).
73 Validating an injection scenario requires numerous numerical simulations,
74 which can be expensive. Machine learning, used as surrogate model, has
75 the potential to expedite these simulations, enhancing decision-making and
76 enabling faster, scalable deployment of carbon capture and storage (CCS).

77 A striking example is found in Chu et al. (2023), where simulations uti-
78 lizing a FNO-based architecture achieved an acceleration of 8000 times com-
79 pared to conventional numerical solvers, while still maintaining significant
80 accuracy. Another innovative architecture, also grounded in FNO, was pre-
81 sented in Wen et al. (2023), showcasing remarkable speed enhancements in
82 CO_2 storage simulations. Furthermore, several other studies highlight con-
83 siderable reductions in the computational costs associated with CO_2 storage
84 numerical simulations (Witte et al., 2023, 2022; Jiang et al., 2023; Shokouhi
85 et al., 2021; Wen et al., 2021).

86 In many cases, the resulting machine learning model serves as a surrogate
87 that replaces the traditional solver. This enables faster evaluation of new in-
88 put parameters compared to the original numerical solver. This advantage is
89 particularly important when the numerical solver is computationally expen-
90 sive and time-consuming, as surrogate models can offer rapid approximations.
91 However, this 'black-box' approach does not guarantee accurate predictions,
92 in contrast to the reliability of numerical solvers. Thus, exploring hybrid ap-
93 proaches that incorporate machine learning models within numerical solvers
94 is valuable for enhancing speed or accuracy while preserving numerical guar-
95 antees for evaluation.

96 A recent hybrid strategy that leverages the predictive capabilities of neu-
 97 ral networks to accelerate numerical solvers is known as 'Hybrid Newton's
 98 Method' (Aghili et al., 2024; Bhattacharyya and Vyas, 2022; Huang et al.,
 99 2020). This approach employs a prediction-correction strategy in which a
 100 neural network predicts a new initial guess that is closer to the solution than
 101 the standard one, followed by applying Newton's method to refine the solu-
 102 tion to the desired accuracy. Ideally, the prediction should fall within the
 103 quadratic convergence zone of Newton's method, requiring fewer iterations
 104 for convergence.

105 We propose a direct extension of the global hybrid Newton method, as in-
 106 troduced by (Lechevallier et al., 2023; Lechevallier, 2024), which addresses the
 107 primary limitations of the original approach. This proof-of-concept method-
 108 ology aims to adapt the hybrid Newton method by predicting a local initial-
 109 ization in the near-well region to improve handling of well events. Although
 110 the physical model we use is relatively simple, it still presents numerical
 111 issues that are typical in more complex, realistic models. All reservoir sim-
 112 ulations are conducted using an open-source Python framework developed
 113 for this purpose. For pressure initialization, we employ a linear approxima-
 114 tion, while for saturation, we use a nonlinear prediction obtained through
 115 supervised learning with a Fourier Neural Operator.

116 The main advantage of this local initialization, compared to a global
 117 approach, lies in its flexibility. It can be applied to any well location, offering
 118 reduced data generation and training costs while still achieving a notable
 119 acceleration in Newton's method convergence.

120 We compare two different saturation predictors trained on two distinct
 121 datasets. The first dataset is generated using a single well location with a
 122 wide range of well events, resulting in a model called the "Single Location
 123 Well Model" (SLWM). The second dataset is designed to move toward a
 124 "Generalized Well Model" (GWM), an accurate saturation predictor capable
 125 of handling any well event, at any location, and under varying discretizations.
 126 In this study, we use a constant discretization and generate a dataset for the
 127 GWM to be trained on by simulating a synthetic small reservoir with a broad
 128 range of well events, allowing for fast data generation and training.

129 Once trained, the SLWM and GWM are integrated into the hybrid New-
 130 ton preconditioning strategy, and their numerical performances are compared
 131 based on the number of Newton iterations required to achieve convergence
 132 given a well event scenario. Our results indicate that the SLWM accelerates
 133 Newton's method by about 45% when applied within its training or testing

distributions. In contrast, the GWM achieves a 28% speedup regardless of well location, while the SLWM's acceleration drops to 15% when applied to different well locations, highlighting the GWM's superior generalization capabilities.

Finally, we introduce a quality measure based on the Wasserstein distance to assess the suitability of the synthetic dataset used to train the GWM for application to more realistic physical datasets. Our findings show that this quality measure correlates well with the observed numerical performance.

We first detail in section 2 the selected physical model and its numerical resolution. Then in section 3 we detail the construction of the hybrid Newton preconditioners and, in section 4, the data generation process that we use for training and inference of the methodology. Finally we expose in section 5 the training process of the preconditioning predictors, the numerical performances obtained on the different datasets and a discussion on means to evaluate the quality of the cheap synthetic dataset used to train the 'Generalized Well Model'.

2. Mathematical model and numerical resolution

2.1. Mathematical Model

We consider the 2D incompressible two-phase flow in porous medium models for the gas g and water w phases. Capillary pressure effects and gravity are neglected. This model can be described using conservation of mass equation (1) and Darcy's law (2) (Hubbert, 1956) for each phase, the closure of the system is ensured by (3):

$$\phi \frac{\partial}{\partial t}(S_w) + \text{div}(v_w) = q_w, \quad \phi \frac{\partial}{\partial t}(S_g) + \text{div}(v_g) = q_g, \quad (1)$$

$$v_w = -\frac{Kkr_w(S_w)}{\mu_w} \nabla P, \quad v_g = -\frac{Kkr_g(S_g)}{\mu_g} \nabla P, \quad (2)$$

$$S_g + S_w = 1, \quad (3)$$

with for $\alpha \in \{w, g\}$, the saturation S_α of the phase α , v_α the mean velocity of phase α and q_α represents the injection or production of phase α induced by sources, sinks or wells. Additionally, kr_α is the relative permeability of phase α and μ_α is the viscosity of the phase α . ϕ represents the porosity, which is assumed to remain constant over time, while K denotes the permeability,

162 which is also constant over time but may vary spatially. In this work, we use
 163 a quadratic relative permeability i.e $kr_\alpha(S_\alpha) = S_\alpha^2$.

164 2.2. Numerical resolution

165 The continuous model is discretized using a two-points finite volume spa-
 166 tial scheme on a cartesian mesh (Eymard et al., 2000; LeVeque, 2002) and an
 167 implicit Euler time scheme. The pressure and saturation are solved simulta-
 168 neously through a fully implicit scheme (Dawson et al., 1997; Palomino Mon-
 169 teagudo and Firoozabadi, 2007; Liu et al., 2013). More precisely, we consider
 170 the system (1) written in the following form:

$$\begin{cases} \phi \frac{\partial}{\partial t}(S_w) + \text{div}(v_w) = q_w, \\ \phi \frac{\partial}{\partial t}(S_g) + \text{div}(v_g) = q_g, \end{cases} \quad (4)$$

171 with $S_w + S_g = 1$, and where v_w, v_g are given by (2) and q_g, q_w are the
 172 water flow and gas flow in the wells. By summing the equations of (4) and
 173 considering $S_w + S_g = 1$, we obtain the elliptic equation in pressure:

$$\text{div}(v_t) = q_t, \quad (5)$$

174 with $v_t = v_g + v_w$, the total velocity inside the reservoir. $q_t = q_g + q_w$ is
 175 the total injection/production flow rate through the wells. The system (4)
 176 can then be rewritten as

$$\begin{cases} \text{div}(v_t) = q_t, \\ \phi \frac{\partial}{\partial t}(S_g) + \text{div}(v_g) = q_g. \end{cases} \quad (6)$$

177 We choose P and $S = S_g$ as primary unknowns and $S_w = 1 - S$ as secondary
 178 unknown.

179 The fully implicit finite volume discretization of the two equations of (4)
 180 in each cell k of the grid can be written as:

$$\begin{cases} -|k|\phi \frac{S_k^{n+1} - S_k^n}{dt} + \sum_{f \in N_k} F_{w_{k,f}}^{n+1} - \sum_{q \in N_{qk}} Q_{w_{k,q}}^{n+1} = 0, \\ |k|\phi \frac{S_k^{n+1} - S_k^n}{dt} + \sum_{f \in N_k} F_{g_{k,f}}^{n+1} - \sum_{q \in N_{qk}} Q_{g_{k,q}}^{n+1} = 0, \end{cases} \quad (7)$$

181 where $|k|$ is the measure of the cell k and $dt = t^{n+1} - t^n$ the time step.
 182 $F_{\alpha_{k,f}}^{n+1}$ is the two points flux approximation of the flux of the phase α between
 183 the cell k and f and $Q_{\alpha_{k,q}}^{n+1}$ is an approximation of the source term q_α in the
 184 cell k . N_k and Nq_k are respectively the set of adjacent cells of k and the set
 185 of wells in the cell k .
 186 Summing the equations of (7) gives the following discrete form of the system
 187 (6):

$$\begin{cases} \sum_{f \in N_k} F_{t_{k,f}}^{n+1} - \sum_{q \in Nq_k} Q_{t_{k,q}}^{n+1} = 0, \\ |k| \phi \frac{S_k^{n+1} - S_k^n}{dt} + \sum_{f \in N_k} F_{g_{k,f}}^{n+1} - \sum_{q \in Nq_k} Q_{g_{k,q}}^{n+1} = 0, \end{cases} \quad (8)$$

188 where $F_{t_{k,f}}^{n+1} = F_{w_{k,f}}^{n+1} + F_{g_{k,f}}^{n+1}$ and $Q_{t_{k,q}}^{n+1} = Q_{w_{k,q}}^{n+1} + Q_{g_{k,q}}^{n+1}$.
 189

190 The system of equations (8) can be written, with R_{α_k} the residual of
 191 phase α in cell k and $R_{t_k} = R_{w_k} + R_{g_k}$, in residual format as:

$$\begin{cases} R_{t_k}(P^{n+1}, S^{n+1}) = 0, \\ R_{g_k}(P^{n+1}, S^{n+1}) = 0. \end{cases} \quad (9)$$

192 Then, the resulting non-linear system of equations is linearized and solved
 193 using Newton's method for $X_i^{n+1} = (P_i^{n+1}, S_i^{n+1})$ and i the Newton iteration:

$$-\frac{\partial R}{\partial X}(X_i^{n+1}) \Delta X_i^{n+1} = R(X_i^{n+1}), \quad X_{i+1}^{n+1} = X_i^{n+1} + D(\Delta X_i^{n+1}), \quad (10)$$

194 with $D(\Delta X^n)$ maintains physical bounds for the unknowns. In practice,
 195 we use Appleyard damping Schlumberger (2013) of the saturation using a
 196 factor of 0.2, and clipping to ensure that the saturation remains between 0
 197 and 1. Regarding the stopping criterion, at the Newton iteration number i ,
 198 we use a criterion based on the infinity norm of the residual:

$$\max_{\alpha \in \{w, g\}} \max_{k=1, N} |R_{\alpha_k}^i| \frac{dt}{|k|} < \epsilon, \quad (11)$$

199 where $|k|$, the cell volume, aims to put into perspective the error between
 200 the large and small cells and *epsilon* the stopping criterion value. In this
 201 work, we select $\epsilon = 10^{-6}$.

202 The fully implicit scheme is unconditionally stable but large time-steps
 203 can prevent Newton’s method from converging. In practice, if Newton’s
 204 method does not converge in N_l iterations, we half the time-step and restart
 205 the step, the effort on this step, the effort already invested in this step is
 206 therefore lost. This process can be time-consuming as each Newton iteration
 207 requires to solve a linear system of equations. Also, the mechanisms allow-
 208 ing to increase the time-step are often more cautious than the mechanisms
 209 reducing the time-step.

210 Well events, such as openings, often prevent Newton’s method from con-
 211 verging within N_l iterations due to their drastic and immediate impact on
 212 pressure and saturation variables (Ahmed et al., 2022). These events can sig-
 213 nificantly increase the total simulation time, as they require a large number
 214 of iterations and time-step reductions for the solver to accurately capture the
 215 variations in saturation and pressure.

216 However, well events are often similar across space and time, and their
 217 characteristics can be described using only a few parameters. With sufficient
 218 data, it should therefore be possible to learn their behavior and improve the
 219 convergence of Newton’s method.

220 3. Methodology

221 3.1. Local hybrid Newton preconditioning

222 We propose to use the hybrid Newton methodology (Aghili et al., 2024;
 223 Bhattacharyya and Vyas, 2022), particularly we adapt the global Hybrid
 224 Newton methodology introduced in Lechevallier (2024) by predicting a local
 225 initialization in the near-well region instead of a global one as in the origi-
 226 nal method. By predicting an initialization closer to the solution than the
 227 standard one, ideally in the quadratic convergence zone, the hybrid Newton
 228 methodology serves as a preconditioning strategy to accelerate the conver-
 229 gence of Newton’s method through well events. This hybrid Newton precon-
 230 ditioning strategy can be seen as a prediction-correction strategy.

231 In practice, the system of equations (9) requires an initialization in satu-
 232 ration and pressure. We propose two different local saturation initialization,
 233 based on machine-learning. We denote by X_{init}^{n+1} Newton’s method initializa-
 234 tion at the time step $n + 1$. The standard or classic initialization is usually
 235 the solution at the previous step: $X_{init}^{n+1} := X^n = (P^n, S^n)$.

236 Now we focus on the construction of a new initial guess X_{HN}^{n+1} for the
 237 hybrid Newton, closer to the solution in the context of well opening.

238 3.2. Initial guess construction

239 We use the same initial guess construction as introduced in (Lechevallier,
240 2024; Lechevallier et al., 2023):

- 241 1. Pressure: linear approximation,
- 242 2. Saturation: Machine-learning regression.

243 The main details regarding each initial guess are described in the following
244 sections.

245 3.2.1. Pressure

246 During a well event, the pressure discontinuity is global throughout the
247 reservoir and occurs instantaneously in time. We use the solution from an
248 implicit pressure solver P_{imp} (IMP) (Sheldon and Cardwell, 1959) at step $n+1$
249 as a prediction guess. The implicit pressure solver solves the linear elliptic
250 equation (5), capturing the main global variations in pressure, but it does
251 not account for the small local variations in pressure caused by saturation
252 changes. Additionally, the implicit pressure solver only requires solving a
253 linear system of size number of grid cells, making it computationally efficient.

254 3.2.2. Saturation

255 In this physical model, saturation follows a non-linear hyperbolic behav-
256 ior, and during a well event, the saturation discontinuity is localized near
257 the well. To capture these non-linear dynamics, we employ a neural net-
258 work to predict the saturation in the near-well region, as neural networks are
259 well-suited for modeling complex non-linear behaviors. Specifically, we train
260 a Fourier Neural Operator in a supervised manner to act as the saturation
261 predictor. The predicted saturation is denoted S_{ML} .

262 In the standard approach, the initial guess is set as $X_{init}^{n+1} = X_{sol}^n =$
263 (P^n, S^n) , while the hybrid approach uses $X_{HN}^{n+1} = (P_{imp}, S_{ML})$ as the initial
264 guess. To provide a fair comparison of the impact of the saturation predictive
265 model, we evaluate the hybrid initialization against a reference initialization.
266 As such, for the reference method used in this article, the initial guess is
267 $X_{init}^{n+1} = (P_{imp}, S^n)$.

268 3.3. Neural Network architecture

269 We use a standard Fourier Neural Operator (FNO) as introduced by Li
270 et al. (2020) and visualized in figure 2. It consists, given a discretized input

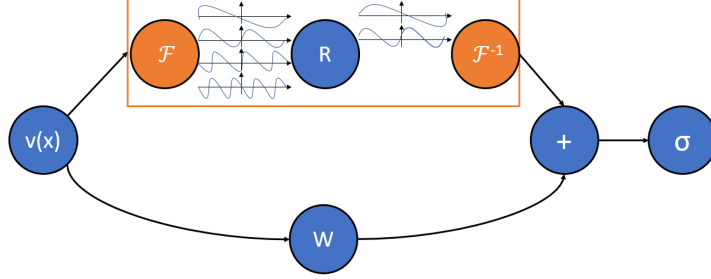


Figure 1: A Fourier layer begins with an input vector v , to which the Fourier transform \mathcal{F} is applied. A linear transformation is then performed on the lower Fourier modes, while the higher modes are filtered out. Afterward, the inverse Fourier transform \mathcal{F}^{-1} is applied. Simultaneously, a local linear transformation w is applied to the original input vector v . The outputs of these two operations are then combined, followed by the application of an activation function σ .

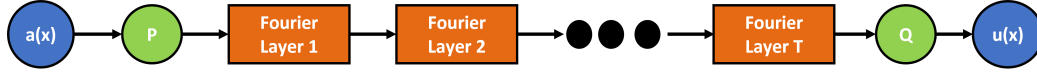


Figure 2: Fourier Neural Operator architecture as introduced by Li et al. (2020), visualization from Lechevallier (2024).

function, of a uplifting dense layer P , T Fourier layers depicted in figure 1
 and a projection dense layer Q . In this study we use $T=4$ Fourier layers. This
 architecture is well-suited for learning mappings between infinite-dimensional
 function spaces and has proven effective as a surrogate for forward reservoir
 simulation modeling (Jiang et al., 2023; Tang et al., 2023; Wen et al., 2023;
 Witte et al., 2023, 2022).

4. Test case and Dataset generation

4.1. SHPCO2 Reservoir test case

We use the SHPCO2 reservoir geometry introduced in Haeberlein (2011)
 as a practical use case as shown in figure 3. It consists in a 2D reservoir of
 size 4 750m in the X axis and 3 000m in the Y axis. We discretize the domain
 using a cartesian mesh with 95 cells in the X axis and 60 cells in the Y axis,
 each cell being of size 50×50 meters. The reservoir contains 3 permeability
 barriers of value $10^{-15} m^2$ and depicted in green on the figure. The remaining
 part of the domain has a permeability of $100 \times 10^{-15} m^2$. The porosity in
 the reservoir is constant and equal to 0.2. Then, three Dirichlet boundary

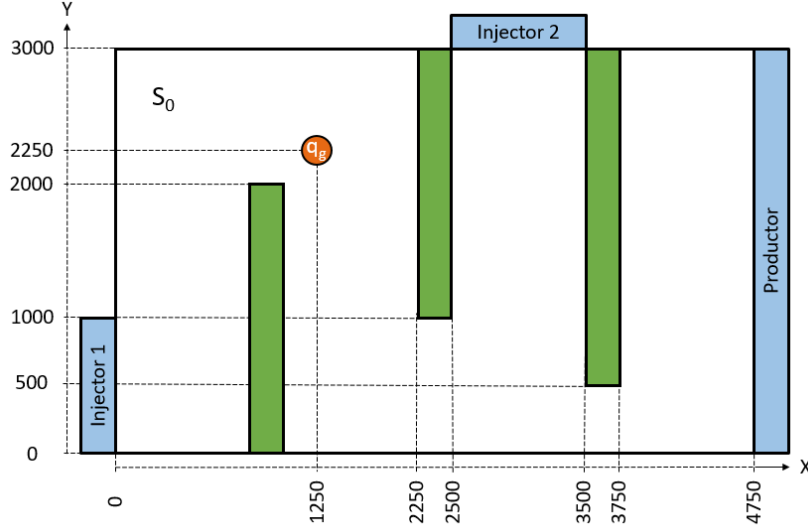


Figure 3: 2D SHPCO2 case geometry

conditions are applied and shown through blue boxes on the border of the reservoir in figure 3. The first one, 'Injector 1', has a pressure of 11 MPa and a gas saturation of 0 i.e water only. The 'Injector 2' has a pressure of 10.5MPa and a gas saturation of 0 also. Finally, the Productor has a pressure of 10MPa and no imposed saturation as the flow leaves the reservoir through this productor. The CO_2 injection well is described by an orange circle on the figure and injects only CO_2 at a rate $q_g \text{ m}^2/\text{s}$.

4.2. Dataset Generation

We consider three experimental setups corresponding to three datasets. The first setup, called 'Single well location,' involves sampling well event parameters at a fixed location in the SHPCO₂ reservoir and then running simulations for each set of parameters.

The second setup, 'Multiple well locations,' is similar to the first but samples well events at different locations in the reservoir. These first two datasets are costly to generate because they require simulations across the entire reservoir to create a local dataset near the wells.

To address this cost issue, the third setup, called 'Synthetic local domain,' focuses on a small local area. Here, we sample well events and run simulations within this smaller domain, which results in a much smaller set of non-linear

equations to solve.

The 'Single well location' and 'Synthetic local domain' datasets are used to train a neural network through supervised learning. This neural network is designed to learn the impact of a well event across a wide range of parameters. For inference, we apply these two models in a hybrid Newton's method preconditioning strategy. We then rerun all the simulations that generated the 'Single well location' and 'Multiple well locations' datasets and compare the numerical performance, specifically the number of Newton iterations needed to converge, between the standard and hybrid Newton's methods.

The local domain size for training and inference in all setups is a constant 9×9 cells, with the well at its center.

All simulations are run using the YADS open-source reservoir simulation python library available at the following GitHub repository: YADS.

4.2.1. Single well location

Well events can be described through an injection flow rate q_g , a step of time dt representing the injection duration, and reservoir properties like the initial gas saturation S_0 . We sample 5004 parameter combinations using a Latin Hypercube Sampling (LHS) (Stein, 1987) strategy within the following ranges: $q_g \in [10^{-5}, 10^{-3}]m^2/s$, $dt \in [0.1, 10]$ years and $S_0 \in [0, 0.6]$. The well location is shown on the following figure:

Generating this dataset took approximately a day on a High Performances Computing (HPC) cluster.

4.2.2. Multiple well locations

In this case, we generate 10 parameter combinations using a LHS strategy within the same range as in the single well location case. Next, we randomly select 200 well locations within possible well locations, ensuring they are not too close to a permeability barrier of the boundary. Finally we run simulations for each well location using the 10 parameter combinations, resulting in 2000 data points.

This dataset was generated in about half a day on a HPC cluster.

4.2.3. Synthetic local domain

We aim to create a synthetic dataset for training that replicates the behavior of well events for any well location or set of parameters. To achieve this, we consider a reservoir of size 9×9 with cell sizes based on SHPCO₂ discretization, with a well positioned at the center of the domain. Although

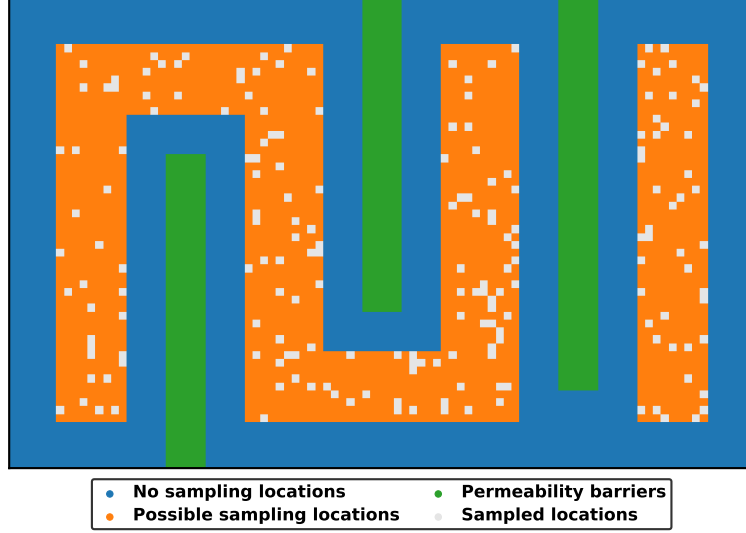


Figure 4: Possible well locations

there is only a single well location in this setup, the key difference between locations lies in the velocity field generated by the pressure field through Darcy's law. Therefore, to build a comprehensive synthetic database that captures a wide range of well event scenarios, we need to generate diverse pressure fields.

We propose using the implicit pressure solver to generate these pressure fields. However, the IMP solver requires pressure boundary conditions. Therefore, the main challenge is to generate these boundary conditions. We suggest the following protocol described in figure 5:

1. Sample N boundaries ($N \in [1, 4]$) and 1 random face on each sampled boundary,
2. Sample pressure $P \in [P_a, P_b]$ for selected face,
3. Apply correction to the sampled pressure, see below
4. Linear interpolation between the corrected pressures to create pressure boundary conditions for all boundary faces,
5. Apply Implicit Pressure Solver to generate pressure field.

The pressure correction is based on the idea that if two sampled pressures are close in location, they should also be close in value. We apply a correction

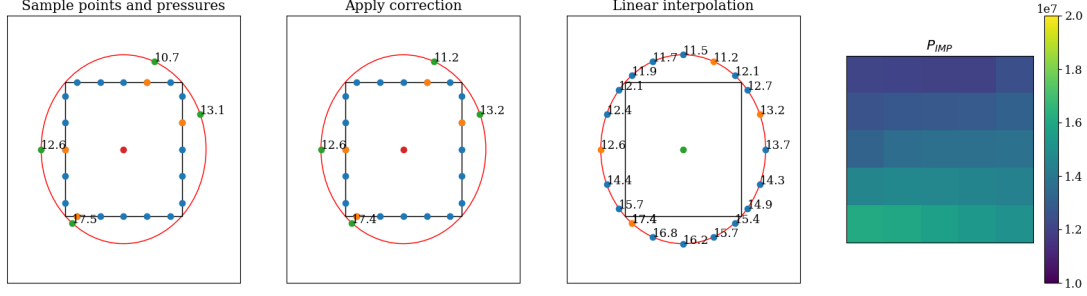


Figure 5: Pressure sampling in MPa protocol and resulting pressure field in Pa (right figure). No wells have been taken into account in this example.

matrix A to regularize the pressure P based on the cartesian distance d between their positions. The resulting corrected pressure is denoted as P^* .

$$P^* = AP, \quad A(i, j) = \text{sign}(P(X_j) - P(X_i)) \exp\left(-\left(\frac{d(X_i, X_j)}{\theta} + \frac{\mu}{|P(X_i) - P(X_j)|}\right)\right)$$

θ and μ can be interpreted respectively as characteristic distance and pressure.

Using this approach we first generate 3100 boundary condition profiles using $P_a = 10MPa$ and $P_b = 20MPa$: 100 with 1 boundary (N=1), 1000 with 2 boundaries (N=2), 1000 with N=3 and 1000 with N=4. We then increase the number of profiles to 243,100 by including all possible rotations for the profiles generated using N=2,3,4 as N=1 yields constant boundary pressures.

Next we sample well event parameter combinations using a LHS for all these pressure boundary condition profiles, using the range of values for q , dt and $S0$ as in the previous dataset. We then remove all cases where the well injection flow rate q is too low to result in an injection well and we run all simulations resulting in 111,600 data points.

This dataset was generated in under two hours on a single CPU.

To summarize, we generated three main datasets, two of them are used for training models: the single well location and the synthetic and are therefore both split in train and test sets with a ratio of 80/20. Then, two of them are used to evaluate the numerical performances of the hybrid newton

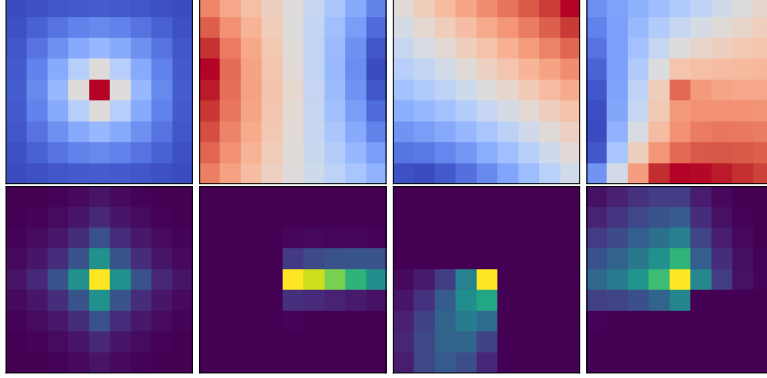


Figure 6: Synthetic dataset qualitative representation of implicit pressure fields (upper figure) and their corresponding saturation fields after a well event (lower figures).

preconditioning strategy for each machine learning model: the single well location and the multiple well location datasets.

5. Results

In this section, we present our main findings. First, we demonstrate that a neural network can learn the near-well behavior across a wide range of well events. Next, we integrate the neural networks into the hybrid Newton method and show that this preconditioning strategy significantly reduces the number of Newton iterations needed for convergence. Finally, we discuss how to assess the quality of a synthetic dataset for training, ensuring that the resulting neural network still performs well in more "physical" scenarios.

5.1. Neural Network training

We train the Fourier Neural Operator presented in 3.3 to learn the following mapping: $\{q, dt, S_0, P_{imp}\} \mapsto S$. In practice, S_0 , P_{imp} and S are used directly as $9 \times 9 \times 1$ images. The scalars q and dt are reshaped as $9 \times 9 \times 1$ images: q image is zero everywhere except at the well location and dt is constant image of value dt . After concatenation, the neural network input has the shape $9 \times 9 \times 4$ and the output has shape $9 \times 9 \times 1$.

Each dataset we train on is split in 80% for training and 20% for testing. We use the relative L2 error as a loss function and Adam optimizer. Also, we keep the model parameters corresponding to the minimum test loss value.

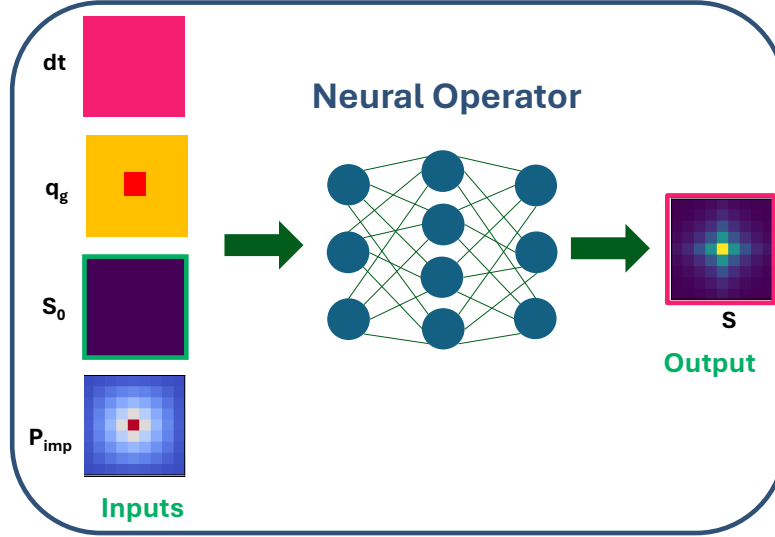


Figure 7: Input and output features of the neural network.

399 5.1.1. Single well location model

400 We train a Fourier Neural Operator on the Single well location dataset
 401 introduced in 4.2.1 resulting in the 'Single well location model' (SLWM). The
 402 training is realised on a NVIDIA A100 GPU using Adam optimizer with a
 403 learning rate of 1×10^{-4} during 1500 epochs and takes approximately 15
 404 minutes. The minimum loss value on the test set is 2.5×10^{-3} and is reached
 405 at the epoch 1368. Its corresponding train loss value is 2.5×10^{-3} . The loss
 406 evolution through epochs is shown on figure 8.

407 5.1.2. Generalized Well Model

408 We train a Fourier Neural Operator on the synthetic local domain dataset
 409 introduced in 4.2.3 resulting in the 'Generalized Well Model' (GWM). The
 410 training is realised on a NVIDIA A100 GPU using Adam optimizer with a
 411 learning rate of 5×10^{-4} during 2500 epochs and takes approximately 20
 412 minutes. The minimum loss value on the test set is 2.5×10^{-3} and is reached
 413 at the epoch 1368. Its corresponding train loss value is 2.5×10^{-3} . The loss
 414 evolution through epochs is shown on figure 9.

415 To visualize the quality of the trained neural network, we use a parity
 416 plot based on the L2 norm of the predicted saturation versus the L2 norm
 417 of true or observed saturation. A parity plot ideally shows a 45-degree line
 418 if predictions perfectly match the data. We show on figure 10 the resulting

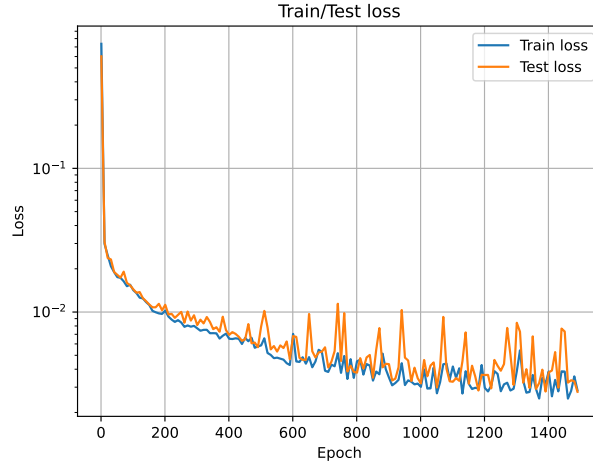


Figure 8: Relative L2 loss over epochs during supervised training of a Fourier Neural Operator on the Single Well Location dataset.

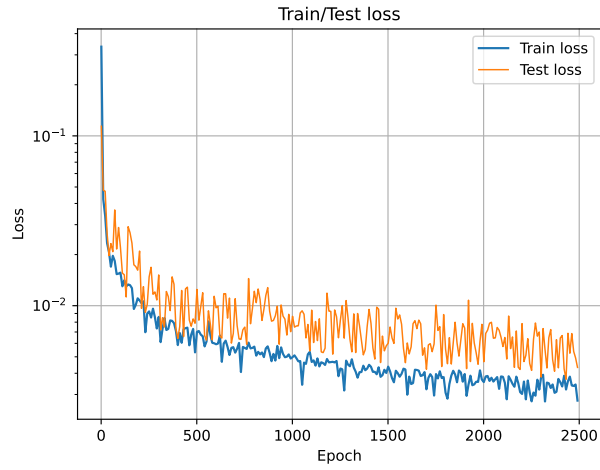


Figure 9: Relative L2 loss over epochs during supervised training of a Fourier Neural Operator on the synthetic local domain dataset.

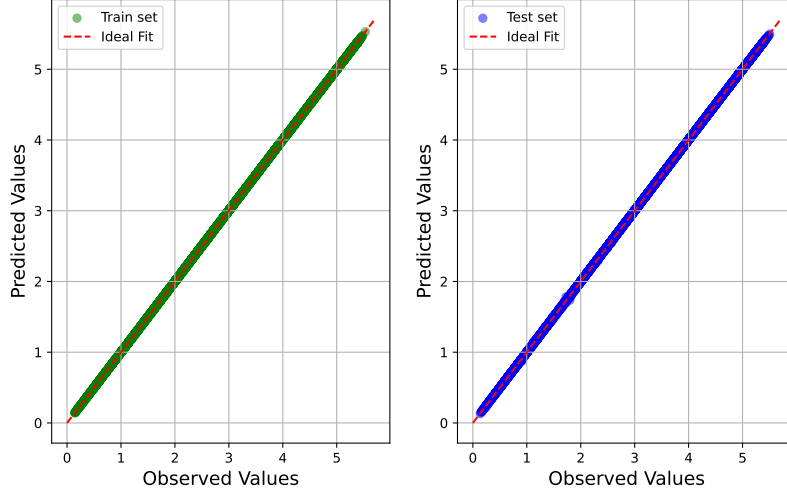


Figure 10: GWM Parity plot

plot for the train and test set, resulting in a close to ideal fit.

5.2. Newton's method performances

We apply the hybrid Newton's non-linear preconditioning strategy to the single well location dataset 4.2.1 on which we trained the single well location model (SWLM) and then to the multiple well location dataset 4.2.2 dedicated for inference. In practice, we use the SLWM and the GWM as saturation predictors and compare the number of newton iterations required to achieve convergence.

We first illustrate on a single example before showing performances over whole datasets. Figure 11 shows the a predicted local saturation (upper left), the global saturation guess (upper right) after concatenation, the solution (lower left) and the qualitative error between the global guess and the solution. We notice that the error is relatively low in the local domain compared to slightly outside the local domain. Without the saturation guess, this well event requires Newton's method 8 iterations to converge while using this saturation guess, it requires only 4 iterations. If the totality of the saturation variations were caught by the local domain, even less newton iterations may be required. Therefore, we derive two main behaviours from this example:

1. Local domain catches all variations of saturation,

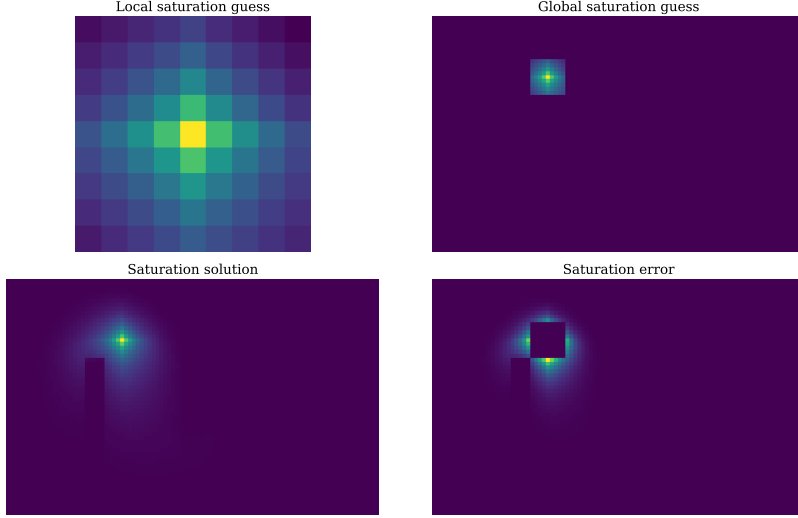


Figure 11: Local saturation machine learning prediction example (upper left), global saturation guess (upper right), saturation solution (lower left) and saturation error between the global guess and the solution (lower right).

438 2. Local domain is too small and does not catch all variations of satura-
 439 tion.

440 In the first case and given a sufficiently accurate prediction, we expect
 441 Newton’s method to be directly initialized in the quadratic convergence zone
 442 while in the second case, we may expect some constant speed up proportional
 443 to the local domain size but not an initialization in the quadratic convergence
 444 zone. These two trends are clearly visible in the next sections.

445 5.2.1. *Single well location dataset*

446 We show in figure 12 and 13 the numerical performances over the whole
 447 training and testing sets using the standard newton initialization versus the
 448 hybrid newton initialization and using respectively the SLWM (see figure
 449 12) and the GWM (see figure 13) as saturation predictors. Overall, we ob-
 450 serve that the hybrid Newton initialization requires lesser or equal number
 451 of Newton iterations to achieve convergence. We notice the two behaviours
 452 depicted in the previous section. The first corresponding to cases where the
 453 variations of saturation induced by the well events caught by the local do-
 454 main, resulting in an initialization in the quadratic convergence zone. The
 455 second one corresponds to cases where the local domain does not catch the

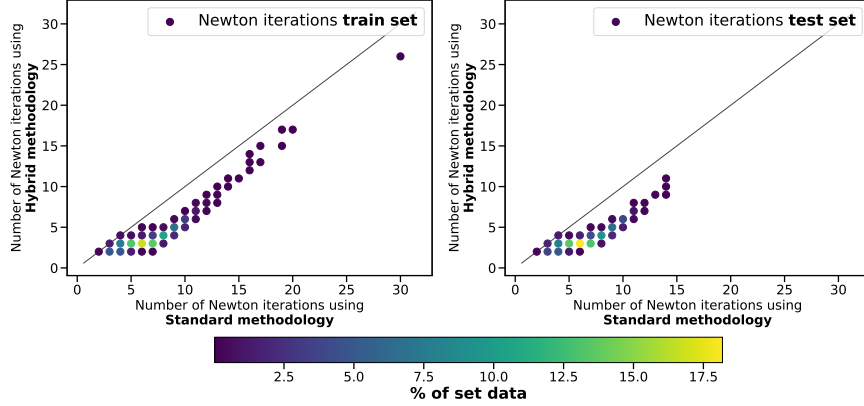


Figure 12: Scatter plots showing the number of Newton iterations required to achieve converge using standard Newton versus using hybrid Newton with SLWM on the single well location train set (left figure) and test set (right figure).

456 totality of saturation variations induced by the well event and results in a
 457 constant reduction of newton iterations required to converge. This reduction
 458 should be proportional to the size of the local domain.

459 5.2.2. Multiple well locations dataset

460 We apply the hybrid newton preconditioning strategy with the SWLM
 461 and the GWM to the multiple well locations dataset, described in 4.2.2, and
 462 compare with the standard newton method in terms of number of iterations
 463 required to converge. Results are shown in figure 14 and 15:

464 Figure 14 shows the performances obtained using the SLWM. We observe
 465 that the hybrid preconditioning strategy requires more Newton iterations
 466 than the standard for a non-negligible portion of data points in both train-
 467 ing and test sets. However, for cases requiring the most number of newton
 468 iterations, the preconditioning seems to have a positive impact. The SLWM
 469 has been trained at a single reservoir location for different well events scenar-
 470 ios, therefore when applying to other locations, the velocity field may be far
 471 from the training distribution and the predictions are far from the solution,
 472 explaining the poor performances.

473 Then for the GWM performances shown in figure 15, the two trends
 474 depicted in the previous section are clear and the Hybrid Newton method
 475 achieves convergence in less iterations than the standard Newton method for

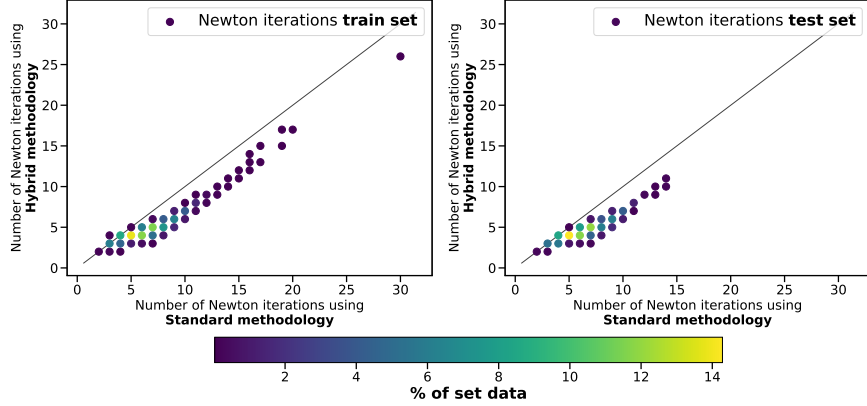


Figure 13: Scatter plots showing the number of Newton iterations required to achieve converge using standard Newton versus using hybrid Newton with GWM on the single well location train set (left figure) and test set (right figure).

most of the data points. In some cases, it requires one more iterations to converge but this is not a big issue as it only appears for cases requiring few iterations to converge already, i.e the standard and hybrid initialization are both in or close to the quadratic convergence zone. These results are important as it shows the capacity of the GWM to speed up convergence for a wide range of well events and well locations.

We summarize the speed up obtained through the hybrid Newton’s method, i.e the reduction of newton iterations required to achieve convergence, for the two models on the datasets in the following table 1:

Hybrid initialization		SLWM	GWM
Single location	Train	46%	28%
	Test	45%	27%
Multiple locations		15%	27%

Table 1: Total acceleration in percent obtained using the hybrid Newton preconditioning strategy for the SLWM and the GWM on each dataset.

We observe that the SLWM applied on the same location significantly outperforms the GWM for the train and test sets. This is expected as the SLWM has been trained at this location and on the same well event distribution. This performance can be interpreted as closest to the best possible for

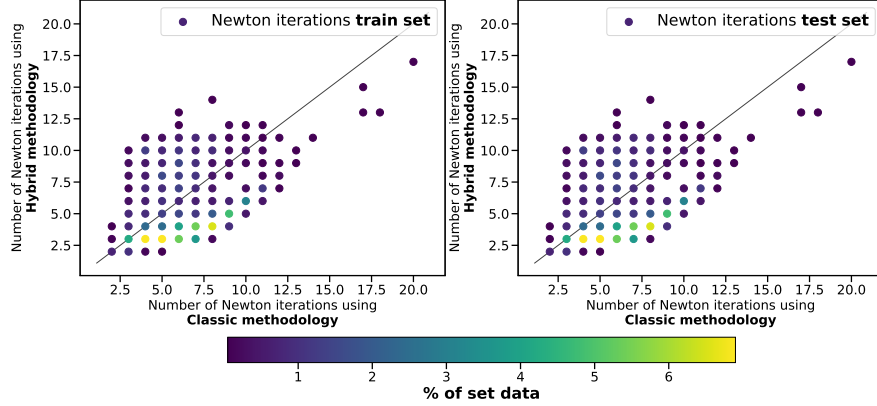


Figure 14: Scatter plots showing the number of Newton iterations required to achieve converge using standard Newton versus using hybrid Newton with SLWM on the multiple well locations train set (left figure) and test set (right figure).

489 this problem if we are able to pay the costly data generation process. Then,
 490 when applying the hybrid newton preconditioning strategy on different well
 491 locations, the GWM significantly outperforms the SWLM in terms of total
 492 number of newton iterations. Moreover, the GWM has similar performances
 493 for all datasets, denoting its 'General' property. Ideally, the GWM should
 494 have performances close to the SLWM applied on the single location dataset
 495 for all dataset. This may be achieved by improving the data generation
 496 process and by quantifying the GWM training set quality.

497 5.3. Dataset quality analysis

498 The Generalized Well Model requires to train a machine learning model
 499 on a synthetic dataset. However, how can we be sure that this synthetic
 500 distribution is close to a "physical" one, and therefore the prediction of the
 501 machine learning model is adapted to the inference case. Ideally, we would
 502 like the synthetic training distribution to cover the whole space of physical
 503 possibilities which is quite challenging considering the variability of the input
 504 parameters. We therefore propose to measure the quality of a synthetic
 505 dataset through a metric based on Wasserstein distance between the synthetic
 506 and the physical implicit pressure datasets as the pressure field contains most
 507 of the variability. We first assess on two examples if Wasserstein distance is
 508 suitable as a quality indicator, then over the whole datasets.

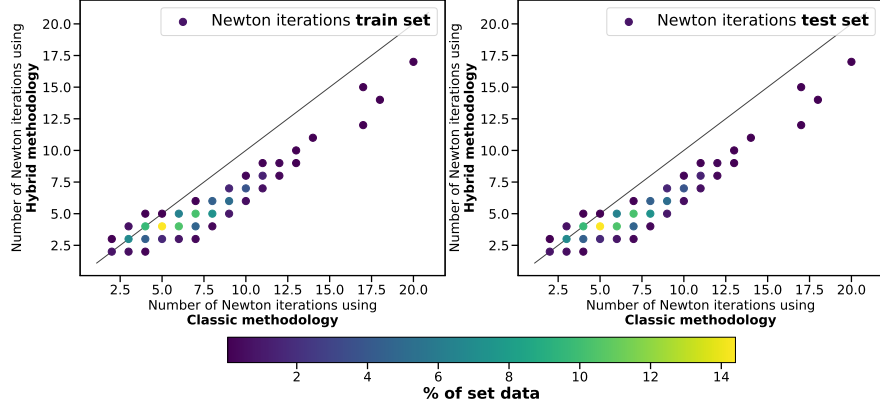


Figure 15: Scatter plots showing the number of Newton iterations required to achieve converge using standard Newton versus using hybrid Newton with GWM on the multiple well locations train set (left figure) and test set (right figure).

5.3.1. Single examples evaluation

We select two samples: one representing the best-case scenario, where the hybrid Newton methodology performs better than the standard Newton method, and another representing the worst-case scenario, where the hybrid Newton methodology requires more iterations than the standard Newton method. In practice, we calculate the minimum distance between the implicit pressure samples, P_{imp}^ϕ , from the best and worst cases (both part of the physical dataset) and all implicit pressure samples from the synthetic dataset P_{imp}^{synth} :

$$\min_{i \in P_{imp}^{synth}} (d(P_{imp}^i, P_{imp}^\phi)), \quad (12)$$

We sum up the quality measurement in table 2:

The relatively lower the quality measurement is, the better the performances should be as there exists at least one training sample in the synthetic dataset close in distance with the sample from the physical dataset. In this case, we observe that all 'Best' samples have a quality measurement lower than the 'Worst' sample quality measures. This indicates that Wasserstein distance may be suitable as a quality measure.

Dataset		Sample	Standard iters	Hybrid iters	Quality measure
Single Well location	Train	Best	10	6	4.4×10^{-3}
		Worst	3	4	8.6×10^{-3}
	Test	Best	9	5	5.2×10^{-3}
		Worst	4	4	8.5×10^{-3}
Multiple Well locations		Best	17	12	5.1×10^{-3}
		Worst	3	4	5.7×10^{-3}

Table 2: Evaluation of quality measurement using Wasserstein distance for best and worst hybrid newton cases between different implicit pressure physical datasets and the synthetic dataset.

5.3.2. Dataset evaluation

We now evaluate the quality of learning datasets with respect to inference physical datasets. We use the following formula as quality measurement with d the Wasserstein distance:

$$\max_{i \in P_{imp}^{\phi}} \left(\min_{j \in P_{imp}^{synth}} (d(P_{imp}^i, P_{imp}^j)) \right), \quad (13)$$

The resulting quality measurements are summed up in table 3:

Dataset		Training sets	
		Single well location	Synthetic
Single well location	Train	—	1.4×10^{-2}
	Test	7.5×10^{-4}	1.4×10^{-2}
Multiple well locations		8.3×10^{-2}	2.3×10^{-2}

Table 3: Evaluation of quality measurement using Wasserstein distance between different implicit pressure physical training sets and inference datasets.

We observe that for the single well location test set, the quality measurement is significantly smaller with respect to the single well location training set than for the synthetic training set. This is expected as this test set is generated from the same distribution as the single well location test set. Then regarding the multiple well locations dataset, the synthetic training set shows the smallest quality measure value. Therefore, the quality measurement matches the numerical performances in terms of total acceleration as presented in table 1.

538 6. Conclusion and perspectives

539 We introduced an innovative proof-of-concept non-linear preconditioning
540 strategy known as the local hybrid Newton strategy. This method initializes
541 Newton’s method closer to the solution compared to standard approaches by
542 employing a linear approximation in pressure and a non-linear approxima-
543 tion in saturation, the latter being obtained through a supervised training
544 of a Fourier Neural Operator. Our results demonstrated that this local pre-
545 conditioning strategy significantly mitigates the impact of well events across
546 a wide range of scenarios and well locations.

547 Additionally, we developed a methodology to reduce the cost of gener-
548 ating training data for the supervised learning of the saturation predictor.
549 This synthetic dataset, when applied to physical datasets, exhibited strong
550 performance across various well events and locations. We also introduced
551 a measurement to assess the quality of the synthetic dataset as a training
552 set. The observed Wasserstein distance quality values were consistent with
553 numerical experiment observations, suggesting its potential for further inves-
554 tinations.

555 This research is still in its early stages, and several avenues remain to
556 be explored. Firstly, the hybrid Newton preconditioning strategy should be
557 adapted and tested on more realistic physical models to fully validate its
558 potential. While local preconditioning accelerates the training phase, it still
559 requires full reservoir simulations to generate data. The synthetic dataset
560 we introduced helps speed up this data generation process, offering a path-
561 way towards the integration of artificial intelligence models into traditional
562 solvers—by generating a cost-effective dataset to expedite expensive simula-
563 tions.

564 Further research is needed to address other aspects such as the opti-
565 mal shape of the local domain, the effects of heterogeneities, and reservoir
566 discretization. Exploring these factors will enhance the applicability and
567 efficiency of our approach in real-world scenarios.

568 7. Acknowledgments

569 The authors thank Sorbonne University, IFP Energies nouvelles, DIM
570 Math Innov and NORCE for supporting financially and physically this re-
571 search work developed as part of Antoine Lechevallier Ph.D thesis work.

572 References

573 Aghili, J., Franck, E., Hild, R., Michel-Dansac, V., Vigon, V., 2024. Acceler-
 574 ating the convergence of newton’s method for nonlinear elliptic pdes using
 575 fourier neural operators. [arXiv:2403.03021](https://arxiv.org/abs/2403.03021).

576 Ahmed, E., Klemetsdal, Ø., Raynaud, X., Møyner, O., Nilsen,
 577 H.M., 2022. Adaptive Timestepping, Linearization, and A Pos-
 578 teriori Error Control for Multiphase Flow of Immiscible Flu-
 579 ids in Porous Media with Wells. *SPE Journal* , 1–21URL:
 580 <https://doi.org/10.2118/203974-PA>, doi:10.2118/203974-PA,
 581 arXiv:[https://onepetro.org/SJ/article-pdf/doi/10.2118/203974-PA/3022338/spe-2039](https://onepetro.org/SJ/article-pdf/doi/10.2118/203974-PA/3022338/spe-203974-PA/1-21.pdf)

582 Bhattacharyya, S., Vyas, A., 2022. A novel methodology for fast
 583 reservoir simulation of single-phase gas reservoirs using machine learn-
 584 ing. *Heliyon* 8, e12067. URL: [https://www.sciencedirect.com/](https://www.sciencedirect.com/science/article/pii/S2405844022033552)
 585 [science/article/pii/S2405844022033552](https://www.sciencedirect.com/science/article/pii/S2405844022033552), doi:[https://doi.org/10.](https://doi.org/10.1016/j.heliyon.2022.e12067)
 586 [1016/j.heliyon.2022.e12067](https://doi.org/10.1016/j.heliyon.2022.e12067).

587 Cai, X.c., Keyes, D., 2001. Nonlinearly preconditioned inexact newton
 588 algorithms. *SIAM Journal on Scientific Computing* 24. doi:10.1137/
 589 S106482750037620X.

590 Cai, X.C., Keyes, D.E., Marcinkowski, L., 2002. Non-linear addi-
 591 tive schwarz preconditioners and application in computational fluid
 592 dynamics. *International Journal for Numerical Methods in Flu-*
 593 *ids* 40, 1463–1470. URL: [https://onlinelibrary.wiley.com/](https://onlinelibrary.wiley.com/doi/abs/10.1002/fld.404)
 594 [doi/abs/10.1002/fld.404](https://onlinelibrary.wiley.com/doi/abs/10.1002/fld.404), doi:<https://doi.org/10.1002/fld.404>,
 595 arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/fld.404>.

596 Chen, Z., Huan, G., Li, B., 2004. An improved impes method for two-
 597 phase flow in porous media. *Transport in Porous Media* 54, 361–376.
 598 doi:10.1023/B:TIPM.0000003667.86625.15.

599 Chen, Z., Huan, G., Ma, Y., 2006. Computational Methods
 600 for Multiphase Flows in Porous Media. volume 2 of *Computa-*
 601 *tional science and engineering*. Society for Industrial and Ap-
 602 plied Mathematics. URL: [https://www.worldcat.org/title/](https://www.worldcat.org/title/computational-methods-for-multiphase-flows-in-porous-media/)
 603 [computational-methods-for-multiphase-flows-in-porous-media/](https://www.worldcat.org/title/computational-methods-for-multiphase-flows-in-porous-media/)
 604 [oclc/7346429515&referer=brief_results](https://www.worldcat.org/title/computational-methods-for-multiphase-flows-in-porous-media/oclc/7346429515&referer=brief_results).

- 605 Chu, A.K., Benson, S.M., Wen, G., 2023. Deep-learning-based flow pre-
 606 diction for co2 storage in shale–sandstone formations. *Energies*
 607 16. URL: <https://www.mdpi.com/1996-1073/16/1/246>, doi:10.3390/
 608 en16010246.
- 609 Dawson, C., Klie, H., Wheeler, M., Woodward, C., 1997. A parallel, im-
 610 plicit, cell-centered method for two-phase flow with a preconditioned new-
 611 ton–krylov solver. *Computational Geosciences* 1, 215–249. doi:10.1023/A:
 612 1011521413158.
- 613 Diab, W., Kobaisi, M.A., 2024. Temporal extrapolation and re-
 614 liable generalization via 2u-nets deep operator network (2u-
 615 deeponet) for time-dependent pdes 2024, 1–16. URL: [https://](https://www.earthdoc.org/content/papers/10.3997/2214-4609.202437052)
 616 www.earthdoc.org/content/papers/10.3997/2214-4609.202437052,
 617 doi:<https://doi.org/10.3997/2214-4609.202437052>.
- 618 Eymard, R., Thierry, G., Herbin, R., 2000. *Finite volume methods* 7. doi:10.
 619 1016/S1570-8659(00)07005-8.
- 620 Haeberlein, F., 2011. Time Space Domain Decomposition Methods for Re-
 621 active Transport — Application to CO2 Geological Storage. Theses. Uni-
 622 versité Paris-Nord - Paris XIII. URL: [https://tel.archives-ouvertes.](https://tel.archives-ouvertes.fr/tel-00634507)
 623 [fr/tel-00634507](https://tel.archives-ouvertes.fr/tel-00634507).
- 624 Huang, J., Wang, H., Yang, H., 2020. Int-deep: A deep learning ini-
 625 tialized iterative method for nonlinear problems. *Journal of Com-
 626 putational Physics* 419, 109675. URL: [https://www.sciencedirect.](https://www.sciencedirect.com/science/article/pii/S0021999120304496)
 627 [com/science/article/pii/S0021999120304496](https://www.sciencedirect.com/science/article/pii/S0021999120304496), doi:[https://doi.org/](https://doi.org/10.1016/j.jcp.2020.109675)
 628 [10.1016/j.jcp.2020.109675](https://doi.org/10.1016/j.jcp.2020.109675).
- 629 Hubbert, M.K., 1956. Darcy’s Law and the Field Equations of the
 630 Flow of Underground Fluids. *Transactions of the AIME* 207, 222–
 631 239. URL: <https://doi.org/10.2118/749-G>, doi:10.2118/749-G,
 632 arXiv:<https://onepetro.org/TRANS/article-pdf/207/01/222/2176441/spe-749-g.pdf>.
- 633 Hwang, F.N., Cai, X.C., 2005. A parallel additive schwartz preconditioned in-
 634 exact newton algorithm for incompressible navier–stokes equations. *Jour-
 635 nal of Computational Physics* 204, 666–691. doi:10.1016/j.jcp.2004.
 636 10.025.

637 Jiang, Z., Zhu, M., Li, D., Li, Q., Yuan, Y.O., Lu, L., 2023. Fourier-mionet:
638 Fourier-enhanced multiple-input neural operators for multiphase modeling
639 of geological carbon sequestration. **arXiv:2303.04778**.

640 Klemetsdal, Ø.S., Møyner, O., Lie, K.A., 2019. Robust Nonlinear New-
641 ton Solver With Adaptive Interface-Localized Trust Regions. *SPE*
642 *Journal* 24, 1576–1594. URL: <https://doi.org/10.2118/195682-PA>,
643 doi:10.2118/195682-PA. eprint: [https://onepetro.org/SJ/article-](https://onepetro.org/SJ/article-pdf/24/04/1576/2123843/spe-195682-pa.pdf)
644 [pdf/24/04/1576/2123843/spe-195682-pa.pdf](https://onepetro.org/SJ/article-pdf/24/04/1576/2123843/spe-195682-pa.pdf).

645 Kompantsev, G., Gildin, E., Rabbani, H., 2024. Novel transfer learning
646 workflow allowing flexible well configurations for physics-aware deep-
647 learning based proxy reservoir simulation models 2024, 1–18. URL:
648 [https://www.earthdoc.org/content/papers/10.3997/2214-4609.](https://www.earthdoc.org/content/papers/10.3997/2214-4609.202437046)
649 202437046, doi:<https://doi.org/10.3997/2214-4609.202437046>.

650 Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K.,
651 Stuart, A., Anandkumar, A., 2021. Neural operator: Learning maps
652 between function spaces. URL: <https://arxiv.org/abs/2108.08481>,
653 doi:10.48550/ARXIV.2108.08481.

654 Lechevallier, A., 2024. Physics Informed Deep Learning : Applications to well
655 opening and closing events. *Theses. Sorbonne Université*. URL: <https://theses.hal.science/tel-04607497>.

657 Lechevallier, A., Desroziers, S., Faney, T., Flauraud, E., Nataf, F., 2023.
658 Hybrid Newton method for the acceleration of well event handling in the
659 simulation of CO2 storage using supervised learning. URL: [https://hal.](https://hal.science/hal-04085358)
660 [science/hal-04085358](https://hal.science/hal-04085358). working paper or preprint.

661 LeVeque, R.J., 2002. *Finite Volume Methods for Hyperbolic Problems*.
662 *Cambridge Texts in Applied Mathematics*, Cambridge University Press.
663 doi:10.1017/CB09780511791253.

664 Li, Z., Kovachki, N.B., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stu-
665 art, A.M., Anandkumar, A., 2020. Fourier neural operator for paramet-
666 ric partial differential equations. *CoRR abs/2010.08895*. URL: [https:](https://arxiv.org/abs/2010.08895)
667 [//arxiv.org/abs/2010.08895](https://arxiv.org/abs/2010.08895), arXiv:2010.08895.

668 Li, Z., Zheng, H., Kovachki, N.B., Jin, D., Chen, H., Liu, B., Azizzade-
669 nesheli, K., Anandkumar, A., 2021. Physics-informed neural operator

- 670 for learning partial differential equations. CoRR abs/2111.03794. URL:
671 <https://arxiv.org/abs/2111.03794>, arXiv:2111.03794.
- 672 Liu, L., Keyes, D., 2015. Field-split preconditioned inexact newton algo-
673 rithms. SIAM Journal on Scientific Computing 37, 1388–1409. doi:10.
674 1137/140970379.
- 675 Liu, L., Keyes, D., Sun, S., 2013. Fully implicit two-phase reservoir simu-
676 lation with the additive schwarz preconditioned inexact newton method,
677 in: SPE Reservoir Characterization and Simulation Conference and Ex-
678 hibition, Society of Petroleum Engineers (SPE). doi:10.2118/166062-ms.
679 kAUST Repository Item: Exported on 2020-10-01.
- 680 Lu, L., Jin, P., Karniadakis, G.E., 2019. Deeponet: Learning nonlin-
681 ear operators for identifying differential equations based on the univer-
682 sal approximation theorem of operators. CoRR abs/1910.03193. URL:
683 <http://arxiv.org/abs/1910.03193>, arXiv:1910.03193.
- 684 Ouaki, F., 2013. Etude de schémas multi-échelles pour la simulation de
685 réservoir. Ph.D. thesis. Ecole Polytechnique X.
- 686 Palomino Monteagudo, J., Firoozabadi, A., 2007. Comparison of fully im-
687 plicit and impes formulations for simulation of water injection in fractured
688 and unfractured media. International Journal for Numerical Methods in
689 Engineering 69, 698 – 728. doi:10.1002/nme.1783.
- 690 Pour, K.M., Voskov, D., Bruhn, D., 2023. Nonlinear solver
691 based on trust region approximation for CO2 utilization and
692 storage in subsurface reservoir. Geoenery Science and En-
693 gineering 225, 211698. URL: <https://www.sciencedirect.com/science/article/pii/S2949891023002853>, doi:<https://doi.org/10.1016/j.geoen.2023.211698>.
- 696 Quandalle, P., Savary, D., 1989. An implicit in pressure and saturations
697 approach to fully compositional simulation, in: SPE Reservoir Simulation
698 Conference?, SPE. pp. SPE-18423.
- 699 Raonić, B., Molinaro, R., Ryck, T.D., Rohner, T., Bartolucci, F., Alaifari,
700 R., Mishra, S., de Bézenac, E., 2023. Convolutional neural operators for
701 robust and accurate learning of pdes. arXiv:2302.01178.

- 702 Sasal, L., Busby, D., Hadid, A., 2024. A graph neural network-based
703 approach for complex reservoirs simulation surrogate modelling 2024,
704 1–13. URL: [https://www.earthdoc.org/content/papers/10.3997/](https://www.earthdoc.org/content/papers/10.3997/2214-4609.202437073)
705 [2214-4609.202437073](https://www.earthdoc.org/content/papers/10.3997/2214-4609.202437073), doi:[https://doi.org/10.3997/2214-4609.](https://doi.org/10.3997/2214-4609.202437073)
706 [202437073](https://doi.org/10.3997/2214-4609.202437073).
- 707 Schlumberger, 2013. ECLIPSE: Technical Description 2013.2. pp. 391–391.
- 708 Seabra, G.S., Mücke, N.T., Silva, V.L.S., Voskov, D., Vossepoel, F., 2024.
709 Advancing data assimilation and uncertainty quantification for co2
710 sequestration through ai-hybrid methods 2024, 1–23. URL: [https://](https://www.earthdoc.org/content/papers/10.3997/2214-4609.202437083)
711 www.earthdoc.org/content/papers/10.3997/2214-4609.202437083,
712 doi:<https://doi.org/10.3997/2214-4609.202437083>.
- 713 Sheldon, J.W., Cardwell, W.T.J., 1959. One-dimensional, incompressible,
714 noncapillary, two-phase fluid flow in a porous medium. Transactions of
715 the AIME 216, 290–296.
- 716 Shokouhi, P., Kumar, V., Prathipati, S., Hosseini, S.A., Giles,
717 C.L., Kifer, D., 2021. Physics-informed deep learning for pre-
718 diction of co2 storage site response. Journal of Contaminant
719 Hydrology 241, 103835. URL: [https://www.sciencedirect.com/](https://www.sciencedirect.com/science/article/pii/S0169772221000747)
720 [science/article/pii/S0169772221000747](https://www.sciencedirect.com/science/article/pii/S0169772221000747), doi:[https://doi.org/10.](https://doi.org/10.1016/j.jconhyd.2021.103835)
721 [1016/j.jconhyd.2021.103835](https://doi.org/10.1016/j.jconhyd.2021.103835).
- 722 Skogestad, J.O., Keilegavlen, E., Nordbotten, J., 2013. Domain decompo-
723 sition strategies for nonlinear flow problems in porous media. Journal of
724 Computational Physics 234, 439–451. doi:[10.1016/j.jcp.2012.10.001](https://doi.org/10.1016/j.jcp.2012.10.001).
- 725 Stein, M., 1987. Large sample properties of simulations us-
726 ing latin hypercube sampling. Technometrics 29, 143–
727 151. URL: [https://www.tandfonline.com/doi/abs/10.1080/](https://www.tandfonline.com/doi/abs/10.1080/00401706.1987.10488205)
728 [00401706.1987.10488205](https://www.tandfonline.com/doi/abs/10.1080/00401706.1987.10488205), doi:[10.1080/00401706.1987.10488205](https://doi.org/10.1080/00401706.1987.10488205),
729 arXiv:<https://www.tandfonline.com/doi/pdf/10.1080/00401706.1987.10488205>.
- 730 Takamoto, M., Praditia, T., Leiteritz, R., MacKinlay, D., Alesiani, F.,
731 Pflüger, D., Niepert, M., 2022. Pdebench: An extensive benchmark
732 for scientific machine learning, in: Koyejo, S., Mohamed, S., Agarwal,
733 A., Belgrave, D., Cho, K., Oh, A. (Eds.), Advances in Neural Inform-
734 ation Processing Systems, Curran Associates, Inc.. pp. 1596–1611.

735 URL: [https://proceedings.neurips.cc/paper_files/paper/2022/](https://proceedings.neurips.cc/paper_files/paper/2022/file/0a9747136d411fb83f0cf81820d44afb-Paper-Datasets_and_)
736 [file/0a9747136d411fb83f0cf81820d44afb-Paper-Datasets_and_](https://proceedings.neurips.cc/paper_files/paper/2022/file/0a9747136d411fb83f0cf81820d44afb-Paper-Datasets_and_)
737 [Benchmarks.pdf](https://proceedings.neurips.cc/paper_files/paper/2022/file/0a9747136d411fb83f0cf81820d44afb-Paper-Datasets_and_).

738 Tang, H., Kong, Q., Morris, J.P., 2023. Multi-fidelity fourier neu-
739 ral operator for fast modeling of large-scale geological carbon storage.
740 [arXiv:2308.09113](https://arxiv.org/abs/2308.09113).

741 Wen, G., Hay, C., Benson, S.M., 2021. Ccsnet: A deep learning
742 modeling suite for co2 storage. *Advances in Water Resources* 155,
743 104009. URL: [https://www.sciencedirect.com/science/article/](https://www.sciencedirect.com/science/article/pii/S0309170821001640)
744 [pii/S0309170821001640](https://www.sciencedirect.com/science/article/pii/S0309170821001640), doi:[https://doi.org/10.1016/j.advwatres.](https://doi.org/10.1016/j.advwatres.2021.104009)
745 [2021.104009](https://doi.org/10.1016/j.advwatres.2021.104009).

746 Wen, G., Li, Z., Azizzadenesheli, K., Anandkumar, A., Benson, S.M.,
747 2022. U-fno—an enhanced fourier neural operator-based deep-learning
748 model for multiphase flow. *Advances in Water Resources* 163,
749 104180. URL: [https://www.sciencedirect.com/science/article/](https://www.sciencedirect.com/science/article/pii/S0309170822000562)
750 [pii/S0309170822000562](https://www.sciencedirect.com/science/article/pii/S0309170822000562), doi:[https://doi.org/10.1016/j.advwatres.](https://doi.org/10.1016/j.advwatres.2022.104180)
751 [2022.104180](https://doi.org/10.1016/j.advwatres.2022.104180).

752 Wen, G., Li, Z., Long, Q., Azizzadenesheli, K., Anandkumar, A., Benson,
753 S.M., 2023. Real-time high-resolution co2 geological storage prediction
754 using nested fourier neural operators. *Energy Environ. Sci.* 16, 1732–
755 1741. URL: <http://dx.doi.org/10.1039/D2EE04204E>, doi:[10.1039/](https://doi.org/10.1039/D2EE04204E)
756 [D2EE04204E](https://doi.org/10.1039/D2EE04204E).

757 Witte, P.A., Hewett, R., Chandra, R., 2022. Industry-scale co2 flow sim-
758 ulations with model-parallel fourier neural operators, in: *NeurIPS 2022*
759 *Workshop on Tackling Climate Change with Machine Learning*. URL:
760 <https://www.climatechange.ai/papers/neurips2022/78>.

761 Witte, P.A., Konuk, T., Skjetne, E., Chandra, R., 2023. Fast co2 sat-
762 uration simulations on large-scale geomodels with artificial intelligence-
763 based wavelet neural operators. *International Journal of Greenhouse*
764 *Gas Control* 126, 103880. URL: [https://www.sciencedirect.com/](https://www.sciencedirect.com/science/article/pii/S1750583623000506)
765 [science/article/pii/S1750583623000506](https://www.sciencedirect.com/science/article/pii/S1750583623000506), doi:[https://doi.org/10.](https://doi.org/10.1016/j.ijggc.2023.103880)
766 [1016/j.ijggc.2023.103880](https://doi.org/10.1016/j.ijggc.2023.103880).

767 Yang, H., Yang, C., Sun, S., 2016. Active-set reduced-space methods with
768 nonlinear elimination for two-phase flow problems in porous media. SIAM
769 Journal on Scientific Computing 38. doi:10.1137/15M1041882.