

Hybrid Physics-Data Driven Approach for Pressure Estimation in CO₂-Rich Gas Injection Operations

MATEUS A. FERNANDES^{1,2}, EDUARDO GILDIN³, and MARCIO A. SAMPAIO¹

¹Dept. of Mining and Petroleum Engineering, University of São Paulo, São Paulo, Brazil

²Dept. of Reservoir and Production Engineering, Petrobras, Santos, Brazil

³Dept. of Petroleum Engineering, Texas A&M University, College Station, TX, USA

Corresponding author: Mateus A. Fernandes (e-mail: matfernand@usp.br).

ABSTRACT

Reliable monitoring of bottom-hole pressure (BHP) in gas injection wells is essential for reservoir management, operational safety, and the success of Enhanced Oil Recovery (EOR) and Carbon Capture, Utilization, and Storage (CCUS) projects. Particularly in offshore oilfields where the carbon dioxide content in the production is relevant, its reinjection mixed with hydrocarbon gases presents unique difficulties arising from the complex physical behavior, since flow often occurs in supercritical state, and measurements from the permanent downhole gauges (PDGs) may be absent. Traditional mechanistic models and purely data-driven machine learning (ML) approaches each have limitations in handling the nonlinearities and uncertainties inherent in these systems. In this work, we address these limitations by developing a Physics-Guided Neural Network (PGNN) that combines first-principles modeling with supervised learning through a hybrid loss formulation. The methodology is validated using a comprehensive dataset from an offshore oilfield in the Brazilian Pre-salt, encompassing multiple injection cycles and varying gas compositions. Results demonstrate that the hybrid approach achieves robust BHP estimation, with error metrics around 1%, outperforming both conventional physical models and standard neural networks, especially when dealing with limited or noisy datasets. The proposed virtual sensor solution enables reliable BHP monitoring in injector wells, minimizing operational expenditures and supporting safe, efficient, and sustainable reservoir management and CO₂ storage.

INDEX TERMS

Petroleum industry, soft sensors, neural networks, carbon capture and storage, natural gas.

I. INTRODUCTION

THROUGHOUT the productive life of an oilfield, we "communicate" with the reservoir through flow and pressure measurements. These form the basis for reservoir characterization, identification of anomalous behaviors, and support studies and actions for reservoir management and operational safety [1].

Flow rates are measured at the surface, while pressures can provide valuable information when measured at different points along the production circuit. Therefore, especially in high-productivity offshore fields, we invest in installing multiple sensors (including pressure and temperature) at the bottom of wells, the Permanent Downhole Gauges (PDGs), and at the wellheads on the seabed, the Temperature and Pressure Transducers (TPTs), enabling continuous monitoring of these variables with real-time signal transmission to platforms and supervisory systems [2].

However, under harsh operational conditions, gauge failures and signal communication issues may occur over the long lifespan of these systems, which can extend to sev-

eral decades [2]. With replacement and repair costs often prohibitive, virtual sensor-based solutions become critical in many cases [3]. In the literature, several studies address bottom-hole pressure (BHP) estimation for oil and gas production wells, employing diverse machine learning (ML) techniques and addressing different production conditions [3]–[11].

On the other hand, the literature still lacks studies addressing ML for BHP estimation in injector wells. In cases involving single-phase flow of water or gas with constant composition, estimating this pressure in injector wells is a low-complexity problem, that can be solved using mechanistic approaches or regression models [9]. The situation changes when dealing with gas injection with variable composition over time, especially when involving carbon dioxide (CO₂) mixed with hydrocarbons. The behavior of CO₂-rich mixtures differs substantially from that of conventional light hydrocarbons, particularly when operating above the critical point in deep wells, what leads to greater variations in compressibility factor and viscosity, as well as increased modeling challenges.

This situation is particularly relevant in offshore fields where CO₂ content is significant alongside oil and natural gas production, as in the Brazilian Pre-salt reservoirs. In such cases, reinjection of this pollutant into the reservoir is the most viable destination, considering economic and environmental considerations [12]. For this purpose, gas separation is performed either at the platform processing plant (using semipermeable membrane systems – compact but with limited efficiency [13]) or in separation systems positioned on the seabed, such as the innovative High Pressure Separation (HISEP) [14].

The reinjection of removed CO₂ can be done mixed with excess natural gas that cannot be exported from a platform, and the operational status of wells, processing plant, and the capacity of export pipelines can result in varying mixture proportions over the field's productive life.

During the lifecycle of an oilfield, gas injection with CO₂ mixture serves a dual purpose: production gains through Enhanced Oil Recovery (EOR) processes and Carbon Capture, Utilization, and Storage (CCUS) initiatives. Both require reliable bottom-hole pressure monitoring to ensure structural integrity and drive reservoir management.

In the case of EOR, controlling gas miscibility relative to oil is critical to maximize volumetric sweep and displacement efficacy through mechanisms such as interfacial tension reduction, viscosity alteration, and wettability changes, ultimately improving the oil recovery factor [15]–[17]. From the perspective of geomechanics, we deal with strict constraints to maintain caprock integrity, which is essential to gas storage. Specifically in the Pre-salt case, the thick evaporite caprock ("salt layer") constitutes a very efficient seal, but at the same time, its ductility and significant thickness variations add great complexity to the stress regime and require additional attention in pressure monitoring [12].

CCUS projects can be combined with EOR during production phases or based on CO₂ injection into depleted reservoirs that are no longer economically viable for production – practices increasingly adopted worldwide as strategies to mitigate atmospheric greenhouse gas emissions. Such injection requires equally stringent geomechanical analysis and pressure monitoring, particularly regarding shear and tensile failure limits, and fault stability [18].

In all these cases, we can face adversities such as the absence or failure of the PDGs, especially when working with mature or previously abandoned plays, making it vital to use alternatives to estimate BHP during injection, including virtual sensors.

II. MOTIVATION AND OBJECTIVES

Obtaining bottom-hole pressure in wells via sensor virtualization can be achieved through different approaches, with traditional methods relying on equation-based modeling that represents underlying physical principles. Deviations from ideal gas behavior, the complexity of physical phenomena, and interactions in mixtures of variable composition, however, can lead to high computational costs and inaccurate

results. Purely data-driven approaches with ML, in turn, lack physical grounding and may not perform well outside their training data domains.

Hybrid approaches have emerged as solutions to these challenges, combining knowledge extraction from real-world datasets while simultaneously incorporating grounding in the known physics of phenomena in the form of terms in loss functions during training. This is the idea behind the family of Physics-Guided Machine Learning methods [19], with particular benefit when training data are scarce or limited in domain [20]. The wide variety of methods within this family has proven successful for various petroleum engineering problems, including reservoir characterization [21], prediction of fluid properties [22], and estimation of downhole variables [9], [11].

In this work, our main objective is to develop and evaluate virtual sensor solutions for estimating flowing bottom-hole pressure in gas injector wells, using hybrid physics-data-driven machine learning.

The specific objectives include: (i) validating a reference phenomenological model based on first principles for the flow, considering hydrostatic and frictional components on the pressure gradient and the influences of gas properties and Equations of State (EoS); (ii) implementing a ML architecture incorporating physical constraints through a hybrid loss function; (iii) conducting a comparative performance evaluation of the proposed model against pure physical models and conventional neural networks; (iv) correlating the characteristics of the obtained model and the problem's physics, by feature importance analysis.

We highlight as novelty the applicability of the proposed solution to wells injecting mixtures of hydrocarbon gases and carbon dioxide, which may have variable proportions over time, a complex modeling scenario that occurs in oilfields with EOR and/or CCUS projects.

The methodology is developed and validated using real-world data from a Brazilian pre-salt offshore field and can be replicated to other similar scenarios, both for retrospective studies and for real-time injection monitoring.

III. PRESSURE DROP MODELING IN GAS FLOW THROUGH WELLBORE TUBING

In this study, we address the challenge of estimating bottom-hole pressure (BHP) in gas injection wells containing mixtures of light hydrocarbons with varying CO₂ content. The problem is defined by complex nonlinear physical relationships involving phenomena such as compressible flow, heat transfer, and pronounced variations in thermophysical properties of the gases as functions of pressure, temperature, and composition.

To establish a reference baseline and also to incorporate physical constraints into neural network training, we develop a first-principles-based model. In this model, we formulate and numerically solve the momentum balance for predominantly vertical single-phase gas flow, including both hydrostatic and frictional components of pressure variation.

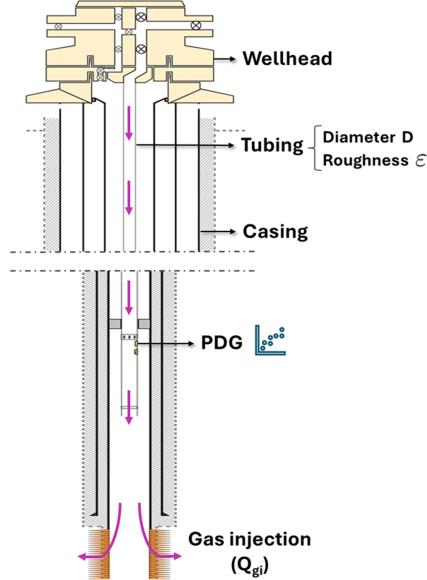


FIGURE 1. Schematic representation of a gas injection well, highlighting the components and variables relevant to our methodology.

We consider as known inputs the injection rate at standard conditions, pressure and temperature at the wellhead, the molar fractions of the gas' main components, and their fundamental tabulated properties (molar weight, critical temperatures/pressures). [17]. Besides that, we obtain wellbore tubing data: length, PDG depth, internal diameter, and roughness [23]. The dependent variable is the bottom-hole pressure. The configuration of a gas injection well, with key components, is shown in Fig. 1.

The governing equation, solving for the pressure gradient with respect to depth [23], [24], is:

$$\frac{dP}{dz} = \rho g + \frac{f}{D} \frac{\rho v^2}{2} + \rho v \frac{dv}{dz} \quad (1)$$

where ρ is the gas density, g the gravitational acceleration, f the Darcy-Weisbach friction factor, D the tubing internal diameter, and v the flow velocity.

The three terms represent, respectively, the gravitational, frictional, and accelerational components. For our problem's geometry, the first term is expected to dominate due to the significant depth (~ 2000 m), while the last term can be neglected given the characteristics of fully developed flow in constant cross-sectional area tubes [24]. Even accounting for changes in density, its low delta values (2-5%) and low velocities (~ 5 m/s) make this term approximately three orders of magnitude smaller than the others.

These components depend on parameters that vary continuously along the well depth, primarily pressure, temperature, and mixture composition, affecting viscosity (μ), density (ρ), and compressibility factor (Z). Particularly in the operational ranges encountered at the Pre-salt oilfields, mixtures with significant CO₂ content are often in the supercritical state,

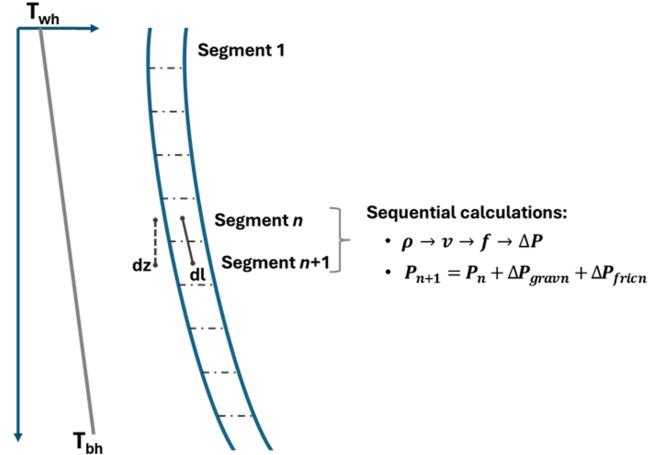


FIGURE 2. Application of the finite segment method to calculate BHP in an injector well.

which implies substantial variations in properties as P and T change.

In our solution, we pursue a compromise between accuracy and computational cost, considering that the training process of our model will require solving this equation thousands of times.

A. FINITE SEGMENT METHOD

To represent these properties with good accuracy along the well, we use numerical integration [23] through a finite segment method (FSM), dividing the wellbore into n uniform-length elements. Wells are predominantly vertical, but some have significant directional deviations; therefore, the gravitational term is based on true vertical depth (TVD), while the frictional term uses the measured depth (MD) actually traveled in the tubing.

For each segment, we sequentially calculate density (depending on Z), velocity, Reynolds number, friction factor, and pressure increment, iteratively updating local conditions. The temperature required for these calculations is obtained assuming a linear gradient along the well, with typical values for the studied field. The schematic in Fig. 2 illustrates this method, highlighting the variables and the sequence of calculation steps.

B. GAS DENSITY

For gas density calculation, we adopt the classical Equation of State (EoS) which describes the volume as a function of P and T at each segment, for real gases [25]:

$$\rho = \frac{PM_{\text{mix}}}{ZRT} \quad (2)$$

where M_{mix} is the mixture molar weight calculated as the molar-fraction-weighted average of the predominant components ($M_{\text{mix}} = \sum y_i M_i$), R the universal gas constant (8.314 J/(mol·K)), and Z the compressibility factor.

Although numerous equations of state and correlations exist in the literature, this formulation is the most widely

adopted in petroleum engineering, requiring, however, reliable determination of the compressibility factor [17], [25], [26].

C. COMPRESSIBILITY FACTOR

For the operating conditions studied, we evaluated several charts and correlations for Z , with results not always satisfactory when using simpler approaches such as Standing-Katz [25], either due to the high CO₂ content or because our operating pressure range places the fluid in or near the supercritical state. We therefore adopted the calculation proposed by Peng-Robinson [27], which is based on van der Waals' concept and accounts for attractive forces and molecular volume, essential when dealing with high pressures. The authors propose the equation of state:

$$P = \frac{RT}{V - b} - \frac{a\alpha}{V(V + b) + b(V - b)} \quad (3)$$

where a and b are substance-specific parameters, and α is a temperature-dependent correction factor, and that can be rewritten as a cubic equation for Z that can be solved iteratively [17], [27]:

$$P = Z^3 - (1-B)Z^2 + (A - 3B^2 - 2B)Z - (AB - B^2 - B^3) = 0 \quad (4)$$

given that:

$$A = \frac{aP}{R^2 T^2} \quad (5)$$

$$B = \frac{bP}{RT} \quad (6)$$

$$Z = \frac{PV}{RT}. \quad (7)$$

Although computationally more expensive, this approach yielded good performance and we validated it comparing with experimental data from the literature [28] and some available chromatographic analysis from the field.

D. FLOW VELOCITY

To calculate flow velocity, we use known surface flow rate measured at standard conditions ($Q_{gi,std}$), reinforcing that mass flow rate \dot{m} is constant all the way to the bottom-hole:

$$\dot{m} = \frac{P_{std} M_{mix} Q_{gi,std}}{RT_{std}}, \quad (8)$$

resulting in:

$$v = \frac{\dot{m}}{\rho A} \quad (9)$$

where A is the tubing cross-sectional area, adding to the previously defined variables.

E. FRICTION FACTOR

The frictional component also depends on the friction factor, which is calculated according to the current flow regime [29]. Flows are turbulent when Reynolds number $Re > 4000$ (in our case, we always operate in a turbulent regime, far above the transition range) [30]. This number is calculated as:

$$Re = \frac{\rho v D}{\mu_{mix}} \quad (10)$$

The equations for the friction factor in laminar and turbulent regimes are, respectively:

$$f_{laminar} = \frac{64}{Re} \quad (11)$$

$$f_{turbulent} = \frac{0.25}{\left[\log_{10} \left(\frac{\varepsilon/D}{3.7} + \frac{5.74}{Re^{0.9}} \right) \right]^2} \quad (12)$$

In the latter equation, we also incorporate the absolute roughness of the tubing (ε), considering typical values for commercial steel in aged wells. This equation, proposed by Swamee and Jain [31], is an explicit formulation that approximates the solution of the Colebrook-White equation (this one, implicit for f) [23]. Among several approximations available in the literature, this is expected to be the most precise for our range of (ε) [30].

F. MIXTURE VISCOSITY

Our final element in the formulation is the dynamic viscosity of the mixture, for which we employ the method proposed by Chung et al. [32], based on a semi-empirical correlation that combines temperature dependence at low pressure, high-pressure correction, and mixing rules for composition. This method is suitable for a wide range of conditions, including high pressures, and shows good alignment with experimental data from the literature [33] and from the field itself, despite the complexity of the equations with their numerous terms and coefficients, as shown in their original paper [32].

G. NUMERICAL INTEGRATION

Finally, we update pressure in each segment [23]:

$$P_{n+1} = P_n + \Delta P_{grav,n} + \Delta P_{fric,n}. \quad (13)$$

Integrating from the wellhead (Christmas tree) to the final depth, we calculate BHP.

IV. HYBRID PHYSICS-DATA DRIVEN APPROACH

Despite their technical rigor, first-principles-based solutions such as the one developed in the previous section frequently face difficulties, including a large number of variables to consider, lack of explicit representation in some equations, and limited applicability of correlations. In our specific problem, we may face imprecision in the measurement or estimation of fundamental quantities with significant impact on calculations, such as the injected gas composition obtained from analyzers and the roughness of tubing – the latter not

having direct measurement and potentially exhibiting highly variable steel corrosion rates for different wells and also along individual wellbores.

We address a problem of complex physics involving multiple coupled correlations and highly intricate formulations: the equation of state is an implicit cubic equation, viscosity correlations contain empirical terms fitted for limited ranges, numerous interdependencies exist among fluid properties, and other effects remain unmodeled, such as deposition, corrosion, and heat transfer. All of this results in observed imperfections and deviations from ideality in the outputs of a purely physics-based model.

On the other hand, we are witnessing an effervescence of purely data-driven methods that are increasingly powerful and capable of learning complex relationships. Deep learning represents the epitome of these methods, leveraging advances in computing power and algorithmic innovations to potentiate the characteristics of neural networks (NN) of pattern extraction from high-dimensional data and learning features with multiple levels of abstraction [34], [35]. However, these approaches remain heavily dependent on the quality and coverage of training data, in the absence of which neural networks are likely to underperform [19]. In a problem with the complexity of supercritical fluid flow, difficulties such as few instrumented wells, noisy data with outliers, and a lack of measurements corresponding to some expected future operating conditions may prevent us from obtaining a reliable virtual sensor.

We aim to merge the best of both worlds, using well-established equations and correlations, that are validated over decades of field operations, to provide regularization for training a supervised data-driven method. This type of integration is an excellent alternative for problems involving complex fluids when the training dataset is not comprehensive enough for a data-intensive Deep Learning model [36]. These solutions can explain phenomena at lower computational cost, learn closure laws, maintain physical consistency, and improve extrapolation beyond the training set domain [36].

We propose incorporating regularization based on the first-principles model as a soft constraint in the NN loss function during training. The numerical solver directs the network weights toward physical feasibility, while the NN learns data-driven corrections to account for the physical model limitations and measurement uncertainties.

According to Faroughi et al. [19], this characterizes a **Physics-Guided Neural Network (PGNN)**, as the physical model is not used as a hard constraint. PGNN solutions are being successfully applied to many areas related to fluid dynamics; applications include analyzing laminar and turbulent flows (e.g., in pipelines and open channels), aerodynamics, viscoelastic flows, and enhancing traditional CFD solvers [19], [36].

Since we aim to estimate pressure only at the bottom-hole rather than its distribution along the entire tubing, we can employ a PGNN with an integral approach, avoiding the need

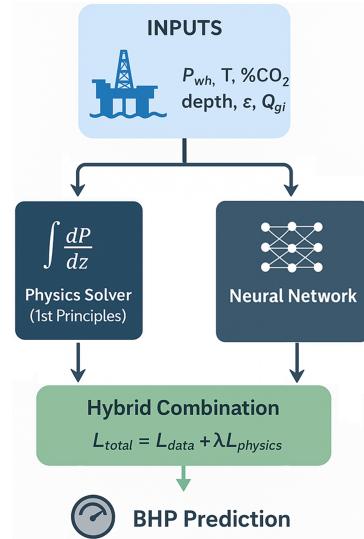


FIGURE 3. Integrated workflow for the hybrid physics-data driven model.

to compute derivatives of complex thermodynamic correlations at thousands of collocation points during each training iteration. This approach differs from Physics-Informed Neural Networks (PINNs), which use automatic differentiation to enforce Partial Differential Equations (PDEs) representative of the phenomena [20], what's not mandatory in our problem. Moreover, the strong coupling among governing equations and the implicit nature of certain correlations (e.g., Peng-Robinson's EoS) would pose additional challenges for automatic differentiation in a PINN framework.

Our approach evaluates the solver once per training example, numerically integrating the momentum balance equations from the wellhead (Christmas tree) to the bottom-hole through finite segments, computing the expected final pressure (BHP_{physics}) for the inputs given in the current sample. This pressure is then directly compared with the neural network prediction, penalizing deviations using a physics loss term, which considers the squared error:

$$\mathcal{L}_{\text{physics}} = (\text{BHP}_{\text{pred}} - \text{BHP}_{\text{physics}})^2. \quad (14)$$

In the same training example, we also compare the current network output with the corresponding measured BHP value, computing the data loss:

$$\mathcal{L}_{\text{data}} = (\text{BHP}_{\text{pred}} - \text{BHP}_{\text{measured}})^2. \quad (15)$$

We obtain the total loss by adding both, with an adjustable weight for the physics term:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \lambda \cdot \mathcal{L}_{\text{physics}} \quad (16)$$

where λ is a hyperparameter that controls the relative importance of physical consistency versus data fidelity.

This hybrid PGNN formulation is shown schematically in Fig. 3 and offers important advantages for our application:

TABLE 1. List of variables used in our methodology

Variable	Node	Unit
Bottom-hole Pressure (BHP)	Bottom-hole (PDG)	kgf/cm ²
Wellhead Pressure (WHP)	Wellhead (TPT)	kgf/cm ²
Wellhead Temperature (WHT)	Wellhead (TPT)	°C
Gas injection rate	Platform	Mm ³ /d
Gas composition	Platform	molar fraction
PDG depth	-	m

- The physics solver guarantees a physically consistent baseline even in regions of sparse data coverage,
- The neural network compensates for model simplifications and uncertain parameters (e.g., tubing roughness, actual gas composition),
- The approach is computationally efficient, requiring only one solver evaluation per training sample rather than gradient computations at multiple collocation points as in classical PINNs.

V. IMPLEMENTATION

In this section, we detail the implementation of our methodology, from data acquisition to the NN architecture and the evaluation metrics.

A. DATA ACQUISITION

This work is based on real operational data from gas injection wells currently in operation in one of the Pre-salt oilfields off the Brazilian coast. The dataset comprises 11 different injection cycles carried out in a group of wells from a single platform over a period of 8 years. We use hourly averages of the measurements, which are available from the supervisory control system of the operating company and can be loaded into our *Python* application using an appropriate interface. The variables initially used are listed in Table 1, along with their respective measurement nodes and units.

In addition to data from pressure and temperature gauges, we have the composition of the injected gas, determined by analyzers installed upstream of the injection compressor. Those measurements, however, are not as accurate as chromatography, which we use as a reference, though they are far sparser, usually made quarterly.

We also obtained the constant values for the molar weight of the main gas components [25], the tubing internal diameter ($D = 6$ in), and an estimate of the absolute roughness of the tubing, considering it is composed of industrial steel, but degraded by wear and corrosion after some years of operation under harsh conditions ($\varepsilon = 1.0$ mm).

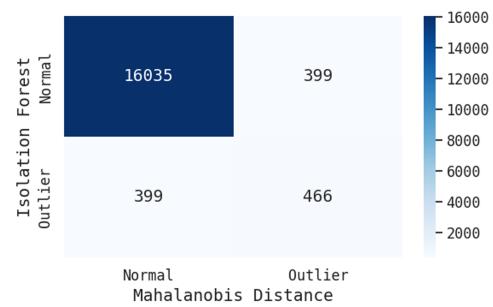
Our dataset comprises 21,073 samples, and we present a statistical summary of the measured variables in Table 2.

B. DATA CONDITIONING

Initially, we set aside part of our dataset for a blind test at the end of the project phase. Those correspond to two of the injection cycles, not using a random division, which is essential to represent real usage conditions for the virtual sensor, given that the cycles can have different characteristics

TABLE 2. Statistical summary of the dataset.

Stats.	Q _{gi} (M m ³ /d)	CO ₂ (%)	WHP (kgf/cm ²)	WHT (°C)	BHP (kgf/cm ²)
mean	2184.4	38.57	499.2	30.61	643.7
std-dv	227.7	4.74	36.5	6.71	44.2
min	887.4	17.46	429.5	8.43	563.9
25%	2117.4	35.68	463.6	26.42	599.7
50%	2208.5	37.00	512.1	32.04	653.4
75%	2302.2	41.00	536.6	34.05	685.2
max	3561.6	78.95	591.1	81.10	708.2

**FIGURE 4.** Confusion matrix showing the overlap of the multivariate outlier detection methods.

regarding average pressure, gas rate, and CO₂ concentration, reflecting a particular moment in the operational setup.

The remaining samples constitute our training/validation dataset, which serves as the foundation for our data conditioning procedures and the 5-fold cross-validation performed during the model definition and tuning phases.

Once we prevented information leakage from the test dataset, we performed outlier detection and removal. We started by checking individual variables, basing our outlier filter on the Inter-Quartile Range (IQR), a non-parametric method using statistical distribution [3], [37].

As we can face imprecision and some temporal mismatch among variables measured at different nodes—e.g., those caused by “frozen” sensor data—we included a combination of multivariable outlier detection techniques to increase rigor. We used two different methods:

- Mahalanobis distance: Outliers can be identified by selecting data points with the largest distances from the group center; in this case, using a calculation which accounts for correlations and variances among variables by scaling each variable’s contribution appropriately [37].
- Isolation Forest: This method consists of an unsupervised machine learning algorithm where an ensemble of decision trees is built for the dataset, and the anomalies are the instances that require fewer splits to be isolated (shortest average path lengths) [38].

We defined the 95% percentile threshold for both multivariate methods and the results had some overlap, as shown in the confusion matrix in Fig. 4. We filtered out samples marked as outliers by either method, the union of sets. At the end of this step, our training/validation dataset contained 16,035 samples.

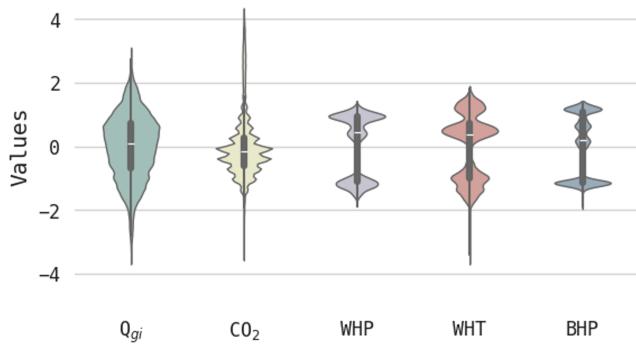


FIGURE 5. Violin plots with the distribution of the measured variables after data conditioning, applied to our training/validation dataset.

Finally, we performed data normalization, evaluating both the Gaussian and MinMax methods. While the former is better for handling broad value ranges, the latter limits all variables to a defined interval, which is better for the NN, though it can compress relevant values into a narrow range [39]. Violin plots showing the distribution of the variables with standard normalization are shown in Fig. 5.

C. MODEL ARCHITECTURE

The implementation of the physics-based model closely follows the step-by-step procedure we described in Section III, using discrete integration, equation of state, and correlations to compute gas properties for each segment and update pressure along the wellbore.

When the physics-based model is evaluated in isolation, and only a single integration is required, we use 100 segments and solve the equations of state through a full iterative procedure. During PGNN training, however, these calculations must be repeated many thousands of times, since the integration is performed for every sample at each training epoch. To reduce the associated computational cost, we adopted a coarser discretization with 10 segments, noting that the resulting approximation error is small and can be compensated by the data-driven component of the model.

An additional measure was the use of a precomputed lookup table for the gas compressibility factor Z , replacing the iterative solution of the Peng-Robinson EoS at each integration step. To do that, we calculated 1 million Z values, forming a multidimensional grid spanning the pressure, temperature, and gas composition ranges observed in our dataset. Then, we can interpolate from the grid as necessary. This strategy reduces the computational cost of evaluating Z by approximately two orders of magnitude, from an average of 178.3 μ s to 0.9 μ s, with virtually no differences in the results.

The Machine Learning model we chose is a Neural Network, based on the traditional Multilayer Perceptron (MLP) configuration, composed of a sequence of fully connected layers. This is a classical solution, robust, validated in both academia and industry, and offers the ability to solve nonlinear problems and versatility to adapt to our problem's complexity [35].

The flexibility of the NN's architecture allows us to find a good balance between specialization (good accuracy) and generalization (avoiding overfitting). In order to accomplish that, we analyze the most relevant hyperparameters defining our model [35], [40], so we can find the most appropriate for our purposes:

- Number of layers: Number of sequential connections in a network. We tested configurations with 1 to 3 layers.
- Number of neurons: Quantity of nodes in each one of the fully connected layers. We evaluated from 20 to 128, either in uniform or non-uniform distribution.
- Activation function: The mathematical function applied to the weighted sum of the inputs at each node. We tested Tanh, ReLU, and Sigmoid.
- Learning rate: Size of the “step” while correcting the weights during training. We employ an adaptive learning rate, reducing its value when progress slows, for greater efficiency.
- Dropout: A strategy to mitigate overfitting, by randomly eliminating connections (setting their weights to zero). We evaluated values ranging between 0 and 50%.
- Weight decay: Another way to prevent overfitting, by applying a regularization factor that penalizes large connection weights. We tested from 10^{-4} to 10^{-1} .
- Physics Loss weight (λ): In our proposed hybrid loss function (16), controls the balance between adhering to the physics modeling and fitting to the measured data. We analyzed values from 0.01 to 1.0. The purely data-driven model utilizes $\lambda = 0$.
- Optimizer: We used the default Adam algorithm.

The models were implemented in *Python*, using as main libraries *PyTorch* and *Scikit-Learn*. We executed them in a cluster with 48 CPUs and 248 GB RAM.

D. EVALUATION METRICS

We evaluate the method's performance using the metrics Mean Absolute Percentage Error (MAPE), Symmetric Mean Absolute Percentage Error (SMAPE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2), according to the following equations:

$$MAPE = \frac{1}{n} \cdot \sum_j \left| \frac{y_j - \hat{y}_j}{y_j} \right| \quad (17)$$

$$SMAPE = \frac{1}{n} \cdot \sum_j \frac{|y_j - \hat{y}_j|}{\frac{(y_j + \hat{y}_j)}{2}} \quad (18)$$

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_j (y_j - \hat{y}_j)^2} \quad (19)$$

$$R^2 = 1 - \frac{\sum_j (y_j - \hat{y}_j)^2}{\sum_j (y_j - \bar{y}_j)^2} \quad (20)$$

where y_j are the predicted values, \hat{y}_j the actual values, and \bar{y}_j the average values of the dependent variable [41], BHP in our case.

TABLE 3. Results obtained with the most representative models for the 5-fold validation.

Model	RMSE	MAE	R2	MAPE	SMAPE
Physics	11.397	9.009	0.930	1.417	1.402
NN	9.031	6.680	0.9421	1.036	1.028
PGNN	8.796	6.177	0.9561	0.963	0.959

VI. RESULTS AND DISCUSSION

In this section, we present the results of our experiments, including model validation, blind tests, diagnostic plots, and importance of variables study, along with a critical analysis.

A. MODEL TUNING AND VALIDATION

The first evaluation step was conducted to define the neural network architecture and tune the hyperparameters using a 5-fold cross-validation approach. In this procedure, the 16,035 samples from the training/validation dataset were divided into five subsets, preserving their temporal relation and considering the injection cycles. We then performed a sequence of training rounds, in which the NN was trained using four subsets and evaluated on the fifth. We repeated this process for each evaluated hyperparameter combination, and computed the corresponding performance metrics (as described in Section V.D) from the concatenated results across all folds. Due to the stochastic nature of neural network training, we performed multiple runs for each configuration to assess consistency.

During the tuning process, a relatively low sensitivity to the hyperparameter choices was observed. The best performance was obtained with the following configuration:

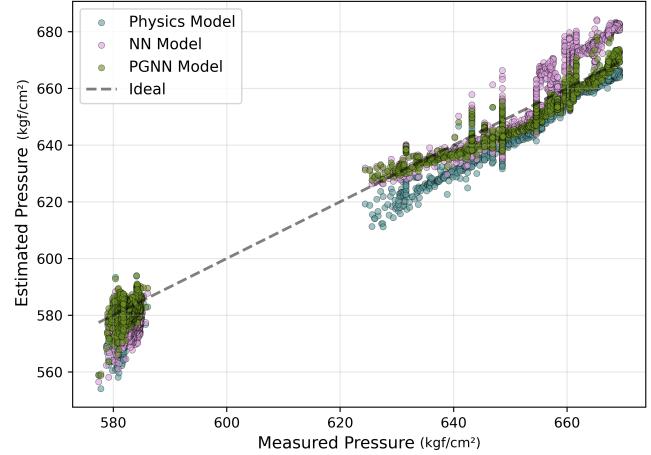
- Fully connected network with two hidden layers of 64 neurons each.
- Standard scaling of input features.
- Hyperbolic tangent activation function.
- Dropout rate of 30% and weight decay of 10^{-3} .
- Adaptive learning rate, starting at 10^{-3} .
- Early stop, configured to monitor validation loss to prevent overfitting.

A fundamental part of the proposed modeling framework is the selection of the weighting factor λ , which controls the balance between the physics-guided and data-driven loss terms, as defined in 16. Using the base NN model, we plotted a sensitivity analysis with respect to this parameter, supporting the choice of $\lambda = 0.3$, what implies a weight for the physics loss that is 30% of the weight of the data-driven loss.

Table 3 summarizes the error metrics obtained from the 5-fold cross-validation using the tuned PGNN model. For comparison, results are also reported for the standalone physics-based model implemented via the finite segment method and for a purely data-driven neural network with the same architecture. The improved performance achieved by the proposed approach validates the selected configuration, which is subsequently employed in the blind test.

TABLE 4. Results obtained with the most representative models for the blind test.

Model	RMSE	MAE	R2	MAPE	SMAPE
Physics	6.736	5.999	0.965	0.997	0.991
NN	7.067	5.585	0.961	0.889	0.887
PGNN	4.643	3.802	0.984	0.611	0.612

**FIGURE 6.** Correlation between estimated and measured pressure for the Physics-only, standard NN, and hybrid PGNN models. The dotted line represents perfect correlation.

B. BLIND TEST

In the next stage, the tuned model from the previous section was used to evaluate performance on the data we set aside for blind testing. The normalization and outlier-filtering parameters previously defined for the training/validation dataset were applied to this test set. The neural network was then trained using the whole training/validation dataset, this time in a single-fold setup.

This dataset includes data from two different injection cycles, deliberately selected as the most recent ones for a pair of wells. This choice allows evaluating model performance under conditions that occurred after the training cutoff date, which often correspond to situations at our domain's boundary (or even beyond it) due to concept drift, a phenomenon particularly relevant in petroleum production [3]. For instance, the current and future pressure distribution and fluid saturation in the reservoir may differ substantially from historical conditions.

The resulting performance metrics are reported in Table 4 and complemented by the plot shown in Fig. 6 for clearer interpretation. Once again, the hybrid PGNN model exhibits better performance, effectively combining the strengths of both the physics-based and data-driven components and approaching the ideal theoretical BHP estimate. The overall performance remained consistent.

Finally, we present in Figs. 7 and 8 the BHP time series for the wells in the test set, enabling a better alignment with how we would use it in practical field applications. In the first case, despite some noise as consequence of the dynamic characteristics of the input variables, the performance gain

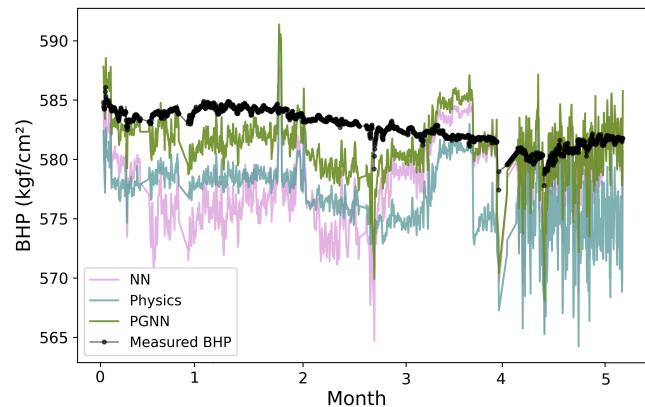


FIGURE 7. Time series of measured pressures and corresponding estimates from the Physics-only, standard NN, and hybrid PGNN models for the first well in the blind test set.

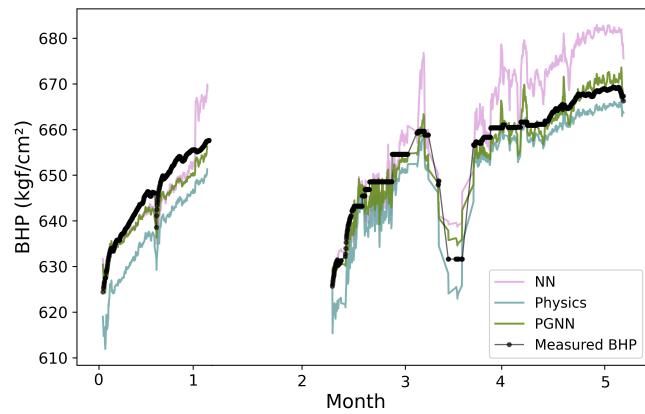


FIGURE 8. Time series of measured pressures and corresponding estimates from the Physics-only, standard NN, and hybrid PGNN models for the second well in the blind test set.

and quality of fit are evident, with error and oscillations always under 1%. In the second case, the quality remains, and the model responded adequately even after a gap in the well's operation and a return with varying conditions.

C. DISCUSSION

Based on the presented results, we bring some key points for discussion.

1) Performance improvement: In almost all the evaluated scenarios, our PGNN consistently outperformed the baseline approaches in estimating BHP during gas injection. Although the absolute improvement is modest given the baselines' already strong performance, relative reductions exceeding 30% were observed in both absolute and squared error metrics. Such improvements are practically relevant for applications involving the multifaceted dynamics of gas mixtures injected into underground rock formations.

2) Use of a complete physics model: The adoption of a complete physics-based model proved beneficial. Simplified formulations that did not include friction effects and pipe roughness, or that used simpler EoS were insufficient to properly represent the problem and were therefore discarded

during preliminary analyses.

3) Training and execution time: It is important to highlight that training with the full physics model is computationally demanding, due to the need to solve the finite-segment iterative process (including auxiliary calculations such as for Z , f , and μ) for each sample. Even with the proposed acceleration strategy for the EoS calculations, a single training epoch may take 1-3 minutes on the available hardware. Once trained, however, the neural network operates like a standard data-driven model, as the learned physical knowledge is embedded in its weights. This leverages our model as a real-time tool, given that inference is performed in milliseconds and the sensor data acquisition frequency in offshore supervisory systems is typically 1 Hz.

4) Applicability limits and bias: Observing the error magnitude as a function of pressure and CO₂ content (Fig. 9), no systematic bias is observed. Nevertheless, certain operating conditions remain more challenging for accurate estimation, mostly the rarest in the training dataset, and associated with temporary or transitional conditions. Model performance could be improved by acquiring more training data or further refining the physics for such conditions. We can also notice from the plot that the number of samples with errors greater than 1% is negligible.

5) Feature Importance: To verify the correspondence between the proposed hybrid model and the underlying physical behavior of the problem, we conducted a variable importance analysis using SHAP (SHapley Additive exPlanations), a model-agnostic solution that is adequate for our case. This approach is based on cooperative game theory and adopts an additive feature attribution formulation, in which model inferences are decomposed into individual feature contributions relative to a reference value [42]. Fig. 10 presents the SHAP summary plot, using a representative group of samples. Additionally, we bring in Fig. 11 the variables influence ranking, based on the mean absolute SHAP values. The results show the clear and expected dominance of wellhead pressure in the ranking of influences. An interesting observation is the broader distribution of higher-value points, which may indicate nonlinear modulation by flow-dependent pressure relations. CO₂ content ranks second, with positive values associated with higher BHP. Depth presents more clustered SHAP values, which is consistent with its primary role in the hydrostatic pressure term. Another relevant observation is that the gas rate contributes inversely, as a result of increasing the friction term. Overall, these results are closely aligned with the physical interpretation.

6) Hybridization as a key factor: The hybrid nature of the PGNN plays a central role in the obtained performance. The physics-based component provides robustness under operating conditions underrepresented in the training dataset (which may arise from concept drift) and significantly reduces the variance of predictions, leading to smoother estimates. On the other hand, measured data can compensate for limitations of the first-principles model, such as simplifications in gas composition and the evolution of pipe roughness caused by

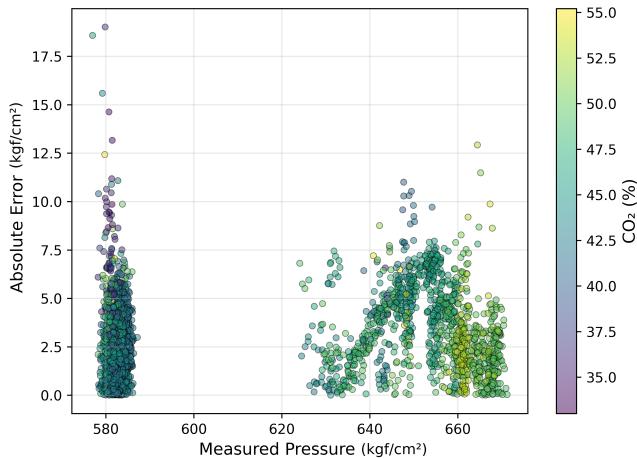


FIGURE 9. Diagnostic plot illustrating the absolute error of the PGNN model as a function of real pressure (x-axis), with CO₂ content represented by the color scale.

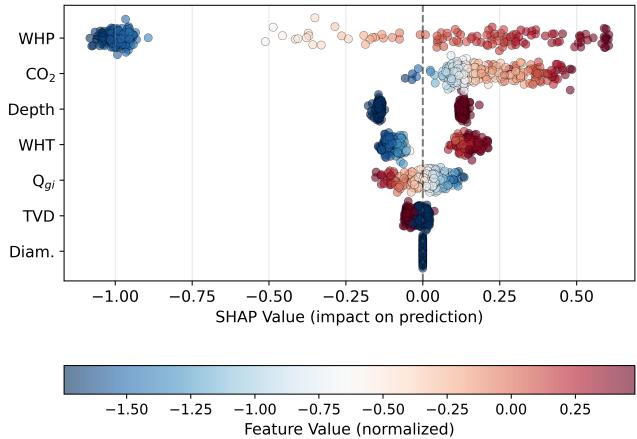


FIGURE 10. SHAP summary plot showing the global influence of input variables on BHP predictions. Each point represents a sample, with color indicating the feature value and horizontal position corresponding to its contribution to the model output.

corrosion, paraffin deposition, or scaling—effects that cannot be directly measured or continuously updated in the physical model parameters.

VII. CONCLUSIONS

In this work, we successfully implemented and evaluated a Physics-Guided Neural Network intended to estimate the flowing bottom-hole pressure in gas injection wells operating with mixtures of hydrocarbons and CO₂. The goal was to obtain a virtual sensor capable of substituting absent or faulty PDGs.

We addressed a practical application that can benefit the Oil & Gas industry and CCUS initiatives, and offered a new methodology to address a relevant problem that still lacks innovative approaches.

We evaluated our framework in actual field conditions, using data from various gas injection cycles in one of the offshore platforms operating in the Brazilian Pre-salt basin.

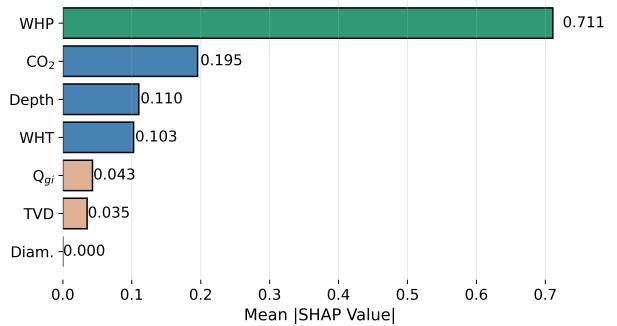


FIGURE 11. Ranking of feature importance obtained through SHAP values, with higher ones indicating greater influence on the output.

Our results show a relevant benefit of incorporating known physics into NN training alongside measured data, which can be valuable in cases where the latter are scarce or domain-limited. Error metrics around 1% indicate robustness and provide confidence for application in practical situations, including real-time monitoring.

The virtual sensor ensures the availability of essential BHP data for injection monitoring, reducing expenditures with PDG installation or replacement during both the productive life of an oilfield and the injection phase of CO₂ storage projects in old or abandoned wells/reservoirs. In these cases, pressure control is essential for safe and efficient operations, leveraging mitigation of geological and operational risks, EOR optimization, and maximization of CO₂ retention, with the outcomes of increasing project value and potentiating decarbonization.

ACKNOWLEDGMENT

The authors would like to sincerely thank the Laboratory of Reservoir Simulation and Management (LASG) at the University of São Paulo (USP) and Texas A&M University for their invaluable support throughout this work, and also acknowledge Petróleo Brasileiro S.A. (PETROBRAS) for granting permission to share this research and for their ongoing encouragement in the pursuit of innovative solutions.

REFERENCES

- [1] R. N. Horne, “Listening to the reservoir—interpreting data from permanent downhole gauges,” *Journal of Petroleum Technology*, vol. 59, no. 12, pp. 78–86, 2007.
- [2] M. J. Martins de Souza, A. H. Martins Silva, J. R. Pelaquim Mendes, and M. A. Jaculli, “Reliability of permanent downhole systems: Minimum sample and quality index,” *Petroleum Research*, vol. 9, no. 3, pp. 472–480, 2024.
- [3] M. A. Fernandes, E. Gildin, and M. A. Sampaio, “Machine learning-based virtual sensor for bottom-hole pressure estimation in petroleum wells,” *Eng*, vol. 6, no. 11, 2025.
- [4] M. A. Fernandes, E. Gildin, and M. A. Sampaio, “Data-driven estimation of flowing bottom-hole pressure in petroleum wells using long short-term memory,” in *2024 International Conference on Machine Learning and Applications (ICMLA)*, pp. 1530–1537, IEEE, 2024.
- [5] H. Zheng, B. Lin, J. Jiang, Y. Jin, and L. Peng, “Knowledge-guided machine learning method for downhole gauge record prediction in deep water gas field,” in *Offshore Technology Conference Asia*, vol. Offshore Technology Conference Asia, p. D031S020R004, 02 2024.

- [6] O. E. Agwu, S. Alatefi, A. Alkouh, and R. R. Suppiah, "Modelling the flowing bottom hole pressure of oil and gas wells using multivariate adaptive regression splines," *Journal of Petroleum Exploration and Production Technology*, vol. 15, no. 2, p. 22, 2025.
- [7] J. He, M. Avent, M. Muller, and L. Bordessa, "Downhole pressure prediction for deepwater gas reservoirs using physics-based and machine learning models," *SPE Journal*, vol. 28, pp. 371–380, 02 2023.
- [8] G. A. Sowek, R. Y. T. L. Chira, F. P. Munerato, and A. J. G. Exposito, "Predicting bottom-hole pressure in wells using machine learning: A cost-effective solution for missing gauge data," in *Offshore Technology Conference Brasil*, vol. OTC Brasil, p. D021S016R001, 10 2025.
- [9] M. Jin and H. Emami-Meybodi, "An integrated machine learning algorithm for unconventional flowing bottomhole pressure prediction during dynamic gas lift operation," *SPE Journal*, vol. 30, pp. 2726–2737, 05 2025.
- [10] N. A. Sami and D. S. Ibrahim, "Forecasting multiphase flowing bottom-hole pressure of vertical oil wells using three machine learning techniques," *Petroleum Research*, vol. 6, no. 4, pp. 417–422, 2021.
- [11] D. Molinari and S. Sankaran, "Merging physics and data-driven methods for field-wide bottomhole pressure estimation in unconventional wells," in *SPE/AAPG/SEG Unconventional Resources Technology Conference*, vol. SPE/AAPG/SEG Unconventional Resources Technology Conference, p. D031S074R004, 07 2021.
- [12] J. P. Pereira Nunes, G. S. Seabra, and L. C. de Sousa, "A review of co₂-injection projects in the brazilian pre-salt — storage capacity and geomechanical constraints," *International Journal of Greenhouse Gas Control*, vol. 137, p. 104232, 2024.
- [13] Z. Dai and L. Deng, "Membranes for co₂ capture and separation: Progress in research and development for industrial applications," *Separation and Purification Technology*, vol. 335, p. 126022, 2024.
- [14] F. M. Passarelli, A. M. Bidart, M. J. Spelta, D. A. Moura, N. E. Silva, A. J. Vieira, A. M. Oliveira, and M. A. Roberto, "Revolutionizing production in high co₂ and high gor fields, paving the way to a low carbon world," in *OTC Offshore Technology Conference*, vol. Offshore Technology Conference, p. D031S034R003, 05 2024.
- [15] M. Hasanzadeh, A. Izadpanahi, and A. Naghizadeh, "Carbon dioxide huff-n-puff," in *Gas injection methods*, pp. 171–198, Elsevier, 2023.
- [16] N. Kumar, M. Augusto Sampaio, K. Ojha, H. Hoteit, and A. Mandal, "Fundamental aspects, mechanisms and emerging possibilities of co₂ miscible flooding in enhanced oil recovery: A review," *Fuel*, vol. 330, p. 125633, 2022.
- [17] C. Whitson, M. Brulé, and S. of Petroleum Engineers of AIME., *Phase Behavior*. Henry L. Doherty Series, Henry L. Doherty Memorial Fund of AIME, Society of Petroleum Engineers, 2000.
- [18] A. Younessi, E. Kärstad, P. Basu, M. Larsen, and O. Burachok, "Geomechanical analysis of caprock integrity and fault stability for greensand co₂ storage project feasibility," in *ARMA US Rock Mechanics/Geomechanics Symposium*, p. D032S038R001, ARMA, 2024.
- [19] S. A. Faroughi, N. Pawar, C. Fernandes, M. Raissi, S. Das, N. K. Kalantari, and S. K. Mahjour, "Physics-guided, physics-informed, and physics-encoded neural networks in scientific computing," *arXiv preprint arXiv:2211.07377*, 2022.
- [20] M. Raissi, P. Perdikaris, and G. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.
- [21] J. Zhang, U. Braga-Neto, and E. Gildin, "Physics-informed neural networks for multiphase flow in porous media considering dual shocks and interphase solubility," *Energy & Fuels*, vol. 38, no. 18, pp. 17781–17795, 2024.
- [22] K. Ghorayeb, K. Mogensen, N. El Droubi, C. K. Kloucha, and H. Mustapha, "Physics-enhanced machine-learning-based prediction of fluid properties for gas injection – focus on co₂ injection," *Gas Science and Engineering*, vol. 123, p. 205228, 2024.
- [23] L. biao Ouyang and K. Aziz, "Steady-state gas flow in pipes," *Journal of Petroleum Science and Engineering*, vol. 14, no. 3, pp. 137–158, 1996.
- [24] G. Yadigaroglu, G. Hetstroni, and G. F. Hewitt, "Flow regimes," in *Introduction to Multiphase Flow: Basic Concepts, Applications and Modelling* (G. Yadigaroglu and G. F. Hewitt, eds.), pp. 95–140, Cham: Springer International Publishing, 2018.
- [25] W. McCain, *The Properties of Petroleum Fluids*. PennWell Books, 1990.
- [26] E. Mohagheghian, A. Bahadori, and L. A. James, "Carbon dioxide compressibility factor determination using a robust intelligent method," *The Journal of Supercritical Fluids*, vol. 101, pp. 140–149, 2015.
- [27] D.-Y. Peng and D. B. Robinson, "A new two-constant equation of state," *Industrial & Engineering Chemistry Fundamentals*, vol. 15, no. 1, pp. 59–64, 1976.
- [28] X. Zhang, S. Yang, G. Li, W. Li, and H. Chen, "Compressibility factor of gas with high content of co₂ in changshen gas reservoir," in *2011 International Conference on Computational and Information Sciences*, pp. 542–545, 2011.
- [29] Y. Bai and Q. Bai, "Chapter 13 - hydraulics," in *Subsea Engineering Handbook* (Y. Bai and Q. Bai, eds.), pp. 349–399, Boston: Gulf Professional Publishing, 2010.
- [30] D. Brkić, "Review of explicit approximations to the colebrook relation for flow friction," *Journal of Petroleum Science and Engineering*, vol. 77, no. 1, pp. 34–48, 2011.
- [31] P. K. Swamee and A. K. Jain, "Explicit equations for pipe-flow problems," *Journal of the hydraulics division*, vol. 102, no. 5, pp. 657–664, 1976.
- [32] T. H. Chung, M. Ajlan, L. L. Lee, and K. E. Starling, "Generalized multiparameter correlation for nonpolar and polar fluid transport properties," *Industrial & Engineering Chemistry Research*, vol. 27, no. 4, pp. 671–679, 1988.
- [33] A. Chapoy, F. J. Owuna, R. Burgass, P. Ahmadi, and P. Stringari, "Viscosity of the co₂ + ch₄ binary systems from 238 to 423 k at pressures up to 80 mpa," *Journal of Chemical & Engineering Data*, vol. 69, no. 6, pp. 2152–2166, 2024.
- [34] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [35] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Pearson, 4th ed., 2021.
- [36] S. A. Faroughi, A. I. Roriz, and C. Fernandes, "A meta-model to predict the drag coefficient of a particle translating in viscoelastic fluids: A machine learning approach," *Polymers*, vol. 14, no. 3, 2022.
- [37] H. Ghorbani, "Mahalanobis distance and its application for detecting multivariate outliers," *Facta Universitatis, Series: Mathematics and Informatics*, pp. 583–595, 2019.
- [38] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *2008 Eighth IEEE International Conference on Data Mining*, pp. 413–422, 2008.
- [39] V. G. Raju, K. P. Lakshmi, V. M. Jain, A. Kalidindi, and V. Padma, "Study the influence of normalization/transformation process on the accuracy of supervised classification," in *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 729–735, IEEE, 2020.
- [40] A. Zhang, Z. Lipton, M. Li, and A. Smola, *Dive Into Deep Learning*. Cambridge University Press, 2023.
- [41] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation," *Peerj computer science*, vol. 7, p. e623, 2021.
- [42] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17*, (Red Hook, NY, USA), p. 4768–4777, Curran Associates Inc., 2017.



MATEUS A. FERNANDES received the B.S. degree in electrical engineering from the Federal University of Minas Gerais, Belo Horizonte, Brazil, in 2006 and the M.S. degree in aerospace engineering from the Aeronautics Institute of Technology, São José dos Campos, Brazil, in 2009. He is currently pursuing his Ph.D. in Mining and Petroleum Engineering at the University of São Paulo, São Paulo, Brazil, with research focused on data-driven models for forecasting and decision support in petroleum reservoirs.

Since 2009, he has been a Petroleum Engineer at Petrobras, Brazil, specializing in reservoir simulation and management, production monitoring, and reserves estimation for onshore and offshore oil and gas fields. He has extensive experience as a professor, having taught oil and gas-related disciplines at academic institutions in Brazil. He also served as a visiting scholar at Texas A&M University, College Station, USA. He has authored numerous papers in international conferences and journals, contributing to advancements in reservoir management and computational techniques, and is also a reviewer for relevant journals in those areas.

Mr. Fernandes is a member of the Society of Petroleum Engineers (SPE).



EDUARDO GILDIN received the B.S. degree in mechanical engineering from Faculdade de Engenharia Industrial, São Bernardo do Campo, Brazil, in 1995, the M.S. degree in mechatronics from the University of São Paulo, São Paulo, Brazil, in 1998, and the PhD degree in Aerospace Engineering from The University of Texas at Austin, USA, in 2006.

He is currently a Professor, Associate Department Head, and Graduate Program Director at the Petroleum Eng. Department at Texas A&M University. He is the holder of the Rob L. Adams '40 Professorship in Petroleum Engineering. He teaches undergraduate and graduate courses in the mathematics of reservoir simulation, data analytics, and reduced-order modeling and has published more than 150 papers in peer-reviewed journals and conferences. His research has been supported by grants from NSF, DOE, DOD, NASA, and Industry.

Dr. Gildin was inducted into the Society of Petroleum Engineers (SPE) Distinguished Membership in 2021.



MARCIO A. SAMPAIO received the B.S. degree in physics in 2005, the M.S. degree in Electrical Engineering in 2007, and the PhD in petroleum science and engineering in 2013, all from the University of Campinas, Campinas, Brazil.

He is currently a full-time Professor at the Mining and Petroleum Engineering department at the University of São Paulo, São Paulo, Brazil. He teaches undergraduate and graduate courses in reservoir engineering, numerical simulation, data science, and machine learning applied to oil and gas industry. He is the founder and coordinator of the Reservoir Simulation and Management Laboratory (LASG) at the University of São Paulo, is involved in scientific and technological research and development projects, and supervises graduate students in the Mineral and Petroleum Engineering Program and the Naval and Ocean Engineering Program. He also serves as a reviewer for leading journals in the field.

• • •