

Application of Artificial Intelligence in Reservoir Simulation for Characterizing Multi-Phase Fluid Flow through Petroleum Reservoirs

Rudarsko-geološko-naftni zbornik
(The Mining-Geology-Petroleum Engineering Bulletin)
DOI: 10.17794/rgn.2025.5.3

Review scientific paper



Viswakanth Kandala¹ , Suresh Kumar Govindarajan¹ , Tummuri Naga Venkata Pavan^{1*} ,
Swaminathan Ponmani² , Srinivasa Reddy Devarapu³

¹ Reservoir Simulation Laboratory, Petroleum Engineering Division, Department of Ocean Engineering, Indian Institute of Technology, Madras, India.

² Department of Petroleum Engineering, Academy of Maritime Education and Training, Kanathur, Chennai, India.

³ Department of Petroleum Engineering and Earth Sciences, UPES, Dehradun, India.

Abstract

Artificial intelligence (AI) has rapidly advanced and influenced all scientific fields, including the petroleum industry, where it is no longer a novel concept. This article explores the evolving role of AI and its integration with traditional reservoir simulation approaches. The first section highlights the diminishing emphasis on conceptual and mathematical modelling in reservoir simulation, with a growing focus on sophisticated numerical solution techniques. This shift often neglects fundamental reservoir physics and mathematics, leading to the superficial characterization of multi-phase fluid flow in petroleum reservoirs. The second section examines the prevalent use of petroleum software packages, which heavily rely on input data without accounting for variations in data scales. These tools treat underlying programming as a black box, often bypassing critical basic sciences, such as reservoir conceptualization, applied mathematics, and numerical techniques, resulting in incomplete reservoir characterization. The third section discusses the role of machine learning (ML) and AI in reservoir applications. While data science plays a pivotal role, the lack of integration with fundamental reservoir physics reduces fluid flow analysis to an art devoid of scientific rigour. The final section proposes a hybrid approach that couples AI/ML with traditional reservoir simulation. This integration bridges the gap between science-based reservoir simulation and AI-driven fluid flow characterization, enabling the petroleum industry to achieve a new paradigm for multi-phase fluid flow analysis. By combining fundamental science with advanced AI techniques, this approach offers a comprehensive framework for accurate reservoir characterization and improved hydrocarbon production.

Keywords:

artificial intelligence, machine learning, reservoir simulation, conceptual model, mathematical model, numerical model

1. Introduction

Post-pandemic, the world is gradually returning to pre-pandemic normalcy in terms of hydrocarbon consumption, despite ongoing geopolitical challenges. Studies project an average global GDP growth rate of 3% until 2030, with emerging economies playing a crucial role. As hydrocarbons continue to be a key driver of this growth, their exploration and utilisation must be optimised for maximum efficiency. In this evolving landscape, advanced technologies such as Artificial Intelligence (AI) and Machine Learning (ML) are rapidly progressing, facilitating the integration of the physical and digital domains on an unprecedented scale, with far-reaching impact.

The adoption of AI and ML in the upstream industry is steadily increasing, with numerous applications showcasing their potential. The AI market in the upstream sector is expected to grow from \$2.8 billion in 2023 to \$5.1 billion by 2028. AI primarily focuses on creating machines that mimic human intelligence, encompassing a broad range of technologies. Within AI, ML represents a subset that builds models on pre-trained data, identifying patterns to predict outcomes (Bhattacharya, 2021). Deep Learning (DL), a further specialized subset of ML, processes large volumes of data using complex algorithms to make decisions. Data Science (DS), which intersects with AI, ML, and DL, combines statistics, programming, and domain expertise to extract meaningful insight and knowledge from both structured and unstructured data.

In a producing hydrocarbon reservoir, vast amounts of real-time data are generated from various sources such as sensors, gauges, and meters. This data is characterized by its volume, velocity, variety, veracity, and

* Corresponding author: Tummuri Naga Venkata Pavan
e-mail address: pavantummuri@gmail.com

Received: 5 December 2024. Accepted: 24 April 2025.

Available online: 21 October 2025

Table 1. Applications of AI/ML in Reservoir Simulation Available in Literature

Application	ML Methods	References
Hydraulic fracture design	ANN, Deep learning	Mohaghegh et al., 1999; Liu et al., 2021; Jia et al., 2024
Uncertainty analysis	ANN	Mohaghegh, 2011
Waterflooding design	ANN, Multi objective optimizations	Farahi et al., 2021; Zhong et al., 2020
Phase stability calculations	ANN, SVM	Li et al., 2019; Zhang et al., 2020
Shale gas simulations	ANN, XGBoost, CatBoost	Kalantari-Dahaghi and Mohaghegh, 2011; Wang et al., 2023
History Matching	ANN, GAN	Jo et al., 2021; Ramgulam et al., 2007; Srinivasan et al., 2021
Oil production prediction	NN	Muradkhanli, 2018; Zhang et al., 2022; Li et al., 2022

ANN (Artificial Neural Network), NN (Neural Network), GAN (Generative Adversarial Network), SVM (Support Vector Machine), XGBoost (Extreme Gradient Boosting), CatBoost (Categorical Boosting).

value – attributes that classify it as ‘big data.’ Such big data plays a critical role in subsurface characterization, reservoir performance analysis, and optimization.

In the context of reservoir management, AI facilitates critical decision-making and the characterization of multi-phase fluid flow within petroleum reservoirs. Some of these applications, as reported in the literature, are summarized in **Table 1**. While a strong foundation in mathematics is essential, particularly in areas related to reservoir simulation, domain expertise is less critical when using AI compared to traditional reservoir simulation methods. Building on the foundation of a single-layer perceptron, or a machine capable of independent thinking, we have now reached a stage where AI/ML approaches driven by pattern recognition and creativity, including those in the petroleum industry, have become indispensable. At the same time, the practice of scientific-based reservoir simulation remains crucial for effective reservoir characterization. Thus, the current article attempts to discuss these two distinct approaches, science-based and AI-based reservoir simulation, individually and proposes leveraging their combined effects to successfully characterize multi-phase fluid flow in petroleum reservoirs with minimal uncertainty.

In the petroleum industry, three main types of models are commonly used: physical models, empirical models, and mathematical models (Noshi and Schubert, 2018). Physical models involve scaled-down versions of actual field-scale reservoirs (Pavan et al., 2024; Reddya and Kumarb, 2014). However, the extent to which the laboratory setup accurately mimics real field conditions is often questionable. This approach is not only costly and time-consuming but also presents significant challenges in upscaling, as the fundamental physics and the reservoir geometry at the laboratory scale differ substantially from those in field conditions. Empirical models, on the other hand, are based on insight derived from experimental observations, such as Darcy’s law. While useful, these models are prone to human error or measurement

inaccuracies, which can affect their reliability. Furthermore, they cannot be generalized, as they are not directly deduced from fundamental physical principles. Mathematical models address some of these limitations by deriving non-linearly coupled partial differential equations (PDEs) from classical physical principles (Ansari and Govindarajan, 2022, 2024; Kandala and Govindarajan, 2023). However, this approach typically involves numerous assumptions and simplifications to manage the mathematical complexity, which may compromise the model’s fidelity. Despite these limitations, reservoir simulation continues to play a critical role in petroleum reservoir management. AI is now transforming the industry by addressing many challenges associated with these traditional models, including reservoir simulation. With its ability to analyse and derive insight from vast datasets, capturing complex relationships between rock, fluid, and rock-fluid properties it is increasingly being recognized as a powerful tool. It enhances decision-making, enables deeper understanding, and provides timely insight required by field reservoir engineers, thereby complementing and potentially surpassing traditional modelling approaches.

With the petroleum industry rapidly transitioning to oil-field digitization through the application of data-driven modelling, the role of AI has become critical (Solomatine and Ostfeld, 2008). The key question is no longer whether to adopt AI, but how to maximize its potential within the petroleum industry to achieve continuous improvements in operational performance (Al-Rbeawi, 2023; Kronberger et al., 2020). In this context, this article proposes coupling reservoir simulation with AI to harness the strengths of both approaches, ensuring maximum benefits with minimal uncertainty. This methodology assigns equal weight to reservoir simulation and AI, leveraging their complementary capabilities. Moreover, the integration of AI/ML within the framework of the big data revolution is expected to significantly reduce reservoir operation costs without compromising safety standards.

This work will significantly benefit the new generation of petroleum engineers pursuing careers in reservoir engineering and simulation, particularly at the intersection of AI and ML. It provides valuable insight into the critical aspects that must be considered on the reservoir side before embarking on model development and apply ML models. This includes understanding the intricacies of reservoir properties, data acquisition, and preprocessing, all of which are essential for building accurate and reliable models. Moreover, thought-provoking questions have been put forward for the readers, which can inspire further research in this field.

2. Reservoir Simulation

The fundamental aspect of reservoir simulation essentially involves four basic stages (see **Figure 1**) a) conceptual modelling, b) mathematical modelling, c) numerical modelling and d) simulation using packages, through which multi-dimensional, multi-phase compressible fluid flow in a petroleum reservoir is characterized. However, even after meticulously following these four stages, the simulation can only partially reflect the reality of the actual reservoir. The degree to which the simulation replicates the real field conditions whether it approaches 95% accuracy or not depends significantly on an individual's in-depth knowledge of reservoir geology, petro-physics, fluid dynamics, thermodynamics, geo-mechanics, differential calculus, along with a strong understanding of the fundamental drainage principles of a hydrocarbon reservoir below and above the bubble point pressure.

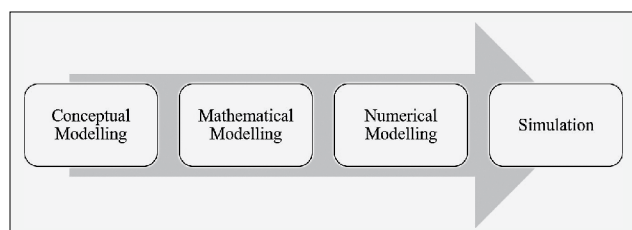


Figure 1. Key steps in reservoir simulation

2.1. Conceptual Modelling

The first stage of reservoir simulation involves the formulation of a conceptual model, as illustrated in **Figure 2**. Unfortunately, this crucial aspect is often undervalued due to the significant conceptual understanding it demands about a reservoir. This stage involves visualising the actual, complex three-dimensional petroleum reservoir within a conceptual framework, highlighting intricate details such as delineation of reservoir boundaries, conformities, and heterogeneities; well patterns and the locations of injection and production wells; the feasibility of pseudo-steady-state and transient fluid flow; phase change aspects; the restructuring of the three-dimensional solid grain network; variations in Reynolds

number as a function of distance from injection or production wells; and the maintenance of laminar flow streamlines, including identifying the location and time where inertial effects, if any, may arise. Additionally, the transport of oil, water, and gas through complex pore networks, the interplay of capillary, viscous, and gravity forces, and their collective influence on fluid flow and oil-water contacts within this three-dimensional domain need to be considered. All these aspects must be conceptually (virtually) brought to life for further analysis.

Once this conceptualisation is achieved, the next step is to identify and list the (a) physical processes, (b) chemical processes (**Devarapu et al., 2023; Dinesh et al., 2024; Govindarajan et al., 2022**), and (c) biological processes (**Chakraborty et al., 2020**) individually associated with the reservoir. Following this, the feasibility of coupled processes – (a) between physical and chemical processes, (b) between chemical and biological processes, and (c) between biological and physical processes must be assessed. From these lists, the dominant and sensitive individual processes, as well as their associated coupled processes, must be identified and documented. This detailed understanding forms the output of the conceptual modelling stage.

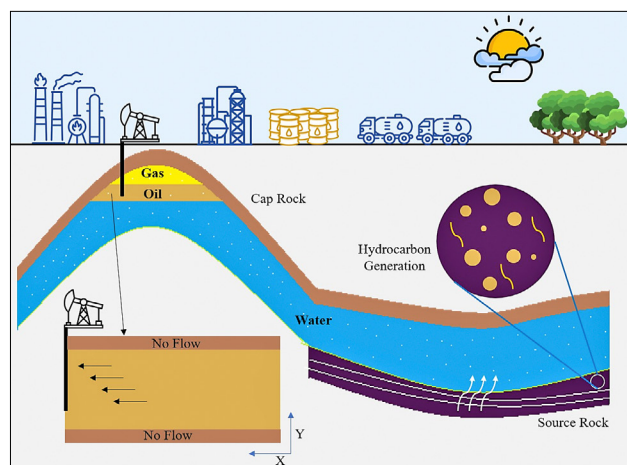


Figure 2. Typical petroleum reservoir schematic illustrating the deduction of a 2D conceptual model, showcasing the no-flow boundaries, wells, and flow dynamics.

2.2. Mathematical Modelling

The second step, mathematical modelling requires a strong understanding of mathematics, especially differential calculus. In this stage, the results from conceptual modelling must be translated into mathematical equations. A solid mathematical foundation is necessary to proceed, as it's important to know when to use linear or non-linear algebraic equations, ordinary differential equations (ODEs), or partial differential equations (PDEs). Basic knowledge of elliptic, parabolic, and hyperbolic PDEs is essential (as shown in **Table 1**). For instance, elliptic PDEs are not relevant for transient-

Table 2. Overview of PDE properties, highlighting the nature of solutions, time dependence, and equilibrium behaviour.

Property	Elliptic PDE	Parabolic PDE	Hyperbolic PDE
Form	$au_{xx} + bu_{xy} + cu_{yy} + du_x + eu_y + fu = g(x,y)$	$u_t = au_{xx} + bu_x + cu + f(x,t)$	$u_{tt} = au_{xx} + bu_x + cu + f(x,t)$
Time Dependence	Time Independent	Evolves over time in single direction	Evolves over time like waves in both directions
Nature of Solutions	Smooth and bounded	Gradual varies with time	Propagates as waves with finite speed
Stability	Unstable to perturbations (no growth of solutions)	Dissipates over time	Sensitive to initial conditions (can lead to shock waves)
Propagation of Information	Global (affects entire domain)	Parabolic diffusion-like propagation	Local (finite speed, well-defined regions of influence)
Equilibrium Behaviour	Steady-state	Diffusion to equilibrium	Oscillatory or wave-like
Examples	Steady-state flow	Transient pressure diffusion	Seismic wave propagation

state fluid flow problems. In contrast, parabolic and hyperbolic PDEs are important in transient reservoir physics. A system where pressure changes over time and eventually reaches a steady state can be modelled with a parabolic diffusivity equation (Ansari and Govindarajan, 2023; Devarapu et al., 2023; Pavan et al., 2023; Pavan and Govindarajan, 2023; Sivasankar and Suresh Kumar, 2018). However, if the system shows waves or heterogeneities, hyperbolic PDEs may also be needed. Once the conceptual model is converted into a set of coupled PDEs, it is crucial to determine the correct initial and boundary conditions for the equations. The solution should be well-posed, meaning numerical solutions should be unique and stable when solving them computationally.

2.3. Numerical Modelling

Numerical modelling involves solving mathematical models that are comprised of PDEs or ODEs, using computer algorithms when exact solutions are difficult to obtain. This requires a good understanding of various numerical methods, especially for non-linear and coupled PDEs.

For example, if the model is a simple parabolic diffusivity equation, finding a numerical solution is easier. In such cases, errors from initial or boundary conditions usually disappear over time as the system reaches a steady state. However, for models with hyperbolic equations, like wave equations, finding a stable solution is harder. This is because hyperbolic PDEs amplify any errors at the start, leading to significant convergence problems.

In petroleum reservoir modelling, we often deal with both parabolic and hyperbolic PDEs. For instance, when modelling fluid flow in a reservoir, we use mass and momentum conservation equations, which result in a mix of parabolic PDEs and additional hyperbolic terms. These hyperbolic terms are sometimes ignored, assuming a simple, homogeneous reservoir, which is rarely the case in reality.

To solve the mathematical models, it's important to understand: (a) the number of equations, (b) dependent

and independent variables, (c) constants and variable coefficients, (d) knowns and (e) unknowns. This understanding is crucial when linearizing non-linear PDEs for numerical solutions. Linear systems can be solved using direct methods (giving exact solutions) or iterative methods (giving approximate solutions). Direct methods compute exact answers in a finite number of steps, while iterative methods start with an initial guess and refine it over time. Examples of iterative methods include Jacobi's, Gauss-Seidel, and Relaxation methods, which converge to the solution after a few iterations, as shown in **Figure 3** (Srinivasa Reddy and Suresh Kumar, 2015). Having figured out an appropriate and efficient numerical solution technique, a clear flow chart indicating the details of input variables, initial unknowns, initial guesses, and the equation to be solved is prepared. If convergence is achieved, it means it will go to the next step; otherwise, it will go back to the same old step until convergence is achieved. Common numerical methods include finite difference (FDM), finite element (FEM), and finite volume (FVM) techniques. The FDM approximates solutions using Taylor's series but cannot capture reservoir heterogeneities between nodes. Larger cell widths can miss crucial details, making it less effective for complex reservoirs. The FEM improves this by introducing elements between nodes, giving better control over variable variations. However, FEM struggles with fluid mass conservation, making it less ideal for petroleum reservoir simulations. The FVM is based on fluid mass conservation and better suited for handling steep gradients and shocks, like flow in fractured reservoirs or shale gas systems (Kudapa et al., 2017).

To ensure accurate results, the numerical model must achieve both numerical convergence and mathematical convergence. Programming, often in Fortran or C or Python, begins with simple PDEs and known initial and boundary conditions to verify the code before tackling real complex equations. The results are validated against existing solutions, experimental data, and field data to ensure reliability. In practice, many skip the foundational steps of conceptual and mathematical modelling, directly

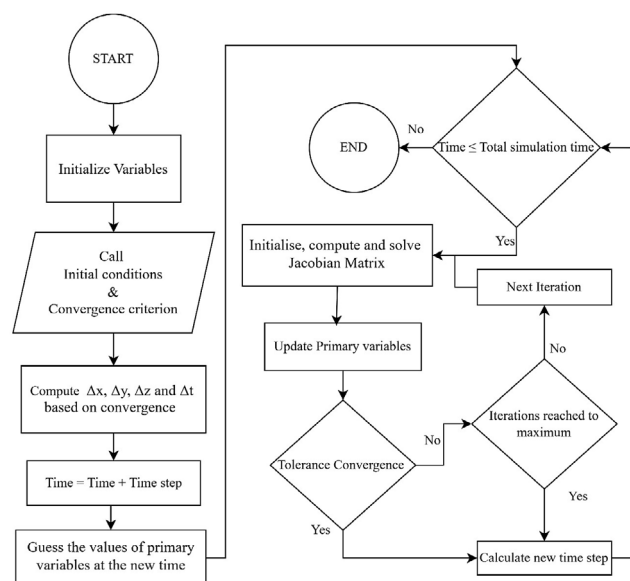


Figure 3. Algorithm describing an iterative approach for solving a system of non-linear PDE

using pre-existing equations from literature with minor modifications. This approach often neglects reservoir physics and mass conservation, leading to poor characterization of multi-phase fluid flow in heterogeneous reservoirs. A reservoir simulation engineer must possess a strong foundation in reservoir physics, computational fluid dynamics, and geology to interpret numerical results confidently and make accurate decisions in the field.

2.4. Reservoir Simulation using Packages

The concept of reservoir fluid flow characterization using existing petroleum software packages remains highly useful for field engineers in the petroleum industry as they may not be able to spend time on the fundamental stages of modelling discussed in previous sections. Thus, in this approach, most of the energy is spent only on gathering the required data from various sources. In this approach, even the data gathering remains not justified completely because of some of the following reasons: (a) the data towards rock property, fluid property, and rock-fluid interaction property do not generally tend to be uniform in nature with reference to the amount of data gathered; (b) most of the rock property, fluid property and rock-fluid interaction property data pertain to a core-scale (done at lab; or from PVT cell data), while, only very few data pertain to a field-scale; (c) some of the fluid property and rock-fluid interaction property data pertain to a much smaller scale (like contact angle and interfacial curvature data) than from the required Representative Elementary Volume (REV); (d) incorporating data at different scales with extreme variations on the same input platform, such as interfacial tension data at the sub-pore scale alongside permeability data at the large field scale, thereby deviating from the continuum hypothesis-based Darcian approach. (Venkata Pavan et al., 2024); (e) feeding more data pertain-

ing to reservoir statics rather than feeding the required reservoir dynamics data; (f) securing laboratory-scale capillary pressure data with ease, while the concept of equilibrium capillary pressure may take a very long time at the field scale; (g) obtaining relative permeability data at the laboratory scale based solely on water saturation may not reflect real field conditions, where hysteresis plays a significant role, and fluid flow is often characterized by partial drainage and partial imbibition. Thus, data gathering alone is insufficient; a thorough understanding of data scales, uniformity, and the number of data points for each rock, fluid, and rock-fluid interaction property is crucial. Many engineers may lack a strong foundation in reservoir physics (conceptual model), applied reservoir mathematics (mathematical model), and numerical solution techniques (numerical model) used in petroleum software. Without a clear idea of initial and boundary conditions and their stability criteria, simply inputting cell width and time step may not produce meaningful results, leading to misinterpretation. Therefore, relying solely on input data without formulating conceptual, mathematical, and numerical models is unsuitable for academic purposes. However, industry professionals, having undergone these modelling stages, can interpret results effectively while considering limitations. Unlike fresh graduates, they have the expertise to critically analyse simulations, ensuring informed decision-making. This approach, in most cases, lacks a strong foundation in fundamental reservoir sciences.

3. AI in Petroleum Industry

AI is advancing towards mimicking human decision-making processes. ML, a subset of AI, enables computers to respond beyond their programmed behaviour by utilizing external data. It primarily helps in the extraction of actionable insight from big data. ML can be categorized into distinct types, as illustrated in Figure 4. The reliability and generalization of a model largely depend on the quantity and diversity of data used during its development. Various statistical and graphical approaches are employed to analyse the performance of predictive models in the petroleum industry. For example, in reservoir simulation, data from multiple sources are integrated without extensive preprocessing. In contrast, ML-based approaches emphasize data preprocessing, refining datasets before applying graphical and statistical analyses. This refinement process involves eliminating unreliable data points, such as outliers or incorrect entries, to enhance the accuracy and reliability of predictive tools. Error analysis is a critical component in this approach, as it evaluates the performance and accuracy of predictive models. The core of error analysis involves quantifying deviations between predicted and actual data points using mathematical formulations. In AI applications, unlike in reservoir simulations, data preprocessing is pivotal, with several essential steps undertaken before using the data for modelling purposes.

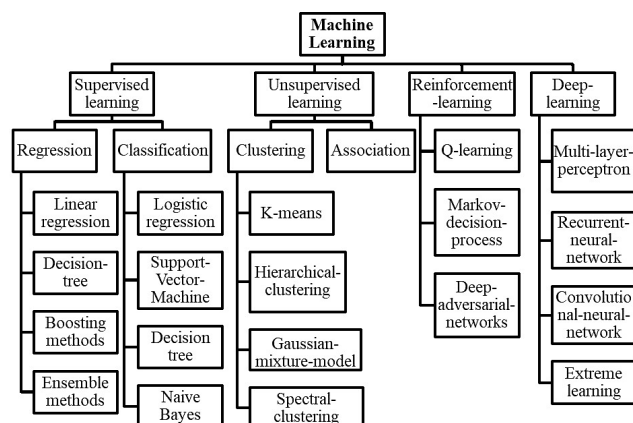


Figure 4. Classification of Machine Learning Techniques

Data preprocessing, as shown in **Figure 5**, primarily involves the following steps: (a) ‘Data cleaning’, which comprises eliminating inconsistencies, smoothing noisy data, and imputation of missing data points; (b) ‘Data integration’, where diverse data representations are combined to deduce a unique and consistent representation; (c) ‘Data transformation’, which includes normalisation, generalisation, and aggregation of data; (d) ‘Data reduction’, involving the reduction of data representation within a database; (e) ‘Data discretisation’, where data points within the same interval are averaged; and (f) ‘Data statistics’, covering metrics such as Skewness (indicating the asymmetry of the probability distribution of a random variable about its mean) and Kurtosis (indicating the degree of tailed-ness in a probability distribution).

Once preprocessing is complete, data processing begins, comprising (i) ‘Data training’, where approximately three-fourths of the dataset is used to train the model, and (ii) ‘Data validation and testing’, aimed at evaluating the model’s ability to predict new data points. Post-processing of data involves evaluating the model using either ‘statistical error analysis’ – which assesses performance through metrics such as average percent relative error, average absolute percent relative error, root mean square error, standard deviation, or the coefficient of determination – or ‘graphical error analysis’, which visualises performance through tools like error distribution curves or cross plots. The applicability domain of the model is then finalised by identifying outliers within the dataset.

Finally, ‘sensitivity analysis’ is conducted to determine how uncertainties in the model’s inputs influence the uncertainties in its outputs. This involves performing relevancy factor analysis before the model is deemed ready for application in the petroleum industry, where intelligent models are increasingly employed.

The development of intelligent models began with the introduction of artificial neural networks (ANNs), inspired by biological neurons and the human brain. Key ANN architectures come in different types based on their structure and purpose. Feedforward neural net-

works (FNN) are the basic, where data transfers in uni-direction from input to output. Convolutional neural networks (CNN) are used for image and spatial data processing, recognizing patterns through filter layers. Recurrent neural networks (RNN) are for sequential data, where past information influences the current output, making them useful in forecasting and language processing. Long short-term memory (LSTM) networks, a type of RNN, help retain information for longer periods. Generative adversarial networks (GAN) consist of two networks, one for data generation and the other for evaluation, commonly used in image creation. Radial basis function (RBF) networks use mathematical functions to improve classification and pattern recognition. Transformer networks, used in modern language models, process information efficiently by focusing on critical parts of the input. Building on these, fuzzy logic systems were introduced to enhance higher-level reasoning and inference, leading to the emergence of adaptive neuro-fuzzy inference systems (ANFIS), which integrate ANN capabilities with fuzzy logic principles.

In reservoir simulation, data-driven neural networks apply these intelligent modelling techniques to learn from historical reservoir and production data for predicting reservoir behaviour. However, their accuracy depends on data availability, and limited datasets can lead to unreliable predictions. To overcome this limitation, physics-informed neural networks (PINNs) incorporate fundamental physical laws, such as conservation of mass and Darcy’s law, ensuring that predictions align with real reservoir flow behaviour even in data-scarce conditions. This integration of data-driven and physics-informed approaches enhances the reliability and applicability of intelligent models in reservoir analysis and decision-making (de la Mata et al., 2023).

The concept of decision trees, particularly in the forms of random forests and extra trees (extremely randomized trees), has become a cornerstone of intelligent models due to their simplicity, interpretability, graphical representation, and low computational cost. More recently, genetic programming, an evolutionary algorithm exploring both program and solution spaces, and gene expression programming, which seeks optimal expression models through chromosome-based encoding and solution reporting, have gained traction. These methods have surpassed the group method of data handling (GMDH) in various applications (Hemmati-Sarapardeh et al., 2020).

Towards the development of training and optimization algorithms, numerous techniques have been introduced in recent years. These span a broad spectrum of optimization approaches, ranging from nature-inspired heuristics to gradient-based deterministic methods, and offer solutions for diverse problem domains. These include: (a) Genetic Algorithm; (b) Differential Evolution; (c) Particle Swarm Optimization; (d) Ant Colony Optimization; (e) Artificial Bee Colony; (f) Firefly Algorithm; (g) Imperialist Competitive Algorithm; (h) Simu-

Table 3. Advantages and disadvantages of application of various ML algorithms in reservoir simulation (Zhou et al., 2024)

Algorithm	Advantages	Disadvantages
Linear Regression	Interpretable and suitable for variable data.	Cannot model nonlinear relationships.
Fuzzy Logic	Handles uncertainty and reduces noise.	Requires domain expertise and algorithm knowledge.
Decision Trees	Easy to interpret using root-to-leaf rules.	Prone to overfitting; pruning is challenging.
Random Forest	Accurate; handles high-dimensional, noisy data.	Limited effectiveness in low-dimensional scenarios.
XGBoost	Efficient, handles missing data, and controls overfitting well.	Poor fit for petroleum scenarios; prone to overfitting.
AdaBoost	Low error, simple coding, no tuning.	Ineffective in scenarios with many outliers.
Support Vector Machine	Accurate; robust to noise and outliers; handles nonlinear data.	Kernel selection is challenging for specific problems.
Evolutionary Algorithms	Few parameters; minimal background knowledge needed.	Easily falls into local optima; requires scenario-specific objectives.
ANN	Handles complex relationships and noisy data effectively.	Difficult to design; requires complex training and many parameters.
RNN	Processes sequential data with memory capabilities.	Struggles with historical data; prone to gradient issues.
LSTM and variants	Manages long dependencies and gradient issues effectively.	Long training, parameter-heavy, and computationally intensive.

lated Annealing; (i) Coupled Simulated Annealing; (j) Gravitational Search Algorithm; (k) Cuckoo Optimization Algorithm; (l) Gray Wolf Optimization; (m) Whale Optimization Algorithm; (n) Levenberg-Marquardt Algorithm; (o) Bayesian Regularization Algorithm; (p) Scaled Conjugate Gradient Algorithm; and (q) Resilient Backpropagation Algorithm.

AI has been extensively used in the upstream industry for various applications, including well testing, production prediction, history matching (Costa et al., 2014), hydrocarbon property estimation, oil field development (Sircar et al., 2021; Ren et al., 2024), and fracture parameter predictions (Ahmed et al., 2019; Nande, 2018). While AI-driven intelligent models have brought significant improvements, their effectiveness needs to be carefully studied. A key challenge lies in selecting appropriate algorithms for reservoir fluid flow analysis, particularly in assessing reservoir rock, fluid, and rock-fluid interaction properties to determine the fraction of enhanced oil recovery. Additionally, well-test analysis and formation damage analysis still require more advanced specialized algorithms. It is crucial to evaluate how accurately AI can deduce rock parameters, fluid parameters, and rock-fluid interaction parameters, ensuring that these deductions align with real field conditions. Even if AI can accurately predict all reservoir properties using advanced algorithms, we still need to check how close these predictions are to actual field scenarios. A key question is the role of detailed mathematical formulations in fluid flow analysis of petroleum reservoirs and how much AI-based models – using pattern recognition, system identification, and cognitive processes – have been able to replicate real-world fluid flow behaviour.

Another important area is understanding how AI performs cross-validation during training and selects the most suitable reservoir model. The process of dividing datasets into training and validation sets, choosing data for learning versus validation, and deciding when to stop training needs further exploration. It is important to consider whether achieving low errors in training and validation sets is sufficient to stop learning or if additional “blind” data is necessary for better evaluation. AI plays a crucial role in addressing disparities in data during the training phase of model development. These intelligent models learn by recognizing hidden patterns between input reservoir properties and output responses, using optimization algorithms to improve performance. While AI is highly effective in handling large datasets, identifying patterns, and detecting trends, a key question remains: can AI fully capture all aspects of reservoir heterogeneity? Another critical consideration is the extent to which AI reduces uncertainty in reservoir performance. Ultimately, the reliability of an AI model depends significantly on the quality and quantity of reservoir data used. For instance, insufficient permeability data can challenge the model’s ability to make accurate predictions. Moreover, in cases where data is limited, dividing the dataset into training, validation, and test subsets may not be feasible, further impacting the robustness and reliability of the intelligent model. These limitations highlight the need for quality data and careful handling during model development.

In essence, introducing AI has essentially translated the reservoir scientific basis into an art form as it essentially hangs around pattern recognition. This particular approach by any user, in the absence of acquiring the actual

reservoir drainage physical principles, while mostly playing with the data set alone, may not help the engineer in the long run, and this approach survives with art as a basis while science seems to be missing completely.

4. Coupled Effect of AI and Reservoir Simulation

As far as the reservoir engineering discipline is concerned, the concept of AI/ML remains widely used in reservoir characterization, in particular, reservoir pressure, temperature and volume estimations, including bubble point pressure, formation volume factor, isothermal compressibility and brine salinity. In addition, ML has also been used in a compositional oil simulator towards phase-equilibrium estimations that include phase-stability tests and phase-splitting calculations (**Mirzaei and Das, 2007**). Reservoir rock properties, including porosity and permeability, have also been estimated using fuzzy logic and SVM (**Koray et al., 2024; Lim, 2005**). An attempt has already been made to forecast the time-lapse saturation profiles at well locations using injection/production data (**Balza and Li, 2020**). In fact, extreme learning machines (ELM) have helped to forecast multiple reservoir parameters, namely lithofacies, shale content, saturation and reservoir porosity (**Lawal et al., 2024; Liu et al., 2021**). Further, ANNs not only helped to forecast bottom hole pressure in vertical wells but also have aided multi-dimensional interpolation of relative permeability to overcome the influences of various parameters during hybrid recovery processes.

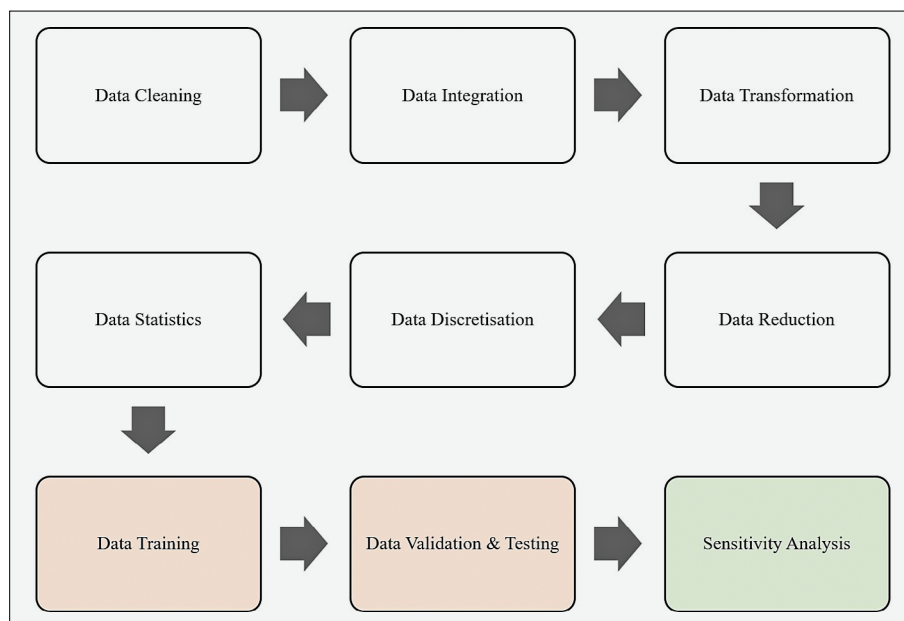
Now, considering three possible approaches for implementing ML algorithms: (a) developing surrogate models for relatively homogeneous and isotropic reservoirs to reduce computational resources and costs; (b) creating ML models for reservoirs where human expertise is crucial, particularly for reservoir planning and field management decisions during hydrocarbon production; and (c) developing ML models for highly heterogeneous and anisotropic reservoirs, which are more complex to understand and model. The first two options can be managed with reservoir simulation, while the third option may still require the help of AI/ML along with reservoir simulation. However, it should be clearly noted that in petroleum reservoir fluid flow analysis, the reservoir environment is mostly heterogeneous, with every decision becoming extremely expensive, while the available data remains highly sporadic in nature. However, the petroleum production period being longer, efficient use of AI/ML could be expected to offer long-term gains when it merges hands with the scientific-based reservoir simulation. And, AI/ML algorithms are expected to play a crucial role in addressing the challenges associated with (a) large amounts of biased data (not all the rock, fluid, rock-fluid data uniformly); (b) various data associated with various scales; (c) a significant amount of inconsistent and inaccurate data; and (d)

high rate of biased data influx (again, not, all the rock, fluid, rock-fluid data influx uniformly) (**Koroteev and Tekic, 2021**). On top of it, now, AI/ML would be able to scrutinize, filter and select the daily produced data from downhole as well as from surface sensors effectively as against the conventional structured and unstructured data used in the petroleum industry to keep track of production, maintenance and safety. Of course, in the petroleum industry, securing accurate data either becomes nearly impossible or extremely expensive and eventually, whether we will be able to provide a sufficient quantity and quality data for training, verification and validation remains a crucial task. The other drawback associated with the AI/ML approach is the time delay associated with the data processing because a significant amount of data is associated with high uncertainty.

In general, conventional scientific engineering analysis using reservoir simulators often faces challenges due to the extensive data requirements for reservoir rock properties, fluid properties, and rock-fluid interactions. Introducing AI techniques can help identify and model the complex non-linear relationships between these parameters, improving our understanding of their interactions. However, the success of AI depends heavily on collecting and filtering credible data from reservoir sites. While AI techniques offer potential, relying solely on them without incorporating reservoir simulation can lead to inaccurate correlations, unrealistic clustering, and biases due to missing data or unstable oil-water fronts (**Ertekin and Sun, 2019**). This highlights the need for a balanced approach. Advanced methods like transfer learning can address this gap by starting with a reservoir simulation model and refining it using AI/ML with specific training data. Combining AI/ML with reservoir simulation can reduce the computational challenges of solving complex coupled PDEs and provide deeper insight into petroleum reservoirs. This hybrid approach offers a more reliable and efficient way to model and analyze reservoirs.

It is now clearly known that reservoir simulation has a very strong scientific basis, while AI has a very strong art basis. If the petroleum industry makes an attempt to couple these two approaches in the best possible way, then the characterization of fluid flow through petroleum reservoirs will end up with minimal uncertainty while gaining the maximum benefit. The concept of reservoir simulation remains well known in the petroleum industry, while the application of AI in the petroleum industry is not in full swing. The way various data needs to be trained and the appropriate selection of algorithms meant for the successful estimation of reservoir rock and fluid parameters require a special skill. In essence, merging science and art elevates multi-phase fluid flow analysis to the next level. By integrating AI with reservoir simulation, we preserve traditional reservoir simulation practices while embracing advancements in AI. This approach enables the petroleum industry to leverage the

Figure 5. Key steps in data preprocessing, processing, training and testing



strengths of both methods, leading to more effective characterization of multi-phase fluid flow in petroleum reservoirs.

5. Conclusions

This article highlights the limitations of the current approach to reservoir simulation, particularly the reliance on petroleum software packages without a solid understanding of reservoir physics and mathematical principles. It also critiques the use of machine learning and artificial intelligence in reservoir characterization, viewing it more as an art than a rigorous science. Finally, the study explores the integration of reservoir simulation and AI, discussing its potential to transform the petroleum industry by enabling more advanced and accurate fluid flow analysis.

The following conclusions have been deduced from the present study.

1. The current approach to reservoir simulation through numerical modelling often ignores the fundamental physics of reservoir behaviour and the formulation of conceptual and mathematical models using applied reservoir mathematics. Instead, the focus is mainly on developing advanced numerical solution techniques, many of which do not follow the basic principle of fluid mass conservation.
2. Using petroleum reservoir simulation software without developing conceptual, mathematical, or numerical models – and relying on input data that is often not justified in terms of scale and uniformity – is not suitable for academic purposes. This method is more appropriate for industry professionals who have the expertise to understand different aspects of modelling. In an academic setting, fresh graduates may lack the necessary
3. knowledge to correctly interpret simulation results. Industry experts, on the other hand, can carefully analyze results while considering the limitations and challenges at each stage, leading to better interpretations.
4. In AI applications, unlike traditional reservoir simulation, pre-processing of data is a very important step where proper measures are taken before using the data for processing. The introduction of AI has made reservoir science more dependent on pattern recognition, turning it into an art rather than pure science. However, depending only on data manipulation without understanding the actual physical principles of reservoir drainage may not be a sustainable approach in the long run, as it moves away from the scientific foundation.
5. By combining “art-based AI” with “science-based reservoir simulation,” we can continue the traditional practice of reservoir simulation while adopting AI advancements. This hybrid approach allows the petroleum industry to benefit from the strengths of both methods, leading to better characterization of multiphase fluid flow in reservoirs. It prevents a complete shift from knowledge-based models to only data-driven AI/ML techniques, ensuring that AI supports rather than replaces traditional reservoir simulation.
5. Finally, this study highlights the importance of fundamental physics in reservoir simulation and its role in developing AI-based models. It also provides guidance for entry-level reservoir engineers on how to approach reservoir simulation problems, explaining the essential steps required for effectively simulating subsurface processes. Additionally, thought-provoking questions have been included to encourage further research in this field.

Acknowledgement

The authors greatly acknowledge the support of the research from the Indian Institute of Technology, Madras.

Conflict of interest

The authors declare that there is no financial or personal interest to research work associated with this paper.

6. References

- Ahmed S, A., Elkatatny, S., Ali, A. Z., Abdulraheem, A., & Mahmoud, M. (2019). Artificial Neural Network ANN Approach to Predict Fracture Pressure. *Presented at the SPE Middle East Oil and Gas Show and Conference*. <https://doi.org/10.2118/194852-ms>
- Al-Rbeawi, S. (2023). A Review of Modern Approaches of Digitalization in Oil and Gas Industry. *Upstream Oil and Gas Technology*, 11, 100098. <https://doi.org/10.1016/j.upstre.2023.100098>
- Ansari, M. I., and Govindarajan, S. K. (2022). One-dimensional heat distribution simulation in heavy oil reservoirs during steam flooding methods. *OCEANS 2022 - Chennai*, 1–8. <https://doi.org/10.1109/OCEANSChennai45887.2022.9775245>
- Ansari, M. I., and Govindarajan, S. K. (2023). Numerical investigation on the impact of initial water saturation distribution on hot water flooding performance under non-isothermal conditions. *Rudarsko-Geološko-Naftni Zbornik*, 38(2), 143–155. <https://doi.org/10.17794/rgn.2023.2.11>
- Ansari, M. I., and Govindarajan, S. K. (2024). Numerical Simulation Study to Analyze the Cold Thermal Front Growth in Discrete Fracture Geothermal Reservoir. *Fifth EAGE Global Energy Transition Conference and Exhibition (GET 2024)*, 1–5. <https://doi.org/10.3997/2214-4609.202421318>
- Balza, A., and Li, Y. (2020). Integrating time-lapse gravity, production, and geologic structure data in a gas reservoir study. *Interpretation*, 8(4), SS129–SS143. <https://doi.org/10.1190/INT-2019-0272.1>
- Bhattacharya, S. (2021). *A Primer on Machine Learning in Subsurface Geosciences*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-71768-1>
- Chakraborty, S., Govindarajan, S. K., and Gummadi, S. N. (2020). Numerical Modeling on the Influence of Effective Porosity, Microbial Kinetics, and Operational Parameters on Enhanced Oil Recovery by Microbial Flooding Within a Sandstone Formation. *SPE Journal*, 25(06), 2932–2961. <https://doi.org/10.2118/200639-PA>
- Costa, L. A. N., Maschio, C., & Schiozer, D. J. (2014). Application of artificial neural networks in a history matching process. *Journal of Petroleum Science and Engineering*, 123, 30–45. <https://doi.org/10.1016/j.petrol.2014.06.004>
- de la Mata, F. F., Gijón, A., Molina-Solana, M., & Gómez-Romero, J. (2023). Physics-informed neural networks for data-driven simulation: Advantages, limitations, and opportunities. *Physica A: Statistical Mechanics and its Applications*, 610, 128415. <https://doi.org/10.1016/j.physa.2022.128415>
- Devarapu, S. R., Dau, T. K. D., Kudapa, V. K., and Govindarajan, S. K. (2023). Compositional numerical analysis of multiphase flow of crude oil in porous media under non-isothermal conditions. *International Journal of Chemical Reactor Engineering*, 21(12), 1465–1476. <https://doi.org/10.1515/ijcre-2023-0019>
- Dinesh, N. S. V., Pavan, P. S., and Sivasankar, P. (2024). Decarbonization of Alkali–Surfactant–Polymer Flooding Enhanced Oil Recovery Process with Hydrogen under Uncertainty. *Energy and Fuels*, 38(12), 11093–11106. <https://doi.org/10.1021/acs.energyfuels.4c01872>
- Ertekin, T., & Sun, Q. (2019). Artificial intelligence applications in reservoir engineering: a status check. *energies*, 12(15), 2897. <https://doi.org/10.3390/en12152897>
- Farahi, M. M., Ahmadi, M., & Dabir, B. (2021). Model-based water-flooding optimization using multi-objective approach for efficient reservoir management. *Journal of Petroleum Science and Engineering*, 196, 107988. <https://doi.org/10.1016/j.petrol.2020.107988>
- Govindarajan, S. K., Kumar, A., Mishra, A., and Pavan, T. N. V. (2022). Conceptual Modelling of Low Salinity Water Flooding in a Fractured Reservoir. *Petroleum and Coal*, 62(4), 1435–1448. https://www.researchgate.net/publication/333133950_STORAGE_TANK_PROTECTION_USING Aspen_HYSYS
- Hemmati-Sarapardeh, A., Larestani, A., Amar, M. N., and Hajirezaie, S. (2020). *Applications of Artificial Intelligence Techniques in the Petroleum Industry*. Elsevier. <https://doi.org/10.1016/C2018-0-04421-7>
- Jia, J., Fan, Q., Jing, J., Lei, K., & Wang, L. (2024). Intelligent hydraulic fracturing under industry 4.0 – a survey and future directions. *Journal of Petroleum Exploration and Production Technology*, 14(12), 3161–3181. <https://doi.org/10.1007/s13202-024-01877-6>
- Jo, H., Pan, W., Santos, J. E., Jung, H., & Pyrcz, M. J. (2021). Machine learning assisted history matching for a deepwater lobe system. *Journal of Petroleum Science and Engineering*, 207, 109086. <https://doi.org/10.1016/j.petrol.2021.109086>
- Kalantari-Dahaghi, A., & Mohaghegh, S. D. (2011, March). Numerical simulation and multiple realizations for sensitivity study of shale gas reservoir. In *SPE Oklahoma City Oil and Gas Symposium/Production and Operations Symposium* (pp. SPE-141058). SPE. <https://doi.org/10.2118/141058-ms>
- Kandala, V., and Govindarajan, S. K. (2023). Numerical Investigations of Low-Salinity Water Flooding in a Saline Sandstone Reservoir. *Journal of Energy Engineering*, 149(5). <https://doi.org/10.1061/JLEED9.EYENG-4893>
- Koray, A.-M., Bui, D., Kubi, E. A., Ampomah, W., and Amosu, A. (2024). Machine learning based reservoir characterization and numerical modeling from integrated well log and core data. *Geoenergy Science and Engineering*, 243, 213296. <https://doi.org/10.1016/j.geoen.2024.213296>
- Koroteev, D., and Tekic, Z. (2021). Artificial intelligence in oil and gas upstream: Trends, challenges, and scenarios for

- the future. *Energy and AI*, 3, 100041. <https://doi.org/10.1016/j.egyai.2020.100041>
- Kronberger, P., Dabrowski, P., Chacon, J., and Bangert, P. (2020, November 2). The Digitalization Journey of the Brage Digital Twin. *Day 2 Tue, November 03, 2020*. <https://doi.org/10.2118/200728-MS>
- Kudapa, V. K., Sharma, P., Kunal, V., and Gupta, D. K. (2017). Modeling and simulation of gas flow behavior in shale reservoirs. *Journal of Petroleum Exploration and Production Technology*, 7(4), 1095–1112. <https://doi.org/10.1007/s13202-017-0324-4>
- Lawal, A., Yang, Y., He, H., and Baisa, N. L. (2024). Machine Learning in Oil and Gas Exploration: A Review. *IEEE Access*, 12, 19035–19058. <https://doi.org/10.1109/ACCESS.2023.3349216>
- Li, W., Wang, L., Dong, Z., Wang, R., & Qu, B. (2022). Reservoir production prediction with optimized artificial neural network and time series approaches. *Journal of Petroleum Science and Engineering*, 215, 110586. <https://doi.org/10.1016/j.petrol.2022.110586>
- Li, W., Zhang, T., Liu, X., Dong, Z., Dong, G., Qian, S., ... & Zhang, T. (2024). Machine learning-based fracturing parameter optimization for horizontal wells in Panke field shale oil. *Scientific Reports*, 14(1), 6046. <https://doi.org/10.1038/s41598-024-56660-8>
- Li, Y., Zhang, T., & Sun, S. (2019). Acceleration of the NVT flash calculation for multicomponent mixtures using deep neural network models. *Industrial & Engineering Chemistry Research*, 58(27), 12312–12322. <https://doi.org/10.1021/acs.iecr.9b00527>
- Lim, J.-S. (2005). Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea. *Journal of Petroleum Science and Engineering*, 49(3–4), 182–192. <https://doi.org/10.1016/j.petrol.2005.05.005>
- Liu, X., Ge, Q., Chen, X., Li, J., and Chen, Y. (2021). Extreme learning machine for multivariate reservoir characterization. *Journal of Petroleum Science and Engineering*, 205, 108869. <https://doi.org/10.1016/j.petrol.2021.108869>
- Mirzaei, M., and Das, D. B. (2007). Dynamic effects in capillary pressure–saturation relationships for two-phase flow in 3D porous media: Implications of micro-heterogeneities. *Chemical Engineering Science*, 62(7), 1927–1947. <https://doi.org/10.1016/j.ces.2006.12.039>
- Mohaghegh, S. D. (2011). Reservoir simulation and modeling based on artificial intelligence and data mining (AI&DM). *Journal of Natural Gas Science and Engineering*, 3(6), 697–705. <https://doi.org/10.1016/j.jngse.2011.08.003>
- Mohaghegh, S., Popa, A., & Ameri, S. (1999, October). Intelligent systems can design optimum fracturing jobs. In *SPE Eastern Regional Meeting* (pp. SPE-57433). SPE. <https://doi.org/10.2118/57433-ms>
- Muradkhanli, L. (2018). Neural networks for prediction of oil production. *IFAC-PapersOnLine*, 51(30), 415–417. <https://doi.org/10.1016/j.ifacol.2018.11.339>
- Nande, S. (2018, September). Application of machine learning for closure pressure determination. In *SPE Annual Technical Conference and Exhibition*. SPE. <https://doi.org/10.2118/194042-stu>
- Noshi, C. I., and Schubert, J. J. (2018, October 7). The Role of Machine Learning in Drilling Operations; A Review. *Day 4 Wed, October 10, 2018*. <https://doi.org/10.2118/191823-18ERM-MS>
- Pavan, T. N. V., Devarapu, S. R., and Govindarajan, S. K. (2024). Characterization of Leaky Deep Saline Aquifer for Storing sc-CO₂. *Energy and Fuels*, 38(12), 11051–11063. <https://doi.org/10.1021/acs.energyfuels.4c01599>
- Pavan, T. N. V., Devarapu, S. R., Kudapa, V. K., and Govindarajan, S. K. (2023). Numerical investigations on sc-CO₂ gas sequestration in layered heterogeneous deep saline aquifers. *International Journal of Chemical Reactor Engineering*, 21(12), 1477–1485. <https://doi.org/10.1515/ijcre-2023-0041>
- Pavan, T. N. V., and Govindarajan, S. K. (2023). Numerical investigations on performance of sc-CO₂ sequestration associated with the evolution of porosity and permeability in low permeable saline aquifers. *Geoenergy Science and Engineering*, 225, 211681. <https://doi.org/10.1016/j.geoen.2023.211681>
- Ramgulum, A., Ertekin, T., & Flemings, P. B. (2007). Utilization of artificial neural networks in the optimization of history matching. In *SPE Latin America and Caribbean Petroleum Engineering Conference*. SPE. <https://doi.org/10.2118/107468-ms>
- Reddya, D. S., and Kumarb, G. S. (2014). A Numerical Investigation on the role of Oil Saturations on Performance of In-Situ Combustion in Porous Media. *International Journal of Scientific T Engineering Research*, 5(5), 531–538.
- Ren, Y., Cai, H., Hou, Y., Bao, H., & Zhang, L. (2024). Oilfield Production Prediction Method Based on PCA and Deep LSTM Neural Network. In *2024 International Conference on New Trends in Computational Intelligence (NTCI)*. IEEE.
- Sircar, A., Yadav, K., Rayavarapu, K., Bist, N., & Oza, H. (2021). Application of machine learning and artificial intelligence in oil and gas industry. *Petroleum Research*, 6(4), 379–391.
- Sivasankar, P., and Suresh Kumar, G. (2018). Modelling the influence of interaction between injection and formation brine salinities on in-situ microbial enhanced oil recovery processes by coupling of multiple-ion exchange transport model with multiphase fluid flow and multi-species reactive transport. *Journal of Petroleum Science and Engineering*, 163(August 2017), 435–452. <https://doi.org/10.1016/j.petrol.2018.01.004>
- Solomatine, D. P., and Ostfeld, A. (2008). Data-driven modeling: some past experiences and new approaches. *Journal of Hydroinformatics*, 10(1), 3–22. <https://doi.org/10.2166/hydro.2008.015>
- Srinivasa Reddy, D., and Suresh Kumar, G. (2015). Numerical simulation of heavy crude oil combustion in porous combustion tube. *Combustion Science and Technology*, 187(12), 1905–1921.
- Srinivasan, S., O'Malley, D., Mudunuru, M. K., Sweeney, M. R., Hyman, J. D., Karra, S., Frash, L., Carey, J. W., Gross, M. R., Guthrie, G. D., Carr, T., Li, L., & Viswanathan, H. S. (2021). A machine learning framework for rapid forecasting and history matching in unconventional reservoirs.

- Scientific Reports*, 11(1), 21730. <https://doi.org/10.1038/s41598-021-01023-w>
- Venkata Pavan, T. N., Devarapu, S. R., and Suresh Kumar, G. (2024). Role of capillary pressure on carbon dioxide sequestration : mathematical analysis. *International Journal of Coal Preparation and Utilization*, 00(00), 1–15. <https://doi.org/10.1080/19392699.2024.2411549>
- Wang, M., Hui, G., Pang, Y., Wang, S., & Chen, S. (2023). Optimization of machine learning approaches for shale gas production forecast. *Geoenery Science and Engineering*, 226, 211719.
- Zhang, K., Wang, X., Ma, X., Wang, J., Yang, Y., Zhang, L., & Yao, J. (2022). The prediction of reservoir production based proxy model considering spatial data and vector data. *Journal of Petroleum Science and Engineering*, 208, 109694. <https://doi.org/10.1016/j.jpetoen.2023.211719>
- Zhang, T., Li, Y., Li, Y., Sun, S., & Gao, X. (2020). A self-adaptive deep learning algorithm for accelerating multi-component flash calculation. *Computer Methods in Applied Mechanics and Engineering*, 369, 113207.
- Zhong, Z., Sun, A. Y., Wang, Y., & Ren, B. (2020). Predicting field production rates for waterflooding using a machine learning-based proxy model. *Journal of Petroleum Science and Engineering*, 194, 107574.
- Zhou, W., Liu, C., Liu, Y., Zhang, Z., Chen, P., and Jiang, L. (2024). Machine Learning in Reservoir Engineering: A Review. *Processes*, 12(6), 1219. <https://doi.org/10.3390/pr12061219>

SAŽETAK

Primjena umjetne inteligencije u simulaciji ležišta za karakterizaciju protoka višefaznoga fluida kroz naftna ležišta

Umjetna inteligencija (AI) brzo napreduje i utječe na sva znanstvena područja, uključujući i naftnu industriju, gdje ona više nije novost. Ovaj članak istražuje razvoj umjetne inteligencije i njezinu integraciju s tradicionalnim pristupima simulacije ležišta. Prvi dio rada ističe sve manji naglasak na konceptualno i matematičko modeliranje u simulaciji ležišta, uz rastući fokus na sofisticirane tehnike numeričkoga rješenja. Ovaj pristup često zanemaruje temeljnu fiziku i matematiku ležišta, što dovodi do površne karakterizacije višefaznoga protoka fluida u naftnim ležištima. U drugome dijelu rada analizirano je često korištenje softverskih paketa u naftnoj industriji, koji se u velikoj mjeri oslanjaju na ulazne podatke, a da pritom ne uzimaju u obzir varijacije u skalama podataka. Ti alati tretiraju temeljno programiranje kao „crnu kutiju”, često zaobilazeći ključne temeljne znanstvene discipline poput konceptualizacije ležišta, primijenjene matematike i numeričkih tehnika, što rezultira nepotpunom karakterizacijom ležišta. U trećemu dijelu rada razmatrana je uloga strojnoga učenja (ML) i umjetne inteligencije u primjenama na ležištima. Dok znanost o podacima igra ključnu ulogu, nedostatak integracije s temeljnim fizičkim zakonitostima ležišta svodi analizu protoka fluida na „umjetnost” bez znanstvene strogosti. U zadnjemu dijelu rada predložen je hibridni pristup koji povezuje AI/ML s tradicionalnom simulacijom ležišta. Ova integracija premošćuje jaz između znanstveno utemeljene simulacije ležišta i analize protoka fluida vođene umjetnom inteligencijom, omogućujući naftnoj industriji da postigne novu paradigmu za analizu višefaznoga protoka fluida. Kombinirajući temeljnu znanost s naprednim tehnikama umjetne inteligencije, ovaj pristup nudi sveobuhvatni okvir za preciznu karakterizaciju ležišta i poboljšanu proizvodnju ugljikovodika.

Ključne riječi:

umjetna inteligencija, strojno učenje, simulacija ležišta, konceptualni model, matematički model, numerički model

Author's Contribution

Viswakanth Kandala (Research Scholar): conceptualization, writing, tables and figures. **Suresh Kumar Govindarajan** (Full Professor): conceptualization, writing the original draft, and supervision. **Tummuri Naga Venkata Pavan** (Research Scholar): writing, figures, editing, and formatting. **Swaminathan Ponamani** (Assistant Professor): tables and editing. **Srinivasa Reddy Devarapu** (Associate Professor): conceptualization, editing. All authors have read and agreed to the published version of the manuscript.