Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Rock Typing in the Digital Age: A Systematic Review of Petrophysical Foundations, Advanced Conventional Formulations, and the Ascent of Physics-Informed Artificial Intelligence

Zainulabdeen Hasan

salahhasan@graduate.utm.my

University of Technology Malaysia

Mohd Zaidi Jaffar

University of Technology Malaysia

Amni Haslinda Alpandi

University of Technology Malaysia

Sharul Nizam

University of Technology Malaysia

Ahmed Khoudhair

University of Baghdad

Systematic Review

Keywords: Petrophysical Rock Typing (PRT), Hydraulic Flow Units (HFU), Flow Zone Indicator (FZI), Digital Rock Physics (DRP), Machine Learning, Physics-Informed Neural Networks (PINN), Explainable AI (XAI), Carbonate Reservoirs, Tight Gas Sandstones

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11428459/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Rock Typing in the Digital Age: A Systematic Review of Petrophysical Foundations, Advanced Conventional Formulations, and the Ascent of Physics-Informed Artificial Intelligence

Zainulabdeen Hasan[1,*], Mohd Zaidi Jaffar[1], Amni Haslinda Alpandi[1], Sharul Nizam[1], and Ahmed Khoudhair[2]

[1] School of Chemical and Energy Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), 81310 Johor Bahru, Johor, Malaysia

[2] Department of Petroleum Engineering, College of Engineering, University of Baghdad, Baghdad, Iraq

** Corresponding author: Zainulabdeen Hasan (salahhasan@graduate.utm.my); Co-authors: mzaidi@utm.my, amnihaslinda@utm.my, sharulnizam@utm.my, Dr.a.alameeny@gmail.com*

Abstract — Petrophysical rock typing (PRT) is the cornerstone of reservoir characterization, bridging microscopic pore-throat morphology and macroscopic dynamic flow behavior to control hydrocarbon-in-place estimates, saturation–height functions, and numerical simulation grid-block properties. For over five decades, reservoir workflows have relied on conventional frameworks rooted in capillary pressure curves (Leverett, 1941; Thomeer, 1960), pore-throat size calibrations (Winland R_{35}; Pittman, 1992), and the modified Kozeny-Carman formulation, most notably the Reservoir Quality Index (RQI) and Flow Zone Indicator (FZI) proposed by Amaefule et al. (1993). However, these classical formulations frequently struggle in complex carbonate systems, tight gas sandstones, and multi-modal reservoirs due to fundamental physical limitations, including the conflation of static capillary storage with dynamic flow capacity, non-Darcy flow regimes, and unresolved multi-scale heterogeneity. This review systematically traces the methodological evolution of rock typing from its historical foundations to the contemporary ascent of Artificial Intelligence (AI) and Machine Learning (ML). We first evaluate advanced conventional advancements that address classical shortcomings, including the rigorous decoupling of Petrophysical Static Rock Types (PSRT) from Petrophysical Dynamic Rock Types (PDRT), the calibration of effective flow porosity (PHIZ) via Purcell-PSKT porosimetry, and fractal bound-water models for low-permeability rocks. We then conduct an in-depth critical analysis of data-driven AI, examining how unsupervised clustering (Self-Organizing Maps, K-Means, Fuzzy C-Means, DBSCAN) automates electrofacies discovery from multi-well wireline logs, and how supervised tree-based ensembles (Random Forest, XGBoost, LightGBM) and deep architectures predict rock classes and continuous permeability. Crucially, we highlight the paradigm shift toward Physics-Informed Machine Learning (PIML), where non-Darcy flow physics and mass conservation are embedded directly into loss functions, and Explainable AI (SHAP, LIME) is leveraged to eliminate black-box opacity. We further analyze the revolution in Digital Rock Physics (DRP) and Image-based Microscale Rock Typing (IMRT), which combines micro-CT, FIB-SEM, and deep learning segmentation (PRRT and TSRT) with Lattice Boltzmann flow simulation to provide verifiable ground truth. Finally, we replace speculative risks with an industry-grounded audit of genuine operational pitfalls in AI petrophysics — including spatial autocorrelation data leakage, out-of-distribution extrapolation failure, core-to-log depth-matching errors, and regulatory compliance — and synthesize these insights into a unified, multi-scale, physics-informed digital rock typing roadmap for resilient reservoir modeling.

Keywords: *Petrophysical Rock Typing (PRT); Hydraulic Flow Units (HFU); Flow Zone Indicator (FZI); Digital Rock Physics (DRP); Machine Learning; Physics-Informed Neural Networks (PINN); Explainable AI (XAI); Carbonate Reservoirs; Tight Gas Sandstones*

1. Introduction

Accurate reservoir characterization is the foundational technical prerequisite for evaluating hydrocarbon assets, optimizing well placement, designing improved oil recovery (IOR) and enhanced oil recovery (EOR) schemes, and de-risking geological carbon storage (GCS) and underground gas storage (UGS) projects (Bachu, 2008; Skalinski & Kenter, 2015; Gholami et al., 2021). At the core of this characterization workflow lies petrophysical rock typing (PRT): the process of segmenting a structurally and texturally heterogeneous formation into a discrete, manageable number of rock classes that share coherent geological attributes, pore-throat size distributions, capillary pressure behavior, and dynamic fluid transport characteristics (Amaefule et al., 1993; Lucia, 1995; Mirzaei-Paiaman et al., 2018). In static geological modeling, rock types govern the spatial propagation of porosity, permeability, and water saturation between sparse well penetrations; in dynamic numerical simulation, they dictate the allocation of saturation–height functions and multiphase relative permeability curves to individual grid blocks (Warren & Root, 1963; Camacho-Velázquez et al., 2005).

A fundamental challenge in reservoir characterization is that subsurface geological media exhibit pronounced spatial heterogeneity spanning more than nine orders of magnitude in length scale (Figure 1 and Figure 4). At the microscopic pore scale (nanometers to micrometers), mineralogy, pore-body geometry, throat aspect ratios, and surface roughness dictate capillary trapping, wettability, and non-Darcy Knudsen diffusion (Clerke et al., 2008; Ji et al., 2022; Wang, 2024). At the laboratory core-plug scale (centimeters), routine core analysis (RCAL) and special core analysis (SCAL) provide bulk measurements of porosity, gas permeability, and drainage capillary pressure (Leverett, 1941; Amyx et al., 1960; Byrnes et al., 1999). At the borehole wireline log scale (meters), logging tools measure volume-averaged physical responses (gamma ray, density, neutron porosity, sonic transit time, resistivity) across a vertical resolution of 0.5 to 2.0 ft, frequently averaging over thin laminations and vuggy clusters (Lucia, 1995; Arifianto et al., 2018). Finally, at the inter-well and 3D seismic scale (tens to hundreds of meters), reservoir engineers must assign effective flow and storage properties to simulation cells that may each exceed 50 m × 50 m × 2 m in dimension (Wu et al., 2019; Ashraf et al., 2024a).

Historically, this multi-scale gap was bridged using semi-empirical petrophysical models, notably the Winland R_{35} empirical method (Kolodzie, 1980), Pittman's (1992) pore-throat aperture equations, and the Flow Zone Indicator (FZI) based on the modified Kozeny-Carman framework (Amaefule et al., 1993). While these classical approaches provided an intuitive, physics-grounded foundation that became universal industry standard, they exhibit severe limitations when applied to complex carbonate formations, tight gas sandstones, and multi-modal fractured-vuggy reservoirs. In particular, classical FZI treats all interconnected pore space as equally conductive, conflating static capillary storage with dynamic flow capacity (Rabiller, 2017; Mirzaei-Paiaman et al., 2018). Furthermore, conventional rock typing is constrained by the scarcity of physical core: in typical development fields, fewer than 10–15% of drilled wells are cored, forcing petrophysicists to

manually propagate core-derived rock types across thousands of feet of uncored intervals using subjective cutoffs and crossplots (Cuddy, 2021; Syed et al., 2022).

Over the past decade, the rapid ascent of Artificial Intelligence (AI) and Machine Learning (ML), combined with the emergence of Digital Rock Physics (DRP), has triggered a profound paradigm shift in petrophysical rock typing (Figure 1). Unsupervised learning algorithms (such as Self-Organizing Maps, K-Means, and DBSCAN) now enable automated, data-driven electrofacies clustering across massive multi-well log databases (Ashraf et al., 2024a, 2024b). Supervised gradient-boosted trees and deep neural networks predict discrete rock classes and continuous permeability profiles in uncored intervals with high statistical fidelity (Chen & Guestrin, 2016; Kumar & Ghosh, 2025). Concurrently, high-resolution micro-computed tomography (μ-CT) and Focused Ion Beam Scanning Electron Microscopy (FIB-SEM) allow direct three-dimensional visualization of pore networks, enabling in-silico flow modeling via the Lattice Boltzmann Method (LBM) and pore-network modeling (PNM) to establish unambiguous physical ground truth (Wang, 2024). More recently, Physics-Informed Machine Learning (PIML) and Explainable AI (XAI) frameworks have emerged to enforce governing hydro-mechanical conservation laws and illuminate the inner decision-making logic of AI surrogates (Raissi et al., 2019; Karniadakis et al., 2021; Lundberg & Lee, 2017; Wang et al., 2024).

Despite these remarkable technological strides, the petrophysical literature remains fragmented across specialized niches. Geologists frequently prioritize depositional and diagenetic facies classifications that fail to map onto distinct flow behaviors; petrophysicists often adhere to classical FZI methods without accounting for modern dynamic corrections; and computer scientists frequently deploy purely data-driven algorithms without evaluating physical consistency, spatial autocorrelation leakage, or out-of-distribution failure modes. This paper provides a comprehensive, systematic review designed to unify these domains. Specifically, this review: (1) critically examines the theoretical foundations, governing equations, and historical evolution of conventional rock typing methods (Section 2); (2) evaluates advanced physical modifications that decouple static and dynamic behavior (Section 3); (3) reviews unsupervised machine learning workflows for multi-well log clustering (Section 4); (4) analyzes supervised and physics-informed ML architectures for property prediction and explainability (Section 5); (5) details the digital rock physics revolution and microscale image-based rock typing (Section 6); (6) provides a multi-dimensional comparative synthesis and rigorous industry risk audit of AI pitfalls (Section 7); (7) outlines a forward-looking roadmap toward multi-scale, physics-informed digital rock typing (Section 8); and (8) concludes with key technical takeaways (Section 9).

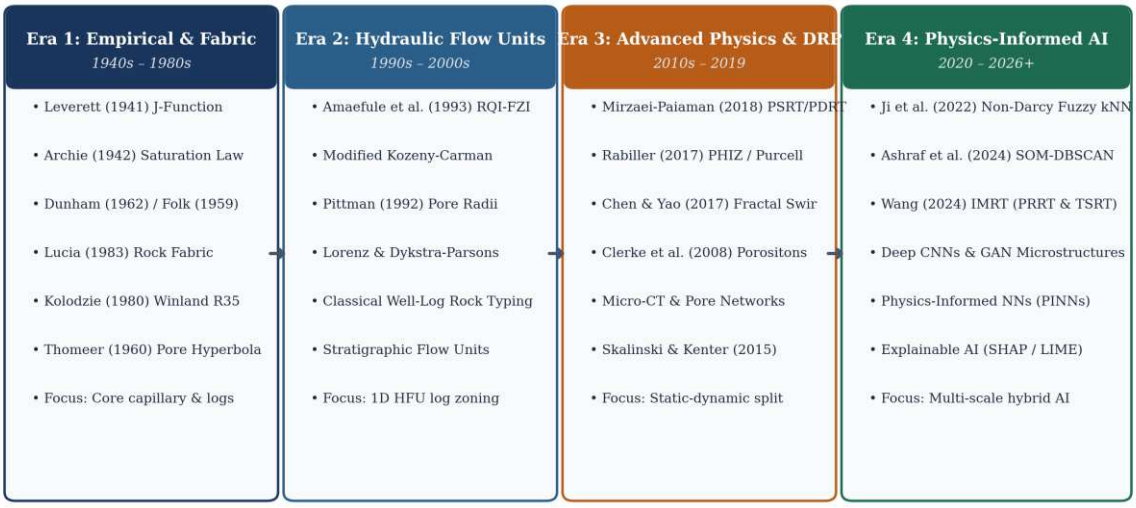


Figure 1. Historical timeline and paradigm shifts in petrophysical rock typing over eight decades (1940s to 2026+), illustrating the transition across four distinct eras: Era 1 (Empirical & Fabric Foundations), Era 2 (Hydraulic Flow Units & Kozeny-Carman), Era 3 (Advanced Physics Refinements & Early DRP), and Era 4 (Physics-Informed AI, Microscale Imaging, and Multi-Scale Digital Workflows).

2. Classical Foundations of Petrophysical Rock Typing

2.1 Geological & Texture-Based Classification Systems

The earliest formal rock typing frameworks originated in sedimentary carbonate petrology. Dunham (1962) and Folk (1959) established depositional texture classifications (mudstone, wackestone, packstone, grainstone, boundstone) based on depositional energy and matrix content. While indispensable for paleo-environmental reconstruction, depositional facies alone correlate poorly with flow properties because subsequent diagenesis (calcite cementation, dolomitization, meteoric leaching, compaction) extensively overprints original pore geometries (Choquette & Pray, 1970; Lucia, 1995; Skalinski & Kenter, 2015). To bridge sedimentology and petrophysics, Lucia (1983, 1995, 1999) formulated the rock-fabric approach, classifying carbonate pore space into two fundamental categories: (i) interparticle (intergranular and intercrystalline) pore space, where pore-throat size tracks particle/crystal size and sorting (subdivided into Petrophysical Classes 1, 2, and 3), and (ii) vuggy pore space, which is explicitly subdivided into separate vugs (interconnected only through the background interparticle matrix, adding storage without dramatically boosting permeability) and touching vugs (forming continuous, interconnected cavernous and fracture conduits that dominate flow capacity). The Lucia framework demonstrated that permeability in carbonates cannot be predicted from total porosity alone without specifying the rock-fabric class and separate-vug porosity fraction.

2.2 Capillary Pressure & Pore-Throat Radii Methods

In parallel with geological texture classifications, reservoir engineers developed capillary-based rock typing techniques grounded in Mercury Injection Capillary Pressure (MICP) porosimetry. Leverett

(1941) introduced the dimensionless J-function to normalize capillary pressure curves across varying porosity and permeability fields:

$$J(S_w) = [P_c / (\sigma \cos \theta)] \sqrt{(k / \varphi)} \qquad (1)$$

where P_c is capillary pressure, σ is interfacial tension, θ is wetting contact angle, k is permeability, and φ is porosity. Archie (1942) established the foundational electrical resistivity equations linking formation factor, porosity, and water saturation. Purcell (1949) and Burdine (1953) demonstrated that absolute permeability and multiphase relative permeability could be calculated directly by integrating the inverse square of capillary pressure over the wetting-phase saturation distribution:

$$k = 0.5 (\sigma \cos \theta)^2 \varphi \int_0^1 [1 / P_c(S_w)^2] \, dS_w \qquad (2)$$

Thomeer (1960, 1983) advanced capillary modeling by fitting MICP curves with an empirical hyperbola controlled by three parameters: the displacement (entry) pressure P_d, the non-wetting phase pore-volume factor B_v∞, and a dimensionless pore-geometrical factor G reflecting pore-throat sorting. In a seminal field study of the Spindle Field in Colorado, Kolodzie (1980) documented the Winland R_{35} method, an empirical multiple regression model developed by H.D. Winland that correlates routine core porosity and permeability to the modal pore-throat aperture radius corresponding to 35% mercury saturation (R_35):

$$\log_{10}(R_{35}) = 0.732 + 0.588 \log_{10}(k) - 0.864 \log_{10}(\varphi_pct) \qquad (3)$$

where k is permeability in mD, φ_pct is porosity in percent (φ × 100), and R_35 is in micrometers (Figure 2b). Martin et al. (1999) popularized Winland R_{35} as an industry standard, dividing reservoirs into mega-porous (R_35 > 10 μm), macro-porous (2 < R_35 ≤ 10 μm), meso-porous (0.5 < R_35 ≤ 2 μm), and micro-porous (R_35 ≤ 0.5 μm) flow units. Pittman (1992) subsequently expanded the Winland equation across a diverse suite of 396 sandstone samples, publishing empirical formulas for pore-throat radii ranging from 10% to 75% mercury saturation (r_10 through r_75), demonstrating that r_35 and r_25 exhibit the strongest statistical correlations with unconfined permeability.

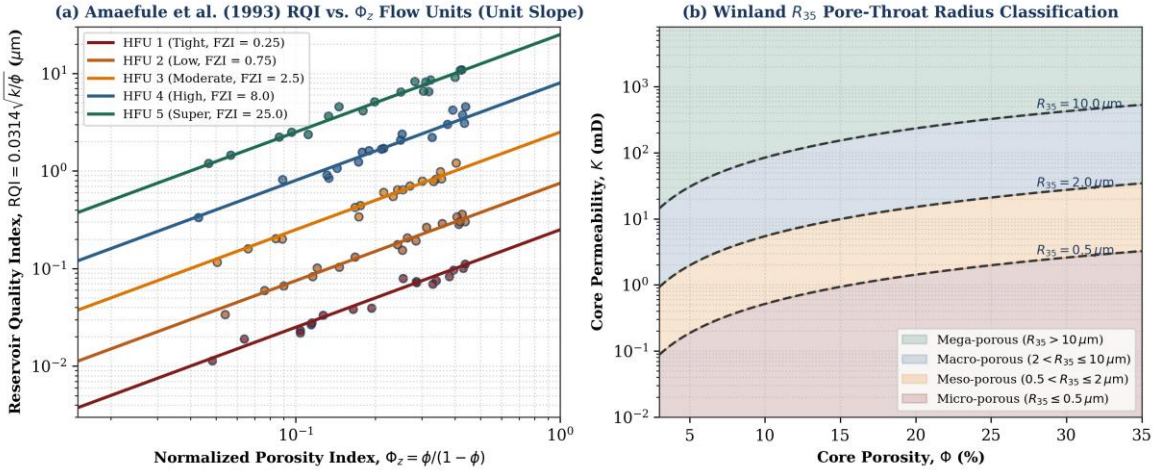


Figure 2. Foundational conventional rock typing frameworks: (a) Log-log diagnostic plot of Reservoir Quality Index (RQI) versus Normalized Porosity Index (Φ_z) based on Amaefule et al. (1993), illustrating five distinct Hydraulic Flow Units (HFUs) exhibiting theoretical unit slopes; (b) Winland R_{35} crossplot of core permeability versus porosity (Kolodzie, 1980; Pittman, 1992), showing the four classic pore-throat aperture domains (Mega-, Macro-, Meso-, and Micro-porous).

2.3 Hydraulic Flow Units (HFU) & The Modified Kozeny-Carman Framework

The seminal contribution of Amaefule et al. (1993) introduced the Reservoir Quality Index (RQI) and Flow Zone Indicator (FZI), establishing the most widely cited quantitative petrophysical rock typing framework in the petroleum industry. Rooted in the Kozeny-Carman equation for laminar fluid flow through a bundle of capillary tortuous tubes (Kozeny, 1927; Carman, 1937, 1956), the generalized flow relationship is formulated as:

$$k = [\varphi_e^3 / (1 - \varphi_e)^2] [1 / (F_s \tau^2 S_{gv}^2)] \quad (4)$$

where φ_e is effective porosity (fraction), F_s is the pore cross-sectional shape factor (typically ~2.0 for circular cylinders), τ is hydraulic tortuosity (ratio of actual fluid path length to straight-line distance), and S_{gv} is the specific surface area per unit grain volume (μm^{-1}). Dividing both sides of Eq. (4) by effective porosity φ_e and taking the square root yields:

$$\sqrt{(k / \varphi_e)} = [\varphi_e / (1 - \varphi_e)] [1 / (\sqrt{(F_s)} \tau S_{gv})] \quad (5)$$

Amaefule et al. (1993) defined the left-hand term as the Reservoir Quality Index (RQI, expressed in micrometers by multiplying by 0.0314 to account for Darcy-to-micrometer unit conversions), the first bracketed right-hand term as the normalized porosity index (Φ_z), and the second bracketed term as the Flow Zone Indicator (FZI):

$$RQI = 0.0314 \sqrt{(k / \varphi_e)}, \quad \Phi_z = \varphi_e / (1 - \varphi_e), \quad FZI = 1 / (\sqrt{(F_s)} \tau S_{gv}) = RQI / \Phi_z \quad (6)$$

Taking the base-10 logarithm of Eq. (6) yields:

$$\log_{10}(RQI) = \log_{10}(\Phi_z) + \log_{10}(FZI) \quad (7)$$

Equation (7) represents a family of straight parallel lines with unit slope (gradient = 1.0) on a log-log plot of RQI versus Φ_z (Figure 2a). All core samples possessing similar pore-throat geometry, tortuosity, and surface area plot along a single line whose intercept at $\Phi_z = 1.0$ equals FZI. Samples clustering within a narrow FZI band define a discrete Hydraulic Flow Unit (HFU). To categorize uncored intervals, petrophysicists traditionally classify cored HFUs using statistical histograms or cumulative probability plots, then develop empirical multi-linear regressions linking FZI to wireline logs.

2.4 Theoretical Constraints and Practical Shortcomings of Classical Frameworks

Despite their universal adoption, classical FZI and Winland methods suffer from four fundamental theoretical and practical limitations: (1) Conflation of storage and flow capacity: Classical FZI treats total effective porosity φ_e as an active contributor to fluid flow. In microporous carbonates and clay-rich shaly sandstones, a large fraction of φ_e resides within ultra-fine pore throats that store hydrocarbons capillary-bound (static storage) without contributing measurably to dynamic permeability (Rabiller, 2017; Mirzaei-Paiaman et al., 2018). (2) Breakdown of Darcy flow assumptions: The Kozeny-Carman derivation assumes viscous laminar Darcy flow through rigid, interconnected cylindrical tubes. In nano-Darcy tight sandstones and organic shales, fluid transport is dominated by gas slippage (Klinkenberg effect), Knudsen diffusion, and multi-layer bound water resistance, rendering the constant FZI assumption invalid (Chen & Yao, 2017; Ji et al., 2022). (3) Subjectivity in HFU boundary selection: The determination of FZI cluster cutoffs using visual inspection of cumulative probability slopes is highly analyst-dependent, leading to inconsistent flow

unit counts across adjacent operators (Syed et al., 2022). (4) Neglect of multi-modal pore systems: Bimodal and trimodal pore structures (such as touching vugs coexisting with chalky micrite) cannot be collapsed onto a single unit-slope line, leading to severe permeability mispredictions.

3. Advanced Conventional Refinements: Decoupling Static and Dynamic Properties

3.1 The Static versus Dynamic Rock Type Paradigm (PSRT vs. PDRT)

To resolve the conflation of storage and transport mechanisms, Mirzaei-Paiaman et al. (2018, 2020) introduced a rigorous conceptual and mathematical distinction between Petrophysical Static Rock Types (PSRT) and Petrophysical Dynamic Rock Types (PDRT). PSRTs are governed by capillary pressure curves that dictate initial fluid saturation distributions above the free-water level (static state), whereas PDRTs are governed by fluid flow capacity and relative permeability characteristics (dynamic state). Re-examining the Young-Laplace and Hagen-Poiseuille equations, Mirzaei-Paiaman et al. (2018) proved that the capillary entry pressure group is controlled by a Petrophysical Static Rock Type Index (PSRTI):

$$PSRTI = 0.0314 \sqrt{[k / (\varphi_e F_s \tau)]} \qquad (8)$$

whereas fluid flow conductance is governed by a revised dynamic flow zone indicator FZI* that removes the invalid normalized porosity group Φ_z:

$$FZI^* = 0.0314 \sqrt{(k / \varphi_e)} \qquad (9)$$

Through multi-field carbonate core analysis, Mirzaei-Paiaman et al. (2018) demonstrated that core plugs exhibiting nearly identical FZI* values can possess markedly different capillary retention curves (different PSRTs), and conversely, samples with identical irreducible saturations can exhibit disparate permeability behaviors (different PDRTs). Decoupling static from dynamic rock types prevents the systematic initialization errors that arise when dynamic flow units are erroneously used to calibrate capillary transition zones.

3.2 Effective Flow Porosity Calibration (PHIZ)

In complex carbonate and microporous reservoirs, Rabiller (2017) addressed the FZI over-prediction problem by introducing PHIZ: the fraction of total porosity associated with the primary, flow-effective pore-throat system. Utilizing Purcell capillary pressure integrals and porosimetry deconvolution, Rabiller (2017) demonstrated that the Purcell-PSKT model can identify the critical threshold throat radius (R_cut) below which pore throats contribute negligibly to total Darcy permeability. By substituting PHIZ in place of total porosity φ_e in Eq. (6), one obtains the "true" effective Flow Zone Indicator (FZI_true):

$$FZI_true = [0.0314 \sqrt{(k / \mathrm{PHIZ})}] / [\mathrm{PHIZ} / (1 - \mathrm{PHIZ})] \qquad (10)$$

Calibrating FZI with PHIZ eliminates the artificial dispersion of carbonate core data across multiple spurious HFUs, collapsing samples with identical macropore geometry onto a single, physically coherent flow unit.

3.3 Non-Linear and Non-Darcy Flow Adjustments in Tight Rocks

In tight gas sandstones and deep unconventional reservoirs (k < 0.1 mD), the irreducible water film adsorbed onto hydrophilic pore walls occupies a substantial fraction of the effective throat radius, while gas molecules experience slip flow along solid surfaces (Klinkenberg effect). Chen & Yao (2017) formulated an enhanced fractal permeability model for low-permeability porous media that modifies the classical Hagen-Poiseuille capillary tube equation to explicitly incorporate the bound water saturation (S_wir) and pore-throat fractal dimension (D_f):

$$k_{fractal} = [\pi D_f / (128 (3 + D_T - D_f))] r_{max}^4 (1 - S_{wir})^{(4 / (2 - D_f))} \qquad (11)$$

where D_T is the tortuosity fractal dimension ($1 \leq D_T \leq 2$) and r_max is the maximum pore-throat radius. Equation (11) demonstrates analytically that conventional Kozeny-Carman and FZI formulations systematically overestimate permeability in tight formations by neglecting the hydraulic throttling effect of immovable capillary water films.

4. The Ascent of Artificial Intelligence: Unsupervised Learning & Facies Discovery

4.1 Unsupervised Algorithms for Multi-Well Borehole Data

The fundamental limitation of core-based rock typing is spatial discontinuity: cored intervals represent sparse, 1D samples covering less than 10% of total reservoir volume. To propagate rock types continuously across thousands of feet of uncored wells, the petroleum industry has increasingly turned to unsupervised machine learning (ML) algorithms (Syed et al., 2022; Ashraf et al., 2024a). Unsupervised learning operates without pre-existing labels, uncovering intrinsic mathematical structures, clusters, and topology within high-dimensional wireline log space (Figure 3a).

Four primary clustering families dominate petrophysical rock typing: (1) K-Means Clustering: A centroid-based partitioning algorithm that iteratively minimizes the within-cluster sum of squares (inertia). While computationally efficient, K-Means assumes spherical clusters of equal variance and is sensitive to outliers. (2) Fuzzy C-Means (FCM): A soft-clustering generalization allowing each log depth interval to belong to multiple rock types with partial membership degrees ($u_{ij} \in [0, 1]$), reflecting the transitional nature of sedimentary facies boundaries. (3) Gaussian Mixture Models (GMM): A probabilistic approach that models data as a linear superposition of multivariate Gaussian distributions, optimized via Expectation-Maximization (EM). (4) Self-Organizing Maps (SOM): An artificial neural network architecture developed by Kohonen (1990) that projects high-dimensional well log features onto a regular low-dimensional (typically 2D) topological grid while preserving neighborhood relationships (Figure 3a). SOM excels at visualizing complex non-linear electrofacies transitions.

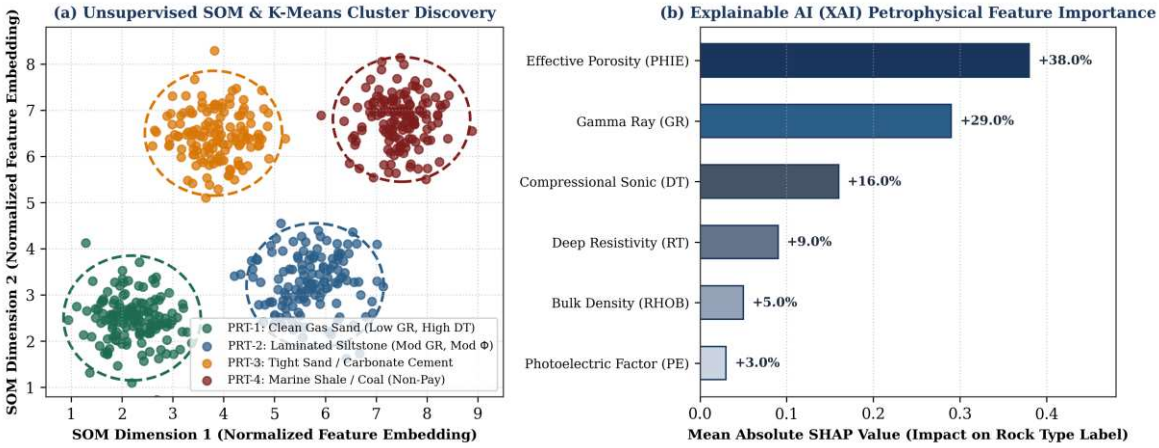


Figure 3. Machine learning well-log workflow and interpretability: (a) 2D Self-Organizing Map (SOM) latent feature space showing unsupervised electrofacies clustering into four distinct petrophysical rock types (Clean Gas Sand, Laminated Siltstone, Tight Sand/Cement, and Marine Shale/Coal); (b) Explainable AI (XAI) feature attribution using mean absolute SHAP values (Lundberg & Lee, 2017; Wang et al., 2024), quantifying the relative physical impact of wireline logs on rock-type classification.

4.2 Multi-Well Feature Selection & Dimensionality Reduction

Raw wireline log suites commonly encompass 10 to 30 distinct measurements, including natural gamma ray (GR), spectral gamma ray (K, U, Th), bulk density (RHOB), neutron porosity (NPHI), compressional sonic transit time (DT), shear sonic (DTS), photoelectric absorption factor (PEF), and multi-depth array induction resistivities (RT). Feeding raw logs directly into clustering algorithms introduces severe multicollinearity and the "curse of dimensionality" (Bishop, 2006). Effective AI workflows employ dimensionality reduction and feature engineering: Principal Component Analysis (PCA) transforms correlated logs into orthogonal principal components, while t-Distributed Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP) preserve non-linear manifold structures for visual cluster inspection. Prior to clustering, rigorous petrophysical log conditioning — including borehole environmental corrections, depth matching against core, mud-filtrate invasion filtering, and robust z-score or min-max normalization — is mandatory to prevent borehole washouts from generating artificial rock types (Cuddy, 2021).

4.3 Integrated Multi-Well Field Workflows

Recent field studies demonstrate the power of hybrid unsupervised workflows that combine multiple clustering algorithms with noise filtering and interpretability checks. Ashraf et al. (2024a, 2024b) implemented an integrated four-stage workflow across heterogeneous clastic and coal-bearing tight gas reservoirs: (i) Self-Organizing Maps (SOM) were first utilized to compress multi-well log suites onto a 2D topological manifold to reveal natural data groupings; (ii) K-Means clustering was executed on the SOM latent space to discretize definitive electrofacies boundaries; (iii) Density-Based Spatial Clustering of Applications with Noise (DBSCAN; Ester et al., 1996) was applied to detect and isolate logging noise, bad-hole intervals, and anomalous outliers; and (iv) the optimal cluster count was validated quantitatively using the Davies-Bouldin index, Calinski-Harabasz score, and Silhouette analysis. This integrated protocol successfully isolated payable, gas-saturated channel sands from

non-pay tight siltstones and correctly distinguished organic-rich coal seams from radioactive shale markers where traditional gamma ray cutoffs failed completely (Ashraf et al., 2024b).

5. Supervised Learning, Physics-Informed Machine Learning, and Property Prediction

5.1 Supervised Classification and Permeability Modeling

Where cored wells provide definitive ground-truth rock-type labels and laboratory permeability measurements, supervised machine learning offers powerful non-linear mapping capabilities to propagate properties into uncored intervals and offset exploration wells. Modern tree-based ensemble algorithms — particularly Random Forest (RF; Breiman, 2001), Extreme Gradient Boosting (XGBoost; Chen & Guestrin, 2016), and LightGBM — routinely outperform traditional empirical multi-linear regressions and conventional multi-layer perceptrons (MLP). Tree ensembles handle non-linear feature interactions, are invariant to monotonic feature scaling, and provide inherent resilience against outliers (Zhang et al., 2021; Kumar & Ghosh, 2025). Support Vector Machines (SVM; Vapnik, 1995) utilizing Radial Basis Function (RBF) kernels excel in high-dimensional classification when training sample sizes are constrained, while Graph Neural Networks (GNNs) have recently been adapted to capture spatial stratigraphical connectivity across multi-well correlations.

5.2 Physics-Informed and Physics-Guided Machine Learning (PIML / PGML)

Despite their statistical accuracy, purely data-driven supervised algorithms remain vulnerable to producing unphysical predictions — such as negative permeabilities, non-monotonic capillary pressure slopes, or water saturations exceeding unity — especially when evaluating out-of-distribution (OOD) depth intervals (Krishnapriyan et al., 2021; Kong et al., 2024). To guarantee physical consistency, the petrophysics community has pioneered Physics-Informed Machine Learning (PIML) and Physics-Guided Neural Networks (PGNN) (Raissi et al., 2019; Tartakovsky et al., 2020; Karniadakis et al., 2021).

Ji et al. (2022) formulated a compelling example of hybrid physics-guided rock typing for tight gas sandstones. Recognizing that classical FZI breaks down under non-Darcy gas transport, they derived three explicit physical flow indices: (i) the Darcy Index (F_i), governing standard viscous flow; (ii) the Gas Non-Darcy Index (K_c), quantifying Klinkenberg slip flow and Knudsen diffusion; and (iii) the Liquid Non-Darcy Index (S_r), accounting for bound-water film drag. These three analytical indices were then fed into a Weighted Fuzzy k-Nearest Neighbor (k-NN) classification algorithm, wherein the AI component optimized the relative weighting of each physical mechanism based on local pore-throat sorting (Ji et al., 2022). More broadly, PIML architectures enforce physical laws by supplementing the empirical data loss (L_data) with penalty loss terms (L_phys) penalizing violations of the Kozeny-Carman equation, mass conservation, or monotonicity constraints:

$$L_{total} = L_{data}(y, \hat{y}) + \lambda_{phys} L_{phys}(\nabla\hat{y}, \Phi, S_w, k) \qquad (12)$$

Embedding domain equations directly into the network's backpropagation loss function ensures that predictions remain physically bounded even in sparse-data extrapolation regimes (Raissi et al., 2019; Amini et al., 2022).

5.3 Explainable AI (XAI) & Petrophysical Interpretability

A primary obstacle impeding the operational adoption of AI rock typing by reservoir engineers and regulatory bodies is the "black box" stigma of deep neural networks (Ribeiro et al., 2016; Cuddy, 2021). Explainable AI (XAI) frameworks — specifically Shapley Additive Explanations (SHAP; Lundberg & Lee, 2017) and Local Interpretable Model-agnostic Explanations (LIME; Ribeiro et al., 2016) — have emerged as mandatory validation tools to audit algorithm decisions. Grounded in cooperative game theory, SHAP computes the marginal contribution of each wireline log to the final rock-type classification for every individual depth point (Figure 3b).

As demonstrated by Ashraf et al. (2024a) and Wang et al. (2024), SHAP feature attribution provides an auditable sanity check: in a valid petrophysical model, effective porosity (PHIE), gamma ray (GR), and sonic transit time (DT) should emerge as dominant drivers for reservoir sandstones, whereas deep resistivity (RT) should dominate hydrocarbon-saturation discrimination (Figure 3b). If a SHAP audit reveals that an AI model is relying heavily on an irrelevant artifact (such as borehole caliper or tool tension) to classify rock types, the model can be diagnosed as overfitted and rejected prior to field deployment.

6. Digital Rock Physics (DRP) and Microscale Image-Based Rock Typing

6.1 High-Resolution Imaging Modalities and Pore-Scale Physics

Digital Rock Physics (DRP) has revolutionized modern petrophysics by digitizing three-dimensional pore-scale microstructures, enabling numerical flow simulations that serve as direct physical ground truth for macroscopic rock typing (Figure 4). DRP utilizes a hierarchy of non-destructive and destructive imaging technologies: (i) Optical Thin-Section Petrography (resolution ~1 μm), providing 2D mineral identification, grain size sorting, and Dunham texture classification; (ii) Scanning Electron Microscopy (SEM) and Focused Ion Beam SEM (FIB-SEM; resolution 5–50 nm), capable of resolving nanoscale microporosity, clay mineral morphologies, and organic kerogen pores in unconventional shales; (iii) Laboratory micro-computed tomography (μ-CT; resolution 0.5–10 μm), capturing non-destructive 3D voxel volumes of core plugs (typically 5 mm to 38 mm in diameter); and (iv) Synchrotron nano-CT (resolution <50 nm), providing ultra-high-resolution reconstruction of dense micritic carbonate textures and tight sandstone pore throats (Clerke et al., 2008; Wang, 2024).

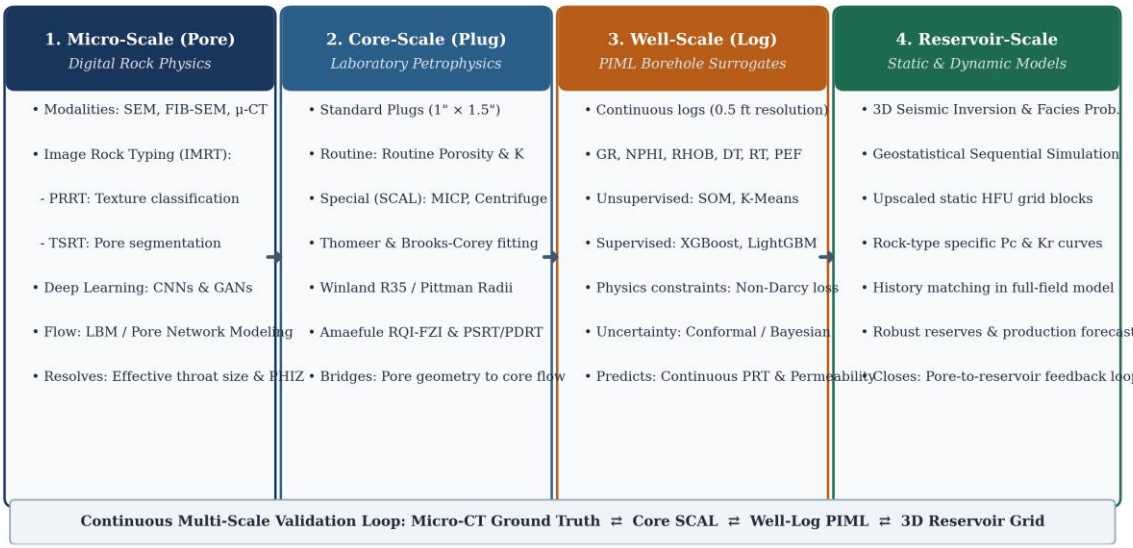


Figure 4. The unified multi-scale digital rock typing integration framework: bridging micro-scale Digital Rock Physics (SEM/μ-CT, IMRT segmentation, LBM flow simulation), meso-scale core analysis (RCAL, SCAL, MICP, Winland/FZI), macro-scale well logs (PIML continuous PRT and permeability prediction), and mega-scale 3D seismic/reservoir models through a continuous, closed-loop validation cycle.

6.2 Pattern Recognition (PRRT) vs. Texture Segmentation (TSRT) Rock Typing

In a comprehensive review, Wang (2024) synthesized Image-based Microscale Rock Typing (IMRT) into two distinct methodological paradigms (Figure 4, Tier 1): (1) Pattern Recognition-related Rock Typing (PRRT): An image-level classification task in which an entire 2D image or 3D sub-volume is assigned to a discrete category (such as Dunham lithofacies, grainstone vs. mudstone, or grain sorting class). PRRT automates traditional petrographic descriptions and standardizes core logging. (2) Texture Segmentation-related Rock Typing (TSRT): A semantic segmentation task wherein a single rock image or CT volume is divided into multiple distinct pore-throat regions, each representing a different hydraulic sub-system within the heterogeneous sample (e.g., separating intergranular macropores, microporous micrite, dissolution vugs, and micro-fractures). TSRT is essential for resolving intra-sample heterogeneity and provides direct visual and quantitative calibration for Rabiller's (2017) effective flow porosity (PHIZ), directly quantifying which portion of total pore volume connects to percolating flow paths.

6.3 Deep Learning for Pore Reconstruction and In-Silico Flow Simulation

Deep learning has become the primary computational engine driving Digital Rock Physics. Convolutional neural networks (CNNs), specifically U-Net and DeepLabV3 architectures, achieve near-human accuracy in segmenting solid grain matrix, clay minerals, and pore phases from noisy CT volumes (Wang, 2024). Generative Adversarial Networks (GANs), particularly CycleGAN and Super-Resolution GANs (SRGAN), are widely utilized to generate super-resolved 3D pore volumes from lower-resolution, large-field-of-view scans, overcoming the fundamental trade-off between imaging resolution and sample representative elementary volume (REV). Once a segmented 3D pore geometry is reconstructed, computational fluid dynamics algorithms — primarily the Lattice Boltzmann Method (LBM) and Direct Pore Network Modeling (PNM) — simulate single-phase and

two-phase fluid flow directly through the digital pore space. In-silico simulation computes absolute permeability tensors, formation resistivity factor (Archie m exponent), mercury injection capillary pressure (MICP) curves, and oil–water relative permeability curves in a fraction of the time required for physical laboratory SCAL, providing an inexhaustible library of synthetic training data for well-log machine learning surrogates (Wang, 2024).

7. Comparative Synthesis, Industry Pitfalls, and Genuine Operational Risks

7.1 Comprehensive Multi-Dimensional Comparison of Methodologies

Table 1 synthesizes the five major petrophysical rock typing paradigms across six operational criteria: input data requirements, effective spatial scale, physical fidelity, level of automation, computational expense, and primary failure modes.

Table 1. Comprehensive multi-dimensional comparison of petrophysical rock typing methodologies from conventional empirical foundations to digital rock physics and AI.

Methodological Paradigm	Key Frameworks & Citations	Input Data Requirements	Spatial Scale	Physical Fidelity & Automation	Primary Failure Modes & Limitations
1. Conventional Empirical & Fabric	Winland R_{35} (Kolodzie, 1980); Pittman (1992); Lucia (1983, 1995)	Routine core porosity & permeability, thin sections, basic logs	Core plug (1" × 1.5") to 1D well log	Moderate physical fidelity; manual heuristics; low automation	Subjective boundary cutoffs; fails in multi-modal carbonates; conflates storage with flow capacity
2. Classical Hydraulic Flow Units	RQI-FZI (Amaefule et al., 1993); Modified Kozeny-Carman	Core plug porosity & permeability, basic wireline log suite	Core plug to 1D well log zone	Analytical Kozeny-Carman basis; manual slope fitting; moderate automation	Assumes laminar Darcy flow; overestimates permeability in tight/shaly formations; requires extensive core calibration
3. Advanced Conventional Refinements	PSRT/PDRT (Mirzaei-Paiaman, 2018); PHIZ (Rabiller, 2017); Fractal S_wir (Chen & Yao, 2017)	Core RCAL, high-pressure MICP, specialized clay/mineralogy data	Core plug to flow unit interval	High physical fidelity; explicitly decouples static storage from dynamic flow; semi-automated	Requires high-quality MICP data; increased analytical complexity; still constrained by core availability

Methodological Paradigm	Key Frameworks & Citations	Input Data Requirements	Spatial Scale	Physical Fidelity & Automation	Primary Failure Modes & Limitations
4. Unsupervised AI (Electrofacies)	SOM (Kohonen, 1990); K-Means; FCM; DBSCAN (Ashraf et al., 2024a, 2024b)	Multi-well raw and conditioned wireline log suites	1D borehole (0.5 ft vertical resolution) across multi-well fields	Data-driven; discovers latent clusters without pre-labels; high automation	Vulnerable to logging noise/washouts; clusters lack intrinsic physical meaning without core ground truth; black-box risk
5. Supervised & Physics-Informed ML	XGBoost, RF, SVM, PIML/PGML (Ji et al., 2022; Amini et al., 2022); SHAP (Lundberg & Lee, 2017)	Paired core-log training databases, physical loss constraints	Borehole log scale with field-wide cross-well propagation	High physical fidelity (under PIML); highly automated; auditable via XAI	Requires labeled core training data; vulnerable to out-of-distribution failure; hyperparameter sensitivity
6. Digital Rock Physics & IMRT	μ-CT, SEM, FIB-SEM; PRRT & TSRT (Wang, 2024); LBM flow simulation	High-resolution 2D/3D rock images, digital micro-plugs	Pore-to-plug scale (nanometers to centimeters)	Direct visual ground truth; full Navier-Stokes/LBM physics; highly automated	High computational & scanning expense; small field of view risks missing representative elementary volume (REV)

7.2 Genuine Operational Risks and Technical Pitfalls in AI Petrophysics

While early discussions of AI in geosciences occasionally indulged in speculative melodrama regarding autonomous runaway algorithms (Cuddy, 2021), real-world petroleum engineering operations face severe, well-documented technical hazards when deploying machine learning for rock typing and property prediction. Five technical pitfalls demand rigorous mitigation: (1) Spatial Autocorrelation and Data Leakage: Standard machine learning validation protocols rely on random k-fold cross-validation, randomly shuffling depth points between training and test sets. In borehole petrophysics, neighboring depth samples separated by 0.5 ft share nearly identical lithology, porosity, and fluid saturations due to stratigraphic continuity and tool sensor response functions. Random splitting allows models to "memorize" adjacent depth intervals, yielding unrealistically optimistic R^2 values (>0.95) that collapse disastrously ($R^2 < 0.40$) when tested on an independent offset well. Petrophysical AI models must strictly enforce spatial Leave-One-Well-Out (LOWO) or Leave-One-Field-Out (LOFO) cross-validation to assess true generalization capability. (2) Out-of-Distribution (OOD) Extrapolation Failures: Machine learning models are non-parametric

interpolators within their training convex hull. When an exploration well encounters an unseen diagenetic facies (e.g., hydrothermal dolomitization, heavy pyrobitumen plugging, or fracture swarms) not present in the cored training database, purely data-driven models output confident but erroneous rock-type labels and permeability estimates (Bishop, 2006). PIML loss regularization and epistemic uncertainty quantification (Bayesian neural networks or deep ensembles; Lakshminarayanan et al., 2017) are mandatory to detect OOD inputs. (3) Physical Inconsistency: Unconstrained neural networks frequently predict unphysical combinations — such as water saturation $S_w < 0$ or $S_w > 1$, negative permeabilities, or non-monotonic capillary pressure profiles where capillary pressure drops as non-wetting saturation increases. (4) Core-to-Log Depth-Shift Discrepancies and Resolution Mismatch: Core depths systematically deviate from drillers' and wireline logger depths by 5 to 30 ft due to drill-pipe stretch, cable tension, and core expansion. Training supervised algorithms on uncorrected depth pairings introduces severe label noise that permanently degrades classifier performance. Furthermore, core plugs sample 1-inch volumes, whereas wireline logs sample 2- to 3-foot spherical volumes, requiring forward-filtering and deconvolution prior to supervised training. (5) Regulatory and Audit Compliance: Subsurface equity determinations, reserves certifications (SEC/SPE-PRMS), and equity redeterminations require auditable, transparent engineering workflows. Deploying opaque black-box neural networks without SHAP feature attribution, conformal uncertainty bounds, and physical cross-checks exposes operating companies to severe commercial and regulatory liabilities.

8. A Unified Multi-Scale Physics-Informed Digital Rock Typing Roadmap

The future of petrophysical rock typing lies not in a dichotomy between conventional methods and artificial intelligence, but in their seamless multi-scale integration (Figure 4). To realize this potential, we articulate six prioritized strategic research directions: (1) Open Multi-Scale Petrophysical Benchmark Compilations: The industry urgently requires standardized, open-access multi-well databases combining registered μ-CT volumes, whole-core photography, MICP/SCAL drainage curves, and complete wireline log suites across varied lithologies (carbonates, clastics, tight gas, shales) with pre-defined spatial train/test splits to enable reproducible algorithmic benchmarking. (2) Self-Supervised Subsurface Foundation Models: Rather than training small bespoke networks for individual fields, the industry should develop large-scale petrophysical foundation models pre-trained via self-supervised masked autoencoding on millions of feet of global well logs and seismic volumes, fine-tuned locally with sparse core calibration. (3) Conditioning-Safe PIML Architectures: Advanced loss-function formulations that seamlessly balance multi-scale Darcy and non-Darcy flow physics while preventing Hessian ill-conditioning across sharp permeability contrasts (Krishnapriyan et al., 2021; Amini et al., 2022). (4) Real-Time Logging-While-Drilling (LWD) Rock Typing: Deploying edge-AI computing on drilling rigs to classify rock types and forecast permeability in real time, enabling geosteering within sweet-spot flow units and automated intelligent completions. (5) Uncertainty-Quantified Multi-Scale Upscaling: Propagating epistemic and aleatoric uncertainty bounds rigorously from pore-scale LBM simulations through wireline log PIML surrogates directly into full-field reservoir simulation grid blocks using conformal prediction intervals. (6) Standardized Open-Source Evaluation Protocols: Establishing standardized reporting requirements — including mandatory spatial cross-validation, SHAP interpretability audits, and

physical consistency checks — across all peer-reviewed petrophysical machine learning publications.

9. Conclusions

Methodological Evolution: 1. Petrophysical rock typing (PRT) has evolved through four major historical eras: from empirical fabric and Winland R_{35} classifications (1940s–1980s), through the modified Kozeny-Carman Hydraulic Flow Unit (HFU/FZI) paradigm (1990s–2000s) and advanced static/dynamic refinements (2010s), to the modern era of Physics-Informed Artificial Intelligence and Digital Rock Physics (2020s–2026+).

Conventional Refinements: 2. Conventional FZI and Winland methods provide essential physical grounding but suffer from fundamental limitations, including the conflation of static storage with dynamic flow capacity and failure in non-Darcy tight formations. Advanced conventional formulations — specifically decoupling Petrophysical Static Rock Types (PSRT) from Dynamic Rock Types (PDRT), calibrating effective flow porosity (PHIZ), and fractal bound-water models — successfully resolve these theoretical shortcomings.

AI & Machine Learning: 3. Artificial intelligence provides transformative capabilities across scales: unsupervised clustering (SOM, K-Means, DBSCAN) automates multi-well electrofacies zoning; supervised tree ensembles (XGBoost, Random Forest) predict continuous rock classes in uncored intervals; and Physics-Informed Machine Learning (PIML) embeds governing flow equations directly into loss functions to prevent unphysical extrapolation.

Digital Rock Physics: 4. Digital Rock Physics (DRP) and Image-based Microscale Rock Typing (IMRT) bridge the pore-to-plug scale, utilizing μ-CT, deep learning segmentation (PRRT and TSRT), and Lattice Boltzmann flow modeling to provide direct visual and numerical ground truth for macroscopic rock types.

Operational Risk Management: 5. Operational deployment of AI in petrophysics demands rigorous risk mitigation: operating companies must replace speculative scenarios with technical safeguards against spatial autocorrelation data leakage (enforcing Leave-One-Well-Out validation), out-of-distribution extrapolation failure, core-to-log depth-shift errors, and black-box opacity through Explainable AI (SHAP/LIME).

Nomenclature

Table 2. Comprehensive nomenclature table of mathematical symbols, physical definitions, units, and petrophysical acronyms.

Symbol / Acronym	Physical & Technical Definition	Dimensions / Standard Units
B_v∞	Thomeer hyperbola non-wetting phase pore-volume factor at infinite pressure	Fraction or bulk volume %
DBSCAN	Density-Based Spatial Clustering of Applications with Noise	Unsupervised ML algorithm

Symbol / Acronym	Physical & Technical Definition	Dimensions / Standard Units
D_f, D_T	Pore-throat fractal dimension and tortuosity fractal dimension	Dimensionless
DRP	Digital Rock Physics	Discipline
DT, DTS	Compressional and shear wave sonic transit times	μs/ft
F_i, K_c, S_r	Darcy index, gas non-Darcy index, and liquid non-Darcy index (Ji et al., 2022)	Dimensionless
F_s	Pore cross-sectional shape factor (Kozeny-Carman)	Dimensionless (~2.0 for cylinders)
FZI, FZI*	Flow Zone Indicator and revised dynamic flow zone indicator (Mirzaei-Paiaman, 2018)	μm
FZI_true	Effective Flow Zone Indicator calibrated with PHIZ (Rabiller, 2017)	μm
G	Thomeer pore-geometrical factor (pore-throat sorting factor)	Dimensionless
GR	Gamma Ray wireline log	API units
HFU	Hydraulic Flow Unit	Classification unit
IMRT	Image-based Microscale Rock Typing (Wang, 2024)	Framework
J(S_w)	Classical dimensionless Leverett J-function (Leverett, 1941)	Dimensionless
k, k_t	Absolute matrix permeability and composite permeability	mD (L^2)
LBM	Lattice Boltzmann Method	CFD numerical method
MICP	Mercury Injection Capillary Pressure	Experimental method
P_c, P_d	Capillary pressure and capillary displacement (entry) pressure	psi, bar, or MPa ($M\ L^{-1}\ T^{-2}$)
PDRT, PSRT	Petrophysical Dynamic Rock Type and Petrophysical Static Rock Type	Classification unit
PEF	Photoelectric absorption factor	b/e (barns/electron)
PHIE, PHIZ	Effective porosity and effective flow-contributing porosity (Rabiller, 2017)	Fraction [0, 1]
PIML / PGML	Physics-Informed Machine Learning / Physics-Guided Machine Learning	Computational paradigm
PRRT, TSRT	Pattern Recognition Rock Typing and Texture Segmentation Rock Typing	IMRT categories
PRT	Petrophysical Rock Typing	Discipline
R_35, r_max	Winland pore-throat aperture radius at 35% mercury saturation and maximum throat radius	μm (L)

Symbol / Acronym	Physical & Technical Definition	Dimensions / Standard Units
RBF	Radial Basis Function kernel (Support Vector Machines)	Mathematical kernel
RHOB, NPHI	Bulk density (g/cm^3) and neutron porosity (fraction/pu) logs	g/cm^3 / fraction
RQI	Reservoir Quality Index (Amaefule et al., 1993)	μm (L)
RT	True deep formation resistivity log	ohm-m (Ω·m)
S_gv	Specific surface area per unit grain volume	μm^{-1} (L^{-1})
SHAP, LIME	Shapley Additive Explanations and Local Interpretable Model-agnostic Explanations	Explainable AI tools
SOM	Self-Organizing Map (Kohonen artificial neural network)	Unsupervised ML algorithm
S_w, S_wir	Wetting-phase (water) saturation and irreducible water saturation	Fraction [0, 1]
τ	Hydraulic tortuosity (ratio of actual fluid path to straight-line distance)	Dimensionless (≥ 1.0)
θ, σ	Fluid contact angle (degrees) and fluid–fluid interfacial tension (dyne/cm or N/m)	Degrees, $M\ T^{-2}$
Φ, Φ_z	Total/effective porosity and normalized porosity index (φ / (1 - φ))	Fraction, Dimensionless
XGBoost	Extreme Gradient Boosting tree ensemble algorithm	Supervised ML algorithm
μ-CT	Micro-Computed Tomography	Imaging modality

Declarations

Data Availability, Funding, and Competing Interests: All datasets and mathematical models reviewed in this study are derived from publicly accessible peer-reviewed literature cited in the References section. No proprietary operator files were accessed. This research received no external funding. The authors declare no competing financial or non-financial interests.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process:

During the preparation of this manuscript, the authors used AI-assisted language tools solely for language editing, stylistic refinement, and improving the text's syntactic clarity. Following these language refinements, the authors comprehensively reviewed, verified, and edited all technical content, mathematical derivations, petrophysical models, and empirical data, taking full personal responsibility for the integrity and accuracy of the publication.

References

[1] Al-Farisi, O., El-Husseiny, A., & Ghedan, S. (2018). Multi-scale rock typing in carbonate reservoirs: A comprehensive review. Journal of Petroleum Science and Engineering, 167, 856–870. https://doi.org/10.1016/j.petrol.2018.04.052

[2] Amaefule, J. O., Altunbay, M., Tiab, D., Kersey, D. G., & Keelan, D. K. (1993). Enhanced reservoir description: Using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells. SPE Annual Technical Conference and Exhibition, Houston, Texas, SPE-26436-MS. https://doi.org/10.2118/26436-MS

[3] Amini, D., Haghighat, E., & Juanes, R. (2022). Physics-informed neural network solution of thermo–hydro–mechanical processes in porous media. Journal of Engineering Mechanics, 148(11), 04022070. https://doi.org/10.1061/(ASCE)EM.1943-7889.0002156

[4] Amyx, J. W., Bass, D. M., & Whiting, R. L. (1960). Petroleum reservoir engineering: Physical properties. McGraw-Hill.

[5] Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. Transactions of the AIME, 146(1), 54–62. https://doi.org/10.2118/942054-G

[6] Arifianto, I., Surjono, S. S., Erlangga, G., Abrar, B., & Yogapurana, E. (2018). Application of flow zone indicator and Leverett J-function to characterise carbonate reservoir and calculate precise water saturation in the Kujung formation, North East Java Basin. Journal of Geophysics and Engineering, 15(4), 1753–1766. https://doi.org/10.1088/1742-2140/aab636

[7] Ashraf, U., Anees, A., Zhang, H., Ali, M., Thanh, H. V., & Yuan, Y. (2024a). Identifying payable cluster distributions for improved reservoir characterization: A robust unsupervised ML strategy for rock typing of depositional facies in heterogeneous rocks. Geomechanics and Geophysics for Geo-Energy and Geo-Resources, 10(1), 131. https://doi.org/10.1007/s40948-024-00845-6

[8] Ashraf, U., Shi, W., Zhang, H., Anees, A., Jiang, R., Ali, M., & Zhang, X. (2024b). Reservoir rock typing assessment in a coal-tight sand-based heterogeneous geological formation through advanced AI methods. Scientific Reports, 14(1), 5659. https://doi.org/10.1038/s41598-024-55823-3

[9] Bachu, S. (2008). CO_2 storage in geological media: Role, means, status and barriers to deployment. Progress in Energy and Combustion Science, 34(2), 254–273. https://doi.org/10.1016/j.pecs.2007.10.001

[10] Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.

[11] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

[12] Burdine, N. T. (1953). Relative permeability calculations from pore size distribution data. Journal of Petroleum Technology, 5(3), 71–78. https://doi.org/10.2118/225-G

[13] Byrnes, A. P., Franseen, E. K., & Steinhauff, D. M. (1999). Integrating plug to well-scale petrophysics with detailed sedimentology to quantify fracture, vug, and matrix properties in carbonate reservoirs: An example from the Arbuckle Group, Kansas (Open-File Report 99-47). Kansas Geological Survey. https://www.kgs.ku.edu/PRS/Poster/1999/99-47/index.html

[14] Camacho-Velázquez, R., Vásquez-Cruz, M., Castrejón-Aivar, R., & Arana-Ortiz, V. (2005). Pressure-transient and decline-curve behaviors in naturally fractured vuggy carbonate reservoirs. SPE Reservoir Evaluation & Engineering, 8(2), 95–112. https://doi.org/10.2118/84979-PA

[15] Carman, P. C. (1937). Fluid flow through granular beds. Transactions of the Institution of Chemical Engineers, 15, 150–166.

[16] Carman, P. C. (1956). Flow of gases through porous media. Butterworths Scientific Publications.

[17] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

[18] Chen, X., & Yao, G. (2017). An improved model for permeability estimation in low-permeability porous media based on fractal geometry and modified Hagen-Poiseuille flow. Fuel, 210, 748–757. https://doi.org/10.1016/j.fuel.2017.09.022

[19] Choquette, P. W., & Pray, L. C. (1970). Geologic nomenclature and classification of porosity in sedimentary carbonates. AAPG Bulletin, 54(2), 207–250. https://doi.org/10.1306/5D25C98B-16C1-11D7-8645000102C1865D

[20] Clerke, E. A., Mueller, H. W., III, Φllips, E. C., Eyvazzadeh, R. Y., Jones, D. H., Ramamoorthy, R., & Srivastava, A. (2008). Application of Thomeer hyperbolas to decode the pore systems, facies and reservoir properties of the Upper Jurassic Arab D limestone, Ghawar field, Saudi Arabia: A "Rosetta Stone" approach. GeoArabia, 13(4), 113–160.

[21] Cuddy, S. (2021). The benefits and dangers of using artificial intelligence in petrophysics. Artificial Intelligence in Geosciences, 2, 1–10. https://doi.org/10.1016/j.aiig.2021.04.001

[22] Dunham, R. J. (1962). Classification of carbonate rocks according to depositional textures. In W. E. Ham (Ed.), Classification of Carbonate Rocks (Memoir 1, pp. 108–121). American Association of Petroleum Geologists.

[23] El-Khatib, N. (2001). Development of a modified capillary pressure J-function. Proceedings of the International Symposium of the Society of Core Analysts, Edinburgh, UK, SCA-2001-35.

[24] Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), 226–231.

[25] Folk, R. L. (1959). Practical petrographic classification of limestones. AAPG Bulletin, 43(1), 1–38. https://doi.org/10.1306/0BDA5C36-16BD-11D7-8645000102C1865D

[26] Gholami, R., Raza, A., & Iglauer, S. (2021). Leakage risk assessment of a CO_2 storage site: A review. Earth-Science Reviews, 223, 103849. https://doi.org/10.1016/j.earscirev.2021.103849

[27] Haghighat, E., Raissi, M., Moure, A., Gomez, H., & Juanes, R. (2021). A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. Computer Methods in Applied Mechanics and Engineering, 379, 113741. https://doi.org/10.1016/j.cma.2021.113741

[28] Ji, L., Lin, M., Jiang, W., Cao, G., Xu, Z., & Hao, F. (2022). An improved rock typing method for tight sandstone based on new rock typing indexes and the weighted fuzzy kNN algorithm. Journal of Petroleum Science and Engineering, 210, 109956. https://doi.org/10.1016/j.petrol.2021.109956

[29] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3(6), 422–440. https://doi.org/10.1038/s42254-021-00314-5

[30] Kohonen, T. (1990). The self-organizing map. Proceedings of the IEEE, 78(9), 1464–1480. https://doi.org/10.1109/5.58325

[31] Kolodzie, S. (1980). Analysis of pore throat size and use of the Waxman-Smits equation to determine OOIP in the Spindle Field, Colorado. SPE Annual Technical Conference and Exhibition, Dallas, Texas, SPE-9382-MS. https://doi.org/10.2118/9382-MS

[32] Kong, Q., Walter, W. R., Wang, R., & Schmandt, B. (2024). Evaluating physics-informed neural network performance for seismic discrimination between earthquakes and explosions. Seismological Research Letters, 96(1), 147–159. https://doi.org/10.1785/0220240097

[33] Kozeny, J. (1927). Über kapillare Leitung des Wassers im Boden. Sitzungsberichte der Akademie der Wissenschaften in Wien, 136, 271–306.

[34] Krishnapriyan, A., Gholami, A., Zhe, S., Kirby, R., & Mahoney, M. W. (2021). Characterizing possible failure modes in physics-informed neural networks. Advances in Neural Information Processing Systems (NeurIPS), 34, 26548–26560.

[35] Kumar, V., & Ghosh, S. (2025). Estimation of vertical permeability of Hugin sandstone from petrophysical well logs using ensemble methods: An enhanced machine learning approach. Arabian Journal for Science and Engineering, 50, 4943–4961. https://doi.org/10.1007/s13369-024-09446-z

[36] Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. Advances in Neural Information Processing Systems (NeurIPS), 30, 6402–6413.

[37] Leverett, M. C. (1941). Capillary behavior in porous solids. Transactions of the AIME, 142(1), 152–169. https://doi.org/10.2118/941152-G

[38] Lucia, F. J. (1983). Petrophysical parameters estimated from visual descriptions of carbonate rocks: A field classification of carbonate pore space. Journal of Petroleum Technology, 35(3), 629–637. https://doi.org/10.2118/10073-PA

[39] Lucia, F. J. (1995). Rock-fabric/petrophysical classification of carbonate pore space for reservoir characterization. AAPG Bulletin, 79(9), 1275–1300.

[40] Lucia, F. J. (1999). Carbonate reservoir characterization. Springer-Verlag. https://doi.org/10.1007/978-3-662-03977-9

[41] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems (NeurIPS), 30, 4765–4774.

[42] Martin, A. J., Solomon, S. T., & Hartmann, D. J. (1999). Characterization of petrophysical flow units in carbonate reservoirs. AAPG Bulletin, 83(7), 1161–1173.

[43] Mirzaei-Paiaman, A., Ostadhassan, M., Rezaee, R., Saboorian-Jooybari, H., & Chen, Z. (2018). A new approach in petrophysical rock typing. Journal of Petroleum Science and Engineering, 166, 445–464. https://doi.org/10.1016/j.petrol.2018.03.075

[44] Mirzaei-Paiaman, A., Saboorian-Jooybari, H., & Chen, Z. (2020). A further verification of FZI* and PSRTI: Newly developed petrophysical rock typing indices. Journal of Petroleum Science and Engineering, 184, 106516. https://doi.org/10.1016/j.petrol.2019.106516

[45] Pittman, E. D. (1992). Relationship of porosity and permeability to various parameters derived from mercury injection-capillary pressure curves for sandstone. AAPG Bulletin, 76(2), 191–198. https://doi.org/10.1306/BDFF87A4-1718-11D7-8645000102C1865D

[46] Purcell, W. R. (1949). Capillary pressures — Their measurement using mercury and the calculation of permeability therefrom. Journal of Petroleum Technology, 1(2), 39–48. https://doi.org/10.2118/949039-G

[47] Rabiller, P. (2017). Combining porosimetry and Purcell permeability modeling to calibrate FZI and define a dynamic permeability cut-off. Proceedings of the International Symposium of the Society of Core Analysts, Vienna, Austria, SCA2017-081.

[48] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

[49] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778

[50] Skalinski, M., & Kenter, J. A. (2015). Carbonate petrophysical rock typing: Integrating geological attributes and petrophysical properties while linking with dynamic behaviour. Geological Society, London, Special Publications, 406(1), 229–259. https://doi.org/10.1144/SP406.6

[51] Syed, F. I., AlShamsi, A., Dahaghi, A. K., & Neghabhan, S. (2022). Application of ML & AI to model petrophysical and geomechanical properties of shale reservoirs — A systematic literature review. Petroleum, 8(2), 158–166. https://doi.org/10.1016/j.petlm.2021.07.001

[52] Tartakovsky, A. M., Marrero, C. O., Perdikaris, P., Tartakovsky, G. D., & Barajas-Solano, D. (2020). Physics-informed deep neural networks for learning parameters and constitutive relationships in subsurface flow problems. Water Resources Research, 56(5), e2019WR026731. https://doi.org/10.1029/2019WR026731

[53] Thomeer, J. H. M. (1960). Introduction of a pore geometrical factor defined by the capillary pressure curve. Journal of Petroleum Technology, 12(3), 73–77. https://doi.org/10.2118/1324-G

[54] Thomeer, J. H. M. (1983). Air permeability as a function of three pore-network parameters. Journal of Petroleum Technology, 35(4), 809–814. https://doi.org/10.2118/10969-PA

[55] Vapnik, V. N. (1995). The nature of statistical learning theory. Springer. https://doi.org/10.1007/978-1-4757-2440-0

[56] Wang, N., Zhang, H., & Chen, Y. (2024). On the use of explainable AI for susceptibility modeling: Examining the spatial pattern of SHAP values. Geoscience Frontiers, 15(4), 101800. https://doi.org/10.1016/j.gsf.2024.101800

[57] Wang, Y. (2024). Image-based microscale rock typing and its application. Journal of Petroleum Exploration and Production Technology, 14(6), 1431–1452. https://doi.org/10.1007/s13202-024-01804-9

[58] Warren, J. E., & Root, P. J. (1963). The behavior of naturally fractured reservoirs. Society of Petroleum Engineers Journal, 3(3), 245–255. https://doi.org/10.2118/426-PA

[59] Wu, X., Liang, L., Shi, Y., & Fomel, S. (2019). FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. Geophysics, 84(3), IM35–IM45. https://doi.org/10.1190/geo2018-0646.1

[60] Zhang, S., Zou, C., Peng, C., Yan, L., & Wu, X. (2022). Pore structure and its effect on acoustic velocity and permeability of reef-shoal carbonates in the Tarim Basin, Northwest China. Journal of Geophysics and Engineering, 19(6), 1340–1354. https://doi.org/10.1093/jge/gxac087

[61] Zhang, W., Wu, C., Zhong, H., Li, Y., & Wang, L. (2021). Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. Geoscience Frontiers, 12(1), 469–477. https://doi.org/10.1016/j.gsf.2020.03.007