IJAME

*Int. J. of Applied Mechanics and Engineering, 2025, vol.30, No.4, pp.137-152*
DOI: 10.59441/ijame/209468

# PREDICTIVE MODELING OF PERMEABILITY LOSS AT HIGH-BARIUM FORMATION USING SYMBOLIC REGRESSION

S. Redha[1*], N. Zeraibi[1], M. Gareche[1], M. Nait Amar[2] and Ch. Benamara[2]
[1]Department of Mining and Petroleum Deposits, Physics Engineering of Hydrocarbons, M'Hamed Bougara University, Boumerdes, ALGERIA
[2]Département Etudes Thermodynamiques, Division Laboratoires, Sonatrach, Boumerdes, ALGERIA
E-mail: r.saifi@univ-boumerdes.dz

Mineral scaling is a common issue in oil recovery, causing severe formation damage and reduced permeability. This primarily results from chemical incompatibility between the injected and formation brines, which induces the precipitation of salts such as calcium carbonate, calcium sulfate, and barium sulfate. Accurate scale prediction is essential for managing operational risks and minimizing associated costs. in, this study, a symbolic regression technique based on evolutionary algorithms was employed to explicit mathematical model linking permeability damage to key flow parameters. A scaling risk map was also produced to distinguish between scaling and non-scaling regions.

A dataset of *431* literature-sourced permeability damage measurements was analyzed using variables such as temperature, pressure differential, Volume of injected water, initial permeability, fluid injection rate, and concentrations of major ionic constituents. For model validation, the dataset was partitioned into training (*80%*) and testing (*20%*) subsets. The model achieved strong predictive accuracy with an $R^2$of *0.99*, showing excellent agreement with experimental observations. Compared to classical thermodynamics-based models, which often neglect kinetic factors, the proposed approach offers improved prediction and control strategies for scale formation during water flooding. This makes it especially valuable for operational planning and real-time decision-making.

In summary, the model developed in this study presents significant practical benefits for researchers and engineers in both academic and industrial settings, enhancing the understanding and mitigation of mineral scaling under mixed salt conditions..

**Key words:** scaling, permeability loss, symbolic regression, sulfate barium, predictive modeling.

## 1. Introduction

Flow assurance is an essential process in oil and gas that ensures the production and transportation of hydrocarbons from the reservoir to processing facilities occurs in an efficient and reliable way. Scale formation represents one of the most prevalent challenges in this area, which occurs when inorganic salts such as calcium carbonate, calcium sulfate, barium sulfate, etc., precipitate inside production equipment. Scaling can form at any stage of the production process and is dependent on parameters like fluid composition, temperature, pressure and flow rate. scale prediction is a typical method for scaling analysis, especially given the widespread and disruptive nature of mineral scale formation in oil and gas production systems. When left unmanaged, scaling can clog and obstruct the natural flow of fluids through wellbores, production tubing, valves, casings, perforations, downhole equipment, and even the reservoir itself [1, 2]. These obstructions not only reduce hydrocarbon recovery rates but also increase maintenance frequency, operational downtime, and overall economic risk. Because scale formation depends on a complex interplay of fluid composition, temperature, pressure, and flow rate, traditional empirical methods often fail to capture the full range of dynamic interactions that drive precipitation. In this context, accurate scale prediction becomes indispensable. It allows engineers to anticipate problem areas, design more resilient systems, and optimize water injection strategies more effectively [3, 4]. Advanced predictive modelling approaches, particularly those employing machine learning

* To whom correspondence should be addressed

techniques such as artificial neural networks, are increasingly recognized as essential step of scale control process for managing scale proactively [5-7]. By enabling early detection and informed decision-making, these models significantly enhance the reliability, efficiency, and cost-effectiveness of oilfield operations.

The formation of inorganic scale is a result of complex geochemical interactions between water and rock, shaped by hydrological conditions and geological history [8]. The American Society for Testing and Materials has identified over $100$ types of deposits in water systems [9]. In oilfield environments, the most common scales are carbonates and sulfates of divalent alkaline-earth metals, such as calcium, strontium, and barium. Additionally, iron-based deposits, including hydroxides, carbonates, and sulfides, can form due to corrosion within the production system, triggered by factors such as sweet or sour environments, oxygen exposure, microbial activity, or the backflow of inadequately inhibited acid stimulation treatments [10, 11] (Tab.1).

Table.1.Representative sample of commonly encountered deposit types.

| Type of Scale | Formula | Acid Soluble |
|---|---|---|
| **Calcium Carbonate** | Calcite $(CaCO_3)$ | Yes |
| | Aragonite $(CaCO_3)$ | Yes |
| | Vaterite $(CaCO_3)$ | Yes |
| **Calcium Sulfate** | Anhydrite $(CaSO_4)$ | No |
| | Hemihydrate $\left(CaSO_4.\frac{1}{2}H_2O\right)$ | No |
| | Gypsum $(CaSO_4.2H_2O)$ | Slightly |
| **Strontium Sulfate** | Celestite $(SrSO_4)$ | No |
| **Barium Sulfate** | Barite $(BaSO_4)$ | No |
| **Iron Oxo Species** | $FeO_4$, $FeOOH$, $Fe(OH)_2$, etc. | Yes |
| **Iron Sulfides** | $FeS_y$ | Highly dependent on $Fe:S$ ratio |
| **Iron Carbonates** | e.g., Magnesite $(MgCO_3)$ | Yes |
| **Other Hydroxides** | e.g., Brucite $(Mg(OH)_2)$ | Yes |
| **Halite** | $NaCl$ | Soluble in fresh water |
| **Lead Sulfide** | Galena $(PbS)$ | Yes |
| **Silica and Metal** | Amorphous $(SiO_2)$ | No |
| **Silicates** | Various metal silicates | - |

In oil and gas operations, sulfate scales, along with carbonate scales, represent one of two major categories of inorganic scales. Common sulfate minerals: anhydrite $(CaSO_4)$, barite $(BaSO_4)$, celestite $(SrSO_4)$, gypsum $(CaSO_4.2H_2O)$. Of these, barium sulfate (barite) is less common than calcium sulfate and calcium carbonate scales, but is of major concern due to its widespread occurrence worldwide, especially in systems treating subsurface waters. As water production volumes increase with nonconventional petroleum extraction, and water flooding(as a widely used secondary recovery method) becomes a more common approach, the problem has become increasingly important for the petroleum industry [12, 13]. Barium sulfate scaling is more difficult to tackle than other scales. It is less soluble on the whole in all the solvents under most conditions, therefore it makes it one of the most insoluble scale-forming species, in contrast to calcium carbonate or calcium sulfate scales can hardly be removed chemically, even using the most advanced modern

technologies. Consequently, barium sulfate deposits usually need to be removed by mechanical means, or the impacted equipment needs to be fully changed, resulting in higher operating costs and downtime.

During waterflooding, there is often a significant risk of formation damage. A major contributor to this damage is the precipitation of inorganic mineral scales, which occurs when chemically incompatible waters (such as sulfate-rich injection brines and formation waters containing divalent cations like barium, calcium, and strontium) mix under varying temperature and pressure conditions. The resulting scales, primarily barium sulfate $(BaSO_4)$ and strontium sulfate $(SrSO_4)$, are highly insoluble and tend to deposit along the flow path, within the pore spaces of reservoir rocks, and throughout injection and production infrastructure [14].
Several factors exacerbate the risk of permeability damage from scale deposition:

- Variations in *pH*, temperature, and pressure during water injection and subsequent fluid flow.
- Incompatibility between injected and formation waters, especially in heterogeneous or layered reservoirs.
- Recirculation and mixing of fluids across pressure gradients, which promote the formation of solid solutions that restrict flow.

Permeability impairment is especially pronounced near the injection wellbore and in regions where pressure and temperature gradients facilitate the mixing of incompatible waters. In stratified formations, early water breakthrough and uneven flow distribution further intensify local supersaturation conditions, accelerating scale precipitation. The buildup of these minerals within the pore network reduces injectivity and hampers oil displacement efficiency, ultimately decreasing sweep performance and overall recovery. This issue is particularly severe in high-temperature, high-salinity environments or during low-salinity water injection (LSWI), where ionic interactions and thermal instability further enhance precipitation risks. [15]

The chemical equilibrium governing barium sulfate dissolution and precipitation is described by the following reaction:

$$BaSO_4(s) \leftrightarrow Ba^{2+}(aq) + SO_4^{2-}(aq). \tag{2.1}$$

The corresponding solubility product constant $(K_{sp})$ is given by:

$$K_{sp} = \{Ba^{2+}\}\{SO_4^{2-}\}. \tag{2.2}$$

Where $\{\}$ denotes the activity of the species, Which denotes the effective ionic concentration in the solution which represents the effective ionic concentration in solution. Activities can be calculated using activity coefficients, which are derived from experimental data or estimated using models such as the Debye-Hückel equation for dilute solutions or the Pitzer equations for more concentrated solutions. At $25°C$, barium sulfate has a solubility of $2.3\,mg/L$ in water and a solubility product $(K_{sp})$ of $10^{-9.99}$ [16, 17]. In comparison, gypsum $(CaSO_4.2H_2O)$ exhibits a much higher solubility of $2080\,mg/L$ at the same temperature [19], highlighting the extreme insolubility of barium sulfate. A key complicating factor in scale formation is the phenomenon of supersaturation, where the concentrations of cations and anions in a solution exceed the equilibrium solubility predicted by thermodynamics. Supersaturated solutions can remain stable for extended periods without forming scale, unless triggered by factors such as the presence of seed crystals or suitable surfaces for nucleation. The saturation ratio $(SR)$ is a key parameter employed to quantify the degree of supersaturation:

$$SR = \frac{\{Ba^{2+}\}\{SO_4^{2-}\}}{K_{sp}}. \tag{2.3}$$

Precipitation occurs only when the $SR$ exceeds $1.0$, indicating supersaturation. The equilibrium constant $K_{sp}$ varies with temperature, pressure, and ionic strength, making these factors critical in scale prediction and management.

To evaluate the scaling tendency, one of the most widely used approaches for calculating the Saturation Index $(SI)$ is the Oddo-Tomson method. This method provides a reliable assessment of scaling risks by incorporating key parameters such as ion concentrations, temperature, and pressure. The $SI$ is determined using the following equation:

$$SI = log\left\{\left[Ca^{2+}\right]\left[Mg^{2+}\right]\right\} + A + BT + CT^2 + Dp + ES_i^{1/2} + FS_i + GS_i^{1/2}T\,. \tag{2.4}$$

The constant values for different sulfate scales, determined based on environmental temperature, are provided in Tab.2 [20, 21].

Table 2. Scaling types and corresponding parameter values for different sulfate minerals as a function of environmental temperature.

| Scaling type | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| Calcium sulfate | | | | | | | |
| *Gypsum ($< 80$℃) | $3.47$ | $1.8\text{E}-3$ | $2.5\text{E}-6$ | $-55.5\text{E}-5$ | $1.13$ | $0.37$ | $-2.0\text{E}-3$ |
| #Hemihydrate ($80$℃$-121$℃) | $4.04$ | $-1.9\text{E}-3$ | $11.9\text{E}-6$ | $-000.5\text{E}-5$ | $-1.66$ | $0.49$ | $-0.66\text{E}-3$ |
| Anhydrite ($> 121$℃) | $2.52$ | $9.98\text{E}-3$ | $-0.97\text{E}-6$ | $-445.5\text{E}-5$ | $-1.09$ | $0.50$ | $-3.3\text{E}-3$ |
| Strontium Sulfate (Celestite) | $6.11$ | $2.0\text{E}-3$ | $6.4\text{E}-6$ | $-667\text{E}-5$ | $-1.89$ | $0.67$ | $-1.9\text{E}-3$ |
| Barium Sulfate (Barite) | $10.03$ | $-4.8\text{E}-3$ | $11.4\text{E}-6$ | $-696\text{E}-5$ | $-2.62$ | $0.89$ | $-2.0\text{E}-3$ |

Scale prediction is typical method for scaling analysis in the oil and gas industry. Numerous models have been developed to address this issue. Thermodynamic models have long been the foundation for predicting scale formation in oilfield systems, providing insights into equilibrium conditions and the potential for mineral precipitation. These models typically estimate the degree of saturation using the Saturation Index $(SI)$, defined as the logarithm of the ratio between the ion activity product $(IAP)$ of scale-forming ions and the solubility product $(K_{sp})$:

$$SI = log\left(\frac{IAP}{K_{sp}}\right). \tag{2.5}$$

An $SI$ greater than zero indicates supersaturation and a high likelihood of scale precipitation. An $SI$ of zero suggests equilibrium, while a negative $SI$ implies undersaturation and potential scale dissolution [23, 24]. Common software tools such as PHREEQC, OLI Studio, and MULTISCALE utilize this approach to identify scaling tendencies based on water chemistry under given pressure and temperature conditions. Studies have shown that pressure and temperature changes, especially near the wellbore, can drastically reduce the solubility of sulfate salts and promote scale formation [25].

However, while thermodynamic models are valuable for identifying the conditions under which scaling is possible, they fail to capture the rate at which scaling actually occurs. This limitation is particularly critical in operational systems, where scale formation can occur rapidly, often well before thermodynamic equilibrium is established. Factors such as temperature, flow regime, and turbulence affect not just whether

scale forms, but how fast and where it accumulates. For example, elevated temperatures enhance reaction kinetics, while flow restrictions and high-shear zones promote nucleation and localized precipitation [26].

To address these limitations, researchers have turned to kinetic modeling and experimental methods such as dynamic coreflood tests. These studies simulate scaling behavior in porous media under flow conditions representative of the reservoir. Results have revealed how variables like contact time, ion concentration ratios, and residence time influence precipitation dynamics and permeability impairment [27, 28].

In parallel, data-driven approaches such as machine learning (ML) have gained prominence for their ability to model complex, nonlinear relationships between operating parameters and scale-related damage. Artificial neural networks (ANNs) and least squares support vector machines have been used to predict permeability decline, with inputs such as initial permeability and ion concentrations showing strong predictive power [6]. Other models have used variables including pressure drop, temperature, sulfate and barium concentrations, and injected volume, achieving prediction errors as low as $2.03\%$ [7]. More recently, generalized regression neural networks have identified injected volume and sulfate concentration as dominant features driving permeability loss [29].

To mitigate the detrimental effects of mineral scale deposition, chemical scale inhibitors are routinely employed during waterflooding operations, typically through squeeze treatments where inhibitors are injected into the near-wellbore region. These inhibitors function by disrupting scale formation pathways either by preventing nucleation or inhibiting crystal growth. Conventional inhibitors include inorganic polyphosphates, phosphonates, organic phosphate esters, and amino-phosphonates [30-32]. The efficacy of these treatments depends on various reservoir and operational parameters such as ion content, temperature, $pH$ and rock adsorption characteristics [33-35]. For instance, the squeeze lifetime of an inhibitor is directly influenced by its adsorption and desorption kinetics under reservoir conditions.

In this study, we have used a dataset of about $341$ experimental data points that were originally collected by Merdhah [27]. The purpose of these experiments was to study the loss of permeability caused by barium sulfate deposition in porous media under different conditions. We mathematically derive the underlying scaling mechanisms by explicitly evolving symbolic regression $(SR)$. $SR$ provides a distinct advantage by generating clear and interpretable equations that reveal scaling behavior and accurately predict critical thresholds. Moreover, it effectively captures nonlinear interactions and complex relationships between scaling factors. As a result, our work has enabled real-time monitoring of water injection to prevent salt deposition within the reservoir.

## 2. Data gathering

In this study, a comprehensive and reliable databank was compiled from open-source literature [36] to develop general models for barium sulfate $(BaSO_4)$ scaling prediction. The experimental methodology included two main procedures: the beaker and the core flooding. In the beaker test, synthetic formation water and seawater were combined in specified volumes under controlled conditions (Tab.3), heated to specific temperatures, and analyzed using atomic absorption spectrometry to determine $BaSO_4$ solubility at temperatures varying between $40°C$ to $90°C$. The core flooding test involved the use of sandstone cores ($3$ inches long, $1$ inch in diameter), which were vacuum-saturated with formation water before injecting seawater at different temperatures $(50-80°C)$ and differential pressures $(100-200\ psig)$. The permeability reduction due to scale deposition was continuously monitored with a pressure transducer and Scanning Electron Microscopy was used to examine the morphology and particle size of the deposited $BaSO_4$ crystals. Key parameters investigated included temperature, differential pressure, brine ionic concentration, permeability reduction, reaction rate constant, pore volume, and flow rate, with residence time being computed based on the injection flow rate (Tab.4). The assembled databank comprises $431$ measured data points, each containing values for ionic concentrations of strontium, calcium, barium, and sulfate, as well as initial permeability $(K_i)$, differential pressure $(\Delta P)$, temperature $(T)$, pore volume $(V_p)$, flow rate $(Q)$, and damaged permeability $(K_d)$.

Table 2. Experimental data summary.

| Parameter | Description | Range/Values |
|---|---|---|
| **Temperature** $(T)$ | Fluid temperature during tests | $40 – 90°C$ (Beaker)<br>$50 – 80°C$ (Core Flooding) |
| **Differential Pressure** $(\Delta P)$ | Applied pressure difference in core flooding | $100 – 200$ *psig* |
| **Brine Composition** | Ionic concentrations in formation water and seawater | Strontium, Calcium, Barium, Sulfate |
| **Initial Permeability** $(K_i)$ | Permeability of core before scaling | Measured in *mD* |
| **Damaged Permeability** $(K_d)$ | Permeability reduction due to scaling | Measured in *mD* |
| **Pore Volume** $(V_p)$ | Volume of voids in sandstone cores | Measured in $cm^3$ |
| **Flow Rate** $(Q)$ | Brine injection rate during tests | Measured in $cm^3$ / min |
| **Residence Time** $(\Delta t)$ | Time brine spends in the core | Computed from $V_p / Q$ |

Table 3. Brine composition (ionic concentrations in *ppm*).

| Ion | Normal Formation Water | High-Barium Formation Water | Barton Seawater | Angsi Seawater |
|---|---|---|---|---|
| **Sodium** $(Na^+)$ | *42,707* | *42,707* | *9,749* | *10,804.5* |
| **Potassium** $(K^+)$ | *1,972* | *1,972* | *340* | *375.05* |
| **Magnesium** $(Mg^{2+})$ | *102* | *102* | *1,060* | *1,295.25* |
| **Calcium** $(Ca^{2+})$ | *780* | *780* | *384* | *429.2* |
| **Strontium** $(Sr^{2+})$ | *370* | *370* | *5.4* | *6.577* |
| **Barium** $(Ba^{2+})$ | *250* | *2,200* | $< 0.2$ | – |
| **Chloride** $(Cl^-)$ | *66,706* | *67,713* | *17,218* | *19,307.45* |
| **Sulfate** $(SO_4^{2-})$ | *5* | *5* | *2,960* | *2,750* |
| **Bicarbonate** $(HCO_3^-)$ | *2,140* | *2,140* | *136* | *158.8* |

## 3. Symbolic regression

Symbolic regression seeks to discover explicit mathematical expressions that best represent underlying patterns in data. It finds analytical expressions that accurately fit a dataset. Pioneering work in this field includes contributions by Langley *et al*. [37-39]. At the core of symbolic regression lies a search algorithm designed to explore different function trees-mathematical structures representing potential equations (Fig.1). The algorithm iteratively adds nodes, refines constants, and optimizes functions to improve accuracy in predicting observed data.

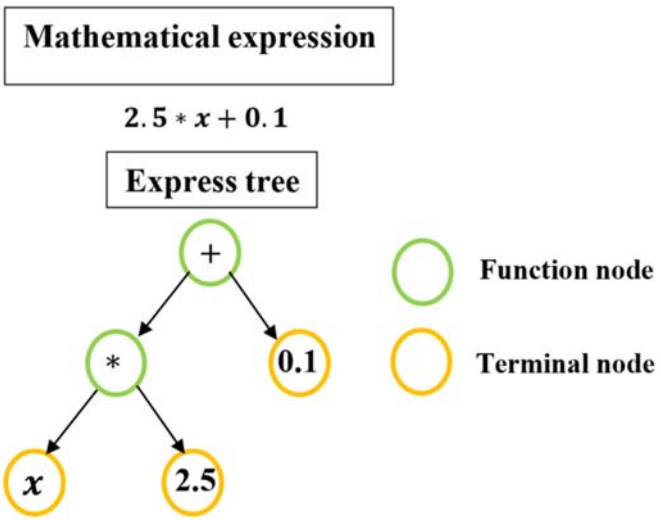


Fig.1. Expression tree representation of a mathematical formula.

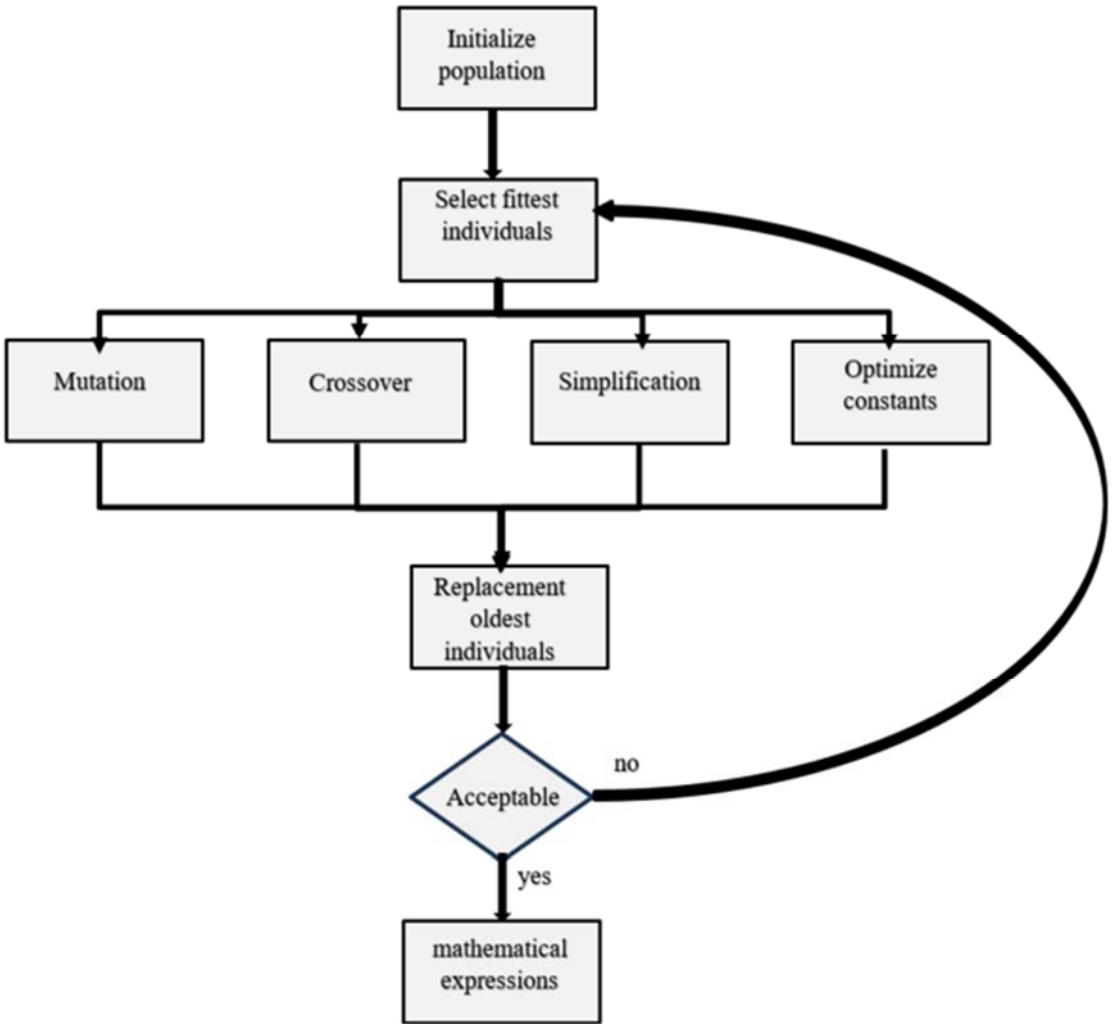


Fig.2. Evolutionary process of symbolic regression.

Unlike conventional machine learning methods, which rely on complex neural networks or statistical models, symbolic regression prioritizes both accuracy and interpretability. This makes it particularly valuable in fields such as physics, engineering, bioinformatics, and materials science. Figure 2 presents a structured workflow diagram illustrating the methodology for developing a Symbolic Regression Process. This process follows a systematic approach to iteratively refine mathematical expressions that best capture underlying data patterns. It begins by randomly subsampling the dataset, selecting a subset for evaluation. The algorithm then identifies the fittest equations based on a predefined fitness criterion, such as accuracy or complexity. To enhance performance, genetic operations are applied, including mutation (modifying coefficients or exponents), crossover (combining elements from top-performing equations), simplification (removing redundant terms), and constant optimization (adjusting numerical values for a better fit). Finally, older equations are replaced with new candidates to maintain diversity and improve predictive accuracy. This iterative approach ensures that symbolic regression not only discovers highly accurate models but also maintains interpretability, making it a valuable tool in scientific and engineering applications. In the present work, we have used PySR to write our script for modeling permeability degradation as a function of important reservoir parameters. PySR (**Py**thon **S**ymbolic **R**egression) is the most popular and reliable symbolic regression tool initially conceived by Cranmer [40] and currently maintained as an open-source project with contributions from researchers around the globe. This package allows scientists from a variety of disciplines to get interpretable, human-readable equations from data, providing a game changing alternative to black-box machine learning models.

## 4. Damaged permeability prediction

Mineral scale formation in porous media is controlled by thermodynamics, where supersaturation and pressure influence solubility, and kinetics, where temperature, viscosity, and hydrodynamics regulate precipitation rates. Figure 3 illustrates the relationship between the sulfate scale saturation indices for $BaSO_4$ (Barium Sulfate), $SrSO_4$ (Strontium Sulfate), and $CaSO_4$ (Calcium Sulfate) and the relative permeability reduction. The weak correlation coefficients suggest an almost negligible positive relationship, indicating that fluctuations in sulfate saturation indices have little to no direct effect on permeability impairment. This finding is likely due to the fact that all tested conditions remained within the supersaturation zone, where only minor variations in saturation levels occur. Consequently, the extent of permeability damage is more strongly governed by reaction kinetics and induction time, which control the nucleation, growth, and deposition of scale within the porous medium, rather than by the absolute degree of supersaturation itself. Therefore, the proposed SR-based model is extended, as follows:

$$K_d = f\left(\Delta p, T, Q_i, \Delta t, K_i\right) \tag{4.1}$$

where, the symbols $K_d$, $\Delta p$, $T$, $Q_i$ $\Delta t$, $K_i$ indicate the damaged permeability, differential pressure, temperature, flowrate, initial permeability.

The dataset is partitioned into training $(80\%)$ and testing $(20\%)$ subsets to ensure unbiased model evaluation. The input features $(X)$ include five parameters $(dp, T, Q, dt, K_i)$, while the target variable $(y)$ is $\log(K_d)$, representing permeability damage. A symbolic regression model is initialized using PySR (Python Symbolic Regression), which employs an evolutionary algorithm to discover an interpretable mathematical equation. The model is configured with binary operators $(*, +, -, /)$ and unary operators (square, cube, exp, log, inverse) to optimize equation discovery while maintaining a balance between accuracy and simplicity. Constraints on complexity, nesting, and operator usage are applied to control model interpretability. The training process is iterated $500$ times, with each iteration generating a new symbolic equation. To assess model performance, predictions are made on the test dataset, and the Mean Squared Error $(\mathrm{MSE})$ and $R^2$ score are

computed to evaluate predictive accuracy. Here are some example discovered equations that satisfy the conditions of training MSE below $10^{-4}$ and $R^2$ exceeding $97\%$.

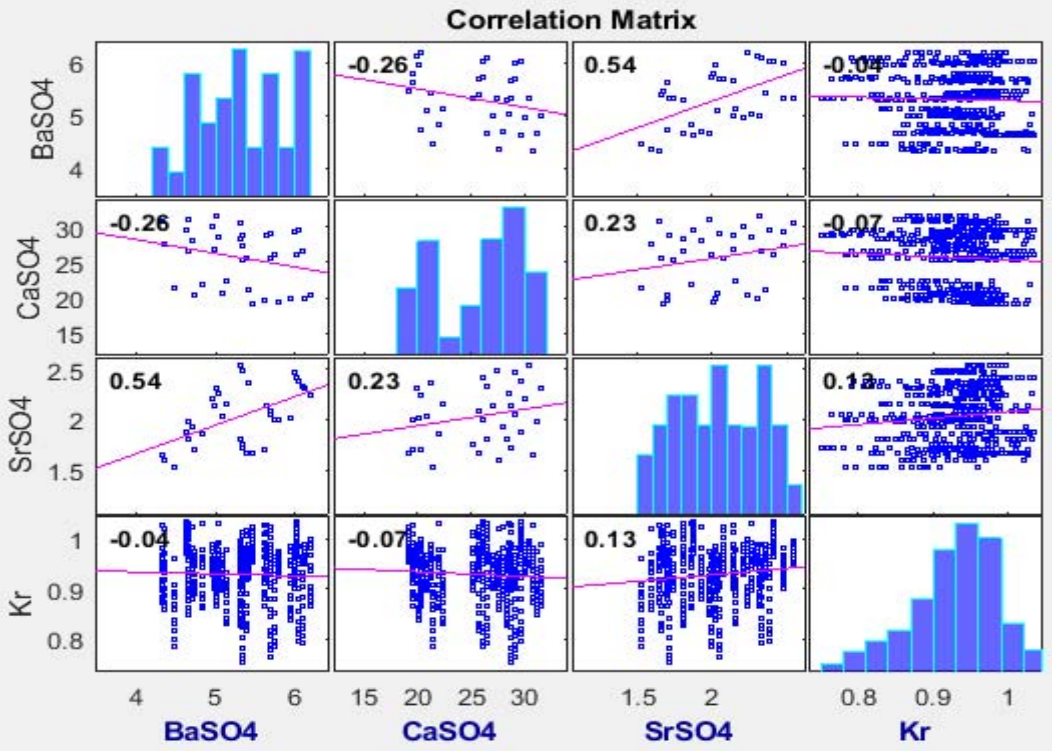


Fig.3. Pairwise correlation matrix of sulfate compounds and permeability ratio.

Table 4. Some symbolic regression equations for permeability damage prediction.

| N° | Equation | Test MSE | Test R2 |
|---|---|---|---|
| **1** | $\log(K_d) = \frac{291.0Q_i}{(\Delta p + 1.23)(T + 16.8)} + 0.669$ | $2.18\text{E}-05$ | $0.9729$ |
| **2** | $\log(K_d) = \frac{-0.000607\Delta p\left(T + \log(K_i)^3\right)}{Q_i} - \frac{1}{K_i} + 1.59$ | $2.06\text{E}-5$ | $0.9744$ |
| **3** | $log(K_d) = -\frac{\Delta p}{Q_i}(0.000581T + 0.0104) - 0.000301\Delta t + +9.9910^{-7}(K_i - 0.000825)^3 + 1.49$ | $1.47\text{E}-05$ | $0.9818$ |
| **4** | $log(K_d) = 1.58e^{-\left(\frac{0.000535\Delta pT + \frac{0.00966\Delta p^2 + log(\Delta t)}{\Delta p}}{Q_i}\right)}$ | $1.75\text{E}-05$ | $0.9782$ |
| **5** | $log(K_d) = -0.000718T\left(\frac{\Delta p}{Q_i} - 1.92\right) + \frac{1.45K_i - 0.704}{K_i}$ | $2.11\text{E}-05$ | $0.9738$ |

The discovered equations suggest (Tab.5.) that permeability damage in porous media is primarily governed by the ratio of pressure drop to flow rate $\frac{\Delta p}{Q}$, which represents the hydraulic resistance of the system (injectivity index). This ratio is a key indicator of how easily fluid moves through the porous structure, with higher values suggesting increased flow resistance and potential scaling or formation damage. The ratio $\frac{\Delta p}{Q_i}$ (pressure drop per unit flow rate) is directly related to the skin factor $(S)$ and near-wellbore effects, both of which influence injectivity and permeability near the well. In a damaged formation, the skin factor represents additional flow resistance due to scaling, fines migration, or other permeability-reducing mechanisms. The injectivity index is defined as $\frac{Q_i}{\Delta p}$, meaning that an increase in $\frac{\Delta p}{Q}$ corresponds to a decrease in injectivity, indicating formation damage or restricted flow paths. The relationship between skin factor and pressure drop can be expressed using the linear flow equation:

$$I = \frac{Q_i}{\Delta p} = \frac{0.001127 K_i A}{\mu (L + S)} \tag{4.2}$$

where $K_i$ represents the original permeability of the porous medium, measured in millidarcies ($mD$). $h$ is the cross-sectional area ($ft^2$) through which the fluid moves, influencing the overall flow capacity. $\mu$ denotes the fluid viscosity $(cP)$, which affects resistance to flow, with higher viscosity leading to greater pressure drops. $L$ refers to the flow length ($ft$), representing the distance the fluid travels through the medium. Finally, $S$ is the skin factor, accounting for damage effects, where positive values indicate formation damage (reducing injectivity) and negative values suggest stimulation (enhancing injectivity). Despite their complexity, these equations demonstrate high predictive accuracy, with $R^2$ values exceeding $95\%$ and training MSE below $10^{-4}$, validating their effectiveness in capturing system dynamics. However, their interpretability remains a challenge, necessitating further simplifications to extract a more generalized governing equation that highlights the dominant physical mechanisms.

Reducing these expressions into a single function of $I = \frac{Q_i}{\Delta p}$ could provide deeper insights into permeability evolution and offer a practical step of scale control process for predicting flow behavior in porous media. Figure 4 presents a plot of $\log(K_d)$ versus $I$, revealing three distinct linear trends, each with $R^2$ values exceeding $98\%$, corresponding to different operating or experimental conditions. Within each trend, $\log(K_d)$ increases with $I$, indicating that a higher injectivity index correlates with greater permeability damage, or equivalently, less relative damage. The clear separation between these lines is attributed to temperature, which shifts the overall relationship, resulting in different intercepts or slopes under varying conditions. The tight clustering within each trend suggests consistent behavior within a specific temperature range, while the offsets between trends highlight the multifactorial nature of permeability damage and the necessity of incorporating additional variables into a comprehensive predictive model. After removing the outliers from the dataset (five tests removed), which can be observed in Fig.4, as they deviate significantly from the trendline, we will take advantage of one of the key strengths of symbolic regression: its ability to discover the optimal equation structure. In this case, our objective is to derive an equation of the form:

$$\log(K_d) = f(T) * I + g(T, \Delta t). \tag{4.3}$$

This approach allows us to identify the best-fit parameters while ensuring the model captures the underlying relationship effectively. In total, we have generated approximately $50$ equations. Here are some example equations that were discovered (Tab.6).

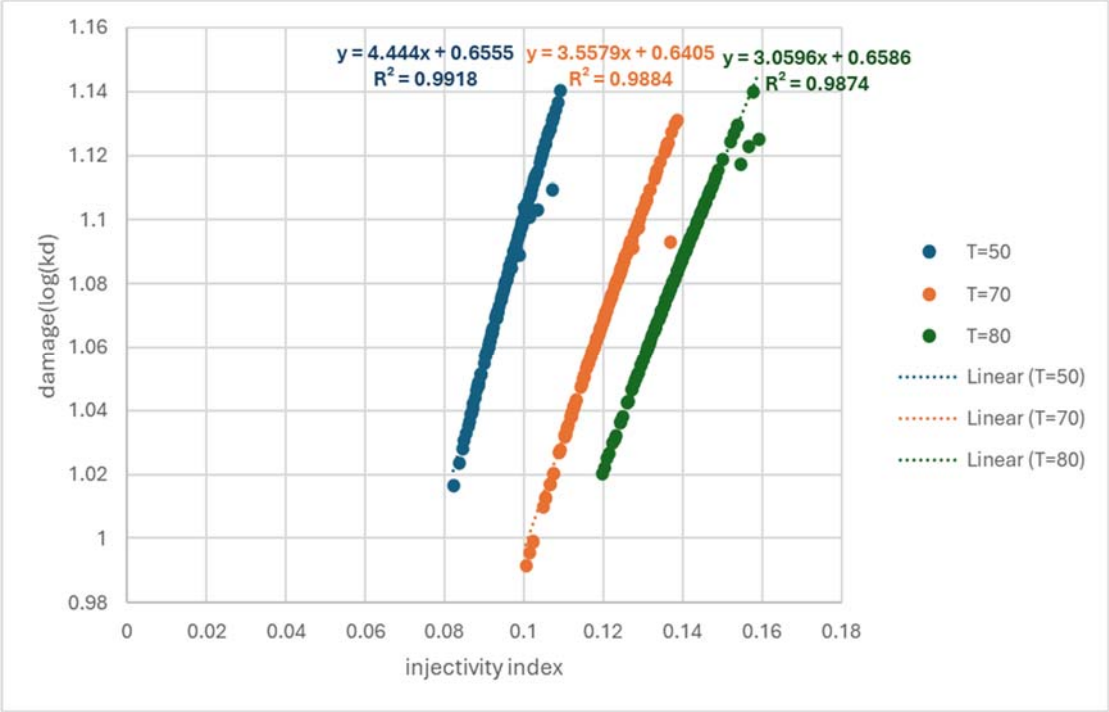


Fig.4. Effect of injectivity index on permeability damage at different temperatures.

Table 5. Performance evaluation of f and g equations for scaling prediction.

| $f$ Equation | $g$ Equation | Test $MSE$ | Test $R^2$ |
|---|---|---|---|
| $f = \frac{313.54}{T + 18.20}$ | $g = 0.6409$ | $2.05\text{E}-6$ | $0.997$ |
| $f = \frac{313.33}{T + 18.24}$ | $g = 0.6414$ | $2.05\text{E}-6$ | $0.995$ |
| $f = \frac{312.08}{T + 18.20}$ | $g = 0.6429$ | $2.05\text{E}-6$ | $0.994$ |

We note the function $f$ predominantly follows an inverse relationship, expressed as

$$f = \frac{A}{T + B}. \tag{4.4}$$

These equations achieve remarkably low test MSE values $\left(10^{-6}\right)$ and consistently high $R^2$ scores $(0.99)$, demonstrating strong predictive performance. On the other hand, $g$ remains relatively stable across different equations, typically represented as a nearly constant value close to $0.64$. in some cases, minor variations appear, but they do not significantly impact the overall predictive accuracy. So, finally we can conclude that the final model to predict damage permeability has this form:

$$\log(K_d) = \frac{A}{T+B} I + C \,. \tag{4.5}$$

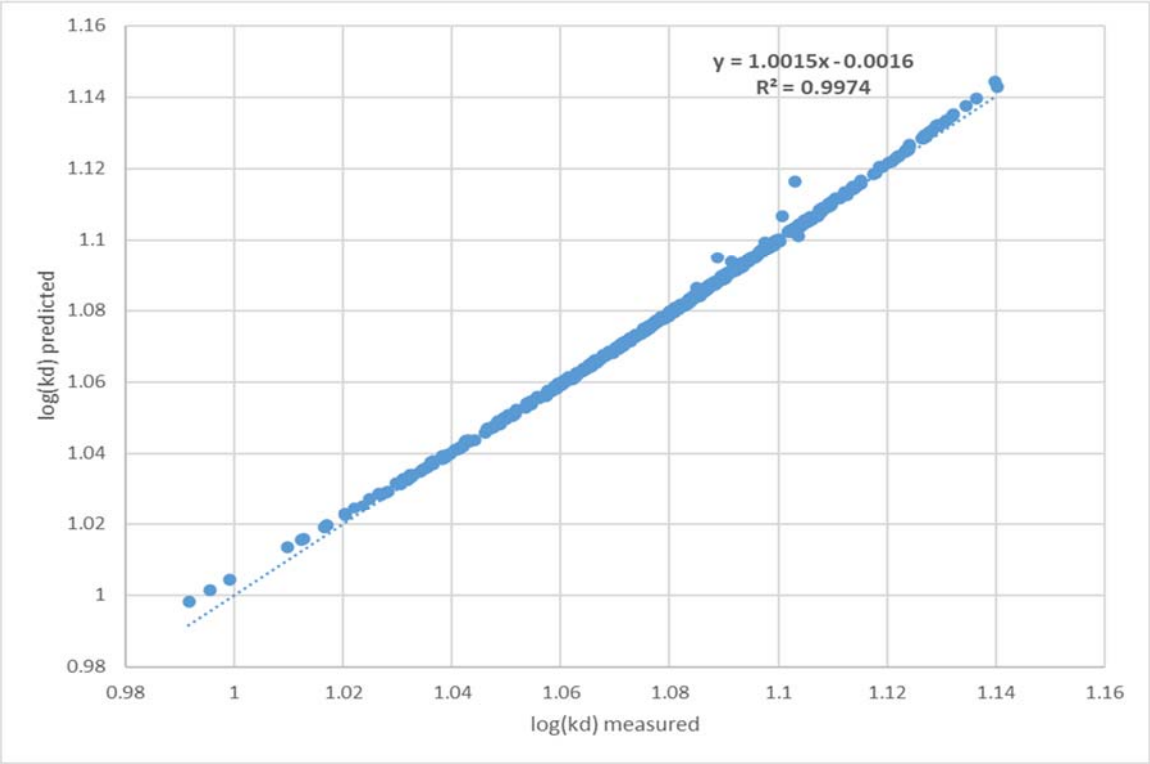


Fig.5. Regression performance: predicted vs. measured $\log(K_d)$.

We can also note that residence time does not appear in this equation. However, this does not imply that it has no influence. According to the author's experimental notes [15], scale formation was most significant near the formation water inlets, while minimal scaling was observed further from these entry points. This suggests that precipitation occurs rapidly at the front of the core, with residence time being much shorter than the total injection duration. This observation highlights the need for researchers to develop models that account for the declining trend of permeability over time. Permeability decreases sharply soon after the mixing of waters within the pores, followed by a slower decline that eventually stabilizes after a significant reduction in permeability. The scatter plot (Fig.5) demonstrates a strong correlation between measured and predicted $\log(K_d)$ values, with a near-unity regression slope $(1.0015)$ and an excellent fit $\left(R^2 = 0.9974\right)$, indicating high model accuracy in predicting permeability damage due to scaling.

Figure 6 shows that the errors are approximately normally distributed, with a Gaussian fit indicating a mean $\mu$ of $-4.51\times10^{-5}$ and a standard deviation $\sigma$ of $1.09\times10^{-3}$. The distribution is centered around zero, suggesting minimal bias in the predictions, with the majority of errors falling within $\pm2\times10^{-3}$. This indicates that the model exhibits a relatively high level of accuracy with minimal bias in the predicted values. The concentration of errors near zero and the relatively narrow standard deviation further suggest that the model is effective in accurately predicting most of the data points. However, the presence of some outliers points to potential areas for improvement in the model's ability to handle occasional extreme discrepancies.

Figure 7 presents a scatter plot highlighting the scaling region in red, where $K_d/K_i < 1$. This plot is generated for a fixed temperature, illustrating how permeability damage varies with injectivity index $I$ on the

$x$-axis and initial permeability on the $y$-axis. As observed in Fig.6, distinct clusters appear for different temperature conditions. The boundary between the scaling and non-scaling regions is defined by the threshold equation

$$I = \frac{(B+T)\left(log\left(K_i\right)-C\right)}{A}. \tag{4.6}$$

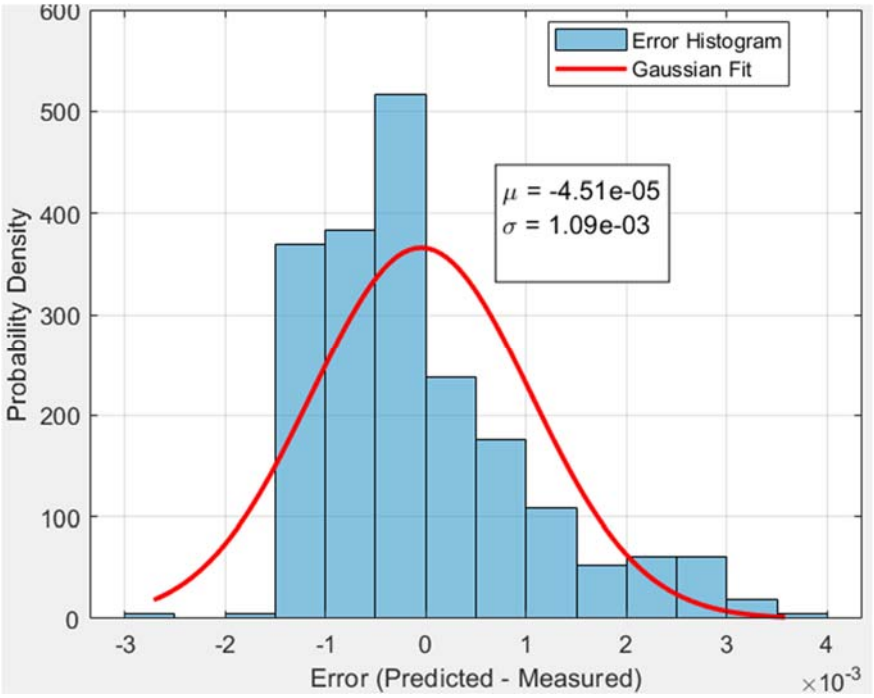


Fig.6. Histogram of prediction error with gaussian fit.

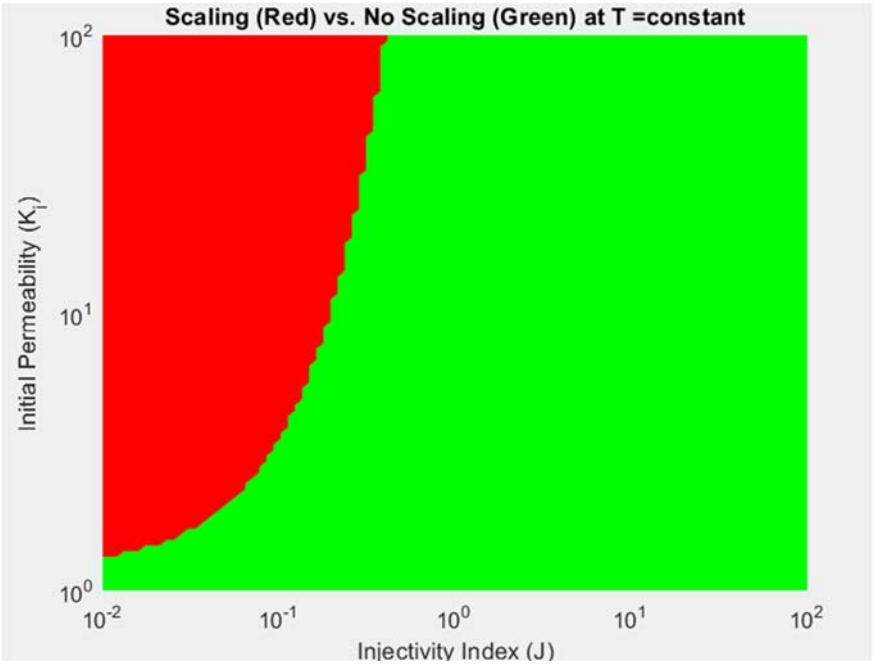


Fig.7. Scaling risk map: effect of injectivity index and initial permeability at constant temperature.

Note that scaling occurs primarily at low $I$ and $K_i$ values, where flow restrictions promote precipitation. In contrast, higher $I$ and $K_i$ reduce scaling risks. The curved boundary separating both regions represents a critical $I$ threshold, beyond which scaling is prevented. Maintaining $I$ above this threshold is crucial for minimizing permeability damage, emphasizing the need for optimized injection conditions and potential scale inhibitors in low-permeability formations.

## 5. Conclusion

In conclusion, this study successfully developed and evaluated a mathematical model for predicting permeability loss using symbolic regression. The model demonstrated high predictive accuracy, with an average absolute error below $10^{-4}$ and a correlation coefficient of $0.99$, indicating strong agreement with experimental data. Additionally, a risk map was generated to assess the likelihood of scaling under dynamic flow conditions, providing a valuable tool for monitoring and identifying scaling risks. The model's reliability was ensured through a rigorous development and validation process, utilizing an $80/20$ training-testing split. By uncovering explicit mathematical relationships within the data, symbolic regression proved to be highly effective in capturing complex scaling patterns. These results highlight the model's potential as a powerful and practical step of scale control process for permeability loss prediction and scale control in oilfield operations, contributing to improved flow assurance strategies.

## Acknowledgment

This research was funded by the Direction Générale de la Recherche Scientifique et du Développement Technologique (DGRSDT) of the Algerian Ministry of Higher Education and Scientific Research under the PNR (Projet National de Recherche) program, through the project entitled "Development of Decision-Making Tools for Controlling and Optimizing Production from Petroleum Fields."

## Nomenclature

$I$ – injectivity index $\frac{cm^3}{psi.S}$
$IAP$ – Ion Activity Product (same unit as $K_{sp}$)
$K_d$ – damaged permeability $mD$
$K_{sp}$ – solubility product $mol^2 / L^2$
$K_i$ – initial permeability $mD$
$(\text{MSE})$ – Mean Squared Error
PySR – Python Symbolic Regression (modeling tool)
$(Q)$ – flow rate $cm^3 / s$
$R^2$ – coefficient of determination (dimensionless)
$SI$ – Saturation Index (dimensionless)
$SR$ – Supersaturation Ratio, defined as $IAP / K_{sp}$ (dimensionless)
$T$ – temperature $°C$
$V_p$ – core pore volume $cm^3$
$\{X\}$ – activity of species $X$ (dimensionless)
$\Delta p$ – differential pressure $psi$
$\Delta t$ – residence time $s$

## References


[1] Crabtree M., Eslinger D., Fletcher P., Miller M., Johnson A. and King G. (1999): *Fighting scale-removal and prevention*.– Oilfield Review, vol.11, No.3, pp.30-45.

[2] Moghadasi J., Müller-Steinhagen H., Sharif A., Jamialahmadi M. and Mota M. (2004): *Scale formation in oil reservoir and production equipment during water injection (a review of recent research achievements*.– Desalination, vol.167, pp.235-248.

[3] Yuan M. and Todd A.C. (1991): *Mechanisms for calcium carbonate scale formation in gas wells and lines*.– SPE Production Engineering, vol.6, No.4, pp.446-452.

[4] Kan A.T., Fu G. and Tomson M.B. (2002): *Scale prediction for oilfield scale control*.– SPE International Symposium on Oilfield Scale, SPE-74563-MS.

[5] Khormali A., Ahmadi S. and Aleksandrov A.N. (2025): *Analysis of reservoir rock permeability changes due to solid precipitation during waterflooding using artificial neural network*.– Journal of Petroleum Exploration and Production Technology, vol.15, No.1, pp.1-18.

[6] Al-Shayji K., Al-Dousari M. and Al-Mutairi A. (2014): *Prediction of scale formation in oilfield applications using artificial neural network*.– International Journal of Chemical Engineering and Applications, vol.5, No.3, pp.215-220.

[7] Yousef A.A. and Al-Jawad M.S. (2015): *Prediction of formation damage due to scale deposition using machine learning*.– Journal of Petroleum Science and Engineering, vol.133, pp.206-214.

[8] Mitchell A.C. and Ferris F.G. (2005): *The influence of Bacillus pasteurii on the nucleation and growth of calcium carbonate*.– Geomicrobiology Journal, vol.22, No.1-2, pp.133-144.

[9] ASTM D1125 - 95(2020): *Standard Test Methods for Electrical Conductivity and Resistivity of Water*.– American Society for Testing and Materials.

[10] Yuan M., Kan A.T. and Tomson M.B. (2001): *Squeeze lifetime predictions for mixed inhibitors*.– SPE International Symposium on Oilfield Chemistry, SPE-65036-MS.

[11] Nasr-El-Din H.A., Al-Humaidi A.I. and Al-Harthy S.S. (2005): *Formation damage due to iron precipitates resulting from seawater injection*.– SPE Production & Operations, vol.20, No.2, pp.135-144.

[12] Kelland M.A. (2014): *Production Chemicals for the Oil and Gas Industry*.– (2nd ed.), CRC Press.

[13] Bénézeth P., Palmer D.A., Wesolowski D.J. and Mroczek E.K. (2007): *Solubility and thermodynamics of barium sulfate (barite) at 0.1 MPa from 0 to 90°C*.– Geochimica et Cosmochimica Acta, vol.71, No.3, pp.637-650.

[14] Flett M.A. and Mutch R.A. (2003): *Formation damage due to sulfate scaling during water injection*.– Journal of Petroleum Science and Engineering, vol.38, No.1-2, pp.31-48.

[15] Khormali A., Koochi M.R., Varfolomeev M.A. and Ahmadi S. (2023): *Experimental study of the low salinity water injection process in the presence of scale inhibitor and various nanoparticles*.– Journal of Petroleum Exploration and Production Technology, vol.13, No.3, pp.903-916.

[16] Söhnel O. and Garside J. (1992): *Precipitation: Basic Principles and Industrial Applications*.– Butterworth-Heinemann.

[17] Stumm W. and Morgan J.J. (1996): *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*.– (3rd ed.), Wiley-Interscience.

[18] Zhang Y. and Dawe R.A. (2000): *The kinetics of carbonate scale precipitation from formation water and seawater mixtures*.– Journal of Petroleum Science and Engineering, vol.25, No.3-4, pp.243-253.

[19] Sadiq M. (1992): *Toxic Metal Chemistry in Marine Environments*.– CRC Press.

[20] Oddo J.E. and Tomson M.B. (1994): *Why scale forms in the oil field and methods to predict it*.– SPE Production & Facilities, vol.9, No.1, pp.47-54.

[21] Tomson M.B. and Kan A.T. (2005): *Scale prediction for oil and gas production*.– SPE International Symposium on Oilfield Scale, SPE-96021-MS.

[22] Amjad Z. (1995): *Calcium Phosphates in Biological and Industrial Systems*.– Springer.

[23] Appelo C.A.J. and Postma D. (2005): *Geochemistry, Groundwater and Pollution*.– (2nd ed.), CRC Press.

[24] OLI Systems Inc. (2021): *OLI Studio: Stream Analyzer User Guide*.

[25] Khormali A., Petrakov D.G., Lamidi A.L.B. and Rastegar R. (2015): *Prevention of calcium carbonate precipitation during water injection into high-pressure high-temperature wells*.– In SPE European Formation Damage Conference and Exhibition, pp.SPE-174277.

[26] Mackay E.J. and Jordan M.M. (2003): *A kinetic model for barium sulfate scale formation*.– SPE Production & Facilities, vol.18, No.1, pp.11-17.

[27] Merdhah A.B. and Yassin A.A.M. (2009): *Formation damage due to barium sulfate scale formation in sandstone rock*.– Journal of Applied Sciences, vol.9, No.22, pp.4110-4114.

[28] Al-Rawajfeh A.E. and Al-Shamaileh E. (2007): *Kinetics and thermodynamics of scale formation in RO desalination plants*.– Desalination, vol.206, No.1-3, pp.395-402.

[29] Khan F., Al-Fawzan M. and Al-Rashed A. (2021): *Predicting permeability loss due to scale using GRNN*.– SPE International Oilfield Scale Conference and Exhibition, SPE-200747-MS.

[30] Kan A.T., Fu G. and Tomson M.B. (2005): A*dsorption and precipitation of scale inhibitors: Effects of temperature and pH*.– SPE Production & Facilities, vol.20, No.2, pp.144-150.

[31] Sorbie K.S. (1991): *Polymer-Improved Oil Recovery*.– CRC Press.

[32] Tomson M.B., Kan A.T., Fu G. and Zhang Q. (2002): *Inhibitor and water chemistry impacts on scale deposition*.– SPE International Symposium on Oilfield Scale, SPE-74689-MS.

[33] Liu X., Kan A.T. and Tomson M.B. (2004): *Adsorption/desorption of inhibitors on rock*.– SPE International Symposium on Oilfield Chemistry, SPE-86679-MS.

[34] Wang Y. and Papenguth H. (2000): *Field-scale modeling of squeeze treatments for scale control*.– SPE Production & Facilities, vol.15, No.2, pp.112-119.

[35] Khormali A. and Ahmadi S. (2023): *Synergistic effect between oleic imidazoline and 2-mercaptobenzimidazole for increasing the corrosion inhibition performance in carbon steel samples*.– Iranian Journal of Chemistry and Chemical Engineering, vol.42, No.1, pp.321-336.

[36] Rostami A., Shokrollahi A., Shahbazi K. and Ghazanfari M.H. (2019): *Application of a new approach for modeling the oil field formation damage due to mineral scaling*.– Oil & Gas Science and Technology-Revue d'IFP Energies Nouvelles, vol.74, pp.62.

[37] Langley P. (1981): *Data-driven discovery of physical laws*.– Cognitive Science, vol.5, No.1, pp.31-54.

[38] Koza J.R. (1990): *Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems*.– Stanford, CA: Stanford University, Department of Computer Science, vol.34.

[39] Schmidt M. and Lipson H. (2009): *Symbolic regression of implicit equations*.– In Genetic programming theory and practice, vol.7, pp.73-85, Boston, MA: Springer US.

[40] Cranmer M. (2023): *Interpretable machine learning for science with PySR and Symbolic Regression*.– jl. arXiv preprint arXiv:2305.01582.

Received: April 6, 2025
Revised: August 13, 2025