

Machine Learning Models to Predict Total Skin Factor in Perforated Wells

S. Thabet and A. Elhadidy, Khalda Petroleum Co., Cairo, Egypt.; M. Elshielh, Gupco Petroleum Co., Cairo, Egypt.; A. Taman, Khalda Petroleum Co., Cairo, Egypt.; A. Helmy, Gupco Petroleum Co., Cairo, Egypt.; H. Elnaggar, Badr Petroleum Co., Cairo, Egypt.; T. Yehia, Texas A&M University, College Station, Texas, USA.

Copyright 2024, Society of Petroleum Engineers DOI 10.2118/218838-MS

This paper was prepared for presentation at the SPE Western Regional Meeting held in Palo Alto, California, USA, 16 - 18 April 2024.

This paper was selected for presentation by an SPE program committee following review of information contained in an abstract submitted by the author(s). Contents of the paper have not been reviewed by the Society of Petroleum Engineers and are subject to correction by the author(s). The material does not necessarily reflect any position of the Society of Petroleum Engineers, its officers, or members. Electronic reproduction, distribution, or storage of any part of this paper without the written consent of the Society of Petroleum Engineers is prohibited. Permission to reproduce in print is restricted to an abstract of not more than 300 words; illustrations may not be copied. The abstract must contain conspicuous acknowledgment of SPE copyright.

Abstract

An accurate total skin factor prediction for an oil well is critical for the evaluation of the inflow performance relationship, and the optimization of the appropriate stimulation treatment such as acidizing and hydraulic fracturing. Performing well testing regularly is not economically feasible, and the equations used for total skin damage may not be accurate. In this work, the goal is to build machine learning (ML) models that can predict the total skin factor in perforated wells using accessible field data. Nine distinct ML algorithms such as Gradient Boosting (GB), Adaptive Boosting (AdaBoost), Random Forest (RF), Support Vector Machines (SVMs), Decision Trees (DT), K-Nearest Neighbor (KNN), Linear Regression (LR), Stochastic Gradient Descent (SGD), and Artificial Neural Network (ANN) are meticulously developed and fine-tuned using a substantial dataset derived from 1,088 wells. The dataset encompasses 19,040 data points, thoughtfully split into two subsets: 70% (13,328 data points) for training the algorithms, and 30% (5,712 data points) for testing their predictions. The parameters used are mostly gathered during well completion and conventional well testing operations, including liquid flow rate, water cut, gas oil ratio, bottomhole flowing pressure, reservoir pressure, reservoir temperature, reservoir permeability, reservoir thickness, perforations diameter, perforations density, perforations penetration depth, well deviation, and penetrated portion of the net pay thickness. In this study, the total skin factor acquired from conventional well test analysis serves as the model's output. K-fold cross-validation and repeated random sampling validation techniques are used to assess the performance of the models against the total skin obtained from the conventional well test analysis. The K-fold cross-validation outcomes of the top-performing ML models, specifically GB, AdaBoost, RF, DT, and KNN, reveal remarkably low mean absolute percentage error values reported as 3.2%, 3.2%, 2.9%, 3.3%, and 3.8%, respectively. Additionally, the correlation coefficients (R^2) for these models are notably high, with values of 0.972, 0.968, 0.975, 0.964, and 0.956, respectively. In conclusion, ML models demonstrated their ability to predict total skin factor for different reservoir fluid properties, well geometries, and completion configurations. ML models offer a more efficient, quick, and cost-effective alternative to the conventional well-testing analysis.