

Machine Learning Models to Predict Production Rate of Sucker Rod Pump Wells

S. Thabet, H. Zidan, A. Elhadidy, and 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 10.2118/218857-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

The design, operation, and optimization of Sucker Rod Pumping (SRP) systems necessitate the utilization of production data. However, forecasting fluid flow rates at the surface of SRP artificially lifted wells usually poses a challenge, especially in instances where traditional separators and multiphase flowmeters are not universally available. Consequently, this study introduces nine machine learning (ML) models employing real data sourced from 598 wells with a production history exceeding three years. The dataset, comprising 8,372 data points, undergoes a random split allocating around 80% of the data (6,697 data points) for training, while around 20% (1,675 data points) are used for testing. The ML models encompass Gradient Boosting (GB), Adaptive Boosting (AdaBoost), Random Forest (RF), Support Vector Machines (SVMs), Decision Tree (DT), K-Nearest Neighbor (KNN), Linear Regression (LR), Artificial Neural Network (ANN), and Stochastic Gradient Descent (SGD). The chosen input features for the models are readily accessible during any SRP well-lifting process, and these inputs include various variables such as wellhead flowing pressure, casing pressure, predicted bottom hole fluid production rate, predicted bottom hole oil production rate, net liquid head above the pump, pump size, pump clearance, stroke length, pump speed, pump setting depth, the temperature at the pump depth, oil gravity, and water viscosity. Evaluation of the different ML models' performance is carried out by two methodologies: K-fold cross-validation, and repeated random sampling. The findings reveal that the top-performing models are GB, AdaBoost, RF, LR, and SGD, exhibiting mean absolute percentage errors of 3.6%, 3.4%, 3.4%, 4.0%, and 4.4% respectively, and correlation coefficients (R^2) of 0.937, 0.934, 0.935, 0.921, and 0.915, respectively. In practical field application, these models are deployed on a well within Egypt's Western Desert fields, demonstrating excellent agreement between actual fluid rates and model predictions. In conclusion, across diverse pumping scenarios and completion configurations, the ML models could effectively forecast production rates for different SRP wells. This capability facilitates continuous monitoring, optimization, and performance analysis of SRP wells, enabling swift responses to operational challenges since the proposed ML models offer an accessible, rapid, and cost-effective alternative to conventional separators and multiphase flowmeters.