Petroleum 3 (2017) 280–285

Contents lists available at ScienceDirect

Petroleum

journal homepage: www.keaipublishing.com/en/journals/petlm

Prediction of sand production onset in petroleum reservoirs using a reliable classification approach

Farhad Gharagheizi [a], Amir H. Mohammadi [b, c, d, *], Milad Arabloo [e], Amin Shokrollahi [e]

[a] Department of Chemical Engineering, Texas Tech University, Lubbock, TX, United States
[b] Thermodynamics Research Unit, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban, 4041, South Africa
[c] Institut de Recherche en Génie Chimique et Pétrolier (IRGCP), Paris Cedex, France
[d] Département de Génie des Mines, de la Métallurgie et des Matériaux, Faculté des Sciences et de Génie, Université Laval, Québec (QC), G1V 0A6, Canada
[e] Young Researchers and Elites Club, North Tehran Branch, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Article history:
Received 11 October 2015
Received in revised form
31 January 2016
Accepted 3 February 2016

Keywords:
Sand production
Least square SVM
ROC graph
Classification description
Modeling
Sanding onset

ABSTRACT

Controlling sand production in the petroleum industry has been a long-standing problem for more than 70 years. To provide technical support for sand control strategy, it is necessary to predict the conditions at which sanding occurs. To this end, for the first time, least square support machine (LSSVM) classification approach, as a novel technique, is applied to identify the conditions under which sand production occurs. The model presented in this communication takes into account different parameters that may play a role in sanding. The performance of proposed LSSVM model is examined using field data reported in open literature.

It is shown that the developed model can accurately predict the sand production in a real field. The results of this study indicates that implementation of LSSVM modeling can effectively help completion designers to make an on time sand control plan with least deterioration of production.

1. Introduction

The sand production prediction is one of the long-standing and complex problems in the petroleum industry [1]. From the phenomenological viewpoint, sand production can occur when the formation do not present sufficient strength to withstand destabilizing forces generated during the flow of reservoir fluid [2–4]. A wide variety of problems such as production loss due to plugging the perforations [2,3,5,6], wellbore instability [6–8], equipment erosion [2,8–11], and additional costs for waste sand depositional [8,9,11] can occur due to sand production. This may result in reduced production time and increased maintenance and operating costs. It is a complex phenomenon which depends on various factors such as lithology of the formation, stress distribution around the wellbore, reservoir characteristics (i.e., rock and fluid properties), wellbore/completion geometry, and production conditions. Due to the importance of sand production prediction in the petroleum industry, great efforts have been directed to develop robust and reliable methods for sand production prediction. Morita et al. [12] proposed an analytical approach to study the effects of many parameters on sand production. According to their research following parameters may affect sand production: wellbore pressure and stress distribution around well; drag forces induced by fluid flow; rock strength of formation; shot density and perforation geometry; cyclic loading history. Morita and Boyd [13] presented analyses of five typical sand problems commonly observed in the field. The first is the sear-type sand problem in poorly consolidated sand formation in

* Corresponding author. Thermodynamics Research Unit, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban, 4041, South Africa.
E-mail addresses: a.h.m@irgcp.fr, amir_h_mohammadi@yahoo.com (A.H. Mohammadi).

Peer review under responsibility of Southwest Petroleum University.

Production and Hosting by Elsevier on behalf of KeAi

http://dx.doi.org/10.1016/j.petlm.2016.02.001

Alaska. The second is the sands produced from intermediate strength formation following water break through. Loss of capillary pressure holding the sand species together is the primary cause of this sand production [13]. The third is from North Sea reservoirs where sand was produced from consolidated formations with reservoir pressure depletion. The fourth type of sand observed in consolidated formations located in California close to the San Andrea fault with high horizontal tectonic forces. Finally, the fifth sand production type was produced with high pressure gradient around cavity surface due to multiple perforations with high perforation depth. Doan et al. [14] developed numerical modeling of gravitational deposition of sand in a horizontal well in heavy oil reservoirs. Examination of simulation results revealed that oil viscosity and flow rate along with sand particle size play important roles in the gravity deposition of sands inside horizontal heavy oil reservoirs. Bianco and Halleck [9] experimentally evaluated the behavior, morphology, and stability of sand arches near the wellbore in two-phase saturated sand samples. They analyzed the effects of changes in fluid flow velocity and water saturation during the sand production process. With the growing use of artificial intelligence modeling in petroleum engineering, Kanj and Abousleiman [15] developed an artificial neural network (ANN) modeling to predict important sanding indication parameters for gas wells of Northern Adriatic Basin. Later, another ANN model was developed by Azad et al. [16] to predict critical bottom hole flowing pressure inhibiting sand production using 38 datasets collected from three different oilfields wells located in the southeast of Iran. The results showed that the developed model provides better prediction than other analytical models. Although the ANN-based models have been generally proven to provide high accuracy for prediction problems [17–20], they often have some limitations and shortcomings. They may have the disadvantages of non-reproducibility of results, over-fitting problems, getting stuck in local minima, setting too many controlling parameters, and etc [21–25]. In this communication, for the first time, a new computer-based algorithm namely least square support vector machine (LSSVM) modeling approach is presented for identifying the sanding condition.

2. Field data and model development

2.1. Database collection

A data set of 31 wells of Northern Adriatic Basin reported by Moricca et al. [10] was used for modeling. Adriatic Sea is geographically bounded by Italy, Albania, and the former Yugoslavian republics. The northern and central portions of this area are known to form one geological unit commonly referred to as the Northern Adriatic Basin [10]. The Northern Adriatic Basin is considered Italy's main supplier in natural gas. Based on extensive experimental tests on core samples as well as drilled cuttings in different fields of the Northern Adriatic Basin, the following data were obtained [10]:

1. The producing formation porosity varies between 10 and 40% while it has very low permeability ranging from 10 to 100 mD.
2. The lithology of vertical column is composed of sand shale alterations. While the producing formation contains uncemented clays. The cohesive strength averages to 19 kg/cm^2.

According to AGIP's research group [10], out of all studied wells, 23 wells considered as problematic due to sand production, and the remaining 8 wells considered as sand free. Reported main factors affecting sand production were: total vertical depth (TVD), transmit time (TT), cohesive strength of the formation (COH), water and gas flow rates (Q_w & Q_g), bottom hole flowing pressure (BHFP), drawdown pressure (DD), effective overburden stress (EOVS), shut per foot (SPF), perforation interval (H_{perf}). This dataset was examined for incomplete data and as the result two sets were removed. The finalized data set consisting of 29 data points is reported in Table 1.

As previously mentioned, our objective is to develop a mathematical non-linear relationship between the available data which considered as inputs (i.e., TVD, TT, COH, Q_w, Q_g, BHFP, DD, EOVS, SPF, and H_{perf}) and the output which is sanding indication. For this purpose, mathematical background of LSSVM is presented.

2.2. Model development

The aim of this section is a review on some basic concept on support vector machines (SVM) for classification problems. SVM can be considered as a non-probabilistic binary linear classifier using regression analysis. This algorithm that is supervised learning manner for pattern recognition and data analysis has been studied extensively for both classification and regression analysis [1,26–34]. The convergence of SVM method to global optimum and detect a quick solution by standard algorithm is very probable It has no need for multiple adjustable parameters and network topology determination in advance (it will be determined automatically at the end of training process) [35–38].

The SVM algorithm builds a separating hyper-surface in the input space. This process is performed as follows [1,35,38–40]:

1) It maps the input patterns into a higher dimensional feature space through nonlinear mapping.
2) Builds a separating hyper-plane with maximum margin.

By considering a training sample set of N data points $\{x_k, y_k\}_{k=1}^{N}$, in which $x_k \in R^n$ is the kth input pattern and $y_k \in R^n$ is the kth output pattern, the SVM algorithm is reduced to constructing a classifier of the below form for classification problems:

$$y(x) = sign\left[\sum_{k=1}^{N} \alpha_k y_k \Psi(x, x_k) + b\right] \tag{1}$$

where α_k positive real constants and b is a real constant. $\Psi(.,.)$ is a function that has some forms such as linear SVM, polynomial SVM, two layer neural SVM, and Radial Basis Function (RBF) SVM. Recent studies show that RBF SVM is hemmore reliable that other form of SVM [33,41–43].

The RBF for of this function is as follow:

$$\Psi(x, x_k) = exp\left\{-\frac{\|x - x_k\|_2^2}{\sigma^2}\right\} \tag{2}$$

where σ^2 is the squared variance of the Gaussian function which should be optimized by the user to obtain the support vector.

Construction of classifier based on SVM approach is as follow. One assumes that

$$y_k\left[w^T \varphi(x_k) + b\right] \geq 1 \quad k = 1, ..., N \tag{3}$$

$\varphi(.)$ is a nonlinear function which maps the input space into a higher dimensional space. The extension of linear SVMs to non-separable case was also made by Cortrs and Vapnik in 1995 [35].