Journal of Petroleum Science and Engineering 135 (2015) 201–205

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: www.elsevier.com/locate/petrol

Robust method based on optimized support vector regression for modeling of asphaltene precipitation

Hamid Reza Ansari *, Amin Gholami

Abadan Faculty of Petroleum Engineering, Petroleum University of Technology, Abadan, Iran

ARTICLE INFO

Article history:
Received 4 November 2014
Received in revised form
4 September 2015
Accepted 6 September 2015
Available online 14 September 2015

Keywords:
Asphaltene precipitation
Titration data
Support vector regression (SVR)
Scaling equation
Imperialist competitive algorithm (ICA)

ABSTRACT

Precipitation of asphaltene largely affects the production rate of crude oil owing to the clogging of transportation pipeline as well as formation damage. Therefore, it is imperative to search for a robust method for the modeling of asphaltene precipitation. This paper aims to propose a new mathematical model for computing asphaltene precipitation as a function of titration data including dilution ratio, temperature, and molecular weight of solvent. This model is constructed based on integrating the support vector regression with the imperialist competitive algorithm. Optimization increases the performance of support vector regression by virtue of determining the optimal values of their parameters. The constructed model is applied to experimental data and its performance is compared with scaling equation which is traditionally employed for modeling of asphaltene precipitation. The results of this study show that the integration of imperialist competitive algorithm and support vector regression has a better performance than the scaling equation. This paper offers that the support vector regression optimized by virtue of the imperialist competitive algorithm is a reliable model for modeling asphaltene precipitation.

1. Introduction

Asphaltene is regarded as the most polar and aromatic part of crude oil (Panuganti et al., 2012). It is defined as a fraction of crude oil which is soluble in toluene but precipitate when normal heptane is added (Shahebrahimi and Zonnouri, 2013). This fraction is the most significant fraction of crude oil, mainly owing to their tendency to separate from crude oil and cause destructive effects in the petroleum industry. Asphaltene is initially dissolved in crude oil due to thermodynamic equilibrium between petroleum fractions (Speight, 1999). Production of crude oil from reservoir alters the thermodynamic conditions. Asphaltene precipitation is a phenomenon at which asphaltene starts to phase-separate from crude oil and form a solid particle in response to aforementioned variation in thermodynamic conditions (Nakhli et al., 2011). The precipitated particles are adhered to each other (flocculation) and generate asphaltene flocs. The flocculated asphaltene is deposited in different media which can occur in upstream and downstream operations. In upstream operations, asphaltene is precipitated and deposited mainly due to pressure depletion during primary recovery as well as crude oil composition change during gas injection process (Nakhli et al., 2011). This phenomenon can cause considerable damages to reservoir rock quality through reducing oil permeability as well as altering the wettability of oil-bearing formation (Kord et al., 2014; Lawal et al., 2012). In downstream operations, precipitation and deposition of asphaltene would result in the clogging of the transportation pipeline and a decrease in the performance of heat exchanger (Lawal et al., 2012).

Modeling of asphaltene precipitation provides an opportunity to find answers for the questions such as when asphaltene starts to precipitate out of crude oil and how much asphaltene is separated at a specific condition. The modeling has been studied by a number of research groups and various strategies had emerged for clarification of this mechanism (Shirani et al., 2012; Mohammadi et al., 2012; Tabatabaei-Nejad and Khodapanah, 2010). Although developed approaches can model the asphaltene precipitation, the phase behavior of asphaltene is not well understood yet, mainly due to the complexity of their properties. Developed models fall into three general groups including molecular thermodynamic model (Mansoori, 1997), colloidal model (Wu et al., 1998), and scaling equation model (Rassamdana et al., 1996). In a molecular thermodynamic model, asphaltene is considered as a particle dissolved in crude oil. In a colloidal model, asphaltene is stabilized in crude oil through peptizing via resin. The interesting feature of scaling-based model is that modeling of asphaltene precipitation is fulfilled while the complicated nature of asphaltene particle and their agglomeration is disregarded. In this molecular thermodynamic model, asphaltene

* Corresponding author.
E-mail addresses: ansari.hamid.r@gmail.com (H.R. Ansari), amingholamiput@yahoo.com (A. Gholami).

http://dx.doi.org/10.1016/j.petrol.2015.09.002
0920-4105/

precipitation is deemed as a reversible process. On the contrary, in colloidal models the precipitation of asphaltene is regarded as an irreversible process.

In addition to aforementioned models, up to now different researchers have employed intelligence-based approaches for modeling the asphaltene precipitation (Ahmadi, 2011; Ahmadi, 2012; Chamkalani et al., 2013; Ashoori et al., 2010; Asoodeh et al., 2014). These studies prove the superiority of intelligence models to scaling-based models and thermodynamic models. The previous intelligence approaches modeled the asphaltene precipitation with good precision, however introducing a robust model for achieving this purpose is valuable.

This study aims to employ the optimized support vector regression (SVR) for making a quantitative formulation between asphaltene precipitation and titration data (dilution ratio, temperature, and molecular weight of solvent). Optimization of the SVR is achieved through a robust evolutionary optimization algorithm entitled as the imperialist competitive algorithm (ICA). The performance of constructed models is compared with the scaling equation usually employed for the estimation of asphaltene precipitation from titration data.

2. Scaling equation model

At 1996 Rassamdana and colleagues implemented a number of titration tests in order to clarify the behavior of asphaltene when precipitated (Rassamdana et al., 1996). They developed a simple approach so-called scaling equation for modeling of asphaltene precipitation based on this assumption that the formation of asphaltene structure is to some extent similar to the aggregation and gelation phenomena. Due to the simplicity and superb performance of this model, it attracted significant attention of researchers. Hu et al. (2000) gathered the experimental data of asphaltene precipitation from open literature source and assessed the ability of Rassamdana et al. (1996) scaling equation for the modeling of aforementioned data. They observed that the scaling equation produces acceptable results. Owing to the titration tests in Rassamdana et al. (1996) paper are carried out on constant temperature, temperature is not included on those model. Hence, a number of researchers performed the titration tests in different temperatures and included a temperature parameter in the scaling formulation (Rassamdana et al., 1999; Hu and Guo, 2001, Ashoori et al., 2010). For detailed study about the formulation of previous scaling models, readers are referred to their original papers.

3. Model description

3.1. SVR

SVR is a new supervised learning method introduced by Vapnik (Vapnik, 1995; Drucker et al. 1996). In the SVR methodology, a kernel function is used to transform a nonlinear learning problem into a linear learning problem. In order to find the best linear approximation function, it employs the structural risk minimization (SRM) principle in their formulation. The SRM controls the errors through a trade-off between model complexity and the quality of fitting the training data (empirical error) by defining two main regulation parameters of cost coefficient (C) and tubing radius of loss function (ε). Other important parameters in SVR are kernel parameters. A Gaussian radial basis function (RBF) is used as a kernel function in this study. The RBF is one of most popular and useful kernel function in application of SVR. The RBF kernel function is a mapping function whose value depends merely on the squared Euclidean distance between the two feature vectors. In summary, C, ε and γ (kernel parameter) are three main parameters of the SVR. The success of the SVR in estimating problem is dependent on the accurate selection of these parameters. Therefore, it is mandatory to use a robust method for determination of these parameters. For this purpose, the ICA is applied as an optimization tool for finding the optimal value of SVR.

3.2. Imperialist competitive algorithm

ICA was introduced by Atashpaz-Gargari and Lucas (2007) based on social political processes of imperialism. This new technique revealed great performance in both achievement of global optima and convergence in several study (Ansari, 2014; Atashpaz-Gargari and Lucas, 2007). Each population in the ICA consists of some countries (solutions) which are divided into two groups: *colonies and imperialists.* Each imperialist and its relative colonies generate an empire. The goal is to create a competition among empires in order to take possession more colonies. The power of countries is evaluated by a *cost* function.

The main operations of ICA are *assimilation*, *revolution* and *competition*. In the ICA terminology, assimilation is a movement of colonies toward the imperialist. The new position of the colonies is selected by a random variable with a uniform distribution between 0 and $\beta \times d$ where d is the distance between a colony and imperialist, and β is a number greater than 1 to create the colonies to get closer to the imperialist. For improving the performance of the algorithm in searching more points around imperialist, a random deviation number θ is selected with a uniform distribution in range of $[-\lambda, \lambda]$. λ is a parameter that sets the deviation from the original direction (Atashpaz-Gargari and Lucas, 2007).

A sudden change in the position of colonies is called revolution. The revolution enhances the performance of the algorithm in exploration and prevents the early convergences into local minimums. The percentage of colonies in each empire that contributes in revolution is selected by a *revolution rate*. An imperialist competition is performed based on the total cost of the empires which is defined by a positive number (ξ) that controls the percentage of contribution of the colonies costs in the total cost. Finally, the ICA is repeated in several *decades* (iterations) until a stop condition is satisfied.

4. Input/output data space

The dataset provided in Hu and Guo paper (2001) is employed for building and assessing the desired model. Based on the scaling equation, amount of asphaltene precipitation can be modeled as a function of titration data including dilution ratio, molecular weight of solvent, and temperature. Therefore titration data are considered as inputs of the optimized model for estimation of amount of asphaltene precipitation. According to Hu and Guo paper (2001), the amount of asphaltene precipitation is measured by adding normal alkane (normal heptan) in different dilution ratio (150) followed by separating by means of a Whatman filter paper No. 42. The crude oil is chosen from Shengli Oil Field in China to carry out the experiments. Experimental implementation is performed in four temperature (293 °k, 308 °k, 323 °k, and 338 °k) in which the pressure of crude oil is kept as a constant value (atmospheric pressure). Composition of crude oil is tabulated in Table 1. Statistical description of data employed in this study is given in Table 2. All data are normalized in range of [0 1]. The normalization task makes dimensionless variable; thus, reduces confusion of the model. In the next step, all data were broken into five subsets, randomly. One of them is used as test data set and other data are used in 4-fold cross validation in training step of SVR model. The advantage of this procedure is that