Petroleum xxx (2016) 1–6

Contents lists available at ScienceDirect

Petroleum

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

Original article

Prediction of asphaltene precipitation using support vector regression tuned with genetic algorithms

Mohammad Ghorbani [a], Ghasem Zargar [a, *], Hooshang Jazayeri-Rad [b]

[a] Department of Petroleum Engineering, Petroleum University of Technology (PUT), Ahwaz, Iran
[b] Department of Instrumentation and Automation, Petroleum University of Technology (PUT), Ahwaz, Iran

ARTICLE INFO

Article history:
Received 15 November 2015
Received in revised form
13 May 2016
Accepted 17 May 2016

Keywords:
Asphaltene precipitation
Prediction
Support Vector Regression (SVR)
Genetic Algorithm (GA)
Parameter optimization

ABSTRACT

Due to the severe and costly problems caused by asphaltene precipitation in petroleum industry, developing a quick and accurate model, to predict the asphaltene precipitation under different conditions, seems crucial. In this study, a new model, namely genetic algorithm – support vector regression (GA-SVR) is proposed, which is applied to predict the amount of asphaltene precipitation. GA is used to select the best optimal values of SVR parameters and kernel parameter, simultaneously, to increase the generalization performance of the SVR. The GA-SVR model is trained and tested on the experimental data sets reported in literature. The performance of the GA-SVR model is compared with two scaling equation models, using statistical error measures and graphical analyses. The results show that the prediction performance of the proposed model, is highly reliable and satisfactory.

Copyright © 2016, Southwest Petroleum University. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

According to the SARA separation technique, Asphaltenes constitute the most aromatic and polar portion of petroleum. Other fractions are Saturates, Aromatics and Resins which have higher molecular weight and aromatic content respectively [1–3]. In general, asphaltenes are defined as a portion of crude oil that can be solved in some aromatic solvents such as toluene, benzene, xylene and pyridine but are insoluble in normal alkane solvents like n-pentane, n-heptane and n-decane [4].

Asphaltene precipitates from crude oil as a result of changes in oil composition, pressure and temperature, however the latter has fewer effect relative to the others. Miscible flooding processes like carbon dioxide and natural gas injection, and microbial enhanced oil recoveries can lead to asphaltene precipitation in petroleum reservoirs by modifying aforementioned parameters [5–7]. This precipitation also may occur in production and refinery facilities [8,9]. Severe problems like wettability alteration, relative permeability reduction in reservoirs and damage and blockage of flow in boreholes and surface pipelines can be results of asphaltene precipitation [10].

So, many researches and studies have been conducted to predict the amount of asphaltene precipitation in both theoretical and experimental approaches. In general, models of asphaltene precipitation can be divide into four categories: 1. Molecular thermodynamic models, which are an evolved type of two statistical and continuous thermodynamic models [11]. The basis of this type is assumption of polymeric structure of asphaltene molecules [12,13]. 2. Colloidal models, which are developed from the previous models, suppose that solubility of asphaltene particles is due to the attachment of resin molecules to their surfaces [14,15]. 3. Scaling equation models, which are simple and user-friendly correlations and are chiefly based on the experimental studies rather than theoretical investigations [16,17]. 4. Artificial Intelligence (AI) based models, which have been developed and are widely used recently based on finding the relation between the value of asphaltene precipitation and

* Corresponding author.
E-mail addresses: ghorbani.u@live.com (M. Ghorbani), zargar@put.ac.ir (G. Zargar).

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.2016.05.006

2405-6561/Copyright © 2016, Southwest Petroleum University. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article in press as: M. Ghorbani, et al., Prediction of asphaltene precipitation using support vector regression tuned with genetic algorithms, Petroleum (2016), http://dx.doi.org/10.1016/j.petlm.2016.05.006

effective parameters without regarding the properties of the chemical process [18,19].

Due to the nature of asphaltene and current vagueness in its phase behavior and also because of diversity of effective parameters causing precipitation, a perfect and robust model cannot be introduced. This problem motivate us to introduce a model without the requirement of a thorough knowledge of nature of asphaltene. The proposed model is just based on experimental data and has a simple and fast applying procedure.

2. Development of the GA-SVR model

2.1. Data acquisition

To obtain a reliable and useful model, generality and accuracy of the method of gathering dataset are essential. So the data which are used in this study, are tried to be verified and accurate [20]. Two types of data have been implemented to develop and train the model. The first dataset, are chosen from Ashoori et al. work [18]. The selected crude oil is an asphaltenic crude oil with a specific gravity of 0.934. The amount of asphaltene precipitation is determined by gravimetric method. Three precipitants (pentane, hexane and heptane) with various dilution ratios at three 30, 50 and 70 °C temperatures are used. All of the experiments are done at the atmosphere pressure [18].

Second dataset are chosen from the study which is performed by Hu and Guo [21]. The oil, under study is the Caoqiao crude oil from an Oil Field in China. Seven n-alkanes (pentane, hexane, heptane, octane, nonane, decane, and dodecane) are used as precipitants in four temperatures including, 19.85, 34.85, 49.85 and 64.85 °C [21].

The input datasets of the proposed model include dilution ratio (Injected n-alkane volume to weight of crude oil), molecular weight of n-alkane and temperature while the corresponding amount of asphaltene precipitation (weight percent) has been considered as the target data of the model. Ranges and average of parameters of both datasets are summarized in Table 1.

2.2. Basic idea of Support Vector Regression

Support Vector Machine (SVM) is a machine learning technique based on the statistical learning theory which is proposed by Vapnic et al. in 1995 [22]. Originally, SVM developed for solving the classification problems but latter, SVR evolved from the SVM for doing regression tasks. Therefore SVM is a general term which can be divided into two subgroups; Support Vector Classification (SVC) and Support Vector Regression. In this study, the latter will be used. Assume that there is a learning samples set, $D = \{(x_i, y_i)\}$ where $x_i \in \mathbb{R}^m$ represent the input values and $y_i \in \mathbb{R}$ are the corresponding output values for $i = 1, 2, \ldots, N$ where N is the number of the samples in the training dataset and m is the dimension of the input dataset. The SVR function is generally defined as follows:

$$f(x) = \langle w, \Phi(x) \rangle + b \tag{1}$$

where $w \in \mathbb{R}^m$ is the weight vector and $b \in \mathbb{R}$ is the threshold and Φ is the mapping function which transfers the input values from a $\mathbb{R}^m$ space to a feature space with higher dimension. To train the SVR, the values of w and b must be found which are done by minimizing the following regularized risk function:

$$R_{reg}(f) = \frac{1}{2}\|w\|^2 + \frac{C}{N}\sum_{i=1}^{N} L(f(x_i), y_i) \tag{2}$$

where L is the error loss function. The term $\frac{1}{N}\sum_{i=1}^{N} L(f(x_i), y_i)$ is the average loss over the training samples, i.e. it represents the empirical risk, and $\frac{1}{2}\|w\|^2$ is the regularization term. C is the error penalty parameter which defines a trade-off between an approximation error and the weights vector norm $\|w\|$, and is chosen by the user.

According to the structural risk minimization principle, minimization of the above regularized risk function leads to the following Quadratic Programming (QP) problem:

$$\min_{w,\xi,\xi^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{3}$$

subject to:

$$y_i - \langle w, \Phi(x_i) \rangle - b \le \varepsilon + \xi, \quad i = 1, \ldots, N$$

$$\langle w, \Phi(x_i) \rangle + b - y_i \le \varepsilon + \xi^*, \quad i = 1, \ldots, N$$

$$\xi \ge 0, \quad i = 1, \ldots, N$$

$$\xi^* \ge 0, \quad i = 1, \ldots, N$$

where ξ and ξ^* are slack variables.

After converting the above problem to a dual Lagrangian problem and solving the dual problem, the regression function is written as:

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\langle \Phi(x_i), \Phi(x) \rangle + b \tag{4}$$

The dot product in the above equation, can be replaced by kernel function $k(x_i, x)$, which is a useful trick to avoid the difficulties encountered with nonlinear mapping into a higher dimension space. This replacement allows to restate the support vector regression function as:

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)k(x_i, x) + b \tag{5}$$

There are various types of kernel functions, from which more common ones are listed in Table 2. Because of the acceptable

Table 1
The ranges of the data are used in the GA-SVR model.

Type	Parameter	Dataset					
		Ashoori et al. work			Hu and Guo work		
		Min	Max	Ave	Min	Max	Ave
Inputs	Dilution ratio, mL/g	0.67	20	7.62	2.3	24.3	12.08
	Temperature, °C	30	70	50	19.85	64.85	39.28
	Molecular weight	72.15	100.2	86.17	72.15	170.33	116.14
Output	Asphaltene precipitation, %	0.5	10.4	4.78	0.12	7.06	2.96

Table 2
Kernel functions.

Kernel function	Formula
Polynomial	$k(x_i, x) = (\langle x_i, x \rangle + 1)^d$
Gaussian Radial Basis	$k(x_i, x) = exp\left(-\frac{\|x_i - x\|^2}{2\gamma^2}\right)$
Sigmoid	$k(x_i, x) = \tanh(\rho\langle x_i, x \rangle + \varrho)$

Please cite this article in press as: M. Ghorbani, et al., Prediction of asphaltene precipitation using support vector regression tuned with genetic algorithms, Petroleum (2016), http://dx.doi.org/10.1016/j.petlm.2016.05.006

Download English Version:

https://daneshyari.com/en/article/5026553

Download Persian Version:

https://daneshyari.com/article/5026553

Daneshyari.com