Journal of Petroleum Science and Engineering 112 (2013) 97–104

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

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

Modeling of vaporization enthalpies of petroleum fractions and pure hydrocarbons using genetic programming

Hossein Parhizgar, Mohammad Reza Dehghani*, Azin Eftekhari

Thermodynamics Research Laboratory, School of Chemical Engineering, Iran University of Science & Technology, Narmak, Tehran 16846-13114, Iran

ARTICLE INFO

Article history:
Received 26 March 2012
Accepted 17 October 2013
Available online 28 October 2013

Keywords:
vaporization enthalpy
petroleum fraction
hydrocarbon
genetic programming

ABSTRACT

In this work utilizing genetic programming, a new model has been presented for prediction of vaporization enthalpy of petroleum fractions as well as pure hydrocarbons. Using genetic programming enabled us to find a new relation between vaporization enthalpy and thermo-physical properties such as specific gravity, molecular weight and boiling point temperature. The results have been compared with previous models. It has been shown that our model can give better results compared to previous ones. In order to check the predictability of the model, vaporization enthalpy of pure hydrocarbons has been excluded from the learning set, it is worth mentioning that finally the presented model predicted them efficiently, this result shows that it can be utilized for prediction in the case of new hydrocarbon fractions.

1. Introduction

Enthalpy of vaporization is simply defined as the heat required for transforming a quantity of liquid substance at its boiling point temperature into gas phase. Enthalpy of vaporization is one of the most important thermo-physical properties in engineering applications, such as equipment design. Estimation of vaporization enthalpy has been paid more attention during recent years (Alessandro, 1979; Riazi and Daubert, 1980;Alessandro, 1995; Gopinathan and Saraf, 2001; Fang et al., 2003; Sergey, 2006; Strechan et al., 2006; Mohammadi and Richon, 2007 ; Cachadiña and Mulero, 2009). Enthalpy of vaporization data are of interest at different temperatures especially for petroleum fractions. In this regard, at least at one temperature for calculation of liquid and vapor enthalpy using isobaric heat capacity, Cp, must be known. Heat of vaporization at normal boiling point is often used for this purpose. Enthalpy of vaporization successfully has been modeled by various computational methods which can be found in the literature (Alessandro, 1979; Riazi and Daubert, 1980; Alessandro, 1995; Gopinathan and Saraf, 2001; Fang et al., 2003; Sergey, 2006; Strechan et al., 2006; Mohammadi and Richon, 2007; Cachadiña and Mulero, 2009). Firstly Trouton found that the entropy of vaporization relatively is a constant value for the most of nonpolar and nonassociating components, however this rule fails in the case of polar components like water. Unfortunately this simple rule has not enough accuracy to be applied for the most of the materials as well as petroleum fractions. The Clausius–Clapeyron equation has been utilized by Kistiakowsky for determination of the temperature dependency of the enthalpy of vaporization (Fang et al., 2003). Using Clausius–Clapeyron equation he could modify Trouton's rule (Fang et al., 2003). Alessandro (1979, 1995) presented two correlation as a function of normal boiling temperature and molecular weight for hydrocarbons. A simple correlation as a function of specific gravity and boiling point was presented by Riazi and Daubert (1980). However the results of their correlation for calculation of vaporization enthalpy of petroleum fractions show that its error is around 7% like Trouton and Kistiakowsky methods. Gopinathan and Saraf (2001), Fang et al. (2003) and Mohammadi and Richon (2007) published some similar works in the continuation of previous works. It is worth mentioning that in most of the previous works, vaporization enthalpies have been correlated as a function of boiling point, molecular weight, and specific gravity. Lyman et al. (1990) presented a good review regarding available models and their applications (A review on previous models have been presented in Table 1). Recently artificial neural network (ANN) as powerful tools has been utilized for prediction of the thermophysical properties. Mohammadi and Richon (2007) used ANN approach and showed that they could reduce the error of prediction up to 2%. Unfortunately they did not report any data about the applied ANN. ANN comprises of a number of complex relations and without network parameters; like biases and weights it is not possible to use it.

Alternatively, genetic programming (GP) has been already used in chemical engineering such as determination of kinetic orders, optimization of complex distillation systems (McKay et al., 1997; Cao et al., 1999; Wang et al., 2008). As an example Shokir (Shokir,

* Corresponding author. Tel.: +98 217 724 0496; fax: +98 217 724 0495.
E-mail address: m_dehghani@iust.ac.ir (M.R. Dehghani).

0920-4105/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.petrol.2013.10.012

Nomenclature

AARD%	average absolute relative deviations
ARD%	absolute relative deviations
ANN	artificial neural network
Cp	Isobaric heat capacity
FS	function set
GP	genetic programming
GEP	gene expression programming
LGP	linear genetic programming
M	molecular weight
MGEP	multi gene expression programming
N	number of experimental data
S	specific gravity
T	temperature
TS	terminals set

Greek symbols

ΔH	vaporization enthalpy

Subscripts

b	normal boiling point

Superscript

Vap	vaporization

2008; Shokir and Dmour, 2009) used genetic programming approach in combination with orthogonal least squares algorithm (GP-OLS) to estimate dew point pressure of gas condensate reservoirs. In another work he used GP-OLS to estimate viscosity of pure and hydrocarbon gas mixtures. In a similar work AlQuraishi (2009) utilized a linear genetic programming to determinate saturation pressure of crude oil. Their results showed that their approach was successful in the estimation of dew point, viscosity and saturation pressure.

In the present work we have utilized genetic programming procedure for determining a new correlation for estimation of vaporization enthalpies of petroleum fractions and pure hydrocarbons. Our new model has been constructed as a function of normal boiling point temperatures (T_b) and specific gravity (S). In the first section the methodology of this work has been presented, finally the results have been presented and compared with previous works.

2. Genetic programming

2.1. Theory

Genetic programming is accounted as a subclass of genetic algorithms and a member of the evolutionary algorithms (Koza, 1992). GP has been developed in an evolutionary algorithm by symbolic optimization (Koza, 1992; Banzhaf, 1998; Langdon and Poli, 2002; Muttil and Lee, 2005). It is based on the tree data structure. Because of its flexibility this structure can reproduce a mathematical equation for prediction of thermo-physical properties. This structure contains the function set and terminal set. The function set can be comprised of the arithmetic operator {−, +, *, /} or even it may also consist of conditional statements ("if…", "then…"), Boolean functions, trigonometric functions, iterative functions, recursive functions and other mathematical functions. The terminals set contain independent variables, numerical constants and other zero-argument in the GP tree data structure. Calculations will start when an initial population is considered for the GP. Each individual of population is represented as a tree data structure. The procedure is similar to the genetic algorithm method. In next step, the fitness of each individual is evaluated. After evaluation of individuals, a new population will generate based on the Darwinian principles such as reproduction, cross over and mutation in current population. Rank and tournament functions are the two usual approaches for reproduction. Selection, replacement, mutation and crossover are the most important genetic operators which are used for the production of new generation with minimum error. Mutation is used to increase the diversity of the expressions and avoid early convergence (Seavey et al., 2010). Crossover transforms structures in the population and exchanges sequences of instructions between two tournament winners (Gandomi et al., 2010). Replacement operator kills old population with maximum error in order to create new individuals. After each generation some part of old generation will survive in new generation. The iterative process will continue to guarantee the effective performance (Gathercole and Ross, 1994; Chakraborty et al., 1996; Madar et al., 2004). We refer the interested reader to the following References: Koza (1992), Banzhaf (1998) and Langdon and Poli (2002).

2.2. Modeling of vaporization enthalpies

In this work, the GP technique has been applied to generate a new equation for the prediction of vaporization enthalpies of petroleum fractions as well as pure hydrocarbons. In Section 1 a complete review regarding different correlations has been presented. Our literature survey shows that in the most of previous works a nonlinear function of normal boiling point temperatures (T_b), the specific gravity (S) at 288.7 K and molecular weight (M) has been considered. As measurement of molecular weight in the case of petroleum cuts is not easy, significant uncertainties are encountered. Due to this, we considered specific gravity and normal boiling point as input variables for genetic programming. In order to start the learning procedure a comprehensive database is needed. In this regard 122 data points (58 data point for petroleum fractions and remaining data point for pure hydrocarbon) have been collected utilizing available literature. Petroleum fractions data are collected using a diversity of literature sources. Data about petroleum fractions are related to four different crude oils (one from Russia, one from Iran, and two from China). Pure hydrocarbon data are collected from the following References: Alessandro (1995), Gopinathan and Saraf (2001), Fang et al. (2003) and Sergey, 2006). In order to check the predictability of the GP model, the experimental data of petroleum fractions were divided into learning and validating sets. The experimental data of pure hydrocarbon have been put to the side for testing the vaporization Enthalpies model. It should be mentioned that the learning and validation data sets are grouped into one set named as training set (Gandomi and Alavi, 2011). In order to prevent the over-fitting phenomenon, the error of learning and validating set in each generation was checked during the modeling procedure. The best model was selected when the error of validation set and learning set was minimized simultaneously.