Fluid Phase Equilibria 390 (2015) 14–22

Contents lists available at ScienceDirect

Fluid Phase Equilibria

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

Computer aided molecular design for undefined petroleum fractions

José M. Valdez-Patrinos, Richart Vázquez-Román [*], Erika Morán-Corona,
José A. Inchaurregui-Méndez, Efraín Quiroz-Pérez

Instituto Tecnológico de Celaya, Departamento de Ingeniería Química, Av. Tecnológico y A.G. Cubas s/n, Celaya 38010, Gto., Mexico

ARTICLE INFO

Article history:
Received 6 October 2014
Received in revised form 12 January 2015
Accepted 13 January 2015
Available online 14 January 2015

Keywords:
Petroleum fraction
Pseudocomponents
CAMD
Cubic equations of state

ABSTRACT

Undefined petroleum fractions typically composed of several components are characterized through a reduced set of components representing the whole mixture. The selected set should appropriately reproduce experimental phase equilibria behavior such as TBP curves as well as mass contents of elementary species such as amounts of carbon and hydrogen. Three approaches have been developed in this work. In the first approach, a predefined number of components are selected from a given database and a numerical evaluation selects the characterizing components by fitting experimental and calculated properties. A synthesis of a predefined number of components is carried out in the second approach using a mathematical programming model for the computer aided molecular design (CAMD). In the third approach, parameters such as critical properties and molecular weight are freely estimated so that characterizing components are not necessarily identifiable or physically existing components; i.e., hypothetical components would be used to characterize petroleum fractions. The Peng–Robinson equation, a typical cubic equation of state, is used to describe the complex phase behavior in the three proposed approaches. The numerical evidence indicates that undefined petroleum fractions can be highly represented in phase equilibria calculations by a properly selected set of components.

1. Introduction

Petroleum is one of the most important substances consumed at present time. It contains different hydrocarbon compounds typically divided into four groups: paraffins, olefins, naphtenes and aromatics. An appropriate petroleum characterization is essential for estimating thermodynamic properties and eventually simulating and designing process units for several processes. A large amount of components are contained in petroleum fractions and the characterizing procedures mainly involve characterization of the so called hydrocarbon-plus fractions, generally expressed as C_{7+} fractions. The C_{7+} fraction has a very wide boiling range. In general, there are two ways to characterize hydrocarbon plus fractions: pseudocomponents or the continuous thermodynamics approach [1,2]. In the pseudocomponents approach, the plus fraction is represented by one or more subfractions or components with known or estimated thermodynamic properties. In the second approach, a probability distribution function is used to describe the composition of the mixture. In particular, the gamma distribution is used to characterize the hydrocarbon mixture by modeling the molecular weight mixture [3]. Thus, crude oil is often characterized by a set of identifiable components and a set of pseudocomponents or a distribution function.

The problem of characterizing petroleum fractions can be seen as finding the molecule or molecules capable to reproduce particular properties of the fraction. An example of such a view has been presented by Marrero and Gani for pure component property estimation based on group-contribution in a molecule [4]. In the group contribution method for pure component property prediction, a molecule is described as a set of groups linked together to end up with a molecular structure. The pure component property estimation requires knowledge of the molecular structure and the groups needed to uniquely represent the molecule. Therefore, having the groups and their contributions together with a set of rules to combine the groups to represent any molecule provide the possibility to “model” the molecule or even a molecular mixture. The needed property is estimated from a set of a priori regressed contributions for the groups representing the molecule. This means that the problem of property estimation via synthesis/design of molecules having desired properties is solved by generating chemically feasible molecular structures and testing for their properties. The problem is known as computer aided molecular design (CAMD) [5]. CAMD has provided means for determining molecules having a desirable set of physicochemical properties [6,7]. Among other applications, a systematic CAMD framework was developed for designing solvents for a reaction

---

[*] Corresponding author. Tel.: +52 461 6117575; fax: +52 461 61 17744.
E-mail address: richart@iqcelaya.itc.mx (R. Vázquez-Román).

http://dx.doi.org/10.1016/j.fluid.2015.01.017
0378-3812/

system to enhance the yield by product extraction using solvent characteristics such as biocompatibility, inertness and ability to cause phase splitting [8]; a large number of CAMD applications on refrigerant design have been proposed to design polymers, drugs, cleaning agents, etc. [9]. More recently, a CAMD tool has been designed to find mixtures with molecular fragments issued from renewable raw materials, able to handle environment-health and safety related properties along with process physicochemical properties [10].

A crude distillation system can be simulated using, for instance, Aspen Plus but an appropriate compositional description for the feed is required. Simulation studies carried out by More et al. [11] ratify that the quality of the solution is difficult to judge without a good crude characterization. Experimental data related to each particular petroleum fraction are always required. True boiling point distillation (TBP) curves have been widely used to characterize crude oils and detect the cumulative mass fraction distilled vs. boiling temperature [12]. Experimental TBP curves were obtained following standards such as ASTM D 2892 or ASTM D 5236 and it is becoming normal to use ASTM D-1160, ASTM D-86 and Hempel distillation. In addition, it has been demonstrated that TBP curves can be acceptably predicted using kinematic viscosity [13]. Another important data in hydrocarbons characterization refers to the mass of elementary species contained in the petroleum fraction. Recent methods such as nuclear magnetic resonance or gel permeation chromatography could be used to characterize crude oils so that elemental composition like carbon and hydrogen percent can be easily determined [14]. Near infrared (NIR) spectroscopy could also be used to estimate TBP curves and API gravity for crude oil [15,16]. The proportion of paraffinic, olefins, naphthenic and aromatic and elementary species has been typically estimated based on correlations [17,18]. Finally, PVT data could also be used in the characterization procedure. A set of experimental PVT and gas injection data for pseudocomponents can be found in [19].

An earlier attempt to estimate properties of petroleum fractions without known composition based on group contribution methods have been developed by Carreón-Calderón et al. [20,21]. They formulate a model based on minimizing a Gibbs function using the IMSL Fortran numerical libraries to solve the optimization problem. Once a molecular structure is determined and described by groups, critical properties and other properties are estimated from group contribution methods. An extensive review on estimation methods of these physical properties is available in earlier editions of the classic book “The properties of gases and liquids” [22].

More recently, the perturbed-chain statistical associating fluid theory (PC-SAFT) equation of state has been used to characterize petroleum fluids [23,24]. Rather than using critical properties, PC-SAFT represents each component through 3 parameters: its number of segments (m), its segment diameter (σ) and its segment energy (ε) representing the maximum attraction between two segments. To adjust PC-SAFT parameters, experimental data used has included saturation point, oil densities, oil formation factors, PVT data and swelling data. In general, parameters for PC-SAFT cannot be directly measured since they are related to the molecular structure and interaction between molecular segments.

When using the pseudocomponents approach, each pseudocomponent can be considered as a representative molecule. A higher number of pseudocomponents would necessarily result in more accurate results. However, increasing the number of pseudocomponents increases the number of variables and it complicates calculations such as those in phase equilibria. Three approaches to characterize petroleum fractions are presented in this work.

2. Problem statement

The pseudocomponents approach is used in this work to characterize a petroleum fluid. It is assumed that experimental TBP curves would be available to achieve the characterization but additional information such as composition of elementary species or PVT data could also be available. The petroleum fraction would then be characterized by a certain number of pseudocomponents and each one will be represented by a set of required parameters. Typical involved calculations are the phase equilibria of bubble point and P–T equilibria. A cubic EOS is used to estimate required thermodynamic properties for any mixture or individual components. Hence parameters for each component such as critical properties, acentric factors, ω, and interaction parameters are required. In general, these parameters are well established, often based on experimental measures, for defined components. Dealing with petroleum fractions presents extra difficulties and the first one starts by determining what components should be included in the mixture. It is clear that the real number of components is extremely high and demands to much computational capability to properly estimate the thermodynamic properties.

The problem of fitting experimental data through assigning values to a set of parameters in a model is typically formulated as a least squares minimization problem. Thus the objective function can be established here as:

$$\min(T_\mathrm{b}-T_\mathrm{b}^{\mathrm{exp}})^2\mu_\mathrm{b}+\sum_j(\psi_{\mathrm{TBP}_j}-\psi_{\mathrm{TBP}_j}^{\mathrm{exp}})^2\mu_{\mathrm{TBP}} + \sum_s(m_s-m_s^{\mathrm{exp}})^2\mu_s \qquad (1)$$

where T_b refers to bubble point temperature, ψ_{TBP_j} is the vaporized mass fraction at the j-boiling temperature taken from the j-TBP point, m_s is the mass of elementary species s,μ refers to Lagrange factors to make terms additive, exp indicates experimental value whereas no indication means calculated value. Any other available experimental property such as PVT could be added in a similar way to the above objective function.

Calculations for Eq. (1) start with the bubble point. A conventional way to calculate this point is by solving the following system of equations:

$$\sum_i y_{b_i}=1 \qquad (2)$$

$$\hat{f}_{\mathrm{b}_i}\mathrm{L}(T_\mathrm{b},P,\mathbf{z})=\hat{f}_{\mathrm{b}_i}\mathrm{L}(T_\mathrm{b},P,\mathbf{y}_\mathrm{b}) \qquad (3)$$

where P is pressure, T is temperature, y_{b_i} is the molar fraction of i-component in the vapor phase at the bubble point, $\mathbf{z}$ is the vector of compositions in the overall mixture, $\mathbf{y}$ is the vector of compositions in vapor phase, and $\hat{f}$ is fugacity where super indexes L refers to liquid and V to vapor and subindex b indicates the bubble point.

Matching experimental TBP demands phase equilibria calculation. In this case, the Gibbs energy function must be minimized for given pressure and temperature. In a single phase, the Gibbs function G is,

$$\frac{nG}{RT}=\frac{1}{RT}\sum_i\tau_i(T)n_i+\sum_i n_i\ln\hat{f}_i \qquad (4)$$

where n is total molar mass, n_i is molar mass of i-component, and τ_i is an integration parameter.

It should be observed that Eq. (4) cannot be evaluated due to unknown values for τ_i. When evaluating two phases, say liquid and vapor, the Gibbs function becomes: