Journal of Petroleum Science and Engineering 174 (2019) 495–513

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

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

EOS lumping optimization using a genetic algorithm and a tabu search

A. Hoffmann

Tieto, Paris, France

ARTICLE INFO

Keywords:
Equation of state
Lumping
Genetic algorithm
Tabu search algorithm
Optimization

ABSTRACT

To reduce CPU time in compositional petroleum simulation models (e.g. compositional reservoir simulations), a minimum number of components should be used in the equation of state (EOS) to describe the fluid phase and volumetric behavior. A ″detailed" EOS model often contains from 20 to 40 components, with the first 10 components representing pure compounds and the remaining components represent a split of the heavier C_{6+} material in single-carbon-number (SCN) fractions. A ″pseudoized" (or lumped) EOS model might contain only 6–9 lumped components. The selection of which components to lump together is difficult because of the huge number of possible combinations.

This paper describes an automated method to find the best pseudoized EOS model based on an initial detailed SCN EOS model. The method is based on (1) a fitness function quantifying the quality of match between a pseudoized EOS model and the detailed SCN EOS model from which it is derived, (2) a genetic algorithm used to obtain a first initial solution and (3) a *tabu* search to refine this initial solution and find the optimal lumping scheme. The method allows for a set of constraints to be imposed on the lumping of components, such as (1) not lumping certain components (e.g. CO_2), (2) forcing lumping of some components (e.g. i– C_4 and n– C_4).

The proposed procedure was successfully able to find the optimal lumped EOS from a detailed SCN EOS with 34 components for three scenarios with different number of components in the lumped EOS (15, 9 and 6). The runtime is ranging from 10 to 45 min.

1. Introduction

1.1. Equation of state

Cubic Equations of State (EOS's) are simple equations relating pressure, volume and temperature. They accurately describe the volumetric and phase behavior of pure compounds and mixtures. Volumetric behavior is calculated by solving a simple cubic equation, usually expressed in terms of $Z = pv/nRT$:

$$Z^3 + A_2 \cdot Z^2 + A_1 \cdot Z + A_0 = 0, \quad (1)$$

where constants A_0, A_1 and A_2 are functions of pressure, temperature, mixture composition, component properties (critical properties and acentric factor) and binary interaction parameters (BIP's). Phase equilibria are calculated by satisfying the condition of chemical equilibrium. For a two-phase system, the chemical potential of each component in the vapor phase μ_{iV} must equal the chemical potential of each component in the liquid phase μ_{iL}, meaning that:

$$\mu_{iL}(x, p, T, v) = \mu_{iV}(y, p, T, v). \quad (2)$$

The most famous EOS used in the oil & gas industry are the Soave-Redlick-Kwong (SRK) EOS (Soave, 1972) and the Peng-Robinson (PR) EOS (Peng and Robinson, 1976).

1.2. EOS fluid characterization

Reservoir fluid is usually sampled at perforation level (bottomhole sample) or at surface (separator sample) and analyzed in a specialized laboratory. Several lab measurements are performed on the sample: chromatography to measure the sample composition (molar fraction of pure components and isomers) and depletion experiments (e.g. Constant Composition Depletion, Constant Volume Depletion) which simulate the reservoir depletion process and measure some gas and oil properties at each stage of the depletion. Some of the component properties (usually boiling point and specific gravity) are measured, but it is very expensive to measure all properties (e.g. critical properties and acentric factor). These properties are usually estimated from correlations, see for example Riazi and Daubert (1980) or Souahi and Kaabeche (2008).

The lab measurements are used to build a detailed SCN EOS model. Some key parameters of the EOS model (e.g. BIP's) are tuned to predict the lab measurements with a good degree of accuracy.

E-mail address: arnaud.hoffmann@tieto.com.

https://doi.org/10.1016/j.petrol.2018.11.021
Received 13 June 2018; Received in revised form 29 September 2018; Accepted 11 November 2018
Available online 23 November 2018
0920-4105/

1.3. EOS lumping

To reduce CPU time in compositional petroleum simulation models (e.g. compositional reservoir simulations), a minimum number of components should be used in the equation of state (EOS) to describe the fluid phase and volumetric behavior. A "detailed" EOS model often contains from 20 to 40 components, with the first 10 components representing pure compounds H_2S, CO_2, N_2, C_1, C_2, C_3, i− C_4, n− C_4, i− C_5, and n− C_5. The remaining components represent a split of the heavier C_{6+} material in single-carbon-number (SCN) fractions such as C_6, C_7, C_8 and C_9, or groups of SCN fractions such as C_{10-12}, C_{13-19}, C_{20-29}, and C_{30+}. Occasionally the light aromatics BTX (benzene, toluene, and xylene isomers) are also kept as separate components for process modeling. Today's typical laboratory compositional analysis provides 50–60 components, including isomers with carbon numbers 6 to 10, SCN fractions out to C_{35} and a residual C_{36+}. This is in contrast to the 11–12 components (through C_{7+}) reported in most commercial laboratory reports pre-1980.

A pseudoized (or lumped) EOS model might contain only 6–9 lumped components e.g. lumping similar components such N_2 and C_1, i− C_4, n− C_4, i− C_5, and n− C_5, and some 3–5 C_{6+} fractions. The selection of which components to lump together is difficult because of the huge number of possible combinations.

Let us assume that a detailed SCN EOS model, EOS*xx*, consisting of N_{xx} components is available and represents with a satisfying degree of accuracy the measured lab data. For some reasons, a lumped EOS, EOS*x*, is required with N_x components.

Creating a lumped EOS with N_x components from a detailed EOS with N_{xx} components consists of choosing where to place $N_x - 1$ *separators* in the $N_{xx} - 1$ possible separation locations, see Figs. 1 and 2. Thus, the number of possible lumping scenarios $N_{\text{scenarios}}$ is given by:

$$N_{\text{scenarios}} = \binom{N_{xx}-1}{N_x-1} = \frac{(N_{xx}-1)!}{(N_x-1)!(N_{xx}-N_x)!}. \tag{3}$$

For example, with $N_{xx} = 34$ and $N_x = 9$, there are $N_{\text{scenarios}} = 13{,}884{,}156$ possible lumping schemes.

The obvious application of pseudoized EOS models is compositional reservoir simulation, where run time is an important issue and fewer components may be important. Lumped EOS are also widely used in well and network simulators for run time reduction. Alavian et al. (2014) note that there might numerous optimal lumped EOS models, each optimal for a particular application, e.g. one for reservoir, another for flow assurance, and yet another for processing facility. The originating detailed SCN EOS model used to develop the different lumped EOS models is likely to be the same, but the number of pseudo-components might vary for each application, together with the sub-space of (*p-T-z*).

On the one hand, lumped EOS enables to lower considerably the runtime of some discipline specific applications (e.g. reservoir simulator). On the other hand, the lumping scheme can have major impact on the accuracy of the EOS predictions, see Fig. 3.This is why finding the optimal EOS lumping scheme is important. For the same number of pseudo-components (i.e. the same runtime in engineering software applications), fluid behavior predictions can be significantly different, see Fig. 3.

Lee et al. (1982) suggest that C_{7+} fractions can be grouped into two pseudo-components according to a characterization factor determined

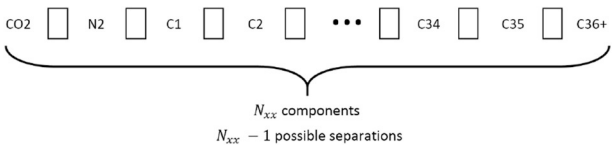

Fig. 1. Example of a detailed EOS with N_{xx} components.

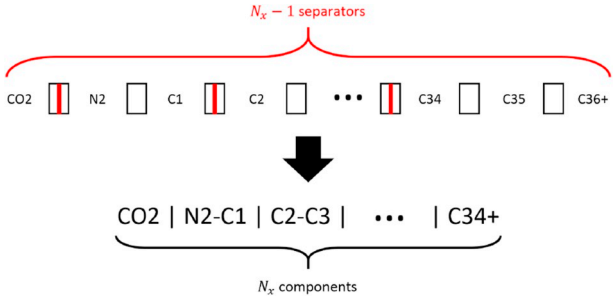

Fig. 2. Example of a lumping scheme from an EOS with N_{xx} components into an EOS with N_x components.

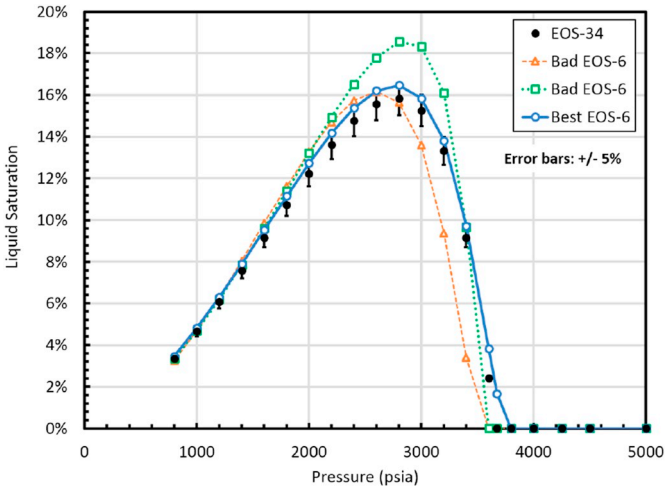

Fig. 3. Comparison of several lumped EOS schemes with the same number of components (6) with different degrees of accuracy compared to the original EOS model (34 components).

by averaging the tangents of fraction properties molecular weight, specific gravity, and Jacoby factor plotted vs. boiling point. Whitson (1983) suggests a method to estimate the number of C_{7+} fractions, and how they should be grouped. Coats (1985) discusses in detail lumping of C_{7+} fractions for modeling the vaporization process in gas condensate cycling. He gives a set of fundamental criteria and methods to calculate the EOS parameters of lumped pseudo-components. The impact of lumping on PVT model quality and reservoir simulation modeling of a gas cycling process is also shown. Behrens and Sandler (1988) suggest a grouping method for C_{7+} fractions based on application of the Gaussian-quadrature method used in continuous thermodynamics. Whitson et al. (1989) show that any C_{7+} molar-distribution can be modeled with the Gamma distribution with any number of pseudo-components. Other pseudoization methods have been proposed by Montel and Gouel (1984), Li et al. (1985), Newley and Merrill (1991), Danesh et al. (1992) and Hustad and Dalen (1993). However, none of them ensures that the lumping scheme is optimal.

Alavian et al. (2014) present a comprehensive method to find an optimal pseudoized EOS model to describe all PVT data that is relevant to a particular reservoir development (e.g. depletion performance, miscible gas injection, compositional variation, surface processing). The method tests all possible lumping scenarios and assess the quality of individual scenarios by quantifying its mismatch from the detailed EOS model using a root mean square (RMS) error (δ_{RMS}):

$$\delta_{\text{RMS}} = \left(\frac{\sum_{i=1}^{N_{\text{data}}} (r_i \cdot w_i)^2}{\sum_{i=1}^{N_{\text{data}}} w_i^2} \right)^{\frac{1}{2}}, \tag{4}$$

with w_i a set of weight factor and r_i the residual for data i defined as:

$$r_i = 100 \cdot \frac{d_{i,x} - d_{i,xx}}{d_{i,ref}}, \quad (5)$$

where $d_{i,x}$ is the PVT data calculated with the lumped EOS, $d_{i,xx}$ with the detailed EOS and $d_{i,ref}$ is a reference value taken as the maximum of all d_{xx} data of a given type (e.g. oil density) in a given simulated lab test (e.g. CCE).

1.4. Genetic algorithm

A genetic algorithm (GA) is a heuristic method inspired by the process of natural selection that belongs to the larger class of evolutionary algorithms (EA) (Mitchell, 1998). A simple GA consists of three operators (see Fig. 4):

Selection: This operator selects chromosomes in the population for reproduction. The fitter the chromosome, the more times it is likely to be selected to reproduce.

Crossover: This operator roughly mimics biological recombination between two organisms.

Mutation: This operator randomly flips some of the bits in a chromosome. Mutation can occur at each bit position in a string with some probability, usually very small (e.g. 0.1–1%).

Turing (1950) proposed a "learning machine" that could follow the principle of evolution. Barricelli (1957) and Fraser (1957) were the first publications about GAs per say. GAs became very popular in the 1970s through the work of Holland (1975) (and his schema theorem). Since then, many papers about GAs have been published. Several selection, crossover and mutation have been tested. Mitchell (1998) gives a comprehensive review of GAs.

GAs have been widely used in the oil & gas industry. For example, Camargo et al. (2010) use a GA to optimize oil production and gas lift injection. Sarvestani et al. (2012) use a specially designed genetic algorithm to search for suitable regression parameters to match the EOS against measured data. The main advantage of the proposed method is (1) its high speed in finding a solution and (2) the multiple final solutions being found, and thereby confining the role of experts to the last stage when selecting among the multiple final solutions. To the author's knowledge, GA was never used to find the optimal lumping scheme.

1.5. Tabu search algorithm

Tabu[1] search algorithm, first introduced by Glover (1986) and formalized by Glover (1989) and Glover (1990), is an heuristic search method employing local search methods used for mathematical optimization.

Tabu search uses a local (or neighborhood search) procedure to iteratively move from one potential solution $\mathbf{x}$ to an improved solution $\mathbf{x}'$ in the neighborhood $\mathcal{N}(\mathbf{x})$ of $\mathbf{x}$, until some stopping criterion has been satisfied. The neighborhood $\mathcal{N}(\mathbf{x})$ usually consists of the closest admissible point around the current point $\mathbf{x}$. Local search procedures (e.g. gradient descent) often become stuck at a local minimum. In order to explore regions of the search space that would be left unexplored by other local search procedures, tabu search allows to choose neighbors $\mathbf{x}'$ that worsens the objective (if no improving move is available).

To avoid looping between two solutions, some neighbors are banned and stored in the tabu list $\mathcal{L}_T$. Usually, the tabu list contains the points that have been previously visited. The length of the tabu list $\mathcal{L}_T$ defines the *memory* of the algorithm and is usually a setting parameter.

The stopping criterion is usually an attempt limit or a score threshold or a maximum number of iterations without improvement. Fig. 5 shows the typical workflow of a tabu search and Fig. 6 shows a schematic example of a typical tabu search.

[1] The word *tabu* comes from the Tongan word to indicate things that cannot be touched because they are sacred.

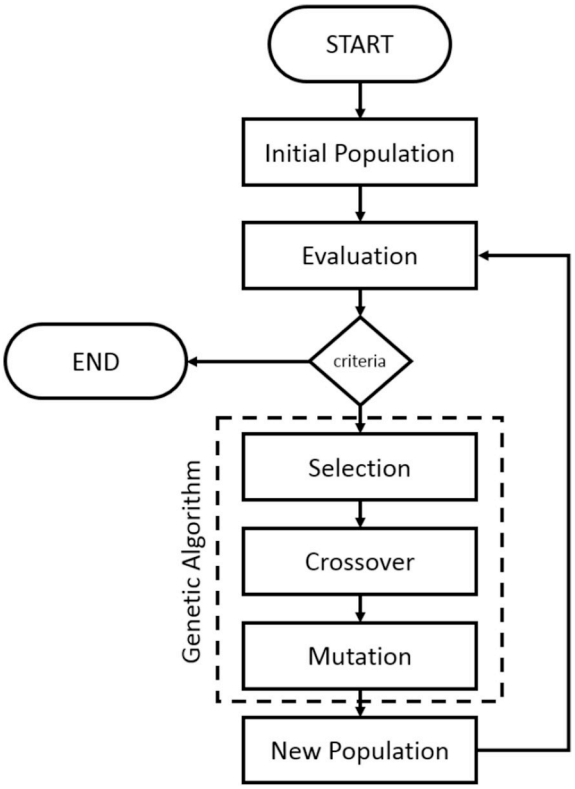

Fig. 4. Workflow of a genetic algorithm (adapted from Liao et al., 2001). The criterion can be (1) a number of generations or (2) a time criteria or (3) a minimum (or maximum) value for the objective function to maximize (or minimize).

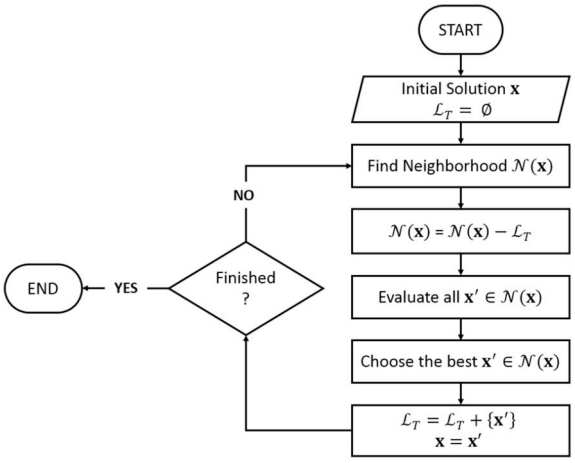

Fig. 5. Workflow of a typical tabu search.

Tabu search algorithms have been used extensively for integer programming purposes over the years. For example, Hansen et al. (1992) discuss the application of a tabu search algorithm to optimally find the location and sizing of offshore platforms.

1.6. Proposed approach

This paper is the first attempt to propose an optimization algorithm to optimally lump components from a detailed SCN EOS model (EOS*xx*) into a simplified EOS model (EOS*x*). It is assumed in this paper that the pseudo-components in the initial SCN EOS model are ordered by ascending molecular weight. The proposed approach uses a GA to quickly

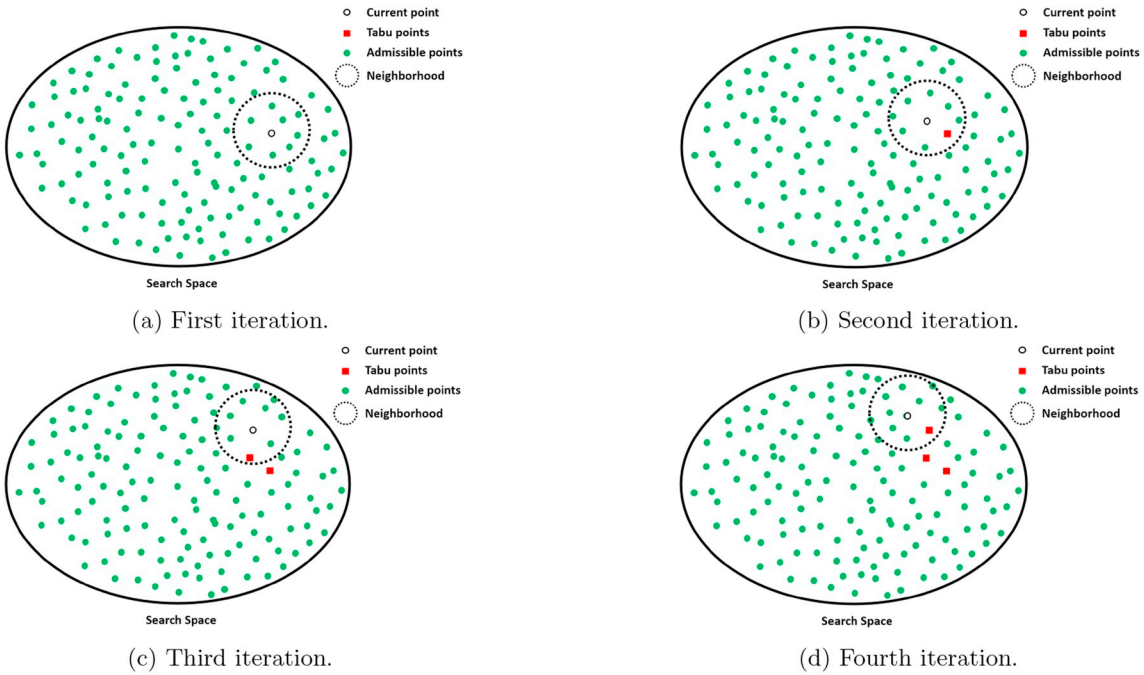

Fig. 6. Schematic illustration of a tabu search.

obtain an initial good solution. This initial solution is then refined using a tabu search algorithm.

2. Proposed algorithm

2.1. Chromosome

In this section, a lumping scheme will be called a *chromosome*. For a starting EOS with N_{xx} components, chromosomes have $N_{xx} - 1$ bits (b_k with $k \in [1, N_{xx} - 1]$). Each bit represent the separation between two sub-sequent components in the detailed EOS, see Fig. 7: (1) $b_k = 1$ means that there is a separation between components k and $k + 1$, (2) $b_k = 0$ means that the sub-sequent components k and $k + 1$ are lumped into a single pseudo-component.

The final EOS having N_x pseudo-components,

$$\sum_{k=1}^{N_{xx}-1} b_k = N_x - 1 \tag{6}$$

This is a *survival criteria* for chromosomes.

2.2. Fitness function

In this paper, we use the inverse of the RMS as fitness function:

$$f = \frac{1}{\delta_{\text{RMS}}}, \tag{7}$$

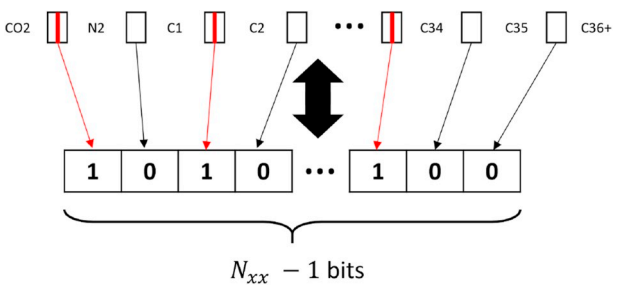

Fig. 7. Example of a chromosome with $N_{xx} - 1$ bits.

with δ_{RMS} define in Eqs. (4) and (5). It is computed based on 1200 lab measurements and EOS predictions obtained with the lumped EOS model (EOSx). Lab measurements include a CCE and a separator test for 5 different fluid (lean and rich gas condensate, critical oil, volatile and less volatile oil). All lab measurements can be found in Appendix A.

2.3. Proposed genetic algorithm

2.3.1. Selection method

Fitness Proportionate. In this paper, we use a Fitness Proportionate Selection (FPS) method In this selection method, also known as *roulette wheel selection*, chromosome i has a probability p_i to be selected:

$$p_i = \frac{f_i}{\sum_{j=1}^{N_c} f_j}, \tag{8}$$

where f_i is the fitness of chromosome i and N_c is the number of chromosomes in the population.

Fitness proportionate selection (FPS) was originally proposed by Holland (1975). The major issue with FPS lies in the early stage of the search: fitness variance in the population is high and a small number of chromosomes are much fitter than others. With FPS, these fitter chromosomes will multiply quickly preventing the GA to explore further the search space. This is known as *premature convergence* (Mitchell, 1998).

Mitchell (1998) notes that the selection operator needs to be balanced between exploitation and exploration: a too-strong selection can lead to suboptimal solutions while too-weak selection results in too-slow evolution. Mitchell (1998) reviews the most utilized selection methods implemented for genetic algorithm. Goldberg and Deb (1991), Bäck and Hoffmeister (1991), De la Maza et al. (1991) and Hancock (1994) provide a rigorous comparison of different selection methods. No other selection method was tested in this paper.

Elitism. De Jong (1975) introduce an additional logic that forces the GA to select a certain number of the best chromosomes at each generation. Elitism can be applied to any selection method and could significantly improve the GA's performances. In this paper, elitism is used together with FPS.

Population Sampling. The simplest sampling method consist of choosing a random number $r \in [0,1]$ and selecting the chromosome i

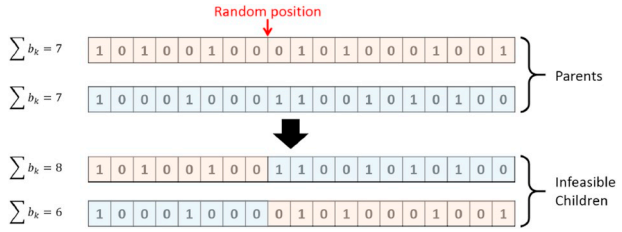

Fig. 8. Example of a single-point crossover between two parent chromosomes. The resulting children do not verify the survival criteria defined by Eq. (6).

verifying:

$$\sum_{j=1}^{i} p_j \geq r \quad \text{and} \quad \sum_{j=1}^{i-1} p_j < r \tag{9}$$

This procedure is applied N_p times to select N_p parents. This stochastic method should normally respects the probability distribution (p_i). However, with the relatively small population used in GAs, it is usually not the case. Baker (1987) propose a different sampling method (Stochastic Universal Sampling or SUS) that solves this issue. Instead of spinning the wheel N_p times, SUS spins the wheel once with N_p equally separated pointers which are used to select the N_p parents. SUS is used in this paper.

2.3.2. Crossover

Many crossover operators haven been proposed in the literature. Single-point crossover (see Fig. 8) is the simplest form (Mitchell, 1998). Other more complex crossover operators (two-point crossover, uniform) have been covered and analyzed (Gwiazda, 2006).

However, in the problem of EOS lumping optimization, Eq. (6) must be verified for all chromosomes. Most of the crossover strategies are not suited for this additional constraint, see Fig. 8. Simple crossover strategies likely yield infeasible children which must be discarded. This eventually can lead to reduced exploration of the search space.

The proposed crossover algorithm uses two parents P_1 and P_2 among the pool of parents $\mathscr{S}_p$ and generate one child chromosome that verifies Eq. (6), see Fig. 9. Bits that are identical in both parents are transferred to the child. These bits are represented in blue in Fig. 9. The remaining "1" necessary to honor Eq. (6) are randomly chosen among the following set:

$$k \in \{1, \ldots, N_{xx} - 1\} \quad \begin{cases} b_k(P_1) = 1 \text{ or } b_k(P_2) = 1 \\ \text{and} \\ b_k(P_1) \neq b_k(P_2) \end{cases} \tag{10}$$

where $b_k(P_i)$ is the bit k of parent P_i. These bits are represented in red in Fig. 8. Appendix B gives the proposed crossover algorithm. The proposed algorithm is somewhat similar to the uniform crossover, but it always ensures the creation of a feasible child.

2.3.3. Mutation

Mutation is an operator ensuring the population against fixation at any particular locus (Mitchell, 1998). Crossover and mutation have the same ability for population disruption. Some comparative studies have been performed on the power of mutation versus crossover, see Spears et al. (1992), Luke and Spector (1998) and White and Poulding (2009).

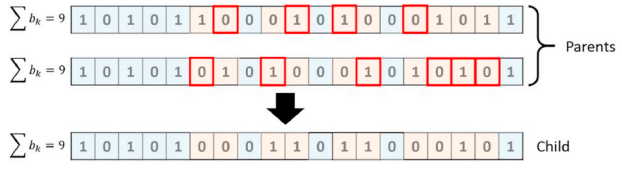

Fig. 9. Illustration of the proposed crossover operator. The resulting child *always* verifies the survival criteria given by Eq. (6).

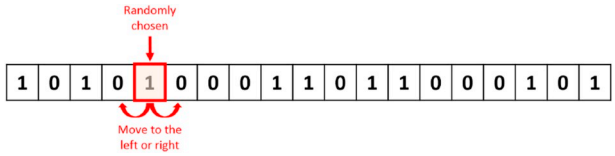

Fig. 10. Illustration of the proposed mutation operator. The resulting child *always* verifies the survival criteria given by Eq. (6).

Table 1
Parameters used in the proposed GA.

Parameter	Value
Population size	50
Mutation rate	5%
Elit size	10
Selection method	FPS
Number of generations	10

In the problem of EOS lumping optimization, a chromosome must honor Eq. (6) after mutation. Classical mutation operator are likely to transform feasible chromosomes into infeasible chromosomes. A different approach is used in this study. If a chromosome is selected for mutation, one of its bits b_k verifying (1) $b_k = 1$ and (2) $b_{k-1} = 0$ or $b_{k+1} = 0$ is randomly chosen. The chosen bit b_k is set to 0 and the 1 is transferred to one of its neighbor bits (b_{k-1} or b_{k+1}), see Fig. 10. Appendix B gives the proposed algorithm.

2.3.4. GA parameters

Table 1 gives the main settings of the GA used in this study.

2.4. Tabu search algorithm

The Tabu search algorithm starts from the best solution obtained by the GA after 10 generations. The tabu list is empty at the beginning ($\mathscr{L}_T$).

At each iteration, the neighbors of the current point have to be determined. A neighbor is cerated by moving one bit b_k whose value is one, to the left or to the right, see Fig. 11. When creating a lumped EOSx with N_x components, a maximum of $2{\cdot}(N_x - 1)$ such neighbors are possible.

Neighbors that are part of the tabu list are discarded. The remaining neighbors are then evaluated and the neighbor giving the lowest RMS is chosen for the next iteration.

The stopping criterion is expressed as a limit number of iterations without objective improvements. In this study, this parameter is set to 10. In addition, the limit of the tabu list is not constrained (i.e. infinite). This ensures that no point will be chosen two times.

3. Results and discussion

In this section, we assume that a detailed SCN EOS consisting of 34 components, EOS-34, represents with a good degree of accuracy the lab data. Three lumping scenarios are tested:

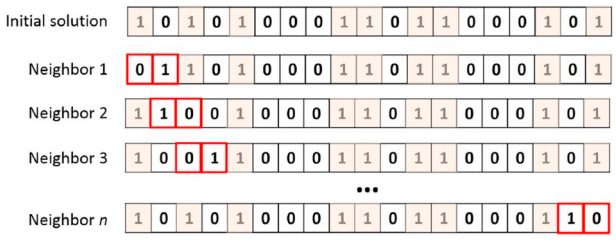

Fig. 11. Example of neighbors of a given solution.

EOS34	EOS15	EOS9	EOS6
N2	N2	N2	N2
C1	C1	C1	C1
CO2	CO2	CO2	CO2
C2	C2	C2	C2
C3	C3	C3	C3
i-C4	i-C4	i-C4	i-C4
n-C4	n-C4	n-C4	n-C4
i-C5	i-C5	i-C5	i-C5
n-C5	n-C5	n-C5	n-C5
C6	C6	C6	C6
C7	C7	C7	C7
C8	C8	C8	C8
C9	C9	C9	C9
C10	C10	C10	C10
C11	C11	C11	C11
C12	C12	C12	C12
C13	C13	C13	C13
C14	C14	C14	C14
C15	C15	C15	C15
C16	C16	C16	C16
C17	C17	C17	C17
C18	C18	C18	C18
C19	C19	C19	C19
C20	C20	C20	C20
C21	C21	C21	C21
C22	C22	C22	C22
C23	C23	C23	C23
C24	C24	C24	C24
C25	C25	C25	C25
C26	C26	C26	C26
C27	C27	C27	C27
C28	C28	C28	C28
C29	C29	C29	C29
C30+	C30+	C30+	C30+
RMS	**1.249%**	**1.265%**	**2.016%**

Fig. 12. Optimal solutions for the three scenarios. User-defined pseudo-components are shown in green, and optimized pseudo-components are shown in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

1. Lumping EOS-34 into EOS-15 (10 light components fixed)
2. Lumping EOS-34 into EOS-9 (3 light components fixed)
3. Lumping EOS-34 into EOS-6 (no component fixed)

The optimal solutions are shown in Fig. 12. The same optimal solution from Alavian et al. (2014) are obtained. Note that only the weighted RMS is used in the fitness function, while Alavian et al. (2014) also consider the minimum miscibility pressure (MMP).

The key PVT property predictions (1200 measurements) using the optimal lumped EOS*x* models from Fig. 12 are shown in A. These can be compared with the original, detailed EOS*xx* model calculations. Overall the optimal lumped EOS*x* models provide very accurate PVT predictions.

Table 2
Summary of the three cases.

Key Performance Index	Case-1	Case-2	Case-3
Number of simulation	30	30	100
Convergence towards the best solution	26	26	92
Success rate	86.7%	86.7%	92%
Best RMS	1.249%	1.265%	2.016%
Average RMS when non-convergence	1.26%	1.81%	4.55%
Average total execution time	45 min 43s	22 min 20s	14 min 24s
Average number of model evaluations (GA)	510	510	510
Average number of model evaluations (Tabu)	95	133	165
Average total number of model evaluations	605	643	675
Number of possible lumping schemes	8855	142506	237336
Proportion of possible lumping schemes being run	6.83%	0.45%	0.28%

The performances of the proposed optimization solution are shown in Table 2. The proposed solution converges approximately 90% of the times towards the best lumping solution (known from Alavian et al. (2014)). In approximately 10% of the cases, the optimization algorithm is attracted by a local minimum and the proposed settings of the GA and tabu search do not allow to exit the attraction basin. A solution to improve the convergence rate consists of increasing the limit number of iterations without objective improvement in the tabu search. In this study, this parameter was held constant and equal to 10. Note that this will also impact the average runtime of the solution.

Compared to Alavian et al. (2014), the number of cases being run is independent of the initial EOS (EOS*xx*) and final EOS (EOS*x*). For case 3, only 0.28% of the possible lumping schemes are actually being run, leading to significant savings in execution time. The execution time is less than an hour, allowing the user to re-run the solution. If run two times, the solution has a 0.01% probability of not converging towards the best solution. In general, if run N times, the solution has a 0.1^N probability of not finding the optimal lumping scheme.

The observed difference in execution time (ranging from 15 min to 45 min) is primarily due to the PVT simulations being more time-consuming when the number of components increases.

Fig. 13 shows the evolution of the genetic algorithm. The average RMS is slowly decreasing, as fittest lumping schemes are chosen and combined together. However, the best lumping scheme after 10

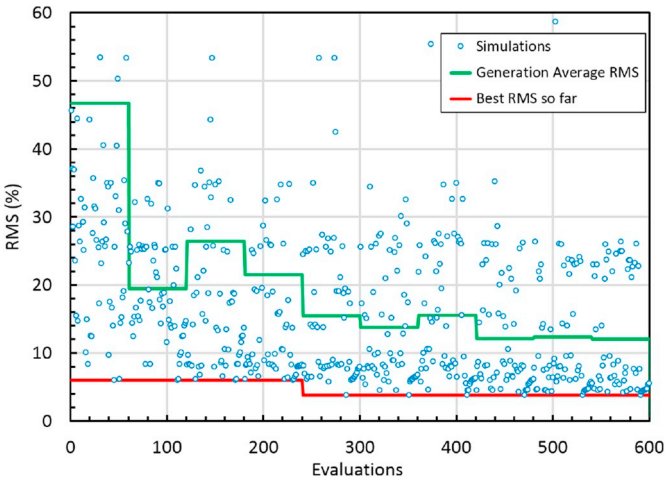

Fig. 13. Evolution of the genetic algorithm for case 3 (lumping of EOS-34 into EOS-6).

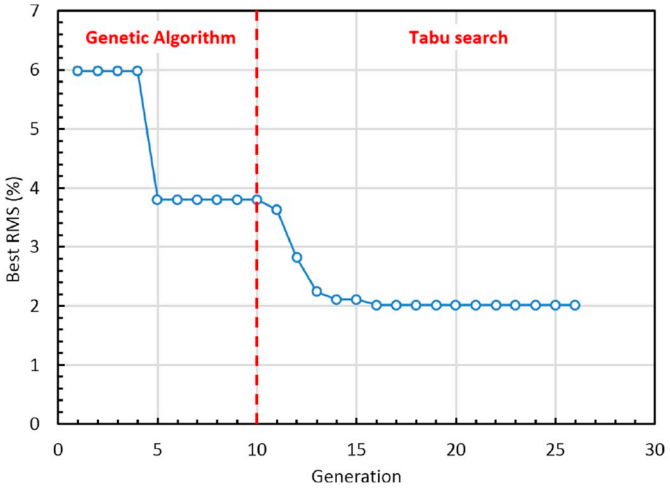

Fig. 14. Solution improvement using a tabu search for case 3 (lumping of EOS-34 into EOS-6).

generations is relatively far from the best solution (RMS of 3.8% vs. 2.016% for the best solution).

This slow GA convergence was the primary motivation to include a tabu search. The best solution obtained from the GA is used as starting point for the tabu search, see Fig. 14. Convergence towards the best solution is achieved within a couple of iterations. Note that each tabu search generation only evaluates the neighbors around the current search point. Therefore, a tabu search iteration is considerably less costly than a GA generation (evaluation of the entire population).

Fig. 15 shows an example of the exploration of the search space by the tabu search algorithm.

In addition, tabu search has the capability of exiting local minimum attraction basin and continuing exploration. This characteristic is very useful in the case of EOS lumping, as many lumping schemes can be local RMS minima, see Fig. 16.

Fig. 17 shows how the tabu search can exit a local attraction basin and continue the exploration to finally converge towards the best solution.

4. Conclusions

- We proposed a new optimization algorithm to optimally lump a detailed SCN EOS (EOS*xx*) into a simplified EOS (EOS*x*). The algorithm consists of a genetic algorithm and a tacu search. The genetic algorithm is used to find a good starting point for the tabu search which is then refined by a tabu search.
- The proposed solution is tested for three lumping scenarios. It converges in 90% of the cases. The number of lumping schemes being evaluated is independent of the initial and final EOS. Only a small proportion (ranging from 0.2 to 7%) of the total number of possible lumping schemes is evaluated, leading to significantly reduced run time.
- Future improvement could include multiple starting points for the tabu search to improve the convergence rate and more complex fitness functions (e.g. including MMP). More complex PVT simulations will surely impact the runtime of the solution.

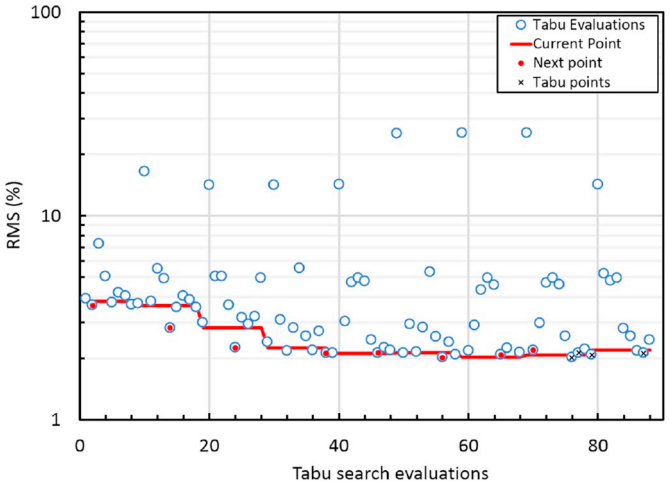

Fig. 15. Exploration of the variable space by the tabu search algorithm (lumping of EOS-34 into EOS-6).

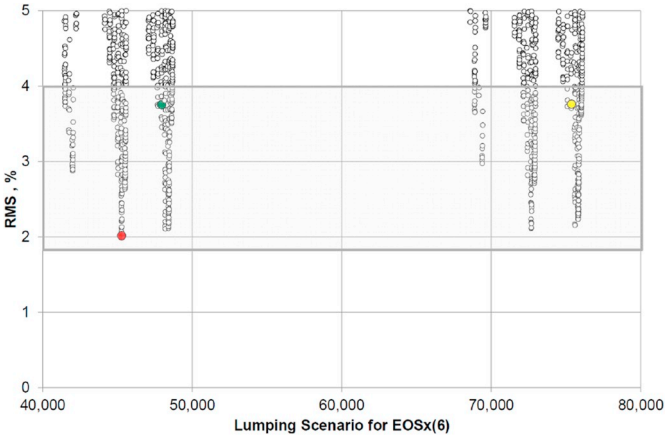

Fig. 16. Example of local minima when lumping EOS-34 into EOS-6 (after Alavian et al. (2014)).

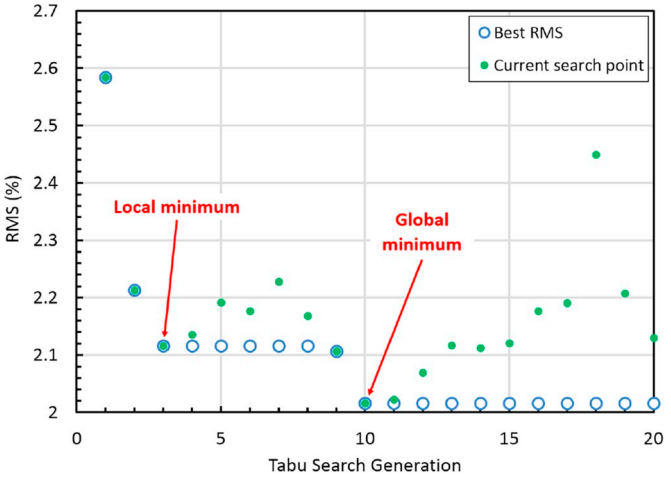

Fig. 17. Example of local attraction basin and the tabu search's capability to continue the exploration.

Unit Conversion

Field	Conversion	S.I.
1 psi	= 6.894757	Pa
1 bbl	= 0.1589873	m^3
1 cf	= 0.028316846592	m^3
1 cp	= 0.001	Pa.s
°API	$\rho = 141.5/(131.5 + \gamma_{API})$	g/cm^3
°F	$°C = (°F - 32)\cdot\frac{5}{9}$	°C
°R	$K = °R\cdot\frac{5}{9}$	K

Acknowledgment

I would like to acknowledge Zick Technologies for the free usage of its PVT simulation software, *PhazeComp*.

Nomenclature

δ_{RMS}	Root mean square error [%]
$\mathscr{L}_T$	Tabu list
$\mathscr{N}(\mathbf{x})$	Neighborhood of point $\mathbf{x}$ in the tabu search
$\mu_{i\pi}$	Chemical potential of component i in phase π
b_k	Bit k in a chromosome
$d_{i,\mathrm{ref}}$	Reference value for PVT data i
$d_{i,xx}$	PVT data i calculated with EOSxx
$d_{i,x}$	PVT data i calculated with EOSx
f_i	Fitness of chromosome i
N	Quantity of moles
N_c	Number of chromosomes
N_P	Number of parents
N_x	Number of components in EOSx
N_{xx}	Number of components in EOSxx
P	Pressure [bar or psi]
p_i	Probability of chromosome i to be selected
R	Ideal gas constant
r_i	Residual relative error for PVT data i [%]
T	Temperature [°F or °R]
V	Molar volume [m^3/kmol]
w_i	Weight factor for PVT data i
Z	Z factor of a mixture
z_i	Molar fraction of component i [mol-%]
BIP	Binary Interaction Parameter
BTX	Benzene, toluene, and xylene isomers
CCE	Constant Composition Expansion
CVD	Constant Volume Depletion
EOS	Equation of State
EOSx	Lumped EOS model with $N_x$components
EOSxx	Detailed SCN EOS model with N_{xx} components
FPS	Fitness-proportionate selection
GA	Genetic Algorithm
GC	Gas condensate
HC	Hydrocarbon
PVT	Pressure, Volume, Temperature
RMS	Root mean square
SCN	Single carbon number
SUS	Stochastic Universal Sampling

Appendix A. PVT Data

Appendix A.1. PVT models

Table A.4 gives the EOS parameters of the original detailed SCN EOS model consisting of 34 components. Table A.5 give the compositions of the 5 fluids used in the PVT experiments.

Appendix A.2. PVT experiments

A total of 1200 PVT measurements were used in this study, The reference point for comparison is the original detailed EOS model (EOS-34). The weighting factors used in the RMS are given in Table. A.3.

Table A.3
Global weighting factors used in the RMS (Alavian et al., 2014).

Fluid property	Weighting factor
Liquid Volumes	3
Liquid Saturation	3
Liquid Density	2
Gas-Oil Ratio	2
Condensate-Gas Ratio	2
Relative Volume	1
Gas Specific Gravity	1
Gas Density	1
Gas Z-factor	1

(*continued on next page*)

Table A.3 (*continued*)

	Weighting
Liquid API	1
Liquid Viscosity	0
Gas Viscosity	0

Figs A.19-A.33 compare the predictions using the different EOS models (EOS-15, EOS-9 and EOS-6, see Fig. 12) for several CCE experiments and several reservoir fluids (see Table. A.5). Table A.9 compares the separator test predictions for the different EOS models. Fig. 18 gives the conditions of the separator test performed for each fluid of Table. A.5.

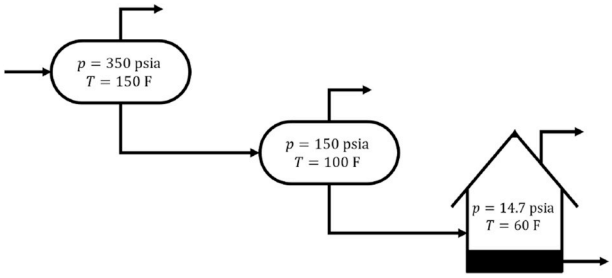

Fig. A.18. Separator test experiment used in this paper.

Table A.4
Peng-Robinson Equation of State EOS-34 (Alavian et al., 2014). Only C_1/C_{7+} and HC/non-HC BIP's are non-zeros.

Components	Molecular	Critical	Critical	Acentric	Volume	Boiling	Specific	LBC	BIP's		
	Weight	Temperature	Pressure	Factor	Shift	Point	Gravity	Z-factor	N2	C1	CO2
	M	T_c (R)	p_c (psia)	ω	s	T_b (R)	γ	Z_c			1
N2	28.01	227.16	492.84	0.037	−0.1676	139.00	0.2834	0.2918	0.000	0.025	0.000
C1	16.04	343.01	667.03	0.011	−0.1500	200.80	0.1461	0.2862	0.025	0.000	0.105
CO2	44.01	547.42	1069.51	0.225	0.0019	332.80	0.7619	0.2743	0.000	0.105	0.000
C2	30.07	549.58	706.62	0.099	−0.0628	332.00	0.3298	0.2792	0.010	0.000	0.130
C3	44.10	665.69	616.12	0.152	−0.0638	415.70	0.5098	0.2763	0.090	0.000	0.125
I−C4	58.12	734.13	527.94	0.186	−0.0620	470.70	0.5704	0.2820	0.095	0.000	0.120
N−C4	58.12	765.22	550.56	0.200	−0.0539	490.70	0.5906	0.2739	0.095	0.000	0.115
I−C5	72.15	828.70	490.37	0.229	−0.0565	542.10	0.6295	0.2723	0.100	0.000	0.115
N−C5	72.15	845.46	488.78	0.252	−0.0293	556.80	0.6359	0.2684	0.110	0.000	0.115
C6	82.42	924.03	489.98	0.240	−0.0026	606.40	0.7028	0.2702	0.110	0.000	0.115
C7	96.10	990.77	454.06	0.275	0.0137	661.20	0.7370	0.2655	0.110	0.026	0.115
C8	108.94	1043.64	421.23	0.311	0.0279	707.70	0.7583	0.2613	0.110	0.030	0.115
C9	122.09	1093.72	388.43	0.352	0.0506	754.30	0.7750	0.2570	0.110	0.034	0.115
C10	135.01	1138.11	360.17	0.392	0.0716	797.10	0.7884	0.2532	0.110	0.038	0.115
C11	147.85	1178.35	335.50	0.431	0.0912	837.10	0.7997	0.2497	0.110	0.042	0.115
C12	160.59	1215.08	313.89	0.471	0.1093	874.50	0.8094	0.2465	0.110	0.045	0.115
C13	173.24	1248.80	294.88	0.506	0.1261	909.70	0.8180	0.2434	0.110	0.048	0.115
C14	185.78	1279.91	278.07	0.542	0.1417	942.80	0.8257	0.2404	0.110	0.051	0.115
C15	198.22	1308.75	263.14	0.578	0.1560	974.10	0.8327	0.2375	0.110	0.054	0.115
C16	210.55	1335.60	249.83	0.614	0.1692	1003.60	0.8391	0.2347	0.110	0.056	0.115
C17	222.77	1360.66	237.91	0.648	0.1813	1031.60	0.8450	0.2320	0.110	0.059	0.115
C18	234.88	1384.16	227.20	0.682	0.1924	1058.10	0.8504	0.2294	0.110	0.061	0.115
C19	246.87	1406.24	217.53	0.715	0.2025	1083.20	0.8555	0.2268	0.110	0.063	0.115
C20	258.75	1427.05	208.78	0.747	0.2118	1107.20	0.8602	0.2243	0.110	0.065	0.115
C21	270.52	1446.72	200.83	0.778	0.2203	1130.00	0.8647	0.2218	0.110	0.067	0.115
C22	282.18	1465.35	193.58	0.809	0.2281	1151.70	0.8688	0.2195	0.110	0.068	0.115
C23	293.73	1483.03	186.96	0.839	0.2352	1172.50	0.8728	0.2171	0.110	0.070	0.115
C24	305.17	1499.86	180.89	0.868	0.2417	1192.30	0.8766	0.2149	0.110	0.071	0.115
C25	316.50	1515.89	175.31	0.897	0.2476	1211.30	0.8801	0.2127	0.110	0.073	0.115
C26	327.73	1531.20	170.17	0.925	0.2529	1229.50	0.8835	0.2105	0.110	0.074	0.115
C27	338.86	1545.84	165.41	0.953	0.2578	1247.00	0.8868	0.2084	0.110	0.075	0.115
C28	349.89	1559.87	161.01	0.979	0.2623	1263.70	0.8899	0.2064	0.110	0.076	0.115
C29	360.82	1573.32	156.92	1.006	0.2663	1279.90	0.8929	0.2044	0.110	0.077	0.115
C30+	496.34	1711.84	121.96	1.298	0.2940	1446.50	0.9235	0.1831	0.110	0.087	0.115

Table A.5
Compositions of five reservoir fluids in EOS-34 used in the PVT experiments (Alavian et al., 2014).

Components	Lean GC		Rich GC		Near Critical Oil		More Volatile Oil		Less Volatile Oil	
	mol-%	mass-%	mol-%	mass-%	mol-%	mass-%	mol-%	mass-%	mol-%	mass-%
N2	0.157	0.161	0.138	0.124	0.116	0.085	0.091	0.047	0.072	0.024
CO2	0.182	0.294	0.182	0.256	0.176	0.203	0.166	0.136	0.160	0.085
C1	68.839	40.510	64.168	32.954	58.010	24.381	48.735	14.594	38.249	7.409
C2	13.596	14.997	13.988	13.465	14.155	11.151	13.770	7.729	12.461	4.524
C3	7.361	11.907	8.084	11.412	8.738	10.095	9.135	7.519	8.729	4.648
I−C4	0.745	1.589	0.857	1.594	0.966	1.470	1.051	1.141	1.036	0.727
N−C4	2.747	5.858	3.219	5.989	3.712	5.652	4.163	4.516	4.198	2.946
I−C5	0.634	1.678	0.779	1.799	0.939	1.774	1.100	1.482	1.148	1.000
N−C5	1.097	2.903	1.365	3.152	1.669	3.155	1.990	2.680	2.102	1.831
C6	1.228	3.712	1.614	4.259	2.092	4.517	2.673	4.112	3.023	3.009
C7	0.897	3.161	1.255	3.862	1.738	4.376	2.403	4.311	2.920	3.388
C8	0.703	2.811	1.038	3.621	1.514	4.322	2.216	4.506	2.826	3.717
C9	0.494	2.213	0.771	3.013	1.187	3.797	1.840	4.193	2.462	3.629
C10	0.357	1.770	0.588	2.543	0.954	3.375	1.563	3.939	2.189	3.569
C11	0.259	1.404	0.449	2.127	0.767	2.971	1.326	3.658	1.941	3.464
C12	0.188	1.107	0.344	1.767	0.617	2.596	1.123	3.367	1.717	3.329
C13	0.137	0.869	0.263	1.460	0.497	2.254	0.951	3.077	1.517	3.173
C14	0.100	0.679	0.202	1.202	0.400	1.948	0.806	2.795	1.339	3.003
C15	0.073	0.530	0.156	0.987	0.323	1.677	0.683	2.526	1.181	2.827
C16	0.053	0.413	0.120	0.808	0.261	1.439	0.579	2.274	1.042	2.648
C17	0.039	0.321	0.093	0.660	0.211	1.232	0.491	2.041	0.919	2.471
C18	0.029	0.250	0.072	0.539	0.171	1.053	0.417	1.827	0.811	2.299
C19	0.021	0.194	0.056	0.439	0.139	0.898	0.354	1.632	0.716	2.133
C20	0.016	0.151	0.043	0.358	0.113	0.765	0.301	1.455	0.632	1.974
C21	0.012	0.117	0.034	0.292	0.092	0.652	0.257	1.295	0.559	1.825
C22	0.009	0.091	0.026	0.238	0.075	0.555	0.219	1.152	0.494	1.684
C23	0.007	0.071	0.021	0.194	0.061	0.472	0.187	1.024	0.438	1.553
C24	0.005	0.055	0.016	0.158	0.050	0.402	0.160	0.910	0.388	1.431
C25	0.004	0.043	0.013	0.129	0.041	0.342	0.137	0.808	0.345	1.317
C26	0.003	0.034	0.010	0.105	0.034	0.291	0.117	0.718	0.306	1.212
C27	0.002	0.026	0.008	0.086	0.028	0.248	0.101	0.637	0.273	1.116
C28	0.002	0.021	0.006	0.070	0.023	0.211	0.087	0.566	0.243	1.027
C29	0.001	0.016	0.005	0.057	0.019	0.180	0.075	0.503	0.217	0.944
C30 +	0.003	0.047	0.018	0.284	0.113	1.464	0.737	6.831	3.348	20.065

Table A.6
Peng-Robinson Equation of State EOS-15. Only C_1/C_{7+} and HC/non-HC BIP's are non-zeros.

Components	Molecular	Critical	Critical	Acentric	Volume	Boiling	Specific	LBC	BIP's		
	Weight	Temperature	Pressure	Factor	Shift	Point	Gravity	Z-factor	N2	C1	CO2
	M	T_c (R)	p_c (psia)	ω	s	T_b (R)	γ	Z_c			
CO2	44.01	547.42	1069.51	0.225	0.0019	332.79	0.7619	0.2743	0.000	0.105	0.000
N2	28.01	227.16	492.84	0.037	−0.1676	139.04	0.2834	0.2918	0.000	0.025	0.000
C1	16.04	343.01	667.03	0.011	−0.1500	200.80	0.1461	0.2862	0.025	0.000	0.105
C2	30.07	549.58	706.62	0.099	−0.0628	331.98	0.3298	0.2792	0.010	0.000	0.130
C3	44.10	665.69	616.12	0.152	−0.0638	415.68	0.5098	0.2763	0.090	0.000	0.125
I−C4	58.12	734.13	527.94	0.186	−0.0620	470.69	0.5704	0.2820	0.095	0.000	0.120
N−C4	58.12	765.22	550.56	0.200	−0.0539	490.74	0.5906	0.2739	0.095	0.000	0.115
I−C5	72.15	828.70	490.37	0.229	−0.0565	542.09	0.6295	0.2723	0.100	0.000	0.115
N−C5	72.15	845.46	488.78	0.252	−0.0293	556.81	0.6359	0.2684	0.110	0.000	0.115
C6	82.42	924.03	489.98	0.240	−0.0026	606.37	0.7028	0.2702	0.110	0.000	0.115
C7-10	114.12	1063.29	405.39	0.329	0.0418	726.73	0.7647	0.2591	0.110	0.032	0.115
C11-15	170.29	1240.47	298.46	0.498	0.1240	901.50	0.8161	0.2437	0.110	0.048	0.115
C16-21	236.73	1387.28	225.46	0.687	0.1946	1061.89	0.8513	0.2286	0.110	0.061	0.115
C22-29	315.53	1514.31	175.79	0.895	0.2470	1209.52	0.8800	0.2125	0.110	0.073	0.115
C30+	496.34	1711.84	121.96	1.298	0.2940	1446.49	0.9235	0.1831	0.110	0.087	0.115

Table A.7
Peng-Robinson Equation of State EOS-9. Only the non-zero BIP's are shown in the table.

Components	Molecular	Critical	Critical	Acentric	Volume	Boiling	Specific	LBC	BIP's		
	Weight	Temperature	Pressure	Factor	Shift	Point	Gravity	Z-factor	N2	C1	CO2
	M	T_c (R)	p_c (psia)	ω	s	T_b (R)	γ	Z_c			
CO2	44.01	547.42	1069.51	0.225	0.0019	332.79	0.762	0.2743	0.000	0.105	0.000
N2	28.01	227.16	492.84	0.037	−0.1676	139.04	0.283	0.2918	0.000	0.025	0.000
C1	16.04	343.01	667.03	0.011	−0.1500	200.80	0.146	0.2862	0.025	0.000	0.105
C2-3	35.89	599.88	664.14	0.122	−0.0633	367.80	0.429	0.2778	0.049	0.000	0.128
C4-7	74.32	864.00	498.58	0.233	−0.0214	564.99	0.662	0.2704	0.105	0.007	0.115
C8-10	121.00	1089.20	390.22	0.349	0.0501	750.41	0.774	0.2572	0.110	0.034	0.115
C11-15	170.29	1240.47	298.46	0.498	0.1240	901.50	0.816	0.2437	0.110	0.048	0.115
C16-25	253.36	1416.97	212.36	0.733	0.2088	1096.11	0.858	0.2245	0.110	0.064	0.115
C26-30+	463.58	1681.89	128.77	1.230	0.2887	1410.15	0.918	0.1869	0.110	0.085	0.115

Table A.8
Peng-Robinson Equation of State EOS-6. Only the non-zero BIP's are shown in the table.

Components	Molecular	Critical	Critical	Acentric	Volume	Boiling	Specific	LBC	BIP's	
	Weight	Temperature	Pressure	Factor	Shift	Point	Gravity	Z-factor	N2−C1	CO2−C3
	M	T_c (R)	p_c (psia)	ω	s	T_b (R)	γ	Z_c		
N2−C1	16.07	342.750	666.664	0.011	−0.1500	200.67	0.146	0.2862	0.0000	−0.0001
CO2−C3	35.95	598.820	665.219	0.122	−0.0630	367.11	0.430	0.2778	−0.0001	0.0000
C4-7	74.32	864.000	498.579	0.233	−0.0214	564.99	0.662	0.2704	0.0069	−0.0001
C8-13	137.09	1143.170	353.829	0.400	0.0792	803.34	0.790	0.2519	0.0392	−0.0001
C14-23	229.29	1372.390	231.419	0.668	0.1893	1045.61	0.848	0.2295	0.0603	−0.0001
C24 +	444.01	1662.890	133.316	1.189	0.2848	1387.16	0.914	0.1892	0.0838	−0.0001

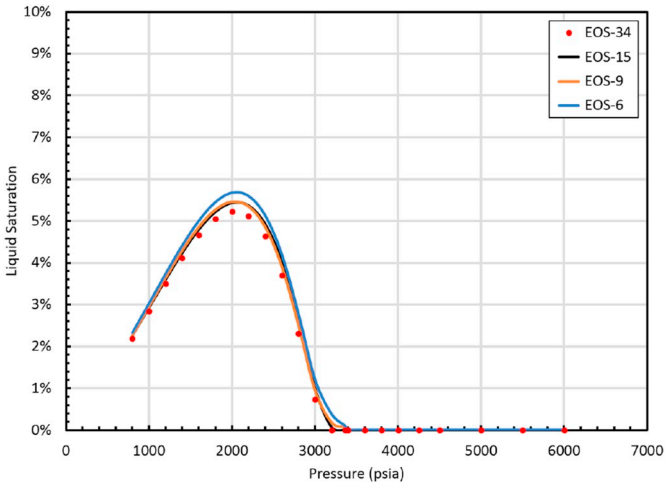

Fig. A.19. Predictions of liquid saturation for a lean gas condensate. CCE performed at 200 F.

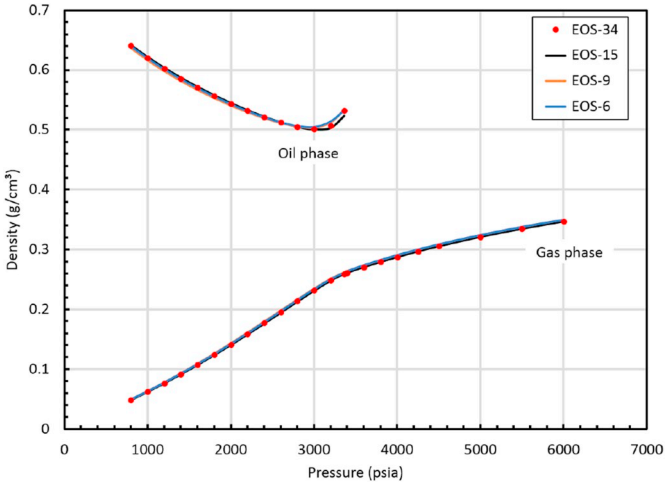

Fig. A.20. Predictions of liquid and gas density for a lean gas condensate. CCE performed at 200 F.

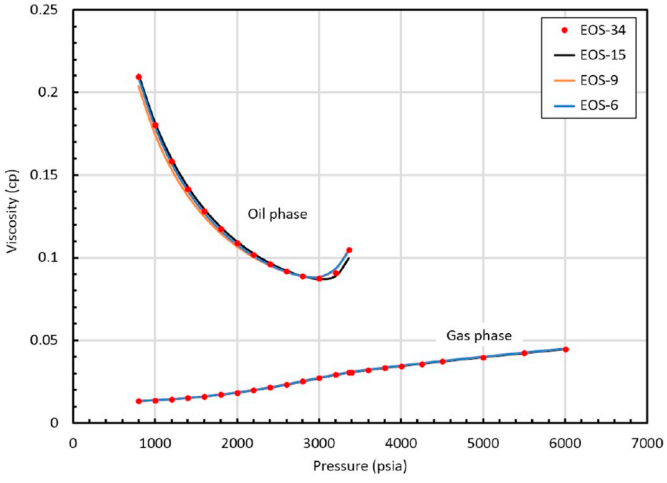

Fig. A.21. Predictions of liquid and gas viscosity for a lean gas condensate. CCE performed at 200 F.

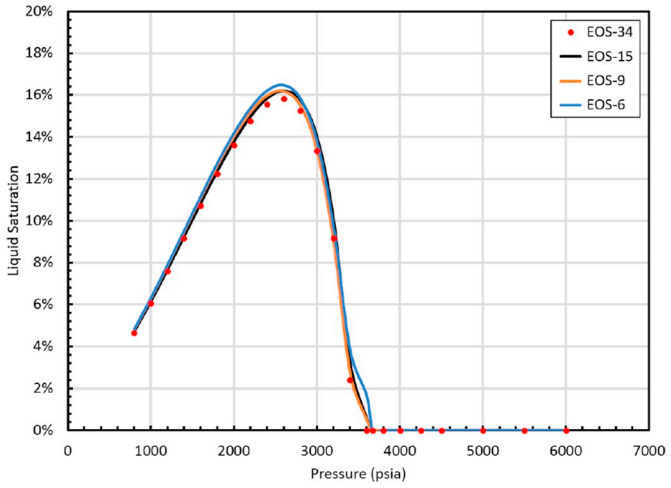

Fig. A.22. Predictions of liquid saturation for a richer gas condensate. CCE performed at 200 F.

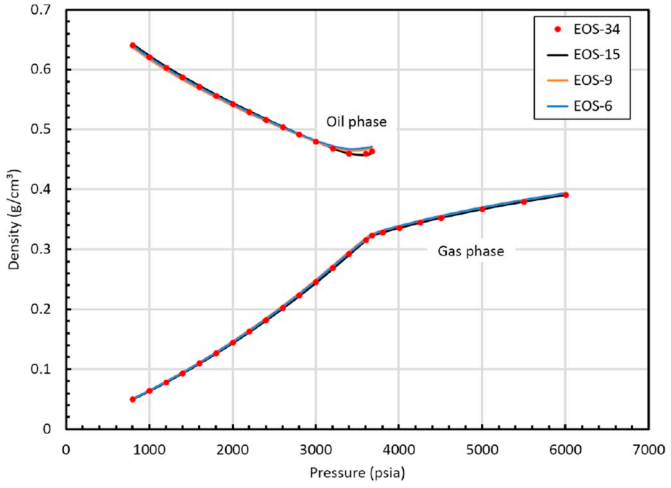

Fig. A.23. Predictions of liquid and gas density for a richer gas condensate. CCE performed at 200 F.

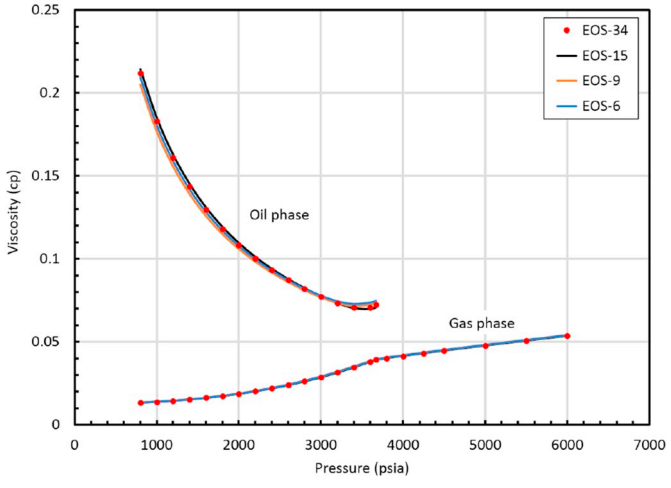

Fig. A.24. Predictions of liquid and gas viscosity for a richer gas condensate. CCE performed at 200 F.

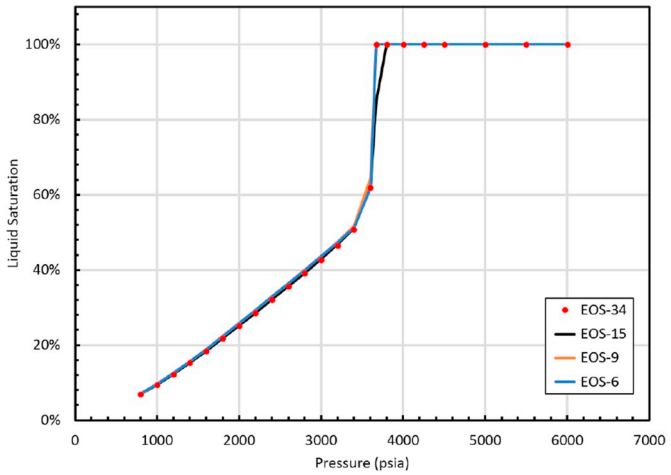

Fig. A.25. Predictions of liquid saturation for a near-critical oil. CCE performed at 200 F.

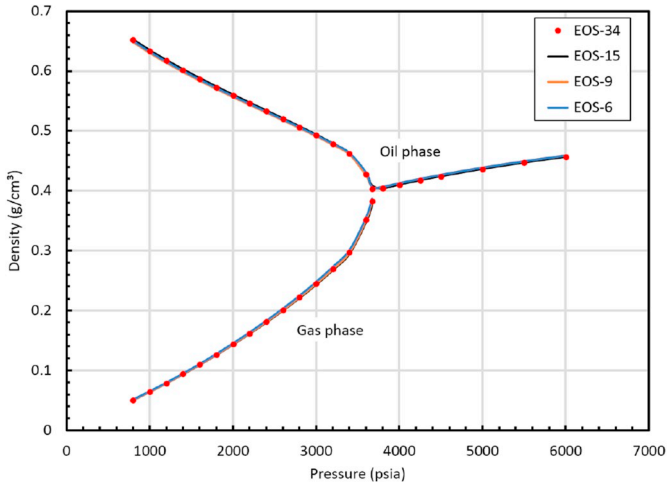

Fig. A.26. Predictions of liquid and gas density for a near-critical oil. CCE performed at 200 F.

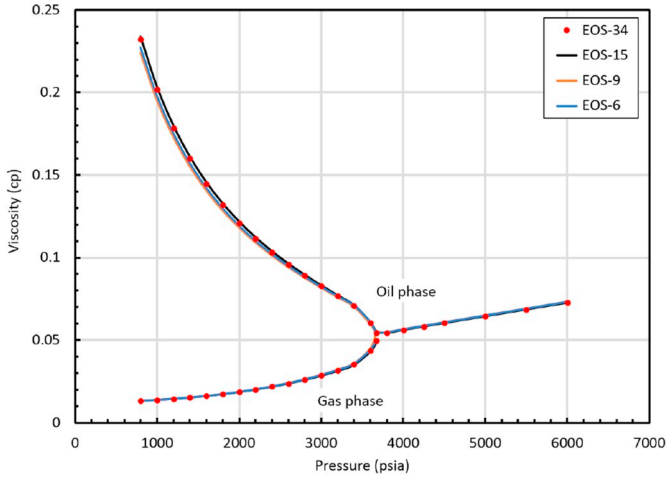

Fig. A.27. Predictions of liquid and gas viscosity for a near-critical oil. CCE performed at 200 F.

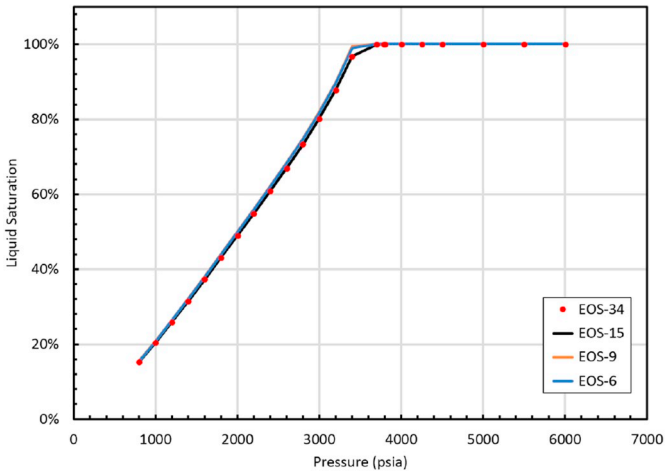

Fig. A.28. Predictions of liquid saturation for a more volatile oil. CCE performed at 200 F.

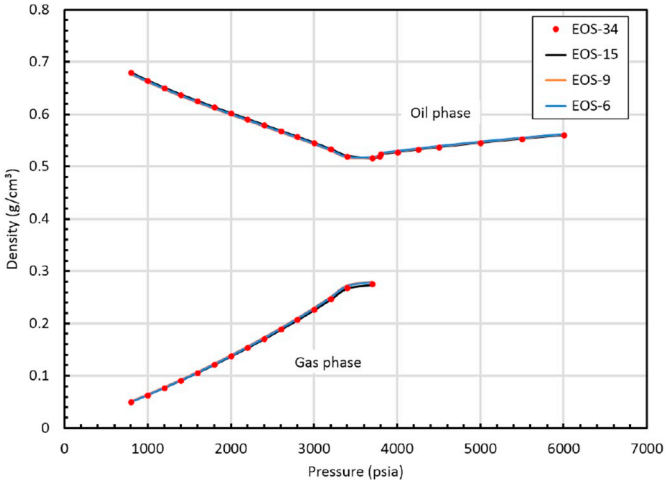

Fig. A.29. Predictions of liquid and gas density for a more volatile oil. CCE performed at 200 F.

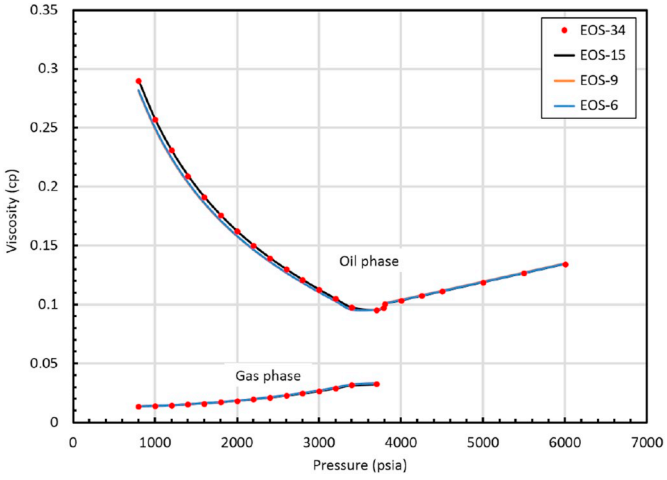

Fig. A.30. Predictions of liquid and gas viscosity for a more volatile oil. CCE performed at 200 F.

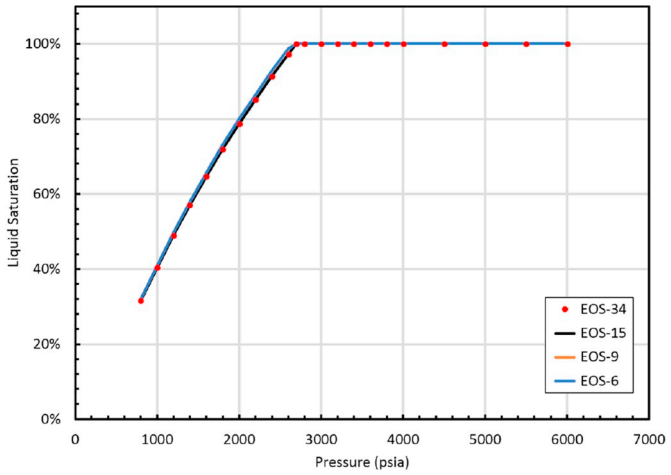

Fig. A.31. Predictions of liquid saturation for a less volatile oil. CCE performed at 200 F.

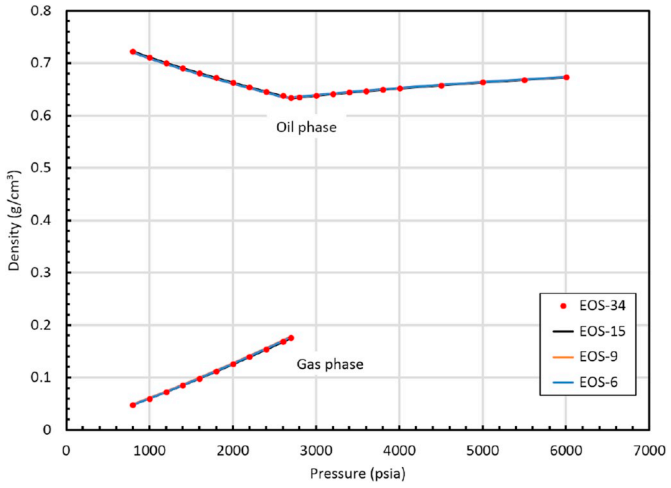

Fig. A.32. Predictions of liquid and gas density for a less volatile oil. CCE performed at 200 F.

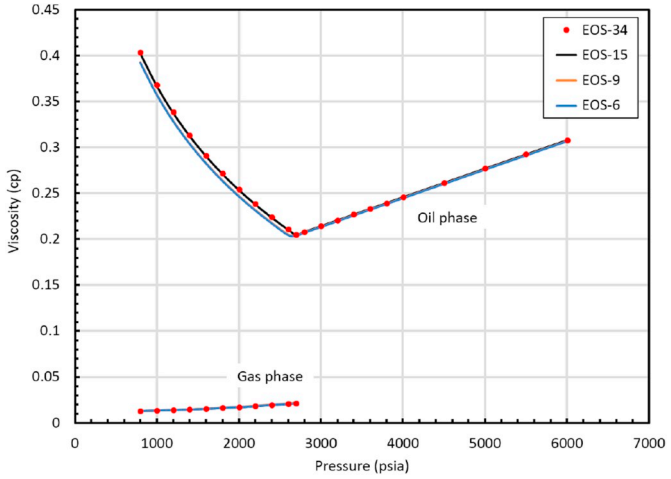

Fig. A.33. Predictions of liquid and gas viscosity for a less volatile oil. CCE performed at 200 F.

Table A.9
Separator test results.

Fluid	Lean	Rich	Near	More	Less
			Critical	Volatile	Volatile
Properties	GC	GC	Oil	Oil	Oil
EOS34 GOR [scf/STB]	–	–	**5175.0**	**2282.8**	**985.6**
EOS-15	–	–	−0.8%	−0.4%	−0.2%
EOS-9	–	–	−3.1%	−2.4%	−1.4%
EOS-6	–	–	−3.1%	−2.4%	−1.4%
EOS34 OGR [STB/MMscf]	**50.6**	**98.4**	–	–	–
EOS-15	3.5%	1.7%	–	–	–
EOS-9	1.7%	2.8%	–	–	–
EOS-6	3.3%	3.3%	–	–	–
EOS34 API [°API]	**56.1**	**55.1**	**52.7**	**48.0**	**41.7**
EOS-15	−0.6%	−0.6%	−0.4%	−0.2%	0.0%
EOS-9	2.0%	2.2%	2.0%	1.4%	0.9%
EOS-6	0.7%	1.2%	1.2%	0.9%	0.8%

Appendix B. Proposed GA Operators

Input: Set of parent chromoses $\mathcal{S}_p$, N_{xx}, N_x
Output: A feasible child c
Choose two random parents P_1 and P_2 among $\mathcal{S}_p$;
var k:=0;
array tab;
var s:=0;
for i =1 ... $N_{xx} - 1$ **do**
  if *P1[i]=1* **and** *P2[i]* =1 **then**
    c[i]:=1;
    k:=k+1;
  else
    if *P1[i]=1* **or** *P2[i]* =1 **then**
      tab[s] = i;
      s:=s+1;
    end
    c[i]:=0;
  end
end
while $k \leq N_x - 1$ **do**
  Choose a random integer $r \in [1; s]$;
  c[tab[r]]:=1;
  k:=k+1;
end
return c;

Algorithm 1: Proposed crossover operation.

```
Input: Population of chromoses S_c, N_xx, p_mut
Output: A modified population of chromosomes S_c
foreach c ∈ S_c do
    Choose a random number p ∈ [0; 1];
    if p ≤ p_mut then
        array tab;
        var s:=0;
        for i=1...N_xx − 2 do
            if c[i]=1 then
                tab[s]:= i;
                s:=s+1;
            end
        end
        var hasMuted:=false;
        while hasMuted=false do
            Choose a random integer r ∈ [1; s];
            var k:= tab[r];
            if c[k] = 1 then
                if c[k-1] = 0 and c[k+1] = 0 then
                    Choose a random number q ∈ [0; 1];
                    if q ≤ 0.5 then
                        c[k-1]:= 1;
                        c[k]:= 0;
                        hasMuted:=true;
                    else
                        c[k+1]:= 1;
                        c[k]:= 0;
                        hasMuted:=true;
                    end
                else if c[k-1] = 0 then
                    c[k-1]:= 1;
                    c[k]:= 0;
                    hasMuted:=true;
                else if c[k+1] = 0 then
                    c[k+1]:= 1;
                    c[k]"= 0;
                    hasMuted:=true;
                else
                    hasMuted:=false;
                end
            end
        end
    end
end
return S_c;
```

Algorithm 2:Proposed mutation operation.

References

Alavian, S.A., Whitson, C.H., Martinsen, S., 2014. Global component lumping for EOS calculations. In: Presented at the SPE Annual Technical Conference and Exhibition, 27-29 October. Society of Petroleum Engineers, Amsterdam, The Netherlands. https://doi.org/10.2118/170912-MS.

Bäck, T., Hoffmeister, F., 1991. Extended Selection Mechanisms in Genetic Algorithms. Morgan Kaufmann, pp. 92–99.

Baker, J.E., 1987. Reducing bias and inefficiency in the selection algorithm. In: Proceedings of the Second International Conference on Genetic Algorithms on Genetic Algorithms and Their Application. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, pp. 14–21.

Barricelli, N.A., 1957. Symbiogenetic evolution processes realized by artificial methods. Methodos 9 (35–36), 143–182.

Behrens, R., Sandler, S., 1988. The use of semicontinuous description to model the C7+ fraction in equation of state calculations. SPE Reservoir Eng. 3 (3), 1041–1047.

Camargo, E., Aguilar, J., Ríos, A., Rivas, F., Aguilar-Martin, J., 2010. Optimization model based on genetic algorithms for oil wells. In: Proceedings of the 9th WSEAS International Conference on Computational Intelligence, Man-machine Systems and Cybernetics. CIMMACS '10. World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA, pp. 131–139.

Coats, K.H., 1985. Simulation of gas condensate reservoir performance. J. Petrol. Technol. 37 (10), 1870–1886.

Danesh, A., hai Xu, D., Todd, A.C., 1992. A grouping method to optimize oil description for compositional simulation of gas-injection processes. SPE Reservoir Eng. 7 (3), 343–348.

De Jong, K.A., 1975. An Analysis of the Behavior of a Class of Genetic Adaptive Systems. Ph.D. thesis. Ann Arbor, MI, USA, aAI7609381.

De la Maza, M., Tidor, B., LAB., M. I. O. T. C. A. I., of Technology Artificial Intelligence Laboratory, M. I, 1991. Boltzmann Weighted Selection Improves Performance of Genetic Algorithms. A.I. Memo. Massachusetts Institute of Technology, Artificial Intelligence Laboratory.

Fraser, A.S., 1957. Simulation of genetic systems by automatic digital computers. i. introduction. Aust. J. Biol. Sci. 10, 448–491.

Glover, F., 1986. Future paths for integer programming and links to artificial intelligence. Comput. Oper. Res. 13 (5), 533–549 applications of Integer Programming.

Glover, F., 1989. Tabu search - part I. ORSA J. Comput. 1 (3), 190–206.

Glover, F., 1990. Tabu search - part II. ORSA J. Comput. 2 (1), 4–32.

Goldberg, D.E., Deb, K., 1991. A comparative analysis of selection schemes used in genetic algorithms. In: Foundations of Genetic Algorithms. Morgan Kaufmann, pp. 69–93.

Gwiazda, T.D., 2006. Genetic Algorithms Reference Vol.1 Crossover for Single-objective Numerical Optimization Problems. Lomianki.

Hancock, P.J.B., 1994. An Empirical Comparison of Selection Methods in Evolutionary Algorithms. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 80–94.

Hansen, P., de Luna Pedrosa Filho, E., Ribeiro, C.C., 1992. Location and sizing of offshore platforms for oil exploration. Eur. J. Oper. Res. 58 (2), 202–214 practical

Combinatorial Optimization.
Holland, J.H., 1975. Adaptation in Natural and Artificial Systems. The University of Michigan Press.
Hustad, O., Dalen, V., 1993. An explicit phase-behavior model for pseudocompositional reservoir simulation. SPE Adv. Technol. 1 (1), 17–26.
Lee, S., Jacoby, R., Chen, W., Culham, W., 1982. Experimental and theoretical studies on the fluid properties required for simulation of thermal processes. Soc. Petrol. Eng. J. 21 (5), 535–550.
Li, Y.-K., Nghiem, L.X., Siu, A., 1985. Phase behaviour computations for reservoir fluids: Effect of pseudo-components on phase diagrams and simulation results. J. Can. Petrol. Technol. 24 (6).
Liao, Y.-H., Sun, C.-T., May, 2001. An educational genetic algorithms learning tool. IEEE Trans. Educ. 44 (2) 20 pp.–.
Luke, S., Spector, L., 1998. A revised comparison of crossover and mutation in genetic programming. In: Genetic Programming 1997: Proceedings of the Second Annual Conference. Morgan Kaufmann, pp. 240–248.
Mitchell, M., 1998. An Introduction to Genetic Algorithms. MIT Press, Cambridge, MA, USA.
Montel, F., Gouel, P., 1984. A new lumping scheme of analytical data for compositional studies. In: SPE Annual Technical Conference and Exhibition, 16-19 September. Society of Petroleum Engineers, Houston, Texas. https://doi.org/10.2118/13119-MS.
Newley, T., Merrill, R., 1991. Pseudocomponent selection for compositional simulation. SPE Reservoir Eng. 6 (4), 490–496.
Peng, D.-Y., Robinson, D.B., 1976. A new two-constant equation of state. Ind. Eng. Chem. Fundam. 15 (1), 59–64.
Riazi, M., Daubert, T., 1980. Simplify property prediction. Hydrocarb. Process. (59), 115=116.
Sarvestani, M.T., Sola, B.S., Rashidi, F., 2012. Genetic algorithm application for matching ordinary black oil PVT data. Petrol. Sci. 9 (2), 199–211.
Soave, G., 1972. Equilibrium constants from a modified redlich-kwong equation of state. Chem. Eng. Sci. 27 (6), 1197–1203.
Souahi, F., Kaabeche, H., 2008. Developing correlations for prediction of petroleum fraction properties using genetic algorithms. Oil Gas Sci. Technol. IFP 63 (2), 229–237.
Spears, W.M., mail Pears, E., Mil, A.N.N., 1992. Crossover or mutation? In: Foundations of Genetic Algorithms 2. Morgan Kaufmann, pp. 221–237.
Turing, A.M., 1950. I. Computing machinery and intelligence. Mind LIX (236), 433–460.
White, D.R., Poulding, S., 2009. A Rigorous Evaluation of Crossover and Mutation in Genetic Programming. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 220–231.
Whitson, C., 1983. Characterizing hydrocarbon plus fractions. Soc. Petrol. Eng. J. 23 (4), 683–694.
Whitson, C., Anderson, T., Søreide, I., 1989. C7+ characterization of related equilibrium fluids using the gamma distribution. Advances in Thermodynamic 1, 35–56.