

Application of neural network to speed-up equilibrium calculations in compositional reservoir simulation

Wagner Q. Barros, Adolfo P. Pires^{*}

Laboratório de Engenharia e Exploração de Petróleo - Universidade Estadual do Norte Fluminense Darcy Ribeiro, Rodovia Amaral Peixoto, Km 163, Imboassica, Macaé, 27925-310, Rio de Janeiro, Brazil

ARTICLE INFO

Keywords:

Neural network
Compositional simulation
Artificial intelligence
Flash calculation
Reservoir engineering

ABSTRACT

Compositional reservoir simulation is an important tool to model fluid flow in oil and gas reservoirs. Important investment decisions regarding oil recovery methods are based on simulation results, where hundred or even thousand of different runs are performed. In this work, a new methodology using artificial intelligence to learn the thermodynamic equilibrium is proposed. This algorithm is used to replace the classical equilibrium workflow in reservoir simulation. The new method avoids the stability test for single-phase cells in most cases and provides an accurate two-phase flash initial estimate. The classical and the new workflow are compared for a gas-oil mixing case, showing a simulation time speed-up of approximately 50%. The new method can be used in compositional reservoir simulations.

1. Introduction

Compositional reservoir simulation is a numerical model used to solve the fluid transport in porous media coupled with thermodynamic phase equilibrium (Fussell and Fussell, 1979; Coats, 1980). To describe the reservoir accurately, numerical grid may be composed of millions of cells, and along a compositional simulation the phase equilibrium is solved for each cell every time-step. Thus, most part of computational effort is spent in flash calculations. There are cases where equilibrium calculations reach more than 50% of simulation time (Okuno et al., 2009; Wang et al., 2019a,b).

Incipient compositional simulators used a K-value approach, where the thermodynamic equilibrium was calculated a priori and stored in tables or correlations of K-values, where K_i is the ratio between the mole fraction of component i in gas and in oil phases ($K_i = y_i/x_i$) (Kniazeff and Naville, 1965; Price and Donohue, 1967; Culham et al., 1969; Roebuck et al., 1969). This approach does not require much memory or processing capacity. The main problem of such model is that K-values may lead to phase composition errors, specially near to critical points.

The increase of computers processing capacity allowed a new compositional simulation technique, based on equations of state (EOS). In this case, the phase equilibrium is solved rigorously, using cubic EOS (Fussell and Fussell, 1979; Coats, 1980; Thele et al., 1983; Okuno et al., 2009; Moortgat and Firoozabadi, 2013; Schmall et al., 2013). This

approach is accurate, and does not cause numerical instability near the critical point. However, $2N_c + 1$ equations for each grid cell must be solved every time step, where N_c is the number of components. Then, for large reservoir grids, computational cost becomes very expensive (Todd and Longstaff, 1972; Wong et al., 1987; El-Banbi et al., 2000). To overcome this limitation, lumping techniques may be used to reduce the number of components in fluid description (Joergensen and Stenby, 1995; Jessen and Moghadam, 2009; Alavian et al., 2014).

For every time-step, appearance and disappearance of a phase must be evaluated in all cells. The phase disappearance is a simple problem (Whitson and Michelsen, 1989; Li and Firoozabadi, 2012), but phase appearance check is not so trivial, and it is a computational time-consuming process.

In the phase appearance algorithm, the stability of a single-phase cell is checked. One technique compares the saturation pressure of the fluid in the cell with the cell pressure (Fussell and Yanosik, 1978; Fussell and Fussell, 1979; Coats, 1980). Modern simulators use the tangent plane stability test to determine if a given single-phase composition is stable (Baker et al., 1982; Michelsen, 1982a; Sun and Seider, 1995; Firoozabadi and Pan, 2000; Li and Firoozabadi, 2012). In both cases we must solve a system of non-linear equations by iterative methods. If the phase is not stable, a two-phase flash calculation is performed to determine the phase composition and saturations in each cell (Fussell and Yanosik, 1978; Boston and Britt, 1978; Hirose et al., 1978; Michelsen, 1982b; Abhmani

^{*} Corresponding author.

E-mail addresses: wagnerqb@gmail.com (W.Q. Barros), puime@lenep.uenf.br (A.P. Pires).

<https://doi.org/10.1016/j.aiig.2022.03.004>

Received 27 December 2021; Received in revised form 28 March 2022; Accepted 30 March 2022

Available online 8 April 2022

2666-5441/© 2022 The Authors. Publishing Services 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/>).

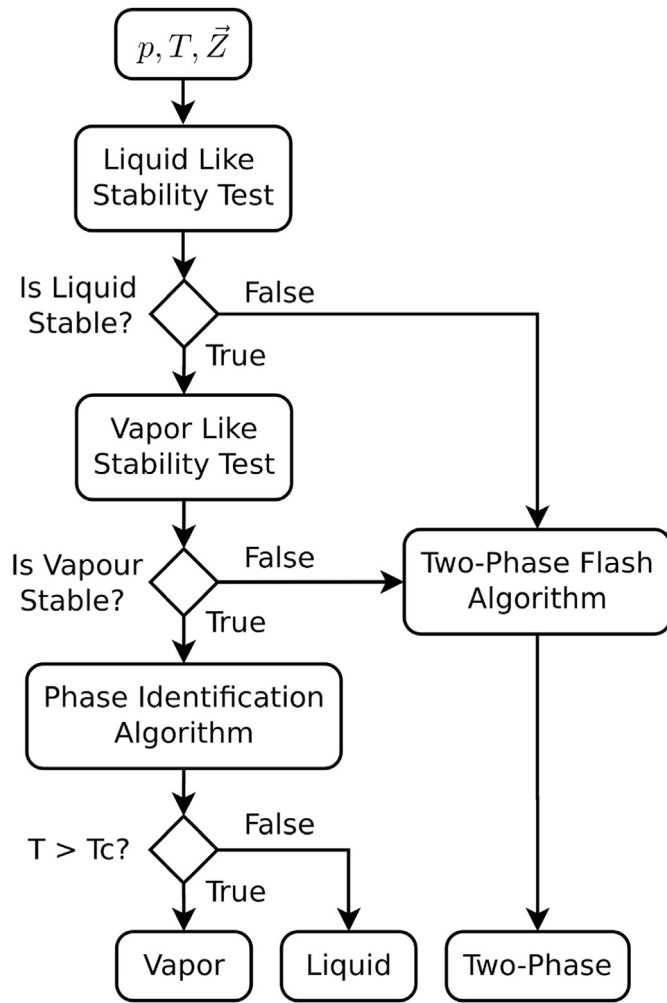
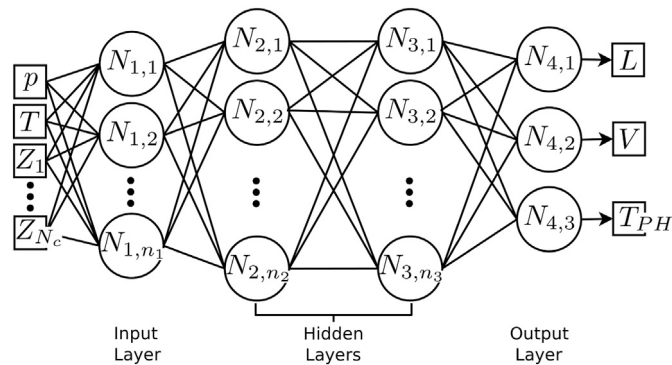


Fig. 1. Classical phase equilibrium workflow.

Fig. 2. Phase identification neural network with 4 layers, where n_1 , n_2 and n_3 are the number of neurons in layers 1, 2 and 3.

and Beaumont, 1987; Li and Johns, 2006).

A technique called Compositional Space Adaptive Tabulation (CSAT) was proposed to decrease the number of stability tests and two-phase flash calculations during reservoir simulations. In this approach, thermodynamic equilibrium data is stored in tables. During the simulation, initial estimates for flash calculations are obtained interpolating the stored data (Voskov and Tchelepi, 2009b,a). This approach provides accurate initial estimates for flash algorithms, decreasing the number of iterations. However, the more components are used in fluid description,

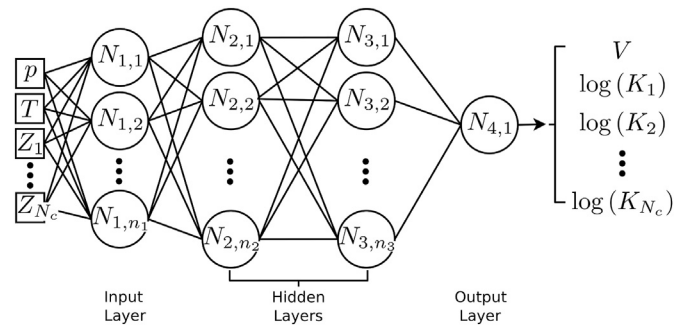
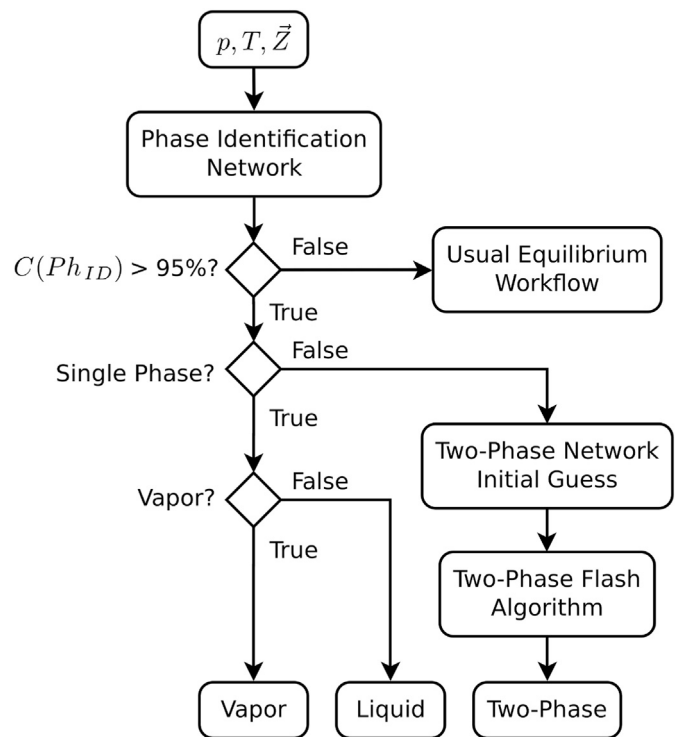
Fig. 3. Two phase initial guess estimate neural network with 4 layers, where n_1 , n_2 and n_3 are the number of neurons in layers 1, 2 and 3.

Fig. 4. Neural network (NN) phase equilibrium workflow.

the more computational time is spent to interpolate the data, decreasing the method efficiency (Rezaveisi et al., 2014).

In this work we propose a similar technique to provide initial estimates for flash calculations, using artificial neural networks instead of tie-line interpolation. A neural network model is based on the human brain and is applied in recognition and classification problems. Mathematically, the neural network is known as a universal function approximator, able to adjust any continuous multivariate function (Ojha et al., 2017). There are many applications of artificial intelligence in thermodynamic equilibrium problems. Goda et al. (2003) trained a neural network to predict black-oil type fluid properties with data from more than 160 different fluids from Middle East reservoirs.

Neural network was also applied to predict CO_2 solubility in ionic liquids (Baghban et al., 2015). Using 728 experimental data from 14 different ionic liquids, a solubility model was built as a function of the critical properties of the pure components, pressure and temperature. The network correlation coefficient was $R^2 = 0.9336$, better than the ones obtained using PR-EOS (Robinson and Peng, 1978) ($R^2 = 0.7323$) and the SRK-EOS (Soave, 1972) ($R^2 = 0.6989$). Baghban et al. (2016a,b) developed several models to predict air dew point and hydrate formation

Table 1
SPE3 fluid characterization.

	\bar{Z}	M_w	P_c	T_c	ω	k_{CO_2}	k_{N_2}	k_{C_1}	k_{C_2}	k_{C_3}
		(g/mol)	(atm)	(K)						
CO_2	0.0121	44.010	72.8	304.2	0.225					
N_2	0.0194	28.013	33.5	126.2	0.040	−0.02000				
C_1	0.6599	16.043	45.4	190.6	0.008	0.10000	0.03600			
C_2	0.0869	30.070	48.2	305.4	0.098	0.13000	0.05000	0.00000		
C_3	0.0591	44.097	41.9	369.8	0.152	0.13500	0.08000	0.00000	0.00000	
C_4	0.0967	58.124	37.5	425.2	0.193	0.12770	0.10020	0.09281	0.00000	0.00000
C_{71}	0.0472	110.900	25.6	572.5	0.332	0.10000	0.10000	0.00000	0.00385	0.00385
C_{72}	0.0153	170.900	16.7	630.2	0.495	0.10000	0.10000	0.00000	0.00630	0.00630
C_{73}	0.0034	282.100	8.5	862.6	0.833	0.10000	0.10000	0.13920	0.00600	0.00600

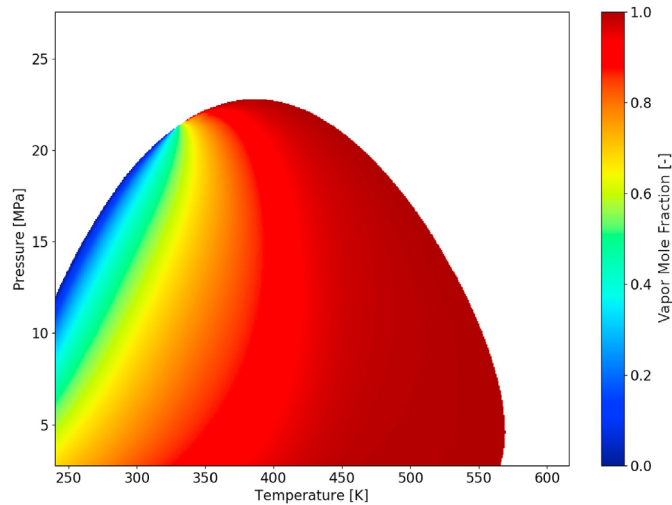


Fig. 5. SPE3 phase envelope filled with vapor phase molar fraction values.

in different scenarios. The enthalpy of vaporization of pure substances based on its boiling temperature, specific gravity and the molecular weight was also modeled using neural networks (Eghtedaei et al., 2019).

Artificial intelligence can be trained with a particular numerical reservoir model generating a proxy that is much faster to evaluate. This proxy model is used to optimize production alternatives and to evaluate uncertainty through Monte Carlo simulations, where thousands of runs are needed to achieve satisfactory results (Mohaghegh et al., 2012; Alakeely and Horne, 2021b). Multiphase flow rate in producing wells can be estimated using deep learning techniques (Tian and Horne, 2017; Li et al., 2019; Alakeely and Horne, 2021a). These artificial intelligence

methods are trained using the production history, and are accurate enough to be used in real-time monitoring tools. A Physics-Informed Neural Network (PINN) can be used to create a meshless reservoir model that can provide accurate results in higher dimensional problems when compared to traditional simulation techniques (Cheung et al., 2018; Coutinho et al., 2021, 2022; Almajid and Abu-alsaud, 2022).

Wang et al. (2019a,b) created a neural network model to predict the bubble point of hydrocarbon mixtures and provide an initial guess for the two-phase flash algorithm. The total reservoir simulation time was reduced in 30%, mainly because the stability test algorithm was replaced by the comparison between the cell pressure and the bubble point predicted by the neural network.

In this work we present a neural network method to provide accurate initial estimates for flash calculations instead of the phase stability test. This procedure may be applied to compositional reservoir simulation saving computational time. The proposed method tests the phase stability and classifies it as a liquid or vapor in the single phase case. In Section 2 we discuss the thermodynamic equilibrium routines, and compare the classical simulation workflow with a new one, based on neural network. Then, in Section 3 the new technique is evaluated, checking the accuracy of the predicted thermodynamic calculations and comparing the computational time between the classical approach and the one proposed in this work. Finally, some conclusions are addressed in Section 4.

2. Two-phase equilibrium

This section presents the mathematical description of the flash equilibrium problem and describes how artificial intelligence can be used to improve these routines.

A closed system at constant pressure, temperature and global composition is at equilibrium if the total Gibbs free energy (G) is minimal:

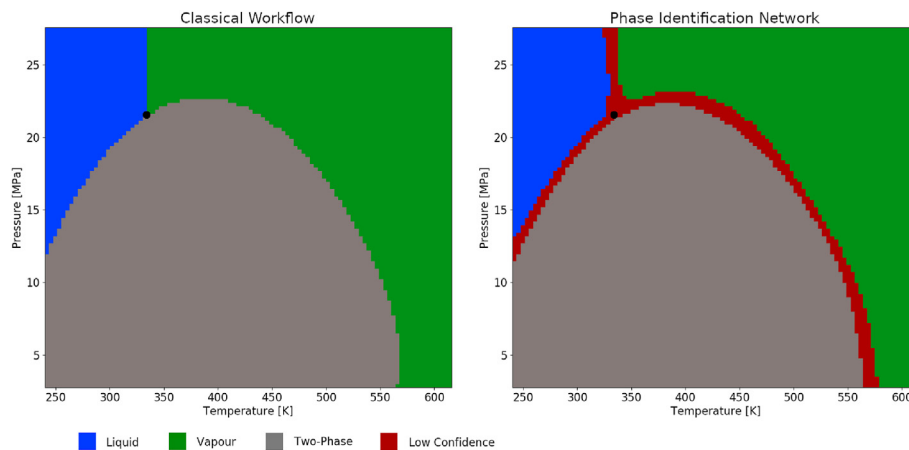


Fig. 6. SPE3 phase envelope with phase identification labels generated by the classical workflow and by the neural network workflow, where the black dot marks the critical point.

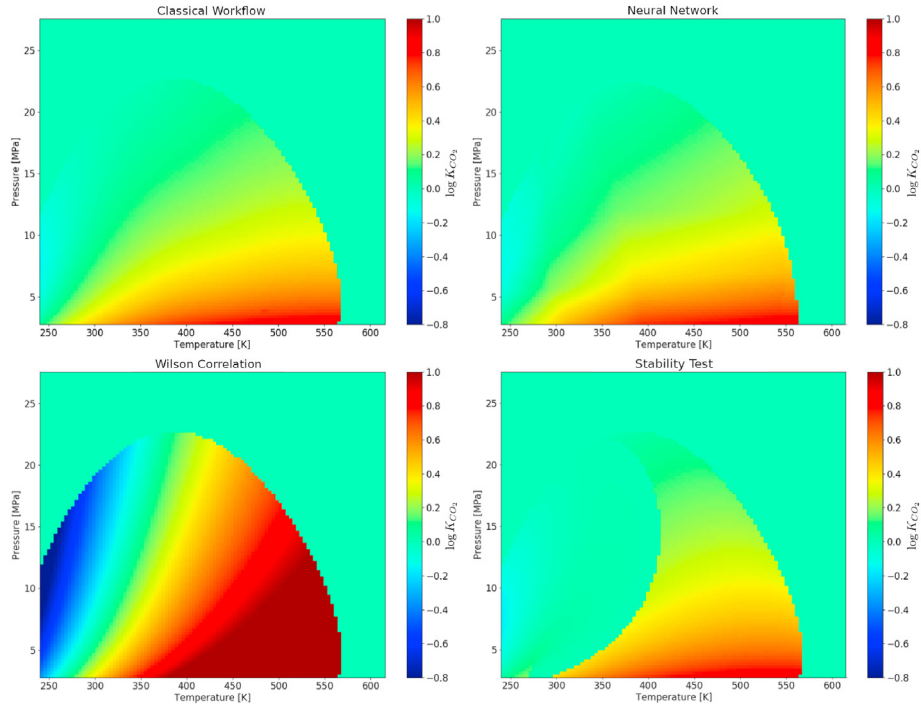


Fig. 7. Phase envelope with $\log K_{CO_2}$ values: (a) equilibrium values, (b) estimated by the neural network, (c) estimated by Wilson correlation, (d) estimated by the stability test.

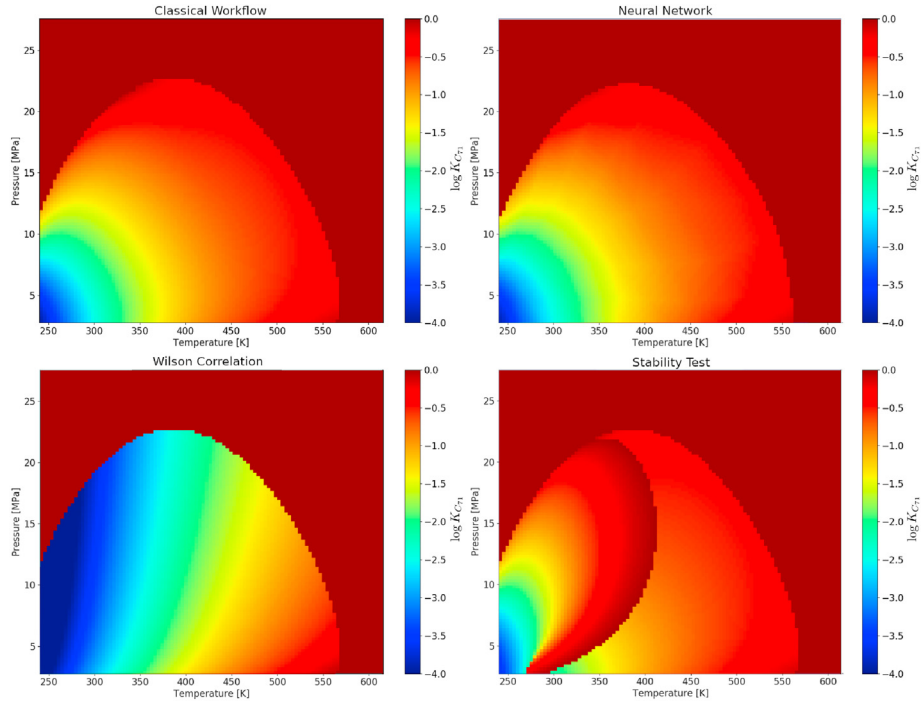


Fig. 8. Phase envelope with $\log K_{C_{71}}$ values: (a) equilibrium values, (b) estimated by the neural network, (c) estimated by Wilson correlation, (d) estimated by the stability test.

$$dG = \sum_{\pi=1}^{N_{\pi}} \left[\left(\frac{\partial G}{\partial p} \right)_{T, \eta_{c,\pi}} dp + \left(\frac{\partial G}{\partial T} \right)_{p, \eta_{c,\pi}} dT + \sum_{c=1}^{N_c} \left(\frac{\partial G}{\partial \eta_{c,\pi}} \right)_{T,p,\eta_{j \neq c, \pi}} d\eta_{c,\pi} \right] = 0 \quad (1)$$

where $\eta_{c,\pi}$ is the number of moles of component c in phase π and N_c and

N_{π} are the total number of components and phases.

2.1. Phase identification workflow

In compositional reservoir simulation, the phase stability test (Michelsen, 1982a; Firoozabadi and Pan, 2000; Li and Firoozabadi, 2012) is performed in single phase cells to evaluate if the Gibbs free

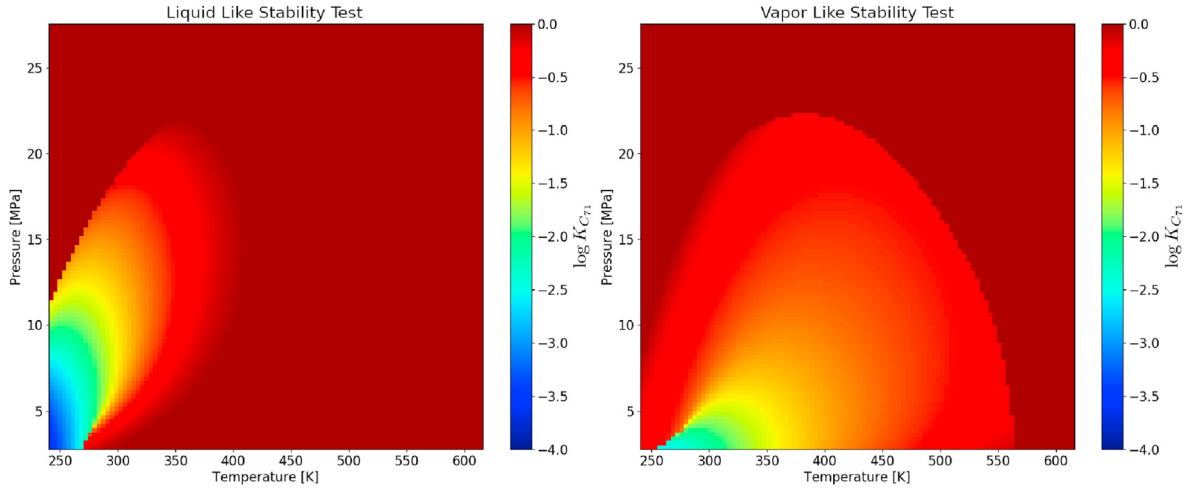


Fig. 9. Phase envelope filled with $\log K_{C_{71}}$ values after convergence of stability test, where: (a) starting with liquid like composition and (b) starting with vapor like composition.

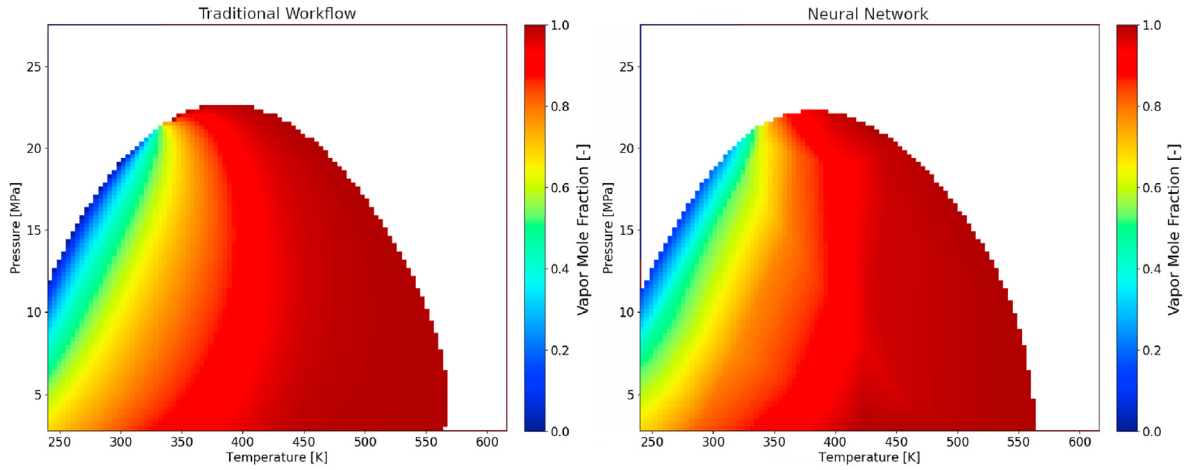


Fig. 10. Phase envelope with vapor phase mole fraction: (a) calculated by the EOS, (b) initial guess estimated by the neural network.

Table 2
Swelling experiment composition of SPE3 oil and its gas.

	Oil Mole Fraction (\bar{x})	Gas Mole Fraction (\bar{y})
CO_2	0.0121	0.0129
N_2	0.0194	0.0206
C_1	0.6599	0.7016
C_2	0.0869	0.0922
C_3	0.0591	0.0623
C_4	0.0967	0.0993
C_{71}	0.0472	0.0110
C_{72}	0.0153	0.0001
C_{73}	0.0034	0.0000

energy is minimal (stable) or if the phase splits in two (unstable). The difference of the Gibbs free energy between a two phase system and a reference phase is given by:

$$\Delta G = (G^I + G^{II}) - G_0 \quad (2)$$

where G^I and G^{II} are the Gibbs free energy of phases I and II , and G_0 is the Gibbs free energy of the reference phase with composition \bar{Z} . Considering that the reference phase splits in two, where the new phase has an

infinitesimal mole number (ϵ) and composition $\bar{\zeta}$, Equation (2) can be rewritten in molar basis as:

$$\Delta G = \epsilon \sum_{i=0}^{N_c} \zeta_i (\mu_i(p, T, \bar{\zeta}) - \mu_i(p, T, \bar{Z})) \quad (3)$$

in which ζ_i is the component i mole fraction and μ_i is the chemical potential of component i , given by $\mu_i = \left(\frac{\partial G}{\partial n_i} \right)_{p, T, n_{j \neq i}}$.

Defining $F(p, T, \bar{\zeta})$ by:

$$F(p, T, \bar{\zeta}) = \sum_{i=0}^{N_c} \zeta_i (\mu_i(p, T, \bar{\zeta}) - \mu_i(p, T, \bar{Z})) \quad (4)$$

it is known that if $F \geq 0$ for any possible $\bar{\zeta}$ composition, then the reference phase \bar{Z} is stable as a single-phase mixture. Instead of evaluating this equation for all possible $\bar{\zeta}$ compositions, the stability test searches for critical points, where the Gibbs free energy is a local minimum. A critical point can be found using a direct substitution method (Michelsen, 1982a):

$$\ln Y_i^{t+1} = \ln Z_i + \ln \varphi_i(P, T, \bar{Z}) - \ln \varphi_i(P, T, \bar{\zeta}^t) \quad (5)$$

where φ_i is the fugacity coefficient of component i , calculated using any

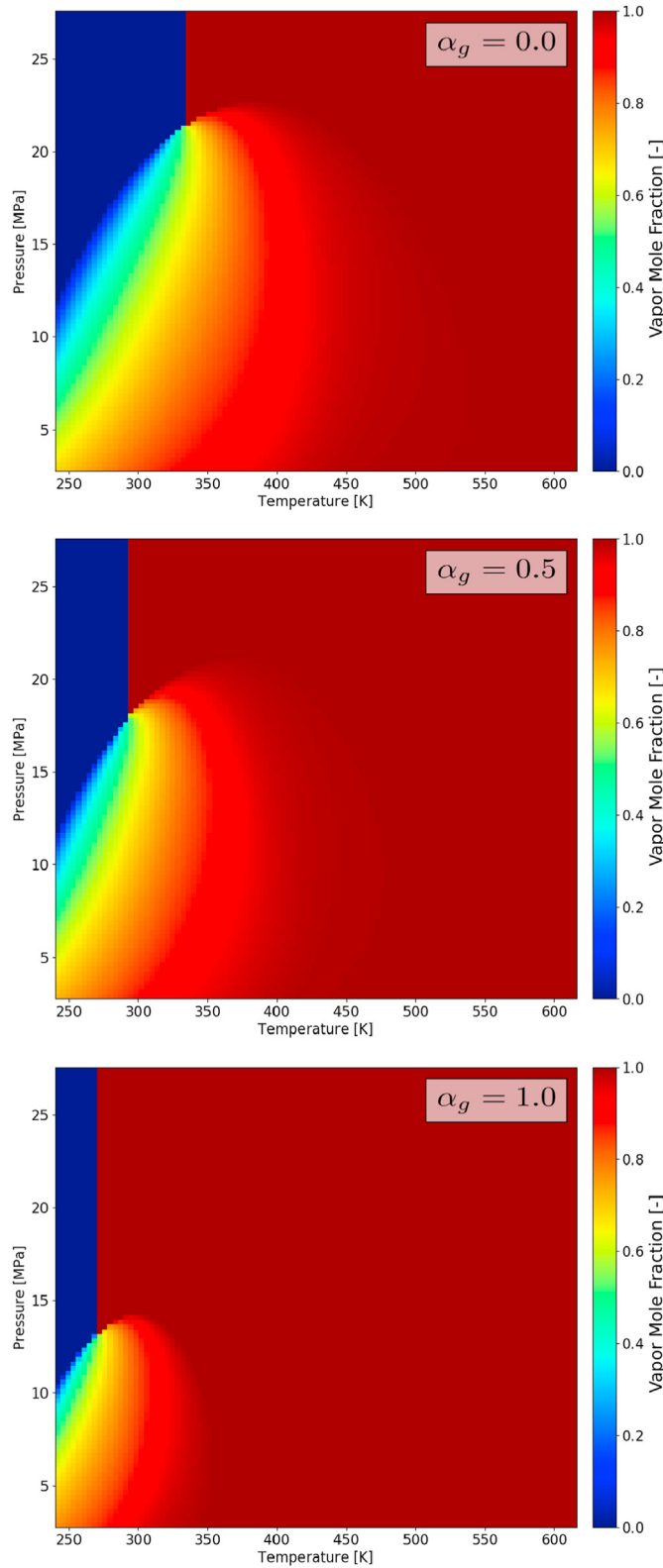


Fig. 11. Phase envelope for different gas fractions (α_g).

equation of state. The composition in step $t + 1$ is found by $\zeta_i^{t+1} = \frac{y_i}{\sum_{j=0}^{N_c} y_j}$.

The initial iteration ζ^0 for hydrocarbons is usually calculated using Wilson's correlation (Wilson, 1968):

$$K_i = \frac{p_{c,i}}{p} \exp \left[5.37(1 + \omega_i) \left(1 - \frac{T_{c,i}}{T} \right) \right] \quad (6)$$

where $p_{c,i}$, $T_{c,i}$ and ω_i are the critical pressure, critical temperature and acentric factor of component i . The K-value relates the vapor (y_i) and liquid (x_i) molar fractions:

$$K_i = \frac{y_i}{x_i} \quad (7)$$

then, if a liquid-like phase is evaluated the initial trial composition is given by $\zeta_i = \frac{z_i}{K_i}$, and if a vapor-like phase is tested the initial trial composition is $\zeta_i = K_i z_i$.

If the phase is stable a phase identification algorithm to determine if it is a liquid or gas must be performed. When the system temperature T is greater than the mixture critical temperature ($T > T_c$) the phase is labeled as a vapor, otherwise as a liquid. The critical temperature depends on the mixture composition and requires great computational effort. In compositional reservoir simulation, the critical temperature is usually approximated by the modified Li mixing rule (Li, 1971):

$$T_c = r_1 \frac{\sum_{i=0}^{N_c} x_i V_{c,i} T_{c,i}}{\sum_{i=0}^{N_c} x_i V_{c,i}} \quad (8)$$

where $V_{c,i}$ is the critical volume of component i and r_1 is an adjustable parameter used to tune this algorithm.

The previously described process will be called classical phase equilibrium algorithm, which is used in EOS based compositional reservoir simulators (Fig. 1). Note that the fluid is only considered single-phase after two stability tests, for both liquid and vapor initial estimates, and then performing a phase identification algorithm. Along a full field compositional reservoir simulation, this routine is called for each grid cell every time step, making this algorithm responsible for about 50% of processing time, where the stability test is the most consuming part (Wang et al., 2019a,b).

2.2. Neural network applied to phase identification

To decrease the number of calls of the stability test and phase labeling algorithms we developed a neural network specialized in phase identification (Fig. 2). The Multi Layer Perceptron (MLP) was the network layout chosen. In Fig. 2, each circle is a perceptron neuron and each layer is directly connected with the next (Haykin, 1999). For the phase identification network, all neurons have the same sigmoidal activation function, given by:

$$N(\vec{x}) = \frac{1}{1 + e^{-(\vec{x} \cdot \vec{w} - b)}} \quad (9)$$

where \vec{x} is the neuron input vector (lines in Fig. 2), \vec{w} is the neuron weight vector and b the constant bias. Note that the neuron output is limited by the interval $N(\vec{x}) \in (0, 1)$. So, training a network is an optimization problem to define the values of \vec{w} and b of each neuron that minimize the error between the network response and the output for a given training data-set.

For the phase identification network, defining as input the pressure, temperature and global composition (p, T, \vec{Z}), this algorithm returns three real numbers in the interval $(0, 1)$: liquid (L), vapor (V) or two-phase (T_{PH}). The phase identification (Ph_{ID}) is performed through a voting system, in which the neural network response is the greatest value among the three outputs:

$$Ph_{ID} = \max(L, V, T_{PH}) \quad (10)$$

To check if the network response is valid we estimate the confidence of Ph_{ID} by:

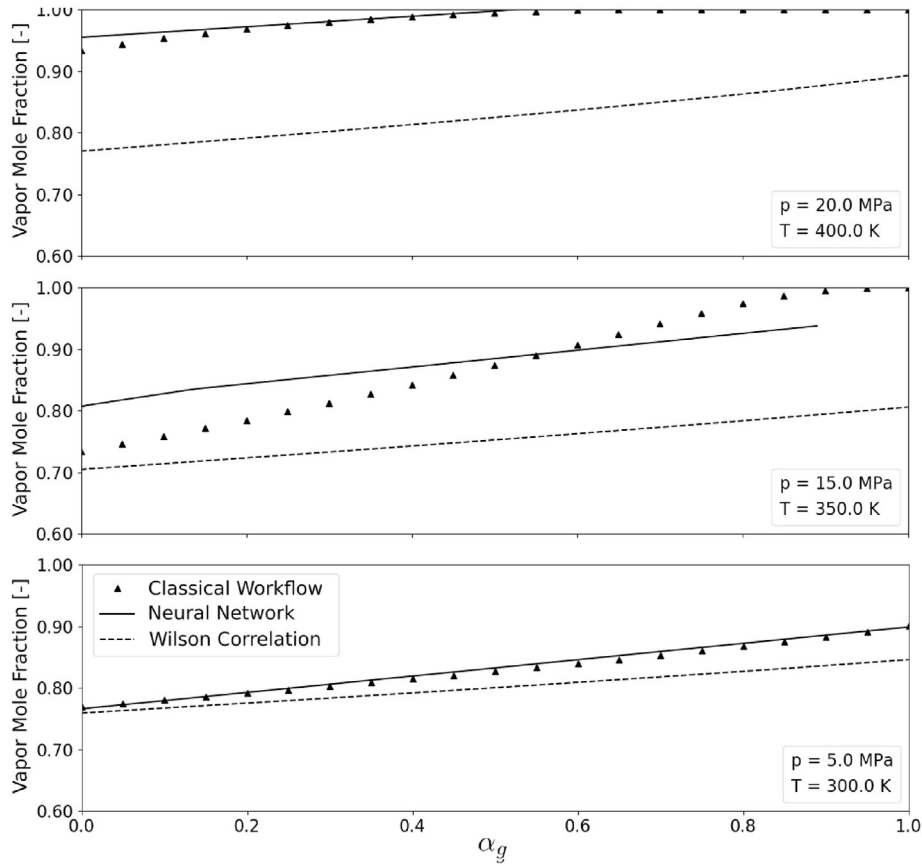


Fig. 12. Vapor mole fraction calculated by an EOS and estimated by the neural network and by Wilson correlation for three different (p, T) conditions and gas fractions.

$$C(Ph_{ID}) = \frac{\max(L, V, T_{PH})}{(L + V + T_{PH})} \quad (11)$$

The phase identified by the neural network is accepted if its confidence is greater than 95%, otherwise the traditional phase identification workflow is used (Section 2.1).

The introduction of the confidence concept and labeling it as a liquid or vapor in the single phase case are some improvements of the technique presented in this work when compared with other methods reported in literature (Wang et al., 2019a,b). Its importance raises when miscible processes are simulated, because the fluid phase may change from liquid to vapor without crossing the two phase boundary or even crossing the two-phase envelope. Thus, the use of the proposed method reduces the number of stability tests calls.

2.3. Two-phase flash algorithm

The phase equilibrium conditions for a two-phase system is given by (Prausnitz et al., 1998; Sandler, 2006):

$$\begin{cases} p^l = p^v \\ T^l = T^v \\ x_1 \varphi_1(p, T, \vec{x}) = y_1 \varphi_1(p, T, \vec{y}) \\ \vdots \\ x_{N_c} \varphi_{N_c}(p, T, \vec{x}) = y_{N_c} \varphi_{N_c}(p, T, \vec{y}) \end{cases} \quad (12)$$

where φ_c is the fugacity coefficient of a component in a mixture, calculated using any equation of state and the superscripts l and v relate the liquid and vapor phases. The fugacity condition equation for each component can be rewritten in a residual form:

$$\log(\varphi_i(p, T, \vec{y})) - \log(\varphi_i(p, T, \vec{x})) + \log(K_i) = R_i, \quad i = 1, \dots, N_c \quad (13)$$

and solved using the Newton-Raphson's method. Usually the K -values obtained at the end of the phase stability test (Equation (5)) are used as the initial estimate. After convergence the vapor phase mole fraction (V) is determined solving the Rachford-Rice equation (Rachford and Rice, 1952):

$$\sum_{i=1}^{N_c} \frac{Z_i(K_i - 1)}{1 + (K_i - 1)V} = 0 \quad (14)$$

and the composition of each phase is calculated through:

$$\begin{cases} x_i = \frac{Z_i}{1 + (K_i - 1)V} \\ y_i = \frac{Z_i K_i}{1 + (K_i - 1)V} \end{cases}, \quad i = 1, \dots, N_c \quad (15)$$

2.4. Neural network to estimate initial flash calculations

When the phase identification neural network is used to determine the number of phases in the system, the phase stability algorithm is not run and initial estimates for the two phase flash algorithm are not available (Equation (13)). To overcome this limitation, artificial intelligence is used to generate initial estimates to Equation (13).

For this problem, one MLP network is trained for each flash variable ($V, \log(K_1), \dots, \log(K_{N_c})$) (Fig. 3). The input is the same of the stability test and phase identification (p, T, \vec{Z}) and there is one output neuron in each network. The output neuron uses the linear activation function:

$$N(\vec{x}) = \vec{x} \cdot \vec{w} - b \quad (16)$$

while all other neurons of the input and hidden layers use the rectified

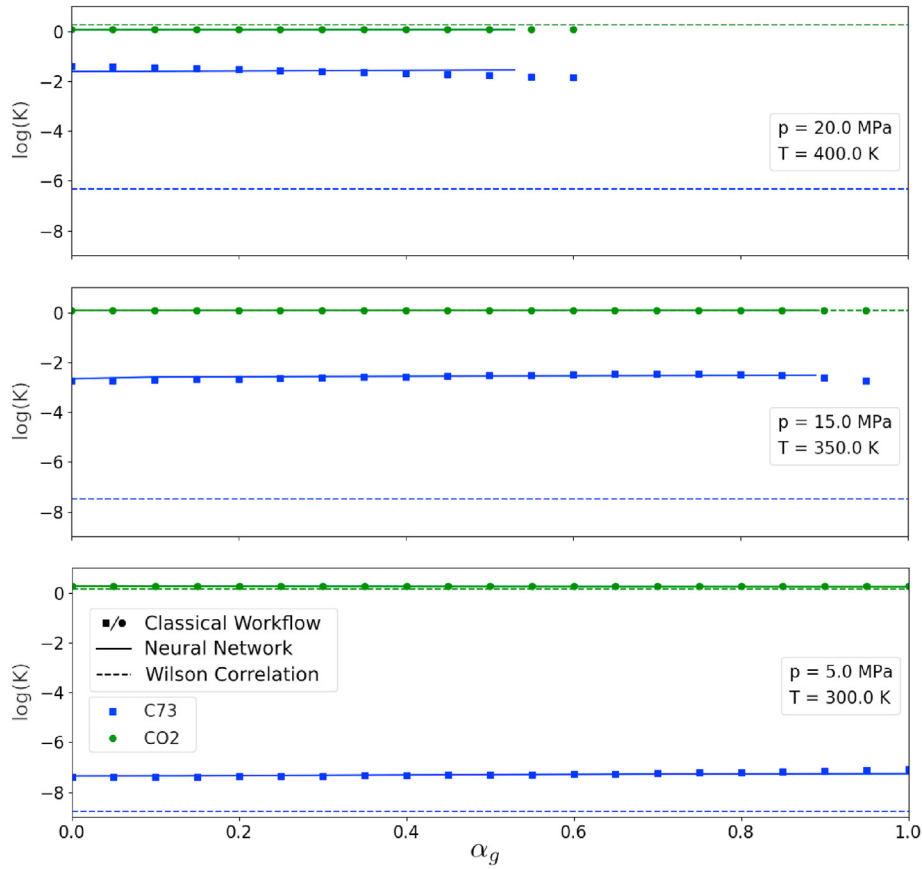


Fig. 13. $\log K_{CO_2}$ and $\log K_{C73}$ calculated by an EOS and estimated by the neural network and by Wilson correlation for three different (p, T) conditions and gas fractions.

linear activation function:

$$N(\vec{x}) = \max(0, \vec{x} \cdot \vec{w} - b) \quad (17)$$

It should be noted that a different neural network for each flash variable ($V, \log(K_1), \dots, \log(K_N)$) must be built. This architecture is based on the specialization concept, where the reduction of the number of output variables decreases the size of the optimization problem and increases its accuracy (Haykin, 1999).

2.5. Improved two-phase equilibrium model

Fig. 4 shows the proposed phase equilibrium workflow that uses artificial intelligence to replace part of the classical flash calculations (Fig. 1). The first step uses the phase identification network to generate an estimate of the confidence and identify the fluid phase (Equation (11)). If $C(Ph_{ID}) < 95\%$ the result is not considered and the usual stability test is performed. Otherwise, the flash parameters using the estimate network for the two-phase region are generated (Fig. 3).

The proposed methodology does not replace completely the flash calculation in compositional simulators. It avoids stability tests and provides initial estimates very close to the solution of the flash. As it will be discussed in the next section, the neural network workflow, combined with the confidence calculation, ensures that the exact solution of two-phase equilibrium will be found.

3. Results and discussion

The new phase equilibrium workflow (Fig. 4) was compared with the classical workflow (Fig. 1) for the SPE3 mixture (Kenyon and Behie, 1987; Haugen and Beckner, 2013). Table 1 presents the SPE3 fluid

characterization and Fig. 5 shows its phase envelope, simulated with the Peng-Robinson equation of state (Robinson and Peng, 1978). The k_i value is the binary interaction coefficient between the row component and i component. This fluid can be characterized as a light oil with 51° API and 850 sm^3/sm^3 solution gas-oil ratio.

This section is divided in three parts: in the first two sub-sections we check if the artificial intelligence algorithm learns two-phase thermodynamic equilibrium results from constant global composition calculations and from a gasflooding recovery process. Last section compares the computational time between the proposed neural network workflow and the classical workflow.

3.1. Constant composition

For the constant composition case 10,000 classical flash calculations for SPE3 fluid (Table 1) were performed changing pressure in the interval $p \in [2, 27]$ MPa and temperature in the interval $T \in [240, 620]$ K. These results were used as the training data set.

3.1.1. Phase identification network

The best phase identification network topology is composed by 4 layers with a total of 43 neurons, whose structure is (8, 16, 16, 3), where each number indicates the number of neurons per layer. This optimum topology was found by testing different architectures to maximize the network accuracy using the minimum number of layers and neurons.

Fig. 6 compares the phase identification envelope generated by the classical workflow (used as training data) and the envelope obtained by the neural network after training. Note that the uncertainty points (red) appear in the phase transition region, and are less than 5.23% of total.

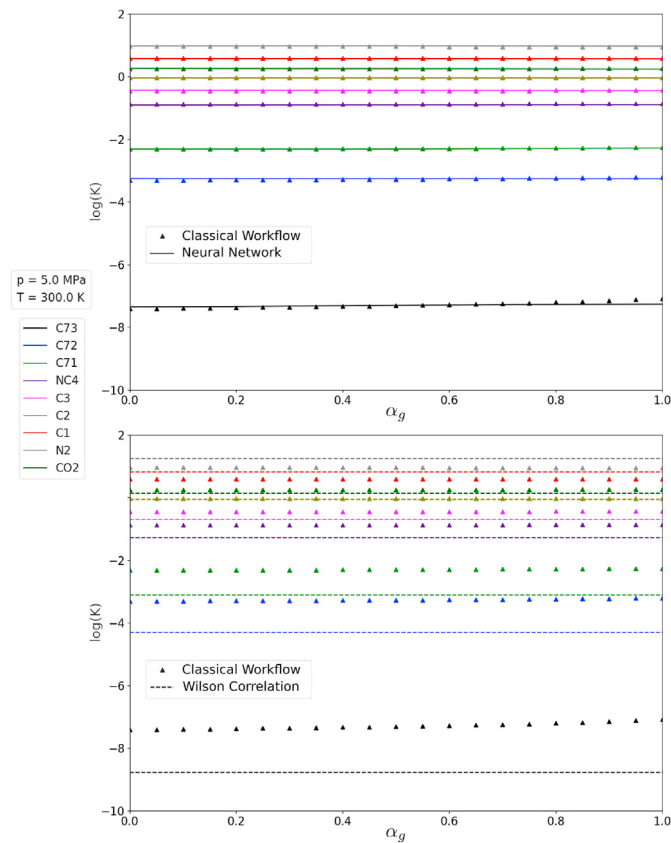


Fig. 14. $\log K_i$ calculated by an EOS and estimated by the neural network and by Wilson correlation for $p = 5$ MPa, $T = 300$ K and several gas fractions.

3.1.2. Two-phase flash initial estimate network

The topology of the network that generates initial estimates for flash calculations is similar to the topology of the phase identification network, with 41 neurons distributed in 4 layers (8,16,16,1). The Newton-Raphson method depends on the initial estimates. So, to reduce the number of iterations of the flash algorithm (Equation (13)), it is important to provide an estimate of $\log K_i$ for each component as close as possible to the solution. As an example, we compare the $\log K_{CO_2}$ and $\log K_{C_{71}}$ obtained after equilibrium by the classical workflow, and estimated by the proposed neural network method, the stability test (Equation (5)), and the Wilson correlation (Equation (6)) (Figs. 7 and 8).

The neural network can learn the K-values of all components, generating initial estimates for the flash problem very close to the exact solution, and reducing the number of flash iterations. On the other hand Wilson correlation provides poor estimates at high pressures causing divergence when applied directly in the flash algorithm. Note that there is a discontinuity in the calculations using the stability test.

The discontinuous region in stability calculations is related with the local minima in Equation (4). Depending if the algorithm starts the tests with liquid like or vapor like composition ($\bar{Z} = \bar{x}$ or $\bar{Z} = \bar{y}$), the stability algorithm may converge to different $\log K_i$ (Fig. 9). Fig. 9 presents $\log K_{C_{71}}$ after the convergence of liquid like and vapor like stability tests. Note that there is an area where both algorithms result in unstable phases with different compositions. Choice of vapor like composition in the region of high liquid fractions (low temperature) can result in poor initial estimates for Newton-Raphson algorithm. Note that the unstable region for either liquid or vapor like stability tests are always smaller than the phase envelope region, the main reason it is necessary to run both tests in the classical workflow (Fig. 1).

Finally, Fig. 10 compares the vapor mole fraction calculated by the EOS and the initial guess provided by the neural network. The initial

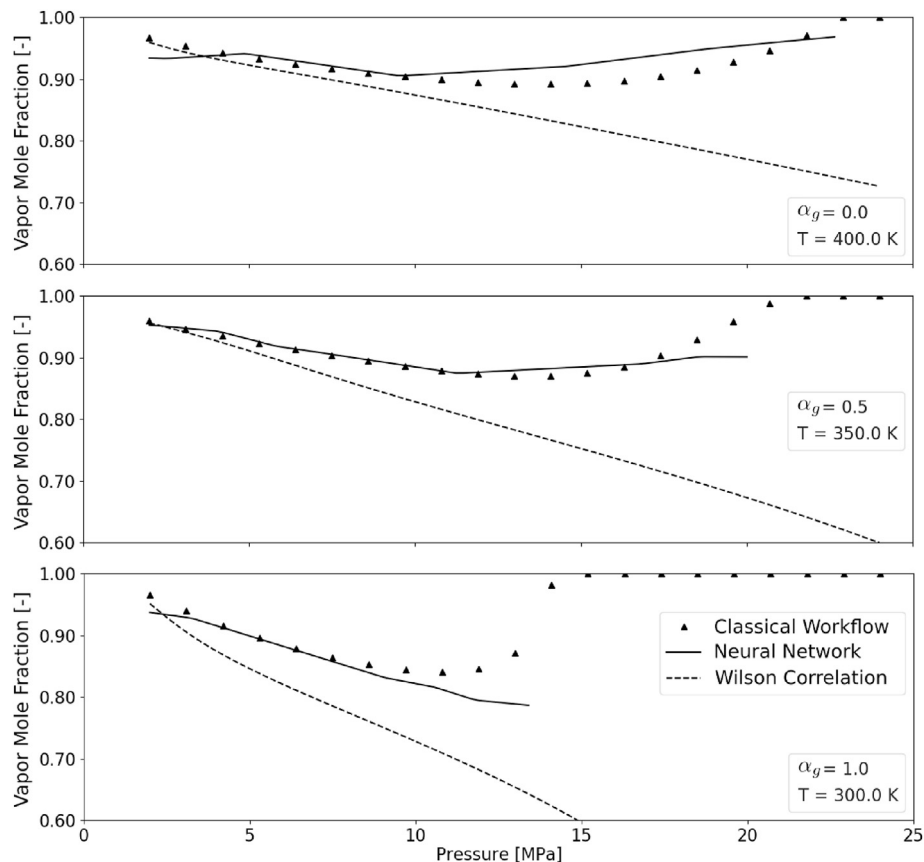


Fig. 15. Vapor mole fraction calculated by an EOS and estimated by the neural network and by Wilson correlation for three different (α_g, T) conditions and different system pressures.

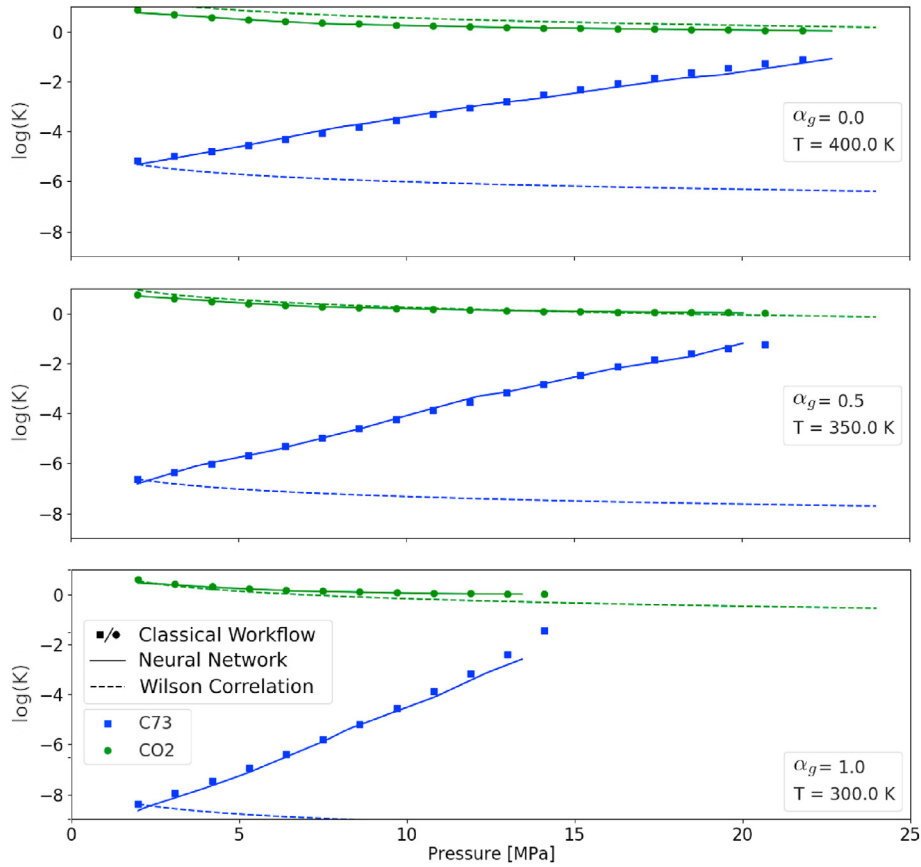


Fig. 16. $\log K_{CO_2}$ and $\log K_{C_{73}}$ calculated by an EOS and estimated by the neural network and by Wilson correlation for three different (α_g, T) conditions and different system pressures.

guess is close to the flash by the EOS, an excellent approximation for the flash calculation even in regions close to the critical point.

3.2. Variable composition

Gas injection is a recovery process of liquid hydrocarbons that changes the global composition of the reservoir fluid. It increases numerical simulation complexity because of the mass transfer. For example, the composition and pressure variation in small cells close to injection well may change the phase from oil to gas in one time step. So, it is necessary to perform the stability test every time step using liquid-like and vapor-like estimates.

To train the proposed neural network workflow for gasflooding simulations scenario a numerical swelling test was created. This test consists of mixing gas and oil at fixed ratios, and measuring the volume and composition of each phase after equilibrium.

The training data-set for the variable composition case was the mixture of SPE3 oil (Table 1) with its own dissolved gas after a flash separation at standard conditions (Table 2). We created 11 different gas fractions (α_g), from pure oil ($\alpha_g = 0$) to pure gas ($\alpha_g = 1$) at constant steps of $\alpha_g = 0.1$ (Fig. 11). The same pressure and temperature regions ($p \in [2, 27]$ MPa and $T \in [240, 620]$ K) were used for each mixture. Note that for a fixed gas fraction, the global composition is calculated by:

$$\vec{Z} = (1 - \alpha_g) \vec{x} + \alpha_g \vec{y} \quad (18)$$

where \vec{x} and \vec{y} are the liquid and gas compositions (Table 2). The best network topology for the swelling case is a (9, 27, 27, 3) structure for the phase identification network and (9, 27, 27, 1) for each initial estimate for flash calculation. These networks require more neurons when

compared with the constant composition case because of the variation of the global composition input.

First we compare the vapor mole fraction (after the convergence of the flash algorithm using an EOS) with the estimates given by the neural network and the result of the Rachford-Rice equation using Wilson correlation (Equations (14) and (6)) for three different pressure and temperature, changing the gas fraction (Fig. 12). As the system pressure increases, Wilson correlation results are poorer, increasing the number of iterations of the stability algorithm (Equation (5)). On the other hand, the neural network avoids the stability test and provides excellent initial estimates for the two-phase flash algorithm. The missing points in the neural network curve for $\alpha_g > 0.9$, $p = 15$ MPa and $T = 350$ K are caused by the low confidence in the phase identification network.

Fig. 13 presents similar results for $\log K_{CO_2}$ and $\log K_{C_{73}}$ and Fig. 14 shows all $\log K_i$ values for a single p, T condition. For CO_2 Wilson correlation can generate good results, however it fails for heavy components, like C_{73} . The neural network provides good estimates for $\log K_i$ in all temperatures and pressures for all components. In many cases the flash calculations can be replaced by the neural network estimates. For $\alpha_g > 0.6$ at $p = 20$ MPa and $T = 400$ K, the mixture becomes single phase.

Figs. 15 and 16 compare $V, \log K_{CO_2}$ and $\log K_{C_{73}}$ for three different (α_g, T) conditions changing system pressure, while Figs. 17 and 18 compare the same variables for three different (α_g, p) conditions changing system temperature. It is clear that Wilson correlation does not provide good estimates for high (p, T) systems, leading to many iterations until convergence of stability tests in the classical workflow, and slowing compositional simulations. Use of proposed technique avoids the stability test for almost 95% of the cases, and generates an initial estimate for the flash algorithm very close to the solution.

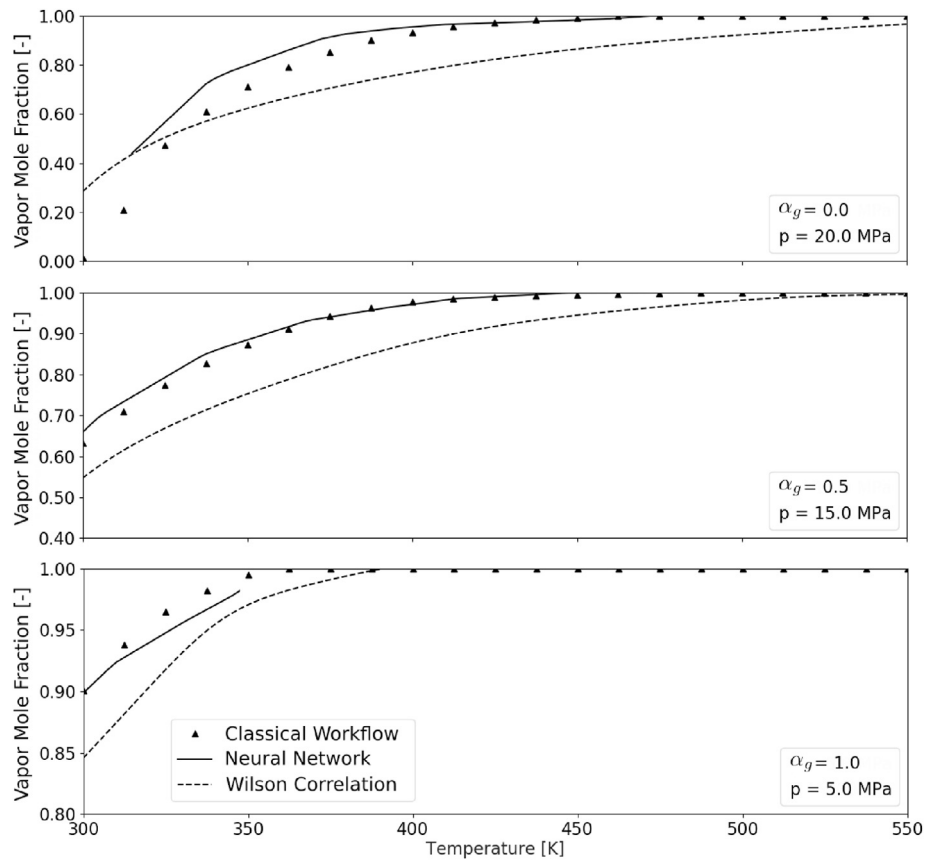


Fig. 17. Vapor mole fraction calculated by an EOS and estimated by the neural network and by Wilson correlation for three different (α_g, p) conditions and different system temperatures.

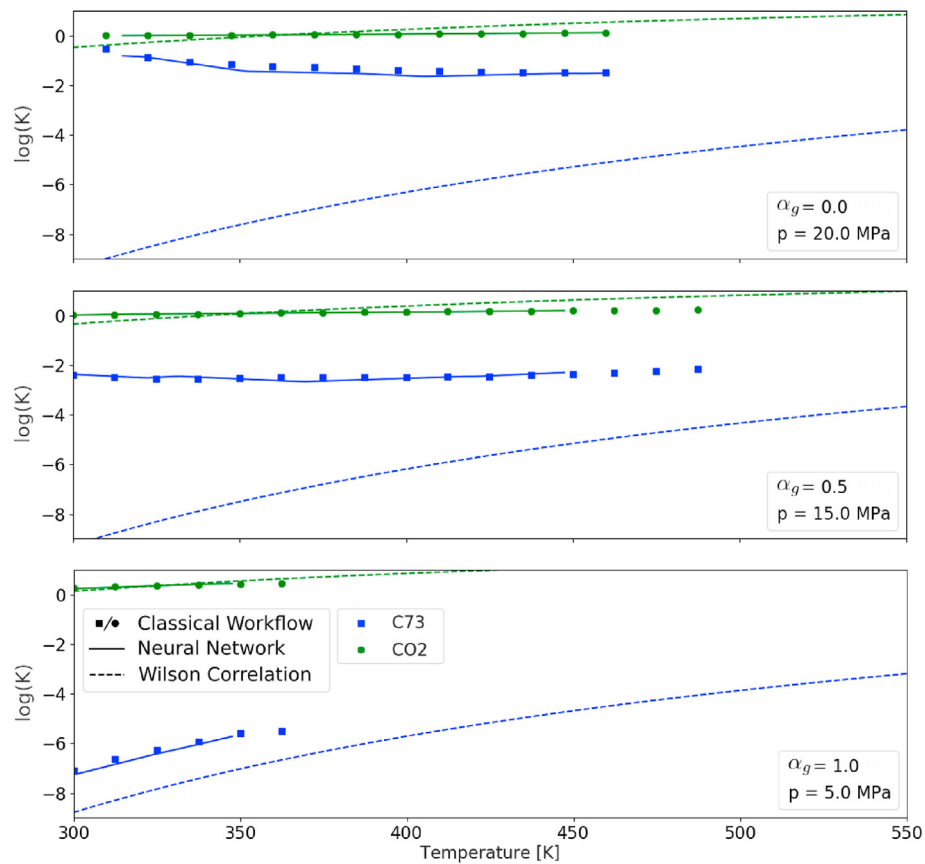


Fig. 18. $\log K_{CO_2}$ and $\log K_{C_{73}}$ calculated by an EOS and estimated by the neural network and by Wilson correlation for three different (α_g, p) conditions and different system temperatures.

Table 3

Number of iterations for stability test and flash algorithms for the classical and neural network workflows, and computational time decrease.

Grid Size	Vapor Fraction	Stability Iterations		Flash Iterations		Execution Time (%)
		EOS	NN + EOS	EOS	NN + EOS	
20x20	$\alpha_v = 0.15$	36,591	4922	4601	1716	–57.4%
40x40	$\alpha_v = 0.75$	105,356	24,796	21,532	10,400	–49.8%
100x100	$\alpha_v = 0.25$	1,148,991	243,202	112,588	43,375	–54.3%
200x200	$\alpha_v = 0.55$	3,581,620	927,840	421,225	199,950	–48.7%
400x400	$\alpha_v = 0.35$	16,936,843	3,910,518	1,844,125	786,628	–42.0%

3.3. Speed-up results

The first two sections showed that it is possible to use artificial intelligence to learn thermodynamic equilibrium variables. As discussed, the proposed equilibrium workflow avoids the stability test for most cases and provides initial estimates for the flash calculation very close to the solution. For the tested cases, the computational time spent in stability test is reduced by approximately 75%, and in the flash algorithm by 50%.

Table 3 shows the number of iterations for some cases. These cases where generated changing system pressure and temperature (grid size) and calling both workflows for each one. The processing time reduction was greater than 40% in all cases. For the flash algorithm, the neural network generated initial estimates so close to the final result that in some cases the solution was obtained after one iteration.

4. Conclusions

In this work we propose a new phase equilibrium workflow using artificial intelligence to decrease the number and in most cases avoid the stability test and speed-up the traditional flash algorithm. The developed technique generates initial estimates for the flash calculation close to the solution. This workflow may be applied in compositional reservoir simulators, where the classical flash equilibrium calculations are performed for each cell every time-step.

It was shown that small neural networks are necessary to learn the phase equilibrium parameters, and the time spent to evaluate this network is much smaller than the reduction of the flash iterations and avoiding stability calls. The neural network and the final flash results are so close that in many cases the flash routine is not necessary. However, we built a robust workflow where the accuracy of equilibrium equations is achieved in every run, avoiding error propagation in reservoir simulation. We compared the processing time between the classical and the neural network workflows, the use of artificial intelligence decreased the time around 50%.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge Universidade Estadual do Norte Fluminense (UENF) for financial support. This study was also supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

- Abhvan, A.S., Beaumont, D.N., 1987. Development of an efficient algorithm for the calculation of two-phase flash equilibria. *SPE Reservoir Eng.* 2, 695–702. <https://doi.org/10.2118/13951-pa>.

- Alakeely, A., Horne, R., 2021a. Application of deep learning methods to estimate multiphase flow rate in producing wells using surface measurements. *J. Petrol. Sci. Eng.* 205. <https://doi.org/10.1016/j.petrol.2021.108936>.
- Alakeely, A.A., Horne, R.N., 2021b. Simulating multiphase flow in reservoirs with generative deep learning. In: *SPE Annual Technical Conference and Exhibition, SPE*. <https://doi.org/10.2118/206126-ms>.
- Alavian, S.A., Whitson, C.H., Martinsen, S.O., 2014. Global component lumping for EOS calculations. In: *SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers*. <https://doi.org/10.2118/170912-ms>.
- Almajid, M.M., Abu-Isa, M.O., 2022. Prediction of fluid flow in porous media using physics informed neural networks. *J. Petrol. Sci. Eng.* 208. <https://doi.org/10.1016/j.petrol.2021.109205>.
- Baghban, A., Ahmadi, M.A., Shahraki, B.H., 2015. Prediction carbon dioxide solubility in presence of various ionic liquids using computational intelligence approaches. *J. Supercrit. Fluids* 98, 50–64. <https://doi.org/10.1016/j.supflu.2015.01.002>.
- Baghban, A., Bahadori, M., Rozyn, J., Lee, M., Abbas, A., Bahadori, A., Rahimani, A., 2016a. Estimation of air dew point temperature using computational intelligence schemes. *Appl. Therm. Eng.* 93, 1043–1052. <https://doi.org/10.1016/j.applthermaleng.2015.10.056>.
- Baghban, A., Namvarrechi, S., Phung, L.T.K., Lee, M., Bahadori, A., Kashiwao, T., 2016b. Phase equilibrium modelling of natural gas hydrate formation conditions using LSSVM approach. *Petrol. Sci. Technol.* 34, 1431–1438. <https://doi.org/10.1080/10916466.2016.1202966>.
- Baker, L.E., Pierce, A.C., Luks, K.D., 1982. Gibbs energy analysis of phase equilibria. *Soc. Petrol. Eng. J.* 22, 731–742. <https://doi.org/10.2118/9806-pa>.
- Boston, J., Britt, H., 1978. A radically different formulation and solution of the single-stage flash problem. *Comput. Chem. Eng.* 2, 109–122. [https://doi.org/10.1016/0098-1354\(78\)80015-5](https://doi.org/10.1016/0098-1354(78)80015-5).
- Cheung, S.W., Chung, E.T., Efendiev, Y., Gildin, E., Wang, Y., 2018. Deep global model reduction learning. *arXiv*. <https://doi.org/10.48550/arXiv.1807.09335>.
- Coats, K.H., 1980. An equation of state compositional model. *Soc. Petrol. Eng. J.* 20, 363–376. <https://doi.org/10.2118/8284-pa>.
- Coutinho, E.J.R., Aqua, M.J., Gildin, E., 2021. Physics-aware deep-learning-based proxy reservoir simulation model equipped with state and well output prediction. In: *SPE Reservoir Simulation Conference*. <https://doi.org/10.2118/203994-ms>. SPE.
- Coutinho, E.J.R., Dall'Aqua, M., McClenny, L., Zhong, M., Braga-Neto, U., Gildin, E., 2022. Physics-informed neural networks with adaptive localized artificial viscosity. *arXiv*. <https://doi.org/10.48550/arXiv.2203.08802>.
- Culham, W., Ali, S.F., Stahl, C., 1969. Experimental and numerical simulation of two-phase flow with interphase mass transfer in one and two dimensions. *Soc. Petrol. Eng. J.* 9, 323–337. <https://doi.org/10.2118/2187-pa>.
- Eghtedaei, R., Moghaddam, N.K., Sarlak, V., deldar, A.N., Baghban, A., 2019. Neural computing approach for predicting vaporization enthalpy of pure hydrocarbons and petroleum fractions. *Petrol. Sci. Technol.* 37, 1898–1904. <https://doi.org/10.1080/10916466.2017.1315727>.
- El-Banbi, A.H., Forrest, J., Fan, L., McCain, W., 2000. Producing rich-gas-condensate reservoirs—case history and comparison between compositional and modified black-oil approaches. In: *SPE International Petroleum Conference*. <https://doi.org/10.2118/58988-ms>. SPE.
- Firoozabadi, A., Pan, H., 2000. Fast and robust algorithm for compositional modeling: Part i - stability analysis testing. In: *SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers*. <https://doi.org/10.2118/63083-ms>.
- Fussell, D., Yanosik, J., 1978. An iterative sequence for phase-equilibria calculations incorporating the redlich-kwong equation of state. *Soc. Petrol. Eng. J.* 18, 173–182. <https://doi.org/10.2118/6050-pa>.
- Fussell, L., Fussell, D., 1979. An iterative technique for compositional reservoir models. *Soc. Petrol. Eng. J.* 19, 211–220. <https://doi.org/10.2118/6891-pa>.
- Goda, H.M., Shokir, E.M.E.M., Fattah, K.A., Sayyoud, M.H., 2003. Prediction of the PVT data using neural network computing theory. In: *Nigeria Annual International Conference and Exhibition. Society of Petroleum Engineers*. <https://doi.org/10.2118/85650-ms>.
- Haugen, K.B., Beckner, B., 2013. A critical comparison of reduced and conventional EOS algorithms. *SPE J.* 18, 378–388. <https://doi.org/10.2118/141399-pa>.
- Haykin, S.S., 1999. *Neural Networks: A Comprehensive Foundation*, International edition. Prentice Hall.
- Hirose, Y., Kawase, Y., Kudoh, M., 1978. General flash calculation by the Newton-raphson method. *J. Chem. Eng. Jpn.* 11, 150–152.
- Jessen, K., Moghaddam, R.R., 2009. A flow based lumping approach for compositional reservoir simulation. In: *SPE Reservoir Simulation Symposium, SPE*. <https://doi.org/10.2118/119160-ms>.

- Joergensen, M., Stenby, E., 1995. Optimization of pseudo-component selection for compositional studies of reservoir fluids. In: SPE Technical Conference, SPE. <https://doi.org/10.2118/30789-ms>.
- Kenyon, D., Behie, G.A., 1987. Third SPE comparative solution project: gas cycling of retrograde condensate reservoirs. *J. Petrol. Technol.* 39, 981–997. <https://doi.org/10.2118/12278-pa>.
- Kniazeff, V., Naville, S., 1965. Two-phase flow of volatile hydrocarbons. *Soc. Petrol. Eng. J.* 5, 37–44. <https://doi.org/10.2118/962-pa>.
- Li, C.C., 1971. Critical temperature estimation for simple mixtures. *Can. J. Chem. Eng.* 49, 709–710. <https://doi.org/10.1002/cjce.5450490529>.
- Li, Y., Johns, R.T., 2006. Rapid flash calculations for compositional simulation. *SPE Reservoir Eval. Eng.* 9, 521–529. <https://doi.org/10.2118/95732-pa>.
- Li, Y., Sun, R., Horne, R., 2019. Deep learning for well data history analysis. In: SPE Annual Technical Conference and Exhibition, SPE. <https://doi.org/10.2118/196011-ms>.
- Li, Z., Firoozabadi, A., 2012. General strategy for stability testing and phase-split calculation in two and three phases. *SPE J.* 17, 1096–1107. <https://doi.org/10.2118/129844-pa>.
- Michelsen, M.L., 1982a. The isothermal flash problem. part i. stability. *Fluid Phase Equil.* 9, 1–19. [https://doi.org/10.1016/0378-3812\(82\)85001-2](https://doi.org/10.1016/0378-3812(82)85001-2).
- Michelsen, M.L., 1982b. The isothermal flash problem. part II. phase-split calculation. *Fluid Phase Equil.* 9, 21–40. [https://doi.org/10.1016/0378-3812\(82\)85002-4](https://doi.org/10.1016/0378-3812(82)85002-4).
- Mohaghegh, S.D., Liu, J., Gaskari, R., Maysami, M., Olukoko, O., 2012. Application of well-based surrogate reservoir models (SRMs) to two offshore fields in Saudi Arabia, case study. In: SPE Western Regional Meeting. <https://doi.org/10.2118/153845-ms>.
- Moortgat, J., Firoozabadi, A., 2013. Higher-order compositional modeling of three-phase flow in 3d fractured porous media based on cross-flow equilibrium. *J. Comput. Phys.* 250, 425–445. <https://doi.org/10.1016/j.jcp.2013.05.009>.
- Ojha, V.K., Abraham, A., Snášel, V., 2017. Metaheuristic design of feedforward neural networks: a review of two decades of research. *Eng. Appl. Artif. Intell.* 60, 97–116. <https://doi.org/10.1016/j.engappai.2017.01.013>.
- Okuno, R., Johns, R.T., Sepehrnoori, K., 2009. Application of a reduced method in compositional simulation. *SPE J.* 15, 39–49. <https://doi.org/10.2118/119657-pa>.
- Prausnitz, J., Lichtenthaler, R., de Azevedo, E., 1998. *Molecular thermodynamics of fluid-phase equilibria*. In: *Physical and Chemical Engineering Sciences*, third ed. ed. Pearson Education.
- Price, H., Donohue, D., 1967. Isothermal displacement processes with interphase mass transfer. *Soc. Petrol. Eng. J.* 7, 205–220. <https://doi.org/10.2118/1533-pa>.
- Rachford, J., Rice, J., 1952. Procedure for use of electronic digital computers in calculating flash vaporization hydrocarbon equilibrium. *J. Petrol. Technol.* 4. <https://doi.org/10.2118/952327-g>, 19–3.
- Rezaveisi, M., Sepehrnoori, K., Johns, R.T., 2014. Tie-simplex-based phase-behavior modeling in an IMPEC reservoir simulator. *SPE Journal* 19, 327–339. <https://doi.org/10.2118/163676-pa>.
- Robinson, D., Peng, D., 1978. *The Characterization of the Heptanes and Heavier Fractions for the GPA Peng-Robinson Programs*. Volume RR-28 of *Research Report (Gas Processors Association)*. Gas Processors Association.
- Roebuck, I., Henderson, G., Douglas, J., Ford, W., 1969. The compositional reservoir simulator: case i - the linear model. *Society of Petroleum Engineers Journal* 9, 115–130. <https://doi.org/10.2118/2033-pa>.
- Sandler, S., 2006. *Chemical, Biochemical, and Engineering Thermodynamics*. John Wiley & Sons.
- Schmall, L., Varavei, A., Sepehrnoori, K., 2013. A comparison of various formulations for compositional reservoir simulation. In: SPE Reservoir Simulation Symposium, SPE. <https://doi.org/10.2118/163630-ms>.
- Soave, G., 1972. Equilibrium constants from a modified Redlich-Kwong equation of state. *Chemical Engineering Science* 27, 1197–1203. [https://doi.org/10.1016/0009-2509\(72\)80096-4](https://doi.org/10.1016/0009-2509(72)80096-4).
- Sun, A.C., Seider, W.D., 1995. Homotopy-continuation method for stability analysis in the global minimization of the gibbs free energy. *Fluid Phase Equilibria* 103, 213–249. [https://doi.org/10.1016/0378-3812\(94\)02579-p](https://doi.org/10.1016/0378-3812(94)02579-p).
- Thele, K., Lake, L., Sepehrnoori, K., 1983. A comparison of three equation-of-state compositional simulators. In: Reservoir Simulation Symposium, SPE. <https://doi.org/10.2118/12245-ms>.
- Tian, C., Horne, R.N., 2017. Recurrent neural networks for permanent downhole gauge data analysis. In: SPE Annual Technical Conference and Exhibition, SPE. <https://doi.org/10.2118/187181-ms>.
- Todd, M., Longstaff, W., 1972. The development, testing, and application of a numerical simulator for predicting miscible flood performance. *Journal of Petroleum Technology* 24, 874–882. <https://doi.org/10.2118/3484-pa>.
- Voskov, D.V., Tchepeli, H.A., 2009a. Compositional space parameterization: multicontact miscible displacements and extension to multiple phases. *SPE Journal* 14, 441–449. <https://doi.org/10.2118/113492-pa>.
- Voskov, D.V., Tchepeli, H.A., 2009b. Compositional space parameterization: theory and application for immiscible displacements. *SPE Journal* 14, 431–440. <https://doi.org/10.2118/106029-pa>. URL:
- Wang, K., Luo, J., Wei, Y., Wu, K., Li, J., Chen, Z., 2019a. Artificial neural network assisted two-phase flash calculations in isothermal and thermal compositional simulations. *Fluid Phase Equilibria* 486, 59–79. <https://doi.org/10.1016/j.fluid.2019.01.002>.
- Wang, K., Luo, J., Yan, L., Wei, Y., Wu, K., Li, J., Chen, F., Dong, X., Chen, Z., 2019b. Artificial neural network accelerated flash calculation for compositional simulations. In: SPE Reservoir Simulation Conference. Society of Petroleum Engineers. <https://doi.org/10.2118/193896-ms>.
- Whitson, C.H., Michelsen, M.L., 1989. The negative flash. *Fluid Phase Equilibria* 53, 51–71. [https://doi.org/10.1016/0378-3812\(89\)80072-x](https://doi.org/10.1016/0378-3812(89)80072-x).
- Wilson, G., 1968. *A Modified Redlich-Kwong Equation of State Applicable to General Physical Data Calculations*. American Institute of Chemical Engineers National Meeting. Paper 15c.
- Wong, T., Firoozabadi, A., Nutakki, R., Aziz, K., 1987. A comparison of two approaches to compositional and black oil simulation. In: SPE Symposium on Reservoir Simulation. Society of Petroleum Engineers. <https://doi.org/10.2118/15999-ms>.