

New Fitting Method for Vapour-Liquid Equilibrium Data and its Application for Ethanol-Water Distillation Process Modeling

CSABA DEZSŐ ANDRÁS^{1*}, ÉVA MOLNOS¹, LÁSZLÓ MÁTYÁS², ALEXANDRU SZÉP¹

¹Sapientia Hungarian University of Transylvania, Faculty of Economics, Socio-Human Sciences and Engineering, Department of Food Science, 530104 – HR, 1 Libertatii Sq., Miercurea Ciuc, Romania

²Sapientia Hungarian University of Transylvania, Faculty of Economics, Socio-Human Sciences and Engineering, Department of Bioengineering, 530104 – HR, 1 Libertatii Sq., Miercurea Ciuc, Romania

Abstract: For modeling and simulation of distillation processes, the main information required is the vapour-liquid equilibrium (VLE) data. These are available as discrete values, but for modeling purpose, an analytical function would be more suitable. We developed a simple calculation protocol to obtain a single continuous function for the VLE curve on the entire concentration domain, without using thermodynamic functions. The fitting parameters of binary ethanol-water mixture VLE curve were determined. The methodology is suitable for fitting continuous functions on VLE data of any binary or multicomponent, non-ideal (even azeotropic) mixtures.

Keywords: distillation, vapour-liquid equilibria, curve fitting, ethanol-water binary mixture

1. Introduction

The modeling of separation processes in chemical engineering is indispensable for the modern design of chemical equipment. The process simulation and modeling in the industry are limited many times by the lack of thermodynamic knowledge and/or data. Poor thermophysical property or equilibrium data lead to significant calculation errors and conduct to erroneous equipment design [1]. Any improvements in this direction increase the accuracy of the predictions. All design procedures of separation processes are based on mass balances that can be calculated from equilibrium data. In particular, for distillation, these are the vapour-liquid equilibrium curves. These data are available in tabular form from databases, as discrete values, or when needed experimentally determined. To obtain continuous curves, a fitting method is needed, but in many cases (mainly for highly non-ideal, azeotropic mixtures), which also presents a concave region, this could be a challenging task. For this purpose, equations of state [2] or Gibbs excess energy method with different activity coefficient (NRTL, Wilson, UNIQUAC) or thermodynamic models (Margules or van Laar equations) are commonly used [3].

Distillation is undoubtedly one of the most important separation unit operation in the chemical processes [4, 5], and alongside with the solid-liquid extraction even in the food industry. The distillation of fermented mash for the production of food grade or bio-ethanol and for obtaining the distilled spirits with high alcohol content [6, 7] is the most important industrial distillation process from the agri-food sector. One of the reasons for choosing the water-ethanol system as subject to our study was the practical importance of the mixture, another, that this binary system is highly non-ideal. The behavior of azeotropic mixtures, with inflexion point and a concave region, makes the choice of fitting function more challenging. In this condition, inasmuch as a reliable fitting method could be developed, this will likely be applicable for further difficult systems too.

Our method is based on the concept of relative volatility (α) of the two components that connect the molar fraction of the volatile component in the vapour (y) and the liquid phase (x). In simple models, the relative volatility is considered constant, but this, even in the case of ideal mixtures is an oversimplification, giving considerable errors. In technical batch distillation calculation is usual practice to consider the relative volatility as constant, their value being the average of initial and final values or equal to the value at the average temperature [8].

*email: andrascaba@uni.sapientia.com



A better fitting with wider applicability is obtained by using the inverse linear function (eq. 1), mainly for the temperature dependence of relative volatility of near-ideal mixtures [9]. A more general approximation for VLE is the generalized constant relative volatility model proposed by Gmehling (eq. 2) [10]. This is considered the best of the mentioned models, but with some limitations, namely, their validity is restrained only to minimum boiling azeotropic binary mixtures [11].

$$\alpha = \alpha_0 + \frac{ax}{1+bx} \quad (1)$$

$$y = \frac{\alpha x}{1 + (\alpha - 1)x} + bx(1 - x) \quad (2)$$

As the mass balance equation of batch distillation (Rayleigh integral) contains the difference of the volatile component concentration in the two phases (the driving force of the process), the simplicity of the relative volatility function increase the possibility that the Rayleigh formula (eq. 3) is analytically integrable.

$$\ln \frac{F}{W} = \int_{x_W}^{x_F} \frac{dx}{y - x} = I_R \quad (3)$$

Based on this information, we try to find simple functions for relative volatility with a minimal error that makes the value of Rayleigh integral (I_R) more precise or even ensure their analytical integrability.

2. Materials and methods

As our work is related to process modeling, it was made on a computer, using dedicated software. For data analysis and graphical presentation, we used Statistica 8.0 (StatSoft, Inc., Oklahoma, USA). The VLE data for ethanol-water binary mixtures at atmospheric pressure was taken from Banu (2006) [6], since this dataset contains the largest number of values. To find an appropriate fitting function for the relative volatility (RV), firstly, we obtained their values from the VLE data, using the defining equation (eq. 4) and making their graphical representation in function of ethanol molar concentration in the liquid phase (x). Based on the graphical representation, the possible adequate function was selected.

$$\alpha = \frac{y}{1-y} \cdot \frac{1-x}{x} \quad (4)$$

The second strategy was to try to choose a transformation for the denominator of the I_R (the driving force), for the resulting fitted form to become analytically integrable. However, to achieve this goal is still necessary for the appropriate choice of the fitting function. For this reason, besides the relative volatility (α), two new functions (β and γ) were defined (eq. 5 and eq. 6):

$$\beta = \frac{y}{x} - 1 \quad (5)$$

$$\gamma = \frac{y}{x} \quad (6)$$

when for the β -function a rational polynomial with unit numerator (reverse polynomial function) (eq. 7) are set, can be easily demonstrated, that the I_R is integrable (eq. 8), where the x_F and x_W represent the initial batch composition and the waste composition, respectively.

$$f(x) = \left(\sum_{k=1}^n a_k x^k \right)^{-1} \quad (7)$$

$$I_R = \int_{x_W}^{x_F} \frac{dx}{y-x} = \int_{x_W}^{x_F} \frac{dx}{\beta x} = \int_{x_W}^{x_F} \left(\frac{1}{x} \sum_{k=0}^n a_k x^k \right) dx = \int_{x_W}^{x_F} \left(\frac{a_0}{x} + \sum_{k=1}^n a_k x^k \right) dx =$$

$$= \left(a_0 \ln x + \sum_{k=1}^n \frac{a_k}{k+1} x^{k+1} \right) \Big|_{x_W}^{x_F} = a_0 \ln \frac{x_F}{x_W} + \sum_{k=1}^n \frac{a_k}{k+1} (x_F^{k+1} - x_W^{k+1}) \quad (8)$$

The differences between the observed data and calculated values are expressed as relative errors (eq. 9):

$$e = \frac{y_{obs} - y_{pred}}{y_{obs}} \quad (9)$$

where: e – relative error (deviation), y_{obs} – the observed data, y_{pred} – predicted value.

For the fitting procedure, we used two nonlinear estimation mode, namely the least square (with Gauss-Newton estimation) and custom loss function (minimization of the square of the difference of the observed and predicted values).

3. Results and discussions

The selection of an adequate fitting functions is a somewhat challenging task, therefore no universal model for VLE of non-ideal mixtures exists. Taken into account the above considerations, we have the following results. Primarily, for fitting the VLE curve to data of ethanol-water mixtures at atmospheric pressure, we tried to find a probe function. Firstly we used the function proposed by *Said and Al-Ameeri* (1987) [8] for calculations in chemical engineering, especially for distillation processes (eq. 10):

$$f(x) = \frac{ax}{1 + (a-1)x^n} \quad (10)$$

Figure 1. represents the values of RV (α) as a function of the liquid mole fraction of ethanol (x) together with the fitting function.

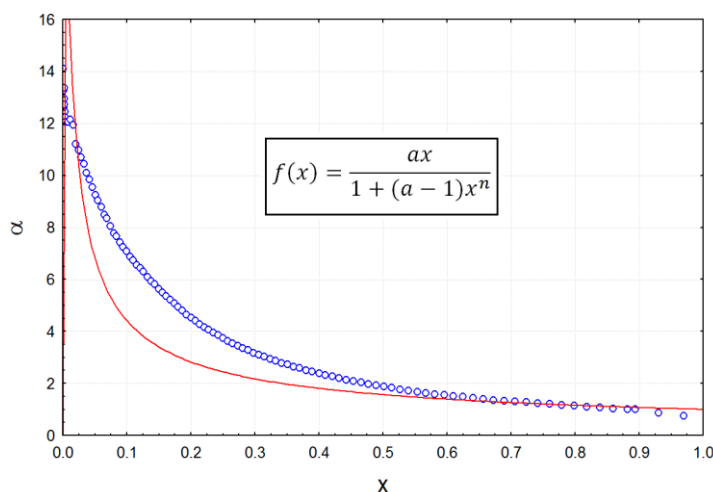


Figure 1. Probe function fitted on relative volatility (α) data:
 \circ – observed values, — – predicted curve

As can be seen in Figure 1, the function does not fit well on the data (a large difference between the observed and predicted values), but the trend (asymptotic decrease) is correct. The predicted values are mainly below the observed data, especially at lower x values. The reason for this behavior is given by

the presence of x in the numerator, which decreases the function value at lower molar concentrations. To compensate this effect, we probe, functions with unitary numerators. The tested functions, the resulted fitting parameters, and the correlation coefficients (R -values) are summarized in Table 1.

Table 1. Fitting probe functions for the relative volatility (α).

Function	Fitting parameters			R
	a	b	n	
$\alpha_1(x) = \frac{1}{a + bx^n}$	0.772700	0.985950	1.1782	0.99891
$\alpha_2(x) = \frac{1}{a + x^n}$	0.077337	–	1.8558	0.99891
$\alpha_3(x) = \frac{1}{(a + bx)^n}$	0.197520	0.942936	1.5856	0.99919

For all the proposed functions, the fitting is tight, the value of the correlation coefficients are very high (closed to 1). Based on the R -values one cannot be evaluating the best function, therefore we represented the relative deviation (error) between the observed and predicted values for all the three functions on the same diagram (Figure 2).

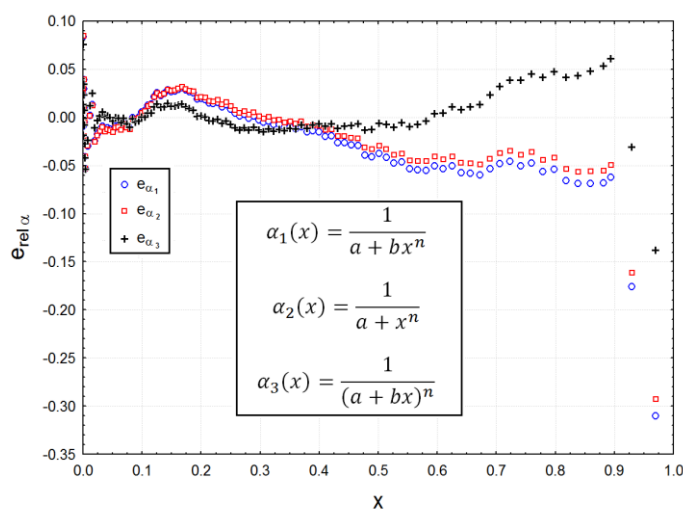


Figure 2. The variation of $e_{rel\alpha}$ as a function of x for the three probe function

In the large interval of x , the relative deviation remains below 5% (0.05). Near the azeotropic point, ($x \geq 0.9$) the deviation increase drastically. The lowest value of relative deviation can be observed for the α_3 -function, named in the following as power-law function.

The observed and the predicted value of the relative volatility by the power-law function are given in Figure 3.

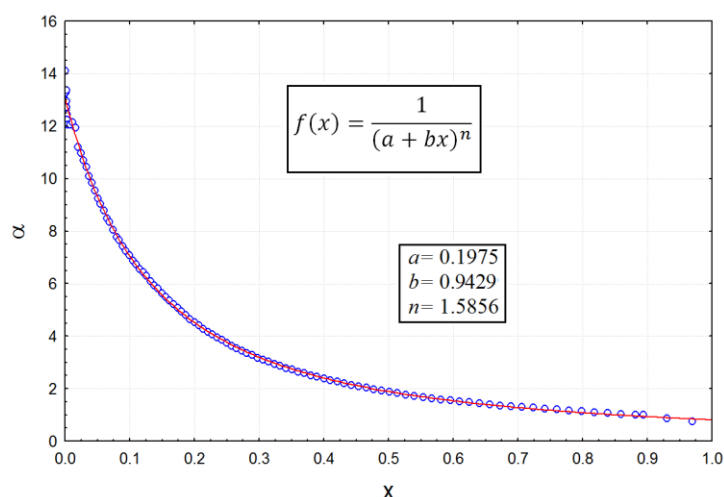


Figure 3. Representation of the observed and predicted RV with the power-law function with the particular parameters (a , b , n)

Based on the relative deviation values (Figure 2., curve e_{a3}), it can be concluded that our proposed power-law function is adequate as an empirical model for the relative volatility curve. Thereafter, the VLE values given by the three empirical models, namely the reverse linear [8], the Gmehling [10–12], and the power-law functions were compared (Figure 4). The parameters of the three empirical models are given in Table 2. and were obtained by function fitting for the reverse linear and power-law function from the literature [12]. Moreover, the Gmehling model is somehow different from the other two models, i.e., use constant relative volatility (α , RV) and give directly the VLE curve ($y=f(x, \alpha)$).

Table 2. Parameter values for the tested empirical models

Function name	Function form	Fitting parameters				Obs.
		a	b	n	α	
Reverse linear	$\alpha = a + \frac{bx}{1+nx}$	13.1662	-109.710	7.7633	–	variable RV
Gmehling*	$y = \frac{\alpha x}{1 + (\alpha - 1)x} + bx(1 - x)$	–	-1	–	9.5	constant RV
Power law	$\alpha = \frac{1}{(a + bx)^n}$	0.19752	0.942936	1.5856	–	variable RV

*Parameter values are taken from the literature [12]

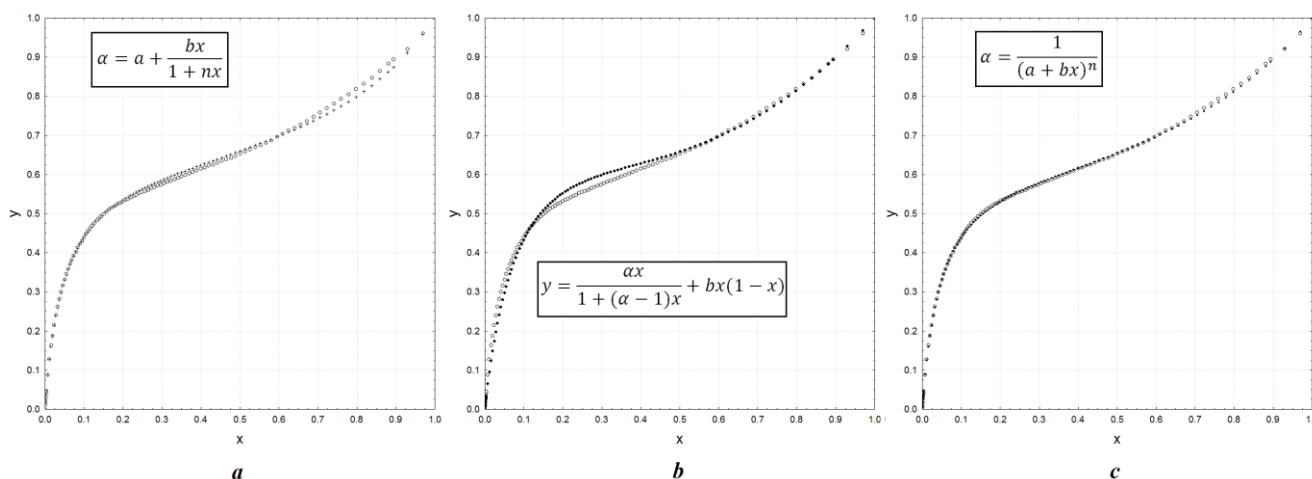


Figure 4. The observed and predicted equilibrium curves for the ethanol-water system: **a.** reverse linear model (\circ – observed values, $+$ – predicted values); **b.** Gmehling model (\circ – observed values, \diamond – predicted values); **c.** power-law model (\circ – observed values, \blacktriangle – predicted values)

Some differences can be observed between Figures 4. *a–c*. Still, to compare the fitting quality of the models, we used as previously the relative deviation values represented graphically in Figure 5. The best fitting is given by the power-law model proposed by us, furthermore, this model has three fitting parameters, too.

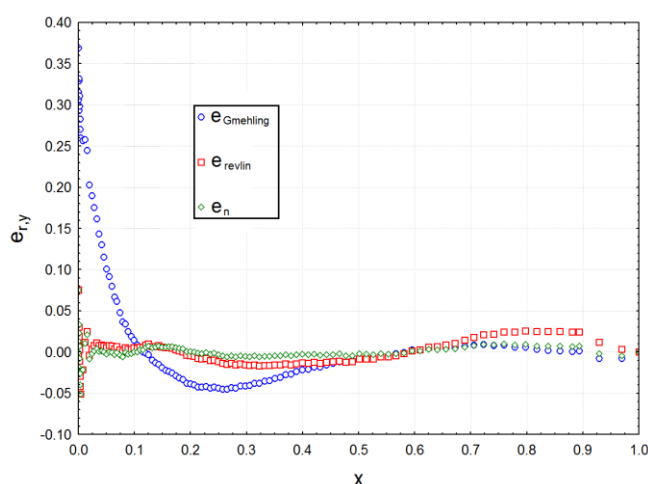


Figure 5. Relative deviations of the observed and predicted equilibrium curve for the three compared fitting functions (\circ – Gmehling, \square – reverse linear, \diamond – power-law function)

The second strategy was to choose an adequate model for the VLE curve to obtain an analytically integrable form of Rayleigh equation. As demonstrated in eq. 7., that choosing a reverse polynomial form for the $(y-x)/x$ fraction (hereinafter referred to as β -function, in conformity with eq. 5.), fulfills the proposed requirement. However, to achieve a tight fitting, it is still necessary the appropriate choice of the polynomial degree. Besides the relative volatility (α), two new functions (β and γ) were defined (eq. 5 and eq. 6).

By gradual increase the polynomial degree, it was demonstrated that the best fitting is given by the 5th degree polynomials. The further increase of the degree leads to oscillatory behavior of the fitting function, moreover, no further improvement in the relative deviation was observed (results not shown).

In the denominator of the argument of the Rayleigh integral (eq. 3), the difference of the phase composition appears, which is represented in Figure 6.

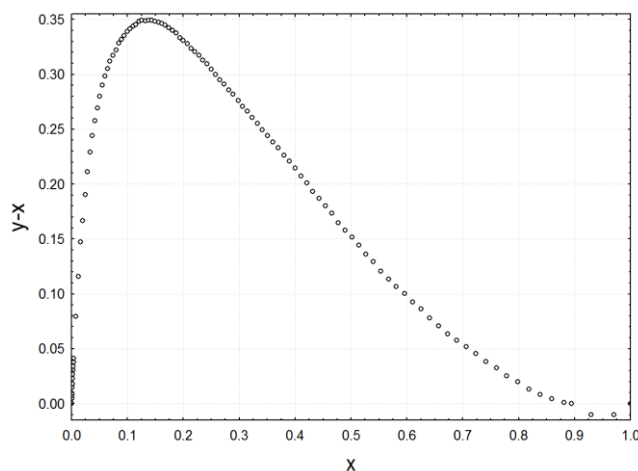


Figure 6. The difference of ethanol mole fractions in the two phases (vapour-liquid phases) in function of mole fraction in the liquid phase (x)

It would be very expedient to be able to select the fitting function based on the visual examination of Figure 6. Despite that it can be established that the function has a probability distribution shape, we were unable to define a proper function type that fits well to these values. To simplify the problem, we introduced the intermediary functions: the fractions of β and γ (eq. 5 and 6). After that, we tried to find fitting functions for them, which can be used further to calculate the VLE curve. This procedure basically is the same that the method using α -function, but the equilibrium value of the y has more simplified form (eq. 11–13):

$$y_{\alpha} = \frac{\alpha x}{1 + (\alpha - 1)x} \quad (11)$$

$$y_{\beta} = (\beta + 1)x \quad (12)$$

$$y_{\gamma} = \gamma \cdot x \quad (13)$$

These three intermediary functions are represented in the same diagram (Figure 7.).

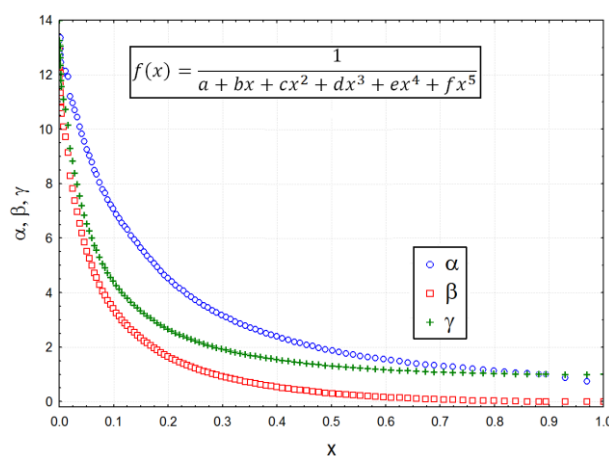


Figure 7. Graphical representation of the intermediary function

All these functions have the same shape (hyperbolic-like, asymptotic curve), it seems reasonable that they could be described with the same fitting function. The aforementioned 5th degree reverse polynomial function fits β -function, therefore we performed the fitting on both α - and γ intermediary functions as well. Assuming that the selected function fit tight on γ -function, it is worth to try a direct fitting on VLE values as follows (eq. 14):

$$y_{dir} = x\gamma = x \left(\sum_{k=1}^5 a_k x^k \right)^{-1} \quad (14)$$

The value of the fitting parameters for all three intermediary functions, as well as for the direct fitting on the VLE curve (y_{dir}) are summarized in Table 3.

Table 3. Fitting parameters of intermediary and direct VLE curve fitting functions, $f(x)$.

$f(x)$	Form of the fitting function	Fitting parameters						R
		a	b	c	d	e	f	
		a_0	a_1	a_2	a_3	a_4	a_5	
α	$\left(\sum_{k=1}^5 a_k x^k \right)^{-1}$	0.076243	0.617056	0.036437	3.663368	-7.1152	4.06665	0.99922
β		0.082212	1.933	-2.0015	50.0203	-150.22	190.781	0.99926
γ		0.076179	1.54997	-0.57268	2.37391	-5.0299	2.63274	0.99925
y_{dir}	$\frac{x}{\sum_{k=0}^5 a_k x^k}$	0.077936	1.45624	0.489814	-1.5746	0.5497	0.006211	0.99998

From the equations on the three intermediary functions (eq. 11–13) could be calculate the fitted models of the VLE curves using the fitted parameters from the table.3, In addition, using eq. 14, the direct fitting function for the same curve can be obtained (Figure 8).

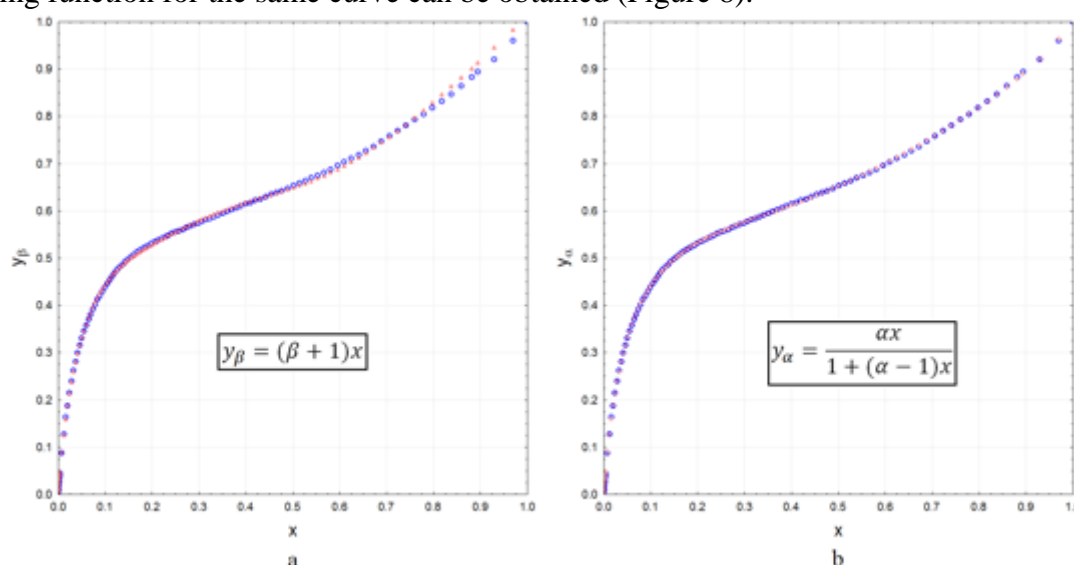


Figure 8. The observed and predicted VLE curves for ethanol-water system:

- a.** using the β -function (\circ – observed values, Δ – predicted values);
- b.** using the α -function (\circ – observed values, \diamond – predicted values)

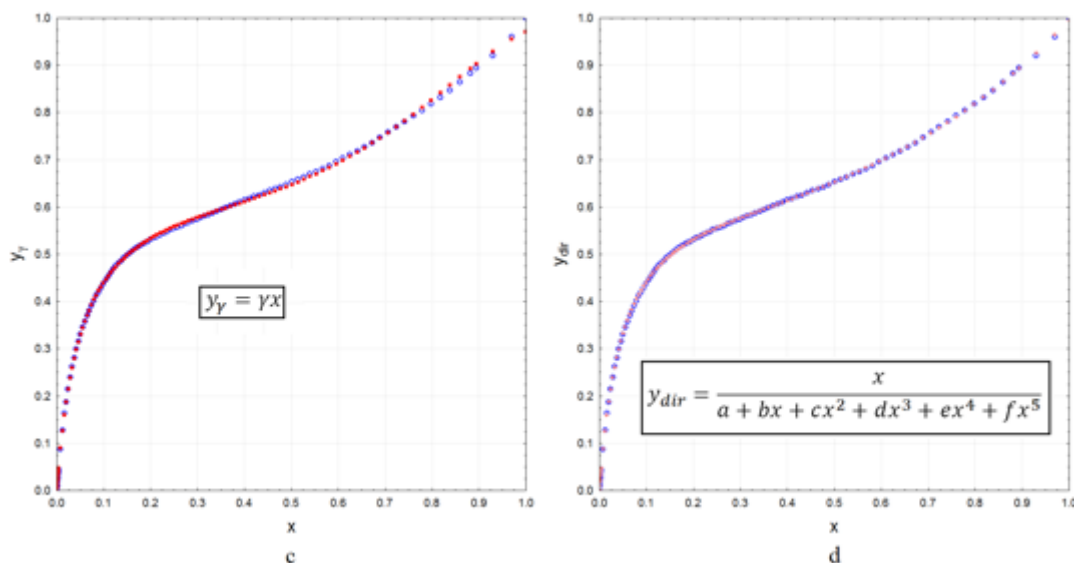


Figure 8. The observed and predicted VLE curves for ethanol-water system:
c. using the γ -function (\circ – observed values, \blacksquare – predicted values);
d. using the y_{dir} -function (\circ – observed values, $+$ – predicted values)

Visually, all the predicted values fit tight with the observed VLE values taken from the literature [6].

In order to compare the fitting quality of the four empirical models, we proceed similarly as previously, i.e., we represent graphically on the same diagram (Figure 9) the empirical model deviation from the tabulated data, and evaluate the graphic.

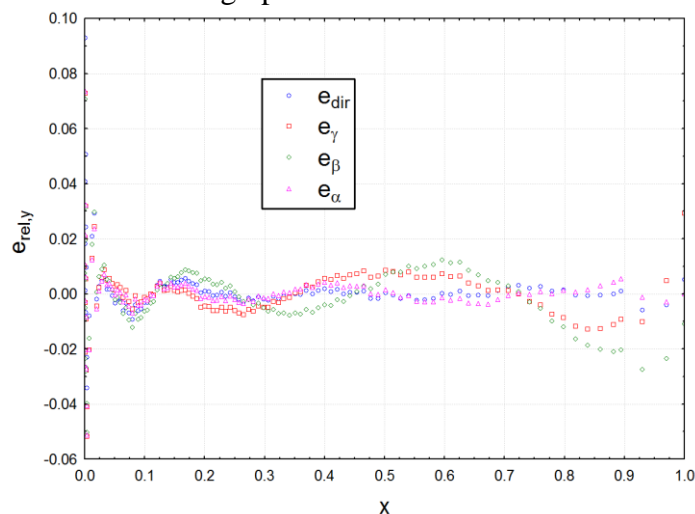


Figure 9. The profile of $e_{rel,y}$ as a function of x for the intermediary and the direct fitting functions

The evaluation of the VLE empirical models shows that at a very low liquid molar fraction ($x < 0.01$) all the models give very poor results, but from an engineering point of view, this domain does not present any real interest. Strictly, taking into account only the value of the deviation in the whole concentration domain, the ranking is as follows: the best model is the y_{dir} function, the second is the traditional y_α , the third is the y_γ , and the worst result is offered by the y_β model. This is not very encouraging for using the y_β model in the batch distillation calculus, but the situation is not so bad as it seems to be since the deviation problems appear only at higher molar fractions (about $x > 0.85$). Below, $x < 0.85$, the relative deviation band is narrow ($\pm 2\%$), suitable for precise engineering estimations.

Moreover, the zone above $x > 0.85$, do not present practical importance as well as the initial concentration, x_F , of the batch, very rarely exceeds the value of $x = 0.2$ (considering the fermentative technology). This study shows that the y_a -based calculus are eligible. The direct fitting models can give excellent results, but in general, is difficult to guess directly the corresponding function form.

Finally, we compared to the best results obtained from the two strategies, namely the power-law function obtained in the first part and the results obtained with the direct fitting in the second part. The results are presented in Figure 10.

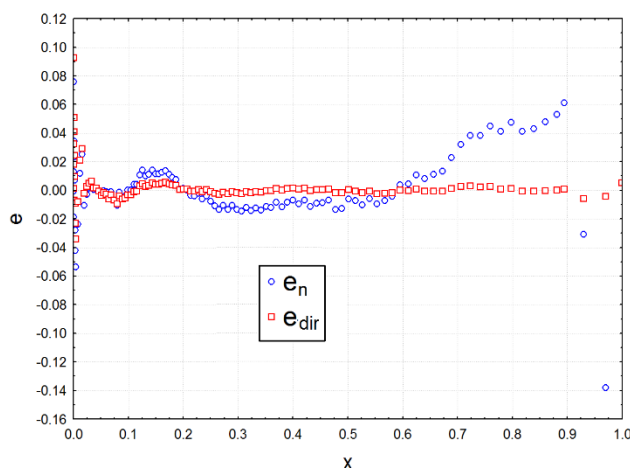


Figure 10. Comparison of the relative errors of the fitting for the direct function (e_{dir}) and power-law function (e_n)

Though the analysis of the figure, it can be observed that at high concentrations and after the azeotropic point the direct fitting gives a better result in comparison with the power-law function. It can be concluded, that the best model among the studied ones was the y_{dir} based direct model, using as fitting function the 5th degree reverse polynomial. The proposed power-law function can be used if the concentration in the liquid phase does not exceed $x = 0.80$ (the relative deviation, in this case, is under 4%). The advantage of the power-law function consists in their simplicity, forasmuch had only 3 fitting parameters, as well as their mathematical handling could be more obvious, does not present major difficulties.

The difference of the molar fractions ($y - x$) in the two phases may be expressed in terms of a power series (eq. 15), named the Ocon active fraction (Z_T) [13–15].

$$\frac{y-x}{x(1-x)} = \sum_{k=0} A_k Z_T^k = \sum_{k=0} A_k \left[\frac{x}{x + R_V(1-x)} \right]^k \quad (15)$$

Beyond mole related quantities, the Ocon active fraction contains another constant R_V , which should be determined by fitting [14], or from thermodynamic considerations [15]. We propose a similar, although a slightly different, more direct method, which makes possible a relatively straightforward way for the fitting procedure.

We also tried to achieve the integrability of the Rayleigh integral, that criteria is not fulfilled by the power series method proposed. However, the advantage of the above method is their applicability in the fitting of other thermodynamic functions and properties [13].



4. Conclusions

The modeling of innovative distillation methods (cyclic distillation) with general nonlinear VLE behavior requires more accurate non-iterative design algorithms [16], and the presented fast, and accurate method could be a part of the modeling arsenal. Using our method, even the most recent programming algorithm based on the classical McCabe-Thiele graphical method [17] could be improved and simplified. This would be possible as the above-mentioned algorithm for VLE estimation uses continuous piecewise linear approximation, with inherent inaccuracy, when the number or position of the data points are chosen inadequately. Even using professional modeling software packages, like Aspen Hysys and Pro/II great difficulty could meet to obtain a convergent solution when the relative volatility presents large variations and the VLE data are given in discrete form [18].

Due to very tight fitting of the proposed functions on the observed data, with minor deviation, the new VLE empirical models are suitable for modeling and simulation purposes, resolving the major source of a calculation error in aforementioned modeling cases.

Acknowledgement. The author are grateful to Szilárd Bucs (KAUST, Thuwal) for valuable comments and useful discussions.

References

1. GMEHLING, J., KLEIBER, M., KOLBE, B., RAREY, J., Chemical Thermodynamics for Process Simulation, 2nd completely revised and enlarged edition, Wiley-VCH, Weinheim, 2019.
2. GEANĂ, D., FEROIU, V., Ecuatii de Stare. Aplicații la Echilibre de Faze, Editura Tehnică, București, 2000.
3. GÓRAK, A., SORENSEN, E. (eds), Distillation: Fundamentals and Principles, Elsevier Academic Press, Amsterdam, 2014.
4. FAIR, J.R., Distillation: King in separations, *Chem. Processing*, **9**, 1990, 23–31.
5. KISS, A.A., Distillation technology-still young and full of breakthrough opportunities, *J. Chem. Technol. Biotechnol.*, **91**, 2014, 479–498. <https://doi.org/10.1002/jctb.4262>
6. BANU, C. (ed.), Bioetanolul - Combustibilul Viitorului, Editura AGIR, București, 2006.
7. HOROBA, E., Alcool Etilic: Obținere, Carburant, Băuturi Alcoolice, Editura PIM, Iași, 2010.
8. SAID, A.S., AL-AMEERI, R.S., Curve fitting in science and technology, *Sep. Sci. Technol.*, **22**(1), 1987, 65–84. <https://doi.org/10.1080/01496398708056158>
9. SAID, A.S., AL-HADDAD, A.A., Curve fitting in science and technology. II. Application to relative volatility, bubble point, and dew point of ideal binary mixtures, *Sep. Sci. Technol.*, **22**(4), 1987, 1199–1218. <https://doi.org/10.1080/01496398708057175>
10. STOJKOVIC, M., GERBAUD, V., SHCHERBAKOVA, N., Batch distillation of binary mixtures: preliminary analysis of optimal control, *IFAC-PapersOnLine*, **50**(1), 2017, 4899–4904. <https://doi.org/10.1016/j.ifacol.2017.08.743>
11. BINOUS, H., AL-HARTHI, M.A., Simple batch distillation of a binary mixture, *Inc. Comput. Appl. Eng. Educ.*, **22**, 2014, 649–657. <https://doi.org/10.1002/cae.21556>
12. DOHERTY, M.F., MALONE, M.D., Conceptual Design of Distillation Systems, McGraw-Hill, New York, 2001.
13. SUSIAL, P., RODRÍGUEZ-HENRÍQUEZ, J.J., APOLINARIO, J.C., CASTILLO, V.D., ESTUPINAN, E.J., Vapour-liquid equilibrium of propyl acetate+ethanol and propyl acetate+2-propanol at 0.6 MPa, *Rev. Chim.*, **62**, 2011, 1115–1118.
14. SUSIAL, P., SOSA-ROSARIO, A., RODRÍGUEZ-HENRÍQUEZ, J.J., RIOS-SANTANA, R., Vapor pressure and VLE data measurements on ethyl acetate/ethanol binary system at 0.1, 0.5, and 0.7 MPa, *J. Chem. Eng. Jpn.*, **44**(3), 2011, 155–163. <https://doi.org/10.1252/jcej.10we176>



15. SUSIAL BADAJOZ, P., GARCÍA-VERA, D., PADRÓN-GUERRA, N., MUJICA-GONZÁLEZ, T., Measurement of VLE data by using an experimental installation with automatic control: Modeling of binary systems of methyl acetate or ethyl acetate with n-heptane or 2,2,4-trimethylpentane at both 0.1 and 1.5 MPa, *J. Chem. Eng. Data*, **64**(12), 2019, 5591–5608.

<https://doi.org/10.1021/acs.jced.9b00668>

16. BÎLDEA, C.S., PĂTRUȚ, C., JØRGENSEN, S.B., ABILDSKOV, J., KISS, A.A., Cyclic distillation technology-a mini-review, *J. Chem. Technol. Biotechnol.*, **91**, 2016, 1215–1223.

<https://doi.org/10.1002/jctb.4887>

17. KONG, L., MARAVELIAS, C.T., From graphical to model-based distillation column design: A McCabe-Thiele-inspired mathematical programming approach, *AIChE J.*, **65**(11), 2019, e16731.

<https://doi.org/10.1002/aic.1673>

18. FAKEEHA, A., Simulation of distillation of a large relative volatility mixture, *J. King Saud Univ. Eng. Sci.*, **15**(1), 2003, 13–26. [https://doi.org/10.1016/S1018-3639\(18\)30759-1](https://doi.org/10.1016/S1018-3639(18)30759-1)

Manuscript received: 4.06.2020