

Commentary on “Uncover the effect of solvent and temperature on solid-liquid equilibrium behavior of L-norvaline”

William E. Acree, Jr.*

Department of Chemistry, 1155 Union Circle Drive #305070, University of North Texas,
Denton, TX 76203, USA

Abstract

Several sets of reported curve-fit parameters reported by Wang and coworkers [J. Mol. Liq. 243 (2017) 273-284] for the Modified Apelblat model do not correctly back-calculate the observed solubility behavior of L-norvaline in the binary aqueous-alcohol solvent mixtures studied by the authors. Too much rounding is likely the reason for the failure of the curve-fit parameters to yield the back-calculated mole fraction solubilities that the authors reported in the published paper.

Key Words and Phrases: L-norvaline; binary aqueous-alcohol solvent mixtures; Modified Apelblat model; mathematical representation

*To whom correspondence should be addressed. (E-mail: acree@unt.edu); fax: 940-565-4318.

In a recent paper appearing in *This Journal* Wang and coworkers [1] reported the solubility of L-norvaline in three different binary aqueous-alcohol solvent mixtures from 283.15 K to 318.18 K. Solubilities were determined by a gravimetric method that involved transferring a known aliquot of the saturated solution into tarred glass dish. The solvent was evaporated at 323.15 K in a vacuum drying oven. The concentration of L-norvaline in the saturated solution was then calculated from the mass of the saturated solution taken for analysis and the mass of the solid residue that remained after the solvent had completely evaporated. The authors correlated the measured mole fraction solubilities in the binary aqueous-methanol, aqueous-ethanol, and aqueous-1-propanol solvent mixtures using the Modified Apelblat equation, the Buchowski-Ksiazczak (λh) equation and the NRTL model.

The purpose of this communication is not to criticize the fine work of Wang and coworkers [1], but to point out what can happen in publishing mathematical correlations if one is not careful. As my illustrational example, I will use the Apelblat equation:

$$\ln x_1 = A + \frac{B}{T} + C \ln T \quad (1)$$

and the curve-fit coefficients of $A = -166.20$; $B = 5832.00$; and $C = 24.80$ that the authors reported in Table 8 of their published paper [1] for L-norvaline dissolved in binary aqueous-methanol solvent mixtures at $x_{\text{methanol}} = 0.20$. Substituting the curve-fit coefficients into Eqn. 1:

$$\ln x_1 = -166.20 + \frac{5832.00}{298.15} + 24.80 \ln 298.15 \quad (2)$$

$$\ln x_1 = -166.20 + 19.56062 + 24.80 \times 5.697597 \quad (3)$$

$$\ln x_1 = -5.338978 \quad (4)$$

I calculate a mole fraction solubility of $x_1 = 0.004801$ for the solubility of L-norvaline at 298.15 K. The back-calculated value that the authors give in Table 2 [1] for this temperature and solvent composition is $x_1 = 0.006304$, which is significantly larger than the value that I calculated with the tabulated curve-fit equation coefficients. Similar discrepancies can be found for several other sets of tabulated equation coefficients. Several sets of tabulated Apelblat equation coefficients fail to describe the observed solubility, and as a result would have very limited applicability in calculating the solubility of L-norvaline, even at the temperatures studied by the authors.

The question that one might ask is why don't the tabulated curve-fit equation coefficients give the numerical back-calculated x_1 values that the authors tabulated in Table 2. I suspect that the reason deals with rounding of the curve-fit equation coefficients. In this case, a slight rounding of the C equation coefficient in the second decimal place can have a rather pronounced effect on the back-calculated value of x_1 . In examining all of the authors' tabulated values of C , I noticed that every C value ended in a zero in the second decimal place. What is the probability that 27 C values determined by regression analysis would all end in a zero in the second decimal? The calculated values were likely rounded to the tenths place, in which case the calculated value of C would fall between $C = 24.75$ and $24.849+$. Let me now perform the back-calculation again, this time using $C = 24.849$. Substituting the numerical values of the curve-fit parameters into Eqn. 1:

$$\ln x_1 = -166.20 + \frac{5832.00}{298.15} + 24.849 \ln 298.15 \quad (5)$$

$$\ln x_1 = -166.20 + 19.56062 + 24.849 \times 5.697597 \quad (6)$$

$$\ln x_1 = -5.059792 \quad (7)$$

I calculate a much larger mole fraction solubility, $x_1 = 0.006347$, than before. The back-calculated value using $C = 24.949$ differs by less than 1 % from the back-calculated value of $x_1 = 0.006304$

that the authors gave in Table 2. Here what may seem as a small degree of rounding had a fairly significant effect on the back-calculated value.

Authors need to remember that the purpose for publishing mathematical representations for describing the variation of mole fraction solubility with temperature is to allow readers to estimate the solubility at other temperatures. This purpose is nullified when authors round the calculated curve-fit parameters to the point where the parameters no longer back-calculate the observed solubility data. Having published several papers I know that reviewers may question the number of significant figures that authors give in reporting curve-fit parameters. Often authors are asked to reduce the number of significant figures in curve-fit parameters to a specified number or to a reasonable number. What I suggest is that in deciding how many significant figures to use in expressing curve-fit parameters, is that one carefully examine the back-calculated values as one systematically reduces the number of significant digits. When the curve-fit equation coefficients no longer provide a satisfactory back-calculation of the measured solubility data then one has gone too far in rounding.

As an illustrational example, I will consider the experimental mole fraction solubility for L-norvaline in binary methanol (2) + water (3) solvent mixtures taken from Table 3 of the paper by Wang and coworkers [1]. The authors arranged the experimental data by temperature. I have reproduced in Table 1 of this commentary the experimental mole fraction solubilities, x_1^{exp} , at $x_{\text{methanol}} = 0.10$. What I have done is regressed the numerical $\ln x_1^{\text{exp}}$ values according to the Modified Apelblat equation. The curve-fit equation coefficients that I obtained have been substituted into Eqn. 1 to yield the following mathematical representation:

$$\ln x_1 = -173.149(13.746) + \frac{6488.487(615.092)}{T} + 25.7280(2.050) \ln T \quad (8)$$

The IBM SPSS statistical software was used in the regression analysis. I have deliberately expressed each equation coefficient to more significant figures than would be justified by the number of significant digits in the experimental data set. For this particular data set every one of the x_1^{exp} values was given to at least 4 significant digits. I now round every coefficient to five significant digits:

$$\ln x_1 = -173.15 + \frac{6488.5}{T} + 25.728 \ln T \quad (9)$$

I calculate the value of $x_1^{\text{calc,eqn 9}}$, and record the numerical values in the third column of Table 1 along with the average absolute percent relative deviation, AAPRD, based on Eqn. 10

$$AAPRD = (1/8) \sum_{i=1}^N \left| \frac{x_{1,i}^{\text{calc,eqn 9}} - x_{1,i}^{\text{exp}}}{x_{1,i}^{\text{exp}}} \right| \quad (10)$$

There are 8 experimental values in the data set. Careful examination of the numerical values in the second, third and fourth columns of Table 1 reveals that Eqns. 8 and 9 both provide a very mathematical representation of the experimental solubility data.

Table 1. Comparison between experimental mole fraction solubilities, x_1^{exp} , and back-calculated values based on Modified Apelblat Equation with different rounding of equation coefficients for binary methanol + water solvent mixtures at $x_{\text{methanol}} = 0.1$

T/K	$10^4 x_1^{\text{exp}}$	$10^4 x_1^{\text{calc,eqn 8}}$	$10^4 x_1^{\text{calc,eqn 9}}$	$10^4 x_1^{\text{calc,eqn 11}}$
283.15	69.26	69.16	69.09	73.33
288.15	72.80	72.90	72.83	77.31

293.15	76.76	77.30	77.22	81.97
298.15	82.56	82.41	82.32	87.39
303.15	88.18	88.28	88.20	93.63
308.15	95.21	95.02	94.93	100.79
313.15	102.79	102.72	102.62	108.96
318.15	111.11	111.49	111.39	118.27
AAPRD		0.225	0.237	6.15

I now reduce the number of significant figures in each coefficient by one:

$$\ln x_1 = -173.1 + \frac{6488}{T} + 25.73 \ln T \quad (11)$$

and calculate the value of $x_1^{\text{calc,eqn 11}}$ and the associated AAPRD value. The respective numerical values are tabulated in the fifth column of Table 1. As shown in Table 1 the deviations between the experimental mole fraction solubilities and back-calculated values based on the calculated curve-fit equation coefficients have increased significantly, from AAPRD = 0.24 (for Eqn. 9) to AAPRD = 6.15 (for Eqn. 11). Too much rounding has taken place. Equation 9 provides a much better mathematical representation of observed solubility data, and is preferred over Eqn. 11. In this case five significant figures are needed for an acceptable mathematical representation. Five significant digits is one more than the least number of significant digits in the data set being mathematically described. Each data set and each solution model will be different. It is recommended that authors carefully examine the back-calculated mole fraction solubilities in

deciding how many significant figures to use when reporting their calculated curve-fit equation coefficients.

References

- [1] Yaping Wang, Ya Liu, Peng Shi, Shichao Du, Yumin Liu, Danan Han, Panpan Sun, Mengmeng Sun, Shijie Xu, Junbo Gong, Uncover the effect of solvent and temperature on solid-liquid equilibrium behavior of l-norvaline, *J. Mol. Liq.* 243 (2017) 273-284.

HIGHLIGHTS

- Modified Apelblat constants found not to calculated mole fraction solubilities
- Modified Apelblat constants reported to too few significant digits
- Rounding of curve-fit constants discussed