Critical Evaluation of Vapor-Liquid Equilibrium, Heat of Mixing, and Volume Change of Mixing Data. General Procedures

Cite as: Journal of Physical and Chemical Reference Data **11**, 941 (1982); https://doi.org/10.1063/1.555667
Published Online: 15 October 2009

Buford D. Smith, Ol Muthu, Ashok Dewan, and Matthew Gierlach

Critical Evaluation of Vapor-Liquid Equilibrium, Heat of Mixing, and Volume Change of Mixing Data. General Procedures

Buford D. Smith, Ol Muthu, Ashok Dewan, and Matthew Gierlach

Thermodynamics Research Laboratory, Box 1144, Washington University, St. Louis, Missouri 63130

This paper is the first in a series of reports on the critical evaluation of vapor-liquid equilibrium, heat of mixing, and volume change of mixing data for binary liquid mixtures of nonelectrolytes. The specific evaluation procedures for each property will be covered in subsequent articles. This paper describes the general procedures used to support the mixture evaluation work. The areas covered include the procedures used to cover the primary and secondary literature, the computer program libraries developed for pure compound and mixture data processing, the procedures used to evalute and correlate the pure compound property data for use by the mixture programs, and the efforts made to make the best equation of state available to the vapor-liquid equilibrium data reduction programs. Improvements are suggested for the presentation of mixture data in the literature.

Key words: equations of state; heat of mixing; liquid density; mixtures; second virial coefficients, vapor-liquid equilibrium; vapor pressure; volume change of mixing.

Contents

List of Tables

List of Figures

1. Introduction

The Thermodynamics Research Laboratory has undertaken the task of evaluating, tabulating, and disseminating the published vapor-liquid equilibrium (VLE), heat of mixing (H^E), and volume change of mixing (V^E) data for mixtures of compounds containing the following elements:

Carbon
Hydrogen
Oxygen
Nitrogen
Sulfur
Halogens (F, Cl, Br, I)
Rare Gases (He, Ne, Ar, Kr, Xe)
Silicon

All organic compounds plus a large number of inorganic compounds are included.

The ultimate objective of the evaluation program is the

creation of a mixture data bank MDATA. All the good binary VLE, H^E, and V^E data identified in the evaluation program are being stored in MDATAl in correlated form. The binary VLE, H^E, and V^E data are correlated simultaneously as a function of temperature and composition using only four adjustable correlation constants [1, 2, 3][1]. The original experimental data can be retrieved from MDATAl within experimental accuracy, and accurate extrapolation with respect to temperature is possible.

Before the mixture data evaluation program could be started, it was necessary to complete four major tasks:

1. All the relevant literature documents had to be identified, retrieved, listed on a computer tape, and filed in a way which made them easily available to the data processers and evaluators.

2. A large specialized computer program library was needed to process, evaluate, tabulate, and store the pure compound data needed for the mixture evaluation work.

3. A similar program library was necessary for the mixture data processing.

4. Enough pure compounds had to be processed and stored in the pure compound data bank to support the mixture data processing.

The fact that two of the four major preliminary tasks involved pure compound data illustrates the importance of the pure compound data in the mixture evaluation work. The Laboratory has gone to extreme lengths to make the best possible pure compound data available to the mixture programs. Besides the retrieval and critical evaluation of the pure compound data, it has been necessary to store those data in a data bank (CDATAl) which can be automatically accessed by any mixture program needing pure compound data. That work has been done about as well as possible and there is no obvious way whereby the quality of the compound data delivered to the mixture programs could be appreciably improved other than to undertake new experimental measurements. CDATAl is now a unique resource which not only makes possible large-scale processing of mixture data but can provide high-quality pure compound data to any user program, or be accessed interactively via a computer terminal.

Another critical area for the evaluation of VLE data is the equation of state used to model the equilibrium vapor phase. The Laboratory has expended a great deal of manpower in that area also but it still remains the weakest link in the processing chain. Some improvements in our procedures are still possible but the basic problem is the weakness of the current equation-of-state technology. The equations and mixing rules available do not permit us to process large numbers of mixtures with complete confidence. This weakness has been an important factor in the choice of systems to be processed initially.

This paper discusses how the four major preliminary tasks were handled. It then describes in some detail what has been done to provide the best possible equations of state available to the VLE data processing programs. Finally, the manner in which the mixture data evaluations will be published is described.

An appendix to this paper presents some recommendations for the improvement of future publications of mixture data.

2. Relevant Literature Documents

The following four procedures were used to identify relevant documents:

1. The thirty-five journals listed in table 1 were "clean-swept". A trained searcher leafed through each journal page-by-page examining each article for relevant data. The title and abstract were not assumed to be sufficient to establish relevancy or nonrelevancy.

2. About one hundred compilations, bibliographies, review articles, etc., dealing with the properties and compounds of interest were searched for relevant literature citations.

3. The relevant sections in all the Chemical Abstract volumes from Volume 1 (1907) through Volume 64 (1965) were searched for relevant documents. After 1965, the Bulletin of Chemical Thermodynamics (included in the second procedure above) was used instead of Chemical Abstracts.

4. The literature references in each retrieved document were inspected to identify other documents which were probably relevant.

Each retrieved document was assigned a Master Reference List (MRL) number and its citation stored on the MRL tapes. The MRL number is used as the identification number

TABLE 1. Journals which have been clean-swept

1. Acta Chemica Scandinavica
2. AIChE Journal
3. Australian Journal of Chemistry
4. Berichte der Bungengesellschaft fuer Physikalische Chemie
5. Canadian Journal of Chemistry
6. Chemical Engineering Progress
7. Chemical Engineering Progress, Symposium Series
8. Chemical Engineering Science
9. Chemical Reviews
10. Chemical Society Reviews
11. Collection Czechoslovak Chemical Communications
12. Discussions of the Faraday Society
13. Fluid Phase Equilibria
14. Industrial and Engineering Chemistry
15. Industrial and Engineering Chemistry, Fundamentals
16. Industrial and Engineering Chemistry, Process Design and Development
17. Journal of American Chemical Society
18. Journal of Applied Chemistry (USSR)
19. Journal of Chemical and Engineering Data
20. Journal of Chemical Physics
21. Journal of Chemical Society, Faraday Transactions I
22. Journal of Chemical Society, Faraday Transactions II
23. Journal of Chemical Society (London)
24. Journal of Chemical Thermodynamics
25. Journal de Chimie Physique
26. Journal fuer Praktische Chemie
27. Journal Physical and Chemical Reference Data
28. Journal of Physical Chemistry
29. Thermochimica Acta
30. Pure and Applied Chemistry
31. Russian Journal of Physical Chemistry
32. Transactions of American Institute of Chemical Engineers
33. Transactions of the Faraday Society
34. Zeitschrift fuer Physikalische Chemie (Frankfurt)
35. Zeitschrift fuer Physikalische Chemie (Leipzig)

[1]Figures in brackets indicate literature references.

for each document cited in the evaluation reports. That number serves as the connecting link between the data tables, the summary lists, and the bibliographies.

Since 1975, documents skimmed from the newly-published literature have been assigned MRL numbers above 40,000. However, other post-1975 documents retrieved in one of the four ways described above were numbered in the regular sequence. It is estimated that there are currently about 27,000 documents relevant to this evaluation project of which over 25,000 have been retrieved.

3. Computer Program Libraries

Currently, fifty-six computer programs are in use for the evaluation of pure compound and mixture data. One group of dual-useage programs is used to list relevant documents for a given project and to prepare bibliographies of various kinds for either pure compound or mixture projects.

Another group of programs performs the following functions for pure compound projects: conversion of units to SI units, correction of temperatures to the IPTS-68 scale, plotting of property data in residual form for visual comparison of data sets, fitting to selected correlation equations for screening purposes, final correlation of selected data for each compound-property combination, and storage of the final correlations and other compound information in the pure compound data bank CDATA1.

The group of programs used for mixture projects performs functions which are somewhat more varied. The conversion to SI units is accomplished for all three mixture properties but conversion to the IPTS-68 temperature scale is necessary only for the vapor-liquid equilibrium data. Also, the vapor-liquid equilibrium data must be reduced to activity coefficient and excess Gibbs function values. Once those steps are accomplished, the "plotting" programs prepare all the plots needed for the evaluations, the "evaluation" programs perform all the calculations for the evalution tests, the "quality" programs assign a quality rating to each set of data, and the "table" programs prepare the final tabulations.

4. Pure Compound Data

The evaluation of the mixture data can be no better than the pure compound data used. The liquid-phase activity coefficients calculated from vapor-liquid equilibrium data are very sensitive to the pure compound vapor pressure data and, because the Gibbs-Duhem thermodynamic consistency test must use the activity coefficients, the consistency test results are equally sensitive. Excess volume values calculated from mixture density or volume data are equally sensitive to the pure compound density values used; in fact, if the experimenter does not report pure compound density values and values from another source must be used, the calculated V^E values are seldom good.

The three experimentally-measured pure compound properties stored in CDATA1 for use in the mixture evaluations are the vapor pressure, the liquid density, and the second virial coefficient. The third virial coefficient can also be stored but is seldom available. If the virial coefficients are not available, an empirical equation of state must be available from CDATA1. The pure compound data evaluation procedures and the correlation forms in current use are described briefly in this section.

The reader should note that the objective of the pure compound work is not the evaluation of individual sets of pure compound data per se, but rather the development of correlation equations which will represent in CDATA1 the evaluator's best definition of true property-temperature curves. That "best correlation" will usually use all or parts of several sets of data from the literature; a record of which literature data sets contribute to the correlation is provided by the sets of "Used" and "Not Used" MRL numbers stored with the correlation in CDATA1. This approach to the pure compound data evaluation reflects the fact that the primary mission of the Laboratory's Data Center is the evaluation of mixture data, and that mission requires highly accurate pure compound data to be available automatically to the mixture programs from the computer itself.

4.1. Initial Processing Steps

All the retrieved data for the given property (regardless of its quality or source) for the compound class being processed are transcribed, keypunched, and loaded into computer disk files. The formats of those files are fixed by the input requirements of the various CON (for unit conversion) programs which are used to convert all the data to SI units and to correct temperatures to the IPTS-68 scale if necessary. Before conversion, the disk files are edited using EDIT programs which identify obvious errors. Printouts of the corrected files are then carefully proofread agains the original literature documents to eliminate all remaining transcription and keypunching errors. Verification of the correctness of the disk files by the evaluator marks the end of the document processing part of a compound project.

The data processing phase begins with conversion of the literature data to SI units (as necessary) and storage in new disk files. The CON programs also calculate a residual (deviation from a correlation equation) for each data point, and those residuals are stored with the converted data.

The residuals are plotted by a CalComp plotter on compound plots which are nine inches high and up to thirty inches in length. Each set of data (literature source) has its own symbol and the symbols are keyed on the plot to the Master Reference List (MRL) numbers of the documents. The MRL number is the document's file number and relates the point on the residual plot to the original literature document. The plotting program selects the ordinate and abscissa scale factors for each compound-property combination to give the maximum possible magnification of the scatter within the size limitations. If one or more "wild" points depress the ordinate scale, or if the number of data points in a certain temperature range for a given compound-property combination is so great that the points plot as a solid band, new plots are made with narrower ordinate and abscissa scale ranges in order to get the optimum magnification of the scatter.

Once all the wild points (if any) have been corrected or verified to be the values actually reported in the literature, the data files are sorted to put the points for each compound-property combination in the order of ascending temperature. The original, converted and sorted disk files for each of

the compound-property combinations are then copied onto magnetic tape for long-term storage. Those files represent all the literature, i.e., no data points have been deleted from those three files. Duplicate points (points reported by the same authors in more than one literature document) have been deleted.

The actual evaluation and selection of the data begins at this point in the processing. The evaluator has all the available data stored on disk in SI units, with all transcription and keypunch errors eliminated, and with the data for each compound ordered with respect to temperature. Also, the data for each compound have been plotted in residual form at a magnification which permits distinction between data sets. The procedure followed from this point on differs with the property being processed.

4.2. Vapor Pressure Data

Data points which deviate widely from the bulk of the reported data are obvious on the residual plots. The literature documents from which those points come are checked and usually provide justification for deletion. Consequently, it is usually easy to trim all the outriders from the combined data set when there are sufficient good data to define the "groove" on the plot where the true property curve probably lies.

Often, there will not be a large number of data sets for a given compound and the location of the "groove" on the residual plot is not obvious. A typical situation involves two sets of data which cover the same temperature range but which disagree. Sometimes, a study of the individual literature documents makes a choice between the sets obvious, but there is often no obvious basis for a choice. (The evaluator must be wary of placing too much reliance on the stated accuracies in the documents.) Sometimes, other data sets in lower and higher temperature ranges will give guidance as to which set of data falls closer to the probable groove. If no justification can be found to eliminate either set, the evaluator may decide to retain both and let them both influence the final correlation equation, according to the weights which can be assigned to reflect any differences in the levels of assumed accuracies.

After a thorough study of the residual plots and the literature documents from which the points came, the evaluator marks the points to be deleted from the disk file. After those deletions are made, screening with selected fitting equations begins. Reliable equations are used to magnify scatter further beyond that feasible on the residual plots. The percent deviation of each point from the equation is displayed, along with the deviation of each point in terms of a multiple of the root-mean-squared deviation (RMSD) for the entire fit. The percent error values are useful at low vapor pressure values while the RMSD multiples are more useful at high pressures. The points with the largest deviations are trimmed away in successive screening fits. Each successive fit magnifies the scatter displayed due to the decrease in the RMSD. The deletion of data points continues until the evaluator reaches the stage where further deletions will not help locate the true curve. A simple reduction of the RMSD is not the objective; often points which disagrees with each other (i.e., scatter) are retained because there is no way to tell which points are more accurate and retention of them all increases the probability that the fitting equation will represent the true curve.

The screening fits are not made over the entire temperature range from the melting point to the critical point. The fits for regions 2 and 4 defined below are used.

Once the evaluator has reduced the set of retained data points for each compound to that set which he believes best represents the true property curve, final fits of the selected data are made over each of the following temperature regions (if there are data points in the region):

Region	Temperature range
1	T_m to T_c
2	T_m to $T_b + 20$
3	T_m to $0.85\ T_c$
4	$T_b - 20$ to T_c

The T_m, T_b, and T_c represent the melting point, normal boiling point, and critical point temperatures, respectively. The six equations available for use by the evaluator for any of the four regions are given in table 2. The use of the narrower temperature ranges in regions 2, 3, and 4 permit better representation of the data than can be obtained in a region 1 fit. The regions overlap widely so that any given request for vapor pressure data will always fall within one temperature region. Whenever a user program goes to the pure compound data bank (CDATA1) for vapor pressure data, the hierarchy of choice is as follows: region 2, region

TABLE 2. Vapor pressure correlations in current use

Name	Equation
Riedel [4]	$\ln_e P' = A + B \ln_e T + \frac{C}{T} + DT^6$
Frost-Kalkwarf [5, 6]	$\ln_e P' = A + B \ln_e T + \frac{C}{T} + \frac{DP'}{T^2}$
Riedel-Plank-Miller-3	$\ln_e P' = A + \frac{B}{T} + CT + DT^2 + ET^3$
Riedel-Plank-Miller-2 [7, 8]	$\ln_e P' = A + \frac{B}{T} + CT + DT^2$
Vapres-2	$\ln_e P' = A + \frac{B}{T} + CT + DT^2 + E \ln_e T$
Wagner [9]	$\ln_e P'_r = A\frac{1 - T_r}{T_r} + B\frac{(1 - T_r)^{1.5}}{T_r} + C\frac{(1 - T_r)^3}{T_r} + D\frac{(1 - T_r)^6}{T_r}$

TABLE 3. Liquid density correlations in current use

Name	Equation
Rackett [10, 11, 12, 13, 14]	$\frac{V^L}{FW}=\frac{1}{d}=\frac{8.31441T_c}{(FW)P_c}Z_{RA}^{[1+(1-T_r)^{2/7}]}$
Francis, $<(T_c-20)$ [15]	$d=A-BT-\frac{C}{D-T}$
Francis, $>(T_c-50)$ [15]	$d=[A(T_c-T)]^{1/B}+d_c$

4, region 3.

When obtaining the final fits for the four regions, the evaluator will use those two or three equations for each region which experience has shown to work best, and then choose the one for each region which represents the selected data best. The Wagner equation is normally used for the region 1 and region 4 fits when the critical point is available because it is a reduced equation and must pass through the critical point. For regions 2 and 3, whatever equation works best is chosen. The constants for the selected equations for each compound are then loaded into CDATA1 as part of the compound's record. The equations' eight character identifiers are loaded with their constants so that a user program will know which equation has been stored for a given region.

4.3. Liquid Density Data

The procedure for liquid density data is the same as for the vapor pressure except that the following temperature regions are used:

Region	Temperature range
1	T_m to T_c
2	T_m to T_b+20
3	T_m to T_c-30
4	T_c-50 to T_c

The fitting regions for liquid density data differ from those used for vapor pressure data because of the different shapes of the two property curves and differences in the correlation equations which are available. The equations in current use for PCFILE are listed in table 3.

4.4. Second Virial Coefficients

The second virial coefficient data are not subjected to the least-squares screening fits used for vapor pressure and liquid density. Instead of using just one equation to calculate

TABLE 4. Second virial coefficient correlations in current use

Kreglewski [16, 17]	$\frac{B}{2V^*}=R_1-(1-R_1)[\exp(u/kT)-1]-(R_2-1)[\exp(-u/kT)-1]$ $-(R_3-R_2)[\exp(-0.21\ u/kT)-1]$ where $u/k=a\left[1+\frac{b}{T}\right]$
Pitzer-Curl [18]	$\frac{BP_c}{82.0568T_c}=f^{(0)}(T_r)+\omega_i f^{(1)}(T_r)$ $f^{(0)}(T_r)=0.1445-0.330T_r^{-1}-0.1385T_r^{-2}-0.0121T_r^{-3}$ $f^{(1)}(T_r)=0.073+0.46T_r^{-1}-0.5T_r^{-2}-0.097T_r^{-3}-0.0073T_r^{-8}$
O'Connell-Prausnitz [19]	$\frac{BP_c}{82.0568T_c}=f^{(0)}(T_r)+\omega_H f^{(1)}(T_r)+f_\mu(\mu_r,T_r)-\eta f_a(T_r)$ $f^{(0)}(T_r)=$ same as in Pitzer-Curl. $f^{(1)}(T_r)=$ same as in Pitzer-Curl. $f_\mu(\mu_r,T_r)=-5.237220+5.665807\ln\mu_r$ $-2.133816(\ln\mu_r)^2+0.2525373(\ln\mu_r)^3$ $+[5.769770-6.181427(\ln\mu_r)+2.283270(\ln\mu_r)^2$ $-0.2649074(\ln\mu_r)^3]T_r^{-1}$ $f_a(T_r)=\exp[6.6(0.7-T_r)]$ $\mu_r=\frac{10^5\mu^2P_c}{T_c^2}$
Tsonopoulos [20, 21]	$\frac{BP_c}{82.0568T_c}=f^{(0)}(T_r)+\omega f^{(1)}(T_r)+f^{(2)}(T_r)$ $f^{(0)}(T_r)=0.1445-0.330T_r^{-1}-0.1385T_r^{-2}-0.0121T_r^{-3}-0.000607T_r^{-8}$ $f^{(1)}(T_r)=0.0637+0.331T_r^{-2}-0.423T_r^{-3}-0.008T_r^{-8}$ $f^{(2)}(T_r)=aT_r^{-6}-bT_r^{-8}$
Nothnagel et al. [22]	$B=b-82.0568\ TK$ $-\ln(82.0568\ TK)-\frac{\Delta H}{1.98719}-\frac{\Delta S}{1.98719}$ $\frac{\Delta S}{1.98719}=1.1\frac{\Delta H}{1.98719}+\ln(3.06\ b)$ $-\Delta H=dT_b(8.75+4.576\log_{10}T_b)-1.98719T_b$

residuals in the unit conversion step (CON programs), up to five equations can be used for the second virial coefficient data. (The five equations in current use are shown in table 4.) The correlation which best represents all the literature data for a compound (based on the residual plot) is then chosen to represent that compound in CDATAI. There is no fitting of the experimental data, and there is only superficial evaluation of those data.

5. Equations of State

The liquid-phase activity coefficients, γ_i, are used to quantify the departure of the liquid phase from ideal solution behavior, while the gas-phase fugacity coefficients, $\hat{\phi}_{i,P}$, and ϕ_{i,P'_i}, serve the same purpose for the gas phase. The experimentally determined liquid-phase activity coefficient is necessary because we cannot predict with any reasonable degree of accuracy the fugacity coefficient for a component in a liquid mixture. Program PTXY2 uses the equation

$$\gamma_i = \frac{y_i P \hat{\phi}_{i,P}}{x_i P'_i \phi_{i,P'_i}} \exp\left[-\frac{V_i^L (P - P'_i)}{RT} \right] \quad (1)$$

to calculate the liquid-phase activity coefficients from experimental P, T, x, y data. The fugacity coefficients $\hat{\phi}_{i,P}$ and ϕ_{i,P'_i} must be predicted by an assumed equation of state. Fortunately, the fugacity coefficients appear in a ratio and, because the errors in their individual values will usually be in the same direction, the errors tend to cancel. Nevertheless, the numerical values of the γ_i's calculated from the experimental data are a function of the equation of state assumed. They will be strong functions of the assumed equation in systems where the liquid mixture is relatively ideal (γ_i values close to 1.0) and the system pressure is high (ϕ_i values deviate widely from 1.0). Numerical examples of that dependence are given later.

5.1. Equations of State in Current Use

The equations of state currently in use in CDATAI are as follows:

1. Virial equation through the C terms.
2. Virial equation through the B terms.
3. Redlich-Kwong equation:
   a. Unmodified form [23]
   b. Chueh-Prausnitz modification [24]
   c. Lu modification [25]
   d. Peng-Robinson equation [26]
4. Benedict-Webb-Rubin equation:
   a. Unmodified form [27]
   b. Sood-Haselden modification [28]
   c. Starling modification [29, 30]

There is no dependable predictive correlation for the virial C coefficients and very few experimental values are available; hence the virial equation usually has to be truncated to the B terms. The correlations used to represent the B values in CDATAI are given in table 4.

Two forms of both the Redlich-Kwong and Benedict-Webb-Rubin equations are usually stored in CDATAI. One form is stored in the "First Choice" location and the other in the "Second Choice" location. The First Choice form is the one which should represent the pure compound properties best. The Second Choice form is added (when necessary) to ensure a possible compatible match between the two components in a binary mixture being processed by program PTXY2. "Compatible match" means the two equations of state can use the same set of mixing rules used to calculate the cross terms. The Starling modification of the BWR cannot be "mixed" with the unmodified or Sood-Haselden forms in a binary mixture. The Sood-Haselden and the unmodified BWR forms have the same mixing rules.

For the Redlich-Kwong forms, the Lu and Chueh-Prausnitz forms are compatible. The unmodified form and the Peng-Robinson equation are not compatible with each other, and neither can be mixed with the Lu or the Chueh-Prausnitz forms. If program PTXY2 cannot find a compatible match between the RK-type of equations stored in CDATAI, it uses the unmodified form for both components as a default.

Among the second virial coefficient correlations, only the Pitzer-Curl and the O'Connell-Prausnitz equations are compatible insofar as the mixing rules are concerned. When a Pitzer-Curl plus Tsonopoulos, or an O'Connell-Prausnitz plus Tsonopoulos, match is found, the Tsonopoulos correlation is used for both components. Whenever PTXY2 must use correlations which are not compatible, the First Choice entries for the two components are used to calculate the B_{11} and B_{22} values, and the Amdur-Mason rule [31] is used to calculate B_{12} from the B_{11} and B_{22} values. That is always done when the Kreglewski equation appears for both components; mixing rules have never been developed for that equation.

5.2. Equation Choice for Mixtures

The problem of choosing the best equation of state for a given binary mixture adds considerable complexity to program PTXY2. Table 5 shows the sequence of the matches which program PTXY2 attempts to make between the various correlations which might be stored for the two components in a binary system. If a match has not been achieved with either the virial equation through the C terms or a pair of compatible BWR forms (first 18 attempts in table 5), the pressure and temperature levels relative to the critical point are checked using the empirical relationship shown after the 18th comparison. Depending upon the result, PTXY2 searches for a match between two compatible Redlich-Kwong forms (comparisons 19A through 36A), or between two compatible B correlations (comparisons 19B through 30B). The A and B paths both require T_c and P_c to be available. If both are not available, the path C options are tried. If P_c is not available but T_c and the normal boiling points are, the D path options are tried. If none of the matches listed in table 5 can be made, program PTXY2 must use the ideal gas equation.

The evaluator runs program PTXY2 once to see what equation-of-state match will be produced for each binary system processed in the project. Each match is investigated to see if empirical correction terms for the mixing rule have been published for that particular binary. (The literature for that area has been summarized in the documentation for program STATE.) If empirical correction factors based on experimental PVT data for the given mixture are available, they can be inputted to PTXY2 for the second (and final) run.

TABLE 5. Sequence of comparisons by program PTXY2 of equation-of-state identifiers stored in CDATA1

Comparison number	Component 1 Choice	Component 1 Stored Equation	Component 2 Choice	Component 2 Stored Equation
1[a]	1	Virial through C's	1	Virial through C's
2	1	Starling BWR	1	Starling BWR
3	1	Sood-Haselden BWR	1	Sood-Haselden BWR
4[b]	1	Sood-Haselden BWR	1	Unmodified BWR
5[b]	1	Unmodified BWR	1	Sood-Haselden BWR
6	1	Unmodified BWR	1	Unmodified BWR
7	1	Sood-Haselden BWR	2	Sood-Haselden BWR
8[b]	1	Sood-Haselden BWR	2	Unmodified BWR
9[b]	1	Unmodified BWR	2	Sood-Haselden BWR
10	1	Unmodified BWR	2	Unmodified BWR
11	2	Sood-Haselden BWR	1	Sood-Haselden BWR
12[b]	2	Sood-Haselden BWR	1	Unmodified BWR
13[b]	2	Unmodified BWR	1	Sood-Haselden BWR
14	2	Unmodified BWR	1	Unmodified BWR
15	2	Sood-Haselden BWR	2	Sood-Haselden BWR
16[b]	2	Sood-Haselden BWR	2	Unmodified BWR
17[b]	2	Unmodified BWR	2	Sood-Haselden BWR
18	2	Unmodified BWR	2	Unmodified BWR
	If $(P/T)\geqslant(P_c/3T_c)$, use *A* path below; if not, switch to *B* path			
19A	1	Lu RK	1	Lu RK
20A[c]	1	Lu RK	1	Chueh-Prausnitz RK
21A[c]	1	Chueh-Prausnitz RK	1	Lu RK
22A	1	Chueh-Prausnitz RK	1	Chueh-Prausnitz RK
23A	1	Peng-Robinson RK	1	Peng-Robinson RK
24A	1	Unmodified RK	1	Unmodified RK
25A[c]	1	Lu RK	2	Chueh-Prausnitz RK
26A	1	Chueh-Prausnitz RK	2	Chueh-Prausnitz RK
27A	1	Peng-Robinson RK	2	Peng-Robinson RK
28A	1	Unmodified RK	2	Unmodified RK
29A[c]	2	Chueh-Prausnitz RK	1	Lu RK
30A	2	Chueh-Prausnitz RK	1	Chueh-Prausnitz RK
31A	2	Peng-Robinson RK	1	Peng-Robinson RK
32A	2	Unmodified RK	1	Unmodified RK
33A	2	Chueh-Prausnitz RK	2	Chueh-Prausnitz RK
34A	2	Peng-Robinson RK	2	Peng-Robinson RK
35A	2	Unmodified RK	2	Unmodified RK
36A[d]		Any RK form		Any RK form
	If $(P/T)\leqslant(P_c/3T_c)$, use virial equation truncated after *B* coefficient and one of the *B* correlation matches shown in the *B* path below			
19B	Pitzer-Curl		Pitzer-Curl	
20B[e]	O'Connell-Prausnitz		Pitzer-Curl	
21B[e]	Pitzer-Curl		O'Connell-Prausnitz	
22B	O'Connell-Prausnitz		O'Connell-Prausnitz	
23B[f]	Tsonopoulos (nonpolar)		Pitzer-Curl	
24B[f]	Pitzer-Curl		Tsonopoulos (nonpolar)	
25B[f]	Tsonopoulos (nonpolar)		O'Connell-Prausnitz	
26B	O'Connell-Prausnitz		Tsonopoulos (nonpolar)	
27B	Tsonopoulos		Tsonopoulos	
28B	Nothnagel		Nothnagel	
29B[g]	Kreglewski		Kreglewski	
30B[h]	Any *B* correlation		Any *B* correlation	
	If both T_c and P_c are nonavailable, follow the *C* path			
19C[g]	Kreglewski		Kreglewski	
20C[h]	Any *B* correlation		Any *B* correlation	
	If T_c is available but P_c is not, follow the *D* path			
19D	Nothnagel		Nothnagel	
20D[g]	Kreglewski		Kreglewski	
21D[h]	Any *B* correlation		Any *B* correlation	

[a] The *C* coefficients from the Chueh-Prausnitz correlation plus the *B* coefficients from whatever correlations are stored.
[b] The Sood-Haselden form is used for both components with the temperature coefficients set equal to zero for the component for which the unmodified form appears.
[c] The Lu modification of the Redlich-Kwong equation is used with the Ω_a and Ω_b values set equal to the Chueh-Prausnitz values for the component for which the Chueh-Prausnitz form was stored.
[d] The unmodified Redlich-Kwong state is used for both compounds.
[e] The *B* values for both components are calculated from the applicable form of the O'Connell-Prausnitz correlation.
[f] The *B* values for both components are calculated from the applicable form of the Tsonopoulos correlation.
[g] The Amdur-Mason rule [31] is used to calculate B_{12} because there are no mixing rules for the Kreglewski equation.
[h] When a *B* correlation has been stored for both components but none of the matches listed (19B through 29B) occurs, PTXY2 uses the stored correlations for B_{11} and B_{22} and calculates B_{12} from the Amdur-Mason rule [31].

5.3. Impact on Calculated Results

It can be seen from the above discussion that considerable effort has been expended in an attempt to provide good fugacity coefficient values to program PTXY2. Despite that effort, the accuracy of the predicted fugacity coefficients can best be characterized as uncertain. Even if the pure component PVT data for the two components have been well correlated by the same equation (e.g., the Starling modification of the BWR equation) the mixture predictions at higher pressures may not be accurate, unless mixture PVT data happen to be available and empirical correction factors for the mixing rules have been determined to force agreement with the experimental data. The amount of mixture PVT data in the literature is not large, and much of what does exist has not been reduced to the parameters needed to make the empirical equations of state dependable for the mixture. Hence, program PTXY2 must usually rely on general mixing rules and their accuracies can only be characterized as spotty.

The equation-of-state problem can be extremely important for mixture classes such as the C_3 hydrocarbons where the activity coefficients are close to 1.0 and the VLE data are often needed at conditions where the gas phase deviates widely from ideal gas behavior. At low pressures, the calculated results are relatively insensitive to the equation of state used. For example, the midpoint G^E values calculated for the 244.85 K (− 28.3 °C) VLE data for propene + propane measured by Hirata and Hakuta [32] using various equations of state are as follows:

Ideal gas	69.60
Unmodified BWR	65.84
Sood-Haselden BWR	64.66
Starling BWR	67.26
Virial through C, B's and C's known	64.53
Virial through B, B's known	64.55
Virial through B, Pitzer-Curl	64.54
Virial through B, O'Connell-Prausnitz	64.55
Virial through B, Tsonopoulos	64.49
Virial through B, Nothnagel	64.81
Unmodified RK	65.53
Chueh-Prausnitz RK	66.58
Lu RK	66.64
Peng-Robinson	65.12

The T_r values are 0.67 and 0.66 for propene and propene at 244.85 K, and the system pressure ranges from 179 to 226 kPa for the Hirata et al. set of data.

The situation changes drastically as the system pressure approaches the component critical pressures. That is illustrated in table 6 for the four Redlich-Kwong forms and the three BWR forms. The G^E values,

$$G^E = RT[x_1 \ln\gamma_1 + x_2 \ln\gamma_2] \quad (2)$$

shown there were calculated from the 160 °F (344.26 K) P, T, x, y values reported by Laurance and Swift [33]. Note that the calculated G^E values vary from 18.54 to 65.85 J mol^{-1} at the midpoint ($x_1 = 0.5$), depending upon the equation of state used.

The VLE data measured by Laurance and Swift were actually total-pressure data; only the P, T, x values were measured experimentally. They also measured the vapor compressibility (PVT data) of the pure components at the conditions of the VLE measurements, and correlated that data with the unmodified Redlich-Kwong equation. The original mixing rules for the Redlich-Kwong equation were used to predict the mixture compressibilities. Laurance and Swift calculated their y values by integration of the isothermal co-existence equation, including the fugacity coefficients which were calculated from the Redlich-Kwong equation. The y values obtained must be thermodynamically consistent with the measured P, T, x values (if the numerical integration of the coexistence equation was correctly performed), but the values obtained are dependent upon the equation of state used.

Presuming the integration of the co-existence equation was correctly performed (and there is no reason to assume otherwise), the P, T, x, y values reported by Laurance and Swift are thermodynamically consistent sets of VLE data. It is interesting to see what happens when we test a set of their

TABLE 6. Dependence of calculated G^E values on the equation of state used. Propene(1) + propane(2) system at 344.26 K. Data of Laurance and Swift [33]

	Calculated G^E values						
x_1	RK	RKCP	RKLU	RKPR	BWR	BWRSOO	BWRSTR
0.1	6.62	10.98	10.88	6.04	9.33	9.34	21.36
0.2	12.12	20.10	19.91	11.07	17.09	17.12	39.05
0.3	16.28	27.05	26.78	14.87	23.00	23.04	52.60
0.4	19.04	31.65	31.34	17.40	26.93	26.97	61.66
0.5	20.27	33.72	33.38	18.54	28.70	28.74	65.85
0.6	19.90	33.09	32.75	18.21	28.18	28.22	64.81
0.7	17.75	29.52	29.22	16.26	25.15	25.18	58.10
0.8	13.83	22.95	22.71	12.68	19.56	19.59	45.41
0.9	7.93	13.14	13.00	7.28	11.21	11.22	26.20

Redlich-Kwong forms:
RK Original equation
RKCP Chueh-Prausnitz modification
RKLU Lu modification
RKPR Peng-Robinson equation

Benedict-Webb-Rubin forms:
BWR Original equation
BWRSOO Sood-Haselden modification
BWRSTR Starling modification

reported P, T, x, y values for consistency with the Gibbs-Duhem equation using different equations-of-state. For this test, the vapor pressure standard state (pure liquid i at the system temperature and a pressure equal to its vapor pressure) will be used for each component. For that standard state, the Gibbs-Duhem equation can be written as

$$\int_{x_1=0}^{x_1=1} \ln(\gamma_1'/\gamma_2')dx_1 + \frac{1}{RT}\int_{P'_2}^{P'_1} V^L dP = 0 \qquad (3)$$

where the primes on γ_1' and γ_2' denote the vapor pressure standard state. The use of that standard state makes the activity coefficients independent of the Poynting term. The correction term can be evaluated accurately because in this case the authors reported experimental V^L values at the temperature of the VLE data.

The agreement of the data with the Gibbs-Duhem equation will be characterized by an area ratio. Evaluation of eq. (3) involves three areas; the first integral has both positive and negative areas, and the second (correction) integral will be either positive or negative depending upon whether P_1' or P_2' is larger. The area for the correction integral is added to the area of like sign from the first integral to give total positive and negative areas. The smaller of the two (positive or negative) is divided by the larger and the sign is dropped. The result is an area ratio which will fall between 0.0 and 1.0, and will be 1.0 when the VLE data agree exactly with the integrated form of the Gibbs-Duhem equation.

The Laurance and Swift data set at 344.26 K (same as in table 6) will be used as an example to demonstrate the effect of the equation of state on the consistency test. At 344.26 K, T_r is 0.94 and 0.93 for propene and propane, respectively, and at those conditions the fugacity coefficients deviate from unity enough to have a major impact on the γ_1 and γ_2 values which are close to unity for this almost ideal liquid mixture. Also, the correction term (second integral) in eq. (3) can become roughly equal in size to the two areas in the first term. (Note. If the total pressure standard state were used, the Gibbs-Duhem equation would be

$$\int_{x_1=0}^{x_1=1} \ln(\gamma_1/\gamma_2)dx_1 + \frac{1}{RT}\int_{P'_2}^{P'_1} V^E dP = 0 \qquad (4)$$

and the correction term would be much smaller. However, each of the activity coefficients would involve the Poynting term,

$$\int_{P_i'}^{P} V_i^L dP \qquad (5)$$

which always involves the use of hypothetical V_i^L values for the component with a P_i' greater than the system pressure P. When experimental V^L data are available at the VLE data set conditions, eq. (3) permits the application of the consistency test without involving hypothetical V_i^L values.)

Figure 1 shows the $\ln\gamma_1'/\gamma_2'$ versus x_1 curves calculated with four of the equations of state listed in table 6. (The Sood-Haselden BWR equation gave essentially the same results as the unmodified BWR. The Lu and the Peng-Robinson forms gave $\ln\gamma_1'/\gamma_2'$ curves which fell between the two Redlich-Kwong curves plotted in figure 1). The value of the correction integral in eq. (3) is 0.018507 at 344.26 K. Adding that area to the positive areas in figure 1 gives the following corrected area ratios:

Starling BWR	0.84
Unmodified BWR	0.96
Chueh-Prausnitz RK	0.84
Unmodified RK	0.79

The unmodified RK equation is the same equation used by Laurance and Swift, but the constants used by program PTXY2 were the general ones calculated from the critical T and P whereas Laurance and Swift determined specific constants for pure propene and propane from their experimental compressibility data. That difference in the Redlich-Kwong constants used is sufficient to change the area ratio from 1.0 (the Laurance and Swift data should be thermodynamically consistent) to 0.79.

This example shows that one cannot test the thermodynamic consistency of almost ideal liquid systems (γ_i's close to 1.0) at pressures where the fugacity coefficients are numerically important, unless one has an equation of state specifically fitted to experimental PVT data for the mixture. It also shows that one should not measure total pressure VLE data for such a system unless one also measures the necessary PVT data for the vapor mixtures, because the y values calculated from the experimental P, T, x values will be affected by the error in the assumed equation of state.

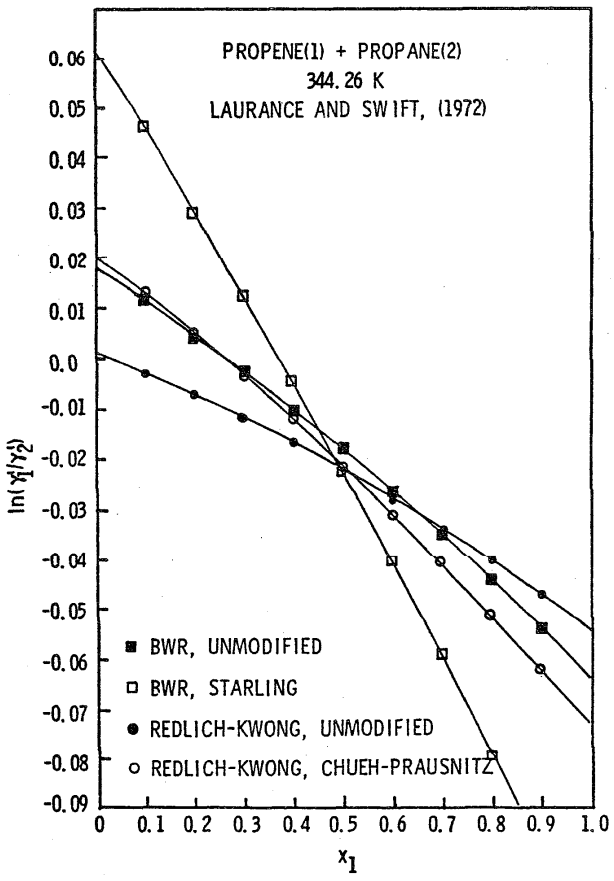

FIGURE 1. Effect of the equation of state on the $\ln(\gamma_1'/\gamma_2')$ versus x_1 plot for the propene(1) + propane(2) system at 344.26 K. Data of Laurance and Swift [33].

Another method often used to check the thermodynamic consistency of VLE data is to calculate y values from the experimentally measured P, T, x values and then compare the calculated values with the experimentally measured values. The calculated y values are consistent with the P, T, x values and the experimental y values are said to be consistent if they agree with the calculated values. The calculated y values are, of course, dependent upon the equation of state used, and for a high pressure set of data such as that used in the example above, the data will contain some undetermined error due to the equation of state. Hence, the method does not avoid the equation of state problem; it just gives results which do not display the possible error as graphically as does the equal area ratio test.

6. Nomenclature

a,b	Empirical constants in Kreglewski correlation for the second virial coefficient.
b	Empirical constant in the Nothnagel-Abrams-Prausnitz correlation for the second virial coefficient.
A,B,C,D,E	Empirical constants in correlation equations.
B	Second virial coefficient, $cm^3\ mol^{-1}$.
d	Differential operator.
d	Density, $g\ cm^{-3}$
d_c	Critical point density, $g\ cm^{-3}$.
FW	Acronym for formula weight.
G^E	Molar excess Gibbs function, $J\ mol^{-1}$.
H^E	Molar excess enthalpy, $J\ mol^{-1}$.
M^E	Molar excess property.
MRL	Acronym for Master Reference List.
P	Pressure, kPa.
P_i'	Vapor pressure of compound i, kPa.
P_c	Critical pressure, MPa.
PTx	The set of variables measured experimentally in total pressure vapor-liquid equilibrium data.
$PTxy$	The complete set of variables which can be measured experimentally in vapor-liquid equilibrium data.
PVT	Acronym for pressure-volume-temperature.
R	Gas constant, $8.31441\ J\ K^{-1}\ mol^{-1}$.
R_1,R_2	Empirical constants in the Kreglewski correlation for the second virial coefficient.
RMSD	Acronym for the root mean squared deviation.
T	Temperature, K.
T_m	Melting temperature, K.
T_b	Boiling temperature at 101.325 kPa.
T_c	Critical temperature, K.
V^L	Liquid molar volume of a mixture.
V_i^L	Liquid molar volume of compound i.
V^E	Molar excess volume, $cm^3\ mol^{-1}$.
V^*	Empirical constant in the Kreglewski correlation for the second virial coefficient.
VLE	Acronym for vapor-liquid equilibrium.
x_i	Mole fraction of component i in the liquid phase.
y_i	Mole fraction of component i in the vapor phase.
Z_{RA}	Empirical constant in the Rackett equation for the liquid density.
γ_i	Liquid-phase activity coefficient of component i.
μ	Dipole moment.
μ_r	Reduced dipole moment.
ϕ_{i,P'_i}	Fugacity coefficient for pure component i at its vapor pressure and the system T.
$\hat{\phi}_{i,P}$	Fugacity coefficient for component i in a gas mixture at the pressure P.
ω	Acentric factor.

Acknowledgments

This work was supported by the Office of Standard Reference Data and by the Industrial Participants in the Thermodynamics Research Laboratory.

8. References

[1] Yuan, W. I.-C., Palmer, D. A., and Smith, B. D., Ind. Eng. Chem., Fundam. **11**, 387 (1972).
[2] Palmer, D. A., Yuan, W. I.-C., and Smith, B. D., Ind. Eng. Chem., Fundam. **11**, 397 (1972).
[3] Stookey, D. J., and Smith, B. D., Ind. Eng. Chem., Process Des. Develop. **12**, 372 (1973).
[4] Riedel, L., Chem.-Ing.-Tech. **26**, 83 (1954).
[5] Frost, A. A., and Kalkwarf, D. R., J. Chem. Phys. **21**, 264 (1953).
[6] Miller, D. G., Ind. Eng. Chem., Fundam. **2**, 78 (1963).
[7] Miller, D. G., Ind. Eng. Chem. **56**, 46 (1964).
[8] Miller, D. G., Lawrence Radiation Laboratory (Livermore) Report UCRL-7292, 1963.
[9] Wagner, W., Cryogenics **13**, 470 (1973).
[10] Rackett, H. G., J. Chem. Eng. Data **15**, 514 (1970).
[11] Rackett, H. G., J. Chem. Eng. Data **16**, 308 (1971).
[12] Spencer, C. F., and Danner, R. P., J. Chem. Eng. Data **17**, 236 (1972).
[13] Spencer, C. F., and Danner, R. P., J. Chem. Eng. Data **18**, 230 (1973).
[14] Yamada, T., and Gunn, R. D., J. Chem. Eng. Data **18**, 234 (1973).
[15] Francis, A. W., Ind. Eng. Chem., Prod. Res. Develop. **49**, 1779 (1957).
[16] Kreglewski, A., J. Phys. Chem. **73**, 608 (1969).
[17] Kreglewski, A., Internal Technical Report of the American Petroleum Institute Project 44 and the Thermodynamics Research Center Data Project, Texas A & M University, College Station, Texas, 1974.
[18] Pitzer, K. S., and Curl, R. F., Jr., J. Amer. Chem. Soc. **79**, 2369 (1957).
[19] O'Connell, J. P., and Prausnitz, J. M., Ind. Eng. Chem., Process. Des. Develop. **6**, 245 (1967).
[20] Tsonopoulos, C., AIChE J. **20**, 263 (1974).
[21] Tsonopoulos, C., AIChE J. **21**, 827 (1975).
[22] Nothnagel, K.-H., Abrams, D. S., and Prausnitz, J. M., Ind. Eng. Chem., Process Des. Develop. **12**, 25 (1973).
[23] Redlich, O., and Kwong, N. S., Chem. Rev. **44**, 233 (1949).
[24] Chueh, P. L., and Prausnitz, J. M., Ind. Eng. Chem. Fundam. **6**, 492 (1967).
[25] Haman, S. E. M., Chung, W. K., Elshayal, I. M., and Lu, B. C. Y., Ind. Eng. Chem., Process Des. Develop. **16**, 51 (1977).
[26] Peng, D. Y., and Robinson, D. B., Ind. Eng. Chem., Fundam. **15**, 59 (1976).
[27] Benedict, M., Webb, G. B., and Rubin, L. C., Chem. Eng. Progr. **47**, 419, 449, 571, 609 (1951).
[28] Sood, S. K., and Haselden, G. G., AIChE J. **16**, 891 (1970).
[29] Cox, K. W., Bono, J.L., Kwok, Y. C., and Starling, K. E., Ind. Eng. Chem., Fundam. **10**, 245 (1971).
[30] Starling, K. E., "Fluid Thermodynamic Properties for Light Petroleum System," Gulf Publishing Company, Houston, 1973.
[31] Mason, E. A., and Spurling, T. H., "The Virial Equation of State," p. 260, The International Encyclopedia of Physical Chemistry and Chemical Physics, Vol. 2, Pergamon Press, London, 1969.
[32] Hirata, M., and Hakuta, T., Mem. Fac. Technol., Tokyo Metrop. Univ., **18**, 1595 (1968).
[33] Laurance, D. R., and Swift, G. W., J. Chem. Eng. Data **17**, 333 (1972).

9. Appendix. Recommendations for Future Publications of Mixture Data

The overall quality of published vapor-liquid equilibrium, heat of mixing, and volume change of mixing data—and the ease with which it can be located and used—will be greatly improved if the following suggestions are followed for future publications of such data.

The experimental values should always be published. There are many reasons why the publication of only a smoothing equation is not an acceptable substitute for the actual data.

The accuracy of a new experimental apparatus should always be verified with a well-established test system before any data for new systems are measured and published.

The pure compound vapor pressure values should always be measured in the device used for measurement of mixture vapor-liquid equilibrium values, and those pure compound values published as the "endpoints" of the mixture data. Similarly, the pure compound values should be measured and reported along with new mixture density or volume data. The accuracy of the pure compound values should be verified whenever possible by comparison to well-established values from the literature.

Because of the extreme importance of the temperature variable in vapor-liquid equilibrium data, the temperature scale used (e.g., IPTS-68) should be reported.

All vapor-liquid equilibrium data in which all four variables (P, T, x, and y) are measured experimentally should be subjected to a Gibbs-Duhem consistency test before being submitted for publication. If vapor-liquid equilibrium data are measured at multiple temperatures for a given system, and if reliable heat of mixing data are also available for the system, the mutual consistency of the two kinds of data should be checked with the Gibbs-Helmholtz consistency check.

Isothermal vapor-liquid equilibrium data are much preferable to isobaric data both for the evaluator and for the designer who must ultimately use the data.

Experimental uncertainties should be kept at the levels permitted by modern instrumentation, and should be realistically evaluated. The variation in temperature can be kept well below $\pm$ 0.005 K in a well-designed liquid bath but the uncertainty in the stated temperature will still be at least 0.02 or 0.03 K when all the possible errors starting from the primary standard are considered. The uncertainty in the reported pressure must include the effect of the temperature variation, the effect of the composition uncertainty, the effect of incomplete degassing or of chemical reaction, in addition to the resolution uncertainties of the pressure measuring device itself. The device used to measure pressure must be described; if it is a mercury manometer, the temperature variation in its environment is important. Ideally, the uncertainty in reported mole fractions should be $\pm$ 0.0005 but that is extremely difficult to obtain; if the mole fractions are obtained by phase analysis, it is difficult to stay below $\pm$ 0.005.

It is important that authors clearly define the components in a mixture. Two problems often arise here. First, a trivial chemical name should never be used without relating it to the index name used by the Chemical Abstracts Service or derived from the rules of International Union of Pure and Applied Chemistry. Second, the numbering of the compounds must be clear. The format

Pentane(1) + 2-Methyl-3-ethylpentane(2)

is preferable to

(x)Pentane + (1–x)2-Methyl-3-ethylpentane

Inclusion of the component number in parentheses after the component name clearly defines the subscripts on the mole fraction symbols wherever they appear, does not interfere with the nomenclature for the chemical name,and is not restricted to binary systems.

Finally, editors are urged to use literature citation formats which provide no chance for error regardless of whether the citation is handwritten, typewritten or printed by a computer. The use of bold-faced characters to identify one of the location numbers may cause trouble; the "boldface" often disappears when the document is copied, and it cannot be reproduced on a typewriter or on a computer. The use of order alone to distinguish between the location numbers is also unsatisfactory; it is often impossible to deduce from a single document what the prescribed order for that particular journal is supposed to be. The use of colons or semi-colons to alleviate these shortcomings is often not helpful, and represents a misuse of those punctuation tools.