Published in final edited form as:

J Chem Thermodyn. 2019 ; 133: 208–222. doi:10.1016/j.jct.2019.01.029.

Validation of thermophysical data for scientific and engineering applications^{1,2}

Vladimir Diky³, Ala Bazyleva, Eugene Paulechka, Joseph W. Magee, Vikina Martinez, Demian Riccardi, Kenneth Kroenlein

Applied Chemicals and Materials Division, National Institute of Standards and Technology, Boulder, Colorado 80305-3337, USA

Abstract

High quality thermophysical property data are essential to many scientific and engineering applications. These data are produced at a high rate and are affected by a range of experimental and reporting error sources that often exceed stated uncertainties. As a result, critical evaluation is required to establish the limits of reliability in a quantified way. The present work describes reporting recommendations and property data validation methods developed and applied at the Thermodynamics Research Center at NIST through the use of the ThermoData Engine (TDE; SRD 103a/b) software. Examples are provided with an emphasis on various consistency checks, which may include the use of equations of state (EOS).

Keywords

Good reporting practices; Thermophysical data; Uncertainties; Models; Equations of state (EOS)

1. Introduction.

The metrology of thermophysical and thermochemical properties is a branch of science with a long history, and the data are long lived. The oldest scientific paper stored in the SOURCE database [1] reported, in 1779 [2], the boiling temperatures of water at pressures in the range (90 to 104) kPa. The range of chemical systems of interest to industrial and scientific pursuit has grown dramatically since then, and the field has evolved to keep pace with this broadening of relevant temperatures, pressures, and chemical functionalities.

Thermophysical property data are required by researchers and engineers for modern applications. Without high-quality data, theoretical methods cannot be validated, physical trends and correlations remain hidden, and property prediction methods have no proven efficacy. At this time, the majority of thermodynamic and transport properties cannot be predicted with an accuracy that competes with properly conducted experiments. As a result, the amount of measured data published per year, along with the number of people and

¹This contribution of the National Institute of Standards and Technology is not subject to copyright in the United States

²Trade names are used for information purpose and do not imply endorsement by the National Institute of Standards and Technology. Similar products by other manufacturers may work as well or better.

³Corresponding author. vladimir.diky@nist.gov.

laboratories involved, continues to increase [3]. Any measurement is affected by random and systematic errors that are reflected in the uncertainty of the measured property. Thus, any datum not accompanied by a reasonable estimate of uncertainty is misleading and may be harmful.

The importance of accurate data for process design is well established, cf. [4,5,6,7,8]. However, errors in the data submitted for publication are ubiquitous [9] and often go unrecognized by the peer review process. The high frequency of erroneous publications has raised concerns that have initiated several studies [10,11,12,13,14]. Even accurate measurements with proper uncertainty estimates can be erroneously reported in publications. The information may be incorrectly transferred or interpreted. Numerous errata and comments to published papers reveal problems and sometimes suggest corrections. Additional validation efforts are thus prudent before relying on any piece of data.

Critical evaluation of thermophysical data is central to the activities at the Thermodynamics Research Center (TRC) at the National Institute of Standards and Technology (NIST). Resultant reliability assessments are constantly performed for data submitted to five cooperating journals before publication [15,16,17,18,19], for each data set added to the SOURCE database [1], and to all stored data during periodic re-evaluation of thermodynamic data for pure compounds, chemical reactions, and mixtures at TRC. The present study describes validation methods used in the TRC data operation and gives recommendations to data users.

2. Data validation methods.

2.1. Content analysis.

Measured properties are useful when published in numerical form along with uncertainty estimates and associated metadata. Metadata is the information that ensures reproducibility of the measurement and must provide identification and purity assessment of studied substances, methods of uncertainty assessment, and other information regarding measurement technique and procedures. Considerations of stability (*e.g.*, thermal, against hydrolysis) along with proof of the equilibrium state should be included. Most importantly, the metadata should be properly supported; for example, proof of identity is needed for optically active compounds, solvates, uncommon compounds or those synthesized by the authors, and other potentially ambiguous substances.

Certain properties are very sensitive to impurities, such as the viscosity of ionic liquids and solubility of salts in organic solvents. Other examples include the enthalpy of formation derived from combustion experiments, which varies with the amount of moisture, or electrical conductivity of molecular compounds, which change with trace amounts of electrolyte impurities. Justifications of purity and uncertainty assessments require descriptions of the method or basis and analysis of contributing factors. The effect of the purity on the targeted property should be estimated.

Each property/method combination requires reporting specific details known to experts. A common example in bomb calorimetry is the reduction of sample mass to that under vacuum

conditions. If the buoyancy corrections are not applied, the error in combustion enthalpy can approach 0.1 %, exceeding reasonable uncertainty for C-H-O-N compounds. A failure to provide the density or information regarding the application of buoyancy corrections renders the data ambiguous and impossible to correct without additional information. When sufficient metadata are provided, TRC experts are sometimes able to correct reported data within the SOURCE database [1].

Considering the above examples, universal recommendations are not possible. Fortunately, guidelines and supporting tools dedicated to several experimental methods are currently under preparation, some of them with participation of TRC staff. Several IUPAC recommendations exist, which can be used for verification of measurement equipment [20], and recommendations for phase equilibrium studies are currently the subject of an IUPAC project [21]. Specific reporting recommendations also exist (*e.g.*, [22]) but, unfortunately, are not consistently enforced. A discussion of more general “Good Reporting Practice” principles has been initiated by TRC that should result in publication of recommendations soon. An example of the failure of good reporting is that of derived or smoothed data (for example, excess volumes derived from measured densities) without reporting directly measured values. Fundamental to measuring reliable property data is the design of a research project. For example, covering a range of conditions (*e.g.*, density as function of temperature) generally provides more reliable results than a single measurement; the same is true when endpoints are measured for mixtures (*e.g.*, binary endpoints for ternary mixtures and, in turn, pure-compound properties for binary mixture data).

Failures of rigorous reporting or experimental design typically indicate lower quality of the reported data, *i.e.*, increased uncertainties in the reported values. Clearly, the extent of possible error is difficult to assess. Analysis of experimental papers requires expertise in appropriate areas and is laborious. Rigorous vetting is performed at TRC for manuscripts submitted to the cooperating journals [15–19]. One of the goals of that cooperation is to assure rigorous experimental reporting. If the report is not properly rigorous, higher uncertainties should be expected and are assigned in the SOURCE database as in the following examples:

- When combustion data fail to specify mass buoyancy corrections, as described above, a 0.1 % uncertainty is added to whatever uncertainty has been assessed at that point.
- When vapor pressures are reported only in a smoothed form, their weight in fitting is reduced to the extent corresponding to 5 % uncertainty.
- When endpoint data are not included for a given mixture, the recorded uncertainty in excess volume increased by 0.5 % with a minimum of 0.04 cm³·mol⁻¹.

Furthermore, even properly reported experiments may contain errors. Errata and comments sometimes amend scientific publications. Thus, even a rigorously reported experiment must be evaluated critically. Errors propagate; data users must carry out continuous searches to ensure the data that is central to their research or applications, especially those at industrial

scale. For the SOURCE database, such searches are performed every year to ensure that errata have been incorporated and properly cited.

Scatter assessment is the first order evaluation of numeric values. Scatter indicates the level of random contributions to the uncertainty and is straightforward for single values measured in repeated experiments [23,24]. According to statistical theory, the random contributions to uncertainty can be reduced to zero by analyzing an infinite number of repeated experiments [23], but systematic errors are not controllable in this way. Analysis of scatter often reveals systematic discrepancies between literature references resulting from common reasons such as substance instability, difficulties to achieve the equilibrium state, or the details of the experimental device and protocol. However, scatter cannot be completely separated from modeling in all other cases as discussed in Section 2.5. At TRC, uncertainties determined from scatter analysis are compared to the reported uncertainties, and the larger of the two is recorded in the SOURCE database.

2.2. Literature comparisons.

Literature comparisons require knowledge of sufficient available data in order to have an unbiased representation. Accumulating sufficient available data requires a comprehensive literature search, which is not trivial; completeness is certainly not guaranteed. Comprehensive chemical search systems, such as Chemical Abstracts (SciFinder) [25] and Reaxys [26], may not contain all information about the content of the publications. General-purpose scientific databases such as Web of Science [27] may be also useful, but they are hampered by non-scientific, human language-based search syntax. Internet search engines sometimes give additional information. Older publications that may contain rigorous reports of the most accurate measurements available are especially problematic. The most efficient way to find older data is usually to search in reference books and reviews, but it is extremely labor intensive because of the absence of electronic indexes in most cases; and finding a comprehensive list of reference books is also a non-trivial search task. The often-overlapping content of reference books increases the work needed to execute a comprehensive data search.

Compared to database queries, those built in natural human language using search engines frequently return a huge number of false positives and may miss cases where the manner of expression of the content deviates from the query. Specialized databases, such as the TRC SOURCE database (available through NIST Standard Reference Databases as 103a and 103b) [28], DIPPR [29], or Dortmund Data Bank [30], can be superior to reference books in many areas, but the extent of their coverage for any particular system/property combination is usually unknown. If the data are collected and significant discrepancies are discovered, verification of the citation and analysis of the original reports are needed.

If different reports with no obvious flaws provide inconsistent values, the uncertainty of the knowledge is at least as high as the magnitude of the deviation, unless additional validation methods can refine the judgment. An example of large discrepancies is given by the viscosity of bis(2-ethylhexyl) hydrogen phosphate (Fig. 1). The practice at TRC is to increase the uncertainties to the order of their discrepancy with the competing literature. If

discrepancies in competing sources exceed the factor of 3 times the default uncertainty for that property, the data undergo an expert review that may have one of three outcomes:

1. If obvious errors are found in one source, it is excluded from evaluation.
2. If the level of the discrepancy still leaves the data usable, they stay with revised uncertainties; otherwise, all data on that property are excluded from use until additional information becomes available.
3. When two or more sources are deemed consistent, scatter analysis applied to all consistent data provides the means of assigning uncertainty for those data sets.

Subtle problems accompany legitimate attempts to review the existing data. While disagreement between different sources indicates a problem, agreement between sources can only increase the probability of the correct result; agreement between sources is no guarantee of correctness. For example, the published D-glucose solubility in water is consistent between publications [31] and [32]. However, a broader literature review reveals that these two sources are inconsistent with several others (Fig. 2), the most important being [33,34,35]. The solubilities in [31] were measured at elevated pressures (1.5 MPa to 12.0 MPa), but extrapolation to atmospheric pressure would cause only a minor change (less than 10 % of the solubility values) – much smaller than the extent of the discrepancy. Consistency with [36] is also shown in [31] on the diagram, but is apparently a result of misinterpretation of [36] by the authors of [31]. Such an erroneous claim of consistency with the previous literature is not unique; see, for example, [37] for solubility of sodium chloride in water. The solubilities published in [32] include temperatures higher than the melting point of D-glucose 416 K [38], which is inconsistent with the method used; furthermore, the manner of representation (plots only) precludes accurate transfer of information. In addition, the reference book [39] cited in [32] misinterpreted the original data and provided an erroneous basis for the consistency claimed in [32]. Thus, the seeming data consistency is misleading without a comprehensive treatment. To facilitate comprehensive literature reviews, TRC has made the ThermoLit resource [40] available. ThermoLit does not guarantee completeness, but it will likely accelerate the literature search. Consultation and comparison with this resource should be considered as a minimal expectation for any experimental reports of thermophysical properties in the modern literature.

2.3. Consistency with endpoints.

For mixtures, an important consideration is the comparison to endpoints, such as pure-compound endpoints for binary mixtures and binary endpoints for multi-component mixtures. As mentioned above, including endpoints significantly increases the reliability of a reported measurement. A comprehensive literature search, as described above, is required to collect endpoint data. If the coverage of the composition space with data is dense, comparison to endpoints is a simple interpolation task. If the composition gaps are wide, the comparison requires modeling, either explicit or implicit (extrapolation by human eye). A few examples illustrate the importance of endpoint comparison and different situations that may be encountered:

- Comparisons to binary liquid-liquid equilibrium (LLE) data from the literature reveals switched components in Table 2 of publication [41] that reports ternary

LLE data (Fig. 3a). That publication error has been corrected after our observation [42] (Fig. 3b).

- Solubility of D-galactose in a binary ethanol-water solvent is inconsistent between publication [43] and that reported in [44] and the binary endpoints reported in [45,46] (Fig. 4). The solubility in publication [44] was validated by comparison with solubility of isomeric carbohydrates. While the way the measurements were reported (plots) in publication [43] causes loss of precision, the extent of inconsistency is much larger (order of magnitude). Due to the many discrepancies, the data set from [43] is rejected from any use in the SOURCE database.
- Solubility of lithium bromide in acetone-water mixtures was reported by the same authors in two publications [47,48] without endpoints. Combining them with literature endpoints in acetone [49] and water (Supplement 1) showed imperfect but reasonable agreement (Fig. 5), and the data from [47,48] were retained in the database. However, users should be aware of possible solubility data error in the vicinity of 0.05 mole fraction of acetone. Further refinement of the reliability of those data requires modeling more sophisticated than interpolation. Users are advised to review consistency with endpoints and make their own judgment on this mixture.

Endpoint consistency checks provide a straightforward way to relate and validate mixture data across publications. Thus, good reporting of mixture measurements and the responsible consumption of such data should include these considerations.

2.4. Property prediction.

Property prediction methods vary by the extent of their physical basis and reliance on experimental data used for parameterization. While property prediction cannot typically compete with experimental measurements, they can be used to augment analysis through expedient identification of significant errors. Empirical correlations between different properties of same compounds can also be considered as prediction methods when one property is known, and the other is derived. If experimental data are available for both properties, such correlations can be used for mutual validation. The accuracy of the property predictions may be as good or better than experimental measurements for some properties, such as virial coefficients and heat capacities of simple diluted gases, cf. [50,51], and gas-phase enthalpies of formation, cf. [52]. A variety of property prediction methods are described in [53] and implemented in the ThermoData Engine (TDE) software [28].

More complex offline methods, such as [52], are used for background data validation scans at TRC. Over 60 reaction enthalpies have been rejected in that process to date. An example of resolved ambiguity is 2-oxazolidone [54,55,56,57], for which the reported values of combustion enthalpy in the crystalline state range from $(-1487 \text{ to } -1460) \text{ kJ}\cdot\text{mol}^{-1}$. The value derived from the predicted gas-phase formation enthalpy $(-346.5 \pm 2.7) \text{ kJ}\cdot\text{mol}^{-1}$ [52], the sublimation enthalpy $(86.0 \pm 3.0) \text{ kJ}\cdot\text{mol}^{-1}$ [56,57], and the formation enthalpies of CO₂ and water [58] is $(-1462.6 \pm 4.0) \text{ kJ}\cdot\text{mol}^{-1}$ (see Fig. 6). The two values with the largest

deviations from prediction have been rejected from use and the uncertainty has been assigned according to the scatter in the remaining experimentally measured values.

The uncertainty of prediction methods should be adequately assessed in order to be correctly applied to data validation. The level of uncertainty may be correlated with molecular structure or estimated by covariance-based [59] or Monte Carlo/Bayesian methods [60]. Assessment of uncertainties is an essential element of prediction method development that is frequently ignored; thus, additional research is required.

As mentioned in the previous section, endpoint consistency checks may require additional modeling. An example for a ternary mixture is given in Fig. 7. The LLE data for acetonitrile + benzene + water were correctly published [61], but incorrectly captured to a database (swapped components). The mistake was revealed by a predictive NIST-modified UNIFAC [62] model due to different tie lines (lines connecting the compositions of the coexisting liquids). Remarkably, a similar result is obtained if an activity-coefficient model fitted to binary mixture data is applied to the ternary mixture. This example demonstrates how endpoint consistency can be validated by modeling, which will be in more detail in the next section.

2.5. Modeling.

Models of property data (*e.g.*, mathematical equations) are useful for data validation and practical applications, such as calculating derivatives or property values at conditions other than those having been treated experimentally; modeling can aid endpoint consistency assessment discussed in section 2.3. A consistent model should be physically valid and pass multiple validity rules, whether empirical or theoretical. Such rules define ranges of valid values and shapes of temperature, pressure, and composition dependences. Deviation from models should be considered with respect to scatter (random deviations) or bias (deviations correlated with state variables). Inadequate models or erroneous data may cause bias. Distinguishing scatter and bias is strictly not possible, and determining whether the model or the data are inaccurate is often challenging. While naïve inspection, non-zero average deviation, or average percent deviation may reveal bias in simple cases, more complex correlation analysis is often needed. Known biases of certain models should be taken into consideration. Examples include the representation of density by cubic equations of state with fixed critical parameters and the pre-defined behavior of most activity-coefficient models for mixtures. Comparing alternative models can often help identify the least biased ones. Models with a variable number of parameters, such as the Redlich-Kister equation used for representation of the properties of ethanol + water [63], are flexible and free from pre-defined behavior. However, their flexibility can be easily abused to create models with non-physical fidelity to targets, termed “overfit.”

While we are unaware of strict criteria for identifying when data are overfit, heuristics can be developed to avoid this issue. For example, when the number of parameters becomes larger than that justified by the available data, statistical significance of the parameters decreases due to increased variance/covariance. At a minimum, an F-test applied to a prospective added parameter can indicate whether the parameter contributes significantly to the deviation reduction. If added parameters reduce the deviations below the uncertainty in

the measurements, they should be disregarded. Orthogonal basis functions can facilitate the correlation analysis. Models subject to overfit may acquire extra inflection points in the function or its derivatives that can be revealed visually or mathematically. Unfortunately, the use of separate training and validation data sets is typically not possible because the amount of the available thermophysical property data is insufficient.

Modeling results can be analyzed algorithmically or visually. Numerical criteria are used in TDE to determine experimental sufficiency, deemed “adequacy,” and bias (absolute or relative). Experimental adequacy is estimated by Eq. 22 from publication [64]: $A = (1/n)\sum(d_i/U_i)^2$ where n is the number of data points, d_i is the curve deviation of the i -th data point, and U_i is the expanded combined uncertainty ($k = 2$) associated with it; bias is estimated from signed weighted deviations. More sophisticated criteria involving correlation of deviations are under development. Visual analysis requires an efficient graphical representation. A solution revealing the behavior of the model and data close to endpoints is proposed in [65]. For binary mixtures, it consists of the transformation of the plotted composition variable x (mole or mass fraction) to $\log(x/(1-x))$.

The inverse approach is the evaluation of published models using validated data. Models can be developed and evaluated against validated data using TDE, as described in Section 3. However, consistency with raw data is not the only available validation method for models; additional methods will be described in section 2.6.

2.6. Thermodynamic consistency.

Since thermodynamics fundamentally relates different properties, the validation of each property in isolation is not the only option. Consistency tests are well-known for vapor-liquid equilibrium data (VLE) for mixtures [cf. 66 and references therein]. Similar tests have been proposed for solid-liquid equilibria (SLE) and liquid-liquid equilibria (LLE) [67]. Combinations of data sets for different thermodynamic properties that fail consistency tests may be ignored (traditional approach) or receive a lower quality factor [66] and lower weights in model fitting. Some of the methods from this category require integration using modeling unless the data density is sufficient. Other methods apply common models to different properties. Some consistency tests are limited by various approximations intended to reduce the amount of data necessary for performing the test (*e.g.*, excluding excess volume or excess enthalpy from consideration). VLE and other consistency tests have been widely discussed in the literature, and that discussion is still active [68,69,70].

2.6.1. Equations of state.—Assessment of consistency through modeling for pure compounds as well as the use of equations of state (EOS) will benefit from continued research efforts. A clear obstacle is that different pure-compound properties have different applications and are measured by different methods and in different laboratories. Wide adoption of non-analytical EOS (in terms of [53]) has started fostered computational progress and advances in EOS research. Over 100 high-accuracy EOS for pure compounds are collected in NIST REFPROP software [71] and are also available through TDE. New measurements reported for compounds with REFPROP EOS have the potential to provide mutual validation of the new data and the EOS. Accumulating additional existing knowledge

of relevant properties is almost always helpful. While fitting an accurate EOS for a new compound is non-trivial, developing an EOS with moderate accuracy for mutual validation of available data is frequently possible. Important exceptions are associating compounds, such as alcohols or acids, and high-boiling substances. Fitting a moderate-accuracy EOS is possible using TDE software. The 12-parameter Span-Wagner equations [72,73] with TDE initialization or an existing equation for a substance with similar critical compressibility factor and acentric factor can provide the initial approximation.

Benzonitrile provides an example of the successful application of this method. Inconsistent values are reported for its vaporization enthalpy, speed of sound, heat capacity, and liquid density. Conclusions can be drawn for the heat capacity because one data set deviates from multiple others, and for low-pressure density and vaporization enthalpy on the basis of consistency with high-pressure density and vapor pressure, respectively. However, the most complete conclusions are only possible after fitting an EOS. Fitting the polar Span-Wagner equation [73] to benzonitrile data has been carried out with TDE software; the parameters in the form readable by REFPROP and TDE software are given in Supplement 2 (an explanation of the parameters can be found in the REFPROP manual [74]). The equation represents the available data with accuracy close to experimental measurements and allows discarding a few inconsistent and apparently erroneous data sets (Fig. 8). Dimethyl sulfide provides another interesting case. The existing REFPROP EOS for 1,3-butadiene and dimethyl ether readily scale to dimethyl sulfide by substituting scaling temperature and density (close to the critical values). After some refinement, there is a reasonable consistency between the EOS and the data stored in the SOURCE database (Fig. 9). The EOS parameters are given in Supplement 3.

The success of this method depends on the complexity of intermolecular interactions, the amount of available data, and the extent of their inconsistency. Ethylene carbonate provides an example of an unresolved inconsistency. A recently published low-temperature vapor pressure [75] disagrees with most other reports (Fig. 10). The authors of [75] believe all previous publications were erroneous due to substance instability or other reasons. Our attempts to resolve the ambiguity by fitting an EOS were unsuccessful. Thus, additional experimental data are needed. The power of using equations of state analysis is that different properties can be involved in the analysis. For example, measurements of density and speed of sound would also contribute to refining the vapor pressure.

Use of EOS for resolving inconsistencies for mixtures is more challenging. EOS for mixtures of unlike components are complex, and the existing formulations can accurately represent all properties only for moderately non-ideal mixtures. However, that limited application should not be neglected, and it is used in appropriate cases at TRC. An example showing the consistency of VLE and volumetric data for propane + 2,3,3,3-tetrafluoro-1-propene is given in Fig. 11. The model is XR0 (an abbreviation of the Kunz-Wagner model KW0 [74] with a zero 5th parameter), and the parameters are 1.01207, 0.93687, 1.004557, and 1.013734.

2.6.2. Properties represented by separate equations.—Equations of state describe a single solid phase or fluid phases (gas and liquid). Consistency between different

phases (other than gas and liquid) cannot be verified with a single EOS. However, consistency requirements can be checked and enforced for models associated with different phases. Consistency conditions supported by TDE are described in [76]. The principles of consistent evaluation are similar to those for EOS fitting and based on an objective function that combines the effect of deviations from data and violation of thermodynamic constraints. A consistency monitor has been developed at TRC for visualization of pure-compound consistency checks and refining property equations. The monitor allows representing fluid-state properties with EOS. For example, it can be shown that accurately measured properties for benzene can be represented by consistent equations for all phases (Fig. 12). The EOS representing fluid properties with experimental accuracy is presented in publication [77]; solid-state heat capacity is represented by a spline [78], and sublimation thermodynamics by PVExpansion (sublimation pressure) and polynomial (sublimation enthalpy) equations [79] in the range 180–278.67 K with the parameters 11.0046, –5883.84, 2.89002, and –0.016494 (PV Expansion) and 49.0562, 0.022773, and –0.000134259 (polynomial). Representation of each property by a separate equation is the method used by TDE software. The consistency conditions involving both fluid and solid phases are as follows:

- Consistency between vapor and sublimation pressure at the triple point
- Consistency between enthalpies of fusion, vaporization, and sublimation at the triple point
- Consistency between sublimation pressure and enthalpy (involving densities of both phases)
- Consistency between sublimation pressure and heat capacity difference (gas-solid) [76].

When different properties are represented by independent equations fitted to those properties, the monitor may reveal inconsistencies. In the absence of information that would allow elimination of invalid data, the extent of possible error (uncertainty) is as large as the difference between existing data and property values derived from other properties using consistency equations. For example, vapor and sublimation pressures of vanillin that have been measured by several researchers (See Supplement 1) do not converge at the triple point. In order to achieve property consistency, a balanced departure of the set of property equations from the raw data has been carried out using the TDE consistency monitor (Fig. 13). The resulting equation parameters are 7.92333, –9.20415, 2.81982, –5.34241, and –4.36161 (indexed as per the TDE documentation) for vapor pressure in the range 354.6 K to 782 K (Wagner equation [76]) and 50.04, –12821.8, –3.0 for sublimation pressure in the range 254.6 K to 354.6 K (PVExpansion [76]). The property equations obtained by TDE consistency enforcement agree well with a recent publication [80] that provides more accurate experimental values of vapor and sublimation pressures, as well as enthalpies of vaporization and sublimation (the latter is $99.6 \text{ kJ}\cdot\text{mol}^{-1}$ at $T = 298.15 \text{ K}$). An additional validation has been provided by consistency of the sublimation enthalpy at $T = 298.15 \text{ K}$, (99.2 ± 17.3) $\text{kJ}\cdot\text{mol}^{-1}$ obtained in the TDE procedure with experimental enthalpy of formation in the crystalline state (-476.3 ± 1.0) $\text{kJ}\cdot\text{mol}^{-1}$ [81,82] and ideal-gas enthalpy of formation (-377.7 ± 3.0) $\text{kJ}\cdot\text{mol}^{-1}$ from high-accuracy quantum chemical calculations [52]. Consistency enforcement provides more accurate values than the original inconsistent

reports. However, uncertainties should continue to be large until additional experimental data became available. The vapor pressure value at the triple point obtained by the above procedure deviates by 20% from the one derived from the new experimental data [80]. The 20 % deviation is a significant improvement compared to the initial inconsistency that is equivalent to 170% in logarithmic scale ($100 \cdot \ln(p_1/p_2)$). The consistency enforcement also reconciles fusion, vaporization, and sublimation enthalpies at the triple point.

2.7. Trends in a series of compounds.

A less rigorous but still powerful check can be based on variation of property values in families of compounds with structural differences that are systematic, such as the structural variation with adding CH₂ or functional groups. Less common structural differences can also be harnessed. Constructing sets of substances can be automated [83]. Variation of property values in such sets should be generally smooth, excluding the smallest molecules in the set, odd-even effects, and some property-specific effects (such as formation of plastic crystals and micellization). For example, correlations of bis(2-ethylhexyl) hydrogen phosphate with tris(2-ethylhexyl) phosphate [84,85,86] and phosphoric acid [87] suggest that lower measured values of viscosity may be correct (Fig. 14). However, additional experimental verification with particular attention to reliable identification, purity assessment, and other factors affecting viscosity (chemical stability, moisture) is needed.

In the absence of additional property data, uncertainties for other similar compounds measured with a particular implementation of a method, equipment, and protocol can provide an estimate of the possible level of uncertainty. However, property value refinement may require additional measurements or high-accuracy theoretical calculations. Generally, resolving ambiguities requires more efforts than the initial measurements. Thus, reporting measurements with low accuracy carry a substantial societal cost. A collection of unresolved issues encountering more than 200 cases exists at TRC (see Supplement 4 for examples). A justification of suggested research for one of them is given in Supplement 5. It is often difficult to find a researcher whose interests, skills, and resources match a particular problem. The probability to find a solution may increase if the issue is widely disseminated, as happened in the case of cubane [88]. Its sublimation enthalpy was re-determined after the publication [88] that raised doubts in its previously published value [89], and the new value was determined to be 1.5 times smaller [90]. Unfortunately, there are currently no expedient channels to communicate unresolved issues to the wider community.

3. Validation of models.

Validation of published models (equations) prior to their use is recommended in all cases to avoid errors and to assess the adequacy of the models. Model development should involve several data checks including consistency of the model with available data, valid property values (falling in the expected range and showing expected dependence on temperature, pressure, and composition), consistency with other properties (described above), and consistency of properties for series of compounds. The TDE consistency monitor supports those checks and provides visualization, including absolute and relative deviations. Deviation from the data provides an estimate of possible model uncertainty, and deviation of

measured property values from the values derived from other properties through thermodynamic relations can be used as a way to assess the possible effects of property inconsistencies. If different properties can be calculated by the same model, more advanced methods such as covariance matrices or Monte Carlo can directly account for inconsistencies. All available data (validated with the above described methods) should be involved in a comparison in order to make unbiased conclusions. As mentioned above, finding and curating data is a non-trivial task. TRC provides property data for model validation.

When the accuracy of models representing thermophysical properties is assessed, sensitivity of the target application (such as process simulation) to the variation of those data (represented by numerical values or equations) should be explored within that application, as is shown, for example, in publication [91].

4. Conclusions.

Experimentally measured properties are frequently affected by errors exceeding claimed or expected uncertainties. The event of publication finishes a research, but it starts another stage in the life of the published data, their validation, that needs additional information and lasts as long as it takes to acquire the sufficient validation basis. Good reporting practices minimize the difficulties in validating information for use in research and engineering applications. Robust application and model development require rigorous validation of property data as described in the present work. The current publication presents several approaches to thermophysical data validation, all of which are employed at the Thermodynamic Research Center at NIST and made available through the ThermoData Engine software. Those methods are applied continuously as additional information becomes available in order to refine our knowledge of thermophysical data. Application of the methods described in the present study allows avoiding erroneous data and reveals situations when additional research is needed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- [1]. Frenkel M, Dong Q, Wilhoit RC, Hall KR, TRC SOURCE Database: A Unique Tool for Automatic Production of Data Compilations, *Int. J. Thermophys* 22 (2001) 215–226. 10.1023/A:1006720022161
- [2]. Shuckburgh G, On the Variations of the Temperature of Boiling Water, *Philos. Trans. R. Soc. London* 69 (1779) 362–375.
- [3]. Frenkel M, Global Information Systems in Science: Application to the Field of Thermodynamics, *J. Chem. Eng. Data* 54 (2009) 2411–2428. 10.1021/jc800877f
- [4]. Mathias PM, The role of experimental data in chemical process technology, *Pure Appl. Chem* 81 (2009) 1727–1743. 10.1351/PAC-CON-08-10-06
- [5]. Watanasiri S, Development of on-demand critically evaluated thermophysical properties data in process simulation, *Pure Appl. Chem* 83 (2011) 1255–1281. 10.1351/PAC-CON-10-11-18

- [6]. Hendriks E, Kontogeorgis GM, Dohrn R, de Hemptinne J-C, Economou IG, Fele Zilnik L, Vesovic V, Industrial Requirements for Thermodynamics and Transport Properties, *Ind. Eng. Chem. Res* 49 (2010) 11131–11141. 10.1021/ie101231b
- [7]. Franca JMP, Nieto de Castro CA, Lopes MM, Nunes VMB, Influence of Thermophysical Properties of Ionic Liquids in Chemical Process Design, *J. Chem. Eng. Data* 54 (2009) 2569–2575. 10.1021/jc900107t
- [8]. Mathias PM, Soto A, Fele-Zilnik L, de Hemptinne J-C, Bazyleva A, Abildskov J, Data Quality and Assessment, Validation Methods and Error Propagation through the Simulation Software: Report from the Round-Table Discussion at the 10th World Congress of Chemical Engineering in Barcelona (October 1–5, 2017), *Chem. Eng. Res. Des* 137 (2018) A1–A8. 10.1016/j.cherd.2018.08.010
- [9]. Chirico RD, Frenkel M, Magee JW, Diky V, Muzny CD, Kazakov AF, Kroenlein K, Abdulagatov I, Hardin GR, Acree WE Jr., Brenneke JF, Brown PL, Cummings PT, de Loos TW, Friend DG, Goodwin ARH, Hansen LD, Haynes WM, Koga N, Mandelis A, Marsh KN, Mathias PM, McCabe C, O'Connell JP, Padua A, Rives V, Schick C, Trusler JPM, Vyazovkin S, Weir RD, Wu J, Improvement of Quality in Publication of Experimental Thermophysical Property Data: Challenges, Assessment Tools, Global Implementation, and Online Support, *J. Chem. Eng. Data* 58 (2013) 2699–2716. 10.1021/jc400569s
- [10]. Ioannidis JPA, Why Most Published Research Findings Are False, *PLoS Med* 2(8) (2005) e124 10.1371/journal.pmed.0020124 [PubMed: 16060722]
- [11]. Horton R, Offline: What is medicine's 5 sigma? *Lancet* 385 (2015) 1380 10.1016/S0140-6736(15)60696-1
- [12]. Ioannidis JPA, An Epidemic of False Claims, *Sci. Am* 304(6) (2011) 16 10.1038/scientificamerican0611-16
- [13]. Felaefel M, Salem M, Jaafar R, Jassim G, Edwards H, Rashid-Doubell F, Yousri R, Ali NM, Silverman H, A Cross-Sectional Survey Study to Assess Prevalence and Attitudes Regarding Research Misconduct among Investigators in the Middle East, *J Acad. Ethics* 16 (2018) 71–87. 10.1007/s10805-017-9295-9 [PubMed: 29755305]
- [14]. Belluz J, Plumer B, Resnick B, The 7 biggest problems facing science, according to 270 scientists, *VOX*, 9 7, 2016 <https://www.vox.com/2016/7/14/12016710/science-challenges-research-funding-peer-review-process> (accessed 05 January 2019).
- [15]. Padua A, Joint statement of editors of journals publishing thermophysical property data, *J. Chem. Thermodyn* 41 (2009) 575–576. 10.1016/j.jct.2008.12.023
- [16]. Brenneke JF, Goodwin ARH, Mathias P, Wu J, New Procedures for Articles Reporting Thermophysical Properties, *J. Chem. Eng. Data* 58 (2013) 1069 10.1021/jc4003674
- [17]. Cummings PT, Joint Statement of Editors of Journals Publishing Thermophysical Property Data: Process for Article Submission for The Journal of Chemical Thermodynamics, Fluid Phase Equilibria, International Journal of Thermophysics, Thermochimica Acta, and Journal of Chemical Engineering Data, *Fluid Phase Equilib.* 276 (2009) 165–166. 10.1016/j.fluid.2008.12.005
- [18]. Schick C, Joint statement of Editors of journals publishing thermophysical property data, *Thermochim. Acta* 484 (2009) vii–viii. 10.1016/j.tca.2008.12.024
- [19]. Haynes WM, Friend DG, Mandelis A, Joint Statement of Editors of Journals Publishing Thermophysical Property Data, *Int. J. Thermophys* 30 (2009) 371–373. 10.1007/s10765-009-0571-9
- [20]. Marsh KN (Editor). Recommended Reference Materials for the Realization of Physicochemical Properties (IUPAC Chemical Data). Blackwell, Oxford, 1987 ISBN-10: 063201718X; ISBN-13: 9780632017188.
- [21]. Bazyleva A, Acree WE Jr., Chirico RD, de Loos TW, Diky V, Hefter GT, Jacquemin J, Kang JW, Magalhaes MCF, Magee JW, O'Connell JP, Shaw JM, Trusler JPM, Weir RD, Recommended Reference Materials for Phase Equilibrium Studies, IUPAC project 2011–037-2–100. https://iupac.org/projects/project-details/?project_nr=2011-037-2-100 (accessed 05 January 2019).
- [22]. Chirico RD, de Loos TW, Gmehling J, Goodwin ARH, Gupta S, Haynes WM, Marsh KN, Rives V, Olson JD, Spencer C, Brenneke JF, Trusler JPM, Guidelines for reporting of phase

equilibrium measurements (IUPAC Recommendations 2012), Pure Appl. Chem 84 (2012) 1785–1813. 10.1351/PAC-REC-11-05-02

- [23]. JCGM 100: 2008, Evaluation of measurement data —Guide to the expression of uncertainty in measurement, BIPM, 2008 https://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf (accessed on 05 January 2019)
- [24]. Taylor BN, Kuyatt CE, Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Technical Note 1297, Gaithersburg, MD, 1994 <https://emtoolbox.nist.gov/Publications/NISTTechnicalNote1297s.pdf> (accessed on 15 December 2018).
- [25]. CAS SciFinder (Chemical Abstracts), American Chemical Society <https://scifinder.cas.org> (accessed on 05 January 2019).
- [26]. Reaxys Database, Elsevier <https://www.reaxys.com/> (accessed on 05 January 2019).
- [27]. Web of Science, Clarivate Analytics <https://webofknowledge.com/> (accessed on 05 January 2019).
- [28]. Diky V, Chirico RD, Frenkel M, Bazyleva A, Magee JW, Paulechka E, Kazakov A, Lemmon EW, Muzny CD, Smolyanitsky AY, Townsend S, Kroenlein K, NIST ThermoData Engine, NIST Standard Reference Database 103a/103b, version 10.2, NIST, Standard Reference Data Program, Gaithersburg, MD, 2017 <https://www.nist.gov/mml/acmd/trc/thermodata-engine/srd-nist-tde-103b> (accessed on 15 December 2018).
- [29]. DIPPR 801 Database, AIChE. <https://www.aiche.org/dippr/events-products/801-database> (accessed on 05 January 2019).
- [30]. Dortmund Data Bank (DDB), DDBST GmbH. <http://www.ddbst.com/ddb.html> (accessed on 05 January 2019).
- [31]. Saldana MDA, Alvarez VH, Haldar A, Solubility and physical properties of sugars in pressurized water, J. Chem. Thermodyn 55 (2012) 115–123. 10.1016/j.jct.2012.06.016
- [32]. Zhang D, Montanes F, Srinivas K, Fornari T, Ibanez E, King JW, Measurement and Correlation of the Solubility of Carbohydrates in Subcritical Water, Ind. Eng. Chem. Res 49 (2010) 6691–6698. 10.1021/ie901845b
- [33]. Smelik A, Torok S, Vukov K, New phase diagram of the D-glucose water system, Acta Aliment. 19 (1990) 139–155.
- [34]. Young FE, D-glucose-water phase diagram, J. Phys. Chem 61 (1957) 616–619. 10.1021/j150551a023
- [35]. Jackson RF, Silsbee CG, The solubility of dextrose in water, Sci. Pap. Bur. Stand. (U. S.) 17 (1922) 715–724.
- [36]. Alves LA, Almeida e Silva JB, Giulietti M, Solubility of D-Glucose in Water and Ethanol/Water Mixtures, J. Chem. Eng. Data 52 (2007) 2166–2170. 10.1021/je700177n
- [37]. Leusbrock I, Metz SJ, Rexwinkel G, Versteeg GF, Quantitative approaches for the description of solubilities of inorganic compounds in near-critical and supercritical water, J. Supercrit. Fluids 47 (2008) 117–127. 10.1016/j.supflu.2008.08.011
- [38]. Bagheri SR, Shaw JM, Observation of Liquid-Crystal Formation during Melting of D-(+)-Glucose, J. Agric. Food Chem 59 (2011) 12605–12609. 10.1021/jf202682d [PubMed: 22026437]
- [39]. Yalkowsky SH, He Y, Handbook of Aqueous Solubility Data, CRC Press, Boca Raton, FL, 2003.
- [40]. ThermoLit, NIST Standard Reference Database #171. <https://trc.nist.gov/thermolit/> (accessed on 15 December 2018).
- [41]. Wannachod T, Hronec M, Sotak T, Fulajtarova K, Pancharoen U, Arpornwichanop A, Effects of salt on the LLE and tie-line data for furfuryl alcohol – n-butanol – water at T = 298.15 K, J. Mol. Liq 218 (2016) 50–58. 10.1016/j.molliq.2016.02.027
- [42]. Wannachod T, Hronec M, Sotak T, Fulajtarova K, Pancharoen U, Arpornwichanop A. Corrigendum to “Effects of salt on the LLE and tie-line data for furfuryl alcohol – n-butanol-water at T = 298.15 K” [J. Mol. Liq. 218 (2016) 50–58], J. Mol. Liq 272 (2018) 188 10.1016/j.molliq.2018.08.156
- [43]. Gao D-M, Kobayashi T, Adachi S, Solubility of D-Galactose, D-Talose, and D-Tagatose in Aqueous Ethanol at Low Temperature, Food Sci. Technol. Res 21 (2015) 801–803. 10.3136/fstr.21.801

- [44]. Zhang L, Gong X, Wang Y, Qu H, Solubilities of Protocatechuic Aldehyde, Caffeic Acid, D-Galactose, and D-Raffinose Pentahydrate in Ethanol-Water Solutions, *J. Chem. Eng. Data* 57 (2012) 2018–2022. 10.1021/je300323g
- [45]. Jonsdottir SO, Cooke SA, Macedo EA, Modeling and measurements of solid-liquid and vapor-liquid equilibria of polyols and carbohydrates in aqueous solution, *Carbohydr. Res* 337 (2002) 1563–1571. 10.1016/S0008-6215(02)00213-6 [PubMed: 12350326]
- [46]. Gray MC, Converse AO, Wyman CE, Sugar Monomer and Oligomer Solubility, in: Davison BH, Lee JW, Finkelstein M, McMillan JD (Eds.), *Biotechnology for Fuels and Chemicals, Applied Biochemistry and Biotechnology*, Humana Press, Totowa, NJ, 2003, 179–193. 10.1007/978-1-4612-0057-4_14
- [47]. Jiang B, Li M.-y., Wang L.-s., Measurement and correlation of solubility of LiBr in acetone-water mixed solvent, *Huaxue Gongye Yu Gongcheng* 29 (2012) 23–27.
- [48]. Jiang B, Wang L-S, Li M-Y, Measurement and correlation of solubility of LiBr in binary solvents formed by acetone and water using LIFAC model. *Adv. Mater. Res* 455–456 (2012) 911–918. 10.4028/www.scientific.net/AMR.455-456.911
- [49]. Bell WRG, Rowlands CB, Bamford IJ, Thomas WG, Jones WJ, CCLII. – The Equilibrium between acetone and salts, *J. Chem. Soc* (1930) 1927–1931. 10.1039/JR9300001927
- [50]. Van TP, Deiters UK, Calculation of intermolecular potentials for H₂-H₂ and H₂-O₂ dimers ab initio and prediction of second virial coefficients. *Chem. Phys* 457 (2015) 171–179. 10.1016/j.chemphys.2015.05.025
- [51]. Van TP, Deiters UK, Calculation of second virial coefficients using ab initio intermolecular pair potentials for F₂-F₂ and H₂-F₂ dimers, *Chem. Phys* 485–486 (2017) 67–80. 10.1016/j.chemphys.2017.01.003
- [52]. Paulechka E, Kazakov A, Efficient Estimation of Formation Enthalpies for Closed-Shell Organic Compounds with Local Coupled-Cluster Methods, *J. Chem. Theory Comput* 14 (2018), 5920–5932. 10.1021/acs.jctc.8b00593 [PubMed: 30234978]
- [53]. Poling BE, Prausnitz JM, O'Connell JP, *The Properties of Gases and Liquids*, Fifth Edition, McGraw-Hill, 2001.
- [54]. Lebedev BV, Kulagina TG, Smirnova NN, Veridusova VV, Kusan J, Keul H, Hocker H, Thermodynamics of dimethylene urethane, of its ring-opening polymerization, and of poly(dimethylene urethane) between 0 and 335 K, *Macromol. Chem. Phys* 201 (2000) 2469–2474. 10.1002/1521-3935(20001101)201:17<2469::AID-MACP2469>3.0.CO;2-7
- [55]. Veridusova VV, Kulagina TG, Lebedev BV, The dependences of the thermodynamic properties of cyclic urethanes on their composition and structure, *Zh. Fiz. Khim* 75 (2001) 1560–1565.
- [56]. Emel'yanenko VN, Turovtsev VV, Orlov YD, Thermodynamic Properties of Dimethylene Urethane, *Russ. J. Phys. Chem. A* 89 (2015) 1146–1151. 10.1134/S0036024415070110
- [57]. Flores H, Ledo JM, Hernandez-Perez JM, Camarillo EA, Sandoval-Lira J, Amador MP, Thermochemical and theoretical study of 2-oxazolidinone and 3-acetyl-2-oxazolidinone, *J. Chem. Thermodyn* 102 (2016) 386–391. 10.1016/j.jct.2016.07.038
- [58]. Cox JD, Wagman DD, Medvedev VA, *CODATA Key Values for Thermodynamics*, Hemisphere, New York, 1989.
- [59]. NIST/SEMATECH e-Handbook of Statistical Methods, 2013 <https://www.itl.nist.gov/div898/handbook/index.htm> (accessed on 15 December 2018).
- [60]. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB, *Bayesian Data Analysis*, Third Edition, CRC Press, Taylor & Francis Group, Boca Raton, FL, 2013.
- [61]. Di Cave S, Mazzarotta B, Liquid-liquid equilibria for ternary systems formed by acetonitrile, water, and aromatic hydrocarbons, *J. Chem. Eng. Data* 36 (1991) 298–303. 10.1021/je00003a012
- [62]. Kang JW, Diky V, Frenkel M, New modified UNIFAC parameters using critically evaluated phase equilibrium data, *Fluid Phase Equilib.* 388 (2015) 128–141. 10.1016/j.fluid.2014.12.042
- [63]. Olson J, Excess Gibbs Energy of Mixing (GE) and Vapor-Liquid Equilibria (VLE) of Ethanol + Water: A Critical Evaluation and Recommended Values, 20th Symposium on Thermophysical Properties, Boulder, CO, USA, June 24–29, 2018, Paper 3521 https://thermosymposium.nist.gov/pdf/Abstract_3521.pdf

- [64]. Diky V, Muzny CD, Lemmon EW, Chirico RD, Frenkel M, ThermoData Engine (TDE): Software Implementation of the Dynamic Data Evaluation Concept. 2. Equations of State on Demand and Dynamic Updates over the Web, *J. Chem. Inf. Model.* 47 (2007) 1713–1725. 10.1021/ci700071t [PubMed: 17518462]
- [65]. Diky V, An Efficient Way of Visualization of Mutual Solubility Data in the Whole Range of Compositions, *J. Chem. Eng. Data* 62 (2017) 2920–2926. 10.1021/acs.jced.7b00174
- [66]. Kang JW, Diky V, Chirico RD, Magee JW, Muzny CD, Abdulagatov I, Kazakov AF, Frenkel M, Quality Assessment Algorithm for Vapor-Liquid Equilibrium Data, *J. Chem. Eng. Data* 55 (2010) 3631–3640. 10.1021/jc1002169
- [67]. Kang JW, Diky V, Chirico RD, Magee JW, Muzny CD, Kazakov AF, Kroenlein K, Frenkel M, Algorithmic Framework for Quality Assessment of Phase Equilibrium Data, *J. Chem. Eng. Data* 59 (2014) 2283–2293. 10.1021/jc500327k
- [68]. Olson JD, Thermodynamic consistency testing of PTx-data via the Gibbs-Helmholtz equation 2.: Methodology and limitations, *Fluid Phase Equilib.* 418 (2016) 50–56. 10.1016/j.fluid.2015.09.040
- [69]. Carrero-Mantilla JI, Ramirez-Ramirez D.d.J., Suarez-Cifuentes JF, Thermodynamic and statistical consistency of vapor-liquid equilibrium data, *Fluid Phase Equilib.* 412 (2016) 158–167. 10.1016/j.fluid.2015.12.035
- [70]. Wisniak J, Ortega J, Fernandez L, A fresh look at the thermodynamic consistency of vapour-liquid equilibria data, *J. Chem. Thermodyn* 105 (2017) 385–389. 10.1016/j.jct.2016.10.038
- [71]. NIST Reference Fluid Thermodynamic and Transport Properties Database (REFPROP), NIST Standard Reference Database #23. <https://www.nist.gov/srd/refprop> (accessed on 15 December 2018).
- [72]. Span R, Wagner W, Equations of State for Technical Applications. I. Simultaneously Optimized Functional Forms for Nonpolar and Polar Fluids, *Int. J. Thermophys* 24 (2003) 1–39. 10.1023/A:1022390430888
- [73]. Span R, Wagner W, Equations of State for Technical Applications. II. Results for Nonpolar Fluids, *Int. J. Thermophys* 24 (2003) 41–109. 10.1023/A:1022310214958
- [74]. Lemmon EW, Bell IH, Huber ML, McLinden MO, REFPROP Documentation, Release 10.0, <https://trc.nist.gov/refprop/REFPROP.PDF> (accessed on 15 December 2018).
- [75]. Pokorny V, Stejfa V, Fulem M, Cervinka C, Ruzicka K, Vapor Pressures and Thermophysical Properties of Ethylene Carbonate, Propylene Carbonate, γ -Valerolactone, and γ -Butyrolactone, *J. Chem. Eng. Data* 62 (2017) 4174–4186. 10.1021/acs.jced.7b00578
- [76]. Diky V, Chirico RD, Muzny CD, Kazakov AF, Kroenlein K, Magee JW, Abdulagatov I, Frenkel M, ThermoData Engine (TDE): Software Implementation of the Dynamic Data Evaluation Concept. 9. Extensible Thermodynamic Constraints for Pure Compounds and New Model Developments, *J. Chem. Inf. Model* 53 (2013) 3418–3430. 10.1021/ci4005699 [PubMed: 24245860]
- [77]. Thol M, Lemmon EW, Span R, Equation of state for benzene, REFPROP, NIST Standard Reference Database 23. <https://www.nist.gov/srd/refprop> (accessed on 15 December 2018).
- [78]. Kawaji H, Study of molecular dynamics in some crystals using calorimetry, M.Sc. Thesis, Graduate School of Science Osaka University, 1986.
- [79]. NIST ThermoData Engine (TDE103b v10.1) User's Guide. <https://trc.nist.gov/TDE/Help/TDE103b/> (accessed on 15 December 2018).
- [80]. Almeida ARR, Freitas VLS, Campos JIS, Ribeiro da Silva MDMC, Monte MJS, Volatility and thermodynamic stability of vanillin, *J. Chem. Thermodyn* 128 (2019) 45–54. 10.1016/j.jct.2018.07.023
- [81]. Lebedeva ND, Ryadnenko VL, Gutner NM, Kiseleva IN, Formation enthalpies of substituted benzenes, *Termodin. Org. Soedin* 5 (1976) 12.
- [82]. Maksimuk Y, Ponomarev D, Sushkova A, Krouk V, Vasarenko I, Antonava Z, Standard molar enthalpy of formation of vanillin, *J. Therm. Anal. Calorim* 131 (2018) 1721–1733. 10.1007/s10973-017-6651-3

- [83]. Diky V, Kazakov A, Kroenlein K, Grid Evaluation of Pure-Compound Properties, AIChE Annual Meeting, Minneapolis, MN, USA, October 29 - November 3, 2017, Paper 365d. <https://aiche.confex.com/aiche/2017/meetingapp.cgi/Paper/500012>
- [84]. Swern D, Palm WE, Sasin R, Witnauer LP, Viscosity Characteristics of Long-Chain Phosphorus Compounds, J. Chem. Eng. Data 5 (1960) 486–488. 10.1021/jc60008a021
- [85]. Haertel GH, Low-Volatility Polar Organic Solvents for Sulfur Dioxide, Hydrogen Sulfide, and Carbonyl Sulfide, J. Chem. Eng. Data 30 (1985) 57–61. 10.1021/jc00039a019
- [86]. Fang S, Zhao C-X, He C-H, Liu J-Q, Sun J-H, Densities and Viscosities of Binary Mixtures of Tris(2-ethylhexyl) Phosphate + Cyclohexane or n-Hexane at T = (293.15, 298.15, and 303.15) K and p = 0.1 MPa, J. Chem. Eng. Data 53 (2008) 2718–2720. 10.1021/jc8006138
- [87]. Greenwood NN, Thompson A, The Mechanism of Electrical Conduction in Fused Phosphoric and Trideuterophosphoric Acids, J. Chem. Soc (1959) 3485–3492. 10.1039/JR9590003485
- [88]. Diky VV, Frenkel M, Karpushenkava LS, Thermodynamics of sublimation of cubane: natural anomaly or experimental error? Thermochim. Acta 408 (2003) 115–121. 10.1016/S0040-6031(03)00318-6
- [89]. Kybett BD, Carroll S, Natalis P, Bonnell DW, Margrave JL, Franklin JL, Thermodynamic Properties of Cubane, J. Am. Chem. Soc 88 (1966) 626. 10.1021/ja00955a056
- [90]. Bashir-Hashemi A, Chickos JS, Hanshaw W, Zhao H, Farivar BS, Liebman JF, The enthalpy of sublimation of cubane, Thermochim. Acta 424 (2004) 91–97. 10.1016/j.tca.2004.05.022
- [91]. Burger L, Schwarz CE, Sensitivity of process design to phase equilibrium uncertainty: Study of the isopropanol + DIPE + 2-methoxyethanol system, Fluid Phase Equilib. 458 (2018) 234–242. 10.1016/j.fluid.2017.11.027
- [92]. Waring W, Form of a Wide-Range Vapor Pressure Equation, Ind. Eng. Chem 46 (1954) 762–763. 10.1021/ie50532a042
- [93]. Swain N, Singh SK, Panda D, Chakravorty V, Viscosities and densities of ternary liquid mixtures of DEHPA, benzene and ortho-xylene, J. Mol. Liq 85 (2000) 321–330. 10.1016/S0167-7322(00)89016-9
- [94]. Koekemoer LR, Badenhorst MJG, Everson RC, Determination of Viscosity and Density of Di-(2-ethylhexyl) Phosphoric Acid + Aliphatic Kerosene, J. Chem. Eng. Data 50 (2005) 587–590. 10.1021/jc0496645
- [95]. Dalai B, Dash SK, Singh SK, Swain N, Swain BB, Physico-chemical properties of di-(2-ethylhexyl) phosphoric acid with apolar solvents from ultrasonic studies, Phys. Chem. Liq 50 (2012) 242–253. 10.1080/00319104.2011.629420
- [96]. Srirachat W, Wannachod T, Pancharoen U, Kheawhom S, Effect of polarity and temperature on the binary interaction between D2EHPA extractant and organic solvents (kerosene, n-heptane, chlorobenzene and 1-octanol): Experimental and thermodynamics, Fluid Phase Equilib. 434 (2017) 117–129. 10.1016/j.fluid.2016.11.029

Figure 1.
Viscosity (η) data for bis(2-ethylhexyl) hydrogen phosphate at low pressures in logarithmic scale as a function of reciprocal temperature ($1/T$). Data sources are listed in Supplement 1. It is not clear what sources should be preferred at this point.

Figure 2. Stable-phase mole-fraction solubility of D-glucose in water at different temperatures (T_{SLE}) [33–36] along with the data from [31] (red) and [32] (orange) at 1.5–12.0 MPa [31]. The data from [31] and [32] are believed to be incorrect, though mutually consistent.

a)

b)

Figure 3.

Revealing and fixing a publication error for a ternary liquid-liquid equilibrium in 1-butanol + furfuryl alcohol + water. a) Comparison of the original publication [41] (red dots) with literature endpoints (black dots) at $T = 298.15$ K that revealed swapped components. b) LLE data from [41] after correction. In both figures, the LLE diagram is given in mass fractions.

Figure 4.

Mole-fraction solubility of D-galactose in ethanol + water in the temperature range $T = (243 \text{ to } 273) \text{ K}$. The data from [43] highlighted by red color are believed to be erroneous due to inconsistency with another source [44] and endpoints [45, 46] (solubility of D-galactose in water validated separately).

Figure 5. Mole-fraction solubility of lithium bromide in acetone + water in the temperature range $T = (293.15 \text{ to } 333.15) \text{ K}$ [47, 48] (red points) with literature endpoints [49 and Supplement 1] (black). The consistency is reasonable though a higher uncertainty of the solubility in acetone-rich solvent is expected.

Figure 6.

Reported combustion enthalpies ($\Delta_c H^\circ$) for 2-oxazolidone at 298.15 K, [54–57] (left to right) vs. the value based on quantum-chemical calculations (represented by the line). The rejected data [54–55] are shown in orange color. The combustion reaction is as follows:

$$\text{C}_3\text{H}_5\text{NO}_2(\text{cr}) + 3.25 \text{O}_2(\text{g}) = 3 \text{CO}_2(\text{g}) + 0.5 \text{N}_2(\text{g}) + 2.5 \text{H}_2\text{O}(\text{liq})$$

Figure 7. Incorrectly interpreted LLE data (in mole fractions) for acetonitrile + benzene + water at temperatures $T = 298.15$ K and 318.15 K [61] (red) vs NIST-modified UNIFAC [62] prediction at $T = 298.15$ K (blue).

a)

b)

c)

d)

e)

f)

Figure 8.

Consistency of benzonitrile properties in a wide range of temperatures (T) verified with a Span-Wagner type EOS (Supplement 2). a) Saturated vapor pressure over liquid (p_{sat}); b) Liquid density (ρ); c) Enthalpy of vaporization ($_{\text{vap}}H$); d) Liquid heat capacity at vapor saturation pressure (C_s); e) Liquid heat capacity at constant pressure (C_p) at pressure $p = 101.325$ kPa; f) Speed of sound in liquid (u) at pressure $p = 101.325$ kPa. Black points represent unsmoothed experimental values, green points represent smoothed experimental values (stored in the SOURCE database when the unsmoothed values are not reported), orange circles show rejected data, curves represent calculations with the EOS. Data sources are listed in Supplement 1.

a)

b)

c)

d)

Figure 9.

Consistency of dimethyl sulfide properties in a wide range of temperatures (T) verified with a Span-Wagner type EOS. a) Saturated vapor pressure over liquid (p_{sat}); b) Liquid density at vapor-saturation pressure (ρ_s); c) Enthalpy of vaporization ($_{\text{vap}}H$); d) Liquid heat capacity at constant pressure (C_p) at pressure $p = 101.325$ kPa. Black points represent unsmoothed experimental values, green points represent smoothed experimental values (stored in the SOURCE database when the unsmoothed values are not reported), curves represent calculations with the EOS. Data sources are listed in Supplement 1.

a)

b)

Figure 10.

Unresolved inconsistency of saturated vapor pressure for liquid ethylene carbonate: a) Saturated vapor pressure (p_{sat}) vs. temperature (T); b) deviation of the experimental vapor pressure data ($p_{\text{sat,exp.}}$) from the values calculated with a PVExpansion equation ($p_{\text{sat,calc.}}$). The data from [75] are shown in red color; other data sources are listed in Supplement 1.

a)

b)

c)

Figure 11.

Consistency of experimental VLE and volumetric data (dots) for propane + 2,3,3,3-tetrafluoro-1-propene proven with a REFPROP/TDE EOS (curves). a) VLE pressure p as a function of mole fraction of propane in the liquid phase x_{propane} (isotherms in the temperature range from 253.15 K to 300.15 K); b) VLE mole-fraction gas composition y_{propane} as a function of mole fraction of propane in the liquid phase x_{propane} ; c) Liquid-phase excess volume V^E in the temperature range (255 to 300) K; d) Gas density at vapor-saturation pressure $\rho_s(\text{gas})$ – deviations between experimental and calculated data. Data sources are listed in Supplement 1.

a)

b)

c)

d)

e)

f)

g)

Figure 12.

Consistency of solid benzene properties with fluid properties expressed through an EOS. a) Temperature dependence of sublimation pressure (p_{sat}) – experiment (dots) vs PVExpansion equation (curve); b) Deviation of the experimental sublimation pressure data from the PVExpansion equation; c) Sublimation enthalpy ($\Delta_{\text{sub}}H$) – experiment (dot) vs polynomial equation (curve); d) Sublimation enthalpy calculated by the polynomial equation (black dots) and derived from the sublimation pressure calculated by the PVExpansion equation (red dots covering black dots); e) Heat-capacity difference between gas and crystal (C_p^{cr}) directly calculated from the corresponding separate equations (black dots) and derived from sublimation pressure calculated by the PVExpansion equation (red dots); f) Saturated vapor

pressure (line) and sublimation pressure (dot) at the triple point; g) Enthalpy of fusion H_{fus} (line) and the difference between the sublimation and vaporization enthalpies at the triple point. All fluid (liquid and gas) properties are represented by the REFPROP equation of state. Black points represent unsmoothed experimental values, green points represent smoothed experimental values (stored in the SOURCE database when the unsmoothed values are not reported). Data sources are given in Supplement 1.

a)

b)

c)

Figure 13.

Property consistency enforcement for vanillin. a) Saturated vapor and sublimation pressures (p_{sat}) in logarithmic scale; b) Waring criterion [92] $W = RT^2 d \ln(p_{\text{sat}})/dT$ (the gas constant multiplied by the negative derivative of vapor-pressure logarithm by reciprocal temperature) equal to the vaporization (sublimation) enthalpy derived from saturated vapor (sublimation) pressure assuming ideal-gas behavior and negligible volume of the condensed phase; c) Deviation of the experimental vapor and sublimation pressure data from the consistent equations. Black points represent unsmoothed experimental values, green points represent smoothed experimental values (stored in the SOURCE database when the unsmoothed values are not reported), curves represent calculations.

Figure 14.

Viscosity comparison for a set of compounds including bis(2-ethylhexyl) hydrogen phosphate at temperatures around 313.15 K and pressure $p = 101.325$ kPa. Viscosity is represented as a decimal logarithm; the compounds are: phosphoric acid (N=0), bis(2-ethylhexyl) hydrogen phosphate (N=2), and tris(2-ethylhexyl) phosphate (N=3). Data sources are [87] (A), [84] (B), [85] (C), [93] (D), [94] (E), [95] (F), and [96] (G).