

Good Reporting Practice for Thermophysical and Thermochemical Property Measurements (IUPAC Technical Report) [†]

Ala Bazyleva ^{a,*}, Jens Abildskov ^b, Andrzej Anderko ^c, Olivier Baudouin ^d, Yury Chernyak ^e, Jean-Charles de Hemptinne ^f, Vladimir Diky ^a, Ralf Dohrn ^g, J. Richard Elliott ^h, Johan Jacquemin ^{i,j}, Jean-Noel Jaubert ^k, Kevin G. Joback ^l, Ursula R. Kattner ^m, Georgios Kontogeorgis ^b, Herbert Loria ⁿ, Paul M. Mathias ^o, John P. O'Connell ^p, Wolfram Schröer ^q, G. Jeffrey Smith ^r, Ana Soto ^s, Shu Wang ^t, Ronald D. Weir ^u

^a Applied Chemicals and Materials Division, National Institute of Standards and Technology, 325 Broadway, Mailstop 647.01, Boulder, CO 80305-3337, United States

^b Department of Chemical and Biochemical Engineering, Technical University of Denmark (DTU), Søtofts Plads, Building 229, 2800 Kgs. Lyngby, Denmark

^c OLI Systems Inc., 2 Gatehall Dr, Suite 1D, Parsippany, NJ 07054, United States

^d ProSim SA, Immeuble Strategie A, 51 rue Ampere, F-31670 Labège, France

^e Huntsman Corporation, 8600 Gosling Rd., The Woodlands, TX 77381, United States

^f IFP Energies Nouvelles, 1&4 Avenue de Bois-Préau; F-92852 Rueil Malmaison Cedex, France

^g Bayer AG, Engineering & Technology, Building E41, 51368 Leverkusen, Germany

^h Chemical and Biomolecular Engineering Department, The University of Akron, Akron, OH 44325-3906, United States

ⁱ Laboratoire PCM2E, Faculté des Sciences, Université de Tours, Parc Grandmont, 37200 Tours, France

^j Materials Science and Nano-Engineering, Mohammed VI Polytechnic University, Lot 660-Hay Moulay Rachid, 43150, Ben Guerir, Morocco

^k Université de Lorraine, École Nationale Supérieure des Industries Chimiques, Laboratoire Réactions et Génie des Procédés (UMR CNRS 7274), 1 rue Grandville, 54000 Nancy, France

^l Molecular Knowledge Systems, Inc., PO Box 10755, Bedford, NH 03110-0755, United States

^m Materials Science and Engineering Division, National Institute of Standards and Technology, 100 Bureau Dr., Stop 8555, Gaithersburg, MD 20899, United States

ⁿ Virtual Materials Group, A Schlumberger Technology, #300, 3553-31 Street NW Calgary, T2L 2K7 Alberta, Canada

^o Fluor Corporation, 3 Polaris Way, Aliso Viejo, CA 92656, United States

^p University of Virginia (Retired), Nipomo, CA 93444 United States

^q Fachbereich 2 Chemie-Biologie, Universität Bremen, Leobener Straße NWII, 28359 Bremen, Germany

^r Eastman Chemical Company, Building 54D, 200 South Wilcox Drive, Kingsport, TN 37660, United States

^s Cretus Institute, Department of Chemical Engineering, Universidad de Santiago de Compostela, 15782 Santiago de Compostela, Spain

^t AspenTech, 20 Crosby Drive, Bedford, MA 01730, United States

^u Royal Military College of Canada, Department of Chemistry and Chemical Engineering, P.O. Box 17000, Stn Forces, Kingston, K7K 7B4 Ontario, Canada

*Corresponding author: A. Bazyleva – Tel.: +1-303-497-5981; E-mail address: ala.bazyleva@nist.gov.

[†]The authors are given in the following order: Task Group Chair, Task Group Members in alphabetical order.

Abstract: Scientific projects frequently involve measurements of thermophysical, thermochemical, and other related properties of chemical compounds and materials. These measured property data have significant potential value for the scientific community, but incomplete and inaccurate reporting often hampers their utilization. The present IUPAC Technical Report summarizes the needs of chemical engineers and researchers as consumers of these data and shows how publishing practices can improve information transfer. In the Report, general principles of *Good Reporting Practice* are developed together with examples illustrating typical cases of reporting issues. Adoption of these principles will improve the quality, reproducibility, and usefulness of experimental data, bring a better level of consistency to results, and increase the efficiency and impact of research. Closely related to *Good Reporting Practice*, basic elements of *Good Research Practice* are also introduced with a goal to reduce the number of ambiguities and unresolved problems within the thermophysical property data domain.

Keywords: Good Reporting Practice, Thermophysical and thermochemical properties, Material properties; Publication; Reporting principles

1 Introduction

Scientific research frequently involves measurement of properties (thermophysical, thermochemical, *etc.*) of individual chemical compounds, of their mixtures, and of complex materials. Motivations for performing such measurements can have different goals and objectives, which may include characterization of new or industrially important materials, testing or support of a novel theory, correction or improvement in existing measurements, and/or creation of a correlation or prediction model. Unfortunately, these experimental results and, hence, the value of this research can be lost to the broader technical community because of problems with their communication. Proper reporting is essential to ensure interoperability and reusability, which are two of the four FAIR principles for data research (Findability, Accessibility, Interoperability, Reusability) necessary for efficient use of data [1]. Based on a review of the present situation, this Technical Report justifies the need for a set of data reporting principles (*Good Reporting Practice*) that, when accepted by funding agencies, publishers, and authors, will assure high quality in reporting and promote the ultimate effectiveness of property research. The concepts are based on the experience of the authors with property data for substances and mixtures related to thermophysical sciences and chemical engineering. Therefore, some recommendations are common for science in general, while others specifically apply to particular areas of material science, although the principles presented here apply to many scientific areas. Similar recommendations have been published [2-4], but they have a more restricted scope and do not address the whole range of the issues raised here.

The goal of this guide to *Good Reporting Practice* (GRP) is to assure the complete and unambiguous description of research results in a form convenient for subsequent use. We only consider reporting of results of measurements that have already been completed. A small section of the Report will also include a few closely related recommendations covering the planning and conducting of experiments to optimally reduce the number of ambiguities and unresolved problems within the thermophysical domain. Such an initial analysis might form a basis of a more complete guide for *Good Research Practice* (GResP). As with any human endeavor, further refinement of such guides, based upon new discoveries and experience, should be expected.

The information on GRP here is provided in three parts. The first part (General Principles) lists the proposed principles, the second one (Details) provides explanation and justification, and the last one (Appendix) shows application of the GRP principles in different situations. The reader may be interested in particular examples close to his/her activities. An additional part (*Elements of Good Research Practice*), touches on a closely related topic.

2 General Principles

The proposed GRP guidelines are based on the following principles:

1. Measured property data should be published in a numerical format (at least in a supplement).
2. Published data should be well defined (including system, state, and property).
3. All published data should be traceable to their origin.
4. Observations should be distinguished from interpretation.
5. Auxiliary (calibration) data should be identified and provided.
6. Necessary details of experimental methods or computation procedures should be given.
7. Uncertainty in each measured value should be reported and justified.
8. Importing reported data into analysis software should be easy and straightforward.
9. Complex mathematical equations should be provided in a machine-readable form.

3 Details

3.1 Principle 1: Measured property data should be published in a numerical format (at least in a supplement)

When data are measured and used in support of other goals, their numerical values are often not published or are only represented in plots or equations, which happens in the material property measurements field and may happen in other fields. However, Principle 1 directs that this practice is not acceptable – all primary measured property data should be provided in a numerical form (either as tables or as stand-alone values). For instance, if excess molar volume is derived

from density measurements, densities are the primary data and thus should be reported (perhaps, in a supplement) even if excess molar volumes were the main target of the study.

If the measuring device output (*e.g.*, electrical potential) is other than the target property (*e.g.*, enthalpy change), the necessity of reporting those raw data in addition to the target property is determined by the specifics of the experiment. Generally, the measured value may be converted several times, and the first quantity, which can be identified as a substance or material property, should be considered as directly measured and, hence, should be reported. This principle does not preclude reporting of any properties derived from the directly measured properties if needed, as well as equations and plots. Additional discussion is given in Case 1a in the Appendix. If the property value is a composite of experimental and computational values, all contributing components should be provided.

Four reasons to explicitly report the numerical values of a measured quantity in their original form are:

- I. Measurements may be labor intensive and require costly equipment, highly pure samples of substances or materials, and qualified personnel, so reproducing them would be expensive, and needless replication is not affordable for the community.
- II. Numerical values are required for many applications, such as development of models (equations) or training and testing of theoretical and predictive methods. Data published as plots lose accuracy of representation and require labor-intensive, error-prone processing. Data reported only as an equation or smoothed data lose statistical information necessary for repeatability-based Type A Uncertainty evaluation [5], such as scatter (deviations of the measured values from smoothing curves or equations) and data density (distribution of data points in the multi-dimensional space of state variables). In addition, a single error in published equation parameters causes the loss of the entire data set.
- III. The data measured in publicly funded research are often required to be made public according to the conditions of grants. Obfuscating these data through poor presentation does not respect the public interest. An exception can be commercially funded work that may have restrictions about what may be released for publication.
- IV. If the data provide the basis for any derivations, developments, or conclusions in scientific papers, not publishing them leaves those developments and conclusions untestable and unjustified. In addition, providing raw experimental data allows corroboration of derived properties and all further calculations done by the authors.

Providing equations or smoothed values, even with their uncertainty estimates, cannot obviate the need for the raw data. One reason: many existing methods and tools handle distinct data points rather than equations and use summation rather than integration. Another reason is that it is always questionable, whether an equation or a set of smoothed values with estimated uncertainties correctly and completely reflects the density and the scatter of the raw data. In addition, there is no single preferred way to represent equation uncertainty.

Reporting should be unbiased and include all results. When property data are reported, the values should be easily identified and not hidden in the text or footnotes to tables or figures, even if they only have an auxiliary role. Interested readers can miss the existence of such data or need to spend enormous time searching for each piece of reported data. Hence, Principle 1 includes not only the requirement of reporting property data in a numerical form, but also the necessity to **visibly** report property data (*i.e.*, in a form which would allow readers to find those values easily without reading the whole paper and supporting information). Separation of all experimental results into a single section, not mixed with discussion, is strongly suggested.

3.2 Principle 2: Published data should be well defined (including system, state, and property)

Well-defined data imply:

- a) identification of the substance or material,
- b) description of the studied samples,
- c) identification of the properties and states (phases) to which they are related,
- d) all phases in equilibrium and state variables to completely determine the conditions at which the properties were measured, and
- e) unambiguous statement of units for every property and state variable.

Identification of substances typically includes their chemical nature, composition of mixtures, and stereoisomer composition. It is advised to give at least two chemical identifiers for each compound (*e.g.*, systematic chemical name, CAS number, structure, IUPAC PIN (Preferred IUPAC Names) [6], InChI [7]). In the cases when complete chemical

identification is impossible, sufficient information should be provided to specify the subject material. Some of such cases are discussed in Case 2b in the Appendix.

Sample description is an essential part of the reported data, because it indicates how much the actual studied system deviates from the nominal one (*e.g.*, pure compound). It should include the source of compounds, purification methods used, and final purity of the substances with description of the purity analysis method. If components are taken from suppliers without further purification, this fact should be stated together with the numerical purity claimed by the supplier (not just purity grades). If the purity of a sample is determined by the authors, the procedure for determining the purity should be described. Purities should be based on well-established analytical methods, not on a simple comparison of the measured physical properties to published literature values. Claimed purities should be justified by sensitivity and uncertainty of the analytical methods. If the substance can change its composition during measurements and handling, that should be also reflected in the report by providing the final composition or, at least, the protocol of handling, which can provide the basis for future evaluations (see an example in the “Elements of *Good Research Practice*” section). It is also suitable to record the age of the sample, since some substances (*e.g.*, hydrocarbons) are known to change properties due to aging.

Property names should be commonly accepted, preferably defined in official publications (*e.g.*, [8]). The terminology should be correctly used. If property names are ambiguous and/or have various definitions, the exact definition that is used should be given, preferably with the corresponding defining equation (*e.g.*, Henry’s law constant, osmotic or activity coefficient based on mole fractions or molalities). Units (preferably, SI units) should be specified. Some properties imply additional information that should always be given: for example, molality in a multi-component system requires definition of both the solvent and the solute, and enthalpy of dilution requires a proper definition of the quantity to which the enthalpy is referred (more details are given in Case 2t in the Appendix). Several properties should always be accompanied by the corresponding non-state variables, such as wavelength or frequency (*e.g.*, speed of sound requires the corresponding sound-wave frequency used in the measurements or refractive index needs a specified light wavelength). Equilibrium and non-equilibrium states should be distinguished with great care [9]. For equilibrium properties, the number of independent state variables for any phase state or phase equilibrium is defined by the Gibbs phase rule. The phases present should be identified and quantified. If the complexity of the system (such as biological materials or petroleum fractions) precludes complete identification of all variables, a discussion of unspecified components and their importance should be given.

It may be difficult to verify whether the data set is defined in a sufficient and self-consistent way. Submission of data in machine-readable formats (see Principle 8) and parsing by computer software can greatly facilitate that task. Automated analysis can also reveal and prevent typographical errors. While we cannot recommend any publicly available convenient data entry tools, such development is highly encouraged.

3.3 Principle 3: All published data should be traceable to their origin

If the data are obtained from work reported by the authors, sufficient details of their origin (experiment or calculation) should be given. If the data are taken from other references (including previous publications of the same authors), a citation should be given along with identification of their source (experimental, calculated, or evaluated). If the information combines elements of different nature, all the elements should be identified (see Principle 5).

A well-defined citation should identify the original source sufficiently for easy access. The type of publication should be specified (journal article, dissertation, report, *etc.*). Web references should be complemented by other measures of access, when the latter exist. If the cited document exists only online, the date it was accessed should be indicated. Specific identifiers (DOI, ISSN, ISBN, URL, WorldCat, secondary sources) may be appropriate, but their use is usually regulated by publishers. Thus, publishers should engage in *Good Reporting Practice* to provide the best possible identification of cited sources [10]. Full journal names and place of publication rather than abbreviations should be given for uncommon or hard-to-access journals.

3.4 Principle 4: Observations should be distinguished from interpretation

Many experimental observations are not straightforward (*e.g.*, similar observations can be attributed to different phenomena). Hence, what is actually seen in the experiment and how it was interpreted by the authors should be clearly identified and separated. For example, observations such as peaks in DSC (differential scanning calorimetry) experiments or turbidity during phase equilibria measurements frequently use additional information for assignment to a

phenomenon, such as a certain phase transition or equilibrium. That information may be obtained in the same study (*e.g.*, structural analysis of the solid phases), taken from the literature (*e.g.*, phase diagram), or merely assumed. The basis of any interpretation should always be clearly stated.

3.5 Principle 5: Auxiliary (calibration) data should be identified and provided

If established properties of a reference substance were used for calibration, their values or a reference to their source should be given. If auxiliary properties of the studied compounds/materials were used in deriving a target property (*e.g.*, proper derivation of enthalpies of combustion requires the knowledge of heat capacity, density, and thermal expansion coefficient of the studied and all auxiliary compounds), their values and reference to their source should be provided. Another example is indirect composition measurements for phase equilibrium experiments (typically, for vapor-liquid equilibria and liquid-liquid equilibria), when compositions of co-existing phases are determined through density or refractive-index calibration curves: in this case, the calibration results should be provided in numerical form as in a table. If additional data have been used for modifying the measured data (*e.g.*, virial coefficients for obtaining fugacities or density data for transforming composition units), they should be also identified in detail. The principles of metrological traceability of measurement results are discussed in detail in [11].

3.6 Principle 6: Necessary details of experimental methods or computation procedures should be given

Each experimental method has a specific set of important details which should be reported along with the results, in order to provide evidence of correct implementation and allow verification of certain operations. All necessary information should be given so one can repeat the measurements in one's own laboratory. Specific recommendations are available for several methods, in particular, from the Experimental Thermodynamics Book Series, Volumes I-X (*e.g.*, [12]) and IUPAC reports published in the Pure and Applied Chemistry journal. It is acceptable to cite previous references, where complete details of the applied method have been published, but a summary of the method used should still be provided. If a method is new or was described in a difficult-to-find publication (*e.g.*, in a very old source, conference proceedings, thesis, or country-specific local journal) or requires a sophisticated language translation, it is strongly suggested that comprehensive experimental details should be given or another more accessible publication should be cited as well.

In some situations, calibrations can be done in a variety of ways, even if the same reference substances are used; in such cases, the calibration procedure should be fully described (see Case 6d in the Appendix).

3.7 Principle 7: Uncertainty in each measured value should be reported and justified

Giving uncertainties is becoming more common but giving them without justification does not provide credible information. Such justification should cover various factors such as purity of the samples, departure from the equilibrium state, accuracy of maintaining the experimental conditions, and uncertainties of measuring devices and calibration data. Several recommendations on assessment and expression of uncertainties have been already published to guide scientists [5, 13]. Validation of the method and equipment with recommended reference materials may also provide a basis for uncertainty estimation, as discussed below in the "Elements of *Good Research Practice*" section. Just as the value of the measurand needs considerable information as to how it was obtained, so does the measurement uncertainty. The uncertainty budget (VIM 2.33 [14]) should be sufficient and justified.

One of the possible ways to estimate a possible ambiguity of the property data knowledge is comparison with previously published values. Authors are expected to discuss the past knowledge (preceding publications reporting the same or related properties), though finding relevant prior publications can be a non-trivial task. For moderate-size organic molecules and common inorganic compounds, a free online tool, ThermoLit [15], is available to identify relevant literature sources in the NIST (National Institute of Standards and Technology) data archive with which to compare new results. ThermoLit is intended to aid researchers and reviewers in determining relevant literature sources for a given experimental measurement, though it cannot replace a comprehensive literature review that should be performed by researchers. In order to provide the best value to the readers, the authors should identify how the literature search was

done, especially in review papers. It should be noted that previously published values may not be always correct, so they may not be considered as reference values that should be reproduced. Unbiased reporting of the new results and analysis of differences and their possible reasons are expected from the authors.

3.8 Principle 8: Importing reported data into analysis software should be easy and straightforward

A good publication should make its reading and use easy. Modern publication techniques allow all data to be in an easy-to-parse form, perhaps in a supplement. Plain tables representing all state variables in a single flow (see Table 1, as an example) are better than “good-looking” tables consisting of small pieces with additional information in sub-headers. All symbols should be defined in a table header or footnote, so that they are self-explanatory and self-sufficient as should all equation symbols and terms. As suggested above, collecting all measured data in a separate section of the article or report facilitates their identification and complete capture by readers.

Table 1: Density of liquid water (T is temperature, P is pressure, ρ is density) ^a

T/K	P/kPa	$\rho/(kg\cdot m^{-3})$
300.0	1000	997.0
300.0	2000	997.4
300.0	3000	997.9
300.0	4000	998.3
300.0	5000	998.7
350.0	1000	974.1
350.0	2000	974.6
350.0	3000	975.0
350.0	4000	975.5
350.0	5000	975.9
400.0	1000	937.9
400.0	2000	938.4
400.0	3000	938.9
400.0	4000	939.4
400.0	5000	939.9

^a Standard uncertainties u are $u(T) = 0.1$ K, $u(P) = 10$ kPa, $u(\rho) = 0.2$ kg·m⁻³.

This and the following principle are coherent with OPEN Government Data Act of the United States [16] and, probably, other similar national and international acts.

3.9 Principle 9: Complex mathematical equations should be provided in a machine-readable form

Providing complicated equations in a machine-readable form (symbolic mathematics or programming languages) not only saves time for users, it also prevents errors that frequently appear when those equations are typed in the text. An example of a machine-readable representation can be found in [17], where a C++ code and a Python script are given in the attachment. This principle does not preclude reporting also in a human-readable form. In addition, providing check values for verification of calculations by equations is suggested.

3.10 Additional note

Standing separate from those nine Principles and applicable to all of them, immediate submitting of *corrigenda* and making them visible from the main publication Web pages, as soon as errors are detected, is also part of *Good Reporting Practice*.

Several examples and specific suggestions for the described principles can be found in the Appendix to this Technical Report. It should be noted that similar requirements and recommendations have already been accepted by some journals [18,19].

4 Elements of Good Research Practice

Upon preparation of the guide on GRP, it has appeared that reporting issues are frequently tied with the design of experiments. In a significant number of cases, best reporting would require conducting additional experimental measurements. That is why we decided to include some elements of *Good Research Practice* (GResP) in this Technical Report. However, GResP is much broader than the current project, and we do not attempt to cover it in a comprehensive way.

Well-designed research should solve problems rather than create them. Sometimes problems can be revealed but not immediately solved; however, well-designed research should minimize the number of remaining questions. Following these elements of *Good Research Practice* should achieve that goal.

4.1 GResP 1: Content (planning)

More complex research should also address simpler questions, which are limiting behavior providing a basis for consistency checks and model development. A few examples are as follows:

- a) A measurement for a binary mixture should typically involve pure components as limiting cases (endpoints), and a measurement for a multicomponent mixture should similarly include the corresponding binary mixtures. This information is valuable for verification of the consistency of complex data with less complex data, for modeling purposes (*e.g.*, for derivation of binary interaction parameters), and for identification and compensation of systematic offsets. In addition, the accuracy of the applied experimental method/apparatus can be assessed by comparison with available literature data, which are usually more readily available than for more complex mixtures. For example, vapor-liquid equilibrium (VLE) measurements for binary mixtures should be normally accompanied by the pure-component vapor pressures measured with the same apparatus [20], and pure-component melting properties should be presented as part of binary SLE (solid-liquid equilibrium) phase diagrams.
- b) The reporting of high-pressure measurements (such as for density or speed of sound) should also include measurements from low-pressure regions when feasible.
- c) Related properties necessary for analyzing the main data should be measured, if possible (*e.g.*, melting temperatures and enthalpies to complement solid solubility data).

Measuring over a range of conditions (temperature, *etc.*) is more valuable than for a single point. If the range is limited by special points (for example, liquid-liquid equilibrium in a ternary system limited by an isothermal critical-mixing-point, *i.e.*, plait point, or solid-liquid equilibrium lines in mixtures limited by eutectic or peritectic points), the limiting behavior should be studied as well when feasible.

4.2 GResP 2: Methodology (experimental procedure)

The method and apparatus should be confirmed on well-studied test systems. Several reference materials have been recommended for verification of the methods [21-24]. Some experimental methods used in the past were shown to give erroneous results due to the nature of the experiment. Authors should avoid using such methods. One example is the isothermal evaporation method to determine SLE: it has been demonstrated that this method does not yield either stable or metastable solubilities [25].

A wide variety of methods are available for the measurement of phase equilibria. The choice of the most suitable method depends on the properties of the components and on the phenomena to be investigated. A detailed understanding of the methods is fundamental not only for the choice of the best method, but also for the evaluation of the experimental data [9].

If properties of a material or substance are studied, the stability of the material should be assessed. Particularly during long-lasting experiments at higher temperatures, thermolabile substances can partly degrade, leading to different thermophysical property results. In these cases, the sample should be analyzed both before and after the experiment with the comparisons being reported [26].

If equilibrium properties are studied, the equilibrium state should be demonstrated, for example, by conducting the experiment over different times and showing when there is no change of state.

Hygroscopic materials (*e.g.*, ionic liquids, DES (deep eutectic solvents)) require a water-content analysis, especially when properties are sensitive to water impurity. Limitations of the Karl-Fischer method should be pointed out when that method is used for water content determination [27]. Furthermore, in the case of ionic liquids, the content of other potential impurities such as metals and/or halides should be quantified to properly define their purity.

All newly synthesized compounds should be accompanied by the proof of their identity, *e.g.*, NMR (nuclear magnetic resonance) spectra, high resolution mass spectra, or elemental analysis for organic compounds, or XRF (X-ray fluorescence) or a similar element-sensitive method for inorganic materials. It is suggested that authors use a combination of methods if a single method cannot provide an unambiguous proof. One common example is protic ionic liquids: they do not consist entirely of ions, but may contain a fraction of molecular species that can unevenly evaporate upon exposure to vacuum conditions; this would lead to non-stoichiometric composition, which requires proper characterization.

Measurements for verification of the method or apparatus should be performed over the entire temperature range of interest, since many experimental errors are due to inhomogeneities of the temperature within the apparatus, a factor that might not be obvious if test measurements are performed only at ambient temperatures.

4.3 GResP 3: Validation

Steps toward validation of the measured values are strongly encouraged whenever possible. For example, Gibbs-Duhem equation-based consistency checks exist for vapor-liquid equilibria (VLE). Other checks are being developed for other phase equilibria, such as the publicly available tool for the analysis of Gibbs energy surfaces of organic systems [28]. Good practice should include such techniques when appropriate, even while recognizing their limitations [29].

If new data are inconsistent with previously published results, the difference should be discussed, as well as all concerns about the correctness of the new results. In addition to the common search resources, there may be others, such as the already mentioned ThermoLit for thermodynamic properties [15].

5 Conclusions

Basic principles of effective reporting of property measurement results have been summarized, with detailed guidelines and some case studies provided for the areas of thermodynamics and thermophysics. Acceptance of those principles would assist in more complete transfer of the knowledge acquired in scientific research. In addition, greater efficiency would be made of the funds distributed for that research to benefit the entire community of taxpayers, readers, and users, as well as the researchers themselves.

Note

Trade names are provided only to specify procedures adequately and do not imply endorsement by the National Institute of Standards and Technology. Similar products by other manufacturers may be found to work as well or better.

Acknowledgement

This manuscript was prepared within the framework of IUPAC, project # 2019-013-1-100, “Good Reporting Practice for Thermophysical and Thermochemical Property Measurements”.

Membership of the sponsoring body

Membership of the IUPAC Physical and Biophysical Chemistry Division for the period 2020–2021 was as follows:

President: Tim Wallington (USA); **Past President:** Ronald Weir (Canada); **Vice President:** Pierangelo Metrangolo (Italy); **Secretary:** Attila Császár (Hungary); **Titular Members:** Jeremy G. Frey (UK); Frances Separovic (Australia); Zhigang Shuai (China); Hiroko Tokoro (Japan); Ilya Vorotyntsev (Russia); Angela K. Wilson (USA); **Associate Members:** Modou Fall (Senegal); Seung-Joon Jeon (South Korea); Theo Christian Kurtén (Finland); Joaquim Luís Bernades Martins de Faria (Portugal); Luis A. Montero-Cabrera (Cuba); Vessela Tsakova (Bulgaria); **National Representatives:** Gordana Ciric-Marjanovic (Serbia); Mohamed Deyab (Egypt); Majdi Hochlaf (France); Lynda C. Ngozi-Olehi (Nigeria); Renata Orinakova (Slovakia); Ilya Vorotyntsev (Russia).

Appendix – Application of the GRP principles in specific cases

Principle 1 (Measured property data should be published in a numerical format)

Case 1a (measured data to report): Temperature rises and any corrections are to be reported for combustion experiments along with the target property, enthalpy of combustion. The minimum requirement is to provide the calibration procedure and its results (see Principle 5), so that the reader can evaluate the quality of the methodology. In contrast, adiabatic calorimetry usually suffices reporting temperature increments per each heat capacity value. If the target value is determined by a break in the slope of directly measured data, reporting those primary data as Supporting Information is strongly suggested.

Case 1b (number of significant digits): The precision of numerical representation should be sufficient to keep the deemed measurement precision. For example, compositions should not be rounded below the accuracy of the mixture preparation (*e.g.*, mole fraction 0.9011, not mole fraction 0.9, if the claimed standard uncertainty of mole fraction x , $u(x)$, is 0.0002) or temperature measurements (*e.g.*, 299.11 K, not 299 K, if the claimed standard uncertainty of temperature is $u(T) = 0.03$ K). The number of significant digits in representation is still an area of active discussions in different fields (*e.g.*, [30]).

Case 1c (reporting method validation results): In addition to the main work, properties of additional systems are frequently measured in order to validate the applied instrumentation. However, those values are typically not reported (only a statement of good consistency is given) or provided as comparison graphs. Those newly measured data constitute a valuable piece of information, irrespective of how well that additional system has been studied in the literature. Such data should also be provided in numerical form.

Case 1d (reporting measured endpoints): When mixture properties are reported, some authors skip limiting values (endpoints), *i.e.*, pure-component properties for binary mixtures, or binary-mixture properties for ternary or higher-order mixtures. Typical examples are vapor pressures or densities for pure components in VLE or volumetric experiments for mixtures. Actually, these properties are often measured by the authors for use in obtaining derived properties, such as activity coefficients or excess volumes. This information is important for assessing systematic errors and reliability of the applied methods, so these data also should be reported together with the mixture data.

Case 1e (completeness of vapor-liquid equilibrium (VLE) data reporting): Some authors measure VLE pressure–total composition data but report the data in either pTx format (assuming that the total composition z is equal to the equilibrium liquid composition x) or $pTxy$ format, where x and y (equilibrium liquid and gas-phase compositions) are calculated from the total composition rather than directly measured. It is acceptable to make those operations, but the measured pTz data should also be published along with additional information on the cell and initial liquid-phase volumes (thus creating an extended $pT\rho z$ data set), so that readers can do an independent verification.

Case 1f (completeness of data reporting in micelle-formation studies): The critical micelle concentration (cmc) for surfactants is often derived from a change in slope of electrical conductivity or surface tension with surfactant addition. However, often no directly measured property values *vs.* concentration are reported. This is not consistent with Principle

1 – all newly measured properties, which are used in support of cmc calculations should be provided in numerical form with corresponding uncertainties.

Case 1g (reporting mixture compositions): When reporting the complete mixture compositions, it is important to indicate which components are measured and which remaining component was obtained by subtracting the sum of the other components from 1 or 100 %. If the absolute amounts of all components were measured and then divided by the sum, it should be also indicated.

Principle 2 (Published data should be well defined)

Case 2a (correct terminology): Appropriate terminology should be used. For example, the term “deep eutectic solvent” has been too casually used in the literature, referring to any mixture exhibiting eutectic behavior. Few organic mixtures form truly deep eutectics, *i.e.*, lowering the liquid temperature below that predicted by ideal liquid mixture modeling. Another example is interchangeable use of “electrical conductivity” and “electrical conductance”, which are, in fact, two different properties. Old terms, such as “latent heat” or “Gibbs free energy”, should be also avoided. The term “heat of mixing” is ambiguous; “enthalpy of mixing” should be used instead. Readers are referred to IUPAC Green Book [8] for currently accepted definitions in the field.

Case 2b (identification of materials): If the phase state of the material (*e.g.*, alloy) is not known, the procedure of its preparation and treatment should be given, which allows reproducing the state. If the material is a natural or industrial sample, its nature, origin, time and place of obtaining may identify it.

Case 2c (reporting experimental pressure and temperature): If measurements are conducted under atmospheric pressure, the actual experimental pressure should be provided, even if the studied property only weakly depends on pressure (*e.g.*, liquid/solid density). This is needed to comply with the Gibbs phase rule to report the correct number of state variables and to avoid any ambiguity about the experimental conditions. The term “atmospheric pressure” should not be used, since it is not well defined: it varies significantly with elevation and weather conditions (compare 101 kPa at sea level vs about 80 kPa at an elevation of 2000 m). The term “room temperature” should be also avoided.

Case 2d (reporting the phase): A number of chemical compounds, mixtures, and materials can form supercooled liquids. When properties of such materials are reported, the corresponding phase (with a clear note that the liquid is supercooled) should be explicitly specified to avoid incorrect interpretation. It is strongly recommended that the method of inducing supercooling should also be indicated.

Case 2e (identification of the studied phase equilibrium type): When studying phase equilibria, the authors are strongly encouraged to clearly identify the type of equilibrium (solid-liquid, liquid-liquid, bubble or dew point, *etc.*), classify their system according to common classification schemes, such as those of van Konynenburg and Scott [31], Privat and Jaubert [32], or Rhines [33], and identify the position of the experimental conditions on the phase diagram (especially for high-pressure measurements).

Case 2f (providing molecular structures): Although standard names and various unique identifiers are often sufficient to unambiguously identify a compound, a drawn molecular structure is the best way to clearly identify a chemical, especially for a complex compound. Drawn molecular structures are essential when the data are associated with substructures, *e.g.*, group contribution values or specific charge configurations.

Case 2g (distinguishing calibration and test results): Calibrations and equipment tests should be distinguished. Calibration means the use of a property value obtained from another source to establish the relation between the measured signal and the target property. Calibration does not produce new knowledge about the calibration substance. Test measurements, when other known values of the measured property are not used, produce additional property values, which are then used to validate instrument accuracy by consistency with previously reported values. In particular, when vaporization is characterized by correlation gas chromatography, test substances and studied substances should be explicitly distinguished. Properties of the reference substances (vapor pressures, vaporization enthalpies) are used to produce the correlation equations. Though application of the correlation equations to the chromatography experiments may return values different from the ones used for correlation, they are not new experimental values.

Case 2h (reporting mixture compositions): Ambiguity of reported composition can be caused by:

- i) not identifying the mode of representation (mole or mass fraction),
- ii) not defining the basis of ppm or ppb (mass or amount of substance),
- iii) not identifying the component for which the fraction is reported,
- iv) giving volume-based composition, such as molarities, normalities, volume fractions, or mass of solute per volume,
- v) not identifying the type of composition: apparent or actual (for example, NO₂ as a single species or as an equilibrium mixture of species including NO₂ and N₂O₄), in the overall system (all coexisting phases) or for a particular phase in a multi-phase system.

In general, volume-based compositions are not recommended, since they are not sufficient to derive the amount fractions of the components. This can be illustrated as follows: molarity (as well as volume fraction, vapor composition in mass per volume, and other volume-related expressions of concentration) of the same mixture changes with temperature and pressure, which requires an exact specification of the temperature and pressure to which the molarity is attributed. Mass of solute per volume, in addition, requires the specification of whether it is related to the volume of solution or solvent. In all cases, mass densities of solution or solvent are needed in order to establish the proportion of the components, and such information is frequently unavailable. If needed for any reason, volume-based compositions can be given, but they should be accompanied by a mole- or mass-based expression. This is consistent with the Journal of Chemical and Engineering Data and Journal of Chemical Thermodynamics requirements [19,34]. An exception can be when the original measured value is volume-based (*e.g.*, from light absorption) and the authors do not have density data; that should be made clear in publication.

One additional ambiguity is related to rounding. Compositions given in fractional or percentage units should total to 1.0 or 100.0 %, respectively. Many standard software packages used to prepare data for publication will round each component composition separately to a certain number of decimal places but not impose such a restriction on the total values. See more details in Case 1g.

Case 2i (data reporting for mixtures with pseudo-components): If the concept of mole is used for a substance not consisting of one kind of molecule (*e.g.*, alloys, deep eutectic solvents (DES), or pseudo-components), a definition of the corresponding entity (an atom, a molecule, or a larger entity formed by a group of atoms or molecules) should be given. The actual compositions of pseudo-components should be given if possible; otherwise, the complete identifying information is needed (source, treatment, applicable analytical characterization), so that the observation can be repeated with only the presented information.

Pseudo-components can be treated as single species for single-phase properties (*e.g.*, liquid-phase density). However, their components are expected to be unevenly distributed between phases in phase equilibrium measurements. Two situations are possible in this case. When it is not experimentally feasible to obtain detailed molecular information on the pseudo-components (*e.g.*, representation of oils as an artificial pseudo-component mixture based on distillation curves), it is acceptable to treat pseudo-components as single components. In all other cases in in phase equilibrium experiments, pseudo-components should be considered as mixtures.

Case 2j (stereoisomers): When studying thermophysical and thermochemical properties, authors sometimes do not consider stereoisomeric composition for chiral compounds, probably, assuming that the properties of stereoisomers are identical. This assumption is not always true even for enantiomers. For example, the solid-liquid and solid-vapor equilibrium properties of an enantiomer mixture depend on the composition of the mixture. Hence, the stereoisomeric composition should always be provided (if unknown, that fact should be disclosed).

Case 2k (reporting compositions of solutions prepared from hydrates or solvates): Reporting solubility of a hydrate in water raises the question whether the crystalline water was considered as part of the solute or part of the solvent. Reporting the content of the anhydrous substance with clear indication of the solid phase is preferred. A similar approach should be used for reporting compositions of aqueous solutions prepared from hydrated substances in other types of experiments: the composition of the corresponding anhydrous substance should be reported with a proper adjustment for the water molecules released from the hydrate to become a part of the solution.

Case 2l (reporting solubility data for hydrates and solvates): Solubility of hydrates in an organic or mixed solvent requires additional information. Depending on the conditions, the solid phase may be a hydrate or anhydrous, and the solvent

may contain different amounts of water extracted from the solid. For a full specification of the results, the content of all components (the solute, water, and organic solvent) in the saturated solution should be determined and given in the paper, along with the identity of the final solid phase. It may appear that the final solid form that appeared during equilibration was different from the solid initially taken before the dissolution experiments. If experimentalists are unable to identify the final solid phase, the complete composition of the saturated solution along with the initial material balance (in combination with temperature and pressure) still forms a well-defined data set provided that stable equilibrium has been achieved. Otherwise, an ambiguity remains as to whether the equilibrium is metastable (example: D-glucose + water [35]). If the authors can only give the content of the initial solute, the data set is incompletely defined. One possible, though not optimal, way to completely specify the final state is to include the initial material balance (initial proportion of the solute and solvent); in this case, while the system is not fully characterized, it is fully defined and therefore can be reproduced or modeled by other researchers.

Case 2m (gas hydrates formation equilibria): It should be clearly identified what kind of composition is given (composition of a phase in equilibrium, composition of the initial phase, or global composition of the multiphase system). In the case of the initial phase, the proportion of the mixed initial phases should be reported, because different proportions of the same initial phases are expected to result in different equilibrium compositions.

Case 2n (reporting property data for drugs): Medical substances may be in the form of molecular species or salts, anhydrous materials or hydrates, or stable or metastable polymorphs. Different forms may be sold under the same name: *e.g.*, vitamin C can be marketed as ascorbic acid and as calcium ascorbate; irrespective of its stereo structure, naproxen can be sold as its base form (2-(6-methoxy-2-naphthyl)propanoic acid), its salt form (sodium 2-(6-methoxy-2-naphthyl)propanoate), or its hydrated salt form with different water contents. The studied form should be unambiguously identified (including information indicating whether it is stable or metastable under the studied conditions), and the amount of water in the hydrate should be proven.

Case 2o (reporting excess enthalpies for ternary mixtures): Excess enthalpies for ternary mixtures are frequently obtained by combining the binary excess enthalpy and mixing enthalpy of the third component with a binary mixture. A clear definition of the experimental conditions (composition of the solutions before and after the mixing) is needed, and reporting the measured mixing enthalpy of the added component with the binary solvent is strongly suggested.

Case 2p (reporting standard state): If activity coefficients are reported, the standard state for the corresponding solute should be specified (*e.g.*, pure liquid solute, infinite dilution).

Case 2q (reporting transport properties): In the case of the transport and acoustic properties, results of an analysis of the measured property (such as viscosity or speed of sound) as a function of experimental conditions (shear-, strain-, stress-rates, time, or frequency) should be reported to verify whether this property is condition-independent (*i.e.*, Newtonian fluids or non-dispersive medium). In the case of the viscosity for non-Newtonian fluids (*i.e.*, shear-rate and/or time dependent viscosity) or frequency-dependent speed of sound (*i.e.*, dispersive medium), the measured property should be defined as an apparent property that is only a complete description under the applied experimental conditions. Commercial equipment, such as a classical viscometer or sound velocity meter, do not, generally, offer the possibility of conducting such an analysis (*i.e.*, the property is measured under fixed conditions that cannot be modulated by the users). Such measured results should be analyzed with great care. This example also involves elements of *Good Research Practice* because it may require additional measurements for complete characterization of the system.

Case 2r (application of measured speed-of-sound data): As mentioned above, speed-of-sound data should be accompanied by the frequency used. Additionally, the interpretation of speed-of-sound values and their usability for the determination of related thermodynamic properties can only be done when the absorption coefficient or relaxation regions are known. For example, observed ultrasound absorption spectra of some ionic liquids show frequency dependence with transducers operating in conventional ultrasound devices (such as the speed of sound equipment used in several research groups [36]). In that case, the speed of sound cannot be regarded as the thermodynamic quantity, since the Newton-Laplace equation is not applicable. Thus, related thermodynamic properties (*e.g.*, isentropic and isothermal compressibility, isentropic static bulk modulus) cannot be obtained. This example also involves elements of *Good Research Practice* because it may require additional measurements for complete characterization of the system.

Case 2s (ion exchange in liquid-liquid equilibrium): Phase equilibria in ternary mixtures involving two salts or a salt and an ionic liquid with no common ions cannot be treated as true ternary phase equilibrium because ion exchange can occur. Though the electroneutrality reduces one degree of freedom, the content of 3 out of 4 ions in each phase should be obviously given in order to completely define the compositions. Giving only the content of any initial neutral component in any co-existing phase is misleading, because the amounts of the constituent ions may differ. This does not apply to incipient conditions when the amount of the second phase is small, and does not significantly affect the composition of the first phase.

Case 2t (reporting enthalpies of dilution): Enthalpy of dilution requires a definition of the compound, to the amount of which the enthalpy is assigned. That may be the solvent, the solute, the added solvent, the final solution, *etc.* If a concentrated solution of A + B is diluted with B, the reported enthalpy of dilution of, for example, $0.75 \text{ kJ}\cdot\text{mol}^{-1}$ needs a clarification whether it is given per 1 mole of component A, or per 1 mole of component B in the initial state, or per 1 mole of component B in the final state, or per 1 mole of the added component B, or per 1 mole of the initial solution, or per 1 mole of the final solution.

Case 2u (reporting potentiometric titration results): Speciation of solutions is often determined using potentiometric titrations, primarily in aqueous systems; such measurements are frequently the main source of information about acid-base and complexation equilibria. Equilibrium constants are typically regressed by fitting potentiometric titration data to a previously assumed reaction scheme. Therefore, the derived equilibrium constants can be model-dependent. While the vast majority of studies report equilibrium quotients or constants, the raw titration data are often not reported or are limited to only sample titration curves. The lack of raw titration data prevents the reader not only from reproducing the original experimental data, but also from re-interpreting the data with an alternative reaction scheme. This is extremely important because fitting the original potentiometric data may be the only way to reconcile experimental data from various sources instead of having to deal with disparate equilibrium quotients or equilibrium constants. Authors may be reluctant to include their raw potentiometric data because of its large volume; such data can be placed in supplementary information.

Case 2v (reporting relative permittivity): Relative permittivity (also known as dielectric constant) is a typical example of a property requiring the knowledge of a non-state variable. In most cases, static relative permittivity, *i.e.*, the value at zero frequency, is a desirable property for modeling purposes. However, relative permittivity is measured at some other frequency, and this is often not provided by authors, hindering its use because how to extrapolate to another or zero frequency is not known.

Principle 3 (All published data should be traceable to their origin)

Case 3a (proving the source of auxiliary data): Proper treatment of combustion calorimetry experiments requires knowledge of densities and heat capacities for all materials used in the combustion experiments (main compound, cotton fuse, *etc.*). When the values are reported in a publication, they should be accompanied by the corresponding source (if taken from other publications) and the corresponding experimental method (experimental or predicted).

Case 3b (origin of literature data for comparison): When measured data are compared with the data for similar compounds, the origin of those data (citation) and their nature (measured or predicted) should be also indicated.

Case 3c (distinguishing previously published data): When properties of mixtures are reported and the endpoints (pure-component data or binary-mixture data for ternary mixtures) have been published in the past, all the previously published data should be accompanied by the citation of the existing publication.

Principle 4 (Observations should be distinguished from interpretation)

Case 4a (interpretation of DSC peaks): Peaks in DSC heat-flow profiles indicate only the existence of an exo- or endothermic process in addition to the base heating of a studied material. The peaks do not provide any information about the nature of the process (solid-phase transition, melting, dissolution of impurities, *etc.*). For example, two DSC peaks for a 95% pure substance do not always mean a solid-phase transition followed by melting of the main compound; they

can also correspond to melting of a eutectics between the main compound and impurities followed by shifted melting of the main compound (*i.e.*, solidus and liquidus temperature for the mixture). Hence, all assignments of DSC peaks should be justified by providing the basis of the conclusion (*e.g.*, X-ray diffraction experiments or comparison to a published phase diagram).

Principle 5 (Auxiliary (calibration) data should be identified and provided)

Case 5a (reporting calibration conditions): As previously mentioned, calibration points and experimental conditions conducted during the calibration and measurements and/or well referenced citations to a previous work for each physical measurement (density, viscosity, thermal and electrical conductivity, *etc.*) should be clearly given. For example, in the case of the density measurements conducted by using a vibrating tube, it should be specified whether the impact of the viscosity of the sample on experimental data was analyzed. Even more importantly, if the measured density is outside the calibration range and extrapolation of the calibration is used, this fact should be explicitly noted, while the use of a standard with a higher density than those measured (or used during the calibration) would be preferable.

Principle 6 (Necessary details of experimental methods or computation procedures should be given)

Case 6a (describing degassing): In electrochemical experiments (such as the potentiometric measurements mentioned in Case 2u), it may be important to know whether (and how) air ingress, leading to residual oxygen or, occasionally, moisture or carbon dioxide, was prevented in the measurements. Deaeration procedures should be described in detail because they are essential for the interpretability of the measurements. Description of degassing procedures is also important for many other experiments: *e.g.*, densities, vapor-pressures, VLE, and PVT experiments.

Case 6b (sampling description): If sampling is needed to measure a property, the details of the sampling location, the procedure of sample withdrawal, sample handling, and analysis steps need to be described. In particular, sampling from high-pressure fluid mixtures affects not only the sample itself but also the remaining system due to the inevitable pressure reduction [37]. Details of the sampling procedure, including countermeasures against the pressure reduction, are important for evaluating effects on accuracy.

Case 6c (clear model description): Adequate models can be incorrectly used. For instance, the correlation of equilibrium data expressed in mass fractions with models derived for mole fractions leads to wrong correlation parameters [38]. Hence, proper description of the applied calculation approach is essential for unambiguous use of the model.

Case 6d (calibration of DSC): According to the recommendations developed by the working group “Calibration of Scanning Calorimeters” of the German Society of Thermal Analysis (GEFTA) [39-43] and adopted by the International Confederation for Thermal Analysis and Calorimetry (ICTAC) in 2000, extrapolated onset temperatures should be used as melting temperatures in calibration and main experiments of DSC. However, some authors instead use peak temperatures for specific reasons; doing this requires explicit description of the applied calibration procedures.

Principle 7 (Uncertainty in each measured value should be reported and justified)

Case 7a (reporting uncertainty type): The uncertainty type (standard deviation of single measurement or mean, standard or expanded uncertainty, *etc.*) reported for all properties and state variables should be provided. According to the Guide to the Expression of Uncertainty in Measurement (GUM) [5], the term “uncertainty” is a concept that has no numerical value, unless used with a modifier, such as “standard uncertainty” or “expanded uncertainty with 0.95 level of confidence (coverage factor $k = 2$).” Uncertainties should be reported as standard uncertainties u (with the default level of confidence of 0.68) or expanded uncertainties U (with a typical level of confidence of 0.95, coverage factor $k = 2$). In the case of expanded uncertainties, the level of confidence and/or the coverage factor should be given explicitly. The level of confidence should be also provided for all uncertainties reported as confidence intervals (*e.g.*, “(0.100 ± 0.005) MPa, where expanded uncertainty with 0.95 level of confidence (coverage factor $k = 2$) is reported”). In the case of standard

uncertainties, the level of confidence can be omitted, since it has only one default value. Uncertainties can be reported as absolute uncertainties (symbols – u , U) and relative uncertainties (u_r , U_r). Readers are referred to GUM [5] for more details.

Case 7b (reporting uncertainty type): Uncertainties should be provided for all properties and state variables, *e.g.*, for pressure (including “atmospheric-pressure” experiments, when pressure was not controlled), temperature, and compositions. If amounts of different components in a mixture were determined separately by independent methods, each corresponding component content uncertainty should be given.

Case 7c (purity and calibration contributions to uncertainty): The effect of sample purity and/or calibration on the reported uncertainties should be considered. One example is uncertainty for phase-transition parameters from DSC experiments: the uncertainties from temperature- and sensitivity-calibrations should be taken into account in addition to repeatability. Another example is the effect of purity on density results; in [44], the relative standard uncertainty for densities of a sample with 0.99 mass-fraction purity was estimated to be $u_r(\rho) = 0.001$ or 0.1%. In addition, it was recently shown that a typical two-point calibration of a vibrating-tube densimeter with water and air will not give a standard uncertainty better than $0.15 \text{ kg}\cdot\text{m}^{-3}$ [45].

Principle 8 (Importing reported data into analysis software should be easy and straightforward)

Case 8a (table format): Well-documented tables are the best presentation of property data. The following table is an example of well-presented data. The table below was extracted from [46] and modified for increased clarity. This table exemplifies several of the principles developed in this Technical Report:

- 1) Even though a detailed derivation of uncertainties was presented in the text of [46], the uncertainty values were repeated in the data tables. Since uncertainty values were the same for each state variable and each property, they were summarized in a footnote.
- 2) The pressure was reported even though the data being reported have a weak pressure dependence.
- 3) The units of all values were reported in the table or table heading. The calculus of quantities is obeyed.
- 4) All mass-fraction compositions total to a value of 1.0.
- 5) The data whose source was not from the current work was clearly referenced, *i.e.*, the values taken from [47].
- 6) All state and non-state variables, *e.g.*, the wavelength of applied light, were reported.

Table 2: Experimental binodal-curve data for water (1) + HMF (2) + 1-pentanol (3): Cloud-point compositions expressed as mass fractions of component i (w_i), and corresponding refractive-index values (n_D) at $T = 298.2 \text{ K}$ and $P = 0.100 \text{ MPa}$.^{a,b}

w_1	w_2	w_3	n_D
0.1014	0.0000	0.8986	1.4019
0.1080 ^c	0.0000 ^c	0.8920 ^c	1.4021 ^c
0.2504	0.3372	0.4124	1.4317
0.5088	0.3416	0.1496	1.4137
0.8734	0.0965	0.0301	1.3542
0.9800 ^c	0.0000 ^c	0.0200 ^c	1.3353 ^c

^a Type B standard uncertainties u for temperature and pressure are $u(T) = 0.1 \text{ K}$ and $u(P) = 0.005 \text{ MPa}$. Combined standard uncertainties for mass-fraction compositions and refractive index are $u_c(w_i) = 0.0075$ and $u_c(n_D) = 0.0020$. ^b Refractive-index values are measured at 589 nm . ^c Values taken from [47].

In addition to the data presented in Table 1, Ref. [46] also included:

- a) The full chemical name of each component as well as any abbreviations, *e.g.*, HMF.
- b) The source of compounds, purification methods used, and final purity.
- c) The details of the experimental procedure.
- d) The refractive indices of each pure component and comparisons with previously measured values.
- e) The refractive indices of all binary endpoint mixtures and comparisons with previously measured values.

Principle 9 (Complex mathematical equations should be provided in a machine-readable form)

Case 9a (units in equations): Good practice for equations should be to define dimensionless quantities, so that all parameters in a model are also dimensionless. Those quantities, such as reduced temperature, should be defined separately along with the scaling factor given with its unit.

References

- [1] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons. *Scientific Data* **3**, 160028 (2016); <http://dx.doi.org/10.1038/sdata.2016.18>
- [2] R. D. Chirico, T. W. de Loos, J. Gmehling, A. R. H. Goodwin, S. Gupta, W. M. Haynes, K. N. Marsh, V. Rives, J. D. Olson, C. Spencer, J. F. Brennecke, J. P. M. Trusler. *Pure Appl. Chem.* **84**, 1785 (2012); <http://dx.doi.org/10.1351/PAC-REC-11-05-02>
- [3] L. Gardossi, P. B. Poulsen, A. Ballesteros, K. Hult, V. K. Švedas, Đ. Vasić-Rački, G. Carrea, A. Magnusson, A. Schmid, R. Wohlgemuth, P. J. Halling. *Trends Biotechnol.* **28**, 171 (2010); <http://dx.doi.org/10.1016/j.tibtech.2010.01.001>
- [4] *Guidelines and Examples for the Journal of Chemical and Engineering Data*: <https://trc.nist.gov/JournalCoop/JCED/> (accessed on 11 February 2020).
- [5] Joint Committee for Guides in Metrology. *JCGM 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in measurement*, BIPM, Sèvres, France (2008); https://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf (accessed on 22 November 2019).
- [6] H. A. Favre, W. H. Powell. *Nomenclature of Organic Chemistry: IUPAC Recommendations and Preferred Names 2013, IUPAC Blue Book*, IUPAC & RSC Publishing, Cambridge, U.K. (2014).
- [7] S. R. Heller, A. McNaught, I. Pletnev, S. Stein, D. Tchekhovskoi. *J. Cheminform.* **7**, 23 (2015); <http://dx.doi.org/10.1186/s13321-015-0068-4>
- [8] E. R. Cohen, T. Cvitas, J. G. Frey, B. Holmström, K. Kuchitsu, R. Marquardt, I. Mills, F. Pavese, M. Quack, J. Stohner, H. L. Strauss, M. Takami, A. J. Thor. *Quantities, Units and Symbols in Physical Chemistry, IUPAC Green Book*, Third Edition, Second Printing, IUPAC & RSC Publishing, Cambridge, U.K. (2008).
- [9] R. Dohrn, J. M. S. Fonseca, S. Peper. *Annu. Rev. Chem. Biomol.* **3**, 343 (2012); <http://dx.doi.org/10.1146/annurev-chembioeng-062011-081008>
- [10] S. Raeissi, J. F. Brennecke. *J. Chem. Eng. Data* **63**, 1849 (2018); <http://dx.doi.org/10.1021/acs.jced.8b00441>
- [11] P. De Bièvre, R. Dybkaer, A. Fajgelj, D. B. Hibbert. *Pure Appl. Chem.* **83**, 1873 (2011); <http://dx.doi.org/10.1351/PAC-REP-07-09-39>
- [12] *Experimental Thermodynamics Book Series. Volume X: Non-equilibrium Thermodynamics with Applications*, D. Bedeaux, S. Kjelstrup, J. Sengers (Eds.), IUPAC & RSC Publishing, Cambridge, U.K. (2016).
- [13] B. N. Taylor, C. E. Kuyatt. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Technical Note 1297*, National Institute of Standards and Technology, Gaithersburg, U.S.A. (1994); <https://emtoolbox.nist.gov/Publications/NISTTechnicalNote1297s.pdf> (accessed on 22 November 2019).
- [14] Joint Committee for Guides in Metrology. *JCGM 200:2012, International vocabulary of metrology – Basic and general concepts and associated terms (VIM)*, Third Edition, BIPM, Sèvres, France (2012); https://www.bipm.org/utis/common/documents/jcgm/JCGM_200_2012.pdf (accessed on 03 August 2020)
- [15] *ThermoLit, NIST Standard Reference Database #171*: <https://trc.nist.gov/thermolit/> (accessed on 22 November 2019); <http://dx.doi.org/10.18434/T4VS35>
- [16] *H.R. 1770 — 115th Congress (2017-2018). Open, Public, Electronic, and Necessary Government Data Act*. <https://www.congress.gov/bill/115th-congress/house-bill/1770/text> (accessed on 11 February 2020).
- [17] I. H. Bell, A. Jäger. *J. Res. Natl. Inst. Stan.* **121**, 238 (2016); <http://dx.doi.org/10.6028/jres.121.011>
- [18] *Submission Checklist for the Journal of Chemical and Engineering Data*: <https://trc.nist.gov/JournalCoop/JCED/checklist.html> (accessed on 22 November 2019).
- [19] *Journal of Chemical Thermodynamics Submission Checklist for Reporting New Experimental Data*: https://www.elsevier.com/_data/promis_misc/JCT_checklist%2007%20Dec%202017-updated.docx (accessed on 22 November 2019).
- [20] M. M. Abbott, H. C. Van Ness. *Fluid Phase Equilib.* **1**, 3 (1977); [http://dx.doi.org/10.1016/0378-3812\(77\)80021-6](http://dx.doi.org/10.1016/0378-3812(77)80021-6)
- [21] *IUPAC Recommended Reference Materials for the Realization of Physicochemical Properties*, K. N. Marsh (Ed.), Blackwell Scientific, Boston, U.S.A. (1987).
- [22] R. Sabbah, A. Xu-wu, J. S. Chickos, M. L. Planas Leitão, M. V. Roux, L. A. Torres. *Thermochim. Acta* **331**, 93 (1999); [http://dx.doi.org/10.1016/S0040-6031\(99\)00009-X](http://dx.doi.org/10.1016/S0040-6031(99)00009-X)

- [23] A. Bazyleva, W. E. Acree Jr., R. D. Chirico, T. W. de Loos, V. Diky, G. T. Hefter, J. Jacquemin, J. W. Kang, M. C. F. Magalhães, J. W. Magee, J. P. O'Connell, J. M. Shaw, J. P. M. Trusler, R. Weir. *IUPAC Project 2011-037-2-100: Recommended Reference Materials for Phase Equilibrium Studies*; https://iupac.org/projects/project-details/?project_nr=2011-037-2-100 (accessed on 22 November 2019).
- [24] M. J. Assael, A. E. Kalyva, S. A. Monogenidou, M. L. Huber, R. A. Perkins, D. G. Friend, E. F. May. *J. Phys. Chem. Ref. Data* **47**, 021501 (2018); <http://dx.doi.org/10.1063/1.5036625>
- [25] M. Steiger, W. Voigt. *J. Solution Chem.* **48**, 1009 (2019); <http://dx.doi.org/10.1007/s10953-018-0794-0>
- [26] J. M. S. Fonseca, R. Dohrn, S. Peper. *Fluid Phase Equilib.* **300**, 1 (2011); <http://dx.doi.org/10.1016/j.fluid.2010.09.017>
- [27] C. F. Jones, A. Mayne. *Application and Limitations of the Karl Fischer Technique of Moisture Analysis in Electrical Plant Insulation (paper 7241)*. Proceedings of the 4th International Conference on Properties and Applications of Dielectric Materials, July 3-8, 1994, Brisbane, Australia, 895 (1994); <http://dx.doi.org/10.1109/ICPADM.1994.414156>
- [28] J. A. Labarta, M. d. M. Olaya, A. Marcilla. *Graphical User Interface (GUI) for the analysis of Gibbs Energy surfaces, including LL tie-lines and Hessian matrix (Software)*: <http://hdl.handle.net/10045/51725> (accessed on 22 November 2019).
- [29] D. K. Nordstrom. *Appl. Geochem.* **27**, 1899 (2012); <http://dx.doi.org/10.1016/j.apgeochem.2012.07.007>
- [30] T. J. Cole. *Arch. Dis. Child.* **100**, 608 (2015); <http://dx.doi.org/10.1136/archdischild-2014-307149>
- [31] P. H. van Konynenburg, R. L. Scott. *Philos. T. R. Soc. A* **298** (1442), 495 (1980); <http://dx.doi.org/10.1098/rsta.1980.0266>
- [32] R. Privat, J.-N. Jaubert. *Chem. Eng. Res. Des.* **91**, 1807 (2013); <http://dx.doi.org/10.1016/j.cherd.2013.06.026>
- [33] F. N. Rhines. *Phase diagrams in metallurgy, their development and application*, McGraw-Hill Book Company, New York, U.S.A. (1956).
- [34] *Author Guidelines of the Journal of Chemical and Engineering Data*: http://pubs.acs.org/paragonplus/submission/jceaax/jceaax_authguide.pdf (accessed on 22 November 2019).
- [35] F. E. Young. *J. Phys. Chem.* **61**, 616 (1957); <http://dx.doi.org/10.1021/j150551a023>
- [36] M. Dzida, E. Zorębski, M. Zorębski, M. Żarska, M. Geppert-Rybczyńska, M. Chorażewski, J. Jacquemin, I. Cibulka. *Chem. Rev.* **117**, 3883 (2017); <http://dx.doi.org/10.1021/acs.chemrev.5b00733>
- [37] S. Peper, R. Dohrn. *J. Supercrit. Fluid.* **66**, 2 (2012); <http://dx.doi.org/10.1016/j.supflu.2011.09.021>
- [38] A. Marcilla, J. A. Reyes-Labarta, M. M. Olaya. *Fluid Phase Equilib.* **433**, 243 (2017); <http://dx.doi.org/10.1016/j.fluid.2016.11.009>
- [39] G. W. H. Höhne, H. K. Cammenga, W. Eysel, E. Gmelin, W. Hemminger. *Thermochim. Acta* **160**, 1 (1990); [http://dx.doi.org/10.1016/0040-6031\(90\)80235-Q](http://dx.doi.org/10.1016/0040-6031(90)80235-Q)
- [40] H. K. Cammenga, W. Eysel, E. Gmelin, W. Hemminger, G. W. H. Höhne, S. M. Sarge. *Thermochim. Acta* **219**, 333 (1993); [http://dx.doi.org/10.1016/0040-6031\(93\)80510-H](http://dx.doi.org/10.1016/0040-6031(93)80510-H)
- [41] S. M. Sarge, E. Gmelin, G. W. H. Höhne, H. K. Cammenga, W. Hemminger, W. Eysel. *Thermochim. Acta* **247**, 129 (1994); [http://dx.doi.org/10.1016/0040-6031\(94\)80118-5](http://dx.doi.org/10.1016/0040-6031(94)80118-5)
- [42] E. Gmelin, S. M. Sarge. *Pure Appl. Chem.* **67**, 1789 (1995); <http://dx.doi.org/10.1351/pac199567111789>
- [43] S. M. Sarge, W. Hemminger, E. Gmelin, G. W. H. Höhne, H. K. Cammenga, W. Eysel. *J. Therm. Anal.* **49**, 1125 (1997); <http://dx.doi.org/10.1007/BF01996802>
- [44] R. D. Chirico, M. Frenkel, J. W. Magee, V. Diky, C. D. Muzny, A. F. Kazakov, K. Kroenlein, I. Abdulagatov, G. R. Hardin, W. E. Acree Jr., J. F. Brenneke, P. L. Brown, P. T. Cummings, T. W. de Loos, D. G. Friend, A. R. H. Goodwin, L. D. Hansen, W. M. Haynes, N. Koga, A. Mandelis, K. N. Marsh, P. M. Mathias, C. McCabe, J. P. O'Connell, A. Pádua, V. Rives, C. Schick, J. P. M. Trusler, S. Vyazovkin, R. D. Weir, J. Wu. *J. Chem. Eng. Data* **58**, 2699 (2013); <http://dx.doi.org/10.1021/je400569s>
- [45] D. Fitzgerald. *Technical Assessment of the Anton Paar DMA5000 density meter*, H&D Fitzgerald Ltd., St Asaph, U.K. (2000); https://www.density.co.uk/documents/review_of_5000.pdf (accessed on 22 November 2019).
- [46] K. V. M. Cavalcanti, L. M. Follegatti-Romero, I. Dalmolin, L. A. Follegatti-Romero. *J. Chem. Thermodyn.* **138**, 59 (2019); <http://dx.doi.org/10.1016/j.jct.2019.06.010>
- [47] L. Domingues, P. A. Cussolin, J. L. da Silva Jr., L. H. de Oliveira, M. Aznar. *Fluid Phase Equilib.* **354**, 12 (2013); <http://dx.doi.org/10.1016/j.fluid.2013.06.007>