

New Global Communication Process in Thermodynamics: Impact on Quality of Published Experimental Data

M. Frenkel,^{*,†} R. D. Chirico,[†] V. Diky,[†] C. Muzny,[†] Q. Dong,[†] K. N. Marsh,[‡] J. H. Dymond,[§] W. A. Wakeham,^{||} S. E. Stein,[‡] E. Königsberger,[#] A. R. H. Goodwin,[○] U. W. Magee,^{†§} M. Thijssen,[▽] W. M. Haynes,[†] S. Watanasiri,[◆] M. Satyro,[◇] M. Schmidt,[△] A. I. Johns,^{*} and G. R. Hardin[†]

Physical and Chemical Properties Division, 838, National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305-3328,[†] Department of Chemical and Process Engineering, University of Canterbury, Private Bag 4800, Christchurch, New Zealand,[‡] Chemistry Department, University of Glasgow, Glasgow G12 8QQ, U.K.,[§] School of Engineering Sciences, University of Southampton, Highfield, Southampton SO17 1BJ, U.K.,^{||} Physical and Chemical Properties Division, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, Maryland 20899,[‡] School of Engineering Science, Murdoch University, Murdoch, WA 6150 Australia,[‡] Schlumberger Technology Corporation, 125 Industrial Blvd., Sugar Land, Texas 77478,[○] Physical and Theoretical Chemistry, Elsevier BV, Radarweg 29, 1043 NX Amsterdam, The Netherlands,[§] Aspen Technology, Inc., Ten Canal Park, Cambridge, Massachusetts 02141-2201,[▽] Virtual Materials Group, Inc., 657 Hawkside Mews NW, Calgary, Alberta, Canada T3G 3S1,[◇] Software Development, FIZ CHEMIE Berlin, Franklinstrasse 11, D-10587 Berlin, Germany, and[○] Oil, Gas and Chemicals Division, TUV NEL Ltd., East Kilbride, Glasgow G75 0QU, U.K.

Received May 19, 2006

Thermodynamic data are a key resource in the search for new relationships between properties of chemical systems that constitutes the basis of the scientific discovery process. In addition, thermodynamic information is critical for development and improvement of all chemical process technologies. Historically, peer-reviewed journals are the major source of this information obtained by experimental measurement or prediction. Technological advances in measurement science have propelled enormous growth in the scale of published thermodynamic data (almost doubling every 10 years). This expansion has created new challenges in data validation at all stages of the data delivery process. Despite the peer-review process, problems in data validation have led, in many instances, to publication of data that are grossly erroneous and, at times, inconsistent with the fundamental laws of nature. This article describes a new global data communication process in thermodynamics and its impact in addressing these challenges as well as in streamlining the delivery of the thermodynamic data from “data producers” to “data users”. We believe that the prolific growth of scientific data in numerous and diverse fields outside thermodynamics, together with the demonstrated effectiveness and versatility of the process described in this article, will foster development of such processes in other scientific fields.

1. INTRODUCTION

Thermodynamic data are a key resource for development and improvement of all chemical process technologies. Incompleteness or poor data quality often lead to erroneous equipment selections (pumps, reactors, heat exchangers) which preclude further improvements in process efficiency and safety at the pilot-scale and production-scale stages, resulting in undesirable and possibly enormous economic losses. Often, a lack of thermodynamic information makes it impossible to simulate new chemical processes, neces-

sitating numerous empirical and expensive trial-and-error iterations for process optimization. This commonly results in significant delays and costs without any assurance of finding the true optimal conditions for the process. This situation is quite typical, especially within rapidly developing industries such as pharmaceuticals, specialty chemicals, biotechnology, and defense. In addition to the highly practical field of process development, high quality thermodynamic property data are frequently essential prerequisites for testing theories and searching for new relationships between properties of chemical systems, which constitute the basis of the scientific discovery process.

Historically, peer-reviewed journals are the major source of thermodynamic information obtained either by experimental measurement or prediction. The enormous growth of published thermophysical and thermochemical data (almost doubling every 10 years due to technological advances in measurement science) has created new challenges for data validation at all stages of the data delivery process. Despite the peer-review process, this has led to many instances of

* Corresponding author e-mail: frenkel@boulder.nist.gov.

[†] National Institute of Standards and Technology, Boulder, CO.

[‡] University of Canterbury.

[§] University of Glasgow.

^{||} University of Southampton.

[‡] National Institute of Standards and Technology, Gaithersburg, MD.

[#] Murdoch University.

[○] Schlumberger Technology Corporation.

[▽] Physical and Theoretical Chemistry, Elsevier BV.

[◆] Aspen Technology, Inc.

[◇] Virtual Materials Group, Inc.

[△] Software Development, FIZ CHEMIE Berlin.

^{*} Oil, Gas and Chemicals Division, TUV NEL Ltd.

published data that are either grossly erroneous or even inconsistent with the fundamental laws of nature. This trend also magnifies the problem, among journals publishing scientific and technical data, of an absence of consensus concerning the requirements placed on authors for characterization of overall data quality and reliability. Indeed, even in reporting experimental property data with a relatively simple metadata infrastructure, such as critical temperature, comprehensive uncertainty analyses currently remain rare, particularly with regard to consideration of contributions arising from sample impurities.¹ Additional challenges and opportunities are provided by the availability of published data via the Internet, which creates unprecedented opportunities for development of innovative information delivery processes from "data producers" to "data users".² The present paper describes how these two independent sets of challenges and opportunities are addressed to communicate, capture, exchange, and validate the large-scale flow of thermodynamic property data information. The development of a new global data delivery process is described in this article. The delivery process is based on three major components: the recently established XML-based IUPAC standard for thermodynamic data communications (ThermoML³), software infrastructure^{4,5} to support the process, and a thermodynamic data expert system (ThermoData Engine^{6,7}) that allows generation of recommended data on-demand through implementation of the dynamic data evaluation concept.⁸⁻¹⁰ This new data delivery process has led to a unique cooperation between five major journals in the field, national and international standardization organizations, and leading chemical process design companies.

2. STANDARDIZING THERMODYNAMIC DATA COMMUNICATIONS

A brief review of the standardization efforts for thermodynamic data communications was recently published.¹⁰ Efforts to develop a standard for thermophysical and thermochemical property data exchange¹¹ were first initiated in the early 1980s, reflecting a new trend in data collection through design of electronic databases, which became possible due to the rapid development of computer technology. However, through the end of the 1990s, these projects did not result in the creation of an internationally established standard primarily because of two reasons: inadequacy of the available technology and structural challenges related to the high level of diversity of thermodynamic property data.

In 2002, the International Union of Pure and Applied Chemistry (IUPAC) approved Project 2002-055-3-024, "XML-based IUPAC Standard for Experimental and Critically Evaluated Thermodynamic Property Data Storage and Capture," and established a Task Group¹² to create standardized mechanisms for thermodynamic data communications with XML (Extensible Markup Language) technology. XML technology,¹³ fully developed within the last 5 years, provides significant advantages for the development of standards for data exchange, such as its native interoperability based on ASCII code, its modular nature, and transparent readability by both humans and computers. From a practical standpoint, it is also critical that this technology is currently supported by both the software and hardware industries. As the result of this project, a new IUPAC standard (ThermoML) for

thermodynamic data storage and exchange has been developed.^{3,14}

The ThermoML structure represents a balanced combination of hierarchical and relational elements. The schema explicitly incorporates structural elements related to basic principles of phenomenological thermodynamics: thermodynamic, thermochemical, and transport properties, state variables, system constraints, phases, and units. Meta- and numerical data records are grouped into nested blocks of information corresponding to data sets. Metadata records precede numerical data information, providing a robust foundation for generating header records for any relational database where ThermoML-formatted files might be incorporated. The structural features of the ThermoML metadata records ensure unambiguous interpretation of numerical data and allow data-quality control based on the Gibbs phase rule. Moreover, some detailed information included in the metadata records could serve as a background for independent assessment of uncertainties, which could be propagated into uncertainties of physical parameters for reaction streams, and, consequently, provide an opportunity for numerical characterization of the quality of a chemical process design.¹⁵

ThermoML covers essentially all experimentally determined thermodynamic and transport property data (more than 120 properties) for pure compounds, multicomponent mixtures, and chemical reactions (including change-of-state and equilibrium).¹⁶ ThermoML allows storage and exchange of property data with full allowance for data traceability and interpretation. Comprehensive specification of the data source (bibliographic information), method of property generation (experimental, predicted, critically evaluated), and multiple uncertainty assessments are included. Representations of all quantities related to the expression of uncertainty in ThermoML¹⁷ conform to the *Guide to the Expression of Uncertainty in Measurement* (GUM).¹⁸

The modular nature of XML was exploited to describe regression equations used for analytical representation of a variety of thermodynamic properties, including complex equations of state. That was achieved by the use of the published MathML¹⁹ formats in conjunction with ThermoML.²⁰ The Mathematical Markup Language (MathML) is a low-level specification for describing mathematics as a basis for machine-to-machine communication in terms of both content (mathematical meaning) and presentation (format). Among other X-markup languages, CML (XML for chemistry)²¹ and MatML (XML for primarily mechanical properties of materials)²² are most closely related to ThermoML.

ThermoML consists of four major blocks:³ *Citation* (description of the source of the data); *Compound* (characterization of the chemical system); *PureOrMixtureData* (meta- and numerical property data for a pure compound or multicomponent mixture); and *ReactionData* (meta- and numerical property data for a chemical reaction with thermodynamic state change or in a state of chemical equilibrium). The compound description within the *Compound* block is linked to a description of the sample that includes its initial source and purity, purification methods used, and final purity, with specification of the method(s) of purity determination.

3. SOFTWARE INFRASTRUCTURE

For reasons discussed earlier, there has been a great demand for the establishment of an efficient global data delivery processes. Previous to the work described here, such a process did not exist in the field of thermodynamics. In fact, there are only two well-known processes of this nature outside the field of thermodynamics: submission and retrieval of protein structures from the Protein Data Bank (PDB)²³ and submission and retrieval of crystal structures for smaller molecules from the Cambridge Structural Database (CSD).²⁴ Establishing a global data delivery process for thermodynamic properties is substantially more complex than that for the PDB and CSD because of the necessity to communicate information related to the numerous (>100) thermophysical and thermochemical properties commonly reported. Moreover, communicating such data is further complicated by the extensive system of thermodynamic metadata (variables, constraints, phases, methods, uncertainties) required. This complexity necessitated development of a software infrastructure to support the global delivery process for thermodynamic data. Key pieces of the software infrastructure are described in the following paragraphs: Guided Data Capture software, the NIST SOURCE data archival system, and the NIST ThermoData Engine expert system.

Guided Data Capture (GDC) software was developed^{25,26} and serves as a data-capture expert by guiding extraction of information from the literature, ensuring the completeness of the information extracted, validating the information through data definition, range checks, etc. A key feature of the GDC software is the capture of information in close accord with customary original-document formats. The compiler's main interactions with GDC involve a navigation tree that provides a visual representation in accord with the hierarchical structure of a typical source document. Numerical values are not shown explicitly in the tree but may be accessed through the property-specification nodes. Enumerated field values (journal titles, compound identifiers, properties, units, phases, experimental methods, etc.) are stored in a local database that is part of the GDC software. Selection of field values by the data compiler is achieved through predefined lists, which prevents many common errors. Most numerical values can be captured through electronic means (PDF files, spreadsheets, etc.) and rarely require manual input. All other input is accomplished through predefined menus, check boxes, or other controlled selection processes. GDC also provides simple "one-click" graphical representation of the numerical data. This is a powerful tool for detection of typographical errors with data sets for thermophysical properties that can include data points numbering in the thousands. GDC can be used to output files in ThermoML format.

The NIST SOURCE data archival system²⁷ created and maintained by the Thermodynamics Research Center (TRC) is a large general-purpose relational data storage facility of experimental values of thermodynamic, thermochemical, and transport properties of pure compounds, mixtures of well-defined composition, and chemical reactions. It includes numerical values for an enormous variety of properties for systems involving the wide spectrum of possible phase scenarios. It is critically important that the SOURCE data

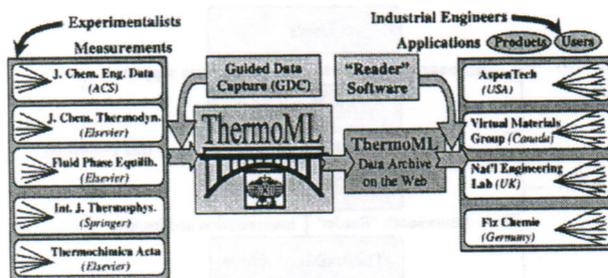


Figure 1. Global experimental thermodynamic data delivery process.³⁴

system contains estimated combined expanded uncertainties for all of the numerical data stored. This allows, in principle, determination of the quality of critically evaluated data generated based on the original experimental data collected in the SOURCE data system.

The new concept of dynamic data evaluation was developed at NIST/TRC.⁸⁻¹⁰ This concept requires the development of large electronic databases capable of storing essentially all relevant experimental data known to date with detailed descriptions of relevant metadata and uncertainties. The combination of these electronic databases with expert-system software, designed to automatically generate recommended data based on available experimental data, leads to the ability to produce critically evaluated data dynamically on-demand. This concept contrasts sharply with static critical data evaluation, which must be initiated far in advance of need. The dynamic data evaluation process dramatically reduces the effort and costs associated with anticipating future needs and keeping static evaluations current. Recently, the concept of dynamic data evaluation has been implemented for the first time in NIST ThermoData Engine.^{6,7}

4. ESTABLISHING A GLOBAL DATA DELIVERY PROCESS

The combination of an established international standard for thermodynamic data communications with the software infrastructure supporting it allows creation of global data delivery processes. Two major types of these processes are illustrated here.

Figure 1 illustrates a unique experimental data communication process, established in cooperation between major journals in the field, leading scientific publishers, national and international standardization organizations, and commercial enterprises, to deliver data from "data producers" (thermodynamicists reporting results of measurements of thermophysical and thermochemical property data in journals in the field) to "data users" (chemical engineers who use engineering software applications, including chemical process design). The GDC software is a key support element for data submission, and ThermoML serves to ensure interoperability for propagation of the data across different platforms. The Web-based "ThermoML Archive"²⁸ currently contains—in ThermoML format—data published by five major journals in the field: *Journal of Chemical and Engineering Data*,²⁹ *The Journal of Chemical Thermodynamics*,³⁰ *Fluid Phase Equilibria*,³¹ *Thermochemica Acta*,³² and *International Journal of Thermophysics*.³³ It is important to emphasize that authors are immediate participants in this process, ensuring

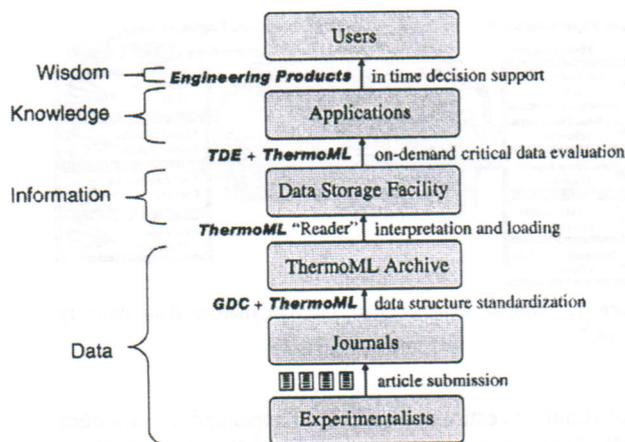


Figure 2. Combining on-demand critical data evaluation with global data delivery.

that the data become available for use essentially simultaneous with publication of the article. The ThermoML Archive has a structure similar to that of the electronic version of the original journal articles providing links (recently, in some instances, reciprocal) to the original sources. ThermoML software “readers” have been developed by a number of organizations³⁴ (shown at the right side of Figure 1) in cooperation with NIST/TRC to transfer data from the ThermoML format to customized formats suitable for application software and databases. This process is supported by standardization efforts with the participation of industry (DIPPR—Design Institute for Physical Properties of the American Institute of Chemical Engineers), IUPAC, and the International Association of Chemical Thermodynamics. Although the “horizontal” data delivery process between “data producers” and “data users” is emphasized in Figure 1, it is also clear that a great variety of “vertical” processes can be established to communicate data within groups of experimentalists, journals, industrial organizations, products, and users.

Figure 2 illustrates a recently implemented, more extended, data delivery process combining Web-based interoperable dissemination of data from the “producers” to “users” with conversion of original experimental data (at the bottom of Figure 2) to the level of *knowledge* and *wisdom* at the stage

of their use in applications (top of Figure 2). In this process, the experimental *data* are collected in the Web-based “ThermoML Archive” and then downloaded into a comprehensive data storage facility (such as the SOURCE data archival system) transforming the numerical *data* into *information*, where each experimental datum is characterized with a comprehensive measure of data quality, i.e., the combined expanded uncertainty. Further, this *information* is critically evaluated with expert-system software (NIST ThermoData Engine; TDE) with enforcement of internal consistency for related properties, thus achieving the status of *knowledge* through generation of on-demand recommended data for use in applications. This *knowledge* allows design of software applications for in-time decisions for chemical process analysis, thus attaining the status of *wisdom*. All communications between the various components of this process are supported by the Guided Data Capture software and the ThermoML structured schema.

5. IMPACT ON QUALITY OF PUBLISHED DATA

The major purpose of development of the new global thermodynamic data communication process was to streamline delivery of data and information from “producers” to “users”. However, this process has proved extremely powerful in enhancing the quality of published data. Figure 3 illustrates the communication flow within the NIST/TRC data facility. After completion of the peer-review process, this flow structure allows for (a) checking of experimental data to be published for integrity (GDC communication line in Figure 3) and (b) consistency with the recommended data evaluated on the basis of the complete body of knowledge available to date (TDE communication line in Figure 3). Consistency is checked for the particular property as well as for those linked through thermodynamic identities and correlations. Below, we illustrate three examples of different types of such testing.

The *GDC communication line* (Figure 3) provides an opportunity to perform initial integrity analysis of the new, soon-to-be-published experimental data. Figure 4 shows analysis with GDC of the originally published mass (specific) density data as a function of temperature for mixtures of chlorobenzene and dimethyl carbonate. The published data³⁵

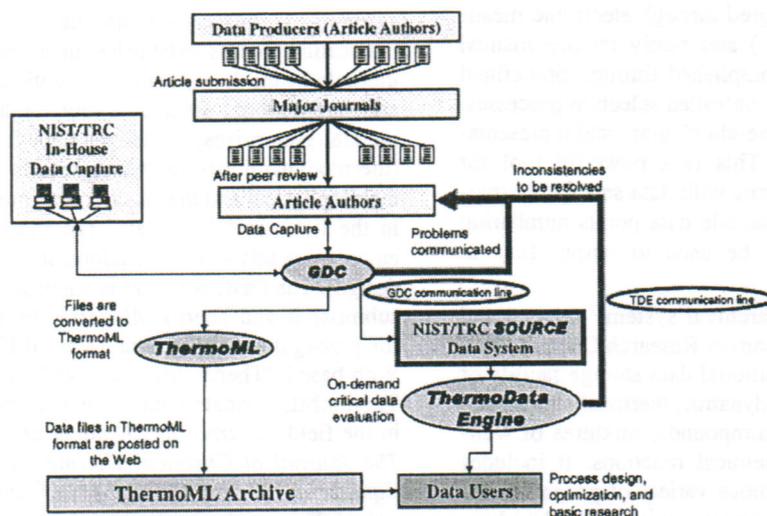


Figure 3. NIST/TRC data facility data flow diagram.

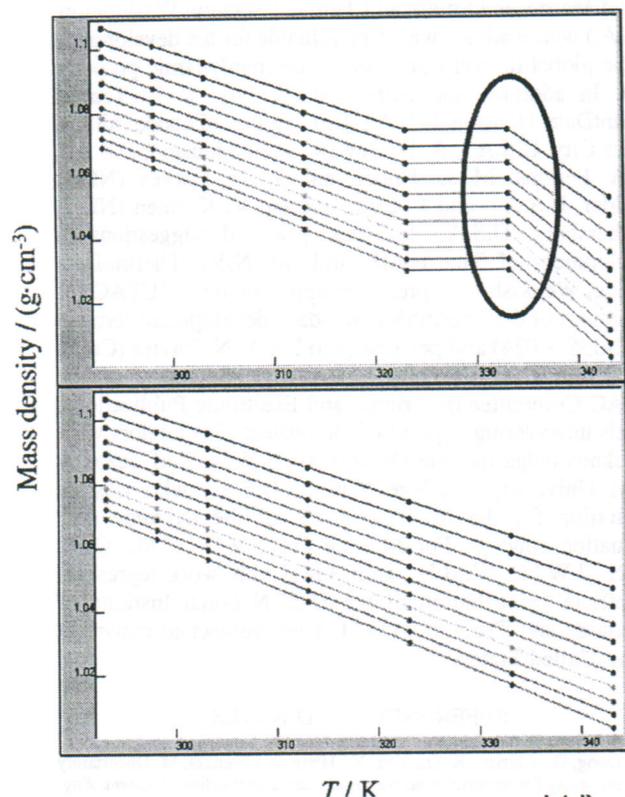


Figure 4. Mass (specific) density as a function of temperature for mixtures of chlorobenzene and dimethyl carbonate. The lines represent results for various mixture compositions. The figures show the reported experimental data before (upper)³⁵ and after (lower)³⁶ correction. The circle indicates the erroneous values in the upper figure. The graphs were created with the GDC software.⁵

showed anomalous behavior (Figure 4, upper figure, circled). The data were later corrected³⁶ (Figure 4, lower figure), once results of the GDC analysis were communicated to the authors.

Figure 5 shows the results of the GDC analysis of reported experimental boiling temperature data for the mixture (tetrahydrofuran + tetrachloroethene) as a function of the mixture composition. The data published originally³⁷ indicated anomalous behavior in the region with mole fraction of tetrahydrofuran greater than 0.5 (Figure 5, upper figure). The erroneous nature of the reported data is further revealed through a comparison of the pure compound boiling temperatures T_b for the mixture components with those critically evaluated by the expert system software TDE (see, *TDE communication line*, Figure 3). Indeed, the reported boiling temperature for tetrahydrofuran at 95.8 kPa [$T_b = (311.45 \pm 0.1)$ K] is significantly lower than that critically evaluated with TDE [$T_b = (337.38 \pm 0.03)$ K]. The originally published data were later corrected³⁸ after the results of the combined GDC and TDE analysis were communicated to the authors (Figure 5, lower figure).

Finally, Figure 6 illustrates the TDE analysis of published experimental density data in the liquid phase for tetrahydrofuran as a function of temperature and pressure. Preliminary GDC analysis did not detect any anomalies in the originally published data.³⁹ However, comparison of the published values with those along the saturation line obtained with on-demand critical data evaluation by TDE showed that the reported single-phase data crossed the saturation line into

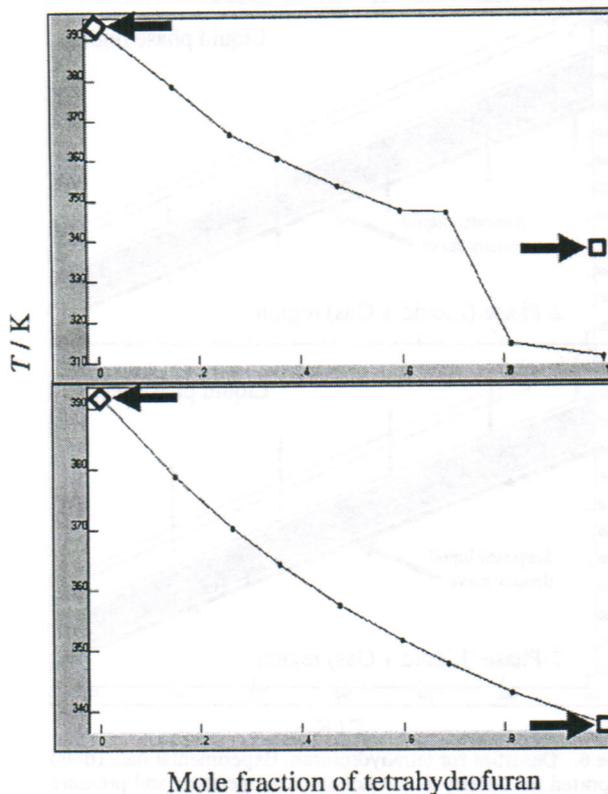


Figure 5. Boiling temperatures T_b at constant pressure for (tetrahydrofuran + tetrachloroethene). The figures show the reported experimental data before (upper)³⁷ and after (lower)³⁸ correction. The arrows indicate boiling temperatures for the pure components critically evaluated with the ThermoData Engine (TDE) software. \square , $T_b = (337.38 \pm 0.03)$ K at $p = 95.8$ kPa for tetrahydrofuran critically evaluated with TDE. \diamond , $T_b = (392.16 \pm 0.06)$ K at $p = 95.8$ kPa for tetrachloroethene critically evaluated with TDE. The graphs were created with the GDC software.⁵

the two-phase region, which is not possible. The originally published experimental data were corrected⁴⁰ after the results of the TDE analysis were reported to the authors. The corrected values are shown in the lower section of Figure 6. These data now appear correctly within the single-phase liquid region.

These examples represent a small fraction of the data analyzed and corrected within the new global delivery process for thermodynamic data. Every effort is made to complete the analysis and necessary data reporting adjustments before articles are published. Nevertheless, based on 2 years of experience with the new data delivery process, we estimate that 10% of articles reporting experimental thermodynamic data for organic compounds contain some erroneous information (the analysis is based on approximately 1000 articles). In a great many instances, these errors would be extremely difficult to detect during the article preparation, submission, and peer-review process. With ever increasing efficiency of data dissemination and use in engineering applications, the consequences of not addressing this situation are very serious. The new thermodynamic data delivery process provides a unique opportunity for curing this problem in a systematic and structural manner. It is a safe assumption that the situation described here with regard to thermodynamic data is typical for many other scientific fields dealing with large arrays of scientific experimental data.

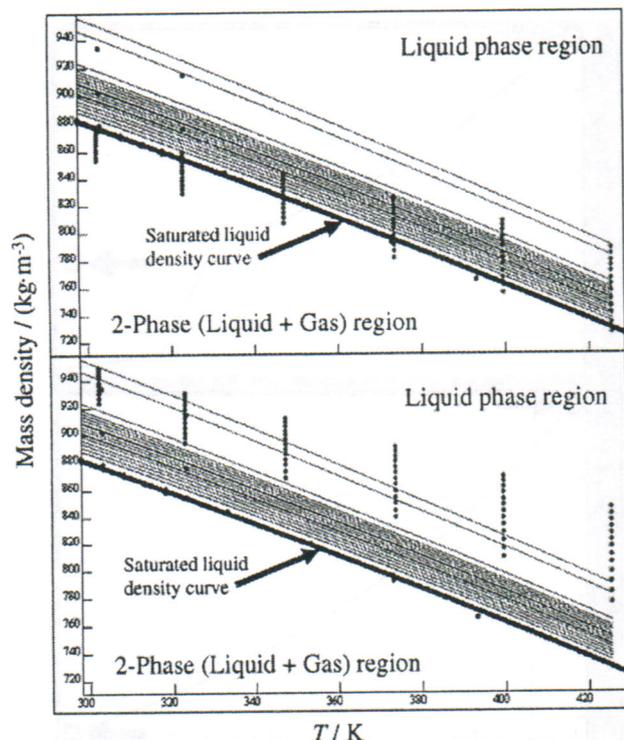


Figure 6. Densities for tetrahydrofuran. Experimental data (dots) are reported for density as a function of temperature and pressure in the liquid phase before (upper)³⁹ and after (lower)⁴⁰ correction. The saturation line (i.e., the line of separation between the one- and two-phase regions) critically evaluated with TDE is shown in the plot. The lines above the saturation density curve correspond to the isobars evaluated by TDE for the single phase region. The graphs were created with the TDE software.⁷

6. FUTURE POSSIBILITIES

The efficiency of the thermodynamic data delivery process described will be significantly further improved when on-demand critical data evaluation is extended to multicomponent mixtures and chemical reactions within the *TDE communication line* (Figure 3). The currently available version 1.0 of the NIST ThermoData Engine⁷ is limited to pure organic compounds. The developed model for the global data delivery process can certainly be extended to other chemical systems, such as biochemicals, ionic systems, and polymers, and can serve as a model for other fields, such as chemical kinetics and properties of materials.

Other future developments could include the design of software tools for strategic measurement planning based on the developed model for thermodynamic data communications. Development of these tools would lead to elimination of unjustified redundant measurements and identification of 'gaps' in the existing measurement space as well as determination of key measurements for validation of new computational and predictive techniques and would, therefore, play a central role in enhancing the existing knowledge base for industrial applications, with ensuing enormous potential economic benefit in the development of new chemical processes.

ACKNOWLEDGMENT

The authors would like to express their appreciation to Dr. Hans J. Koesch (Springer, New York, U.S.A.) and Dr.

David Martinsen (American Chemical Society, Washington, U.S.A.) whose advice was very valuable for the development of the global delivery process for thermodynamic property data. In addition, the authors thank Drs. T. L. Teague (ePlantData, Houston, U.S.A.), D. L. Embry (ConocoPhillips, Ponca City, U.S.A.), A. K. Dewan (Shell, Houston, U.S.A.), R. A. Perkins, M. L. Huber, and A. H. Harvey (NIST, Boulder, U.S.A.), and P. Linstrom and V. K. Shen (NIST, Gaithersburg, U.S.A.) for their practical suggestions in development of ThermoML and the NIST ThermoData Engine. We wish to express our appreciation to IUPAC for its support of the ThermoML standard development (project 2002-055-3-024) and personally to Dr. A. N. Davies (Creon Lab Control AG, Frechen, Germany), Secretary of the IUPAC Committee on Printed and Electronic Publications, for his unwavering support of the project. The authors wish to acknowledge the late Dr. Randolph Wilhoit of Texas A & M University (College Station, U.S.A.) who was an inspiration for the implementation of the dynamic data evaluation concept. For M.F., R.D.C., V.D., C.M., Q.D., S.E.S., J.W.M., W.M.H., and G.R.H. this work represents an official contribution of the U.S. National Institute of Standards and Technology and is not subject to copyright in the United States.

REFERENCES AND NOTES

- (1) Dong, Q.; Chirico, R. D.; Yan, X.; Hong, X.; Frenkel, M. Uncertainty reporting for experimental thermodynamic properties. *J. Chem. Eng. Data* **2005**, *50*, 546–550.
- (2) The future of the electronic scientific literature. *Nature* **2001**, *413* (6851), 1–3.
- (3) Frenkel, M.; Chirico, R. D.; Diky, V. V.; Marsh, K. N.; Dymond, J. H.; Wakeham, W. A.; Stein, S. E.; Königsberger, E.; Goodwin, A. R. H. XML-based IUPAC standard for experimental, predicted, and critically evaluated thermodynamic property data storage and capture (ThermoML): IUPAC Recommendations 2006. *Pure Appl. Chem.* **2006**, *78*, 541–612.
- (4) Diky, V. V.; Chirico, R. D.; Wilhoit, R. C.; Dong, Q.; Frenkel, M. Windows-based Guided Data Capture software for mass-scale thermophysical and thermochemical property data collection. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 15–24.
- (5) Diky, V. V.; Chirico, R. D.; Wilhoit, R. C.; Dong, Q.; Frenkel, M. Guided Data Capture Software, version 3.0; 2005. <http://www.trc.nist.gov/GDC.html> (accessed May 19, 2006).
- (6) Frenkel, M.; Chirico, R. D.; Diky, V. V.; Yan, X.; Dong, Q.; Muzny, C. ThermoData Engine (TDE): software implementation of the dynamic data evaluation concept. *J. Chem. Inf. Model.* **2005**, *45*, 816–838.
- (7) NIST ThermoData Engine. Standard Reference Database 103, version 1.0; National Institute of Standards and Technology, Gaithersburg, 2005. <http://www.nist.gov/srd/nist103.htm> (accessed May 19, 2006).
- (8) Wilhoit, R. C.; Marsh, K. N. Automation of numerical data compilations. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 17–22.
- (9) Frenkel, M. Dynamic compilation: a key concept for future thermophysical data evaluation. In *Report on Forum 2000: Fluid properties for new technologies – connecting virtual design with physical reality*; Rainwater, J. C., Friend, D. G., Hanley, H. J. M., Harvey, A. H., Holcomb, C. D., Laesecke, A., Magee, J. W., Muzny, C., Eds.; NIST Special Publication 975, 83; Gaithersburg, MD, 2001; pp 83–84.
- (10) Frenkel, M. Global communications and expert systems in thermodynamics: connecting property measurement and chemical process design. *Pure Appl. Chem.* **2005**, *77*, 1349–1367.
- (11) Wilhoit, R. C.; Marsh, K. N. *CODATA Standard Thermodynamics. Rules for preparing COSTAT message for transmitting thermodynamic data*; Report to CODATA Task Group on Geothermodynamic Data and Chemical Thermodynamic Tables; Paris, 1987. (b) <http://www-i5.informatik.rwth-aachen.de/lehrstuhl/projects/gco/> (accessed May 19, 2006). Web page contact: Alexander Kuckelberg (kuckelberg@informatik.rwth-aachen.de). (c) <http://www.fiz-karlsruhe.de/dataexplorer/test/iucosped/dataexplorer.html> (accessed May 19, 2006). Maintained by Fachinformationszentrum Karlsruhe; Karlsruhe, Germany. (d) Dewan, A. K.; Embry, D. L.; Willman, T. J. DIPPR/AIChE Project 991 – Thermophysical property data exchange. *Book of Abstracts of the 14th Symposium on Thermophysical Properties*; Boulder, CO, 2000; p 169.

- (12) <http://www.iupac.org/projects/2002/2002-055-3-024.html> (accessed May 19, 2006).
- (13) Finkelstein, C.; Aiken, P. *Building corporate portals with XML*; McGraw-Hill: New York, 1999.
- (14) <http://www.iupac.org/namespaces/ThermoML/index.html> (accessed May 19, 2006).
- (15) Whiting, W. B. Effects of uncertainties in thermodynamic data and models on process calculations. *J. Chem. Eng. Data* **1996**, *41*, 935–941. (b) Vasquez, V. R.; Whiting, W. B. Uncertainty and sensitivity analysis of thermodynamic models using equal probability sampling (EPS). *Comput. Chem. Eng.* **2000**, *23*, 1825.
- (16) Frenkel, M.; Chirico, R. D.; Diky, V. V.; Dong, Q.; Frenkel, S.; Franchois, P. R.; Embry, D. L.; Teague, T. L.; Marsh, K. N.; Wilhoit, R. C. ThermoML—an XML-based approach for storage and exchange of experimental and critically evaluated thermophysical and thermochemical property data. 1. Experimental data. *J. Chem. Eng. Data* **2003**, *48*, 2–13.
- (17) Chirico, R. D.; Frenkel, M.; Diky, V. V.; Marsh, K. N.; Wilhoit, R. C. ThermoML — an XML-based approach for storage and exchange of experimental and critically evaluated thermophysical and thermochemical property data. 2. Uncertainties. *J. Chem. Eng. Data* **2003**, *48*, 1344–1359.
- (18) *Guide to the Expression of Uncertainty in Measurement*; International Organization for Standardization: Geneva, Switzerland, 1993. This Guide was prepared by ISO Technical Advisory Group 4 (TAG 4), Working Group 3 (WG 3). ISO/TAG 4 has as its sponsors the BIPM (Bureau International des Poids et Mesures), IEC (International Electrotechnical Commission), IFCC (International Federation of Clinical Chemistry and Laboratory Medicine), ISO, IUPAC (International Union of Pure and Applied Chemistry), IUPAP (International Union of Pure and Applied Physics), and OIML (Organisation Internationale de Métrologie Légale). Although the individual members of WG 3 were nominated by the BIPM, IEC, ISO, or OIML, the Guide is published by ISO in the name of all seven organizations.
- (19) Sandhu, P. *The MathML Handbook*; Charles River Media, Inc.: Hingham, MA, U.S.A., 2003. See, also: www.w3.org/Math/ (accessed May 19, 2006).
- (20) Frenkel, M.; Chirico, R. D.; Diky, V. V.; Marsh, K. N.; Dymond, J. H.; Wakeham, W. A. ThermoML — an XML-based approach for storage and exchange of experimental and critically evaluated thermophysical and thermochemical property data. 3. Critically evaluated data, predicted data, and equation representation. *J. Chem. Eng. Data* **2004**, *49*, 381–393.
- (21) Murray-Rust, P.; Rzepa, H. S. Chemical markup, XML, and the Worldwide Web. 1. Basic principles. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 928–942.
- (22) <http://www.matml.org/> (accessed May 19, 2006).
- (23) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. See, also: <http://www.rcsb.org/pdb/> (accessed May 19, 2006).
- (24) <http://www.ccdc.cam.ac.uk/products/csd/> (accessed May 19, 2006). Maintained by the Cambridge Crystallographic Data Centre; Cambridge, U. K.
- (25) Diky, V. V.; Chirico, R. D.; Wilhoit, R. C.; Dong, Q.; Frenkel, M. Windows-based Guided Data Capture Software for mass-scale thermophysical and thermochemical property data collection. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 15–24.
- (26) See <http://www.trc.nist.gov/GDC.html> (accessed May 19, 2006).
- (27) Frenkel, M.; Dong, Q.; Wilhoit, R. C.; Hall, K. R. TRC SOURCE database: a unique tool for automatic production of data compilations. *Int. J. Thermophys.* **2001**, *22*, 215–226.
- (28) <http://trc.nist.gov/ThermoML.html> (accessed May 19, 2006). Maintained by NIST/TRC; Boulder, CO, U.S.A.
- (29) Marsh, K. N. New Process for data submission and dissemination. *J. Chem. Eng. Data* **2003**, *48*, 1.
- (30) Electronic data submission to the NIST Thermodynamics Research Center. *J. Chem. Thermodyn.* **2004**, *36*, iv.
- (31) Electronic data submission to NIST Thermodynamics Research Center. *Fluid Phase Equilib.* **2004**, *226*, v.
- (32) Electronic data submission to NIST Thermodynamics Research Center. *Thermochim. Acta* **2004**, *421*, 241.
- (33) Haynes, W. M. Electronic data submission for the International Journal of Thermophysics. *Int. J. Thermophys.* **2005**, *26*, 307.
- (34) Names of commercial products and/or commercial entities are provided for complete scientific description and as a service to the reader of this publication. Such identification neither constitutes nor implies endorsement of such products or companies by NIST or by the U.S. government. Other products or services may be found to serve just as well.
- (35) Changsheng Yang, C.; Xu, W.; Ma, P. Excess Molar Volumes and Viscosities of Binary Mixtures of Dimethyl Carbonate with Chlorobenzene, Hexane, and Heptane from (293.15 to 353.15) K and at Atmospheric Pressure. *J. Chem. Eng. Data* **2004**, *49*, 1802–1808.
- (36) Changsheng Yang, C.; Xu, W.; Ma, P. Correction. *J. Chem. Eng. Data* **2005**, *50*, 1496.
- (37) Prasad, T. E. V.; Raj, E. D. A.; Maheedhar, G.; Reddy, M. S.; Kumar, V. S.; Garapati, S.; Patanjali, V.; Prasad, D. H. L. Bubble-Temperature Measurements on Some Binary Mixtures Formed by Tetrahydrofuran or Amyl Alcohol with Hydrocarbons, Chlorohydrocarbons, or Butanols at (94.6 or 95.8) kPa. *J. Chem. Eng. Data* **2004**, *49*, 746–749.
- (38) Prasad, T. E. V.; Raj, E. D. A.; Maheedhar, G.; Reddy, M. S.; Kumar, V. S.; Garapati, S.; Patanjali, V.; Prasad, D. H. L. Correction. *J. Chem. Eng. Data* **2005**, *50*, 1497.
- (39) Zhang, W.; Kiran, E. (p,V,T) Behaviour and miscibility of (polysulfone+THF+carbon dioxide) at high pressures. *J. Chem. Thermodyn.* **2003**, *35*, 605–624.
- (40) Zhang, W.; Kiran, E. Corrigendum. *J. Chem. Thermodyn.* **2006**, *38*, 504.

CI600208F