**A GLOBAL INFORMATION- EXPERT- SYSTEM for THERMOPHYSICAL PROPERTIES. TRC/NIST THERMOPHYSICAL DATABASE. CONTRIBUTION of the RUSSIAN PUBLICATIONS and ESTIMATING THEIR QUALITY**

Abdulagatov I.M.1,2

1*Geothermal Research Institute of the Dagestan Scientific Center of the Russian Academy of Sciences, Makhachkala, Dagestan, Russia,*

2*Present address: National Institute of Standards and Technology, Boulder, Colorado, U.S.A., e-mail:* [*ilmutdin@boulder.nist.gov*](mailto:ilmutdin@boulder.nist.gov)

Production of research thermophysical data is growing about 7 % annually, doubling every 10 years. Poor quality data are leading to mistakes in equipment selection, wrong or not efficient design of the technological equipment (technological application problems), new processes simulating and discovering problems (scientific application problems), and poor regulatory decisions [1]. The traditional journal’s peer-review process takes very limited time which is not enough to thorough evaluate all new experimental data and correctly estimate their quality. According to results of a 10- year collaboration between the National Institute of Standards and Technology (NIST) and five key thermophysical journals (Int. J. Thermophysics; Fluid Phase Equilibria; J. Chemical Thermodynamics; J. Chem. Eng. Data, and Thermochimca Acta), conventional journal’s peer review process is not enough to ensure reported data quality. Approximately 10-15% of all published articles include typographical errors in the tables of experimental data. Problems: number transposition (“normal” typos); unit errors; component switch for mixtures; wrong compound; erroneous column duplication in tables; and hidden errors (calibration errors); poorly defined variables/constraints; failure to report experimental data; nonsensical uncertainty estimates relative to data scatter. About 30 % of manuscripts submitted to the key thermophysical journals contained erroneous according to a report by NIST and the collaborating journals [1]. Also 20% of all submitted articles have serious problems in the tables of experimental data. After we check and repair data with author cooperation, 30% of the articles are rejected. Rapid development of new chemical processes, compound types, and measurement efficiencies, plus greater globalization of scientific research, has overwhelmed the slow and labor-intensive traditional methods of critical data evaluation, and has created the environment in which a dynamic and responsive approach is essential.

Global Information- Expert- System (GIES) developed at NIST/TRC provides critically evaluated thermophysical property data to users worldwide. GIES includes collection of experimental data from the literature from 5 key International journals mentioned above and from Russian and Chines journals, validation, critical evaluation, and dissemination.A NIST/TRC GIES is the first full-scale software implementation of the concept of dynamic data evaluation for thermophysical properties [2-8]. The GIES includes:

(1) Software tools for mass-scale data capture, Guided Data Capture (GDC) software (for collection of data from the literature);

(2)Comprehensive Data Storage Facility SOURCE Data Archival System;

(3) Data Entry Facility;

(4) Software Expert Systems, Thermodynamic Data Engine (TDE) software(for critical evaluation of the data, quality assessments);

(5) Data Communications Standard, Thermo ML (Store and Exchange Thermophysical Data);

(6) Data Reader Software, ThermoML Opener into Microsoft Excel;

(7) Web Communication Portal, NIST Web-Oracle infrastructure.

The ThermoData Engine (TDE) software expert software is the core component in implementation of the concept of GIES in Science for the field of thermodynamics. The main goal of the GIES is providing critically evaluatedthermophysical property data (boiling, melting, and critical points data; density; vapor-pressure; sublimation pressure; heat capacity; enthalpy; speed of sound; solubility; VLE; LLE; SLE; surface tension; viscosity; thermal conductivity; chemical reaction data; *etc.*) of fluids and fluid mixtures (binary and ternary) for use by industry, academia, and other government agencies for: (1) chemical technology process development and optimization; (2) fundamental research into molecular properties; (3) estimation quality of published thermophyscal data; (4) regulatory decisions. Implementation requires the combination of large electronic thermophysical databases, storing essentially all relevant experimental information known to date with metadata and uncertainties, with expert-system software, designed to automatically generate recommended property values based on available experimental and predicted data. This combination of tools allows production of critically evaluated data on demand.

Initially TDE was focused on thermodynamic property data evaluation for pure fluids. Later versions TDE has included on-demand generation equation of state (EOS), dynamic update of local data resources based on the TRC-SOURCE data storage system, support for binary mixtures, including VLE data, and chemical reaction properties. The most recent version [9] included experimental planning support features and a product design tool for identification of compounds with specific properties [2]. Also included were advanced consistency tests for automated quality assessment for vapor-liquid equilibrium data sets [2]. TDE is used in a broad variety of applications, including: (1) product and process design; (2) data product development; (3) data quality assurance; and (4) validation of new experimental data in advance of publication. Recent additions include a vapor−liquid equilibrium (VLE) data modeling test for assessment of mixtures, involving supercritical components, that complements consistency tests implemented previously for low-pressure/subcritical VLE data, as well as addition of the NIST-KT-UNIFAC group-contribution prediction method for VLE data with automatic decomposition of molecular structures into KT-UNIFAC groups and subgroups. TDE was extended to critical evaluation of thermodynamic properties of ternary mixtures through on-demand dynamic evaluation of the properties of the binary subsystems. The methodological bases of the GIES developed at NIST/TRC, although initially was focused on the field of thermophysics, can be applied to other scientific areas and engineering.

**Contribution of the Russian Publications and Estimating Their Quality**

NIST/TRC DataBase includes collection of experimental thermophysical data from the Russian literature from 30 academic peer reviewing journals (High Temperature; Russian J. Phys. Chem.; Russian J. Appl. Chem.; Inzh. Fiz. Zhurnal; Izv. Vuzov, ser. Neft & Gas; Russian J. Nonorganic Chem.; Dok. Acad. Nauk; Geochemistry; J. Exp. Theor. Phys.; Ukr. Phys. J.; *etc*.). NIST/TRC data entry (GDC) and data quality (TDE) check facilities have been implemented for Russian publications covering the period from 1930 to 2013.

**Statistics of the Russian Articles in the NIST/TRC DataBase by Journals (Russian)**: 1. High temperature -272 articles; J. Phys. Chem. -1943; J. Appl. Chem.-1388; Inzh. Fiz. Zhurnal-251; Izv. Vuzov, ser. Neft & Gas-372; J. Nonorganic Chemistry-319; Dok. Acad. Nauk-430; Geochemistry-13; J. Exp. Theor. Phys.-67; Ukr. Phys. J.-94. In total 5249 Russian articles in the NIST/TRC DataBase.

**Statistics by data points for pure fluids, binary and ternary mixtures**: Total compound-2488; Total pure component data points – 116134; Total binary data points-115471; Total ternary data points-40918; and Reaction data points-941. In total Russian contribution to the NIST/TRC DataBase is 273464 data points.

**Statistics of the data points by year** (from 1930 to 2013): (1930-1950)- **5831**; (1951-1961)-**25357**; (1962-1972)- **69507**; (1973-1983)-**78844**; (1984-1994)-**61929**; (1995-2005)-**24364**; and (2006-2013)-**8640.**

**Total**: 5250 articles (High Temperature-273, J. Phys. Chem.-1943, J. Appl. Chem.-1388, Inzh. Fiz. Zh.-251, Izv. Vuzov, ser. Neft & Gas-372, J. Nonorg. Chem.-319, Dok. AN-430, Uk. Fiz. Zh.-94, J. Exp. Theor. Phys.-67, and Geochem.-13); 273 346 data points (116134 for pure fluids, 115471 for binary mixtures, 272523 for ternary, and 941 for reaction data).

TDE quality test found that about (35-40) % of all Russian published articles include typographical errors in the tables of experimental data and have serious problems in the presentation of experimental data.

**Acknowledgment**

The author thanks the Applied Chemicals and Materials Division at the National Institute of Standards and Technology for the opportunity to work as a Guest Researcher at NIST during the course of this research. I would like also express my appreciation to all TRC staff (Dr. Frenkel, Dr. Chirico, Dr. Diky, Dr. Magee, Dr. Kroeblein, Dr. Kazakov, and Dr. Muzny) for their valuable help during my work at TRC/NIST.

**REFERENCES**

\*1.R.D. Chirico, M. Frenkel,J.W. Magee, V. Diky, Ch.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, Th.W. de Loos, D.G. Friend, A.R.H. Goodwin, Lee D. Hansen, W,M. Haynes, N. Koga, A.M. Kenneth N. Marsh, P.M. Mathias, C. McCabe, J.P. O’Connell, A. Pádua, V.R. Ch.Schick, J. P. M. Trusler, S. Vyazovkin, R.D. Weir, J.Wu., *J. Chem. Eng. Data* **58** (2013) 2699.

2. V. Diky, R.D. Chirico, A.F. Kazakov, Ch.D. Muzny, J.W. Magee, I.M. Abdulagatov, J.Won Kang, K.Kroenlein, M.Frenkel. J. Chem. Inf. Modeling **51** (2011) 181.

3. V. Diky, R. D. Chirico, Ch.D. Muzny, A.F. Kazakov, K. Kroenlein,J.W. Magee,I. Abdulagatov,J.W. Kang, M. Frenkel. J. Chem. Inf. Modeling **52** (2012) 260.

4. Jeong Won Kang, V. Diky, R.D. Chirico, J.W. Magee, Ch.D. Muzny, I.M. Abdulagatov, A.F. Kazakov, M. Frenkel. Fluid Phase Equilib. **309** (2011) 68.

5. A. F. Kazakov, Muzny Ch.D., Kroenlein K.G., Diky V., Chirico R.D., Magee J.W., Abdulagatov I.M., Frenkel M. Int. J. Thermophys. **33** (2012) 22.

6**.** V. Diky, R. D. Chirico, Ch.D. Muzny, A.F. Kazakov, K. Kroenlein,J.W. Magee,I. Abdulagatov, C.A. Diaz-Tovar, J.W. Kang, R. Gani, M. Frenkel. *J. Chem. Inf. Mod.* **53** (2013) 249.

7. Jeong Won Kang, V. Diky, R. D. Chirico, J. W. Magee, Ch. D. Muzny, I.M. Abdulagatov, A. F. Kazakov, M. Frenkel. J. Chem. Eng. Data **55** (2010) 3631.

8. V. Diky, R.D. Chirico, Ch.D. Muzny, A.F. Kazakov, K. Kroenlein, J.W. Magee, I. Abdulagatov, and M. Frenkel. *J. Chem. Inf. Mod.* **53** (2013) 3418.

9. M. Frenkel, R.D. Chirico, V. Diky, Ch.D. Muzny, A.F. Kazakov, J.W. Magee,I. Abdulagatov, J.W. Kang, *NIST ThermoData Engine, NIST Standard Reference DataBase 103b –Pure Compounds, Binary Mixtures, and Chemical Reactions, version 5.0; Standard Reference Data Program, National Institute of Standards and Technology*: Gaithersburg, MD, 2010.