Netsu Sokutei **24** (4) 199-205

論 文

On Quality of Adopted Values in Thermodynamic Databases

Vladimir S. Iorish and Gleb V. Belov

(Received May 5, 1997)

Quality of numerical data in thermodynamic databases is known critically influence upon the results of modeling of various processes. That is why the analysis of data quality and searching for erroneous values are very actual.

This paper outlines automated procedure which allows to find the most considerable mistakes in thermodynamic properties of individual substances adopted in a database. HSC-database is analyzed as an example. Comparison procedure of this database and IVTANTHERMO database is supposed to reveal large discrepancies as a sign of erroneous values in one of two databases. Detail examination of the case allows to find the errors.

The most remarkable discrepancies are analyzed to show their influences upon the result of thermodynamic modeling.

General approach to prevent erroneous values usage is discussed. The data quality in databases could be improved by automated control of input values using a set of conditions. These conditions may be added and corrected taking into account theoretical and experimental data.

The importance of international cooperation in improvement of thermodynamic databases quality is emphasized. The time has come for special program development to agree and combine all databases on thermodynamic properties of individual substances. The ultimate goal of the program may be a joint international database, which could be maintained and developed on "real time" regime.

1. Introduction

Thermodynamic databases are widely used in scientific research, engineering calculations, physico-chemical experiments planning as well as for teaching and other purposes. Progress of recent decades in petrochemistry and metallurgy, creation of rocket engines and modem structural materials, development of traditional and nuclear power engineering and many others could not have been achieved without preliminary thermodynamic analysis of the relevant processes. Successful solution of many global problems facing humankind today depends to some extent on the reliability of theoretical models of complex physico-chemical systems. One of the most important parts of such models is set of thermodynamic properties of the system being studied. The development of new methods of supplying humankind with energy, the choice of the optimum ways for utilization of industrial wastes, solution of the safety problems of nuclear power engineering, as well as solution of many other problems will be impossible without preliminary analysis based on thermodynamic modeling.

Basic part of any complex physicochemical system's model incorporates an idea about individual substances of species. The characteristics of the whole system are largely determined by the properties of substances, which form the system. That is why so necessary are further experimental and theoretical study of thermodynamic properties of individual substances and accumulation of

Thermocenter RAS, Izhorskaya 13/19, Moscow 127412, Russian Federation

this information. This information is intended for scientists and engineers who work in various branches of science and engineering and it must be delivered them in an intelligible form. One may regard the thermodynamic properties of substances (enthalpy of formation, heat capacity, entropy etc.) as fundamental physicochemical properties, which are used for analysis of processes with participation of these substances in a wide range of pressure and temperature. However the array of thermodynamic properties as distinct from fundamental physical constants is practically unlimited, it is determined first of all by multitude of species in various phase states. Because of permanent growth of the experimental and theoretical information about substances it is necessary to systematize it in the form of reference books, as well as in the electronic form (data bases and data banks). For detailed review and critical analysis of such reference books and data bases see.[1]

The evaluation of quality of the information published in reference books and stored in data bases is a very complicated task, Still more difficult and practically unsolvable task is the evaluation of results of the use of unreliable information for thermodynamic analysis of various processes. Main reason of complexity of the first task is that the complete information about the adopted values, *i.e.* all primary data and details of their processing is not available. It should be mentioned here that well known problem of thermodynamic consistency can't be investigated within limited set of data, it's necessary to take into consideration information from other data bases or original data sources. Another reason of the problem complexity is difficulties of algorithmization of the data quality expert analysis procedure, which is usually fulfilled by qualified specialists. At present it is not easy to find financial support for such investigations, and the number of qualified specialists in this area is very small.

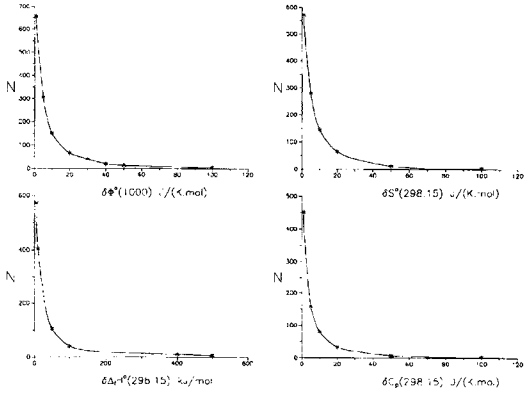

Fig.1 Number of cases (N) for which difference in function exceeds δF.

In present report we are trying to evaluate data quality of one of known data bases by means of computer processing, to carry out the analysis of reasons that caused greatest errors and to discuss possible ways of their preventing.

2. Description of Analysis Procedure, Statistical Data and Examples of Errors Found

In order to evaluate data quality it's proposed a procedure of comparison of adopted thermodynamic property values from two data bases (analyzed data base and comparison data base). We have taken HSC data base[2] as an analyzed one. This data base is uncritical compilation from various data sources, and it isn't internally consistent from the thermodynamic point of view. As a comparison data base we have taken IVTANTHERMO database,[3] all thermodynamic properties of which are evaluated by experts and represent a system of mutually consistent quantities.

We have compared the following thermodynamic properties:

$\Delta_f H^0$ (298.15 K) - enthalpy of formation from elements in standard state;

C_p^0 (298.15 K) - isobaric heat capacity;

S^0 (298.15 K) - entropy;

F^0 (1000 K) $= -[G^0(T) - H^0(298.15)]/T$ - reduced Gibbs energy.

It appeared that the compared data bases contain about 1300 the same individual substances. We have compared the mentioned property values from the two data bases with various levels of discrepancy criterion (δ) and found number of substances for which the absolute value of the difference between selected properties exceeds the value of δ. **Fig.1** presents corresponding graphs for all investigated properties. Similarity of all graphs is noticeable. The attention should be drawn to big number of substances with significant discrepancies of property values. For example,

in case of enthalpy of formation discrepancy exceeds 100 kJ $mole^{-1}$ for 40 substances, for other properties-discrepancy exceeds

20J $K^{-1}mole^{-1}$ in case of heat capacity for 33 substances, in case of entropy for 65 substances, and in case of reduced Gibbs energy for 66 substances. It should be noted that these large discrepancies may occur because of errors in one of the data bases or in both ones. Only detailed investigation could answer the question about the reasons of these discrepancies. No one error was found in comparison data base in this study. However, it is known that there are some misprints and errors in the reference book [19] on which IVTANTHERMO is based.

The found discrepancies of thermodynamic property values are practically uncorrelated as shown on charts **Figs.2-4**. Calculated values of correlation coefficient for each of three pairs of functions are about zero. This fact confirms independence of accuracy of thermochemical data values and thermodynamic functions, which usually takes place.

As a result of our analysis we have discovered two main kinds of errors: random and object-dependent.

Random errors are caused by:

- misprints or mistakes in data sources;
- incorrect use of published data;
- errors of input into the data base.

Object-dependent errors have different origin, that is why it's so difficult to classify them. And nevertheless there are among these errors:

- use of out-of-date data (when there exist up-to-date ones);
- use of unreliable methods of estimation or estimation, based on out-of-date information;
- use of results of the only experimental study, which badly correlate with more reliable data for analogous substances;

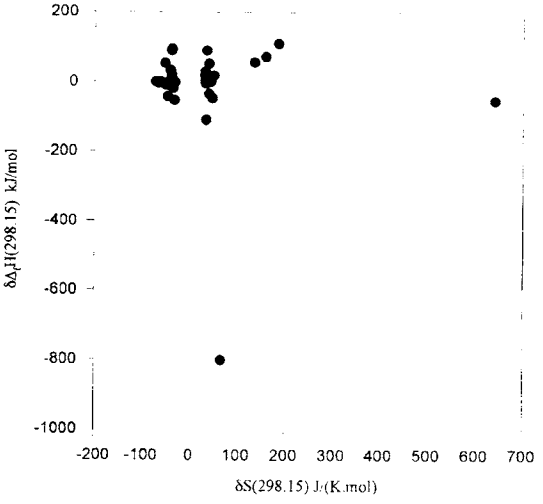

Fig.3 Discrepancies correlation for the cases $\delta S(298.15\ K) > 20\ JK^{-1}\ mol^{-1}$.

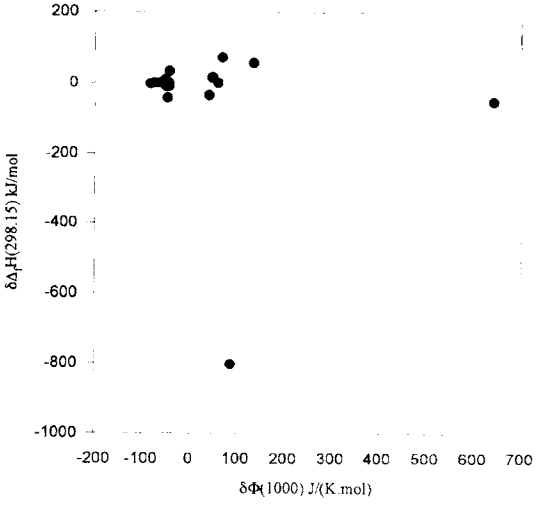

Fig.2 Discrepancies correlation for the cases $\delta\Phi(1000) > 40\ JK^{-1}\ mol^{-1}$.

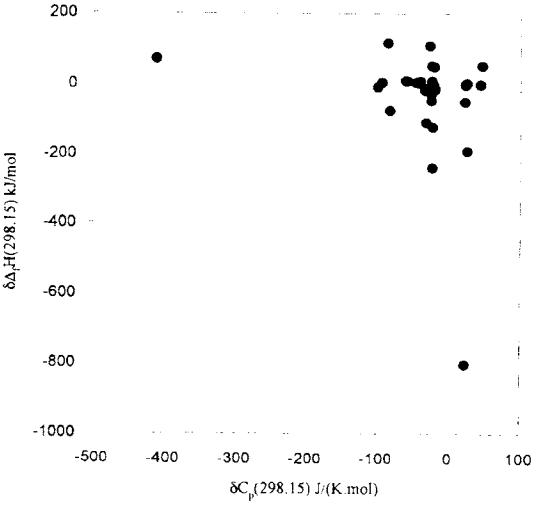

Fig.4 Discrepancies correlation for the cases $\delta C_p(298.15\ K) > 20\ JK^{-1}\ mol^{-1}$.

- acceptance of rough assumption $C_p(T)$ = const. for wide temperature range.

Let us consider several examples, which demonstrate various kinds of error sources.

Example 1 Cs_2O(g)

Data base HSC contains negative (!) value of standard entropy $S^0(298.15\ K) = -318\ J\ K^{-1}\ mole^{-1}$. We had consulted with the data source [4] and have discovered a regrettable misprint in the table heading. It should be noted that the corresponding value in the table itself is correct.

Example 2 Ba(g)

In data base HSC for the temperature range 298.15-3000 K is adopted equation for heat capacity from reference book [5]. However in the data source this equation is valid only for the range 2171-3000 K. When this equation was used at $T = 298.15$ K it gave an absurd value $S^0(298.15K) = 10\ J\ K^{-1}\ mole^{-1}$, what is significantly less than corresponding value for solid barium.

Example 3 $SrSO_4$(cr)

For this substance there exists a tremendous discrepancy for the value of $\Delta_f H^0(298.15\ K)$ - more than 800 kJ mole^{-1}. After the data source [5] had been consulted with it has been found, that in data base HSC were entered values not of $SrSO_4$(cr) but of SrO_2(cr), which are located at the same page.

Example 4 Rb_2(g)

For this substance data in HSC were taken from [5], On examination the tables for Rb_2(g) and Rb(g) one can notice an evidently wrong inequality $S^0(Rb_2(g)) < S^0(Rb(g))$. There are no doubts that the value of entropy for Rb_2(g) is erroneous. Data source [5] refers to the first edition of reference book by Hultgren.[6] We consulted with the data source and have discovered that the respective table refers to 1/2 Rb_2 but not Rb_2.

The examples mentioned above demonstrates errors, which we called random. Let us further discuss two examples of object-dependent errors.

Example 5 P_4O_6(g)

For this substance data base HSC contains value $\Delta_f H^0(298.15\ K) = -2214.173$ kJ mole^{-1}, which is different from the corresponding value adopted in IVTANTHERMO ($\Delta_f H^0(298.15\ K) = -1606.000$ kJ mole^{-1}) by more than 600 kJ mole^{-1}. Source of data in HSC-reference book.[7] This reference book in its turn contains data from the second edition of JANAF,[8] the value was recalculated in order to take into consideration white phosphorus as phosphorus standard state (in [8] standard state for phosphorus is red phosphorus). Corresponding data in [8] are based on experimental study of the heat of formation of phosphorus in nitrogen oxygen, made by Koerner and Daniels.[9] The later investigations showed that complexity of combustion products composition and inaccuracy of its determination had led to gross error in [9]. More reliable value adopted in IVTANTHERMO is based on measurement, made by Hartley and McCoubrey,[10] They have found the value of heat of formation of P_4O_6 in crystal state by burning it P_4O_{10}(cr). It is significant that the value was confirmed in later mass-spectral investigation of reactions equilibria with the participation of P_4O_6(g), made by Smoes and Drowar.[11]

$\Delta_f H^0(298.15\ K) = -2214.173$ kJ mole^{-1},

Example 6 VF_2(cr)

Considerable discrepancies found for this substance in adopted value of $\Delta_f H^0(298.15\ K)$ (> 200 kJ mole^{-1}) and $C_p^0(298.15\ K)$ (> 20 J K^{-1} mole^{-1}). The reasons for both cases are poor data estimations ($\Delta_f H^0(298.15\ K) = -753$ kJ mole^{-1} and $C_p^0(298.15\ K) = 86\ J\ K^{-1}\ mole^{-1}$), taken for HSC from [12]. Meanwhile for a long time there are reliable experimental data about heat of formation ($\Delta_f H^0(298.15\ K) = -992.3$ kJ mole^{-1})[13] and about heat capacity ($C_p^0(298.15\ K) = 63.18\ J\ K^{-1} \cdot mole^{-1}$),[14] which were used for IVTANTHERMO.

3. Influence of Erroneous Data on the Results of Chemical Equilibria Calculations

As was already mentioned above it is impossible to evaluate all consequences of the use of erroneous data in modeling of various processes. We will consider several examples demonstrating the fact that wrong thermodynamic properties of individual substances may cause significant discrepancies between calculated and experimental parameters of the system being studied.

Table 1 contains calculated results of equilibrium composition of zirconium tetraiodide decomposition products at $p = 0.01$ atm and $T = 1600$-1800 K. These conditions approximately correspond to well known zirconium refinement process.[15] As one may see from the **Table 1** with data from HSC zirconium at this conditions remains in compounds with iodine and is not reduced to metal. At the same time results obtained with

Table 1 Equilibrium composition (mole) of zirconium tetraiodide decomposition products at $P=0.01$ atm.

	IVTANTHERMO HSC					
T, K	Zr(cr)	ZrI_4(g)	I(g)	ZrI_3(g)	I_2(g)	ZrI_2(g)
1600	0.0437	0.8994	0.2316	0.0568	0.0001	0.0001
	0	0.0318	1.5505	0.3800	0.0029	0.5880
1700	0.1350	0.7665	0.6380	0.0982	0.0004	0.0004
	0	0.0056	1.8060	0.1800	0.0018	0.8150
1800	0.3934	0.4745	1.7050	0.1310	0.0010	0.0012
	0	0.0010	1.9200	0.0760	0.0010	0.9230

data from IVTANTHERMO are in qualitative agreement with used technological process.

The reason of such a significant discrepancy are wrong data on the heat of formation for ZrI_2(g)(−66.9 kJ mole^{-1}) and ZrI_3(g) (−221.7 kJ mole^{-1}), adopted in HSC, while IVTANTHERMO contains more reliable data (123.3 and −125.7 kJ mole^{-1} respectively). Use of erroneous data led to fictitious stabilization of di- and triiodide and, as a result, the formation of metallic zirconium was suppressed.

In **Table 2** equilibrium pressure-temperature dependence of O_2(g) over $SrSO_4$(cr) is presented, which is calculated with the use data from HSC and IVTANTHERMO, as well as known experimental values.[16] As one may see from the table wrong values of thermodynamic properties for $SrSO_4$(cr) in data base HSC, discussed above, Lead to absolutely unreal big values of oxygen pressure. In other words, existence of solid strontium sulfate is impossible at examined temperature range, the fact that contradicts the experimental results.

Experimental results for SHS titanium silicides obtainment process[17] show that for all temperature range titanium disilicide is more stable than titanium silicide. The same result was obtained from thermodynamic calculation when data from IVTANTHERMO were used. However, when calculation were carried out with data from HSC it appeared that TiSi(c) and Si(c) were present instead of $TiSi_2$(c). The reason of the result is adopted in HSC an out-of-date value $\Delta_f H^0$(298.15 K) = −134 kJ mole^{-1} for titanium disilicide. In the recent work[18] the value −170.9 ± 8.3 kJ mole^{-1} is obtained, which was adopted in IVTANTHERMO.

The examples mentioned above demonstrates how erroneous and unreliable data may cause evident failures of thermodynamic modeling. In many other cases some not so evidently mistaken data may cause not so noticeable errors of modeling. Thus the errors may result in serious loss of time and funds.

Table 2 Composition of calculated and experimental values of the oxygen perssure (in atm) over solid strontium sulfate.

T, K	IVTANTHERMO	HSC	Experimental data [16]
1000	0.46E-16	1.8E6	0.37E-16
1100	0.88E-14	6.1E6	0.71E-14
1200	0.69E-12	1.8E7	0.57E-12
1300	0.27E-10	4.5E7	0.24E-10

4. Development of Criteria for Automated Data Testing in Data Base

In this regard a question arises how to avoid errors and enhance quality of the data stored in a database.

Of course the most evident way is an expert analysis of all new data, their evaluation resulted in recommended values and their uncertainties. Historical experience of creation of critical reference editions and data bases in Russia ([19], [20], IVTANTHERMO) and USA ([4], NBSTHERMO]) has demonstrated that high quality data sets may be obtained only by groups of qualified specialists. The attention should be drawn to the fact that appearance of computerized data bases did not result in quality enhancing of the data concerning the thermodynamic properties of individual substances. Computer's application led to growth of database entries, to development of data estimation techniques for substances, which were not studied experimentally. Computer makes easier access to data and their use for processes modeling. The lack of progress in the quality of the data stored in computer data bases is explained mostly by the fact that main source of information for them is reference books. And in addition to usual shortcomings of the reference books the data bases often have another one caused by data inconsistency.

Data bank IVTANTHERMO was also created on the basis of the third edition of reference book "Thermodynamic properties of individual substances". However during its development the roles were changing and the fourth edition[19] is being published in USA on the basis of IVTANTHERMO. Despite the highly automated work of experts and publication of reference tables from computer files it turned out to be impossible to avoid all errors. The reason of this as well as reasons of discussed errors in HSC data base is in our opinion

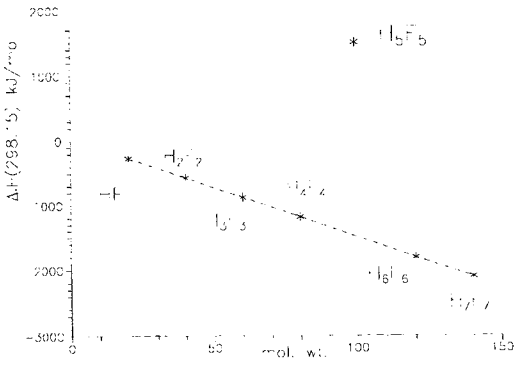

Fig.5 Enthalpies of formation for hydrogen fluoride clusters.

insufficient level of new data input and the data treatment control automation. This control could at least essentially decrease the quantity of random errors and in some cases help to avoid object-dependent ones. Development of more or less complete set of the control criteria may be regarded as "knowledge bank" filling process and it must become an object of further investigations. In present work we will only outline the possible approaches to solution of this task.

We suppose it's useful to examine three groups of checking criteria

1) simple inequalities based on physical meaning of parameters;
2) quantitative inequalities, which reflect possible relations between new data and already existing one
3) checking of agreement with various correlations of the type COMPOSITION -PROPERTY for analogous species in the periodic table, types of crystal lattice etc.

As examples of inequality from the first group may serve the following ones:

$S^0(T) > 0$, $C_p0(T) > 0$; $\Delta_a H^0(T) > 0$, $\Delta_s H^0(T) > 0$, etc.

where a- atomization, s- sublimation.

Examples from the second group are:

$S^0(A_n) > S^0(A)$, $S^0(M(cr)) < S^0(M(g))$, etc.

where A_n- gas cluster, A- cluster fragment, M- individual substance.

Example of correlation testing is shown on **Fig.5**. As one can see from the picture significant deviation of property of one of the hydrogen fluorine clusters from linear dependence allows to discover an error.

5. Conclusion

In conclusion we would like to point out the following:

* quality of adopted values in thermodynamic data bases needs to be improved;
* errors in data bases may result in quantitative and qualitative mistakes in process modeling;
* significant number of errors may be eliminated by means of computer control of the data input;
* "NO" - to data bases compiled from reference books, "YES" - to periodic publications of the information from data bases.

At present the most efficient thing to do is creation of an international data base, supported by leading laboratories from USA, Europe, Russia, China and Japan. Their contributions may contain financial, technical and intellectual components. Making use of advantages of international network INTERNET it would be possible to found an International Data Center on Thermodynamic Properties of Pure Substances (IDC TPPS). On the first stage IDC TPPS could develop unified rules for analysis, processing and evaluation of data sets. Later, after initial subject filling it could function in real time mode expanding the data base with new information and spreading the actual data among interested organizations and persons.

References

1) L. V. Gurvich , The reference books and data bases on the thermodynamic properties of pure substances, 12th IUPAC Conference on Chemical Thermodynamics, Joint meeting with the 47th Calorimetry Conference. Program, Abstracts, and Reports. Snowbird, Utah, U.S.A., p.183 (1992).
2) HSC Chemistry for Windows, Version 1.10, Outokumpu Research Oy, Pori, Finland (1993).
3) IVTANTHERMO - A Thermodynamic Database and Software System for the Per-sonal Computer. User's Guide, L. V. Gurvich, V. S. Iorish, D. V. Chekhovskoi and V. S. Yungman, Boca Raton, CRC Press, Inc. (1993).
4) M. W. Chase, J. L. Curnutt, A. T. Hu, H. Prophet, A. N. Syuerud and L. C. Walker, JANAF Thermochemical Tables. Third Edition (1985).
5) I. Barin, O. Knacke and O. Kubaschtwski, Thermochemical properties of inorganic substances.

Springer-Verlag Berlin Heidelberg New York (1977).

6) R. Hultgren, R. L. Orr, D. Anderso and K. K. Kelley, Selected Values of thermodynamic Properties of Metals and Alloys, New York; Jon Wiley & Sons (1963).

7) L. B. Pankratz, Thermodynamic properties of elements and oxides, Bulletin (United States. Bureau of Mines), 872. Washington, D.C. (1982).

8) JANAF Thermochemical Tables. Second Edition. D.R. Stull and H.Prophet (Eds.), NSRDS-NBS, N 37, Washington (1971).

9) W. E. Koerner and F. Daniels, *J. Chem. Phys.* **20**, 113 (1952)..

10) S. B. Hartley and J. C. McCoubrey, *Nature* **198,** 476 (1963).

11) S. Smoes and J. Drowart, *Faraday Sympos. Chem. Soc.*, 139 (1973).

12) L. P. Ruzinov and B. S. Gulyanitsky, Ravnovesniye prevrascheniya metallurgicheskikh reaktsy. Moscow, "Metallurgiya" (1975).

13) R. A. Kent, K. Zmbov, J. D. McDonald *et al.*, Proc. Conf. Nucl. Appl. Nonfissionable Ceramics, Washington, D.C. (1966).

14) J. W. Stout and W. O. J. Boo, *J. Chem. Phys.* **71**, 1 (1979).

15) A. I. Yevstukhin, Metals of high purity, Moscow, Nauka, p.114 (1976).

16) M. Fredriksson and E. Rosen, *Chem. Scr.* **16,** 34 (1980).

17) T. S. Azatyan, V. M. Mal'tsev, A. G. Merzhanov and V. A. Seleznev, *The Physics of Combustion and Explosion* **15,** 43 (1979).

18) L. Topor and O. J. Kleppa, *Metallurgical Transactions A.* **17A,** 1217 (1986).

19) Thermodynamic Properties of Individual Substances. Fourth Edition. Editors: L. V. Gurvich, I. V. Veyis, C. B. Alcock and V. S. Iorish, Vol.I, Vol.II, Hemisphere Publ. Corp., New York (1989-1990); Vol.III, CRC Press & Begell House, Boca Raton (1994).

20) Termicheskiye konstanti vecshestv. Reference book in I O volumes. (Editors: V. A. Medvedev *et al*). Moscow, VINITI (1967-1982).