

A Program for Linear Regression with a Common Point of Intersection: The Isokinetic Relationship

Carole Ouvrard and Michel Berthelot*

Laboratoire de Spectrochimie, Faculté des Sciences et des Techniques, 2 rue de la Houssinière,
BP 92208 - 44322 Nantes, Cedex 3, France

Thierry Lamer

Laboratoire de Génie des procédés, CRTT, Bd de l'Université, BP 406-44602 Saint-Nazaire, Cedex, France

Otto Exner

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic,
CZ-16610 Praha 6, Czech Republic

Received July 11, 2000

A program is described for statistically correct calculations of the enthalpy–entropy relationship, applicable also generally for linear regression when several regression lines are constrained by a common point of intersection. Examples from the chemical literature show the difference between a correct and incorrect statistical treatment.

INTRODUCTION

The problem of several regression lines intersecting in one point does not appear in common textbooks of statistics but has been encountered several times in treatment of chemical data. The most popular case is the so-called isokinetic relationship¹ (IKR). It may be presented as an empirical relationship between enthalpies and entropies in a series of related reactions and called compensation effect;¹ in another connection it was called the Meyer-Neldel Rule.² The matter has been reviewed in the fields of homogeneous kinetics,^{3,4} heterogeneous catalysis,⁵ macromolecular chemistry,² and biochemistry.⁶ As an enthalpy–entropy relationship, IKR is represented by eq 1 in which the slope β is called isokinetic temperature.

$$\delta\Delta H = \beta\delta\Delta S \quad (1)$$

Eq 1 may apply to kinetics or to thermodynamics: symbols ΔH and ΔS mean either ΔH^\ddagger and ΔS^\ddagger or ΔH° and ΔS° ; δ denotes the difference between an individual item (compound, reaction) and the reference standard. However, eq 1 cannot be used for regression⁷ since the two variables are a priori dependent.^{3,8–12} Several correct statistical solutions,^{10–13} quick graphical tests,⁹ and an approximate procedure¹⁴ were advanced. Warnings against using eq 1 as a regression were repeated many times.^{1,3,8–10} Despite this, innumerable statistically incorrect papers appeared.^{15,16} This occurred repeatedly after these solutions had been published,¹⁶ even recently and in prominent journals.^{16g,h} Many of these papers reached conclusions completely at variance with the original experimental facts. In our opinion, further proliferation of the error should be stopped, and the most promising way is by

providing a versatile computer program for a statistically correct procedure. This is attempted in the present note.

The most straightforward of the solutions described¹⁰ transfers the regression into a priori independent coordinates $\log k$ and T^{-1} ; eq 1 is then mathematically equivalent to eq 2 which describes several regression lines with the slopes b_i intersecting in a common point of intersection (β^{-1} , y°), see Figure 1A.

$$\log k_{ij} = y^\circ + b_i(T_{ij}^{-1} - \beta^{-1}) \quad (2)$$

In eq 2, the i th line of the total of l lines is determined by m_i points with the coordinates $\log k_{ij}$, T_{ij}^{-1} , where $j = 1, 2, \dots, m_i$. There are altogether N points, $N = \sum m_i$. The symbol k relates either to kinetics (rate constant k) or to thermodynamics (equilibrium constant K) and T is the temperature in Kelvin. The same problem of a common point of intersection appeared in several other applications:¹⁸ usually a statistical treatment was not dealt with, or the authors were not aware that a nontrivial statistical problem was encountered.¹⁹

The statistical solution of eq 2 was reported¹⁰ within the framework of the least-squares method. One of the present authors had a program for the Hewlett-Packard calculator 9820.⁴ At least four groups of authors²⁰ have written programs based on this solution, but these have not been published;²¹ even recently, only a graphical solution was presented without a mathematical treatment.²² The program described below is suitable for IKR (eq 2, coordinates T^{-1} , $\log k$) but can be used also quite generally when the regression lines in the coordinates x , y intersect in one point (x° , y°), eq 3.

$$y_{ij} = y^\circ + b_i(x_{ij} - x^\circ) \quad (3)$$

* Corresponding author phone: (33) 2 51 12 54 26; fax: (33) 2 51 12 54 12; e-mail: Michel.Berthelot@chimie.univ-nantes.fr.

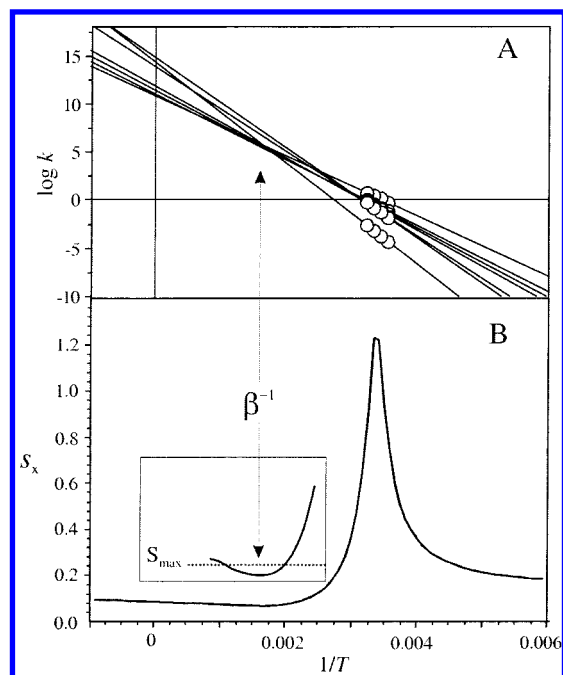


Figure 1. Graphic representation of IKR. A: Arrhenius lines and experimental points in the coordinates $\log_{10} k$ vs reciprocal temperature $1/T$, data from ref 17; B: dependence of the fitting error s_x on the assumed isokinetic temperature (plotted as $1/\beta$), the enlargement (in the direction of the y axis) shows the optimum value of β and its confidence interval.

An alternative statistical treatment¹³ is based on a special regression in the coordinates ΔG and ΔH . Its great defect is that it is applicable only to special cases where all reactions were followed at the same set of temperatures. In these very regular cases, it gives the same results as our solution. As far as we know, no program for this treatment is available.

DESCRIPTION OF THE PROGRAM

The program is based on the least-squares condition for all of the N experimental points taken with the same weight, eq 4. It works by subsequent approximations according to the equations developed^{10b} whether the pattern of experimental data is regular or not. (In the former case, "the special case", an explicit solution is possible^{10a}). The necessary equations were published.^{10b}

$$\sum_{ij} (\log k_{ij} - \log k_{ij(\text{calc})})^2 = \min \quad (4)$$

The program is implemented on a Microsoft EXCEL97 spreadsheet. It uses dedicated macros written with the VBA utilities and the solver, a nonlinear least-squares routine. It can be downloaded by Internet together with the directions for use²³ and copied in any version of EXCEL97 or later. A testing example is available at the same address. The operator chooses between the following possibilities: IKR in thermodynamics, or IKR in kinetics, or a regression in arbitrary variables. Then he introduces data (T in K or t in $^{\circ}\text{C}$; $\log k$ or k , or $\log K$ or K) successively according to the l individual regression lines. The program gives first the result of regressions along the unconstrained lines (Figure 1A, fitting error expressed by the standard deviation s_{00}) and calculates the values of ΔH and ΔS . Then it proceeds to IKR and draws a dependence of the fitting error s_x on the assumed value of

β^{-1} (Figure 1B). The operator suggests an abscissa (T^{-1}) near the anticipated minimum from which the optimization procedure will start: in this way finding the absolute minimum is assured. Then, the optimum value of β is found and regressions through the common point of intersection are carried out (fitting error: s_0). The confidence interval of β (at a confidence level α) is defined by the highest acceptable value of the fitting error s_{max} calculated by a F -test with respect to N and l (eq 5).

$$s_{\text{max}}^2 = s_0^2 \left[1 + \frac{1}{N-l-1} F_{1,N-l-1}(\alpha) \right] \quad (5)$$

Once the s_{max} value is determined, the confidence interval of β is given by the two points of intersection with the curve as shown in the enlargement of Figure 1B. When the accuracy of the IKR is low, the confidence interval of β can extend even to infinity and over it by changing sign. This would mean that all values of β are acceptable except a narrow interval (near the experimental temperatures). In such cases, IKR will be either found invalid or isoenthalpic.

In the next step, the most important quantities are summarized: standard deviation of fitting for different models (unconstrained lines s_{00} , IKR s_0 , isoenthalpic IKR s_{enth} , isentropic IKR s_{entr}). A F -test, shows whether s_{00} is significantly smaller than s_0 , see eq 6 where S_0 and S_{00} stand for the sum of the squares corresponding to s_0 and s_{00} , respectively.

$$F = \frac{S_0 - S_{00}}{S_{00}} \times \frac{N-2l}{l-2} \quad (6)$$

Another problem is whether IKR is acceptable as an empirical relationship of an approximate validity. This can be estimated from the ψ -value²⁴ (eq 7)

$$\psi = \frac{s_0}{s_{\text{data}}} \quad (7)$$

In eq 7, s_{data} is the standard deviation of all the $\log k$ values from their mean. The scale of ψ is purely empirical, and no particular distribution is assumed: the values $\psi < 0.1$ are acceptable for a good relationship approximation.²⁴ Ultimately, reaction enthalpies and entropies are calculated both for the unconstrained Arrhenius lines (ΔH^{\ddagger} and ΔS^{\ddagger}) and for IKR (so-called isokinetic parameters^{10a} ($\Delta H_{\text{iso}}^{\ddagger}$ and $\Delta S_{\text{iso}}^{\ddagger}$)).

EXAMPLES

The following examples should show only how the program works and serve for its testing. They also prove that the procedure criticized, plotting ΔH vs ΔS , yields completely wrong results. However this has been proven in many cases.^{3,4,8,13,14} The chemical problems connected with the following examples were discussed in the original literature.^{16c,e,17}

In an example from the older literature,¹⁷ recombination rates of seven *ortho*-substituted triarylmethyl cations with OH^- were measured at the temperatures between 10 and 40 $^{\circ}\text{C}$. On the basis of wrong statistics (eq 1), IKR with apparent isokinetic temperature $\beta_{\text{app}} = 330$ K was claimed when one substituent was eliminated. Figure 1A shows that the

Arrhenius lines intersect only quite approximately in one point. In our calculations, no substituent was eliminated. IKR with $s_0 = 0.052$ was found and accepted as an approximate relationship ($\psi = 0.044$), but it is not valid within experimental uncertainty ($F = 124.5$; rejected), also for the reason that the experimental precision was unusually high ($s_{00} = 0.009$). Isokinetic temperature $\beta = 599$ K was determined with a relatively great uncertainty, but the confidence interval (488–1021 K, Figure 1B) does not include the isoentropic relationship ($F = 191.6$).

In a more recent example, reaction rates of cis-to-trans isomerization of 10 substituted azobenzenes were measured^{16c} at temperatures between 15 and 60 °C. From a plot of ΔH^\ddagger vs ΔS^\ddagger the apparent isokinetic temperature $\beta_{\text{app}} = 306$ was estimated. According to our treatment, IKR does not exist ($s_{00} = 0.017$; $s_0 = 0.075$; $F = 51.3$; rejected, $\psi = 0.106$; rejected). With the same precision, the relationship could be also isoenthalpic ($s_{\text{enth}} = 0.073$) or isoentropic ($s_{\text{entr}} = 0.078$). No real confidence interval for β can be determined; it would be from $-\infty$ to 192 K and from 1143 K to $+\infty$.

In contrast, the third example shows that even wrong statistics can yield some acceptable results when the accuracy of both IKR and the data is high. Kinetics of thermolysis of six aromatic hyponitrites were measured at temperatures between 43 and 65 °C and IKR with $\beta = 301$ K was claimed.^{16e} Our results confirmed a valid IKR within experimental uncertainty, $s_{00} = 0.012$; $s_0 = 0.015$; $F = 4.19$; accepted; $\psi = 0.031$, accepted. The optimum value of β is 291 K, the confidence interval is from 256 to 304, and it includes also the value determined by the original authors.^{16e}

ACKNOWLEDGMENT

One of the authors (O.E.) thanks for a grant to the Grant Agency of the Czech Republic (Grant No. 203/99/1454).

REFERENCES AND NOTES

- Exner, O. The Enthalpy–Entropy Relationship. *Prog. Phys. Org. Chem.* **1973**, *10*, 411–482.
- Yelon, A.; Movaghar, B.; Branz, H. M. Origin and Consequences of the Compensation (Meyer-Neldel) Rule. *Phys. Rev., Sect. B* **1984**, *46*, 12244–12249.
- Linert, W. Mechanistic and Structural Investigations Based on the Isokinetic Relationship. *Chem. Soc. Rev.* **1994**, *23*, 429–438.
- Exner, O. The Enthalpy–Entropy Relationship in Organic Reactions. *Collect. Czech. Chem. Commun.* **1975**, *40*, 2762–2780.
- Galwey, A. K. Compensation Effect in Heterogeneous Catalysis. *Adv. Catalysis* **1977**, *26*, 247–322.
- Lumry, R.; Rajender, S. Enthalpy–Entropy Compensation Phenomena in Water Solutions of Proteins and Small Molecules: a Ubiquitous Property of Water. *Biopolymers* **1970**, *9*, 1125–1227.
- The enthalpy–entropy relationships belong into this category when ΔH and ΔS have been determined together from measurements at a variable temperature. They may concern reaction parameters ΔH° and ΔS° or activation parameters ΔH^\ddagger and ΔS^\ddagger . The cases when ΔH° was obtained separately by calorimetry represent another statistical problem; both cases were recently reviewed and distinguished: Exner, O. How to Get Wrong Results from Good Experimental Data: a Survey of Incorrect Applications of Regression. *J. Phys. Org. Chem.* **1997**, *10*, 797–813.
- S. Wold, S. The Abnormal Behaviour of Enthalpy–Entropy Plots. *Chem. Scr.* **1972**, *2*, 145–147.
- Exner, O. On the Enthalpy–Entropy Relationship. *Collect. Czech. Chem. Commun.* **1964**, *29*, 1094–1113.
- (a) Exner, O. Statistics of the Enthalpy–Entropy Relationship. I. The Special Case. *Collect. Czech. Chem. Commun.* **1972**, *37*, 1425–1444. (b) Exner, O.; Beránek, V. Statistics of the Enthalpy–Entropy Relationship. II. The General Case. *Collect. Czech. Chem. Commun.* **1973**, *38*, 781–798.
- Shimulis, V. I. Metod diskriminantsii gipotezy o lineinom kompensionnom effekte. *Kinetika Kataliz* **1983**, *24*, 715–719.
- Chen, H. J.; Yuan, S. P. On Concurrence of Several Linear Regressions and Applications to Problems of Enzyme Kinetics. *Commun. Statist.-Theor. Meth.* **1982**, *11*, 395–409.
- Krug, R. R.; Hunter, W. G.; Grieger, R. A. Enthalpy–Entropy Compensation. 1. Some Fundamental Statistical Problems Associated with the van't Hoff and Arrhenius Data. *J. Phys. Chem.* **1976**, *80*, 2335–2341. Krug, R. R.; Hunter, W. G.; Grieger, R. A. Enthalpy–Entropy Compensation. 2. Separation of the Chemical from the Statistical Effect. *J. Phys. Chem.* **1976**, *80*, 2341–2351.
- Linert, W. J. Thermodynamic Implications of Substituent and Solvent Effects on Reactivity. *Chem. Inf. Comput. Sci.* **1992**, *32*, 221–226.
- For instance: Leffler, J. E. The Enthalpy–Entropy Effect. *J. Org. Chem.* **1955**, *20*, 1202–1231. Brown, R. F. The Enthalpy–Entropy Relationship and Its Implications for Organic Chemistry. *J. Org. Chem.* **1962**, *27*, 3015–3026. Beltrame, P.; Simonetta, M. Empirical Relation between Energy and Entropy of Activation. *Gazz. Chim. Ital.* **1959**, *89*, 495–501.
- For instance: (a) Gnanadoss, L. M.; Kalaivani, D. Mechanism and Linear Free Energy Relationships in the Kinetics of Formation of Bicyclo[3.3.1]nonane Derivatives from 1,3,5-Trinitrobenzene, Benzoyl-acetones, and Base. *J. Org. Chem.* **1985**, *50*, 1178–1182. (b) Kim, S. S.; Choi, S. Y.; Kang, C. H. The Reality of Polar Transition States in Photobromination of Toluene by N–Bromosuccinimide (NBS). *J. Am. Chem. Soc.* **1985**, *107*, 4234–4237. (c) Asano, T.; Okada, T.; Shinkai, S.; Shigematsu, K.; Kusano, Y.; Manabe, O. Temperature and Pressure Dependence of Thermal Cis-to-Trans Isomerization of Azobenzenes which Evidence an Inversion Mechanism. *J. Am. Chem. Soc.* **1981**, *103*, 5161–5165. (d) Dorovska-Taran, V.; Momtcheva, R.; Gulubova, N.; Martinek, K. The Specificity in the Elementary Steps of α -Chymotrypsin Catalysis. A Temperature Study with a Series of N-Acetyl-L-amino acid Methyl Esters. *Biochim. Biophys. Acta* **1982**, *702*, 37–53. (e) Quinga, E. M. Y.; Mendenhall, G. D. Substituent Effects and Isokinetic Relations in the Homolysis of Hyponitrite Esters. The Question of Free-Radical Homoconjugation. *J. Org. Chem.* **1985**, *50*, 2836–2840. (f) Palm, V.; Kaasik, S. Reinvestigation of the Substituent, Environmental and Temperature Effects on the Solvolysis Kinetics of Disubstituted Mannich Bases of the General Structure 3-X¹C₆H₄-NHCH₂C(NO₂)₂C₆H₄X²-3. 2. Quantitative Interpretation of Combined Substituent, Solvent and Temperature Effects. *Org. Reactivity (Tartu)* **1996**, *30*, 73–83. (g) Gallicchio, E.; Kubo, M. M.; Levy, R. M. Entropy–Enthalpy Compensation in Solvation and Ligand Binding: a Revision. *J. Am. Chem. Soc.* **1998**, *120*, 4526–4527. (h) Losi, A.; Wegener, A. A.; Engelhard, M.; Bralavsky, S. E. Enthalpy–Entropy Compensation in a Photocycle: the K-to-L Transition in Sensory Rhodopsin II from *Natronobacterium Pharaonis*. *J. Am. Chem. Soc.* **2001**, *123*, 1766–1777.
- Idlis, G. S.; Ginzburg, O. F. Effect of *ortho* Substituents on the Reactivity of Malachite Green Type Dyes. *Reakts. Sposobnost Org. Soedin. (Tartu)* **1965**, *2* (4), 54–60.
- For instance: Johnsen, R. M. On the Treatment of Osmotic Pressure Data. *Chem. Scr.* **1971**, *1*, 149–153. Eley, D. D. Energy Gap and Preexponential Factor in Dark Conduction by Organic Semiconductors. *J. Polymer Sci., Part C* **1967**, *73*–81. Thorn, R. J. Correlation of Enthalpy and Entropy in Sufficiently Equivalent Systematic Errors: Vaporization of Uranium. *J. Chem. Phys.* **1970**, *52*, 474–476.
- (a) Cserháti, T.; Magyar, K. Relationship between the Hydrophobic and Hydrophilic Molecular Parameters of Some Monoamine Oxidase Inhibitory Drugs, Determined by means of Adsorptive and Reversed-Phase Thin-Layer Chromatography. *J. Biochem. Biophys. Methods* **1992**, *24*, 249–264. (b) Vogel, S.; Knoche, W.; Schoeller, W. W. Protonation of 1,3,5-Tripyrrolidinobenzene in Basic Aqueous Solution. Thermodynamics and Kinetics of the Formation of a Stable σ -Complex. *J. Chem. Soc., Perkin Trans. 2* **1986**, 769–772. (c) Armentrout, P. B. Entropy measurements and the Kinetic Method: a statistically Meaningful Approach. *J. Am. Mass Spectrom.* **2000**, *11*, 371–379.
- (a) Müller, P.; Perlberger, J.-C. Isokinetic Relationships in Alcohol Oxidation with Chromic Acid. *Helv. Chim. Acta* **1974**, *57*, 1943–1950. (b) Pytel, O.; Večeřa, M.; Vetešník, P. Calculation of Isokinetic Temperature of Noncatalyzed Hydrolysis of Substituted 3-(N-Methyl-carbamoyl)-1,3-diphenyltriazenes. *Collect. Czech. Chem. Commun.* **1981**, *46*, 898–905. (c) Eyer, G. N.; Cañizo, A. I.; Mateo, C. M.; Alvarez, E. E.; Cafferata, L. F. R. *J. Org. Chem.* **1999**, *64*, 8457–8460. (d) Tarasow, T. M.; Lewis, C.; Hilvert, D. Investigation of Medium Effects in a Family of Decarboxylase Antibodies. *J. Am. Chem. Soc.* **1994**, *116*, 7959–7970.
- It is true that eq 2 or 3 can be treated by any program for nonlinear regression, using a matrix of $2l$ exploratory variables: l variables T_{ij}^{-1} and l dummy variables z_i (equal to 0 or 1). Optimum values of β and b_i can be obtained in this way. Since the parameter β is of particular

importance we prefer the treatment in this paper, following possible solutions for all values of β (Figure 1B). Another merit of our program is that it gives all necessary quantities in one package, including particularly the graphic representation as in Figure 1 which is very important.

- (22) Fernandez, A.; Reyes, C.; Lee, T. Y.; Prock, A.; Giering, W. P.; Haar, C. M.; Nolan, S. P. Assessing the Stereoelectronic Properties of Pyrrolyl Phosphines and Related Ligands. The Quantitative Analysis

of Ligand Effects (QALE). *J. Chem. Soc., Perkin Trans.* **2000**, 2, 1349–1357.

- (23) <http://sciences.univ-nantes.fr/spectro/publicat.htm>.

- (24) Exner, O. Additive Physical Properties. I. General Relationships and Problems of Statistical Nature. *Collect. Czech. Chem. Commun.* **1966**, 31, 3222–3251.

CI000089S