Journal of Thermal Analysis (1985) 831–840

SOME ASPECTS OF MATHEMATICAL STATISTICS AS APPLIED TO NONISOTHERMAL KINETICS

Part I. Regression and correlation analysis

S. V. Vyazovkin and *A. I. Lesnikovich*

INSTITUTE OF PHYSICO-CHEMICAL PROBLEMS, MINSK, USSR

(Received December 3, 1984)

Restrictions of the simplest use of correlation and regression analysis to obtain a single-valued solution to the inverse kinetic problem are considered. The Coats–Redfern method is suggested as a version of nonlinear regression analysis to increase the unambiguity of the solution.

Determination of kinetic parameters from the data of nonisothermal experiments is among the most difficult kinetics problems. Its solution without independent data is generally ambiguous. This ambiguity stems from the fact that the calculation methods as applied to the data on the kinetics of a chemical process with an unknown mechanism result in a great number of kinetic functions and parameters which describe the process equally well. A "single" solution is often obtained from semi-quantitative subjective considerations. It might seem possible to achieve nonambiguity by applying appropriate statistical criteria to estimate the values underlying the choice of kinetic characteristics.

However, the case is much more difficult, as follows from the incomplete statistical analysis of experimental data performed by a majority of investigators. This is probably attributable to the fact that statistical methods do not generally provide the desired unambiguity. Nevertheless, it is unreasonable to renounce the correct statistical estimation of the results as a means of attaining objective unambiguity. We believe that further improvement of the methods of kinetic parameter calculations requires more careful consideration of the statistical aspects of these methods, which is undertaken in this series. The present paper deals with the analysis of individual reported results to show the importance of statistical approaches in nonisothermal kinetics to provide objective characteristics of these results and unambiguous kinetic parameters.

Let us consider some of the well-known methods of calculating kinetic parameters, as suggested by Coats–Redfern [1], Satava [2] and Zsakó [3].

John Wiley & Sons, Limited, Chichester
Akadémiai Kiadó, Budapest

The first two apply correlation analysis to the linear form of the constitutive equation. The kinetic function chosen is generally based on the correlation coefficient.

However, the choice of the kinetic function based on the maximum correlation coefficient is not always statistically well grounded, since the maximum correlation coefficient may differ insignificantly from its lower values. It seems necessary, therefore, to analyse the significance of differences in the correlation coefficients. As the sample correlation coefficient distribution with a small number of samples (less than 30) is of a complex form greatly differing from the normal one [4], the significance of two sample correlation coefficients may be estimated by taking into account the Fischer normalization transform:

$$Z = 0.51n\frac{1+R}{1-R} = \mathrm{arthR} \qquad (1)$$

Comparing

$$T = \frac{|\mathrm{arth}R_1 - \mathrm{arth}R_2|}{\sqrt{\frac{1}{n_1-3}+\frac{1}{n_2-3}}} \qquad (2)$$

(where n_1 and n_2 are the sample volumes, and R_1 and R_2 the corresponding correlation coefficients) with the normalized Laplace function parameter for the chosen significance level (U_p), we may estimate whether the difference between R_1 and R_2 is significant [5].

Considering (1) and (2), let us estimate the significance of R for the data of [6] dealing with the thermal decomposition of $Mg(OH)_2$. The decomposition has been described by 17 kinetic functions g(α), with R between 0.980 and 0.999. The authors chose the mechanism corresonding to $R = 0.999$. In Table 1, compiled from the data

Table 1 Kinetic parameters and functions $Mg(OH)_2$ insignificantly differing in the R value

$g(\alpha)$	$E, \frac{\text{kcal}}{\text{mol}}$	A, s^{-1}	R
$-\ln(1-\alpha)$	50.338	$1.3*10^{14}$	0.997
$(-\ln(1-\alpha))^{\frac{1}{1.5}}$	32.698	$1.4*10^{8}$	0.997
$(-\ln(1-\alpha))^{\frac{1}{2}}$	23.878	$1.2*10^{5}$	0.997
$(-\ln(1-\alpha))^{\frac{1}{3}}$	15.157	$8.9*10^{1}$	0.997
$(-\ln(1-\alpha))^{\frac{1}{4}}$	10.645	<1	0.997
$(1-\alpha)^{-1}-1$	60.716	$6.3*10^{17}$	0.999
$(1+\alpha)^{-\frac{1}{4}}+1$	55.281	$3.7*10^{15}$	0.999

of [6], the kinetic functions are presented along with corresponding kinetic parameters which differ insignificantly in the value of R, in accordance with our calculations by (2). It is important to note that, with insignificantly different R, the kinetic parameters A and E vary widely. By analysing in terms of (2) the correlation coefficients obtained in [6] to calculate the kinetic parameters using the Satava method, we have arrived at similar results. Thus, when (1) and (2) are used, correlation analysis of the data of [6] does not provide statistical grounds why the kinetic functions yielding $R=0.999$ should prevail over those yielding $R=0.997$.

Zsakó has suggested a method for calculating kinetic parameters and choosing kinetic functions based on the minimum standard deviation:

$$\delta = \sqrt{\frac{\Sigma(B_i - \bar{B})^2}{r}} \tag{3}$$

where $B = \log g(\alpha) - \log p(x)$; $g(\alpha)$ is the integral form of the function $p(x) = \frac{e^{-x}}{x} - \int_{+x}^{\infty} \frac{e^{-u}}{u} \mathrm{d}U$; and r is the number of experimental points taken to calculate $\bar{B}$. However, similarly to the previous case, it is not sufficient to find only the minimum value responsible for the choice of $g(\alpha)$.

It is necessary to estimate the statistical significance of the differences in δ against $\delta_{\min}$. We have used the Fischer criterion (F_p) to estimate the significance of the differences in δ from [3]. In all cases, $\delta^2/\delta^2_{\min}$ did not exceed the critical value of F_p. Therefore, the choice of the kinetic function $f(\alpha) = 1 - \alpha$ and the activation energy of 28.9 kcal/mole exemplified in [3] cannot be regarded as statistically well grounded.

In view of the above, it is clear that the estimation of the kinetic parameters and the choice of the kinetic function require assessment of the statistical significance of the values underlying the choice. The procedure of the significance assessment completely excludes subjectivity when choosing the kinetic function. However, such a choice is generally ambiguous. We believe that in a situation such as this, the kinetic parameters might be determined with the use of methods which do not require a knowledge of the kinetic function, e.g. the estimation of invariant Arrhenius parameters [7].

The problem of an unambiguous choice of kinetic functions has two main aspects. First, the difference between various kinetic functions describing a certain process does not often exceed the experimental error. Second, the majority of the methods of kinetic parameter calculation rely on a low-sensitive computational procedure of linearizing different modifications of the basic nonisothermal kinetics equation. To be specific, it is shown in [8] that the linear form of the basic equation suggested by Achar [9] does not permit distinction between Avrami–Erofeev kinetic functions (A_m) with different m values. We have arrived at similar conclusions (cf., for example, Table 1) through assessment of the statistical significance of the values underlying the choice of the kinetic functions in the calculations with linearization.

Renouncing the linearization of equations for kinetic parameters may increase the sensitivity of the calculation method to the form of the kinetic function. It should be noted that linearization, though yielding simple and convenient mathematical expressions, is not a necessary condition for estimating kinetic parameters and choosing kinetic functions. In this connection it is worth considering the potentialities of nonlinear regression analysis for an unambiguous choice of kinetic functions in nonisothermal kinetics.

One of the possible nonlinear approaches to the choice of kinetic functions is suggested in [10]. It implies determination of the minimum standard deviation of the experimental degrees of decomposition from the one calculated from Eq. (4):

$$\hat{\alpha}_1 = A + B(\alpha_2) + C(\alpha_2)^2 + D(\alpha_2)^3 \tag{4}$$

which occurs under certain conditions. α_1 and α_2 are the degrees of decompositions at appropriate heating rates; A, B and C are constants dependent on the form of the kinetic function and the heating rate ratio.

The choice of the kinetic function is based on the minimum standard deviation of the experimental α_1 from $\hat{\alpha}_1$ predicted with (4).

It is to be noted that in this case too the minimum standard deviation may not always differ significantly from the nearest value. Further, if the experimental data do not correspond to any of the kinetic functions of the program, some inadequate kinetic function will correspond to the minimum standard deviation. Thus, when the approach in [10] is used, the adequacy of the chosen kinetic function must be estimated, e.g. by comparing the minimum standard deviation of $g(\alpha)$ with its deviation due to accidental error in the α measurement. This value is the extrinsic dispersion estimate. In this case, F_p may be regarded as an adequacy criterion.

We think it reasonable to use nonlinear versions of kinetic parameter calculation methods in nonisothermal kinetics problems. To be specific, nonlinear regression analysis as applied to the Coats–Redfern approach with the linear Eq. (1):

$$\ln\left(\frac{1-(1-\alpha)^{1-n}}{T^2(1-n)}\right) = \ln\frac{AR}{\beta E}\left(1-\frac{2RT}{E}\right) - \frac{E}{RT} \tag{5}$$

is presented in what follows.

Write the degree of decomposition as a function of temperature:

$$\alpha_i^* = \varphi(T_i) \tag{6}$$

The most exact values of the kinetic parameters A and E may then be found from the condition of the minimum value:

$$S^2 = \frac{1}{K-2}\sum_{i=1}^{K}(\alpha_i - \alpha_i^*)^2 \tag{7}$$

where α_i is the experimental degree of decomposition

Minimum requirement (7) yields the system of nonlinear equations

$$\begin{cases} \dfrac{\partial S^2}{\partial E} = 0 \\ \\ \dfrac{\partial S^2}{\partial A} = 0 \\ \\ \dfrac{\partial S^2}{\partial n} = 0 \end{cases} \tag{8}$$

which may be solved numerically by the gradient method [11].

The calculation below shows that the S^2 dependence on any of the three parameters A, E and n, the other two being constant, is fitted rather well (around the minimum) by some parabole (Fig. 1). The value S_0^2 is the boundary below which the differences between the residual sum of squares are not significant:

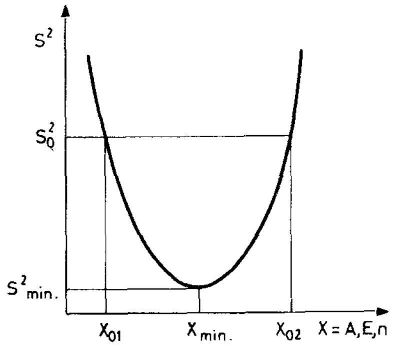

Fig. 1 Schematic illustration of the dependence of the residual sum of squares from any of kinetic parameters (A, E, n)

$$S_0^2 = S_{\min}^2 F_p \tag{9}$$

where F_p is the Fischer criterion.

Following (9), the variation of A, E and n from X_{01} to X_{02} does not involve significant differences of S^2 against $S_{\min}^2$, and hence the values

$X_{02}-X_{\min} = X_{\min}-X_{01}$ are confidence intervals for A, E and n. For example, $E_{\min}-E_{01} = E_{02}-E_{\min} = E \pm \Delta E$ (the same for A and n).

If kinetic functions different from $f(\alpha) = (1-\alpha)^n$ are used, expression (7) becomes:

$$S^2 = \frac{1}{K-2} \sum_{i=1}^{K} (g(\alpha_i) - g^*(\alpha_i))^2 \tag{10}$$

where $g(\alpha)$ is the integral form of the kinetic function. The value of $S^2_{\min}$ is chosen from the set of solutions to the nonlinear equations of the form:

$$\begin{cases} \dfrac{\partial S^2}{\partial E} = 0 \\ \\ \dfrac{\partial S^2}{\partial A} = 0 \end{cases} \tag{11}$$

for each $g(\alpha)$ function. The further procedure of choosing the function $g(\alpha)$ is similar to the choice of the reaction order in Eq. (8).

The significance of the residual sum of squares (10) may be estimated by comparing it with the deviation of the $g(\alpha)$ value due to accidental error in the α measurements and then estimating the adequacy through F_p.

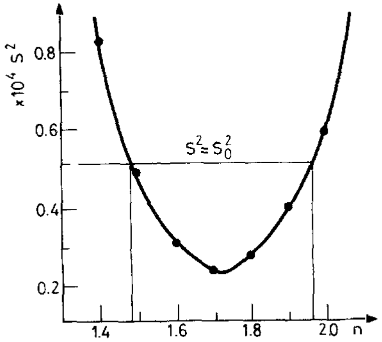

Fig. 2 Dependence of the residual sum of squares from the reaction order value in the case of $Mg(OH)_2$

Returning to the data of [6], find the optimum values of the kinetic parameters and the reaction order using the above approach. The change in the residual sum of squares vs. the reaction order is shown in Fig. 2. The minimum (7) as found by the gradient method is $0.2425 \cdot 10^{-4}$ at $E = 57.706$ kcal/mol, $\log A = 18.441$ and $n = 1.725$.

Confidence intervals for these values may be found in accordance with (9): $S_0^2 = 0.5141 \cdot 10^{-4}$. The projections of the intersection points for the line $S^2 = S_0^2$

and the parabola onto the abscissa represent the boundary values of the reaction order (Fig. 2). Similar relationship hold for S^2 *vs.* and log A (Fig. 3). The final values will be $n = 1.73 \pm 0.24$; $E = 57.71 \pm 2.53$ kcal/mol; log $A = 18.44 \pm 0.89$ (the calculation is made for Trace 1 [6]).

It should be noted that, for the linear form of the Coats–Redfern equation, the reaction orders 1.5 and 2.0 yield equal correlation coefficients (Table 1). For nonlinear regression analysis in accordance with (7), the residual-to-minimum squared sum ratio is equal to 2.05 for $n = 1.5$, which is less than the critical value for 19 degrees of freedom (2.17), while it is in excess of it for $n = 2.0$ (2.47). The kinetic function with $n = 2.0$ must be excluded, therefore, as it does not adequately describe decomposition. The comprehensive interpretation of the process using $n = 1.5$ will be grounded statistically, since this value of n lies within the confidence interval of its formal values $n = 1.73 \pm 0.24$. The predicted kinetic parameters are compared with experimental data in Table 2. The mean relative deviation from the experimental value is 2.27%. Thus, nonlinear regression analysis as applied to the data of [6] enables us to choose one of the two kinetic functions consistent with equal linear correlation coefficients. This points to the higher sensitivity of the nonlinear method to the form of the kinetic function as compared to the linear approach.

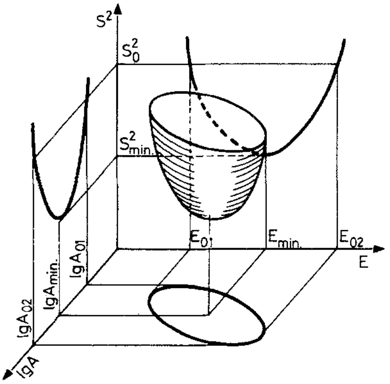

Fig. 3 Grafical illustration of the equations system (8)

Approximation of the curve in Fig. 2 by the parabola

$$S^2 = An^2 + Bn + C \tag{12}$$

is a less accurate, though a more simple way, to find its minimum against the gradient method.

Table 2 Values of the experimental degree of decomposition (α) as in [6], of the theoretical degree of decomposition (α^*) and of the relative deviation for $Mg(OH)_2$

α	α^*	$\frac{(\alpha-\alpha^*)}{\alpha}\cdot 100$, %
.0251	.0268	6.77
.0342	.0364	6.43
.0478	.0456	4.60
.0592	.0569	3.89
.0797	.0759	4.77
.0957	.0936	2.19
.1162	.1147	1.29
.1435	.1398	2.58
.1777	.1799	1.24
.2141	.2153	.59
.2620	.2698	2.98
.2894	.3000	3.66
.3485	.3487	.06
.4100	.4009	2.22
.4715	.4739	.51
.5308	.5299	.17
.5900	.5853	.80
.6538	.6561	.35
.7107	.7055	.73
.7563	.7507	.74
.7950	.8038	1.11

The values of the residual sum of squares and of the reaction order consistent with the minimum are the coordinates of its peak

$$n_{\min} = -\frac{B}{2A} \tag{13}$$

$$S^2_{\min} = \frac{2A \quad 4AC-B^2}{4A} \tag{14}$$

The confidence interval boundaries for n may be found by solving the quadratic equation (12), where $S^2 = S^2_{\min}\cdot F_p$:

$$n_{1,2} = n_{\min} \pm \frac{\sqrt{(4AC-B^2)(F_p-1)}}{2A} \tag{15}$$

The coefficients of (12) calculated using the least-squares method for the data of Table 2 are: $A = 0.00052$; $B = -0.00181$; $C = 0.00159$. The values of $n_{\min}$ and $S^2_{\min}$ predicted from (13) and (14) are 1.730 and $0.2399 \cdot 10^{-4}$, respectively. The confidence interval from (15) ranges between 1.50 and 1.96.

Conclusion

The wide use of statistical methods provides complete objectivity and higher unambiguity in solving inverse nonisothermal kinetics problems and in assessment of the accuracy of the results obtained. The accuracy is often overestimated, because use is made only of the data of linear regression analysis, applying one of the numerous linearization methods of nonisothermal kinetics. A more comprehensive statistical analysis allows conclusions of higher objective value concerning the confidence interval for kinetic parameters. Taking into account the inevitable formalism in describing complex processes such as the transformations of solids in terms of unique and simple kinetic functions, we may state that the real accuracy of the kinetic parameter values obtained by the methods of nonisothermal kinetics cannot be essentially higher than the accuracy found so far in the present work by confidential estimation of the kinetic parameters obtained for the global minimum using the Fischer criterion.

Further improvement of the statistical approaches in nonisothermal kinetics requires, in particular, that the error in argument $1/T$ measurements be allowed for.

The qualitative characteristics of the calculation methods applied in nonisothermal kinetics involve a very important statistical problem, to be considered in Part II.

References

1 A. W. Coats and J. P. Redfern, Nature, 201 (1964) 68.
2 V. Satava, Thermochim. Acta, 2 (1971) 423.
3 J. Zsakó, J. Phys. Chem., 72 (1968) 2406.
4 R. A. Fischer, Statistical Methods for Research Workers, Oliver and Boyd, London, 1954.
5 N. L. Johnson and F. C. Leone, Statistics and Experimental Design in Engineering and the Physical Science, vol. 1, J. Wiley, N.Y., 1977.
6 P. H. Fong and D. T. Y. Chen, Thermochim. Acta, 18 (1977) 273.
7 A. I. Lesnikovich and S. V. Levchik, J. Thermal. Anal., 27 (1983) 89.
8 T. B. Tang, Thermochim. Acta, 58 (1982) 373.
9 B. N. N. Achar, G. W. Brindley and J. H. Sharp, Proc. Int. Clay Conf., Ierusalem, 1 (1966) 67.
10 L. Reich and S. S. Stivala, Thermochim. Acta, 62 (1983) 129.
11 H. H. Rosenbrock and C. Storey, Computational Techniques for Chemical Engineers, Pergamon Press, N.Y., 1966.

Zusammenfassung — Die Beschränkungen der einfachsten Anwendung der Korrelations- und Regressionsanalyse zur eindeutigen Lösung des inversen kinetischen Problems werden erörtert. Die Coats–Redfern-Methode wird als eine Version der nichtlinearen Regressionsanalyse zur Erhöhung der Eindeutigkeit der Lösung vorgeschlagen.

Резюме — Рассмотрены ограничения наиболее простого использования корреляционного и регрессионного анализов для получения единственно правильного решения задачи обратной кинетики. Метод Коутса-Рэдферна предложен в качестве варианта нелинейного регрессионного анализа для увеличения однозначности решения.