ELSEVIER

Chemical Engineering and Processing 36 (1997) 243–249

Chemical Engineering and Processing

Statistical analysis of linear and nonlinear correlation of the Arrhenius equation constants

Neima Brauner [a,*], Mordechai Shacham [b]

[a] School of Engineering, Tel-Aviv University, Tel-Aviv, 69978, Israel
[b] Department of Chemical Engineering, Ben-Gurion University of the Negev, Beer-Sheva, 84105, Israel

Received 15 August 1996; accepted 24 October 1996

Abstract

Engineers must often use correlations that were developed before statistical analysis and verification of the correlation became a routine procedure. In this paper, we use modern statistical techniques to compare the traditional linear regression technique with the modern nonlinear regression as applied to the Arrhenius equation. The objective of the comparison is to determine whether there are basic flaws with the technique used in the past and whether these flaws may render the constants published in the literature untrustworthy.

It is concluded that linear regression, when applied to the Arrhenius expression, is in principle not inferior to nonlinear regression and if the relative error in the data is distributed normally, it can even be superior. Nevertheless, if insufficient data were used for calculation of the constants and/or the experimental data were interpolated or smoothed, the accuracy of the published correlation is unpredictable. © 1997 Elsevier Science S.A.

Keywords: Arrhenius; Linear; Nonlinear; Regression; Statistical analysis

1. Introduction

Consistent and accurate modeling and correlation of experimental data is essential in the era of computer-aided chemical process design. In the days when calculations were done by hand or with a calculator, inappropriate data could be immediately detected and discarded. But, nowadays, if the correlation is included inside a large, say, process simulation program, it is very difficult to detect inaccurate or meaningless results obtained in a particular range of temperature, pressure or composition.

Nowadays, it is customary to carry out a statistical analysis and verification of the model which is being fitted to the data. But, correlations that were developed more than 30 years ago are still being widely used. These correlations were developed in the slide rule and graph paper era, mostly without any statistical analysis. It is, therefore, very difficult to assess their accuracy. If the original data is still available, statistical analysis can provide this missing information. Such an analysis may sometimes indicate that the experimental data cannot justify the model used, see for example, Shacham et al. [1].

The original data is not always available to carry out the statistical analysis. On the other hand, a correlation cannot be discarded off hand, since repeating the experiments is expensive and time consuming. This dilemma can be partly resolved by analyzing the correlation techniques that were used in the past for a particular group of model equations and comparing its accuracy with the techniques that are being used today. If serious flaws in the old techniques are detected, then the model equations and parameter values obtained in the past probably cannot be trusted. But, if no such flaws are found, it is justified to use the old correlation.

We have carried out such an analysis for the Arrhenius equation. In the remainder of this section, some basic concepts relevant to the Arrhenius expression are described.

The Arrhenius equation for correlating the dependence of reaction rate coefficients on temperature is:

$$k = A\exp(-E/RT) \tag{1a}$$

---

* Corresponding author. Tel: +972 3 6408127; fax: +972 3 6429540.

PII S0255-2701(96)04186-4

where k is the rate coefficient, A is the frequency factor, E is activation energy, R is the ideal gas law constant, and T is the absolute temperature. The constants A and E are calculated by regression of experimental data.

Eq. (1a) possesses some undesired properties in the statistical sense (strong correlation between the parameters A and E) and in the numerical sense (nearly singular normal matrices). The following form of the equation is frequently used to overcome these difficulties (see for example Himmelblau [2], Bates and Watts [3] and Kittrell [8]):

$$k = A' \exp[-E/R(1/T - 1/T_0)] \qquad (1b)$$

where T_0 is the average absolute temperature and $A' = A \exp(-E/RT_0)$.

Eq. (1a) or Eq. (1b) can be very conveniently linearized:

$$\ln k = \ln A - E/RT \qquad (2)$$

so that the constants can be calculated using linear regression. This type of linearization, which allows fitting a straight line to the transformed experimental data, has been used for nearly a century (Chen and Aris [4]). Most of the constants reported in the literature were obtained using this method.

An alternative, is to use nonlinear regression techniques with Eq. (1a) or Eq. (1b). Nonlinear regression applied to Eq. (1a) leads to the following minimization problem:

$$\min_{A,E} S = \sum_{i=1}^{n} (k_i - A \exp[-E/RT_i])^2 \qquad (3)$$

where S is the sum of squares of the errors and n is the total number of data points of k vs T.

It should be mentioned that the parameters calculated using either linear and nonlinear regression are only approximate because of the experimental error in the observed k values. The notation used in statistics to differentiate between exact and approximate values will be introduced in Section 2 which deals with the error distribution in the Arrhenius expression.

In order to solve the minimization problem described in Eq. (3), either derivative based methods (such as the Marquardt method [5]) or nonlinear search algorithms (such as the Simplex method, [2]) can be used. Both types of methods may encounter difficulties because of the extreme nonlinearity of the Arrhenius equation. When a derivative based method is used, the matrix of partial derivatives may often turn to be nearly singular (as noted by Himmelblau, [2]). On the other hand, the surface $S(A,E)$, defined by Eq. (3), is often a narrow canyon with steep sides and a very long shallow bottom [4]. Such a shape makes the direct search methods very slow and inefficient. The parameter values obtained using linear regression can be used as initial estimates for nonlinear minimization alleviating the problems associated with using this method.

In this study, linear and nonlinear regression applied to the Arrhenius equations, are compared in order to evaluate the accuracy of data in the literature that was obtained using linear regression.

In the following section, some of the statistical concepts used in the comparison of the linear and nonlinear regression results are introduced. In Section 3 results of a recent study of the subject (Chen and Aris [4]) are analyzed. A large set of experimental data (from Wynkoop and Wilhelm [7]) is used in Section 4 for assessing the accuracy of linear and nonlinear regression results.

Most of the calculations reported in the paper were carried out using the linear and nonlinear regression program in the POLYMATH 3.0 package (Shaham and Cutlip [9]).

2. Error distribution in the linearized and nonlinear form of the Arrhenius equation

The statistical assumption behind the least squares error method for parameter estimation is that the measured value of the dependent variable has a deterministic part and a stochastic part. The stochastic part is often denoted by error (or experimental error), ϵ_i. Thus, Eq. (1a) can be rewritten:

$$k_i = A \exp(-E/RT_i) \pm \epsilon_i \qquad (4)$$

It is further assumed that the error is normally and independently distributed with zero mean and equal variances.

An infinite number of measurements is required in order to obtain the exact values of the parameters E and A. Since a sample always contains a finite number of measurements, the calculated parameters are always an approximation for the true values, and they are denoted with a circumflex. Thus, $\hat{E}$ and $\hat{A}$ are the calculated values of the parameters and $\hat{k}_i$ is the estimate for the dependent variable k_i.

A key indicator for a particular model to represent the data correctly is the error distribution. In order to determine the error distribution before the regression is carried out, replicate measurements (meaning several experimental k values at the same temperature) must be available. In most kinetic studies, no replicates are available. In such a case, inspection of the residuals, after carrying out the regression while using a particular model, provides a clue to whether the error distribution satisfies the underlying assumptions (Bates and Watts [3], p. 24).

When the error distribution is nonhomogeneous and no replicate measurements are available, transformations such as the power transformation proposed by Box and Hill [6], can be used to correct inhomogeneity of the variance. The transformation proposed by Box and Hill [6] is the following:

$$k_i^{\phi} = \begin{cases} \dfrac{k_i^{\phi} - 1}{\phi} & \phi \neq 0 \\ \ln k_i & \phi = 0 \end{cases} \qquad (5)$$

Using this transformation, a particular weight (W_i) is assigned to each k_i value (Box and Hill [6]):

$$W_i = \hat{k}_i^{(2\phi - 2)}$$

Adding a weighting factor to the objective function Eq. (3) and rewriting it in the notation introduced in this section yields:

$$\min_{A,E} S = \sum_{i=1}^{n} (k_i - \hat{A} \exp[-\hat{E}/RT_i])^2 \hat{k}_i^{(2\phi - 2)} \qquad (7)$$

Two cases are of special interest to us. For $\phi = 1$ the weighting factor is $\hat{k}_i^{(2\phi-2)} = 1$, and in this case the sum of squares of the absolute error is minimized. When $\phi = 0$, the transformation ln (thus linearization of the Arrhenius expression) is used. In this case, the weighting factor is $\hat{k}_i^{-2}$. Since the expression in the square brackets in Eq. (7) is an estimate of ϵ_i, the objective function amounts to minimizing $(\epsilon_i/\hat{k}_i)$^2, thus the relative error, $\epsilon_{r_i} = \epsilon_i/\hat{k}_i$.

It should be noted that the residual of the linearized Arrhenius equation (residual of ln k) is equivalent to the relative error, ϵ_{r_i} (this can be easily shown using Taylor series expansion of $\ln(1 + \epsilon_{r_i})$). Thus, the key for selecting between linear and nonlinear regression is observing the error distribution as reflected from the residual plots. When the experimental error, ϵ_i, is normally distributed, the parameters obtained using nonlinear regression should be preferred. But when the relative error is distributed normally, the linear regression results are more appropriate.

If the experimental data is not precise enough or there are not enough data points, there will not be a significant difference between the parameter sets obtained using linear and nonlinear regressions. In order for the parameter values to be significantly different the joint confidence regions (or joint likelihood contours) must be well separated for the two parameter sets. The calculation of the joint confidence region is discussed by Himmelblau [2] and Kittrell [8]. The method of calculation for the Arrhenius expression is shown in Appendix B. For nonlinear models, such as the Arrhenius expression, the joint confidence region is only approximate. A more accurate predictor of the confidence level of the parameter values is the joint likelihood contour which is described by Bates and Watts [3]. In this work we include results of both predictors.

3. Example 1. Synthesis of ethyl acetate

Recently, Chen and Aris [4] compared linear and nonlinear regression techniques using one set of data.

Table 1
Experimental data (Chen and Aris [4])

i	T_i (°C)	k_i
1	30	0.5
2	40	1.1
3	50	2.2
4	60	4.0
5	70	6.0

They concluded, based on their results, that the constants calculated by nonlinear regression are more accurate. Curl [10] concluded, using the same set of data, that linear regression yields more accurate constants.

The data used by Chen and Aris [4] for the comparison is shown in Table 1.

It is quite evident that the data shown in Table 1 is not experimental data. While the independent variable (temperature in this case) can often be set to round numbers (e.g. 60, 70), which are more convenient for calculations, it is quite unusual for the dependent variable, k, to obtain exactly the values of 4.0 and 6.0. Chen and Aris [4] indicate that the data was taken from Saleton and White [11]. This original reference contains much more data than that in Table 1, but Chen and Aris [4] took most of the data from Table 7 in Saleton and White [11]. The pertinent data from this table is shown in Table 2.

Saleton and White [11] noted that the data in Table 2 'was read from smoothed curves'. Thus, even the original data in Table 2 do not represent experimental data, but data which were altered by interpolation and extrapolation. This process introduced changes in the error distribution of the data, the effects of which are impossible to assess without comparing the smoothed data with the raw experimental data. Chen and Aris [4] further changed the error distribution by rounding the numbers that appear in Table 2.

The conclusion from this discussion is that the data in Table 1 does not represent experimental data. Therefore, it is inadequate for comparing different regression techniques. For such a comparison, original, unaltered measured data should be used.

There are only five data points in Table 1. Is such a small number of points large enough to find a significant difference between regression techniques? To answer this question, we have calculated the constants A and E of the Arrhenius equation, using linear and nonlinear regression.

Table 2
Experimental data (Saleton and White [11])

i	T_i (°C)	k_i (l mol^{-1} h^{-1})
1	40	1.09
2	50	2.19
3	61	3.99
4	71	6.0

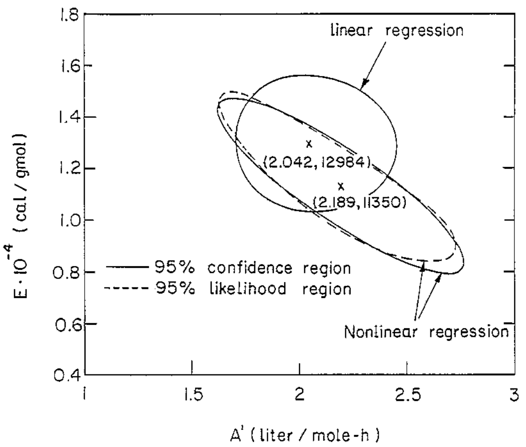

Fig. 1. Joint confidence regions for A and E for the Chen and Aris [4] data.

The Arrhenius constants obtained (including individual 95% confidence intervals) are: For nonlinear regression: $A=(1.0399\pm3.8278)\times10^{8}$ l mol^{-1} h^{-1}, $E=11350\pm2469$ cal/g mol. The sum of squares of errors with these constants is $S=0.1496$. For linear regression: $A=(1.2365\pm24.24$ or $-1.17)\times10^{9}$ l mol^{-1} h^{-1}, $E=12984\pm1935$ cal g^{-1} mol^{-1} and $S=0.455$.

There is no significant difference between these results and the parameter values obtained by Chen and Aris [4]. It can be seen that the confidence intervals on A are very large and when nonlinear regression is used, the confidence interval extends to negative values. In this case, the large confidence intervals are due to strong interaction between the parameters A and E. Using Eq. (1b) instead of Eq. (1a) with $T_0=323.15$ yields the following results. For nonlinear regression: $A'=2.189\pm0.41796$ l mol^{-1} h^{-1}, for linear regression: $A'=(2.0418+0.2897$ or $-0.2537)$ l mol^{-1} h^{-1}. The value of E and S are identical to those obtained when using Eq. (1a).

Eq. (1b) is used for calculating the 95% joint confidence regions for the constants obtained by linear and nonlinear regression. It can be seen (Fig. 1) that the two regions overlap in about half of their area and the minimum of the sum of squares is located in the common region for both types of regression. This indicates that the difference between the constants obtained by linear and nonlinear regression is statistically insignificant. The insignificance is probably a result of using too few data points. In order to obtain statistically significant results, it is imperative to use unaltered experimental data and large enough number of data points.

We have repeated this test using the joint likelihood regions. For the linear case, the joint confidence region and the joint parameter likelihood region are identical. The 95% likelihood contour obtained for the results of nonlinear regression is also shown in Fig. 1 (dashed curve). The differences between the joint confidence and likelihood regions is insignificant, indicating that the linear approximation ellipses correctly describe the parameters likelihood region.

4. Example 2: hydrogenation of ethylene

Wynkoop and Wilhelm [7] studied the catalytic hydrogenation of ethylene over copper magnesia catalyst in a continuous flow tubular reactor. They carried out 75 experiments with average temperature ranging between 13 and 79°C. The results of these measurements were reported with three to five decimal digits of accuracy, and are used here to calculate the reaction rate coefficient k. Because of the variation of the temperature along the reactor, we used the linear averaged temperature as recommended by Winkoop and Wilhelm. From among the 75 experiments, 30 were carried out with water vapor present.

Wynkoop and Wilhelm [7] concluded that water vapor causes reversible poisoning of the catalyst. For the present study, we used only the 45 data points without water vapor present. For the 45 data points of k vs temperature, the constants, A and E, were obtained by linear regression and nonlinear regression using both Eq. (1a) and Eq. (1b). The value $T_0=329$ was used with Eq. (1b). The regression program of POLYMATH [9] was used for the calculations. The calculated constants, sum of squares of errors, and results reported by Wynkoop and Wilhelm [7] are shown in Table 3. (Note that all the sum of squares of errors were calculated using k values, (not ln k)).

There are several facts worth noting in this table. The results obtained here using linear regression (with double precision computation) are very close to the values reported by Wynkoop and Wilhelm [7]. This indicates that when linearization is used, the calculations do not require high precision, since Wynkoop and Wilhelm probably did not have high precision computational tools available in 1950. The sum of squares of errors is somewhat higher when linearization is used, but the confidence interval of the individual constants (relatively to the constant value itself) is much smaller in linear regression, indicating that the uncertainty in the calculated constants is smaller when linearization is used. (Note that the confidence interval in the linearized case is not symmetric because of the transformation from ln k to k).

Fig. 2(a) shows the experimental points and the line of calculated ln k vs $1/T$ when the constants of the linearized Arrhenius expression are used. It can be seen that, indeed, the experimental data is well represented by a straight line. Fig. 2(b) shows the plot of k (calcu-

Table 3
Arrhenius equation constants and sum of squares of errors for the Wynkoop and Wilhelm [7] data

	Linear regression	Nonlinear regression	Wynkoop and Wilhelm
$\hat{A}^{a}$	$5988+3686$ -2281	909 ± 1136	5960
$\hat{A}'$	$(8.2665+0.3004-0.2898)\times10^{-9}$	$(8.7357\pm0.6383)\times10^{-6}$	
$\hat{E}^{a}$	13336 ± 304	12068 ± 859	13320
S	12.84×10^{-11}	9.21×10^{-11}	14.68×10^{-11}

[a] Numbers were rounded to the decimal point.

lated and observed) versus $1/T$ when the nonlinear regression constants are used. The fit does not look as good as for the linear case, especially for the high temperature values (low $1/T$). This, of course, does not provide a clear evidence for the superiority of one of the methods. We have to check whether the observed experimental error is nearly normally distributed with zero mean, as expected from the theory of regression diagnostics. The residual for data point i is:

$$\epsilon_i = k_i - \hat{A}\exp(-\hat{E}/RT_i) \tag{8}$$

and the corresponding relative error is $\epsilon_{r_i} = \epsilon_i/k_i$.

Fig. 3(a) shows a plot of the residuals of k obtained with the linearized Arrhenius equation. There are 21 positive and 24 negative residuals. It can be seen that, in general, the residual, which represents the experimental error, increases with increasing k.

Fig. 3(b) is the residual plot for the nonlinear case. There are 11 positive residuals and 34 negative residuals. The discrepancy between the number of positive and negative residuals is a clear indication that the experimental error is not normally distributed around the calculated curve. As in the linear case, the residual tends to increase with increasing k.

The trend of increasing experimental error with increasing k, can be eliminated by plotting the relative error vs k for the linearized equation (Fig. 4(a)), which corresponds to the residual plot of ln k. It can be seen that, indeed, the relative error is distributed randomly and all the points, except one, lie inside the region of ± 24% error.

Fig. 4(b) shows the relative error versus k for the nonlinear case. It can be seen that now there is a clear trend of increasing relative error with decreasing k and all, except two points, lie inside the region of + 20 and − 45% relative error.

Box and Hill [6] recommended finding an optimal value for the parameter ϕ for the power transformation (defined in Eq. (5)). This optimal value can in principle be different from 0 or 1 (corresponding to minimization of the relative errors or absolute errors respectively). The optimal value of ϕ is estimated by maximizing the likelihood function, L, following the procedure outlined in Draper and Smith [12]. The plot obtained for $L(\phi)$ is shown in Fig. 5, which shows that the likelihood function attains a maximum at $\phi \approx 0$.

All these indicators substantiate the observation that it is the relative error which is normally distributed in the experimental data and the appropriate Arrhenius constants are those obtained by minimizing the relative

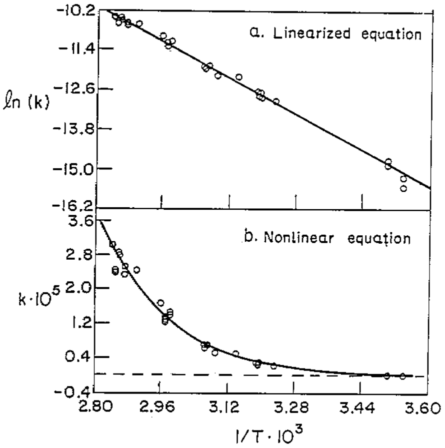

Fig. 2. Experimental data and calculated Arrhenius curve for Wynkoop and Wilhelm [7] data.

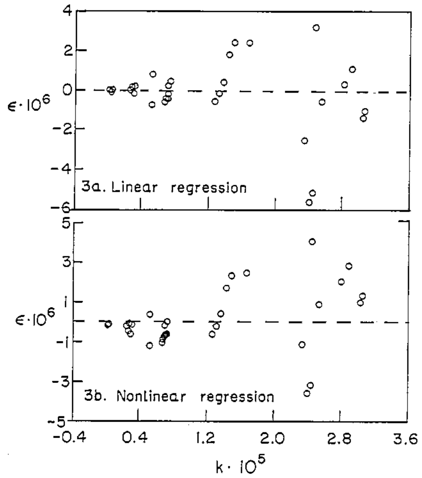

Fig. 3. Residual plot (absolute errors) for the Wynkoop and Wilhelm [7] data.

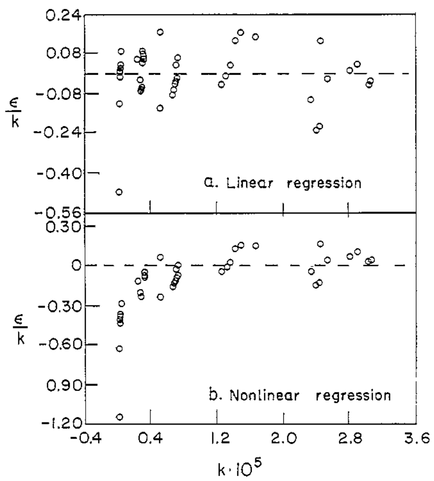

Fig. 4. Residual plot (relative errors) for the Wynkoop and Wilhelm [7] data.

error, as done when the linearized form of the Arrhenius equation is used. The use of relative error can be further justified by considering the measurement precision reported by Wynkoop and Wilhelm [7]. The reaction rate coefficient is calculated from measured flow rate, reactor volume and conversion. Wynkoop and Wilhelm report the approximate errors in measuring these variables as percentage, thus relative errors. In the temperature range of their experiments, the k values change by two orders of magnitude (from about 2×10^{-7} to 3×10^{-5}). Therefore, when the absolute error is minimized, the error at the range of high k values dominate, while the error at the low k values will have only a negligible effect.

Fig. 6 shows the 95% joint confidence regions for the constants obtained by linear and nonlinear regression. It can be seen that for this set of data, the confidence regions are practically separated (there is only a very small overlap on the boundaries). The 95% joint confidence region and 95% likelihood contours (dashed line) are practically identical. Thus, the difference between the parameter values obtained with linear and nonlinear regression in this case is indeed statistically significant.

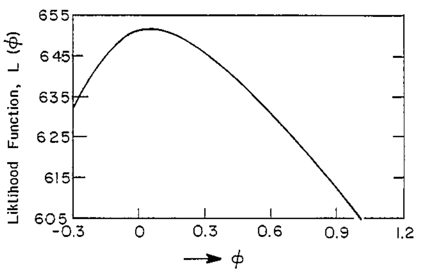

Fig. 5. Plot of the likelihood function $L(\phi)$ for the Wynkoop and Wilhelm [7] data.

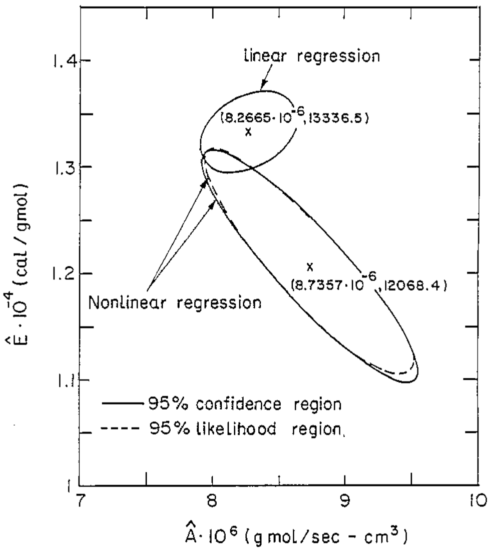

Fig. 6. Joint confidence regions for $\hat{A}$ and $\hat{E}$ obtained for the Wynkoop and Wilhelm [7] data.

5. Conclusions

Advanced statistical techniques and indications, namely: analysis of residuals, the maximum likelihood approach, joint confidence regions and joint likelihood contours were used to compare the traditional linear regression technique for the Arrhenius equation with the new and intuitively superior nonlinear regression technique. The comparison has shown that linear regression is in principle not inferior to nonlinear regression; depending on the nature of the error distribution in the data, it can be even superior (if the relative error is distributed normally, like when the k values change over a range of several orders of magnitude).

It should be emphasized that this conclusion applies only to this particular equation, for other group of equations, a different conclusion can be reached. Shacham et al. [13] has shown, for example, that linearization of equations which represent activity coefficients may cause serious errors.

In spite of the proven accuracy of the linear regression technique, when applied to the Arrhenius equation, the constants of this equation that appear in the literature cannot be blindly trusted. If there are too few data points or/and the original data was interpolated or smoothed before the regression was carried out, the accuracy of the correlation is unpredictable, as shown in Example 1.

Appendix A. Nomenclature

A	frequency factor
A'	modified frequency factor, Eq. (1b)
E	activation energy (kJ mol^{-1})

k	rate coefficient
n	number of data points
R	ideal gas constant (kJ mol^{-1} K^{-1})
S	sum of squares of errors
T	temperature (K)
T_0	average temperature (K)
W	weight assigned to data

Greek letters

ϵ	absolute error
ϵ_r	relative error
ϕ	parameter of transformation, Eq. (5)

Subscripts

i	index of data point
r	relative

Superscripts

$\hat{}$	estimated value

Appendix B. Joint confidence interval for A and k

The approximate joint confidence interval for linear model is defined by (Draper and Smith [12]):

$$(\beta - b)^{\mathrm{T}}(X^{\mathrm{T}}X)(\beta - b) = ps^2F_{1-a}(p, n-p) \tag{B1}$$

where n is the number of observations, p the number of parameters in the model, X is the $(n \times p)$ matrix of observations of the independent variables, β and b are the $(p \times 1)$ vectors of the estimate on expected values for the model parameters and those obtained by the regression of the n observations, respectively. $F_{1\text{-}a}$ is the upper limit of the F distribution for p and $n-p$ degrees of freedom and s^2 is the estimated variance of the experimental error:

$$s^2 = \frac{S}{n-p} \tag{B2}$$

In case of simple linear regression, $p = 2$ and Eq. (B1) reduces to a quadratic algebraic equation, which yields an explicit expression for the joint confidence interval of β_1 and β_0:

$$\tilde{\beta}_1 = \beta_1 - b_1 = -a_{21}^2 \pm \{a_{21}^2\tilde{\beta}_0^2 - a_{22}(a_{11}^2\tilde{\beta}_0^2 - C)\}^{1/2};$$
$$\tilde{\beta}_0 = \beta_0 - b_0 \tag{B3}$$

where

$$C = ps^2F_{1-\mathrm{a}}(2, n-2) \tag{B4}$$

The a_{ij} are the elements of the (2×2) X^TX matrix, and are given by:

$$a_{11} = n;\ a_{12} = a_{21} = \sum_{i=1}^{n} X_i;\ a_{22} = \sum_{i=1}^{n} X_i^2 \tag{B5}$$

Note that Eq. (B3) defines the upper and lower limits for the range of variation of β_o over the joint confidence interval:

$$b_0 - \frac{\sqrt{a_{22}C}}{(a_{22}a_{11} - a_{21}^2)^{1/2}} \leq \beta_0 \leq b_0 + \frac{\sqrt{a_{22}C}}{(a_{22}a_{11} - a_{21}^2)^{1/2}} \tag{B6}$$

Eq. (B3) to Eq. (B6) have been utilized to obtain the joint confidence interval for Arrhenius equation parameters $A(\equiv \beta_0)$ and $E \equiv (\beta_1)$ with $X_i \equiv T_i$ (Figs. 1 and 6).

References

[1] M. Shacham, N. Brauner and M.B. Cutlip, Critical analysis of experimental data, regression models and regression coefficients in data correlation, *AIChE Symp. Ser., 304 (91)* (1995) 305.

[2] D.M. Himmelblau, *Process Analysis by Statistical Methods*, Wiley, New York, 1970.

[3] D.M. Bates and D.G. Watts, *Nonlinear Regression Analysis and its Application*, Wiley, New York, 1988.

[4] N.H. Chen and R. Aris, Determination of Arrhenius constants by linear and nonlinear fitting, *AIChE J., 38* (1992) 626.

[5] D.W. Marquardt, An algorithm for least squares estimation of nonlinear parameters, *J. Soc. Ind. Appl. Math, 11* (1963) 431.

[6] G.E.P. Box and W.J. Hill, Correcting inhomogeneity of variance with power transformation weighting, *Technometrics, 16 (3)* (1974) 385.

[7] R. Wynkoop and R.H. Wilhelm, Kinetics in tubular reactor, hydrogenation of ethylene over copper-magnesia catalyst, *Chem. Eng. Progr., 46 (6)* (1950) 300–310.

[8] J.R. Kittrell, Mathematical modeling of chemical reactions, *Adv. Chem. Eng., 8* (1970) 97–198.

[9] M. Shacham and M.B. Cutlip, *Polymath 3.0 Users' Manual*, CACHE Corporation, Austin, Texas, 1994.

[10] R.L. Curl, Letter to the Editor, *AIChE J., 39 (8)* (1993) 1420.

[11] D.I. Saleton and R.R. White, Cation-exchange resin as a catalyst in the synthesis of ethyl acetate, *Chem. Eng. Prog. Symp. Ser., 48 (4)* (1952) 59–74.

[12] N.R. Draper and H. Smith, *Applied Regression Analysis*, Wiley, New York, 1981.

[13] M. Shacham, J. Wisniak and N. Brauner, Error analysis of linearization methods in regression of data for the Van-Laar and Margules equations, *Ind. Eng. Chem. Res., 32* (1993) 2820.