

## Approximation of Experimental Rate-temperature Data by two Different Extended Arrhenius Equations

SVANTE WOLD

*Department of Organic Chemistry, Umeå University, S-901 87 Umeå, Sweden*

The performance of two different extensions of the Arrhenius equation has been tested on 70 series of rate-temperature data taken from the literature.

One of the equations which frequently has been used for estimation of activation parameters assumes a temperature independent heat capacity of activation ( $\Delta c_p^\ddagger$ ).

The other equation, recently proposed, gives a temperature dependent heat capacity of activation.

No significant difference has been found in the accuracy of fit of the two equations to the data.

Recently, in connection with the development of the Varytemp method,<sup>1,2</sup> a fast routine method to determine activation parameters from one kinetic experiment, a new equation for the description of the temperature dependence of rate constants was proposed. This equation predicts a temperature dependence of the heat capacity of activation.

The main reason for adoption of the new equation in the Varytemp method was that the conventional equation was shown to be numerically unstable in the temperature intervals used in the method.

The two equations were tested on a limited amount of experimental data. The preliminary results indicated that the new equation fitted the data as well or better than the conventional equation. However, no rigorous statistical tests on the performance of the two equations were made.<sup>2</sup>

In this investigation the goodness of fit of the two equations to all relevant literature data easily available has been compared. The test data consisted of 40 series of measurements on first order hydrolysis reactions in water carried out mainly by Robertson and co-workers, and 30 series of first order solvolysis data where the solvent used was aqueous ethanol or aqueous acetone.

The results have been subjected to various straight forward statistical tests to find any difference in the goodness of fit of the two equations.

## THEORY

*Equations.* The equation commonly used to describe the temperature dependence of rate constants ( $k$ ) is the Arrhenius equation

$$\ln k = A + B(1/T - 1/T_0) + \varepsilon_k \quad (1)$$

The natural logarithm is denoted by  $\ln$ .  $A$  and  $B$  are adjustable parameters specific for the reaction in question. The temperature  $T$  is measured in degrees Kelvin and  $\varepsilon_k$  symbolizes a random error in  $\ln k$  due to various causes.  $T_0$  is the reference temperature around which the parameters  $A$  and  $B$  are valid in an interval usually defined by the lowest and highest temperatures  $T_1$  and  $T_N$ .

The need for an extension of eqn. 1 was early pointed out (for reviews on the development of extended Arrhenius equations see Refs. 3-5), and Robertson and Hyne<sup>6,7</sup> suggested the use of the three parameter equation

$$\ln k = A + B(1/T - 1/T_0) + C \ln (T/T_0) + \varepsilon_k \quad (2)$$

when the Arrhenius equation was insufficient to describe the data. This equation can be derived from the transition state theory<sup>8-10</sup>

$$\ln k = \ln(kT/h) - \Delta H^\ddagger/RT + \Delta S^\ddagger/R \quad (3)$$

( $k$  is Boltzmann's constant,  $h$  is Planck's constant,  $\Delta H^\ddagger$  is the enthalpy of activation,  $\Delta S^\ddagger$  is the entropy of activation, and  $R$  is the gas constant) assuming that the heat capacity of activation

$$\Delta c_p^\ddagger = \frac{\partial \Delta H^\ddagger}{\partial T}$$

is constant (for derivation, see for example Ref. 3 or 6 and 7). Fitting eqn. 2 to empirical data, one can derive the following estimates for the activation parameters at the temperature  $T$

$$\Delta H^\ddagger = -R[B + (1-C)T] \quad (4)$$

$$\Delta S^\ddagger = R[A + B(1/T - 1/T_0) + C \ln(T/T_0) - \ln(kT/h)] + \Delta H^\ddagger/T \quad (5)$$

$$\Delta c_p^\ddagger = R(C-1) \quad (6)$$

Eqn. 2 cannot be used directly on data with small temperature intervals since the third term then becomes almost linearly dependent on the first two terms. This can be overcome by a suitable orthogonalization procedure as shown by Clarke and Glew,<sup>11</sup> or by substitution of the third term by one which is not as covariant with the first two. The latter approach led to the proposal of the equation

$$\ln k = A + B(1/T - 1/T_0) + C(1/T - 1/T_0)^2 + \varepsilon_k \quad (7)$$

for the description of rate-temperature data in the Varytemp<sup>1,2</sup> method. This equation has the desirable numerical stability and preliminary tests indicated that it fitted experimental data as well or better than did eqn. 2.

Eqn. 7, interpreted in view of the transition state theory, gives a constant derivative  $\partial \Delta H^\ddagger / \partial (1/T)$  which seems as reasonable as the constant  $\Delta c_p^\ddagger$

obtained from eqn. 2. The estimates for the activation parameters at the temperature  $T$  using eqn. 7 become

$$\Delta H^\ddagger = -R[B + 2C(1/T - 1/T_0) + T] \quad (8)$$

$$\Delta S^\ddagger = R[A - B/T_0 + C(1/T_0^2 - 1/T^2) - \ln(kT/h) - 1] \quad (9)$$

$$\Delta c_p^\ddagger = R[2C/T^2 - 1] \quad (10)$$

It should be noted that eqns. 2 and 7 can be further extended to include a fourth term,  $D(T - T_0)$  and  $D(1/T - 1/T_0)^3$ , respectively.<sup>6,7,2</sup> This makes eqn. 2 even more numerically unstable but does not affect eqn. 7. The expressions for the activation parameters must then, of course, include the fourth parameter  $D$ . However, in only a few of the cases tested in this investigation (Nos. 11, 23, and 25 in Tables 1 and 2) has the degree of fit been significantly increased by inclusion of the fourth terms. Hence, this investigation will be confined to testing eqns. 2 and 7 using only three terms.

Thus, the difference between the two equations tested is the form of the third term. Eqn. 7 gives a temperature dependent  $\Delta c_p^\ddagger$  with the derivative

$$\partial \Delta c_p^\ddagger / \partial T = -4RC/T^3 \simeq -2\Delta c_p^\ddagger / T \quad (11)$$

If  $T$  is around 300°K and  $\Delta c_p^\ddagger$  is of the usual order  $-50$  cal/deg·mole, this quantity is of the order 0.3 cal/deg<sup>2</sup>·mole. With the hitherto largest estimated  $\Delta c_p^\ddagger$  (by eqn. 7) of 96 cal/deg·mole the derivative is 0.64 cal/deg<sup>2</sup>·mole at  $T = 300^\circ\text{K}$  (27°C). These values are well consistent with the limits estimated by Ives and Marsden<sup>12</sup> (0.5 cal/deg<sup>2</sup>·mole) but somewhat large compared to the limits set by Fox and Kohnstam<sup>13</sup> (0.2 cal/deg<sup>2</sup>·mole). Were the latter authors correct, this would probably show up as a significantly better fit of eqn. 2 than of eqn. 7 to experimental data.

The importance of discrimination between eqns. 2 and 7 is understood if it is noted that the  $\Delta c_p^\ddagger$  values estimated by eqn. 2 are independent of the temperature interval used in the experiment; the  $\Delta c_p^\ddagger$  values obtained from experiments with different  $T_0$  values can thus be compared directly. The adoption of eqn. 7 makes it necessary to compare the values of  $\Delta c_p^\ddagger$  at the same temperature. Thus the difference in the  $\Delta c_p^\ddagger$  values estimated by eqn. 7 at 280°K and 360°K is about 60% based on the smaller value (at 360°K).

Since  $\Delta c_p^\ddagger$  values invariably are used for comparison with other  $\Delta c_p^\ddagger$  values, it is important to know whether it is necessary for the values to be normalised to the same temperature.

The  $\Delta c_p^\ddagger$  values estimated at the middle temperature of the interval,  $T_0$ , differ very little as estimated from eqn. 2 or eqn. 7 (or probably from any other three parameter equation). This makes it possible to compute the values at any desired temperature,  $T$ , by simply applying the temperature dependence thought to best represent the actual situation to these values estimated at  $T_0$ . Thus if we think that  $\Delta c_p^\ddagger$  is a function of temperature as  $1/T$ , we get the values at the temperature  $T$  as  $\Delta c_{p,T}^\ddagger(T_0/T)$ .

*Statistical tests.* (The treatment below is statistically simple and can be understood with the help of any standard textbook in statistics, e.g. Ref. 14).

A general measure of the goodness of fit of an equation to data is

$$s(\varepsilon) = \sqrt{(\sum_i \varepsilon_i^2 \cdot \omega_i) / (N - P)} \quad (12)$$

the so called standard deviation of  $\varepsilon$ . The square of this property,  $s^2$ , is called the estimated variance of  $\varepsilon$  and is an equally good measure of the fit. The summation in eqn. 12 is made over all points to which the equation in question is fitted.  $N$  is the number of these points and  $P$  is the number of parameters in the equation. The function  $\omega_i$  is a weighing function incorporated to compensate if  $\varepsilon_i$  has different expected variance in different regions of the experimental interval. In this investigation the value of 1 has been used for  $\omega_i$  throughout the calculations since the errors in  $\ln k$  are in general of constant-variance. (This was also confirmed by the results.)

In general an equation which is much better than another one will give a smaller  $s(\varepsilon)$ . However, in this case the differences are small since most of the variation in  $\ln k$  is described by the first two terms which are equal in the two equations. Thus,  $s(\varepsilon)$  will not invariably be smaller for one equation than for the other, and great care must be taken in the total analysis of the results of all calculations so that relevant measures of fit are compared and analyzed.

In this investigation the different series of measurement concern different reactions and are measured in different temperature intervals which are of different lengths. Also, there is a variation in the accuracy of measurement between the different series.

The expected difference (from the statistical point of view) between the fit of the two equations increases with the temperature interval and with  $\Delta c_p^\ddagger$  (as this is a measure of the deviation from eqn. 1). The dependence on the errors of measurement comes in, not in the size of the difference in fit, but in the variance of this difference. Thus for very inaccurate measurements, possible differences in fit are obscured by fluctuations in the fit due to the random errors. However, it is in principle always possible, regardless of the size of the errors, to discriminate between two equations given a sufficient number of test cases.

In order to (at least partly) correct for these variations in the expected difference of fit in the test material, the variance of this difference (hereafter the difference in variances of  $\varepsilon$  estimated by the two equations) has been estimated for every data series. The corresponding standard deviation, hereafter called  $d_{12}$ , is thus a measure of the sharpness or reliability of the difference in fit. Consequently in the total analysis, the measure

$$R_{12} = (s_I^2 - s_{II}^2)/d_{12} \quad (13)$$

has been used as a measure of the difference in performance of the two equations on an experimental data series.  $d_{12}$  has been estimated by a Monte Carlo type of method. (See experimental part.) In eqn. 13 the subscripts I and II symbolize the use of eqns. 7 and 2, respectively; this notation will be used hereafter.

However,  $R_{12}$  is not corrected for variation due to different size of  $\Delta c_p^\ddagger$  and temperature interval. Because of the expected difference in fit increases with the increase of these factors, the natural thing to do would be to divide  $R_{12}$  by the expected difference  $E(s_I^2 - s_{II}^2) = E_{12}$ . However, it may also be argued that larger weight should be given to series with large  $E_{12}$  since in those cases the difference should have the least chance to be an artefact. Thus

$R_{12}$  should be multiplied by  $E_{12}$  to give a relevant and suitable measure of difference in performance.

In this investigation it has been found that the conclusions based on the analyses of  $R_{12}$ ,  $R_{12}/E_{12}$ , and  $R_{12}E_{12}$  are the same. Since  $R_{12}$  is a compromise between the various arguments, the emphasis will be placed on the analysis of this parameter.

The following statistical tests have been made on the various measures of difference.

*F tests.* These tests give a probability that one variance is equal to or greater than the other. The quantities

$$F_1 = \sum s_1^2 / \sum s_{11}^2 \quad (14)$$

and

$$F_2 = (\sum s_1^2/d_{12}) / (\sum s_{11}^2/d_{12}) \quad (15)$$

with the summations made over all series considered, will thus give a probability that one model is better than the other.  $F$  value larger than one will indicate that eqn. 2 is better;  $F$  values smaller than one will indicate that eqn. 7 is better.

*Significance tests on mean values.* As further tests on the performance of eqn. 2 and eqn. 7, the mean values ( $m$ ) of the different measures  $R_{12}$ ,  $R_{12}/E_{12}$ , and  $R_{12}E_{12}$  have been computed for the series 1–40, 41–70, and 1–70. Confidence intervals for these mean values have been computed in the usual way by the formula

$$G_{0.95} = 2 \ s / \sqrt{N} \quad (16)$$

which gives the 95 % confidence interval for the mean value ( $m$  is the estimate of  $\mu$ )

$$m - G \leq \mu \leq m + G \quad (17)$$

The use of formula 16 is allowed in this case since even distributions having large deviations from the normal distribution give this confidence interval if the number of observations is large. In this case there are 70 observations in the estimates of mean values and variances, this number should be well sufficient for robust estimates, especially since the deviations from normal distributions of the different measures is not large (see below in Table 5).

*Data.* The test data consist of 70 series of measurements on first order solvolysis reactions. In 40 cases (1–40) the solvent used was water, in 30 cases (41–70) it was aqueous acetone or ethanol. All data have been taken from the literature. The relevant references are given together with the results in Tables 1–4.

The data have been analyzed in two groups separately (1–40 and 41–70) and also together in order not to make erroneous conclusions due to possible variations between reactions in pure water and reactions in aqueous solvents.

Most of the rate constants have been determined with great accuracy, however, an estimated  $s(\varepsilon)$  smaller than  $1.5 \times 10^{-3}$  (corresponds to a standard deviation of 0.15 % in the rate constants) is probably due to overfitting of the corresponding equation and should be looked upon with some suspicion. This is concluded from the measurements of Robertson and co-workers (1–40) who have probably made the most accurate rate determinations possible today;

$s(\varepsilon)$  estimated from these data is usually larger than  $1.5 \times 10^{-3}$ . In these data (1–40) there are also more observations in each series (ten or more) than in the data 41–70 (five or six points per series) which make conclusions based on results from the first group more reliable than those based on the series 41–70.

Since all data come from solvolysis reactions in aqueous solvents, the conclusions of this investigation are valid only for such reactions.

## RESULTS

*Activation parameters.* The activation parameters  $\Delta H^\ddagger$ ,  $\Delta S^\ddagger$ , and  $\Delta c_p^\ddagger$  estimated by fitting of the two equations 2 and 7 to a series of measurements do not differ at the middle point ( $T_0$ ) of the interval. Also in the end points of the interval, the estimated values of  $\Delta H^\ddagger$  and  $\Delta S^\ddagger$  are not significantly different as estimated by the two equations. However, the  $\Delta c_p^\ddagger$  values estimated by the two equations are often significantly different well inside the experimental temperature interval. Therefore only the  $\Delta c_p^\ddagger$  values are tabulated (Tables 1–4) for the different data series. Since  $\Delta c_{p\text{II}}^\ddagger$  (estimated by eqn. 2) is constant with temperature, only one column is needed for this quantity, whereas  $\Delta c_{p\text{I}}^\ddagger$  (estimated by eqn. 7 at 25°C) is tabulated for comparison ( $\Delta c_{p\text{I}}^\ddagger = \Delta c_{p\text{II}}^\ddagger$  at  $T_0$ ). Tabulated errors are estimated double standard deviations.

It can be seen that the  $\Delta c_{p\text{I}}^\ddagger$  values at 25°C often give another picture than do the  $\Delta c_{p\text{II}}^\ddagger$  values. Thus the  $\Delta c_p^\ddagger$  values for allylchloride and allylbromide are different as estimated by eqn. 2 but equal as estimated by eqn. 7. The same behaviour of the  $\Delta c_p^\ddagger$  values within a series of reactions with similar substrates is seen in the series of methanesulfonates (31–34, 37 in Table 2) where the  $\Delta c_{p\text{I}}^\ddagger$  values are equal at 25°C but the  $\Delta c_{p\text{II}}^\ddagger$  values are significantly different. In no reaction series has the reverse situation occurred; thus eqn. 7 often predicts less variation of  $\Delta c_p^\ddagger$  (at one temperature) than does eqn. 2.

*Statistical tests on goodness of fit.* Various statistical quantities have been tabulated in Tables 1–4 together with the  $\Delta c_p^\ddagger$  values.  $s_I(\varepsilon)$  and  $s_{\text{II}}(\varepsilon)$  are the estimated standard deviations of the residuals, the squares of these quantities are the corresponding variances. The estimated variance of  $s_I^2$  and  $s_{\text{II}}^2$  ( $d_{12}$ ) is also tabulated as well as the derived quantities from  $s_I^2$ ,  $s_{\text{II}}^2$ ,  $d_{12}$ , and  $E_{12}$  (the expected difference between  $s_I^2$  and  $s_{\text{II}}^2$ ). Histograms of the distribution of  $R_{12}$  (defined in eqn. 13) for the series 1–40, 41–70, and 1–70 are shown in Fig. 1. It can be seen that the three distributions have mean values slightly larger than zero (see Table 5) indicating that eqn. 2 fits the data better than eqn. 7. However, the estimated mean values of  $R_{12}$  are not significantly different from zero, as the confidence intervals of the means all include zero well within the limits. The same is the case for the means of the other quantities  $R_{12}/E_{12}$  and  $R_{12} \cdot E_{12}$  as can be seen in Table 6. To estimate the deviation from normal distribution of the distributions of  $R_{12}$ ,  $\chi^2$  tests have been made. These show (Table 5) that the deviations are significant, but not very large. This shows that it is permissible to estimate the confidence intervals in the way done here.

Table I. Computed  $\Delta c_p^\ddagger$  values for hydrolysis of halides in water. Subscripts I and II refer to eqns. 7 and 2, respectively.  $T_0$  is the middle temperature in the experimental interval.  $T_1$  is the lowest temperature, and  $T_N$  is the highest temperature, in the experimental interval.  $N$  is the number of points in the experimental interval.  $s(\epsilon) = \text{computed standard deviations } \sqrt{\sum \epsilon^2 / (N-3)}$ .  $d_{13}^2$  is estimated variance of  $s^2(\epsilon)$  under the experimental conditions.  $R_{13} = (s_I^2 - s_{II}^2) / d_{13}$ .

No.	Compound	$T_0$ (°C)	$T_1$ (°C)	$T_N$ (°C)	$N$	$-\Delta c_p I^\ddagger$ (25°C)	$-\Delta c_p II^\ddagger$	$s_I(\epsilon) \times 10^3$	$s_{II}(\epsilon) \times 10^3$	$d_{13}^2 \times 10^{10}$	$R_{13}$	Reference
1	Methyl chloride	75	50	100	10	66 ± 10	49 ± 7	5.25	5.15	.217	.219	15
2	Methyl bromide	70	35	100	13	59 ± 3	46 ± 3	3.76	4.15	.0387	-1.566	15
3	Methyl iodide	75	50	100	12	77 ± 6	56 ± 4	3.66	3.43	.0325	.902	15
4	Ethyl bromide	75	60	90	9	66 ± 9	49 ± 7	1.88	2.08	.00641	-.966	16
5	Allyl chloride	60	35	85	12	61 ± 3	49 ± 3	2.15	2.30	.00538	-.915	17
6	Allyl bromide	40	15	65	13	65 ± 3	60 ± 2	2.80	2.29	.00523	3.63	17
7	Allyl iodide	50	25	70	9	52 ± 12	46 ± 11	7.76	7.70	2.61	.0528	17
8	Benzyl chloride	40	15	65	11	46.7 ± 1.5	43 ± 2	1.29	1.70	.00200	-2.75	17
9	Benzyl bromide	75	50	100	11	51 ± 4	38 ± 3	2.09	1.94	.00478	.905	16
10	Isopropyl chloride	50	25	75	13	68 ± 6	59 ± 5	4.35	4.24	.108	.288	16
11	Isopropyl bromide	50	25	75	14	68 ± 5	58 ± 5	4.70	5.31	.146	-1.59	16
12	<i>t</i> -Butyl chloride	10	1	20	20	76 ± 8	83 ± 9	1.88	1.86	.00102	.243	18
13	<i>t</i> -Butyl bromide	5	0	12	8	96 ± 30	112 ± 32	1.83	1.81	.00362	.116	19
14	2,2-Dichloropropane	40	27	55	13	83 ± 14	75 ± 14	4.29	4.30	.0788	-.0137	20
15	2-Chloro-2-bromopropane	25	10	35	13	92 ± 4	94 ± 4	1.07	1.08	.000225	-.0744	20
16	2,2-Dibromopropane	30	20	45	11	85 ± 15	80 ± 15	3.67	3.73	.0192	-.317	20
17	2-Chloro-2-methyl-1-propylmethylether	25	8	40	10	60 ± 6	61 ± 6	2.36	2.24	.00583	.704	19
18	$\alpha$ -Bromoisobutyrate ion	25	10	40	18	73 ± 9	73 ± 9	3.52	3.53	.0326	-.0260	21

220

Sum I-18

Table 2. Hydrolysis of sulfonates and miscellaneous compounds in water. Symbols are the same as in Table 1.

No.	Compound	$T_0$ (°C)	$N$	$T_1$ (°C)	$T_N$ (°C)	$-\Delta C_p^{\ddagger}$ (25°C)	$-\Delta C_p^{\ddagger}$	$s_I(\epsilon) \times 10^3$	$s_{II}(\epsilon) \times 10^3$	$d_{12} \times 10^{10}$	$R_{12}$	Refer- ence
19	Methyl benzenesulfonate	40	11	10	70	37 ± 2	34 ± 2	1.62	2.46	.00461	-5.06	22
20	Ethyl	40	12	10	75	37 ± 2	34 ± 2	2.46	2.38	.00715	.489	22
21	Propyl	50	7	20	80	36 ± 3	31 ± 2	2.57	2.69	.0280	-.378	22
22	Ethyl <i>p</i> -Me-	40	13	10	75	45 ± 4	40 ± 3	5.41	5.56	.153	-.420	16
23	Methyl <i>p</i> -Me-	40	17	0	80	33 ± 3	30 ± 2.5	8.51	6.87	.334	4.35	23
24	Methyl <i>p</i> -MeO-	50	10	35	70	46 ± 13	39 ± 11	4.84	4.89	.139	-.147	24
25	Methyl <i>p</i> -Br-	40	10	15	65	38 ± 5	35 ± 4	4.01	3.51	.122	1.09	24
26	Methyl <i>p</i> -nitro	40	11	25	55	22 ± 12	20 ± 11	3.73	3.69	.0625	.114	24
27	Methyl <i>m</i> -nitro	35	10	15	55	34 ± 15	32 ± 14	7.65	7.68	.231	-.081	24
28	Methyl 3,4-di-Me	50	9	30	80	49 ± 5	41 ± 4	2.75	2.74	.0193	.0399	25
29	Methyl 2,4-di-Me	50	9	30	80	43 ± 5	38 ± 4	2.93	2.81	.0431	.342	25
30	Methyl 2,4,6-tri-Me	50	8	30	80	39 ± 5	32 ± 4	2.95	3.23	.0442	-.828	25
31	Methyl methanesulfonate	30	12	0	60	37 ± 2	37 ± 2	3.39	2.66	.0139	3.76	26
32	Ethyl	50	8	10	80	38 ± 3	35 ± 3	5.27	4.05	.325	1.99	26
33	Propyl	50	15	20	90	34 ± 1.5	28 ± 1.5	2.34	2.91	.0152	-.241	26
34	Butyl	70	10	40	90	34 ± 4	26 ± 3	2.45	2.36	.00512	.617	26
35	Isopropyl benzene- sulfonate	20	15	0	35	39 ± 5	41 ± 5	3.11	3.05	.0141	.271	22
36	Isopropyl <i>p</i> -Me-	20	13	0	40	40 ± 6	42 ± 7	4.32	4.47	0.405	-.655	16
37	Isopropyl methane sulfonate	20	13	5	35	33 ± 3	35 ± 3	1.09	1.15	.000492	-.610	16
38	<i>t</i> -Butyl-dimethyl- sulfonium ion	70	12	45	90	9 ± 2	7 ± 1.6	1.00	0.99	.00012	.200	19
39	Dimethyl sulfate	25	14	5	45	45 ± 5	46 ± 4	2.93	2.74	.00806	1.22	27
40	Diethyl sulfate	25	15	5	45	43 ± 6	43 ± 6	4.19	4.00	.0375	.786	27



Table 3. Solvolysis reactions in aqueous ethanol. Symbols are the same as in Table 1.

No.	Compound	Mole Fraction or vol % EtOH	$T_0$ (°C)	$N$	$T_1$ (°C)	$T_N$ (°C)	$-\Delta C_{PI}^\ddagger$ (25°C)	$-\Delta C_{PI}^\ddagger$	$s_I(e) \times 10^3 s_{II}(e) \times 10^3$	$d_{12} \times 10^{10}$	$R_{12}$	Reference
41	Benzyl chloride	0	50	4	25	80	$50 \pm 20$	$42 \pm 18$	13.6	14.5	-.220	28
42	»	0.040 M	50	4	25	80	$40 \pm 23$	$34 \pm 20$	15.5	16.3	-.257	28
43	»	0.080 M	50	4	25	80	$27 \pm 16$	$23 \pm 14$	11.2	10.7	.280	28
44	»	0.125 M	50	4	25	80	$27 \pm 5$	$23 \pm 5$	3.49	3.96	-.468	28
45	»	0.203 M	50	4	25	80	$41 \pm 26$	$34 \pm 23$	17.6	18.3	-.275	28
46	Ph-CHCl <sub>2</sub>	50 %	40	5	20	60	$64 \pm 6$	$58 \pm 3$	2.44	1.62	3.21	29
47	Ph-CCl <sub>3</sub>	50 %	20	5	0	40	$70 \pm 10$	$73 \pm 8$	5.44	4.71	1.24	29
48	<i>t</i> -BuCl	50 %	20	6	5	30	$32 \pm 6$	$34 \pm 6$	1.31	1.40	-.465	30
49	<i>t</i> -BuCl ( $d_9$ )	50 %	20	6	5	30	$30 \pm 4$	$31 \pm 4$	0.897	.971	-.823	30
50	Ph-CCl <sub>3</sub>	80 %	45	7	25	65	$27 \pm 90$	$24 \pm 80$	40.5	40.4	.0104	29

49

Sum 41-50

Table 4. Solvolysis reactions in aqueous acetone. Symbols are the same as in Table 1.

No.	Compound	Vol % Acetone	$T_0$ (°C)	$N$	$T_1$ (°C)	$T_N$ (°C)	$-\Delta C_{pI}^\ddagger$ (25°C)	$-\Delta C_{pII}^\ddagger$	$s_1(\epsilon) \times 10^3$	$s_{II}(\epsilon) \times 10^3$	$d_{12}^3 \times 10^{10}$	$R_{12}$	Reference
51	Ethyl bromide	50	70	5	50	90	31 ± 6	31 ± 6	1.89	1.74	.0248	.348	31
52	Propyl bromide	50	70	5	50	90	29 ± 6	23 ± 4	1.77	1.70	.00621	.297	31
53	Butyl bromide	50	70	5	50	90	36 ± 6	27 ± 5	1.71	1.95	.00688	-1.06	31
54	<i>t</i> -Butyl chloride	50	40	5	20	60	36 ± 1	33 ± 3	0.548	0.146	.000659	3.44	32
55	Benzyl bromide	50	50	5	30	70	28 ± 4	24 ± 3	1.45	1.38	.00239	.375	31
56	<i>p</i> -Nitro-di- Ph-CHCl		20	5	0	40	26 ± 3	26 ± 2	1.66	1.31	.00467	1.50	32
57	<i>p</i> -Me-Ph-CHCl <sub>2</sub>		20	5	0	40	39 ± 4	40 ± 5	2.17	2.58	.0201	-1.36	32
58	Ph-CHCl <sub>2</sub>		55	6	30	80	36 ± 6	30 ± 5	3.37	3.94	.129	-1.16	29
59	Ph-CCl <sub>3</sub>		20	5	0	40	43 ± 6	44 ± 6	3.52	3.07	.241	.606	29
60	<i>t</i> -Butyl chloride	70	35	5	15	55	40 ± 3	38 ± 2	1.07	0.789	.000977	1.65	31
61	bromide		20	5	0	40	39 ± 6	40 ± 7	3.36	3.70	.0654	-.951	31
62	Benzyl chloride		95	5	75	115	31 ± 3	21 ± 2	0.717	0.772	.000134	-.703	31
63	bromide		65	5	45	85	28 ± 5	22 ± 3	1.53	1.39	.00348	.660	31
64	di-Ph-CHCl		15	5	-10	40	35 ± 6	38 ± 6	5.67	5.40	.797	.337	33
65	<i>p</i> -Nitro-di- Ph-CHCl		70	5	45	95	52 ± 2	40 ± 2	0.924	1.20	.00415	-.909	31
66	Br		35	5	15	55	33 ± 5	31 ± 4	2.22	1.78	.0417	.850	31
67	Phenyl- <i>p</i> -nitro- phenyl-acetate		110	6	80	140	42 ± 1	26 ± 1	0.505	0.480	.0000900	.262	12
68	<i>t</i> -Butyl chloride	80	50	5	30	70	37 ± 5	32 ± 4	2.10	1.70	.00377	2.50	31
69	bromide		20	5	0	40	30 ± 6	31 ± 6	3.13	3.47	.124	-.634	31
70	Di-Ph-CHCl		15	5	-10	40	37 ± 8	40 ± 7	7.68	6.24	.505	2.83	31

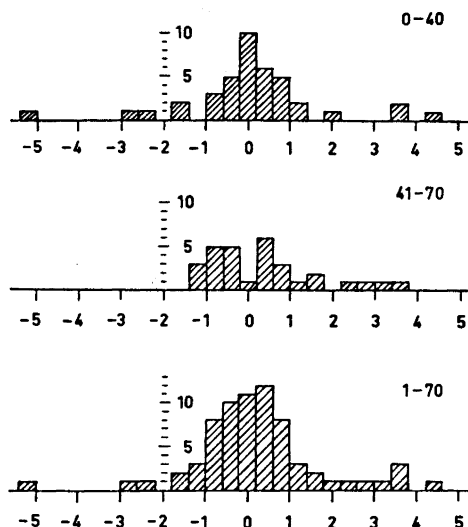


Fig. 1. Distributions of  $R_{12} = (s_I^2 - s_{II}^2) / d_{12}$  for the series 1-40, 41-70, and 1-70.

The  $F$  tests give the same results (Table 7). Eqn. 2 seems to be slightly better but the difference is far from significant.

The fact that the statistical tests give the same results for the measures  $s_I^2 - s_{II}^2$ ,  $R_{12}$ ,  $R_{12}/E_{12}$ , and  $R_{12} \cdot E_{12}$  shows that the estimation of and tests on the latter three quantities were somewhat superfluous. This might have been predicted in advance, it is quite improbable that different measures give different test results when the number of test cases is as large as 70.

However, since the possibility always exists that different measures give different test results, the present author feels that the tests on the quantities

Table 5. Mean values and confidence intervals for  $R_{12} = (s_I^2 - s_{II}^2) / d_{12}$ .

	1-40	41-70	1-70
$U = \sum(s_I^2 - s_{II}^2) / d_{12}$	3.53	11.1	14.6
Mean	0.088	0.370	0.209
Standard dev.	1.63	1.30	1.50
$c/\sqrt{N}$ 90 and 95 %	0.26, 0.31	0.30, 0.36	0.20, 0.24
Confidence interval (90 %)	-0.33...0.50	-0.02...0.76	-0.09...0.51
$\chi^2$ for normal dist.	17.21	8.36	13.44
$\chi^2 \times 95$	11.07	9.49	12.59

Table 6. Mean values and standard deviations for  $Q_1 = (s_I^2 - s_{II}^2) / [d_{12} \cdot E(s_I^2 - s_{II}^2)]$  and  $Q_2 = (s_I^2 - s_{II}^2) E(s_I^2 - s_{II}^2) / d_{12}$ .

	1-40	41-70	1-70
Sum $Q_1$	$+1.01 \times 10^8$	$-6.37 \times 10^7$	$+3.69 \times 10^7$
mean $Q_1$	$+2.52 \times 10^6$	$-2.12 \times 10^6$	$+5.28 \times 10^5$
$s(Q_1)$	$1.60 \times 10^7$	$1.62 \times 10^7$	$1.61 \times 10^7$
Sum $Q_2$	$+2.15 \times 10^{-5}$	$+1.12 \times 10^{-5}$	$+3.27 \times 10^{-5}$
mean $Q_2$	$+5.37 \times 10^{-7}$	$+3.73 \times 10^{-7}$	$+4.67 \times 10^{-7}$
$s(Q_2)$	$+3.56 \times 10^{-6}$	$1.32 \times 10^{-6}$	$2.82 \times 10^{-6}$

Table 7. F-tests.

	1-40	41-70	1-70
$\sum s_I^2$	$5.97 \times 10^{-4}$	$2.72 \times 10^{-3}$	$3.32 \times 10^{-3}$
$\sum s_{II}^2$	$5.64 \times 10^{-4}$	$2.75 \times 10^{-3}$	$3.31 \times 10^{-3}$
F	1.059	0.989	1.003
$\sum s_I^2/d_{12}$	277	110	387
$\sum s_{II}^2/d_{12}$	274	98.5	372
F	1.011	1.117	1.040
Degrees of freedom	354	61	415
F × 95	1.22	1.50	1.16

based on  $R_{12}$  have added further security to the conclusions that the two equations perform equally well on the experimental data.

It is thus concluded that the two equations 2 and 7 fit the data equally well and the small difference found in this investigation can well be the result of chance.

#### DISCUSSION

The two investigated equations fit experimental data equally well. Some results indicate that eqn. 7 is slightly better (*e.g.* the smaller variation within a series of reactions with similar substrates), whereas other results (the slightly better fit of eqn. 2) indicate that eqn. 2 should be preferred. The arguments are, however, far from conclusive. More accurate rate data are obviously needed to make a distinction possible between the two equations.

Since the two equations give different temperature dependence for  $\Delta c_p^\ddagger$  the situation is somewhat unsatisfactory and one cannot at this stage recommend the use of either of these two equations. However, it can be concluded that until an equation is found which is superior to the two equations investi-

gated here,  $\Delta c_p^\ddagger$  values estimated from experiments with differing temperature intervals must be compared with great caution. At present, the best thing to do is probably to take the weighted mean values of  $\Delta c_{pI}^\ddagger$  and  $\Delta c_{pII}^\ddagger$  at a reference temperature (e.g. 25°C) and make the comparisons with these values.

It is the hope of the present author that further investigations (currently under way) of the temperature dependence of rate constants will provide an equation which is better suited for the description of rate-temperature data than the equations investigated here.

### EXPERIMENTAL

*Computer program.* The computer program described in Ref. 2 has been refined. The regression analyses on eqns. 2 and 7 have been performed as described previously.<sup>2</sup>

The estimation of the variance of  $s_I^2$  and  $s_{II}^2$  made necessary the use of a random number generator for generation of synthetic errors of measurement. The random generator used was the sequence

$$U_{i+1} = U_i(2^{12} + 3) + 1 \text{ (modulo } 2^{23}\text{)}$$

which generates rectangularly distributed pseudo random numbers.<sup>32</sup> The tested autocorrelation of the random numbers is less than 0.1 %.

To obtain a normally distributed pseudo random number one computes

$$D_k = \left( \sum_{j=n}^{n+12} U_j \right) - 6$$

which makes  $D_k$  have an almost normal distribution (accuracy > 99 %) with a variance of 1, mean 0, range  $\pm 6$ .

*Monte Carlo estimation of  $d_{12}$ .* This was made in the following way for each reaction series 1–70.

a) Eqn. 2 (subscript II) and eqn. 7 (subscript I) were fitted to the data series giving estimates of the parameters  $A_{II}$ ,  $B_{II}$ ,  $C_{II}$ ,  $s_{II}$ ,  $A_I$ ,  $B_I$ ,  $C_I$  and  $s_I$ , respectively.

b) Using the set  $A_I$ ,  $B_I$ ,  $C_I$  and the experimental  $T$ -values  $T_1 - T_N$ , 100 different synthetic series of data were generated according to eqn. 7, with  $\varepsilon_k$  being  $D_k \cdot s_I$  and  $D_k$  generated as described above.

Fitting these 100 different series to eqn. 2 and eqn. 7 give 100 different  $s'_{II}{}^2$  and  $s'_I{}^2$  values, respectively, (the primes indicate that these variances are different from  $s_{II}^2$  and  $s_I^2$ ).

c) Similarly  $s'_{II}{}^2$  and  $s'_I{}^2$  are computed for 100 different series generated by the parameters estimated by eqn. 2.

d) For the four series of  $s^2$ , the means ( $m$ ) and variances ( $d^2$ ) were computed, and  $d_{12}$  was taken as the square root of the mean of these four variances.

$E_{12}$ , the estimated difference in fit of the two equations (were one of the equations true) is taken as the difference of the means ( $m$ ) as estimated by eqn. 7 and eqn. 2.

*Acknowledgements.* The author is greatly indebted to professor Göran Bergson, professor Herman Wold and fil. lic. Per Ahlberg for valuable discussions, helpful advice and kind encouragement with this work. Valuable advice has been given by Mr. Birger Bjurén of the Statistical Service of the Swedish Natural Science Research Council, Umeå. The access to unpublished primary data in the Depository of Unpublished Data, National Science Library, NRC, Ottawa, Canada, is gratefully acknowledged.

### REFERENCES

1. Ahlberg, P. *Acta Chem. Scand.* **24** (1970) 1883.
2. Wold, S. and Ahlberg, P. *Acta Chem. Scand.* **24** (1970) 618.
3. Kohnstam, G. *Advan. Phys. Org. Chem.* **5** (1967) 121.

4. Robertson, R. E. *Prog. Phys. Org. Chem.* **4** (1967) 213.
5. Hulett, J. R. *Quart. Rev. (London)* **18** (1964) 227.
6. Robertson, R. E. and Hyne, J. B. *Can. J. Chem.* **33** (1955) 1544.
7. Robertson, R. E. *J. Chem. Phys.* **25** (1956) 375.
8. Evans, M. G. and Polyani, M. *Trans. Faraday Soc.* **50** (1955) 875.
9. Eyring, H. *J. Chem. Phys.* **3** (1935) 107.
10. Glasstone, S., Laidler, K. J. and Eyring, H. *The theory of rate processes*, McGraw, N.Y. 1941.
11. Clarke, E. C. W. and Glew, D. H. *Trans. Faraday Soc.* **62** (1966) 539.
12. Ives, D. J. G. and Marsden, P. D. *J. Chem. Soc.* **1965** 649.
13. Fox, J. R. and Kohnstam, G. *J. Chem. Soc.* **1963** 1593.
14. Cf. Kreyzig, E. *Statistische Methoden und ihre Anwendungen*, Van der Hoeck and Ruprecht, Göttingen 1965; and Mandel, J. *The Statistical Analysis of Experimental Data*, Wiley, N.Y. 1964.
15. Hepolette, R. L. and Robertson, R. E. *Proc. Roy. Soc. A* **252** (1959) 273.
16. Hepolette, R. L. and Robertson, R. E. *Can. J. Chem.* **44** (1966) 677.
17. Robertson, R. E. and Scott, J. M. W. *J. Chem. Soc.* **1961** 1597.
18. Moelwyn-Hughes, E. A., Robertson, R. E. and Sugamori, S. E. *J. Chem. Soc.* **1965** 1965.
19. Leffek, K. T., Robertson, R. E. and Sugamori, S. E. *J. Am. Chem. Soc.* **87** (1965) 2097.
20. Queen, A. and Robertson, R. E. *J. Am. Chem. Soc.* **88** (1966) 1363.
21. Hendy, B. N., Redmond, N. A. and Robertson, R. E. *Can. J. Chem.* **45** (1967) 2071.
22. Robertson, R. E. *Can. J. Chem.* **35** (1957) 613.
23. Robertson, R. E. *Can. J. Chem.* **33** (1955) 1536.
24. Robertson, R. E., Stein, A. and Sugamori, S. E. *Can. J. Chem.* **44** (1966) 685.
25. Hamilton, G. A. and Robertson, R. E. *Can. J. Chem.* **37** (1959) 966.
26. Barnard, P. W. C. and Robertson, R. E. *Can. J. Chem.* **39** (1961) 881.
27. Robertson, R. E. and Sugamori, S. E. *Can. J. Chem.* **44** (1966) 1728.
28. Hyne, J. B., Wills, R. and Wonkka, R. E. *J. Am. Chem. Soc.* **84** (1962) 2914.
29. Bensley, B. and Kohnstam, G. *J. Chem. Soc.* **1956** 287.
30. Hakka, L., Queen, A. and Robertson, R. E. *J. Am. Chem. Soc.* **87** (1965) 161.
31. Cowie, G. R., Fitches, H. J. M. and Kohnstam, G. *J. Chem. Soc.* **1963** 1585.
32. Rudal, E. *Program description RNUM-361.00*, University of Umeå Computing Center, June 1965.

Received January 14, 1970.