

## NEW STATISTICAL PARAMETERS FOR DESCRIPTION OF SOLVENT EFFECT

Oldřich PYTELA

*Department of Organic Chemistry, Institute of Chemical Technology, 532 10 Pardubice*

Received February 22, 1988

Accepted March 31, 1988

A new three-parameter statistical scale has been suggested to describe the solvent effect. The parameters were obtained by applying the method of conjugated deviations (suggested earlier by the author) to a set of 378 data series describing the solvent effect in various chemical and physical processes. The first parameter PA describes the solvent acidity in a medium of a given polarity, the second parameter PB predominantly reflects the solvent polarity with a contribution of basicity, and the third parameter PP refers to the polarity with a contribution of polarizability. The correlation equation, though additive by its form, is de facto an additive-multiplicative relation. Deviations of the parameters indicate a great variability of the solvent effect due to specific manifestations in particular systems. The correlations using the definition set of processes were most successful in the case of the data obtained from kinetic measurements and electronic and infrared spectra. Less significant correlations were obtained for the data from NMR and EPR spectroscopies and measurements of equilibrium constants. Distinctly worse results were found for other processes. The comparison with four most significant empirical equations applied to the same data was unambiguously in favour of the scale suggested (except for the interpretation of EPR data).

The previous review<sup>1</sup> was focused on a survey and evaluation of the empirical models used for description of solvent effects on processes in solutions. In detail analyzed were the empirical parameter scales from the point of view of applicability and physical meaning of the parameters used. The parametrization by Kamlet & Taft<sup>2</sup> was found most successful when applied to a set of 368 experimental data series. This three-parameter equation has a linear form with the parameters describing the acidity, basicity, and polarity/polarizability of solvents. The comparison with equations having both more and less parameters showed that the form of linear additive relation with three parameters represents the optimum for the description of solvent effect from both practical and theoretical points of view. The testing involved only one empirical equation with statistical parameters (two of them were defined) suggested by Swain et al.<sup>3</sup> In the context of the other five equations tested, this one assumed the third place. Although the parameters were adjusted at the optimum in the definition set, their number appeared to be insufficient, and, in our opinion, this is the only main reason of the failure. Beside the scale by Swain et al. several other scales were published<sup>4-11</sup> which are of less practical importance; they<sup>3</sup>

were mostly obtained by the method of principal components or by factor analysis. Generally it can be stated that a priori the statistical methods applied to a sufficiently representative set of experimental data can provide a better set of parameters than that obtained with application of model processes and substances.

Therefore the aim of this present work was to obtain a new statistical scale<sup>7</sup> for description of solvent effect by means of the method of conjugated deviations<sup>12</sup> applied to the set of the data series of ref.<sup>1</sup>, to analyze the physico-chemical meaning of the individual parameters, and to evaluate the practical applicability of the scale suggested by comparison with the other empirical scales used.

### THEORETICAL AND CALCULATIONS

For an adequate quantitative empirical description of solvent effect it is necessary to know the form of the correlation equation and the number of parameters. So far practically successful was the linear additive form based on the LFER principle (a survey for solvents see in ref.<sup>1</sup>). The number of parameters cannot be determined until on the basis of calculation (the method of principal components, factor analysis etc.) or it must be given in advance as it is the case with the optimization methods inclusive of the method of conjugated deviations. The method of conjugated deviations, which was successfully used for the parametrization of nucleophilicity<sup>12</sup>, is based on the linear regression. The basic idea of this method is that the closeness of the correlation between a dependent random variable and an independent non-random variable in a linear regression is a function of not only the experimental error of the dependent random variable but also the value of the independent non-random variable. For every set of linear dependences of experimental data on parameters it is possible to find such a set of parameters which will fulfil the condition of the minimum residual sum of squares simultaneously in all the regressions. There are three differences from the similarly defined condition in the method of principal components or in factor analysis. First of all, in the method of the conjugated deviations the optimum parameters are sought simultaneously and not successively. The second difference consists in the exclusion of statistically insignificant regression coefficients and, hence, also parameters in the individual equations. The third difference consists in non-orthogonality of the parameters obtained.

From the facts given it follows that the magnitude and direction of the correction of deviation of independent non-random variable can be derived from the magnitude and direction of the deviation of the dependent variable from the regression straight line. Hence, both quantities are mutually conjugated, wherefrom the name of this method was derived. Let us define a set of regression equations in the form (I)

$$\hat{y}_{kj} = a_{k0} + \sum_{i=1}^p a_{ki} P_{ij}, \quad k = 1, 2, \dots, m; \quad j = 1, 2, \dots, n, \quad (I)$$

where  $\hat{y}_{kj}$  means the estimate of the dependent random variable in the  $k$ -th data set for (in this case) the  $j$ -th solvent,  $P_{ij}$  means the  $i$ -th parameter for the  $j$ -th solvent, and  $a_{k0}$ ,  $a_{ki}$  are statistically significant regression coefficients. The expression (2) proved to be practically advantageous for the correction  $\Delta P_{ijk}$  of the independent variable (parameter),

$$\Delta P_{ijk} = a_i(y_{kj} - \hat{y}_{kj}) / \left( \sum_{i=1}^p |a_i| \right),$$

$$i = 1, 2, \dots, p, \quad j = 1, 2, \dots, n, \quad k = 1, 2, \dots, m, \quad (2)$$

where the meaning of symbols is the same as that in Eq. (1), and  $y_{kj}$  is the experimental value of the dependent variable standardized to the zero average and unit standard deviation. The statistically insignificant regression coefficients (according to the  $t$ -test carried out in each iteration step) are not considered (are taken as being zero). With respect to the standardization of the  $y_{kj}$  variable in Eq. (2), the  $\Delta P_{ijk}$  deviations are comparable, and a new parameter value ( $P_{ij}^{\text{new}}$ ) can be obtained as a correction of the old value ( $P_{ij}^{\text{old}}$ ) according to Eq. (3)

$$P_{ij}^{\text{new}} = P_{ij}^{\text{old}} + \alpha \sum_{k=1}^m \Delta P_{ijk}. \quad (3)$$

Generally, for a given  $i$  the  $P_{ij}$  values must be standardized to the interval  $\langle 0, 1 \rangle$ . The iterative repetition of the procedure described with a suitable choice of the damping term  $\alpha$  in Eq. (3) provides the  $P_{ij}$  parameters fulfilling (in the experimental data set given) the condition of the minimum residual sum of squares. The choice of the initial estimates of  $P_{ij}$  parameters is not substantial for the result, but it can affect the process (in making the divergence predominant over the convergence) and the number of iterations. E.g. the principal components or scores from the factor analysis represent good initial estimates. The method of conjugated deviations (as compared with the method of principal components or the factor analysis) is not sensitive to missing data, the parameters found have physical meaning, it is easily algorithmized, and it has low demands with respect to the computer memory and time. As compared with the other optimizing methods, e.g. that used by Swain et al.<sup>3</sup>, it is more efficient.

For the reasons given, the method of conjugated deviations was used for a construction of parameter scale for description of the solvent effect. On the basis of the above-given analysis we chose a three-parameter equation ( $p = 3$ ) in Eq. (1) for a set of 51 solvents ( $n = 51$ , Table I). The data set consisted of 368 data series described in the previous communication<sup>1</sup> completed by the  $E_T(30)$  parameters by Dimroth & Reichardt<sup>13</sup>, ACITY and BASITY by Swain et al.<sup>3</sup>, Y, P, E, and B by Koppel & Palm<sup>14</sup>, and the parameters  $W$ ,  $f(n^2)$ , and  $W \cdot f(n^2)$  by Bekárek et al.<sup>15</sup> (altogether  $m = 378$ ). For the initial estimate used were the parameters by Kam-

TABLE I

The statistical parameters PA, PB, and PP for description of solvent effect, their deviations,  $s$ , according to Eq. (4), and numbers  $M$  of the statistically significant corrections

No	Solvent	PA $s$	( $M$ )	PB $s$	( $M$ )	PP $s$	( $M$ )
1	Hexane	0.074 0.198	(93)	0.014 0.158	(97)	0.000 0.181	(111)
2	Heptane	0.064 0.147	(39)	0.000 0.181	(35)	0.002 0.131	(49)
3	Cyclohexane	0.084 0.250	(112)	0.036 0.138	(120)	0.051 0.224	(139)
4	Benzene	0.141 0.214	(143)	0.243 0.300	(129)	0.459 0.206	(154)
5	Toluene	0.129 0.166	(85)	0.262 0.094	(77)	0.423 0.151	(102)
6	<i>m</i> -Xylene	0.089 0.270	(20)	0.214 0.144	(24)	0.359 0.246	(24)
7	<i>p</i> -Xylene	0.103 0.279	(41)	0.244 0.116	(40)	0.377 0.234	(58)
8	Mesitylene	0.080 0.226	(28)	0.200 0.165	(29)	0.342 0.186	(41)
9	Tetrachloromethane	0.135 0.221	(128)	0.068 0.249	(121)	0.267 0.208	(145)
10	Chloroform	0.340 0.292	(115)	0.332 0.194	(105)	0.625 0.290	(117)
11	Dichloromethane	0.297 0.270	(111)	0.424 0.171	(113)	0.658 0.257	(123)
12	1,2-Dichloroethane	0.266 0.170	(80)	0.471 0.128	(89)	0.675 0.150	(103)
13	Chlorobenzene	0.192 0.344	(84)	0.316 0.120	(78)	0.577 0.328	(92)
14	Bromobenzene	0.195 0.181	(64)	0.298 0.169	(57)	0.586 0.172	(71)
15	Fluorobenzene	0.204 0.111	(25)	0.296 0.219	(26)	0.534 0.125	(20)
16	Diethyl ether	0.117 0.188	(127)	0.416 0.157	(114)	0.264 0.185	(131)
17	Dibutyl ether	0.099 0.160	(53)	0.338 0.175	(55)	0.213 0.157	(55)

TABLE I  
 (Continued)

No	Solvent	PA <i>s</i>	( <i>M</i> )	PB <i>s</i>	( <i>M</i> )	PP <i>s</i>	( <i>M</i> )
18	Diisopropyl ether	0.093 0.132	(24)	0.408 0.096	(25)	0.238 0.126	(26)
19	Anisole	0.176 0.130	(46)	0.460 0.262	(44)	0.587 0.137	(42)
20	Phenetole	0.234 0.590	(20)	0.337 0.188	(18)	0.424 0.575	(21)
21	Tetrahydrofurane	0.143 0.133	(130)	0.611 0.113	(131)	0.508 0.146	(108)
22	Dioxane	0.165 0.205	(139)	0.531 0.225	(139)	0.500 0.211	(132)
23	Acetone	0.213 0.173	(164)	0.714 0.216	(152)	0.635 0.196	(128)
24	Butanone	0.193 0.146	(62)	0.716 0.201	(55)	0.609 0.153	(57)
25	Cyclohexanone	0.165 0.177	(44)	0.661 0.094	(41)	0.611 0.179	(43)
26	Methyl acetate	0.216 0.247	(38)	0.631 0.198	(49)	0.477 0.234	(42)
27	Ethyl acetate	0.174 0.192	(116)	0.569 0.158	(129)	0.484 0.192	(117)
28	Acetanhydride	0.352 0.208	(24)	0.740 0.164	(21)	0.695 0.236	(18)
29	Formamide	0.659 0.275	(55)	0.872 0.307	(48)	1.000 0.337	(37)
30	N,N-Dimethylformamide	0.220 0.158	(136)	0.871 0.178	(139)	0.760 0.173	(114)
31	N,N-Dimethylacetamide	0.174 0.142	(81)	0.903 0.153	(77)	0.757 0.173	(55)
32	Hexamethylphosphoric triamide	0.061 0.276	(64)	1.000 0.210	(64)	0.765 0.320	(48)
33	Acetonitrile	0.299 0.231	(181)	0.784 0.268	(157)	0.720 0.256	(147)
34	Benzonitrile	0.245 0.190	(64)	0.621 0.129	(57)	0.734 0.219	(48)

TABLE I  
(Continued)

No	Solvent	PA <i>s</i>	( <i>M</i> )	PB <i>s</i>	( <i>M</i> )	PP <i>s</i>	( <i>M</i> )
35	Water	1.000 0.325	(75)	0.898 0.238	(67)	0.885 0.323	(76)
36	Methanol	0.669 0.198	(158)	0.746 0.187	(129)	0.598 0.214	(136)
37	Ethanol	0.578 0.161	(138)	0.724 0.145	(115)	0.550 0.173	(120)
38	1-Butanol	0.526 0.185	(91)	0.708 0.184	(86)	0.509 0.197	(80)
39	2-Propanol	0.513 0.197	(107)	0.712 0.170	(97)	0.522 0.201	(102)
40	tert. Butyl alcohol	0.413 0.205	(73)	0.683 0.251	(64)	0.483 0.211	(69)
41	Benzyl alcohol	0.559 0.291	(51)	0.734 0.169	(51)	0.714 0.278	(56)
42	1,2-Ethanediol	0.736 0.194	(66)	0.845 0.145	(55)	0.805 0.202	(61)
43	2-Methoxyethanol	0.514 0.178	(33)	0.770 0.143	(27)	0.652 0.194	(28)
44	Acetic acid	0.728 0.197	(62)	0.680 0.199	(54)	0.559 0.237	(43)
45	Triethylamine	0.000 0.278	(56)	0.425 0.359	(60)	0.111 0.273	(58)
46	Pyridine	0.174 0.255	(99)	0.772 0.356	(90)	0.707 0.286	(79)
47	Nitromethane	0.345 0.299	(95)	0.704 0.294	(88)	0.812 0.318	(84)
48	Nitrobenzene	0.232 0.194	(72)	0.555 0.133	(65)	0.748 0.194	(72)
49	Dimethyl sulphoxide	0.256 0.369	(128)	0.987 0.190	(139)	0.862 0.402	(108)
50	Sulpholane	0.282 0.129	(23)	0.813 0.200	(26)	0.827 0.143	(19)
51	Carbon disulphide	0.151 0.246	(57)	0.051 0.252	(59)	0.385 0.238	(61)

let & Taft<sup>2</sup>, because a calculation of principal components or factor scores was technically impossible for such a large set. The missing values were assessed from analogy with similar compounds. One hundred iterations were carried out, which reliably ensured the corrections in Eq. (2) to be below 0.001. The variability of parameters in the used set of data series was expressed by the deviation calculated from the correction  $\Delta P_{ijk}$  for the optimum parameters according to Eq. (4)

$$s_{ij} = [(\sum_{k=1}^m \Delta P_{ijk}^2)/(M - 1)]^{1/2}, \quad (4)$$

where  $M$  symbolizes the number of cases for which this correction was statistically significant. All the calculations were realized on an EC 1033 computer according to our own programs in the FORTRAN language.

### RESULTS AND DISCUSSION

Table I summarizes the statistical parameters PA, PB, and PP for description of solvent effect which were obtained by the method of conjugated deviations; moreover it gives their deviations according to Eq. (4) and numbers of the statistically significant corrections  $M$ . Out of the total number of 378 data series seven were not used for the optimization due to lack of data (Nos 219–225 according to the numbering used in the previous communication<sup>1</sup>), and 13 data series exhibited statistically insignificant all regression coefficients in the regression with the optimum parameters.

The parameters given in Table I are relative parameters, being related to the reference solvent with the smallest value. This relativity, which is not uncommon with empirical scales (cf. e.g. the  $\pi^*$  parameter by Kamlet & Taft<sup>2</sup>), is not prejudicial to the correlations used for analysis of the nature of processes and for predictions of results in a certain solvent. However, relative parameters do not make it possible to predict the result of a process in the absence of solvent, i.e. in vacuum. In order to remove this drawback, we must add a constant to every parameter, this constant expressing the difference between the values of the parameter in a reference solvent and in vacuum. For the estimate of the constant in statistical sense it is possible to use the regression of dependence of the relative empirical parameter on some theoretical, semiempirical or even empirical quantity with unequivocally defined value in vacuum. The best suited parameter for an adequate estimate of the shift is the PP parameter which correlates very well with the semiempirical parameters by Bekárek et al.<sup>15</sup> ( $R = 0.992$ ) and gives the shift value  $\Delta PP = 0.89 \pm 0.02$ . A close value  $\Delta PP = 0.95 \pm 0.15$  can be obtained with application of the equation by Koppel & Palm<sup>14</sup>, although the correlation is worse ( $R = 0.958$ ). A less favourable situation is encountered with the remaining two parameters. In both the cases the best correlation is provided with the equation by Koppel & Palm<sup>14</sup>, the shift value obtained for the PA parameter being  $\Delta PA = 0.002 \pm 0.024$  ( $R = 0.983$ ), i.e. zero from the

statistical point of view. For the PB parameter the shift estimate is  $\Delta PB = 0.309 \pm 0.047$  ( $R = 0.958$ ). More precise estimates of the shifts cannot be obtained without the knowledge of the physico-chemical meaning of the parameters presented.

An at least approximative idea about the predominant character of the individual parameters can be obtained on the basis of comparison with other empirical or semi-empirical parameters. The PA parameter correlates best with the E parameters by Koppel & Palm<sup>14</sup> ( $r = 0.962$ ),  $E_T(30)$  by Reichardt & Dimroth<sup>13</sup> ( $r = 0.940$ ), ACITY by Swain et al.<sup>3</sup> ( $r = 0.930$ ), and  $\pi^*$  by Kamlet & Taft<sup>2</sup> ( $r = 0.909$ ). From the survey given it is obvious that the PA parameter first of all describes the solvent acidity. The nature of the PB parameter is very specific, the closest connection was found ( $r = 0.890$ ) with the semiempirical polarity characteristics  $(\epsilon - 1)/(2\epsilon + 1)$ . The PP parameter correlates relatively well ( $r = 0.967$ ) with the  $\pi^*$  parameter by Kamlet & Taft<sup>2</sup> referring to the solvent polarity and polarizability.

A more comprehensive evaluation of the relation between the parameters suggested and other empirical and semiempirical parameters can be based on their dislocation in the factor space. The nearness of individual parameters-objects was evaluated by means of the cluster analysis in the factor space with application of the centroidal method in the Euclid metrics. If the calculation only involved the parameters  $E_T(30)$  (ref.<sup>13</sup>), ACITY, BASITY (ref.<sup>3</sup>),  $\pi^*$ ,  $\alpha$ ,  $\beta$  (ref.<sup>2</sup>),  $W$ ,  $f(n^2)$ ,  $W \cdot f(n^2)$  (ref.<sup>15</sup>),  $Y$ ,  $P$ ,  $E$ ,  $B$  (ref.<sup>14</sup>), and PA, PB, PP, then four stable clusters were obtained. The PA parameter was gradually united with the parameters ACITY,  $\alpha$ ,  $E$ , and  $E_T(30)$ , i.e. those describing predominantly the solvent acidity. The PB parameter was gradually associated with the parameters  $Y$  and  $W$ , and, at a somewhat greater distance, with  $\beta$  and  $B$ . These results indicate a superposition of basicity and predominant polarity. Finally, the PP parameter gradually formed a cluster with the parameters  $\pi^*$ , BASITY, and  $W \cdot f(n^2)$ , the polarity again being dominant but completed with polarizability in this case. If the other parameters were involved (see Table III in ref.<sup>1</sup>), analogous results were obtained. The PA parameter was also associated, beside the parameters mentioned, with  $\delta_H$  by Hildebrand<sup>16</sup>,  $Z$  by Kosower<sup>17</sup>,  $a^{14N}$  by Knauer & Napier<sup>18</sup>,  $S$  by Brownstein<sup>19</sup>,  $\Phi$  by Dubois<sup>20,21</sup>,  $S$  by Zelinskii<sup>22</sup>, AN by Mayer & Gutmann<sup>23</sup>, and some others. Similarly, the PB parameter appeared analogous to  $\epsilon_{Al_2O_3}$  by Snyder<sup>24</sup> and DN by Gutmann et al.<sup>23</sup>. The PP parameter fell into a common cluster with the  $\int_H^{NO}$  parameter by Taft et al.<sup>25</sup> (beside the other parameters mentioned). The  $P$ ,  $f(n^2)$  parameters and connectivity<sup>26</sup> with  $\chi_R$  by Brooker<sup>27</sup> formed separate groups.

From the facts given it is obvious that the PA parameter is specified relatively unequivocally, expressing a measure of solvent acidity in a medium of a given polarity. Less obvious is the physico-chemical nature of the PB and PP parameters. Both the parameters are correlated ( $r = 0.805$ ). The solvents dislocated in the plane of these parameters form an oblong cluster (Figs 1, 2), its main axis being given by the solvent sequence: hexane, heptane, cyclohexane, water, hexamethylphosphoric triamide,



formamide, and dimethyl sulfoxide. This sequence corresponds to increasing solvent polarity. Quantitatively this property can be expressed by the first principal component which is calculated as the arithmetic mean from PB and PP in this case. This quantity correlates relatively well ( $r = 0.940$  without dioxane) with the semiempirical polarity characteristic  $(\epsilon - 1)/(2\epsilon + 1)$ . The inclusion of the remote value of dioxane results in a lowering of the correlation coefficient ( $r = 0.920$ ). Deviations from the main axis towards larger PB values are exhibited by the solvents having relatively large basicity as related to their polarizability (triethylamine), those towards larger PP values are observed with the solvents having relatively large polarizability as related to their basicity (carbon disulfide, chlorinated solvents). Hence, the PB parameter is a measure first of all of polarity completed by basicity, and the PP parameter analogously is a measure of polarity completed by polarizability. The parameters PB and PP are not redundant, because their sum (as well as the arithmetic mean) expresses the polarity, whereas their difference corresponds to the difference between basicity and polarizability.

The already mentioned non-orthogonality between the PB and PP parameters is not an isolated case; similar non-orthogonality also exists between the parameters PA, PB ( $r = 0.551$ ) and PA, PP ( $r = 0.549$ ). A high correlation between parameters indicates either their unsuitable adjustment or their redundancy. On the other hand it is possible to find at least four independent properties described by the parameters to some extent or other, which on the contrary indicates their small number. The explanation of this discrepancy can be found in a more complex model. Our previous communication<sup>1</sup> has already mentioned a possibility of interaction between the individual fundamental manifestations of a solvent. For  $Q$  properties, the structure

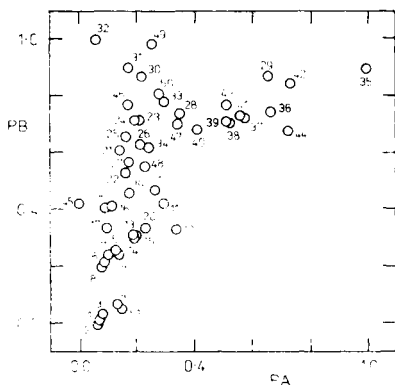


FIG. 1

Dependence of PB parameter on PA parameter. For the numbers of points see Table I

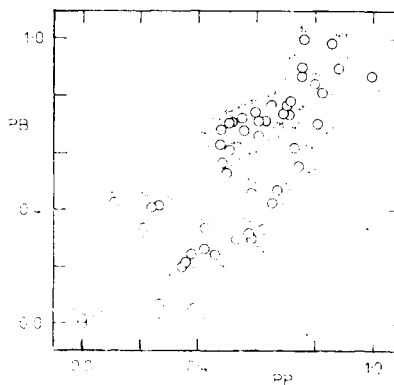


FIG. 2

Dependence of PB parameter on PP parameter. For the numbers of points see Table I

of an empirical parameter can generally be described by the equation

$$P = \pi_0 + \sum_{i=1}^Q \kappa_i \pi_i + \sum_{i=1}^Q \sum_{j=i+1}^Q \kappa_{ij} \pi_i \pi_j + \dots, \quad (5)$$

where  $\pi$  means a quantitative expression of the intrinsic basic property, and  $\kappa$  reflects the measure of its contribution. The  $\pi_0$  coefficient represents the correction for the standard state, vacuum being the best one. The correlations between the suggested parameters PA, PB, and PP can be explained by means of Eq. (5), the  $Q$  quantity being equal to four, if we consider the four usual fundamental properties of solvents – acidity, basicity, polarity, and polarizability. Obviously, some or perhaps the majority of the  $\kappa$  coefficients are insignificant. The complex character of an empirical parameter also has its practical advantages, because a suitable combination (26 such possibilities exist for three parameters if their sign is also taken into account) can describe several fundamental properties, the number of properties being greater than that of the parameters. In the data set tested (see below), each combination was applied three times at least and 32 times at most.

The deviations of parameters given in Table I are by one order of magnitude greater than the same quantities obtained in the parametrization of nucleophilicity<sup>12</sup>. The exclusion of 20% of the data series with the highest residual variance results in a decrease of the parameter deviations to one half. Therefrom it follows that large deviation values are a property of the set parametrized and are not caused by few extremely deviated cases. In terms of the factor analysis it could be said that solvents have a small communality as compared with nucleophiles (i.e. the same character of operation), and, hence, a great specificity (individuality of operation) at a presumed equal experimental error of data. Undoubtedly, this is due to the large variability of the processes in which the solvent effect makes itself felt. On the other hand, specific deviations can represent a source of additional information about the process studied.

Application of the parameters suggested to the test processes of ref.<sup>1</sup> provided correlations whose closeness expressed by the multiple correlation coefficient is given in Table II. From the table it follows that the data were not interpreted in 13 cases (3.6% of the total number), out of which in six cases (1.7%) the data were interpretable by none of the correlation relations tested<sup>1</sup>. The remaining cases involved the chemical shifts in the <sup>19</sup>F NMR spectrum of 1,1,2,2-tetrachloro-3,3,4,4-tetrafluorocyclobutane (No. 172 of the list in ref.<sup>1</sup>), <sup>23</sup>Na NMR spectrum of sodium iodide (No. 196) and sodium rhodanide (No. 198), the logarithm of rate constant of dimerization of pentadiene (No. 267), the logarithm of the rate constants ratio  $\log(k_{\text{tert}}/k_{\text{prim}})$  of photochlorination of 2,3-dimethylbutane (No. 295), and finally the logarithm of equilibrium constants of tautomerization of 5,5-dimethyl-1,3-cyclohexanedione (No. 301) and of formation of complexes between iodine and solvent (No. 308).

TABLE II

The multiple correlation coefficients  $R$  obtained by application of the PA, PB, and PP parameters to the results of test processes; for numbering and description of the processes see ref.<sup>1</sup>

No.	$R$	No.	$R$	No.	$R$	No.	$R$	No.	$R$
Electronic absorption and fluorescence spectroscopy									
1	0.995	2	0.978	3	0.973	4	0.957	5	0.974
6	0.986	7	0.996	8	0.980	9	0.995	10	0.974
11	1.000	12	0.968	13	0.976	14	0.977	15	0.920
16	0.876	17	0.968	18	0.973	19	0.985	20	0.968
21	0.677	22	0.831	23	0.935	24	0.991	25	0.996
26	0.934	27	0.911	28	0.964	29	0.926	30	0.993
31	0.969	32	0.849	33	0.840	34	0.915	35	0.882
36	0.941	37	0.974	38	0.969	39	0.968	40	0.972
41	0.882	42	0.804	43	0.774	44	0.977	45	0.852
46	0.713	47	0.753	48	0.834	49	0.990	50	0.982
51	0.999	52	0.975	53	0.996	54	0.984	55	0.756
56	0.993	57	0.989	58	0.989	59	0.725	60	0.976
61	0.867	62	0.788	63	0.906	64	0.979	65	0.995
66	0.975	67	0.996	68	0.999	69	0.978	70	0.837
71	—	72	0.948	73	0.998	74	0.969	75	0.985
76	0.989	77	0.992	78	0.981	79	0.993	80	0.934
81	0.996	82	0.988	83	0.992	84	0.995	85	0.962
86	0.993	87	0.981	88	0.797	89	0.957	90	0.992
91	0.988	92	0.990	93	0.968	94	0.989	95	0.981
96	0.976	97	0.983	98	0.958	99	0.984	100	0.989
101	0.977	102	0.833	103	0.932	104	0.973	105	0.960
106	0.987	107	0.895	108	0.963	109	0.990	110	0.981
111	0.987	112	0.990	113	0.996	114	0.674	115	0.995
116	0.925	117	0.987	118	0.980	119	0.995	120	0.982
121	0.796	122	0.865						
Infrared spectroscopy									
				123	0.961	124	0.946	125	0.969
126	0.826	127	0.907	128	0.941	129	—	130	0.926
131	0.948	132	0.897	133	0.938	134	0.973	135	0.897
136	0.951	137	0.962	138	0.894	139	0.893	140	0.996
141	0.986	142	0.986	143	0.994	144	0.991	145	0.973
146	0.965	147	0.965	148	0.982	149	0.994	150	0.992
151	0.976	152	0.994	153	0.738	154	0.975		

TABLE II  
(Continued)

No.	R	No.	R	No.	R	No.	R	No.	R
Nuclear magnetic resonance									
156	0.863	157	—	158	—	159	—	155	0.928
161	0.669	162	0.762	163	0.987	164	0.899	160	0.725
166	0.540	167	0.876	168	0.941	169	0.996	165	0.944
171	0.963	172	—	173	0.971	174	0.930	170	0.990
176	0.974	177	0.952	178	0.905	179	0.949	175	0.929
181	0.986	182	0.942	183	0.938	184	0.894	180	0.951
186	0.901	187	0.897	188	0.972	189	0.934	185	0.889
191	0.955	192	0.977	193	0.989	194	0.981	190	0.959
196	—	197	0.775	198	—	199	0.863	195	—
201	0.942	202	0.874	203	0.769	204	0.973	200	0.674
206	0.967	207	0.906					205	0.963
Electron paramagnetic resonance									
				208	0.991	209	0.997	210	0.995
211	0.994	212	0.999	213	0.807	214	0.867	215	0.988
216	0.992	217	0.991	218	0.688				
Kinetic processes									
221	<sup>a</sup>	222	<sup>a</sup>	223	<sup>a</sup>	219	<sup>a</sup>	220	<sup>a</sup>
226	0.983	227	0.992	228	0.997	224	<sup>a</sup>	225	<sup>a</sup>
231	0.991	232	0.986	233	0.997	229	0.995	230	0.996
236	0.992	237	0.961	238	0.978	234	0.996	235	0.982
241	0.993	242	0.985	243	0.994	239	0.988	240	0.979
246	0.968	247	0.981	248	0.973	244	0.992	245	0.946
251	0.994	249	0.938	253	0.949	249	0.938	250	0.911
256	0.971	252	0.973	258	0.939	254	0.968	255	0.990
261	0.993	257	0.951	263	0.998	259	0.930	260	0.984
266	0.978	262	0.995	268	0.936	264	0.932	265	0.844
271	0.976	267	—	273	0.788	269	0.909	270	0.974
276	0.878	272	0.922	278	0.979	274	0.986	275	0.911
281	0.921	277	0.975	283	0.989	279	0.975	280	0.668
286	0.995	282	0.823	288	0.991	284	0.893	285	0.981
291	0.911	287	0.878	293	0.991	289	0.985	290	0.934
296	0.973	292	0.990			294	0.631	295	—
		297	0.940						

TABLE II  
(Continued)

No.	R	No.	R	No.	R	No.	R	No.	R
Equilibrium processes									
				298	0.906	299	0.923	300	0.971
301	—	302	0.826	303	0.896	304	0.814	305	0.731
306	0.973	307	0.972	308	-	309	0.989	310	0.974
311	0.959	312	0.973	313	0.890	314	0.931	315	0.990
316	0.972	317	0.869	318	0.845	319	0.886	320	0.963
321	0.937	322	0.926	323	0.958	324	0.997	325	0.976
326	0.993	327	0.991	328	0.985	329	0.977	330	0.736
331	0.677								
Other processes									
		332	0.823	333	0.826	334	0.887	335	0.892
336	0.929	337	0.925	338	0.867	339	0.870	340	0.852
341	0.956	342	0.853	343	0.979	344	0.958	345	0.984
346	0.953	347	0.986	348	0.988	349	0.927	350	0.599
351	0.920	352	0.828	353	0.855	354	0.875	355	0.686
356	0.843	357	0.931	358	0.917	359	0.911	360	0.877
361	0.845	362	0.895	363	0.940	364	0.667	365	0.948
366	0.708	367	0.652	368	0.801				

<sup>a</sup> Insufficient number of data.

The cases given mostly involve processes in which solvent effects can be expected to a little extent only. Moreover, from Table II it follows that the closest correlations were found for the data obtained from the electron absorption and fluorescence spectroscopy, infrared spectroscopy, or kinetic measurements. Average correlations are obtained for the data from NMR and EPR spectrometry and logarithms of equilibrium constants. Distinctly worse are the correlations for the other non-included processes. First of all, the electronic spectra and kinetic processes form the strongest groups (122 and 79 data series, respectively), hence their effect on the final parameter values will be the greatest one. The next point is that the processes of the first group exhibit the highest sensitivity to the solvent effects, which is documented by the number of the empirical scales suggested on the basis of these methods. Last but not least, the homogeneity of data in the sense of operation of solvent makes itself felt, this homogeneity being naturally the worst in the other processes.

The set of the examples tested can also serve for an estimate of the validity range of the equation and parameters suggested. In the regressions with standardized values of dependent variable (zero mean, unit standard deviation) the following average values of residual variance were obtained for the individual types of processes: the electron spectroscopy 0.109, infrared spectroscopy 0.113, NMR spectrometry 0.185, EPR spectrometry 0.125, kinetic processes 0.102, equilibrium processes 0.156, other processes 0.261. The overall mean value is 0.139. The order of successfulness and its rationalization are identical with those of the correlation analysis. The equation and parameter scale suggested can be applied to a great number of processes in which solvent acts as a medium or also as a reactant. In these cases the results can be interpreted roughly with the accuracy given. The properties of the solvents themselves were not tested, however, it is possible to expect less successful correlations.

The test results can also serve for comparison of quality of the parameters suggested with that of the other empirical scales used<sup>1</sup>. The average order obtained by comparison of residual deviations (which agrees with the order of multiple correlation coefficients in most cases) for the individual types of processes is given in Table III. The data of this table unequivocally show the advantages of the parameter scale suggested both in the definition and tested sets of data series. With respect to the scope and variability of the set tested it can be expected that the parameters suggested will prove competent also for other data within their validity range.

TABLE III

Average orders of successfulness in interpretation of results of the processes tested (ref.<sup>1</sup>) obtained by comparison of residual standard deviations in the regressions with the parameters by Dimroth & Reichardt (DR, ref.<sup>1,3</sup>), Swain et al. (SSPA, ref.<sup>3</sup>), Kamlet, Taft et al. (KAAT, ref.<sup>2</sup>), Koppel & Palm (KP, ref.<sup>1,4</sup>), and with the parameters PA, PB, and PP suggested in this present communication (P)

Process	Empirical equation				
	DR	SSPA	KAAT	KP	P
UV-VIS-FLUOR	4.60	3.22	2.22	2.80	2.17
IR	4.84	2.84	2.19	3.19	1.94
NMR	4.49	3.54	2.17	2.87	1.93
EPR	5.00	1.55	2.73	2.73	3.00
log <i>k</i>	4.41	2.76	2.83	3.03	1.97
log <i>K</i>	4.10	3.28	2.65	3.09	1.88
Others	4.22	3.58	3.04	2.54	1.62
Altogether	4.49	3.13	2.48	2.89	2.02

## REFERENCES

1. Pytela O.: Collect. Czech. Chem. Commun. 53, 1333 (1988).
2. Kamlet M. J., Abboud J.-L. M., Abraham H. M., Taft R. W.: J. Org. Chem. 48, 2877 (1983).
3. Swain C. G., Swain M. S., Powel A. L., Alunni S.: J. Am. Chem. Soc. 105, 502 (1983).
4. Chastrette M.: Tetrahedron 35, 1441 (1989).
5. Chastrette M., Carretto J.: Tetrahedron 38, 1615 (1982).
6. Chastrette M., Carretto J.: Can. J. Chem. 63, 3492 (1985).
7. Cramer R. D.: J. Am. Chem. Soc. 102, 1837 (1980).
8. Cramer R. D.: J. Am. Chem. Soc. 102, 1879 (1980).
9. Sjöström M., Wold S.: J. Chem. Soc., Perkin Trans. 2, 1981, 104.
10. Elguero J., Fruchier A.: Ann. Quim., C 79, 72 (1983).
11. Svoboda P., Pytela O., Večeřa M.: Collect. Czech. Chem. Commun. 48, 3287 (1983).
12. Pytela O., Zima V.: Collect. Czech. Chem. Commun. 54, 117 (1989).
13. Reichardt C., Harbusch-Görnert E.: Justus Liebigs Ann. Chem. 1983, 721.
14. Koppel I. A., Palm V. A. in: *Advances in Linear Free Energy Relationships* (N. B. Chapman and J. Shorter, Eds), p. 203. Plenum Press, London 1972.
15. Bekárek V., Bekárek V.: Collect. Czech. Chem. Commun. 52, 287 (1987).
16. Hildebrand J. H., Prausnitz J. M., Scott R. L.: *Regular and Related Solutions*. Van Nostrand-Reinhold, Princeton 1970.
17. Kosower E. M.: J. Am. Chem. Soc. 80, 3252 (1958).
18. Knauer B. R., Napier J. H.: J. Am. Chem. Soc. 98, 4395 (1976).
19. Brownstein S.: Can. J. Chem. 38, 1590 (1960).
20. Dubois J. E., Goetz E., Bienvenüe A.: Spectrochim. Acta 20, 1815 (1964).
21. Dubois J. E., Bienvenüe A.: J. Chim. Phys. 65, 1259 (1968).
22. Zhmyreva I. A., Zelinskii V. V., Kolobkov V. P., Krasnitskaya I. D.: Dokl. Akad. Nauk SSSR, Ser. Khim. 129, 1089 (1959).
23. Gutmann V.: *The Donor-Acceptor Approach to Molecular Interactions*. Plenum Press, New York 1978.
24. Snyder L. R. in: *Techniques of Chemistry* (E. S. Perry and A. Weissberger, Eds) Vol. 12, Chap. 2. Wiley, New York 1978.
25. Taft R. W., Abboud J.-L. M., Kamlet M. J.: J. Am. Chem. Soc. 103, 1080 (1981).
26. Kier L. B., Hall L. H.: Eur. J. Med. Chem. 12, 307 (1977).
27. Brooker L. G. S., Craig A. C., Heseltine D. W., Jenkins P. W., Lincoln L. L.: J. Am. Chem. Soc. 87, 2443 (1965).

Translated by J. Panchartek.