The Role of *meta* and *para* Benzene Derivatives in the Evaluation of Substituent Effects: a Multivariate Data Analysis

Miroslav Ludwig,[a,*] Svante Wold[b] and Otto Exner[c]

[a] Department of Organic Chemistry, Institute of Chemical Technology, CS-532 10 Pardubice, Czechoslovakia, [b] Department of Organic Chemistry, University of Umeå, S-901 87 Umeå, Sweden and [c] Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, CS-166 10 Prague 6, Czechoslovakia

Ludwig, M., Wold, S. and Exner, O., 1992. The Role of *meta* and *para* Benzene Derivatives in the Evaluation of Substituent Effects: a Multivariate Data Analysis. – Acta Chem. Scand. 46: 549–554.

Five aromatic reactivity data sets, selected to be balanced and representative for different types of substituent effect have been investigated with respect to eight extrathermodynamic relationships (the Hammett equation and various extensions). The five data sets all refer to properties of *meta* and *para* substituted phenyl compounds and comprise ionization in solution and in the gas phase of some acids and bases, NMR shifts, IR spectral frequencies and melting points.

The results of principal components analysis (PCA) and projection to latent structures (PLS) analysis of the data give evidence that the effects of *meta* and *para* substituents are essentially different in nature. The behaviour of *para* derivatives is more complex, requiring more terms in the correlation equation: this complexity is due mainly to the donor substituents. Modelling *meta* and *para* derivatives together has mainly historical reasons; it may be acceptable in the simplest cases or when the small number of data prevent a more correct treatment.

Since the discovery of the Hammett equation[1] (1), *meta* and *para* benzene derivatives have taken a unique position as model compounds for the estimation of substituent effects on organic reactivity.[2] No doubt a greater number of quantitative studies has been devoted to compounds of this type than to all other compounds put together. The restriction to these derivatives was motivated by the early discovery[1] that *ortho* derivatives behave less regularly owing to steric and other short-range effects. On the other hand it has never been proved that the behaviour of *meta* and *para* derivatives is the same or at least so similar that their modelling as a common series is substantiated; in terms of the Hammett equation this corresponds to whether the constant ϱ should be the same.[3,4] A linear regression analysis by Jaffé[4] revealed only small differences, hence *meta* and *para* derivatives have currently been treated together. The problem has not been resolved by using more elaborate models,[5–20] extending eqn. (1) by additional terms, see e.g. later, eqns. (2)–(8). All these models express an observable quantity y (e.g. pK) as a linear function of substituent parameters which may be given a physical meaning. In some of these models *meta* and *para* derivatives are analysed together, in others separately. One difficulty comes from the fact that *meta* derivatives have been less extensively investigated, both in most experimental studies and in some recent mathematical treatments.[10–12,19,20]

In this paper we study the problem of *meta* and *para* compatibility using two methods of multivariate analysis: principal components analysis[21] (PCA) and projection to latent structures[22] (PLS). PCA has previously been applied in this field mainly to determine the necessary number of terms in the correlation equation[13,14,16–20] [see e.g. eqns. (1)–(5) which represent models with an increasing number of terms] and to compare the precision attained (usually in % of the explained sum of squares).

Here we have applied PCA first to *meta* and *para* derivatives separately, then to combined data matrices encompassing both. Such a combined data matrix can be constructed in two ways: either each substituent is taken twice as *m*-X and *p*-X, respectively ($\frac{m}{p}$ matrix:, or each reaction series is divided into two series, *meta* and *para*, if the data for both are available ($m|p$ matrix), see Table 1 and Fig. 1. A comparison of the explained variance obtained in all variants can then answer the question of whether or not the *meta* and *para* derivatives can be considered as belonging together.

PLS describes the relationship between two matrices: it derives one or several independent factors from one matrix and uses these to predict the elements of the other matrix. One can ask about the precision of the prediction (again expressing the explained variance as % of the total sum of squares) and about the number of necessary factors. The method has been applied earlier to *meta* and *para* benzene derivatives, using rather heterogeneous properties.[17]

Investigated models. As mentioned, the models[5–17] used in organic reactivity, often called extrathermodynamic rela-

* To whom correspondence should be addressed.

Table 1. Characteristics of various correlation equations for benzene *meta* and *para* derivatives.[a]

	meta	*para*	PCA	PLS		
			$\frac{mb}{p}$	$m\|p$[c]	$m\rightarrow p$	$p\rightarrow m$
(a) Required for individual models						
Hammett,[1] eqn. (1)	1	1	**1**	(2)	–[d]	–[d]
Yukawa–Tsuno,[5] eqn. (2)	1	2	**2**	(3)	–[d]	–[d]
DSP, Taft *et al.*,[6–8] eqn. (3)	2	2	(4)	**2**	2[~100]	2[~100]
3-term eqn.,[10–13] eqn. (4)	3	3	(6)	**3**	3[~100]	3[~100]
Haldna,[14] eqn. (5)	4	4	**4**	(8)	–[d]	–[d]
Exner,[15] eqn. (6)[e]	1	1	**1**	**1**	1[~100]	1[~100]
Exner–Budĕšínský,[16] eqn. (7)[f]	**1**	**2**	(2)[g]	(2)[g]	1[~80]	1[~100]
Alunni,[17] eqn. (8)	**2**	**2**	(4)	(4)	–[d]	–[d]
(b) Found in this work						
p*K* (1)	1[97]	1[98]	1[98]	2[98][h]	3[97]	3[94]
$\Delta G°$ (g)	1[96]	2[99]	2[99]	3[99]	1[87]	5[99]
^{13}C NMR	1[99]	2[99][i]	2[99][j]	3]99]	1[69]	2[94]
IR freq.	1[78]	1[89]	1[87]	1[83]	1[73]	1[73]
M.p.s	1[79]	1[78]	1[77]	1[74]	1[53]	1[72]
(c) Found in this work for acceptors only						
p*K* (1)	1[99]	1[98]			1[97]	1[96]
$\Delta H°$ (g)	2[96][k]	2[99][l]			2[99]	4[99]
^{13}C NMR	1[99]	2[99]			1[96]	1[96]
IR freq.	1[82]	1[78]			1[93]	1[88]

[a]Bold-face numbers correspond to the original form of the correlation equation; in parentheses is an artificial construction, not compatible with this original form; in brackets is the percentage of the explained sum of squares (in the case of PLS this value refers to the **Y** matrix by using scores from **X** matrix). [b]Every substituent taken twice, as *m*-X and *p*-X. [c]Every reaction series taken twice, separately for *meta* and *para* derivatives. [d]No relation anticipated. [e]The relationship restricted only to substituents without a lone electron pair in the α-position (acceptors mainly). [f]Applied only to ^{13}C NMR shifts. [g]Accuracy slightly lowered. [h]There is a further component with very low explaining ability. [i]83 % for the first PC, 16 % for the second. [j]87 % for the first PC, 12 % for the second. [k]54 % for the first PC, 42 % for the second. [l]93 % for the first PC, 6 % for the second.

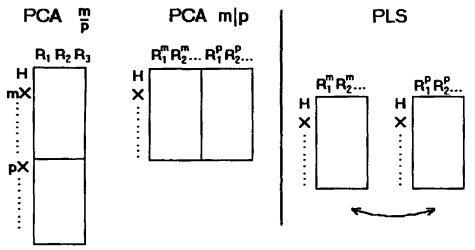

Fig. 1. Arrangement of data matrices in PCA and PLS, explaining the terms 'matrix $\frac{m}{p}$' and 'matrix $m|p$'.

tionships, involve a number of assumptions about the dependence or independence between *meta* and *para* derivatives. These assumptions have not received enough attention which has instead been focused on the mathematical form of the eqns., particularly on the number of necessary terms. Nevertheless, recognition of these assumptions is essential. Their degree of realism is seem in the results of PCA and PLS and is as important as the mathematical form of the equation (Table 1).

$$y_{m,p} = y_o + \varrho\sigma_{m,p} \quad (1)$$

$$y_{m,p} = y_o + \varrho\sigma_{m,p} + \varrho r\,\Delta\sigma_p \quad (2)$$

$$y_{m,p} = y_o + \varrho_I^{m,p}\sigma_I + \varrho_R{}^{m,p}\sigma_R \quad (3)$$

$$y_{m,p} = y_o + L^{m,p}\sigma_I + D^{m,p}\sigma_d + R^{m,p}\sigma_r \quad (4)$$

$$y_{m,p} = y_o + \varrho_1\sigma_{m,p}^{(1)} + \varrho_2\sigma_{m,p}^{(2)} + \varrho_3\sigma_{m,p}^{(3)} + \varrho_4\sigma_{m,p}^{(4)} \quad (5)$$

$$y_{m,p} = y_o + f_{m,p}\,\varrho\sigma \quad (6)$$

$$y_{m,p} = y_o + \varrho_I^{m,p}\,\sigma_{m,p}^{(1)} + \varrho_2{}^{p}\sigma_p^{(2)} \quad (7)$$

$$y_{m,p} = y_o + \varrho_I^{m,p}\,\sigma_{m,p}^{(1)} + \varrho_2^{m,p}\,\sigma_{m,p}^{(2)} \quad (8)$$

(1) *The Hammett equation.* In the original Hammett equation[1] *meta* and *para* derivatives are included in one matrix $\frac{m}{p}$ with different and independent substituent constants. If this is a good model, the fit in this arrangement should be better than in the matrix $m|p$: the latter should require two parameters, one for the *meta* series and one for the *para*

series, one of them always with a zero coefficient. (In the primary result of PCA linear combinations – a rotation – of these factors are obtained, hence the coefficients need not be zero.) The transfer of information between the *m* and *p* matrices in PLS should be small, even zero, in either direction *meta* → *para* or *para* → *meta* (Table 1).

(2) *The Yukawa–Tsuno equation.* The same principle is maintained in the Yukawa–Tsuno equation[5] but the more complex behaviour of *para* derivatives is expressed by two substituent constants in eqn. (2).

(3) *The Taft DSP model.* The widespread Taft DSP (dual substituent parameter) treatment,[6] eqn. (3), is characterized not only by an additional term but also by the separation of the data matrix. This now contains each substituent only once but each reaction series is divided (matrix *m*|*p*). These two features of DSP have seldom been evaluated separately.[16] If the DSP assumptions are valid, it follows that for the matrix $\frac{m}{p}$ four terms would be necessary with two parameters always equal to zero, the other two being equal to σ_I and σ_R. In PLS the transfer of information should approach 100 % in either direction, *meta* → *para* or *para* → *meta* (Table 1).

Table 2. Results of PCA and description of data matrices.

	Matrix	Matrix dimension[a]	Filling (%)	Residual st. deviation		
				PC1	PC2	PC3
p*K*s of benzoic acids in various solvents[c]	*meta*	14× 7	100	0.10	[b]	[b]
	para	14× 7	97	0.11	[b]	[b]
	$\frac{m}{p}$	27× 7	99	0.09	[b]	[b]
	m|*p*	14×14	99	0.14	0.11	0.09
Δ*G*° (g) of various acids[d]	*meta*	11×11	79	1.03	[b]	[b]
	para	11×11	86	1.83	0.91	[b]
	$\frac{m}{p}$	21×11	83	1.62	0.83	[b]
	m|*p*	11×22	83	1.71	1.21	0.72
^{13}C NMR shifts in the alpha position[e]	*meta*	14×10	91	0.12	[b]	[b]
	meta	14×10	92	0.35	0.10	[b]
	$\frac{m}{p}$	27×10	92	0.34	0.11	[b]
	m|*p*	14×20	92	0.28	0.15	0.13
IR frequencies of Various groups[f]	*meta*	23×18	65	4.80	[b]	[b]
	para	23×18	76	5.60	[b]	[b]
	$\frac{m}{p}$	45×18	70	5.00	[b]	[b]
	m|*p*	23×36	70	5.64	[b]	[b]
Melting points of benzene bisderivatives[g]	*meta*	18×11	60	22.7	[b]	[b]
	para	18×11	84	24.5	[b]	[b]
	$\frac{m}{p}$	35×11	72	24.7	[b]	[b]
	m|*p*	18×22	72	25.7	[b]	[b]
(b) Only acceptor data						
p*K*s of benzoic acids in various solvents	*meta*	6× 7	100	0.10	[b]	[b]
	para	6× 7	100	0.06	[b]	[b]
Δ*G*° (g) of various acids	*meta*	6×11	77	4.17	1.38	[b]
	para	6×11	85	1.77	0.84	[b]
^{13}C NMR shifts	*meta*	6×10	93	0.10	[b]	[b]
	para	6×10	95	0.11	[b]	[b]
IR frequencies	*meta*	10×11	70	5.46	[b]	[b]
	para	10×11	82	7.50	[b]	[b]

[a]The two matrices, *meta* and *para*, had always identical both lines and columns. [b]Insignificant terms. [c]Substituents: H, CH_3, $COCH_3$, CN, NO_2, OH, OCH_3, SH, SO_2CH_3, SO_2NH_2, F, Cl, Br, I; solvents: water, methanol, ethanol, acetone, acetonitrile, *N*,*N*-dimethylformamide, sulfolane; Ref. 23; st. deviation in p*K* units. [d]Substituents: H, CH_3, CF_3, CN, $COCH_3$, $COOCH_3$, NO_2, OCH_3, SCH_3, F, Cl; acids: ArCOOH, $ArNH_2$, ArOH, $ArCH(CH_3)_2$, $ArCH_2CN$, ArCH=OH, $ArC(=OH)CH_3$, ArC (=OH)OCH_3, $ArC(=OH)N(CH_3)_2$, $ArNH(CH_3)_2$, PyH; Ref. 11; s.d. in kcal mol^{-1}. [e]Substituents: H, CH_3, CF_3, CN, $COCH_3$, $COOCH_3$, $N(CH_3)_2$, NO_2, OCH_3, F, Cl, Br, I; compounds: $ArCH=CH_2$, $ArC(CH_3)=CH_2$, $ArCH=CHC_6H_5$, ArCH=CHCN, $ArCH=CHNO_2$, ArCHO, $ArCOCH_3$, $ArCOOCH_3$, $ArCOOC_2H_5$, ArCN; data of Ref. 16 extended by the series $ArCOOCH_3$ and $ArCH=CHNO_2$ from Refs. 24, 34, s.d. in ppm. [f]Substituents: H, CH_3, *t*-C_4H_9, C_6H_5, CH_2Cl, CH_2Br, CF_3, CN, CHO, $COCH_3$, COC_6H_5, $COOCH_3$, NH_2, $N(CH_3)_2$, $NHCOCH_3$, NO_2, OH, OCH_3, $OCOCH_3$, F, Cl, Br, I; compounds and vibrations: $ArC(=O)CH_3$, ArC(=O)OH, ArCOO–H, Ar(C=O)OH dimer, $ArCH_2COO$–H, $ArCOO^-$ as and s, ArC=N, $ArNH_2$ as and s, $ArNH_2$ first overtone, as and s, ArN–$HCOCH_3$, $ArNO_2$ as, s and t, ArO–H, ArO–H···dioxane; data collected from Refs. 25–27, 35–38; s.d. in cm^{-1}. [g]Substituents: H, CH_3, *t*-C_4H_9, CF_3, CN, CHO, $COCH_3$ $COOCH_3$, COCl, $N(CH_3)_2$, NO_2, OCH_3, $OCOCH_3$, SO_2CH_3, F, Cl, Br, I; compounds $ArCH_3$, $ArCH_3$, ArCN, ArCHO, $ArCOCH_3$, $ArCOOCH_3$, $ArN(CH_3)_2$, $ArNO_2$, $ArOCH_3$, ArCl, ArBr, ArI; data collected from handbooks; s.d. in K.

In practice, the use of eqns. (2) or (3) is complicated by the possibility of choice among several σ_R constants (σ_R°, σ_R^+, σ_R^-). In terms of Table 1 this means that the actual number of parameters would be greater with some coefficients always equal to zero. This means that Table 1 refers to eqns. (2) and (3) analysed in their simplest forms, i.e. when they are applied to rather homogeneous data (particularly those involving conjugation of the same type).

(4) *Additional models.* The *m*|*p* matrix is also used in various forms of correlations with two terms[7-9] or three terms,[10-13] eqns. (3) and (4), respectively. The individual treatments differ mainly in the choice or estimation of substituent constants. Still closer relation (*viz.* direct proportionality) between *para* and *meta* derivatives follows from the treatment of Exner,[15] eqn. (6), restricted to acceptor substituents (f is a generally valid constant, say f_m = 1, f_p = 1.14). On the other hand, a four-term eqn.[14] (5) returns to the original Hammett concept[1] of independence. In two papers,[16,17] separate treatment of *meta* and *para* derivatives was explicitly advocated, in either case on the basis of PCA. In the special case of ^{13}C NMR shifts a one-term equation was recommended for *meta* derivatives and two terms for *para*,[16] eqn. (7), while for a broader set of mixed data two-term equations were preferred[17] for both, eqn. (8). In these two the parameters for the *meta* and *para* positions were not very different, as shown either from a comparison of components[16] or from PLS.[17]

Data. We will now compare the predictions of Table 1 with the results on five reaction series, each of them chosen to be as homogeneous as possible (Table 2). The pK values of benzoic acids in different solvents[23] represent the classical scope of the Hammett equation, while ΔG° of various acids in the gas phase[11] are an example of the recent, most extended form of correlation equation. The additional two series, ^{13}C NMR shifts and IR frequencies, concern properties in which some irregularities have already been observed,[16,24-27] particularly for *meta* derivatives. The last series of melting points was included just as an example of the quantity generally believed to show no simple dependence on structure.[2]

Calculations and results

The data matrices (Table 2) were constructed separately for *meta* and *para* derivatives, always with the same substituents and the same reaction series, only the filling (amount of missing data) was somewhat different. PCA and PLS were carried out according to standard procedures.[20,21] The data were not scaled but centred, the location parameter for each series corresponding to the parameter y° in eqns. (1)–(8). It follows that the standard deviations are given in the original units (pK units, kcal mol^{-1}, ppm, cm^{-1}, K, see Table 2) which can thus be compared immediately with the experimental error. However, when the calculations were repeated with scaling to unit variance, there was little effect on the final conclusions. The number of components was chosen according to the cross-validation procedure and is given in Table 1, but standard deviations are given for each number of components (Table 2).* PCA was carried out separately in the *meta* and *para* matrix, then in the combined matrices $\frac{m}{p}$ (each substituent taken twice as *m*-X and *p*-X) and *m*|*p* (each series taken twice, for *meta* and *para* derivatives). In some series calculations were repeated with the data matrix restricted only to acceptor substituents.

Discussion

In the following discussion attention is focused on the main problem of whether it is reasonable to combine *meta* and *para* derivatives in one series. Discussion of the values of components and loadings will be limited, particularly concerning their relation to established scales. It was not our intention to suggest the values of components as some new scales of substituent parameters: for such a purpose our database would not be sufficient.

The simplest results were obtained in the first series, pKs of substituted benzoic acids, differing only in the solvent. This set represents the narrowest possible range of validity of the Hammett equation and the restriction to pure solvents eliminates the complications in mixed solvents recently described.[28] From Table 1 it is evident that two procedures are of the same value: separate processing of *meta* and *para* derivatives and an $\frac{m}{p}$ matrix with two independent, unrelated substituent constants (σ_m and σ_p) for each substituent. The latter corresponds to the classical Hammett equation which is thus an economical description of the structural dependence in this simple case, with the standard deviation only moderately exceeding the experimental error. The independence of substituent constants is shown very clearly on the results of PLS: three components would be necessary, the first one describing only 90 % of the sum of squares, much less than in PCA. Note that the PLS results are incompatible even with the DSP model: the transmitted information is too small and the number of terms too high. All of this favours the original Hammett conception[1] of independent σ_m and σ_p. A claimed proportionality between them[3] was based on an inadmissible extension of validity range and was later restricted[15] to substituents without a lone electron pair in the α-position, loosely claimed to be acceptors. It is then expressed[15] by eqn. (6) or in terms of σ constants by the simple eqn. (9). The validity of these eqns. can be confirmed by a separate treatment of acceptor substituents (Table 1, lower part). Compared with the unrestricted set the overall fit is somewhat improved and the much closer similarity of *meta* and *para* derivatives is revealed by PLS results. According to our more recent data[23] the constant (λ) in eqn. (9) would be lower than 1.14 as determined earlier.[29] In particular, it is

* The values of components and loadings are not given but they are available from the authors on request.

evidently lower in dimethylformamide and in sulfolane than in a water. Values of components have already been discussed[30] and compared to the common σ_m and σ_p values. There are some small differences due to changes of solvent.

$$\sigma_p = 1.14\ \sigma_m \qquad (9)$$

Dissociation in the gas phase may be a more complex phenomenon since it is also controlled by the substituent polarizability (spreading of the charge on a larger volume[11,31]). In addition, there are large differences in the functional group within this set which includes neutral and cationic acids, and substituted pyridines in one series. Correlation with a three-term equation was tried[11] with polar, resonance, and polarizability parameters derived from other sources. The data matrix was of the *m|p* type but quite often some terms appeared to be insignificant.[11] Our analysis actually gave three terms for this type of matrix but the results of PLS were far from 100 % with three terms (Table 1). The treatment of *meta* and *para* derivatives separately yielded the best results. The behaviour of *meta* derivatives is simpler and more regular and a one-term eqn. is satisfactory. It has already been shown[16,24] that neglecting the fundamental difference between *meta* and *para* derivatives is the main flaw of the DSP treatment: the *meta* series is simply overparametrized without improving the fit. This reasoning is still more important for the three-term extension of DSP as suggested[11] for the present series. For the theory of gas-phase acidities our results could mean that the polarizability effects – proved convincingly on simpler molecules[31] – may have been overestimated in the case of small substituents on a relatively large benzene nucleus. Restriction to acceptor substituents does not improve the picture but note that the restricted data matrices are not well filled in this case.

The ^{13}C NMR shifts represent a very homogeneous data set with a property quite different from the equilibrium constants, although the structure of the compounds is of the classical Hammett type. A slightly smaller set has previously been analysed[16] by PCA; the present results – completed by PLS – essentially confirm the previous ones. As in the preceding example the *meta* derivatives exhibit a simpler pattern: they can be well described by one parameter and are evidently overparametrized in the common DSP treatment. This is confirmed by PLS, showing better prediction in the direction *para* → *meta* than in *meta* → *para*. Restriction to acceptor substituents brings some simplification: PLS demonstrating particularly the close similarity of *meta* and *para* effects. The values of the components have been correlated[16] with the existing σ scales: there are small but significant deviations.

IR frequencies have sometimes been considered as a particularly suitable physical quantity to be correlated with reactivities, i.e. with constants σ. However, the correlations found in practice have always been of relatively low precision. One important anomaly has been named the *meta*-effect;[25,27] its physical nature has not yet been elucidated. Reasons for further deviations have recently been explained in an excellent review.[32] Our data matrix (Table 2) contains frequencies different in character, some of which evidently do not depend at all on substituent polarity,[27] some others are affected by the mentioned secondary effects.[25,32] In addition the data matrix is of poorer quality than the other examples (more lacking items, with an irregular distribution). As expected, the results in Table 1 reveal a rather poor fit: they described only the part of spectral shifts that is dependent on substituent polar effects. Any improvement of the correlation equation by adding further terms does not help. The results of PLS show equal but small amounts of information transferred in either direction, *meta* → *para* or *para* → *meta*. Our main problem is that the IR data are not very informative: when the correlation is poor, the simpler form of the equation is always preferred and a distinction is not possible. Even in this case, however, restriction to acceptor substituents means a significant improvement, particularly the PLS results. If the less clear results of this example are due to the aforementioned defects of the data matrix, one could improve them by dropping columns of low modelling power. After this reduction, more components are needed for the *para* matrix, while the *meta* matrix is not changed. The results thus become increasingly similar to those of the preceding examples but their statistical weight is much less after data reduction. This applies also to PCA on a smaller matrix restricted only to the carbonyl stretching frequency.[20]

The set of melting points was included in our test samples just for comparison and any significant correlation was not expected. In fact the results were better than anticipated and formally rather similar to the IR series. The systematic part of the variance cannot be ascribed to substituent polar effects but rather to some specific effects of steric origin. The melting points of *meta* derivatives are, as a rule, lower, as follows from their lower symmetry. Interestingly enough, some asymmetry of the model is observed even in this series: according to PLS the behaviour of *para* derivatives is more complex (in spite of their higher symmetry). We believe that some relationships between structure and property could be found even for the melting points.

Conclusions

Our results are more definite and quantitatively different from those of a previous regression analysis[3] which was also restricted only to the original Hammett eqn. It turned out that the arrangement of the data matrix is of the same importance as the number of terms when different correlation equation are compared. Very often the treatment described in the literature is overparametrized, while a simple separation of *para* and *meta* derivatives would produce a more significant improvement. Most frequently, if not always, all the deviations and complex behaviour are caused by donor substituents in the *para* position: the defect of common correlation equations is in that they also extend

the complex treatment to acceptors and to the *meta* position. Introduction of statistically unnecessary terms may also be caused by the intention of using 'theoretically pure' parameters.[10] General statistical problems of multiparameter equations have been discussed in detail.[2,33]

Of the two historically most important treatments, those of Yukawa–Tsuno[3] and Taft,[6] each was right in one point: the Yukawa–Tsuno treatment in that it treats *meta* and *para* derivatives on different levels of complexity, the DSP in that it treats these two series separately. We are of the opinion that these derivatives should be processed separately whenever possible and the number of nesessary terms should be estimated in each series independently. The classical Hammett equation, treating these derivatives together, is strictly valid only in the simplest cases, i.e. in the most similar 'similarity models'. In practice it can be also used when a more exact treatment is precluded by the small number of data.

References

1. Hammett, L. P. *Physicial organic Chemistry,* McGraw-Hill, New York 1940.
2. Exner, O. *Correlation Analysis of Chemical Data,* Plenum, New York 1988.
3. Hine, J. *J. Am. Chem. Soc. 81* (1959) 1126.
4. Jaffé, H. H. *J. Am. Chem. Soc. 81* (1959) 3020.
5. Yukawa, Y. and Tsuno, Y. *Bull. Chem. Soc. Jpn. 32* (1959) 971.
6. Ehrenson, S. R., Brownlee, T. C. and Taft, R. W. *Prog. Phys. Org. Chem. 10* (1973) 1.
7. Swain, C. G. and Lupton, E. C. *J. Am. Chem. Soc. 90* (1968) 4328.
8. Swain, C. G., Unger, S. H., Rosenquist, N. R. and Swain, M. S. *J. Am. Chem. Soc. 105* (1983) 492.
9. Tsuno, Y., Kusuyama, Y., Sawada, M., Fujii, T. and Yukawa, Y. *Bull. Chem. Soc. Jpn. 48* (1975) 3337.
10. Charton, M. *Prog. Phys. Org. Chem. 16* (1987) 287.
11. Taft, R. W. and Topsom, R. D. *Prog. Phys. Org. Chem. 16* (1987) 1.
12. Weeks, G. and Horák, V. *J. Org. Chem. 45* (1980) 2068.
13. Niewdorp, G. H. E., de Ligny, C. L. and van Houwelingen, H. C. *J. Chem. Soc., Perkin Trans. 2* (1979) 537.
14. Haldna, Ü. L., Juga, R. J., Tuulmets, A. V. and Jüriado, T. J. *J. Chem. Soc., Perkin Trans. 2* (1987) 1559.
15. Exner, O. *Collect. Czech. Chem. Commun. 31* (1966) 65.
16. Exner, O. and Buděšínský, M. *Magn. Reson. Chem. 27* (1989) 27.
17. Alunni, S., Clementi, S., Ebert, C., Linda, P., Musumarra, G., Edlund, U., Sjöström, M. and Wold, S. *Gazz. Chim. Ital. 116* (1986) 679.
18. Alunni, S., Clementi, S., Edlund, U., Johnels, D., Hellberg, S., Sjöström, M. and Wold, S. *Acta Chem. Scand., Ser. B 37* (1983) 47.
19. Hnyk, D. *Collect. Czech. Chem. Commun. 55* (1990) 55.
20. Fadhil, G. F. *Collect. Czech. Chem. Commun. 57* (1992). *In press.*
21. Wold, S. and Sjöström, M. In: Kowalski, B. R., Ed. 'Chemometrics: Theory and Applicatio,' *Am. Chem. Soc. Symp. Ser. 52* (1977) 243.
22. Wold, S., Albano, C., Dunn, W. J., Edlund, U., Esbensen, K., Geladi, P., Hellberg, S., Johansson, E., Lindberg, W. and Sjöström, M. In: Kowalski, B. R., Ed., *Chemometrics,* NATO ASI Series, Reidel, Boston 1984.
23. Ludwig, M., Baron, V., Kalfus, K., Pytela, O. and Večeřa, M. *Collect. Czech. Chem. Commun. 51* (1986) 2135.
24. Buděšínský, M. and Exner, O. *Magn. Reson. Chem. 27* (1989) 585.
25. Exner, O. and Boček, K. *Collect. Czech. Chem. Commun. 38* (1973) 50.
26. Krueger, P. J. *Can. J. Chem. 41* (1963) 363.
27. Exner, O., Kováč, S. and Solčáiová, E. *Collect. Czech. Chem. Commun. 37* (1972) 2156.
28. Hoefnagel, A. J. and Wepster, B. M. *J. Chem. Soc., Perkin Trans. 2* (1989) 977.
29. Exner, O. and Kalfus, K. *Collect. Czech. Chem. Commun. 41* (1976) 569.
30. Pytela, O., Ludwig, M. and Večeřa, M. *Collect. Czech. Chem. Commun. 51* (1986) 2143.
31. Taft, R. W., Taagepera, M., Abboud, J. L. M., Wolf, J. F., DeFrees, D. J., Hehre, W. J., Bartmess, J. E. and McIver, R. T., Jr. *J. Am. Chem. Soc. 100* (1978) 7765.
32. Laurence, C. In: Zalewski, R. I., Krygowski, T. M. and Shorter, J., Eds., *Similarity Models in Organic Chemistry, Biochemistry and Related Fields,* Plenum Press, New York 1991.
33. Shorter, J. *Correlation Analysis of Organic Reactivity with Particular Reference to Multiple Regression,* Research Studies Press, Wiley, Chichester 1982.
34. Happer, D. A. R. and Steenson, B. E. *J. Chem. Soc., Perkin Trans. 2* (1989) 19.
35. Laurence, C. and Berthelot, M. *J. Chem. Soc., Perkin Trans. 2* (1979) 98.
36. Laurence, C. and Berthelot, M. *Spectrochim. Acta, Part 34* (1978) 1127.
37. Berthelot, M. and Laurence, C. *C. R. Acad. Sci., Ser. C 276* (1973) 979.
38. Dunn, G. E. and McDonald, R. S. *Can. J. Chem. 47* (1969) 4577.

Received September 23, 1991.