Talanta, Vol. 33, No. 5, pp. 435–441, 1986
Printed in Great Britain. All rights reserved
0039-9140/86 $3.00 + 0.00
Copyright © 1986 Pergamon Press Ltd

MULTIPARAMETRIC CURVE FITTING—IX*

SIMULTANEOUS REGRESSION ESTIMATION OF STOICHIOMETRY AND STABILITY CONSTANTS OF COMPLEXES

JOSEF HAVEL
Department of Analytical Chemistry, College of Science, J. E. Purkyně University, 611 37 Brno, Czechoslovakia

MILAN MELOUN†
Department of Analytical Chemistry, College of Chemical Technology, 532 10 Pardubice, Czechoslovakia

(*Received* 22 *January* 1985. *Accepted* 30 *September* 1985)

Summary—A chemical model (*i.e.*, the number of complexes, their stoichiometry and stability constants with molar absorptivities) in solution equilibria may be established by (*i*) the trial-and-error method in which stability constants are estimated for an assumed set of complexes in the mixture and a fitness test is used to resolve a choice of plausible models to find the true one; (*ii*) the simultaneous estimation of the stoichiometry and stability constants for species divided into "certain" species for which the parameters β_{pqr}, (p, q, r) are known and held constant, and "uncertain" species with unknown parameters which are determined by regression analysis. The interdependence of stability constants and particular sets of stoichiometric indices requires that the computational strategy should be chosen carefully for each particular case. The benefits and limitations of both approaches are compared by means of examples of potentiometric titration data analysis by the POLET(84) program and of spectrophotometric data analysis by the SQUAD(84) program. A strategy for efficient computation is suggested.

Potentiometry and spectrophotometry are often used to study solution equilibria. The chemical model of an equilibrium system consists of (*i*) the number of species in equilibrium, (*ii*) their stoichiometry, and (*iii*) their stability constants or molar absorptivities, respectively. The program MESAK[1] determines the average composition of species from potentiometric data. The program PRCEK[2,3] evaluates the stoichiometric indices p, q, r of an $M_pL_qH_r$ complex from the slopes of linearly transformed relationships, by linear regression. Under simplifying conditions PRCEK may be used to analyse absorbance–pH or absorbance–concentration data for equilibrium systems of even polynuclear complexes[4] but a preliminary data-treatment of the absorbance–concentration curves should be made,[5] as the mathematical model used in PRCEK is limited to two complexes only.[6,7] The program MRLET[8] algorithmically determines the stoichiometric coefficient q of a complex ML_q by regression analysis of an absorbance *vs.* mole–ratio curve, provided this is the predominant complex in the equilibrium mixture.

Other programs use the operator-controlled "trial-and-error" (heuristic) method of stoichiometry determination[9–12] in which the stoichiometric coefficients are part of the hypothesis tested, and are either confirmed by regression analysis or are rejected and replaced by those for a new model. The process is continued until a satisfactory fit of the model and regression analysis is obtained. The initial choice of species stoichiometry is usually based on experience, but automatic selection from a list of possible species and subsequent hypothesis testing has been developed, *e.g.*, the STYRE species-selector introduced by Sillén in LETAGROP[11,12] and the species selector in PSEQUAD,[13] which tests several species hypotheses step by step so that the species from a list are included in or excluded from a model quite automatically.

Even with computer assistance, it is rather difficult to test all possible combinations of stoichiometric coefficients. For example, if p, q and r for a ternary complex $M_pL_qH_r$ are restricted to the ranges 0–3, 0–3, and 0–5, respectively, then, excluding proton–metal species and cases where two or all three coefficients are zero, 68 different species can be obtained by combination of the components M, L and H, 2278 combinations of two different species, and so on. Use of chemical experience can radically lower the number of possible combinations to be tested but it is still difficult to test them all. Usually, only a few tens of combinations are tested,[14] though Varga *et al.*[15] investigated up to 128 different models in a solvent extraction experiment. The problems involved have been discussed in a review.[16]

The present paper describes a direct algorithmic estimation of stoichiometric indices (the ESI approach) based on simultaneous regression deter-

*Part VIII: *Talanta*, 1985, **32**, 973.
†Author for correspondence.

mination of stoichiometric indices and stability constants. The novelty of the proposed approach is that the stoichiometric indices are estimated as *real numbers* simultaneously with the stability constants. The method has been applied to potentiometric data analysis by the program POLET(84)[17] and to spectrophotometric data analysis by the program SQUAD(84),[18] and some preliminary results have been reported.[19]

THEORY

An equilibrium system consisting of metal (M), ligand (L) and proton(s) (H) as basic components combined in a species of general formula $M_pL_qH_r$ may be described by the mass-balance equations (charges omitted for simplicity):

$$c_M = [M] + p \sum_{j=1}^{n_c} (\beta_{pqr}[M]^p[L]^q[H]^r)_j \quad (1)$$

$$c_L = [L] + q \sum_{j=1}^{n_c} (\beta_{pqr}[M]^p[L]^q[H]^r)_j \quad (2)$$

$$c_H = [H] + r \sum_{j=1}^{n_c} (\beta_{pqr}[M]^p[L]^q[H]^r)_j \quad (3)$$

and the overall stability constant

$$\beta_{pqr} = [M_pL_qH_r]/([M]^p[L]^q[H]^r) \quad (4)$$

Solution equilibria are usually studied by potentiometry, spectrophotometry, extraction, *etc.*, the total analytical concentrations of the basic components c_M, c_L and c_H being known and experimentally adjustable, and representing the independent variable in the regression analysis. The dependent variable y_{exp} is represented in potentiometry by the free concentration of a basic component, indicated by an ion-selective electrode (*e.g.*, pH by a glass electrode), and in spectrophotometry by the absorbance. The functional relationship between $\mathbf{y}_{exp}$, an experimentally adjustable vector of the independent variable, $\mathbf{x}_{exp} = \{c_M, c_L, c_H, \text{pH}, etc.\}$ and a vector of the unknown parameters $\boldsymbol{\beta} = \{\beta_{pqr,j}, \epsilon_{pqr,j}, (p, q, r)_j\}$, $j = 1, \ldots, n_c$, is a mathematical model formulated in the form $\mathbf{y} = f(\mathbf{x}; \boldsymbol{\beta})$. Here $\beta_{pqr,j}$ and $\epsilon_{pqr,j}$ represent the stability constant and molar absorptivity of the jth species with integers p, q, r as stoichiometric coefficients. The equilibrium mixture contains n_c species.

The general problem is to find estimates for the unknown parameters which will minimize the residual-square sum function U:

$$U = \sum_{i=1}^{n} w_i[y_{exp,i} - f(x_i; \beta)]^2 \approx \text{minimum} \quad (5)$$

where w_i is a statistical weight, usually taken as unity.

The minimization may be done algorithmically or heuristically. The algorithmic process (*e.g.*, SQUAD[18]) usually finds a global minimum whereas the heuristic process depends more on human control. POLET(84)[17] offers a choice of algorithmic or heuristic processing. The user must decide whether a local or global minimum is required. In computational strategy, restrictions and initial guesses for the parameters and minimization steps for particular parameters should be supplied, and special care paid to parameters that are interdependent in the model.

Some difficulty in a non-linear estimation of β_{pqr} and p, q, r may arise when there is an adverse mutual influence of the estimates of one of the stoichiometric coefficients and of β_{pqr}, *i.e.*, if the one parameter is incorrectly estimated, the other is also incorrectly estimated (biased), but the combined effect of the incorrectly estimated parameters, when introduced into the model, may yield quite reasonable predictions. It is worthwhile calculating the correlation coefficients to obtain information about the interactions of parameters.

The ESI approach is based on treating the stoichiometric indices as real numbers instead of integers and varying them simultaneously with the stability constants in the regression process. Though non-integer coefficients of chemical reaction have no physical meaning, treating the coefficients as real numbers leads to estimates which are close to integers for a true chemical model.

The following approaches may be used to estimate the separable parameters: (*i*) the ESI method, in which the model species are divided into the "certain" species, for which β_{pqr} and p, q, r are held constant, and the "uncertain" species, for which these parameters are directly estimated by regression; (*ii*) the trial-and-error method, in which p, q and r for all species are fixed integers based on a hypothetical chemical model, and only the stability constants β_{pqr} are estimated by regression.

Other possibilities are based on a combination of the two methods in various computational steps, each of which refines only some of the parameters; for example, known stability constants and stoichiometric indices are kept unchanged during a partial regression refinement of other parameters, and in the last step all the parameters, stability constants and stoichiometric indices are refined to confirm the chemical model found. Simultaneous estimation of the stability constant and of all stoichiometric indices of a particular species seems to be a rare case because the interdependence of parameters, indices and stability constant in equations (1)–(4) does not enable an estimation.

Which computational strategy will prove optimal depends on the number of complexes, previous knowledge of some species in the chemical model, and the experimental design for changing the basic components in the equilibrium system, and therefore an *ad hoc* choice is necessary. The experimental strategies possible for estimation of one stoichiometric coefficient by varying the total concentration of the corresponding basic component are surveyed in Table 1.

Table 1. A survey of experimental possibilities of simultaneous estimation of stability constants and stoichiometric indices of $M_pL_qH_r$ species from various spectrophotometric data according to change of the basic components in the experiment: β'_{pqr} denotes the conditional stability constant

Equilibrium	Dependence	Condition	Estimated parameters: Suitable	Estimated parameters: Not suitable
(1) Protonation of monomer: $L + rH = LH_r$	$A = f(\mathrm{pH})$	$c_L = \text{const.}$	β_{01r}, r	—
(2) Protonation of oligomer: $qL + rH = L_qH_r$	$A = f(\mathrm{pH})$	$c_L = \text{const.}$	β'_{0qr}, r	β_{0qr}, q
		$c_L \neq \text{const.}$	β_{0qr}, q, r	—
(3) Complex-forming equilibria: $pM + qL + rH = M_pL_qH_r$				
(i) Mole-ratio plots	$A = f(c_M/c_L)$ or $A = f(c_L/c_M)$	pH const. $c_L \neq \text{const.}$ $c_M \neq \text{const.}$	β'_{pqr}, p, q	r
(ii) A–pH plots for an excess of c_L:	$A = f(\mathrm{pH})$	$c_L \gg c_M$ $c_L = \text{const.}$ $c_M = \text{const.}$	β'_{pqr}, p, r	β_{pqr}, q
(iii) A–pH plots for an excess of c_M:	$A = f(\mathrm{pH})$	$c_L \ll c_M$ $c_L = \text{const.}$ $c_M = \text{const.}$	β'_{pqr}, q, r	β_{pqr}, p
(iv) A–pH plots for varied c_M and c_L:	$A = f(\mathrm{pH})$	$c_L \neq \text{const.}$ $c_M \neq \text{const.}$	β_{pqr}, p, q, r	—

EXPERIMENTAL

Computational strategy

Determination of number of species. A set of potentiometric titration curves is transformed into $\{Z, \mathrm{pH}\}$ normalized co-ordinates and the resulting **Z** matrix is analysed by the factor-analysis algorithm SPECIES.[20] For various integers of matrix rank k, the standard deviation of **Z**, $s_k(\mathbf{Z})$ is calculated and the intersection of two linear parts of the $s_k(\mathbf{Z}) = f(k)$ graph indicates the matrix rank k^* which is equal to the number of species in the equilibrium.

The factor-analysis algorithm FA608[21] is analogously used to analyse the absorbance matrix and the rank is estimated as above. The intersection of two linear parts of the $s_k(\mathbf{A}) = f(k)$ graph indicates the matrix rank k^* equal to the number of light-absorbing species, n_c, in the equilibrium mixture.

Stability constants of "certain" species. The graph of n_c complexes is divided into n_{cert} "certain" species, and n_{uncert} "uncertain" species. The stoichiometry of the "certain" species is known either from previous experimentation or the literature. The stability constants $\beta_{pqr,j}, j = 1, \ldots, n_{\mathrm{cert}}$ (or corresponding molar absorptivities in spectrophotometric data analysis), are estimated from separate experimental data by the trial-and-error method. A goodness-of-fit test[22] establishes the reliability of the parametric estimates found.

Stoichiometry and stability constant of "uncertain" species. Realistic values for the stability constants and stoichiometric coefficients are supplied as initial guesses of the unknown parameters $[(\beta^{(0)}_{pqr}, p^{(0)}, q^{(0)}, r^{(0)})_j; j = 1, \ldots, n_{\mathrm{uncert}}]$, and the regression algorithm is applied. Chemical experience, tables of stability constants and knowledge of co-ordination chemistry should assist in making the initial guesses and choosing the computational strategy for simultaneous estimation of the stoichiometry and stability constants. This estimation sometimes fails if there is a strong interdependence of the indices and the stability constants, resulting in divergence of the minimization algorithm. Care should be taken to choose a minimization routine that will also work for interdependent parameters. The fitness test[22] is again used as a criterion for the reliability of the regression estimation.

Final confirmation of chemical model found. The real-number values of the stoichiometric indices of "uncertain" species are rounded to integers and kept constant in a final refinement of the stability constants (or molar absorptivities) of all the "certain" and "uncertain" species by the trial-and-error method. Before the final refinement, the real-number estimates of the stoichiometric indices should be as close as possible to integers and have low standard deviations. $U_{\min}$ should not change much in the final refinement from the value found with the real-number indices.

Computation

The estimation of stoichiometric indices (the ESI method) has been introduced into the programs POLET[5,10] and SQUAD[24] to give two regression programs: POLET(84) for potentiometric data analysis[17] and SQUAD(84) for spectrophotometric data analysis.[18] Our computations were done with the EC 1033 (500 K) computer in the Department of Computing Technique, J. E. Purkyně University, 611 37 Brno and the Computing Centre, College of Chemical Technology, 532 10 Pardubice Czechoslovakia.

DISCUSSION

The ESI method has been applied to potentiometric and spectrophotometric data analysis, and some examples, benefits and limitations of the method are discussed below.

The first example is concerned with the overlapping protonation equilibrium of a weak acid;[10] the simulated data in normalized co-ordinates $\{Z, \mathrm{pH}\}$ were calculated for the two stability constants, $\log \beta_{011} = 3.50$ and $\log \beta_{012} = 8.00$ and also loaded by random errors calculated on the basis of an arbitrary value of $s_{\mathrm{inst}}(Z) = 0.005$. Application of POLET(84) in two runs with different initial guesses for the stability constants and stoichiometric indices, and of the LETAG minimization algorithm[23] in POLET(84), led to the true values of the parameters, and the goodness-of-fit test showed an excellent fit and hence a reliable parametric estimation (Table 2).

Table 2. Simultaneous estimation of stability constants and stoichiometric indices from potentiometric data analysis by POLET(1984)

Example 1. Determination of two overlapping protonation constants of a weak acid from simulated $Z = f(\mathrm{pH})$ data.[10] Selected values $s_{\mathrm{inst}}(Z) = 0.005$, $\log \beta_{011} = 3.50$ and $\log \beta_{012} = 8.00$

Run		Initial guess		Estimated parameters		Degree of fit test	
		$\log \beta_{01r}^{(0)}$	$r^{(0)}$	$\log \beta_{01r}$	r	$s(Z)$	$R \times 10^2$
1	LH	2.5	1	3.473 ± 0.024	0.993 ± 0.005	0.0060	0.316
	LH_2	7.0	2	7.985 ± 0.020	1.997 ± 0.003		
2	LH	3.5	0	3.502 ± 0.020	1.000 ± 0.005	0.0070	0.368
	LH_2	8.5	0	8.006 ± 0.005	2.000 ± 0.000		

Example 2. Determination of Bi(III) hydrolysis constants and species stoichiometry from simulated $Z = f(\mathrm{pH})$ data.[10] Selected values $s_{\mathrm{inst}}(Z) = 0.010$, $\log \beta_{1,0,-1} = -1.58$, $\log \beta_{6,0,-12} = 0.33$

Run	Species	Initial guess			Estimated parameters		
		$\log \beta_{p0r}^{(0)}$	$p^{(0)}$	$r^{(0)}$	$\log \beta_{p0r}$	p	r
1	Bi(OH)	−2.0	1	−2	-1.562 ± 0.002	1	-0.988 ± 0.001
2	$Bi_6(OH)_{12}$	0.33	6	−10	0.319 ± 0.018	6	-11.999 ± 0.013
3	$Bi_6(OH)_{12}$	0.33	6	−8	0.320 ± 0.002	6	-12.001 ± 0.000
4	$Bi_6(OH)_{12}$	−1.5	6	−8	0.318 ± 0.016	6	-11.999 ± 0.013
5	$Bi_6(OH)_{12}$	−2.0	2	−3	0.392 ± 0.017	6.116 ± 0.012	-12.225 ± 0.026
6	$Bi_6(OH)_{12}$	−2.0	2	−4	0.356 ± 0.044	6.035 ± 0.089	-12.062 ± 0.089

The second example shows that the composition of the complexes might also be determined by the ESI method even when a bad initial guess for the stoichiometric indices is used. For a model consisting of two hydrolytic species, Bi(OH) and $Bi_6(OH)_{12}$, described by the stability constants[10] $\log \beta_{1,0,-1} = -1.58$ and $\log \beta_{6,0,-12} = 0.33$ and an arbitrary value of $s_{\mathrm{inst}}(Z) = 0.010$, the normalized data $\{Z, \mathrm{pH}\}$ were simulated. Six runs of the ESI approach in POLET(84) computations starting with different initial guesses of the stoichiometric indices were performed. In the first run, the parameters $\log \beta_{6,0,-12}^{(0)} = 0.33$ and $q_1^{(0)} = 0$, $p_2^{(0)} = 6$, $q_2^{(0)} = 0$, and $r_2^{(0)} = -12$ were kept constant while $\log \beta_{pqr,1}$, p_1 and r_1 were estimated for the first species. In the following five runs, the parameters $\log \beta_{1,0,-1}^{(0)} = -1.58$ and $p_1^{(0)} = 1$, $q_1^{(0)} = 0$, $r_1^{(0)} = -2$ and $q_2^{(0)} = 0$ were held constant and $\log \beta_{pqr,2}$, p_2 and r_2 for the second species were estimated. Minimization terminated in all runs with the estimated real-number stoichiometric indices quite close to the true integer coefficients, so they could easily be rounded off to integers and the final refinement of the stability constants performed.

The third example involves regression analysis of spectrophotometric data by SQUAD(84). Simulated spectra as described by Leggett and McBryde[24] were calculated for the species LH, LH_2, M(OH), ML, ML_2 and MLH and loaded with a spread of random errors generated for an arbitrary value of the instrumental error, $s_{\mathrm{inst}}(A) = 0.0074$. In the first step, a chemical model was set up with LH, LH_2 and M(OH) as the "certain" species, and ML, ML_2 and MLH as the "uncertain" species. The trial-and-error method was then applied for determination of β_{110}, β_{120} and β_{111} and the corresponding $U_{\min}$ value. The statistical characteristics of the fit achieved serve here as reference values for comparison with those achieved in the subsequent ESI minimization process. As a rather high value of $s_{\mathrm{inst}}(A)$ was chosen in the data simulation, the stability constants found, 9.990, 17.813 and 14.071, should be considered in good agreement with the values used for data simulation, *viz.* 10.00, 17.80, 14.10 (Table 3).

Simultaneous estimation of p, q, r and the corresponding stability constants fails if there is a strong interdependence of the indices. The first three model hypotheses shown for this method in Table 3 differ in the choice of which of p, q, and r is to be varied, the other two being kept constant. This method led to true values for the stability constants and stoichiometry and also degree of fit values the same as those for the reference trial-and-error method. The real-number values of p, q and r are close to integers and may therefore easily be rounded off for refinement of the stability constants.

The last four model hypotheses (Table 3) are all false and demonstrate the ability of the ESI method to distinguish between a true and a false chemical model. For a false model the real-number values found for p, q, r are far from being integers and have no physical meaning, the degree of fit is bad, and the models hypothesized should be rejected.

The protonation equilibrium of 2-(2-thiazolylazo)-4-methoxyphenol (TAMP)[9] is the last example used here for demonstration of the benefits and limitations of simultaneous estimation of stoichiometry and stability constants. Dissociation of r protons from the acid LH_r, which has an overall protonation (stability) constant $\beta_{01r} = [LH_r]/([L][H]^r)$, may be indicated by plots of absorbance *vs.* pH. For n_w wavelengths and

Table 3. Determination of stability constants and stoichiometric indices of three uncertain complexes by simulated absorbance matrix analysis with use of the ESI approach in SQUAD(1984): experimental conditions: absorbance matrix generated[20] for the absorbance error $s_{inst}(A) = 0.007404$ and selected model: $\log \beta_{011} = 9.70$, $\log \beta_{012} = 13.30$, $\log \beta_{10-1} = -3.90$, $\log \beta_{110} = 10.00$, $\log \beta_{120} = 17.80$, $\log \beta_{111} = 14.10$; in the regression analysis, the stability constants, stoichiometry and molar absorptivities of "certain" species [LH, LH_2 and M(OH)] were held constant while those of "uncertain" complexes were estimated

Hypothetic chemical model	Species assumed	Estimated parameters				Fitness test		Hypothesis of model may be
		$\log \beta_{pqr}$	p	q	r	$s(A) \times 10^3$	$R \times 10^2$	
A. The trial-and-error method								
1st	ML	9.990 ± 0.028	1	1	0	7.31	0.995	Accepted
	ML_2	17.813 ± 0.029	1	2	0			
	MLH	14.071 ± 0.018	1	1	1			
B. The method of simultaneous estimation of stoichiometry and stability constants								
2nd	ML	9.727 ± 0.113	0.975 ± 0.015	1	0	7.20	0.974	Accepted
	ML_2	17.389 ± 0.143	0.951 ± 0.017	2	0			
	MLH	14.267 ± 0.086	1.065 ± 0.027	1	1			
3rd	ML	9.654 ± 0.113	1	0.955 ± 0.015	0	7.35	0.994	Accepted
	ML_2	18.254 ± 0.139	1	2.062 ± 0.019	0			
	MLH	14.265 ± 0.073	1	1.014 ± 0.006	1			
4th	ML	9.945 ± 0.106	1	1	0.009 ± 0.015	7.35	0.994	Accepted
	ML_2	17.710 ± 0.194	1	2	0.017 ± 0.30			
	MLH	14.066 ± 0.041	1	1	0.999 ± 0.019			
5th	ML	15.973 ± 0.245	2.254 ± 0.056	1	0	32.85	4.707	Rejected
	ML_2	17.204 ± 0.268	0.948 ± 0.046	2	0			
6th	ML	5.010 ± 0.137	1	0.448 ± 0.014	0	29.75	4.262	Rejected
	ML_2	17.759 ± 0.607	1	2.065 ± 0.079	0			
7th	ML	12.837 ± 0.116	1	1	0.627 ± 0.031	40.72	5.835	Rejected
	ML_2	19.434 ± 0.519	1	2	0.308 ± 0.095			
8th	MLH	12.837 ± 0.116	1	1	0.627 ± 0.031	40.73	5.835	Rejected
	ML_2	19.434 ± 0.519	1	2	0.308 ± 0.095			

n_s solutions differing in pH, the absorbance matrix, of size $(n_w \times n_s)$ represents the input data for SQUAD(84). Separation of the data into individual A–pH curves, and subsequent graphical analysis, may be performed and $\log \beta_{01r}$ and r estimated from the intercept and slope of dependence of $\log (A - \epsilon_{01r} c_L)/(\epsilon_{010} c_L - A) = r\,\mathrm{pH} + \log \beta_{01r}$ (for absorbance measurements in 1-cm cells). However, various linear transformation methods are rather sensitive to the choice of segment of the A–pH curve to be analysed and the parametric estimates are always loaded by systematic error. Therefore the rigorous non-linear regression approach of SQUAD(84)[18] is preferred for examining the possibilities and limitations of the ESI method.

Both computational strategies of SQUAD(84), the classical trial-and-error method and the method of simultaneous stoichiometry estimation were applied and the results are compared in Table 4. The trial-and-error method searches for the best degree of fit for various initial guesses of both $\log \beta_{01r}^{(0)}$ and $r^{(0)}$, with the option of keeping r constant and refining $\log \beta_{01r}$, or keeping $\log \beta_{01r}^{(0)}$ at a realistic guessed value and estimating r. The degree of fit is again here a sensitive resolution tool. For a false initial guess of the stoichiometric indices or for quite unrealistic stability constants, the minimization process fails because of a divergence of the minimization algorithm or because of a poorly developed minimum, and also there is a poor fit.

The shape of the residual-square sum function U for the spectrophotometric data of the TAMP A–pH curve at a single wavelength is interpreted by contours computed by the program MINUIT.[26] It is evident from Fig. 1 that the minimum is a rather skew and very narrow pit which looks like a cleft. Such a minimum is difficult to reach by most minimizing routines when a bad initial guess of parameters is used. Sillén[27] proved that the convergence of a minimization can be speeded up by varying the parameters along the main axis of the skew pit. Therefore POLET(84), which uses this twisting matrix technique, converges better than SQUAD(84), even from bad initial guesses.

Simultaneous estimation of stoichiometry and stability constants is also handicapped if there is a strong interdependence of $\log \beta_{01r}$ and r. The total correlation coefficient for the two parameters $\rho(\log \beta_{01r}, r) = 0.998$ indicates a strong correlation between them and their interdependence may be expressed by the equation $(\log \beta_{01r})_{found} = (r)_{given} \times \log \beta_{01r}$. The U response surface may be visualized by means of a three-dimensional graph of U as a function of both parameters, $(10 - U) = f(\log \beta_{01r}; r)$, and the minimum U_{min} is interpreted as the maximum value of $(10 - U_{min})$. As SQUAD(84) is not able to

Table 4. Simultaneous estimation of protonation constant and one stoichiometric index r of LH_r species by SQUAD(1984) analysis of absorbance matrix[10] of TAMP: experimental conditions: $5.0 \times 10^{-5} M$ TAMP, $I = 0.1$ (KNO_3), 1-cm cell, 25°C

Part 1. Absorbance matrix of dimensions $n_w = 7$, $n_s = 15$

pH	Absorbance at 440 nm	470 nm	510 nm	540 nm	560 nm	580 nm	610 nm
6.05	0.356	0.373	0.230	0.094	0.052	0.035	0.028
6.42	0.354	0.370	0.240	0.106	0.067	0.050	0.035
6.69	0.343	0.368	0.250	0.128	0.091	0.073	0.051
6.91	0.335	0.357	0.257	0.155	0.124	0.102	0.071
7.11	0.325	0.348	0.270	0.188	0.160	0.135	0.095
7.35	0.302	0.333	0.291	0.241	0.223	0.194	0.132
7.57	0.281	0.315	0.312	0.296	0.278	0.245	0.171
7.87	0.245	0.286	0.350	0.391	0.395	0.355	0.237
8.11	0.215	0.263	0.382	0.465	0.484	0.438	0.290
8.29	0.184	0.248	0.410	0.529	0.557	0.502	0.328
8.53	0.170	0.233	0.436	0.583	0.615	0.556	0.362
8.85	0.154	0.221	0.456	0.630	0.669	0.605	0.393
9.28	0.144	0.218	0.470	0.660	0.710	0.635	0.410
9.90	0.136	0.212	0.480	0.680	0.730	0.660	0.415
11.63	0.134	0.210	0.489	0.700	0.744	0.670	0.423

Part 2. Results of the trial-and-error method and the simultaneous estimation of stability constants with stoichiometric index r: XXX means that too big a number was estimated

Parameters varied	Initial guess $\log \beta_{01r}^{(0)}$	$r^{(0)}$	Estimated parameters $\log \beta_{01r}$	r	Degree of fit test $U \times 10^2$	$s(A)$	$R \times 10^2$	Hypothesis testing
A. The trial-and-error method								
β_{01r} only	7.846	1	7.846 ± 0.005	—	0.2288	0.0050	1.18	Accepted
		2	15.790 ± 0.044	—	9.1319	0.0318	5.56	Rejected
		3	23.741 ± 0.074	—	17.9159	0.0446	10.89	Rejected
r only	1.95	1	—	0.2943 ± 0.0068	0.0426	0.0426	10.02	Rejected
	3.90	1	—	0.5136 ± 0.0035	6.3287	0.0045	3.85	Rejected
	7.846	1	—	0.9999 ± 0.0007	0.2287	0.0050	1.18	Accepted
B. The method of simultaneous estimation of stoichiometry and stability constants								
β_{01r}, r	7.846	1	7.464 ± 0.082	0.9520 ± 0.0103	0.1834	0.0045	1.06	Accepted
	7.85	1	7.461 ± 0.079	0.9516 ± 0.0099	0.1834	0.0045	1.06	Accepted
	7.5	1	$6.500 \pm$ XXX	0.6158 ± 0.2817	126.8170	0.1187	27.77	Rejected
	1.95	1	0.950 ± 6.121	0.2184 ± 0.1247	22.1310	0.0496	11.60	Rejected

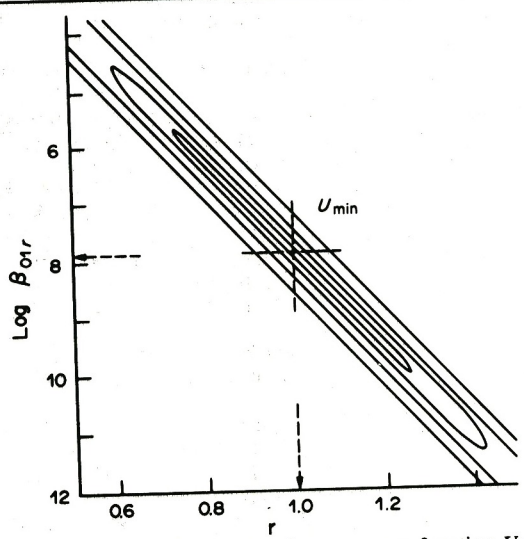

Fig. 1. The contours of residual-squares sum function U on the plot of $\log \beta_{01r}$ *vs.* r demonstrate a skew and narrow pit and the determination of the co-ordinates of U_{min} is rather difficult for many minimization routines. Example of A–pH curve of TAMP from Table 4 at 560 nm, computed by MINUIT program.

map a fine structure and reach a skew cleft-like minimum, the fine structure of the maximum was not calculated and plotted (Fig. 2). The position of the crease for the maximum proves strong covariation of the two parameters, since its main axis is situated at an angle to both co-ordinate axes.

The interdependence of $\log \beta_{01r}$ and r should be kept in mind when choosing a realistic value in the initial guess for both parameters. Either a reasonable initial guess of the stability constant should be used, in which case the pit-mapping algorithm[23] in POLET(84)[17] can minimize both parameters simultaneously with good convergence, or one parameter is kept constant at the value of the initial guess and only the other is refined.

The ESI method can be used to confirm an experimentally determined chemical model. When the stoichiometric indices are varied together with the corresponding stability constants, the estimates found for the stoichiometric indices should *not* differ significantly from the tested integer values. When the chemical model is false, the estimated stoichiometric indices are far from being integers and can sometimes

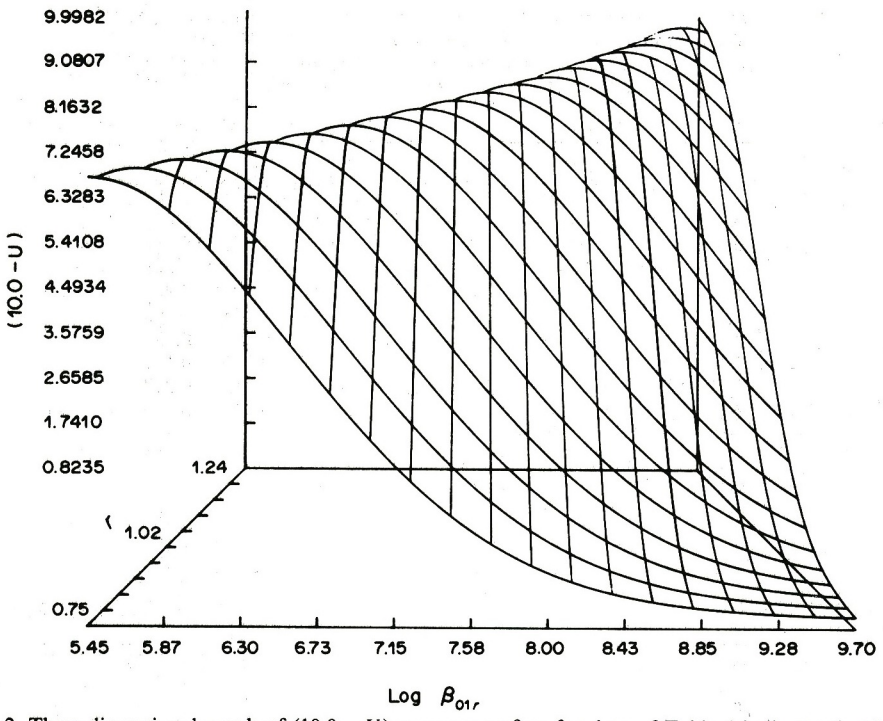

Fig. 2. Three-dimensional graph of $(10.0 - U)$ response-surface for data of Table 4 indicates that both parameters, $\log \beta_{01r}$ and r, are interdependent in a model.

reach quite unrealistic values. When the chemical model tested is incomplete with some species missing, the degree of fit test reliably indicates a very bad fit is achieved.

CONCLUSIONS

It has been demonstrated that stoichiometric indices can be estimated as real numbers simultaneously with stability constants. The proposed method (ESI) can be applied (*i*) for the direct determination of the stoichiometry of "uncertain" species as one step in finding the correct chemical model, and (*ii*) for final confirmation of the species composition. The ESI method is also useful in searching for a chemical model and as a diagnostic tool for judging whether a suggested chemical model is true or false. A reliable minimizing subroutine able to estimate covariant parameters is required. Application of the ESI method is under further development.

REFERENCES

1. L. G. Sillén, *Acta Chem. Scand.*, 1961, **15,** 1981.
2. J. Havel and V. Kubáň, *Scripta Fac. Sci. Nat. UJEP Brun., Chemia 2,* 1971, **1,** 87.
3. V. Kubáň, L. Sommer and J. Havel, *Collection Czech. Chem. Commun.*, 1975, **40,** 604.
4. L. Sommer, V. Kubáň and J. Havel, *Spectrophotometric Studies of Complexation in Solution, Folia Fac. Sci. Nat. Univ. Purk. Brunensis,* Vol. XI, *Chemia* 7, Op. 1 (1970).
5. J. Havel and M. Meloun, in *Computational Methods for the Determination of Formation Constants,* D. J. Leggett (ed.), Plenum Press, New York, 1985.
6. M. Šucmanová-Vondrová, J. Havel and L. Sommer, *Collection Czech. Chem, Commun.*, 1977, **42,** 1812.
7. J. Vošta and J. Havel, *ibid.*, 1977, **42,** 2871.
8. M. Meloun and M. Javůrek, *Talanta*, 1984, **31,** 1083; V. Říha, M. Meloun, M. Franz and J. Čermák, *Sb. Věd. Pr., Vys. Šk. Chemickotechnol., Pardubice*, 1975, **33,** 39.
9. M. Meloun and J. Havel, *Computation of Solution Equilibria,* Part 1, *Spectrophotometry, Folia Fac. Sci. Nat. Univ. Purk. Brunensis,* 1986.
10. M. Meloun and J. Havel, *Computation of Solution Equilibria,* Part 2, *Potentiometry, Folia Fac. Sci. Univ. Purk. Brunensis,* in the press.
11. L. G. Sillén and B. Warnqvist, *Arkiv Kemi*, 1969, **31,** 341.
12. N. Ingri and L. G. Sillén, *Acta Chem. Scand.*, 1962, **16,** 173.
13. I. Nagypál and L. Zékány, in *Computational Methods for the Determination of Formation Constants,* D. J. Leggett (ed.), Plenum Press, New York, 1985.
14. J. Havel, in *Coordination Chemistry in Solution,* E. Högfeldt (ed.), *Kungl. Tek. Hög. Hand., Trans. Royal Inst. Technol.*, 1972, 277, 371.
15. L. P. Varga, W. D. Wakley, L. S. Nicolson, M. L. Madden and J. Patterson, *Anal. Chem.*, 1965, **37,** 1003.
16. A. Vacca, A. Sabatini and M. A. Gristina, *Coord. Chem. Rev.*, 1972, **8,** 45.
17. J. Havel and M. Meloun, *Talanta*, 1986, in press.
18. M. Meloun, M. Javůrek and J. Havel, *ibid.*, 1986, in press.
19. J. Havel and M. Vrchlabský, *Euroanalysis V, Cracow (Poland)*, 26–31 August 1984, Paper XVIII-20, 465.
20. J. Havel and M. Meloun, *Talanta*, 1985, **32,** 171.
21. J. J. Kankare, *Anal. Chem.*, 1970, **42,** 1322.
22. A. Sabatini, A. Vacca and P. Gans, *Talanta*, 1974, **21,** 53.
23. M. Meloun and J. Čermák, *ibid.*, 1984, **31,** 947.
24. D. J. Leggett and W. A. E. McBryde, *Anal. Chem.*, 1975, **47,** 1065.
25. M. Meloun and M. Javůrek, *Talanta*, 1985, **32,** 973.
26. F. James, *Comp. Phys. Commun.*, 1975, **10,** 343.
27. L. G. Sillén, *Acta Chem. Scand.*, 1964, **18,** 1085.