

Evaluation of Equilibrium Constants from Potentiometric Z -pH Normalized Curves Using the Partial Least Squares (PLS) Self-Calibration Method

^aJ. PERŮTKA, ^aJ. HAVEL, ^bL. GALINDO, ^bF. G. MONTELONGO, and ^bJ. J. ARIAS LEÓN

^a*Department of Analytical Chemistry, Faculty of Science,
Masaryk University, CZ-611 37 Brno, Czech Republic*

^b*Department of Analytical Chemistry, Faculty of Chemistry,
University of La Laguna, E-38204 La Laguna, Tenerife, Spain*

Received 22 April 1996

The Partial Least Squares (PLS) multivariate calibration method accompanied by the so-called "self-calibration" approach enables to evaluate equilibrium data and obtain stability or equilibrium constants. The idea of the PLS self-calibration technique is similar to the so-called "normalized graphs" method. The calibration matrix is calculated from proposed initial parameters applying a method of experimental design. The method is validated on examples of simulated and experimental data and it was proved that it converges quite well. The proposed method represents a new approach towards the equilibrium data evaluation. Even though it is demonstrated on examples of potentiometric data, the proposed PLS self-calibration method can be used to evaluate any kind of chemical equilibrium data. A comparison with the "classical" approaches that are used in most of the minimizing programs has been done.

Computer evaluation of equilibrium constants has recently been reviewed [1] and thoroughly discussed elsewhere [2–4]. It is obvious that any new approach to computation of chemical equilibria has not been presented since Lars Gunnar Sillén's LETAGROP [5] (known from 1961). All "general" programs for equilibrium constant computation are based on the use of a nonlinear least squares methods and various minimizing techniques. Briefly, all programs are using "trial and error" algorithm. It means that in almost every case a large number of models (possible equilibria) must be examined so that the best one of all tested could be obtained. Determination of stoichiometric coefficients together with stability constants, *i.e.* ESI (Estimation of Stoichiometric Indices) approach, can effectively help us to choose an acceptable model [6, 7].

There is an increased importance of multivariate calibration methods in multicomponent analysis in analytical chemistry, especially of those that use the partial least squares method (PLS) with decomposition into latent variables [8–10]. It has already been shown that the PLS method gives favourable results in comparison with multiple regression methods for spectrophotometric data [11]. It has also been used in a few other analytical applications [11–15], first in fluorimetry [12], in the evaluation of potentiometric acid-base titrations [13], and even in the evaluation

of ICP-AES data [14], and so on. Recently, the PLS method has also been proposed for evaluation of kinetic data [16–18].

The use of the PLS calibration method is investigated here as an alternative to classical approaches for the evaluation of potentiometric equilibrium data. In contrast to the former application in potentiometry [13], where concentrations were estimated on the basis of experimentally obtained calibration matrix, in this contribution we would like to demonstrate that the PLS calibration approach can also be used to evaluate equilibrium constants by using a calibration matrix with simulated data. There is some analogy with normalized projection maps, introduced in the fifties by Sillén and Biedermann [19–21], that represented an advanced graphical approach. Later on it was shown that they can be plotted and calculated by a computer using NORMEX program [22]. The NORMEX results are comparable to those of LETAGROP. The examples of its use are given also by Ferri and Wahlberg [23]. Using the PLS calibration approach the evaluation of equilibrium constants can be made in an analogous way, however, without the necessity of plotting and/or drawing the diagrams. This paper would like to demonstrate that the PLS calibration method has some advantages not only over graphical methods but also over the classical computer methods and that the technique of normalized graphs can easily be

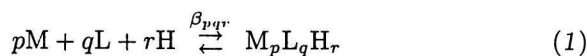
performed without a necessity to make any data transformations.

EXPERIMENTAL

The calibration data were simulated using the POLET program [6]. The estimation of unknown values of parameters was made using the PLS-computer program which was written in FORTRAN 77 according to the published algorithms [8, 9]. A standard PC 486/50 was used for the calculations.

THEORETICAL

The complex formation equilibria, *e.g.* of three components (metal, ligand, and proton), may be described by the reaction (1), stability constant (2), and mass-balance equations (3–5)



$$\beta_{pqr} = [M_p L_q H_r] / ([M]^p [L]^q [H]^r) = c / (m^p l^q h^r) \quad (2)$$

$$c_M = m + \sum_{i=1}^{n_c} p_i c_i \quad (3)$$

$$c_L = l + \sum_{i=1}^{n_c} q_i c_i \quad (4)$$

$$c_H = h + \sum_{i=1}^{n_c} r_i c_i \quad (5)$$

where c is a free concentration of a complex, c_i is a free concentration of an i -th complex, n_c is a number of complexes in an equilibrium mixture, m, l, h are free concentrations of metal, ligand, and proton, c_M, c_L, c_H are total concentrations of metal, ligand, and proton.

Potentiometric data can be normalized into a Z —pH function (dependence of Z on pH), where the variable Z represents *e.g.* the average number of protons bound per one ligand

$$Z = (c_H - h) / c_L = \sum_{i=1}^{n_c} r_i c_i / c_L \quad (6)$$

A linear relationship between Z and the concentration of the i -th complex results from the equation (6). There is an obvious similarity between the above relationship and a dependence of absorbance on concentration of a component in a mixture, thus eqn (6) can be transcribed analogously to the Beer—Lambert law in the matrix form (7) and considering $e_i = r_i / c_L$ as follows

$$\mathbf{Z} = \mathbf{C} \times \mathbf{E} \quad (7)$$

where \mathbf{Z} matrix is ($n_{\text{exp}} \times n_{\text{pH}}$), \mathbf{C} is ($n_{\text{exp}} \times n_c$), and \mathbf{E} is ($n_c \times n_{\text{pH}}$), n_{exp} is a number of experiments (potentiometric titration curves), n_{pH} is a number of

pH values (experimental points), and n_c is a number of complexes in an equilibrium mixture. The n_c can be estimated from matrix \mathbf{Z} by the PCA (Principal Component Analysis) method [24]. If n_c is known or determined, the PLS method can be applied for the prediction of $\log \beta$ of all complexes in the solution.

The PLS method has already been well described, an overview can be found elsewhere [8, 9]. Furthermore the PLS method has been proposed also to determine kinetic parameters [16, 17]. In contrast to the previous use of the PLS calibration method, where the calibration matrix is prepared experimentally, in this contribution it is suggested that (*i*) the equilibrium model is known and thus (*ii*) the calibration matrix and consequently the values of parameters can be calculated from the model without the necessity of an experimental work.

The PLS self-calibration method has two steps – calibration and prediction. The calibration needs two data matrixes: \mathbf{Z} (matrix of calibration solutions) and \mathbf{B} (matrix of experimental design of parameters). There should be a linear relationship between \mathbf{Z} and \mathbf{B} . Matrix \mathbf{Z} ($n_{\text{cal}} \times n_{\text{exp}}$), where n_{cal} is a number of calibration solutions, contains POLET-simulated data. \mathbf{Z} is created on the basis of an experimental design, *i.e.* on the basis of matrix \mathbf{B} which contains several levels of $\log \beta$ of assumed (or probable) complexes in the solution. \mathbf{B} is ($n_{\text{ed}} \times n_c$), where n_{ed} is a number of experiments in the experimental design. Number of calibration solutions is n_{ed} times higher than number of potentiometric titration curves, *i.e.* there are n_{ed} simulated calibration curves for each experimentally obtained curve. In other words, the calibration curves are simulated for the same c_M and c_L for which the experiments were done. **Calibration** is done by decomposition of matrixes \mathbf{Z} and \mathbf{B} into the so-called latent variables

$$\mathbf{Z} = \mathbf{F}_Z \mathbf{L}_Z + \mathbf{E}_Z \quad (8)$$

$$\mathbf{B} = \mathbf{F}_B \mathbf{L}_B + \mathbf{E}_B \quad (9)$$

where \mathbf{F}_Z is the latent \mathbf{Z} -matrix ($n_{\text{cal}} \times d$), *i.e.* latent matrix of calibration solution, d is a number of dimensions (number of abstract component), \mathbf{L}_Z represents the loading \mathbf{Z} -matrix ($d \times n_{\text{exp}}$), \mathbf{F}_B is the latent matrix of experimental design of parameters ($n_{\text{ed}} \times d$), *i.e.* latent matrix of $\log \beta$, \mathbf{L}_B is the loading \mathbf{B} -matrix ($d \times n_c$), \mathbf{E}_Z and \mathbf{E}_B are error matrixes with the same dimensions as the original \mathbf{Z} and \mathbf{B} matrix. The relationship between \mathbf{F}_B and \mathbf{F}_Z expresses a diagonal regression matrix \mathbf{V}

$$\mathbf{F}_B = \mathbf{F}_Z \mathbf{V} + \mathbf{E}_d \quad (10)$$

where \mathbf{E}_d is an error matrix. The matrix \mathbf{V} is necessary for the calculation of unknown values of the parameters ($\log \beta_1 \dots \log \beta_{n_c}$) in the **prediction** step

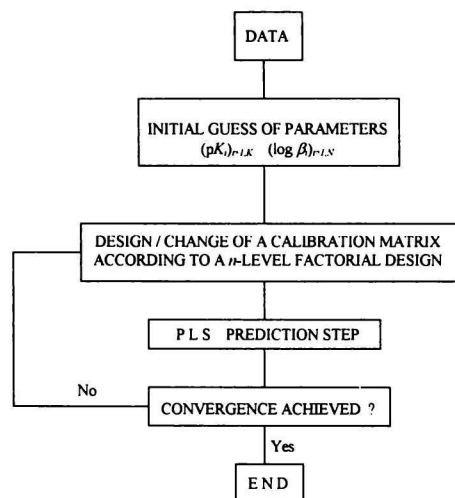


Fig. 1. A flow chart of the proposed PLS self-calibration method.

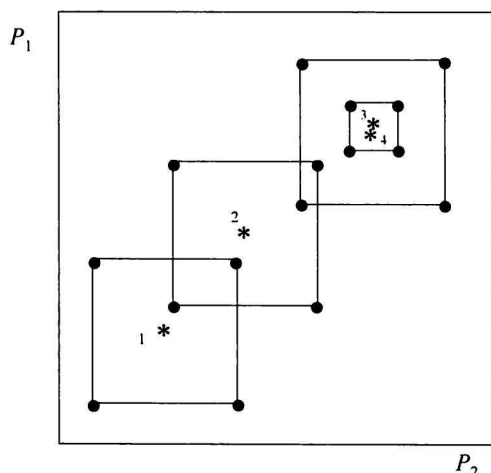


Fig. 2. A graphical presentation of the convergence in the case when two parameters P_1 and P_2 are estimated.

$$\log \beta_0 = z_0(\mathbf{F}_B^t \mathbf{Z})^t \mathbf{V} \mathbf{L}_B \quad (11)$$

where z_0 represents the experimentally obtained potentiometric titration curve, t is the symbol for the transposition of the matrix, $\log \beta_0$ is the vector of estimated parameters ($\log \beta_1 \dots \log \beta_{n_c}$).

Every new prediction of $\log \beta_0$ needs a new experimental design (proposed according to calculated parameters) and also a new POLET-simulated calibration matrix. The predicted $\log \beta_0$ represents the central point of a new experimental design. The convergence is achieved when the newly calculated $\log \beta_0$ changes not more than by the convergence tolerance. The flow chart of the algorithm proposed is given in Fig. 1. Fig. 2 shows the graphical presentation of the convergence in the case when two parameters P_1 and

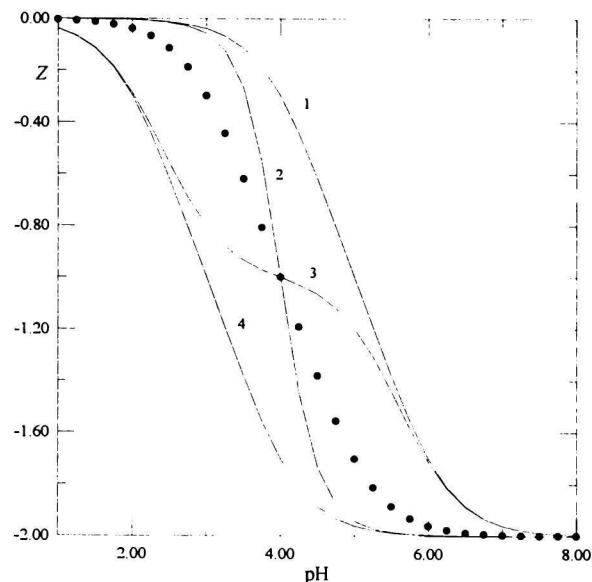


Fig. 3. A graphical presentation of the calibration set. The dots represent the test curve, lines the calibration curves 1—4.

P_2 are estimated. The first square represents a starting experimental design with a central point *1. According to these data the calibration set is simulated and a new central point *2 of a new factorial design is predicted. It also gives a situation when the convergence is achieved: the central point *4 is almost the same as the previous one, number *3.

RESULTS AND DISCUSSION

The potential of multivariate data evaluation using the PLS self-calibration approach will now be demonstrated on several examples.

Case 1. Dissociation of a Weak Acid in Two Steps, Two Near pK Values

Z -pH data were simulated with POLET so that 29 points were calculated for each series of parameters. The Z -values were loaded with random errors of Gaussian distribution with different value of standard deviation s_{inst} . The four calibration mixtures for various combinations of pK_1 and pK_2 levels were calculated and used for the prediction of pK values from the test data. Fig. 3 shows the graphical presentation of the calibration set. The dots represent the test curve, lines the calibration curves. The results obtained are given in Table 1. It is evident that exactly the same values are obtained for which the data were simulated. Therefore in the next examples the use of the PLS self-calibration method will be demonstrated on the obtained experimental data.

Table 1. Results of Estimation of Two Near pK Values from Simulated Normalized Potentiometric Z —pH Data

Method	pK_1	pK_2	Ref.
Correct values	3.40	4.60	This paper
POLET program	3.419 ± 0.001	4.61 ± 0.01	This paper
PLS method	3.413 ± 0.001	4.601 ± 0.001	This paper

Table 3. Results of the Estimation of Two Near Dissociation Constants of 2,2'-Bis(carboxymethylmercapto)diethyl Ether (LH₂)

Method	pK_1	pK_2	Ref.
Graphical method	3.60	3.84	[25]
POLET program	3.407 ± 0.036	4.149 ± 0.052	This paper
PLS method	3.403 ± 0.034	4.151 ± 0.053	This paper

Case 2. Determination of Two Near pK Values of 2,2'-Bis(carboxymethylmercapto)diethyl Ether (LH₂)

50 cm³ of 0.98 mM solution of ligand was titrated with 0.1000 M-NaOH at ionic strength $I(\text{KNO}_3) = 0.1 \text{ mol dm}^{-3}$ at 20°C. The experimental data are given in Table 2, the results of computation with the POLET computer program and the PLS self-calibration method are in Table 3. A comparison with a graphical method [25] shows a very good agreement between results of the general regression POLET program and the PLS self-calibration method proposed.

Case 3. Determination of Two pK Values of 3-(2-Thiazolylazo)-2,6-dihydroxypyridine (TADHP)

150 cm³ of 0.666 mM-TADHP containing 1.733

Table 2. Experimental Data of the Determination of Two Near Dissociation Constants of 2,2'-Bis(carboxymethylmercapto)diethyl Ether (LH₂)

pH	3.325	3.410	3.510	3.616	3.750	3.897	4.052
Z	1.516	1.398	1.275	1.139	1.000	0.848	0.683
pH	4.260	4.520	4.802	4.974	5.196	5.581	6.182
Z	0.514	0.337	0.248	0.193	0.156	0.119	0.080

Table 6. Results of the Determination of $-\log \beta_{pq}$ of Uranium(VI) Hydrolytic Species

Method	$(\text{UO}_2)_2(\text{OH})_2^{2+}$ $-\log \beta_{22}$	$(\text{UO}_2)_3(\text{OH})_5^+$ $-\log \beta_{35}$	$(\text{UO}_2)_3(\text{OH})_4^{2+}$ $-\log \beta_{34}$
LETAGROP program	6.20 ± 0.01	16.75 ± 0.02	13.02 ± 0.07
PLS method	6.19 ± 0.02	16.76 ± 0.02	13.04 ± 0.07

Table 4. Experimental Data in the case of Determination of Two pK Values of TADHP in Dioxane—Water ($\varphi = 50 \text{ vol. \%}$)

pH	Z	pH	Z	pH	Z
4.761	0.025	5.121	0.077	5.376	0.130
5.548	0.182	5.700	0.235	5.818	0.287
5.931	0.340	6.015	0.392	6.114	0.445
6.210	0.497	6.298	0.550	6.394	0.602
6.491	0.655	6.594	0.707	6.714	0.760
6.856	0.812	7.008	0.865	7.206	0.917
7.470	0.970	7.901	1.022	8.678	1.075
9.866	1.180	10.550	1.285	10.910	1.390
11.135	1.495	11.303	1.600	11.480	1.705
11.612	1.810	11.718	1.915	11.843	2.020

Table 5. Results of the Determination of Two pK Values of TADHP in Dioxane—Water ($\varphi = 50 \text{ vol. \%}$)

Method	pK_1	pK_2	Ref.
Graphical method	6.20	11.38	[25]
POLET program	6.223 ± 0.034	11.45 ± 0.14	This paper
PLS method	6.221 ± 0.032	11.50 ± 0.09	This paper

mM perchloric acid was titrated with 0.105 M-NaOH at 25°C. The experimental data are in Table 4, the results of computation with the POLET computer program and the PLS self-calibration method are given in Table 5.

Case 4. Determination of $-\log \beta$ of Uranium(VI) Hydrolytic Species

Experimental data obtained for 3 M-NaClO₄ are given in [26]. The results of computation with the LETAGROP computer program and the PLS self-calibration method are in Table 6.

The PLS self-calibration technique, demonstrated on several examples above, can be an alternative for the evaluation of equilibrium constants from potentiometric titration curves. In contrast to most of commercially available software packages for equilibrium constant computation this new approach can estimate acceptable values of unknown parameters more effectively because of the use of experimental design for initial guess of $\log \beta_0$. In other words, it is possible to test a wide range of unknown parameter values (at several levels of magnitude) and at the same time. This is hardly possible in classical nonlinear methods. Mostly, there also exists an unpleasant sensitivity to the first approximation of $\log \beta_0$. Very often if there is a bad initial guess of $\log \beta_0$, the minimizing techniques do not converge at all in spite of the fact that the tested complexes are present in the equilibrium mixture. These limitations could be overcome with the use of the PLS self-calibration technique.

CONCLUSION

It was shown that with the PLS calibration method and proposed method of self-calibration it is possible to determine acid-base/stability constants from potentiometric normalized Z -pH curves in an analogous way to using normalized maps invented by Sillén but without a necessity of drawing any diagrams.

The method works quite rapidly and the obtained results are satisfactory in comparison with those of standard potentiometric programs like LETAGROP-Z+ETA, ESAB, POLET, SUPERQUAD, and so on.

The PLS self-calibration technique is still tested with various experimental data, the possibilities of ESI approach in this method are studied and the algorithm of calibration step is being improved so that an algorithm as general as possible could be proposed for the evaluation of equilibrium data and calculation of formation constants.

REFERENCES

- Meloun, M., Havel, J., and Högfeltd, E., *Computation of Solution Equilibria, A Guide to Methods in Potentiometry, Extraction and Spectrophotometry*. Ellis Horwood, Chichester, 1985.
- Legget, D. J. (Editor), *Computational Methods for the Determination of Formation Constants*. Plenum Press, New York, 1985.
- Meloun, M. and Havel, J., *Computation of Solution Equilibria 2. Potentiometry. Folia Fac. Sci. Nat. Univ. Purk. Brun. XXVI, Chemia 19*, Op. 9. Pp. 145. Technical University, Brno, 1985.
- Havel, J. and Meloun, M., in *Computational Methods for the Determination of Formation Constants*. (Legget, D. J., Editor.) P. 221—289. Plenum Press, New York, 1985.
- Dyrssen, D., Ingri, N., and Sillén, L. G., *Acta Chem. Scand.*, A 15, 604 (1961).
- Havel, J. and Meloun, M., *Talanta* 33, 525 (1986).
- Havel, J. and Meloun, M., *Talanta* 33, 435 (1986).
- Wold, H., in *System under Indirect Observation*, Vol. 2. (Jöreskog, K. G. and Wold, H., Editors.) North-Holland, Amsterdam, 1982.
- Lorber, A., Wangen, L. E., and Kowalski, B. R., *J. Chemometrics* 1, 19 (1987).
- Lindberg, W., Öhman, J., Wold, S., and Martens, H., *Anal. Chim. Acta* 41, 174 (1985).
- Otto, M. and Wegscheider, W., *Anal. Chem.* 57, 63 (1985).
- Lindberg, W., Persson, J. A., and Wold, S., *Anal. Chem.* 55, 643 (1983).
- Lindberg, W. and Kowalski, B., *Anal. Chim. Acta* 206, 125 (1988).
- Havel, J., Moreno, C., Hrdlička, A., and Valiente, M., *2nd Scand. Symp. on Chemometrics*, 28—31 May, Bergen, Norway, 1991.
- Perůtka, J., Havel, J., and Jančář, L., *Collect. Czech. Chem. Commun.* 59, 1951 (1994).
- Havel, J., Cuesta, F. S., and Jančář, L., *KAC'92, Kinetics Methods in Analytical Chemistry*. Erlangen, Germany, September 27—30, 1992.
- Havel, J., Cuesta, F. S., and Jančář, L., *React. Kinet. Catal. Lett.* 49, 189 (1993).
- Havel, J., Jimenez, F., Bautista, R. D., and Arias, J. J., *Analyst* (London) 118, 1355 (1993).
- Sillén, L. G., *Acta Chem. Scand.*, A 10, 186 (1956).
- Biedermann, G. and Sillén, L. G., *Acta Chem. Scand.*, A 10, 203 (1956).
- Biedermann, G. and Spiro, T. G., *Chem. Scripta* 1, 193 (1971).
- Ferri, D., Wahlberg, O., and Högfeltd, E., *Acta Chem. Scand.*, A27, 3591 (1973).
- Ferri, D. and Wahlberg, O., *Chem. Commun.* (Stockholm), No. XII (1973).
- Havel, J. and Meloun, M., *Talanta* 32, 171 (1985).
- Galindo, L., *Doctoral Thesis*. Universidad de La Laguna, 1982.
- Lubal, P., *Diploma Thesis*. Masaryk University, Brno, 1993.

Translated by J. Havel