References

1 - Mota, A.M.; Santos, M.C.; chap. 5 of *Metal Speciation and Bioavailability in Aquatic Systems*, ed. by A. Tessier and D.R. Turner, IUPAC series, John Wiley & Sons (1995).

2 - Mota, A.M.; Gonçalves, M.L.S.; chap.2 of *Element Speciation in Bioinorganic Chemistry*, Chemical Analysis Series, vol. 135, S. Caroli ed., John Wiley & Sons (1996).

3 - Buffle, J.; in *Complexation Reactions in Aquatic Chemistry*, Ellis Horwood, John Wiley & Sons (1988).

4 - Mota, A.M.; Buffle, J.; Kounaves; S.P., Simões, M.L.; *Analytica Chimica Acta*, **172** (1985) 13.

5 - Bugarin, M.G.; Mota, A.M.; Pinheiro, J.P.; Gonçalves, M.L.S.; *Analytica Chimica Acta*, **294** (1994) 271.

6 - Capelo, S.; Mota, A.M.; Gonçalves, M.L.S.; *Electroanalysis*, 7 (1995) 563.

7 – Wonders, J.; van Leeuween, H.P.; *Electrochimica Acta* (1998), in press.

8 - Filella, M.; Buffle J.; van Leeuwen, H.P.; *Anal. Chim. Acta*, **232** (1990) 209.

9 - Pinheiro, J.P.; *tese de doutoramento*, IST, Lisboa (1996).

10 - Boal, I.; *tese de doutoramento*, Universidade de Aveiro, Aveiro (1997).

11 - Mota, A.M.; Cruz, P.; Gonçalves, M.L.S; *Env. Sc. Technol.*, submitted.

12 - Muller, F.L.L.; Kester, D.R.; *Mar. Chem.*, **33** (1991) 71.

13 - Bruland, K.W.; *Limnol. Oceanogr.*, **34** (1989) 269.

14 - Donat, J.R.; Bruland, K.W.; *Mar. Chem.*, **28** (1990) 301.

15 - Capodaglio, G.; Coale, K.H.; Bruland, K.W.; Mar. Chem., **29** (1990) 221.

16 - Capodaglio, G.; Toscano, G.; Scarponi, G.; Gescon, P.; *Ann. Chim.*, **79** (1989) 543.

17 - Coale, K.H.; Bruland, K.W.; *Limnol. Oceanogr.*, **33** (1988) 1084.

18 - Hanson, A.K.; Sakamoto, A., Jr.; Carole, M.; Hizenga, D.L.; Kester, D.R.; *Mar. Chem.*, **23** (1988) 181.

19 - Van den Berg, C.M.G.; *Mar. Chem.*, **15** (1984) 1.

20 - Scarano, G.; Bramanti, E.; Zirino, A.; *Anal. Chim. Acta*, **264** (1992) 153.

NON-LINEAR MODELING OF POTENTIOMETRIC ACID-BASE TITRATION CURVES BY PARTIAL LEAST SQUARES AND ARTIFICIAL NEURAL NETWORKS

Joaquim C.G. Esteves da Silva* and Adélio A.S..C. Machado
LAQUIPAI, Chemistry Department, Faculdade de Ciências do Porto,
R. Campo Alegre 687, P4150 Porto, Portugal

Abstract. A procedure for the quantitative analysis of mixtures of acids, based in the treatment of data of potentiometric pH titrations by multivariate chemometric calibration methods, is presented. Experimentally, the procedure consists in the titration of a set of mixtures of acids similar to the sample, but with known concentration of the components, which constitute the calibration set. A two step data analysis procedure is used for the development of calibration models: (i) the raw titration curves are adjusted to a polynomial equation of the fifth degree and the inverted titration curves are obtained (volume as function of the pH); (ii) the inverted titration curves are used to develop calibration models, using multivariate chemometric methods like partial least squares and artificial neural networks which can deal with some non-linearity in the basic inverted model. The results showed that adequate predictions of the concentration of the acids of the mixture are obtained.

Key words. Potentiometric titration curves; Mixture of acids; Quantification of acids; Partial least squares; Artificial neural networks.

Introduction

Traditional methods of modeling potentiometric acid-base titration curves are based on equilibrium conditions and require the *a priori* fixation of a physical model expliciting the number, acid-base constants and concentration of the components involved in the titration. In the case of multicomponent samples, such models tend to become very complex [1-4], and the quantitative analysis of mixtures of acids by potentiometric (pH) titration may be a difficult or impossible task. Further problems arise in the accuracy and precision of the results when the acid-base properties of the components (protonation constants and coefficients of activity for correction of

concentrations) and the detailed background composition of the mixture (to calculate ionic strength) are not fully known, as is generally the case.

An alternative and simplified procedure for the analysis of potentiometric acid-base titration curves consists on the development of calibration models using directly their inverse curves (volume as function of the pH). The development of reliable chemometric multicomponent calibration techniques have had an active role in the development of useful calibration and prediction models [5,6]. In this new approach, the exact definition of the physical model is replaced by a two step procedure, consisting of calibration followed by prediction: (i) calibration: analysis of titration curves of mixtures of similar composition to those to be analyzed but with known concentration of the constituents; (ii) prediction: determination of the concentration of acids in mixtures by analysis of the respective titration curves.

This paper demonstrates the potentialities of the non-linear modeling of potentiometric (pH) titration curves, namely of the inverted titration curves (volume of titrant as function of the pH), for the quantitative analysis of aqueous mixtures of mono- and polyprotic acids. Two multivariate calibration chemometric techniques were used for processing experimental data: (i) partial least squares (PLS) [7]; and (ii) artificial neural networks (ANN) [8]. PLS is a technique developed for the analysis of linear models but that can accommodate non-linearity, while ANN is intrinsically developed as a non-linear modeling technique. As example, the results for a binary mixture of succinic and maleic acid are presented and discussed.

Experimental

Reagents. Analytical grade reagents were used. A 0.2 M decarbonated potassium hydroxide was used as titrant of the solutions of acids prepared in deonized water.

Equipment. An automatic titration system was used, which allowed the addition of constant volumes of titrant after the potential measurements achieved a predefined stabilization criteria (in this case, ±0.2 mV after about 3 minutes). The system was composed by equipment similar to that described in previous works [9-11].

Data Analysis. Several home made programs, developed in TurboPascal (Borland, USA), were used in pre-processing and data manipulation,. The following commercial software packages were used: (i) SPSS (The SPSS Company, USA) (Statistical Package for the Social Sciences), for fifth degree polynomial fit of the raw titration curves; (ii)

GRAMS-32 (Galactic Industries Corporation, USA) for PLS calculations; (iii) Neural Works Pro II (Neural Ware Inc., USA) for ANN calculations.

Multivariate calibration chemometric methods

The basic model. The standard way of representing a potentiometric titration curve is by plotting the measured pH electrode response (milivolts or pH values) as function of the volume of added titrant (in this case hydroxide). Fig. 1.a shows a typical example of such a pH=f(volume) curve. To define an instrumental response which is roughly proportional (almost linear relationship) to the concentration of the mixture of acids under titration, the volume of titrant must be estimated as function of the pH [volume=f(pH)], as shown in Fig. 1.b. The meaning of this relationship is due to the fact that the volume of titrant added at a particular pH value is being used in the titration of the acid species with pKa most close to that value of pH.

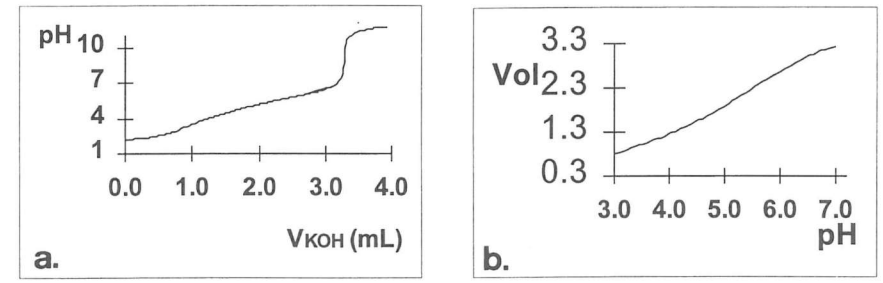

Fig.1 - Raw (a) and inverted (b) titration curves for the titration of a mixture of succinic and maleic acids with potassium hidroxide

By defining a pH range and selecting a pH resolution (depending on the problem under analysis) for the analysis of each data set, the following basic calibration model is obtained (in matrix notation)

$$\mathbf{C} = \mathbf{V}\,\mathbf{P} + \mathbf{E} \qquad (1)$$

where **C**(naxnc) is the matrix of the concentrations (na - number of titration curves; nc - number of component acids in the mixture), **V**(na x np) is the matrix of the volumes (np - number of points per titration curve), **P**(np x nc) is the matrix of proportional constants and **E**(na x nc) is the matrix of the residual error. After calibration, this equation can be

used for predicting the concentration of mixtures of acids with unknown concentration. Eq. 1 can be schematically represented by

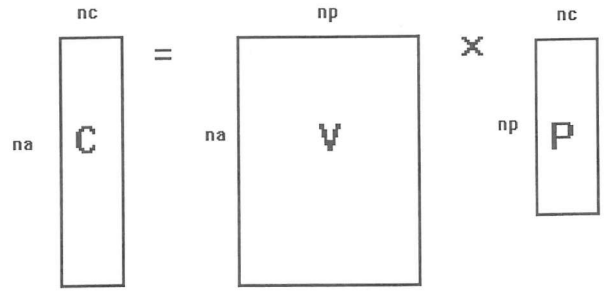

Fig. 2 - Schematical representation of the calibration model

Partial least squares (PLS). This method performs the simultaneous decomposition of both the matrices **V** and **C** into a set of two smaller matrices (the matrices of the scores, **T**(na x nf) and **U**(na x nf), and loadings, $\mathbf{F^T_v}$(nf x np) and $\mathbf{F^T_c}$(nf xn c), respectively - where nf is number of factors detected in the PLS) [7]. The calibration is established between the two score matrices instead of eq.1. In the case of a non-linear relationship between **C** and **V**, a number of factors larger than the number of components is considered. The decomposition process and calibration is schematically represented by

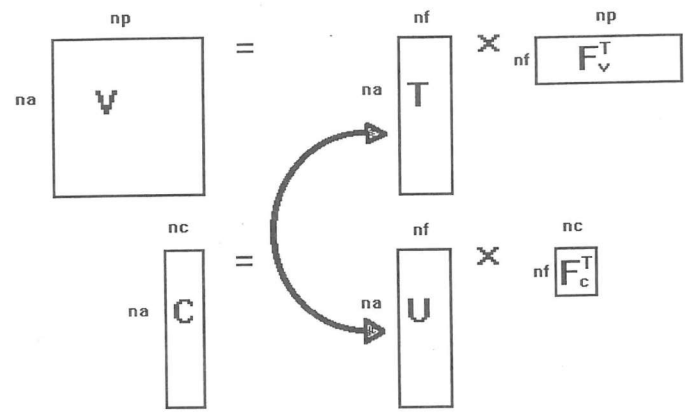

Fig. 3 - Schematical representation of the PLS decomposition and calibration model

Artificial neural networks (ANN). The ANN attempts to simulate and understand what goes in the nervous system (human brain) with the hope of capturing some of the learning and predictive capacities of its biological systems. The basic structures of a ANN are the neurons which are organized in layers. Each neuron in a layer is connected to all the neurons in the next layer with an associate weight. The properties of the ANN depend on the characteristics of the network rather than the neuron and the learning process consists in the iterative modification of the weights until the capacities of learning and prediction are achieved [8]. Fig. 4 shows an optimized ANN for the determination of succinic and maleic acids, constituted by three layers.

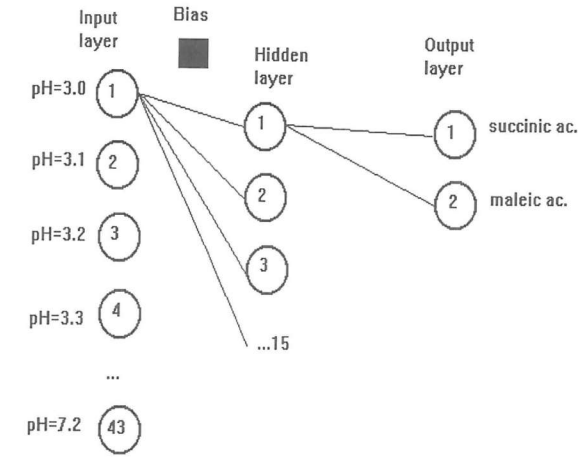

Fig. 4 - Schematical representation of the ANN architecture used in the calibration model developed for the analysis of a mixture of succinic and maleic acids

Results and discussion

Fig. 1.a shows a typical titration curve obtained for a mixture of succinic and maleic acids. Twenty four mixtures of these two acids were obtained using a full factorial design with two factors and five levels corresponding to the following concentrations: 0.00, 0.005, 0.010, 0.020 and 0.050 M. After fitting these titration curves to a fifth degree

Table 1 Results for the prediction of the acids concentration

Sample	True Concentration	PLS		ANN	
		Pred.	Error	Pred.	Error
		Succinic acid			
1	0.00499	0.00562	0.00062	0.00450	0.00049
2	0.00499	0.00617	0.00118	0.00399	0.00100
3	0.00999	0.00999	0.00000	0.00940	0.00059
4	0.00999	0.00908	-.00090	0.00960	0.00039
5	0.01997	0.02077	0.00080	0.02001	-.00004
6	0.04993	0.04796	-.00197	0.05306	-.00314
7	0.04993	0.04961	-.00032	0.04494	-.00499
RMS*			0.00150		0.00228
		Maleic acid			
1	0.00524	0.00428	-.00096	0.00503	0.00021
2	0.02095	0.02112	0.00017	0.02137	-.00142
3	0.00000	-.00108	-.00108	-.00051	-.00051
4	0.01047	0.01152	0.00105	0.00951	0.00096
5	0.00524	0.00433	-.00091	0.00470	-.00054
6	0.00000	0.00055	0.00055	-.00364	0.00364
7	0.01047	0.00979	-.00069	0.01367	-.00032
RMS*			0.00083		0.00196

* Root mean square.

polynomial equation, inverted titration curves were obtained (Fig. 1.b) between pH values of 3.0 to 7.2, with resolution of 0.1 pH units (43 points).

A randomly distributed sub-set of twelve inverted titration curves was used to develop calibration models for PLS and ANN. The PLS model used three factors showing that a non-linear relationship is observed between the volume of potassium hydroxide and the concentration of the acids. An adequate calibration was obtained by PLS, *i.e.*, a good correlation was obtained between the observed and estimated acid concentrations (linear correlation >0.99). Fig. 4 shows the optimized ANN architecture to obtain good calibration and estimation characteristics. Using this ANN architecture the calibration characteristics were similar to these obtained by PLS.

Table 1 shows the prediction of the concentration of the two acids of the test set by PLS and ANN. A global analysis of the results presented in this table suggest that both calibration techniques provided adequate predictions.

Conclusions

The results obtained in this study suggest that complex mixtures of acids (or bases) can be analyzed by direct interpretation of the potentiometric titration curves by non-linear modeling. This analysis is achieved without need of a detailed thermodynamic description of the acid-base properties of the constituents of the mixtures. Multivariate calibration chemometric methods of partial least squares (PLS) and artificial neural networks (ANN) allow to obtain adequate predictions. These calibration techniques can be hyphenated with software for automatic titration control to originate powerful analytical instruments.

Acknowledgments. A Praxis XXI Nr 2/2.1/QUI/294/94 project grant is acknowledged to JNICT (Lisbon).

References

1. F.Ingman, A.Johansson, S.Johansson and R.Karlsson, *Anal. Chim. Acta*, 64 (1973) 113-120.
2. G.Arena, E.Rizzarelli, S.Sammartano and C.Rigano, *Talanta*, 26 (1979) 1.
3. M.Betti, P.Papoff and L.Meites, *Anal. Chim. Acta*, 182 (1986) 133-145.
4. L.Meites, *Anal. Letters*, 15 (1982) 507-517.
5. W.Lindberg e B.Kowalski, *Anal. Chim. Acta*, 206 (1988) 125-135.
6. X.H.Song, J.Xu e R.Q.Yu, *Mikrochim. Acta*, 111 (1993) 199-206.
7. P.Geladi and B.Kowalski, *Anal. Chim. Acta*, 185 (1986) 1-17.
8. J.Zupan and J.Gasteiger, Neural Networks for Chemistry, VCH, Weinheim, 1993.
9. J.C.G.Esteves da Silva and A.A.S.C.Machado, *Anal. Letters*, 28 (1995) 2401-2411.
10. J.C.G.Esteves da Silva and A.A.S.C.Machado, *Analyst* , 120 (1995) 2553-2560.
11. J.C.G.Esteves da Silva, A.A.S.C.Machado and C.S.P.C.O.Silva, *Anal. Chim. Acta*, 318 (1996) 365-372.

Received, December 3, 1997
Revised, January 19, 1998