Scientific Paper

POTENTIOMETRIC TITRATION OF SOME HYDROXYLATED BENZOIC ACIDS AND CINNAMIC ACIDS BY ARTIFICIAL NEURAL NETWORK CALIBRATION

A. Hakan Aktaş,[a]* Samim Yaşar[b]

[a] *Süleyman Demirel University Faculty of Science and Art, Department of Chemistry, 32260 Isparta-Turkey, e-mail: ahakan@fef.sdu.edu.tr*

[b] *Süleyman Demirel University, Faculty of Forestry, 32260 Isparta-Turkey*

Received 04-10-2003

Abstract

In this study, two hydroxycinnamic acids and two hydroxylated benzoic acids, namely 4-hydroxycinnamic acid (*p*-coumaric acid), 4-hydroxy-3,5-dimethoxycinnamic acid (sinapinic acid), 3,4-dihydroxy benzoic acid (vanillic acid) and 3-hydroxy-4-methoxy benzoic acid (izovanillic acid), were titrated potentiometrically using tetrabuthylammonium hydroxide in 2-propanol under a nitrogen atmosphere at 25 °C. An artificial neural network (ANN) was applied for data treatment as a multivariate calibration tool in a potentiometric acid-base titration. The artificial neural network trained by the back-propagation learning was used to model the complex non-linear relationship between the concentration of *p*-coumaric acid (HpC), sinapinic acid (HS), vanillic acid (HV), and izovanillic acid (HiV), and the milivolt (mV) of solutions at different volumes of the added titrant. The principal components of the mV matrix were used as the input of the network. The optimized network predicted the concentrations of acids in synthetic mixtures. The results showed that the ANN used can proceed the titration data with an average relative error of less than 4.18%.

Introduction

A multivariate calibration has historically been a major cornerstone of the chemometrics as applied to the analytical chemistry. Most of the chemometrics involves the multivariate calibration. Some groups have based much of their development over the past two decades primarily on applications of the partial least squares (PLS) algorithm. The PLS is often regarded as the major regression technique for multivariate data. In fact, in many cases it is applied inappropriately and is not justified by the data. In areas outside the main stream of the analytical chemistry, or even biometrics and psychometrics, the PLS certainly is an invaluable tool, because the underlying factors have little or no physical meaning so a linearly additive model in which each underlying factor can be interpreted chemically is not expected.[1] The PLS is the multivariate calibration method that has received considerable attention in the chemometric

literature.[2-4] Recently, much attention has been paid to the application of the novel method of artificial neural networks in chemistry and some satisfactory results have been achieved.[5-10] In the analytical chemistry, ANN have been used for calibration,[11-13] parameter estimation,[14-16] spectrophotometric research,[17-20] and optimization of analytical conditions.[21,22] The corresponding non-linear multivariate maps use a non-linear transformation of the input variable to project inputs on to the designated attribute values in the output space. The strength of the modeling with layered; feed-forward ANN lies in the flexibility of the distributed soft model defined by the weights of the network. Both linear and non-linear mapping functions can be modeled by suitably configuring the network. The multilayer feed-forward neural network trained with back-propagation learning algorithm becomes an increasingly popular technique.[23-26]

The PLS regression, one of the multivariate calibration methods, was applied to potentiometric titration for determination of acid mixture.[27] In this study, the researchers used the volumes needed to reach a given pH as response data. They assumed a linear relationship between the volume of titrant added and the analytes concentrations. The PLS calibration method has been applied to acid-base titration, complexometric titration[28] and potentiometric precipitation titration[29] also by other authors. However, in the complex acid-base systems the interactions between components in the titration vessel becomes complicated. To overcome this problem, Song et al.,[30] Bronjdak-Voncina et al.[31] and Zampronia et al.[32] used ANN to treat potentiometric acid-base titration.

Hydroxycinnamic acids are commonly found in foods such as fruits, vegetables and grains. The highest concentrations of these phenolics are typically found in the surface layer. Accordingly, it has been speculated that these compounds play some role in the natural fungal resistance[33] of these foods. These phenolics acids have also been associated with sour, bitter and astringent flavors found in the vegetable proteins.[34] Hydroxycinnamic acids have also been associated with accelerated browning.[35,36] Because these acids can influence both the color and the flavor of a variety of food products, many analytical procedures have been developed for their determination.

There are number of publications on the titration of carboxylic acids and phenolics in non-aqueous media.[37-42] But only one report with the titration of cinnamic acids in non-aqueous media has appeared.[43]

The non-linear relationship between mV and analyte concentration can be modeled by ANN. In this study, we used a four-layer ANN with back-propagation of error algorithm for modeling the complex relationship between mV and concentration through a multicomponent acid-base titration. In order to decrease the number of data points, the data were factor analyzed before entering into ANN. The original data were used as input of the neural network. The method was applied to simultaneous determination of weak hydroxylated cinnamic acids and hydroxylated benzoic acids (i.e., *p*-coumaric acid (HpC), sinapinic acid (HS), vanillic acid (HV), and izovanillic acid (HiV)) in their quarternary mixtures and satisfactory results were obtained.

Experimental

Apparatus

Electrode potentials were measured using a Hanna HI 9321 Microprocessor pHmeter. A glass-silver-silver chloride electrode system was used and the silver-silver chloride electrode was modified by replacing the saturated aqueous KCl solution with a saturated solution of KCl in methanol.

Reagents

Analytical reagent grade chemicals were used unless indicated otherwise. *p*-coumaric acid, sinapinic acid, vanillic acid and izovanillic acid were purchased from Sigma (99% pure) and used without purification. Tetrabuthylammonium hydroxide (TBAOH) was purchased from Merck as a 0.100 M solution in 2-propanol/methanol and was diluted with pure, dry 2-propanol to give an approximately 0.020 M solution.

Throughout the work ca. 2.0×10^{-3} M solutions of the phenolic compounds were titrated (five runs) with 0.0200 M solutions of titrants.

Procedure

In a typical titration, suitable amounts of individual acids or acids mixture were placed in a 50 mL vessel and 5 mL 2-propanol was added to the solution. The solution was stirred and titrated with 0.0200 M tetrabuthylammonium hydroxide solution using a

micro-burette. The mV was recorded after each 0.02 mL addition of titrant. For each solution, at least 70 data point were recorded.

Methodology

A feed-forward ANN model with four layers of nodes was constructed as in Figure 1.

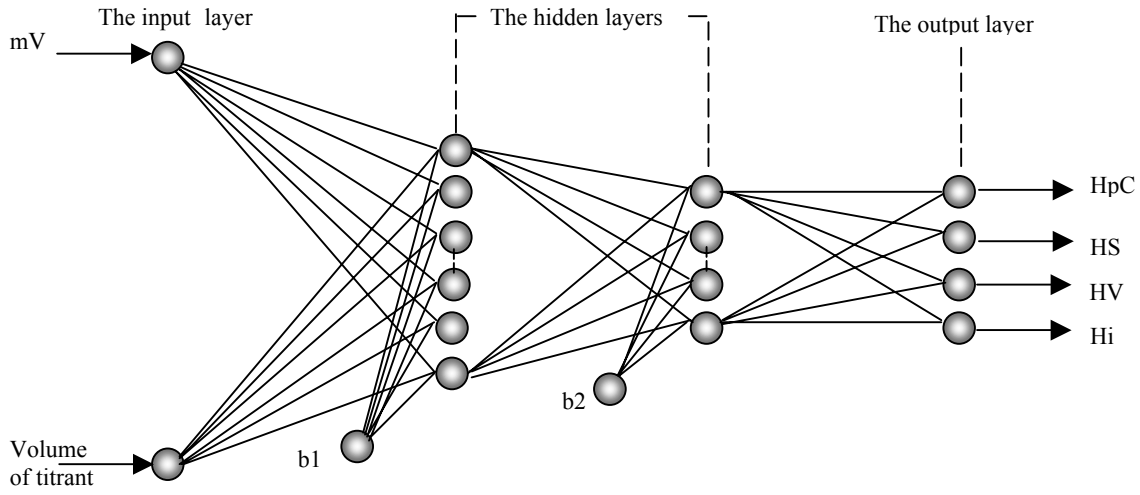

Figure 1. Network architecture used in the potentiometric titration modeling.

The logistic function was used as the activation function in a neural network. The training and testing data sets must be normalized into a range 0.1-0.9. The input and the output data sets were normalized by using following equation:

$$X_N = 0.1 + \frac{0.8(X - X_{\min})}{(X_{\max} - X_{\min})} \quad (1)$$

where X_N is normalized value of a variable (the network input or the network output), X is an original value of the variable, and X_{max} and X_{min} are the maximum and the minimum original values of the variables, respectively. In order to produce sufficient data for training and testing of the model shown in Figure 1, 15 different standard solutions were prepared using different acid concentrations and each standard solutions was subjected to potentiometric titration. Randomly chosen 857 data pairs from these

1070 data pairs were used for training of the neural network, and the rest of the data were used for testing. The root mean square error values were calculated from following equation to prove quantitatively the accuracy of the testing results of neural network models:

$$RMS = \sqrt{0.5N^{-1}\sum_{i=1}^{N}(X_1^{'} - X_1)^2} \qquad (2)$$

where N is the number of testing data and $X_1^{'}$ is target value.

Results and Discussion

The acids used in this study are chemically related compounds with close acidity constants. Figure 2 and 3 show the mV titration curves of these acids and their mixtures. It is obvious that the titration curves of these four acids are overlapped seriously.

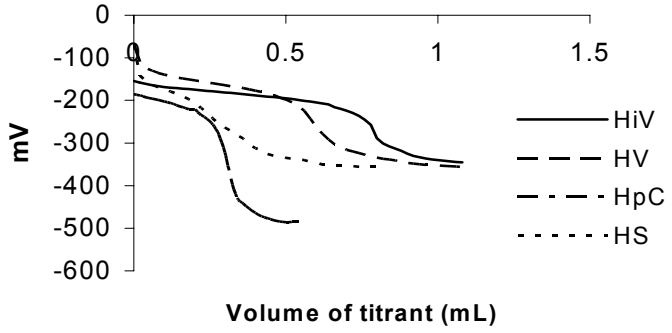

Figure 2. Titration curves for HpC, HS, HV and HiV with tetrabuthylammonium hydroxide.

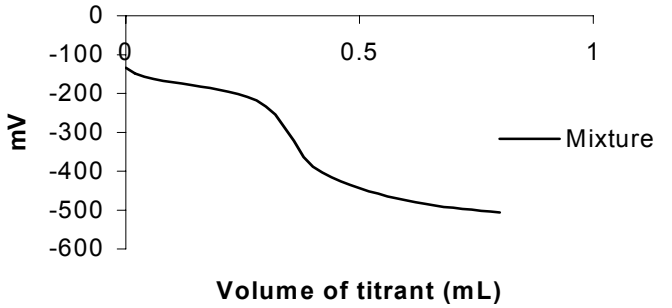

Figure 3. Titration curves equimolar quaternary mixtures with tetrabuthylammonium hydroxide.

To obtain the best network performance, the optimal network architecture and parameters must be chosen. Studies of the network structure include the selection of the number of layers and number of nodes in each layer. The number of layers used for this neural network modeling was four, i.e. an input layer, one or two hidden layers and an output layer. As can be seen from Figure 1, two neurons were used in the input layer, which were the mV and volume of titrant (mL). The titrant volume and mV value of the solution were considered as independent variables of the potentiometric titration method. Therefore, these variables were used as input variables in the network architecture. The neurons in the hidden layer were optimized for each acid and mixture solutions. In Table 1, the concentration of standard solutions is represented (or output data of the network).

Table 1. Concentration of different acids in quaternary standard solutions.

Sample number	Concentration mM			
	HpC	*HS*	*HV*	*HiV*
1	0.50	0.40	0.45	0.40
2	1.00	0.80	0.90	0.95
3	1.50	1.20	1.35	1.60
4	2.00	1.60	1.80	2.00
5	2.00	0.40	0.90	2.00
6	1.50	0.80	0.45	0.95
7	1.00	1.20	1.80	0.40
8	0.50	1.60	1.35	0.90
9	1.00	0.40	1.35	2.00
10	0.50	0.80	1.80	1.60
11	2.00	1.20	0.45	0.90
12	1.50	1.60	0.90	0.45
13	1.50	0.40	1.80	0.90
14	2.00	0.80	1.35	0.45
15	0.50	1.20	0.90	2.00

The various neural network models, which have the logistic function, were trained and tested. In this step, the number of the hidden layer units of the network was determined by performance evaluating of the network models defined in Table 2. According to RMS errors given in Table 2, the **NN 3-17-15-4** model, which performs best on the testing data set, were selected as neural network model to predict the acid concentrations.

Table 2. Comparison of the performances of the neural network models.

Model	*RMS error*							
	HpC		*HS*		*HV*		*HiV*	
	Training	*Testing*	*Training*	*Testing*	*Training*	*Testing*	*Training*	*Testing*
NN1 2-17-11-4	0.1855	0.21184	0.1195	0.18854	0.02572	0.03281	0.24539	0.21965
NN2 2-17-13-4	0.00497	0.00563	0.00351	0.00463	0.0177	0.0693	0.02368	0.02855
NN3 2-17-15-4	0.000394	9.03×10^{-5}	0.000315	7.0276×10^{-5}	0.00039	0.000193	0.000566	0.000229
NN4 2-17-17-4	0.03814	0.1025	0.02479	0.029984	0.04107	0.1874	0.05684	0.23684
NN5 2-17-21-4	7.0276×10^{-5}	0.02401	0.1893	0.2469	9.03×10^{-5}	0.03154	0.006847	0.00899

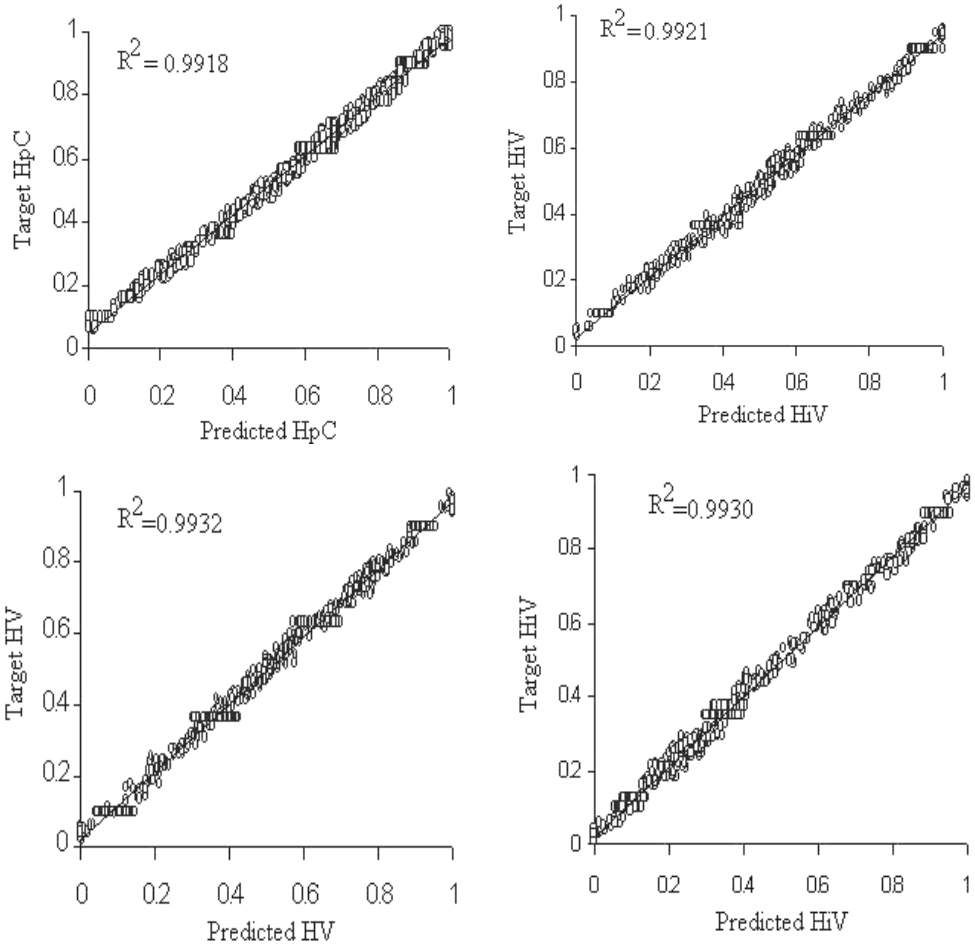

Figure 4. Comparison of predicted results from the NN3 model with target values from the actual values HpC, HS, HV and HiV for the training data set.

Figure 4 shows parity plots of predicted values of acid concentrations from NN3 model and the output values for the training data set. Figure 5 represents correlation between the outputs of the NN3 model and the target values from the actual values for the testing data set. Predictions with a RMS error of less than 0.001 for all acids indicated that each acid concentration in a given solution was accurately predicted by using the NN3 model. In addition to this, there is a good agreement between actual (target) results and predicted results from the model structured (see Figures 4 and 5).

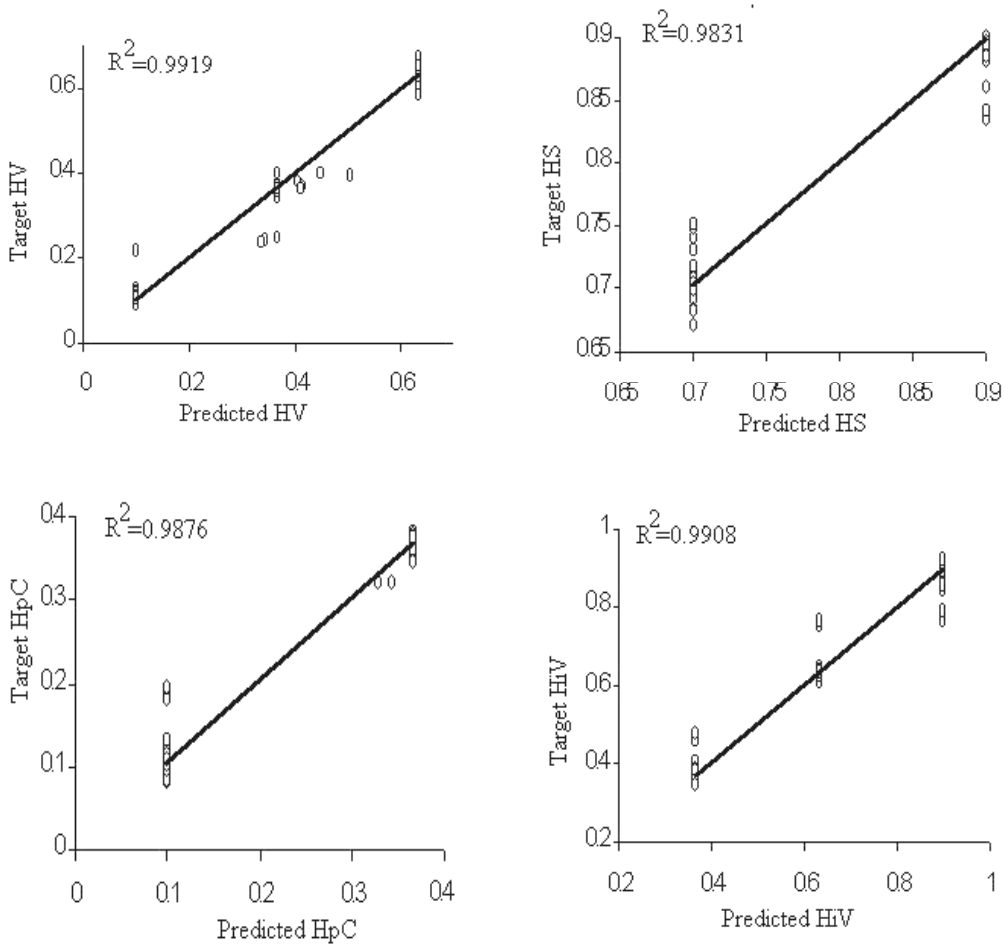

Figure 5. Comparison of predicted results from the NN3 model with target values from the actual values HpC, HS, HV and HiV for the testing data set.

Furthermore, several additional solutions, were prepared and titrated to show validation of the model selected. Experimental results and estimated results from the model were given in Table 3. As can be seen from the table, the error in the obtained estimation is at negligible level. The percent relative standard error of prediction is varied between -12.4 and 7.68. The low average relative error of prediction (<4.18%) indicates that the

networks used can properly process the titration data and model the complex relationship between the concentration of acids in the mixture and mV data at different volumes of the titrant.

Table 3. Statistical parameters calculated for the prediction set using optimized neural network models.

	Acid mixture composition											
	HpC			HS			HV			HiV		
	Actual (mM)	Predicted (mM)	RE %	Actual (mM)	Predicted (mM)	RE %	Actual (mM)	Predicted (mM)	RE %	Actual (mM)	Predicted (mM)	RE %
1	0.616	0.60764	-1.35	0.296	0.3	1.35	0.1	0.094	-6.00	0.491	0.49829	1.48
2	0.553	0.55925	1.13	0.804	0.84271	4.81	0.572	0.56384	-1.42	0.56	0.57515	2.70
3	0.9	0.86005	-4.43	0.816	0.845	3.55	0.9	0.91279	1.42	0.804	0.8515	5.90
4	0.754	0.76745	1.78	0.715	0.72337	1.17	0.214	0.19251	-10.0	0.28	0.25639	-8.43
5	0.697	0.72573	4.12	0.449	0.43093	-4.02	0.367	0.38	3.63	0.484	0.51415	6.22
6	0.264	0.23109	-12.4	0.268	0.28643	7.68	0.9	0.958	6.44	0.732	0.74558	6.20
7	0.9	0.883	-1.88	0.35	0.362	4.85	0.633	0.66183	4.50	0.384	0.396	3.12
8	0.6333	0.61445	-2.97	0.24	0.2376	-1.00	0.9	0.925	2.77	0.547	0.57609	5.31

Conclusions

Acid concentrations of these potentiometric titrations could be estimated by the neural network with an error that might easily be negligible. The neural network modeling could process the non-linear relationship between the mV of solutions at a given volume of titrant, and predict the concentration of acids in unknown sample solutions. For all acids, low prediction errors (<4.18%) and high correlation coefficients (0.9918, 0.9921, 0.9932 and 0.9930 for HpC, HS, HV and HiV, respectively) emphasize the high linear relationship between the predicted and actual concentrations.

Acknowledgements

The authors would like to thank Dr. E. Cafer Çilek from Süleyman Demirel University, Faculty of Engineering, Department of Mining.

References

1. R. G. Brereton, *Analyst* **2000**, *125*, 2125–2154.
2. P. Geladi, B. R. Kowalski, *Anal. Chim. Acta* **1988**, *185*, 1–17.
3. D. M. Haaland, E. V. Thomas, *Anal. Chem*. **1988**, *60*, 1193–1202.
4. K. P. Beebe, B. R. Kowalski, *Anal. Chem*. **1987**, *59*, 1007–1017.
5. G. Barka, J. Hlavay, *Talanta* **1997**, *44*, 2237–2245.
6. F. Despagne, D. L. Massart, *Analyst* **1998**, *123*, 157R–178R.
7. J. Zupan, J. Gasteiger, *Anal. Chim. Acta* **1991**, *148*, 1–30.
8. Z. Li, Z. N. Cheng, X. Li, T. H. Li, *Anal. Chim. Acta* **1993**, *65*, 393–396.
9. J. M. Andrews, S. H. Lieberman, *Anal. Chim. Acta* **1994**, *285*, 237–246.
10. S. Ventura, M. Silva, D. Perez-Bendito, *Anal. Chem*. **1995**, *67*, 1521–1525.

11. B. Hitzmann, T. Kullich, *Anal. Chim. Acta* **1994**, *294*, 243–249.
12. P. J. Gemperline, J. R. Long, V. G. Gregorion, *Anal. Chem*. **1993**, *63*, 2313–2323.
13. T. B. Blank, S. D. Brown, *Anal. Chem.* **1993**, *65*, 3081–3089.
14. L. S. Anker, P. C. Jurs, *Anal. Chem*. **1992**, *64*, 1157–1164.
15. M. D. Wessel, P. C. Jurs, *Anal. Chem*. **1994**, *66*, 2480–2497.
16. J. Han, R. R. Ruan, C. H. Park, *Biotechnol. Tech*. **1995**, *9*, 637–642.
17. J. W. Ball, P. C. Jurs, *Anal. Chem*. **1993**, *65*, 505–512.
18. J. W. Ball, P. C. Jurs, *Anal. Chem*. **1993**, *65*, 3615–3621.
19. C. R. Mittermayr, D. Drouen, M. Otto, M. Grasserbauer, *Anal. Chim. Acta* **1994**, *294*, 227–242.
20. R. Bro, *J. Chem*. **1995**, *9(5)*, 413–430.
21. H. Mo, B. Deng, *Chin. Chem. Lett.* **1996**, *3*, 265–266.
22. J. Zupan, J. Gasteiger, *Nueural Network for Chemists*, VCH Weinheim, ***1993***, p 197.
23. S. Seculic, M. B. Seasholtz, Z. Y. Wang, B. R. Kowalski, S. E. Lee, B. R. Holt, *Anal. Chem*. **1993**, *65*, A835–A845.
24. B. Walczak, *Anal. Chim. Acta* **1996**, *322*, 21–29.
25. P. A. Jansson, *Anal. Chem*. **1991**, *63*, 357A–362A.
26. C. Borggoard, H. H. Thodberg, *Anal. Chem*. **1992**, *64*, 545–551.
27. W. Lindberg, B. R. Kowalski, *Anal. Chim. Acta* **1988**, *206*, 125–135.
28. Y. Ni, Y. Wu, *Anal. Chim. Acta* **1997**, *354*, 232–240.
29. Y. Ni, Y. Wu, *Anal. Chim. Acta* **1999**, *390*, 117–123.
30. X. H.Song, J. Xu, R. Q. Yu, *Mikrochim. Acta* **1993**, *111*, 199–206.
31. D. B. Voncina, D. Dobcnik, M. Novic, J. Zupan, *Chemometr. Intell. Lab.* **1999**, *47*, 79–88.
32. C. G. Zampronio, J. J. R. Rohwedder, R. J. Poppi, *Chemometr. Intell. Lab.* **2002**, *62*, 17–24.
33. D. H. Hahn, J. M. Faubion, L. W. Rooney, *Cereal. Chem*. **1993**, *60(4)*, 255–260.
34. C. J. Huang, J. F. Zayas, *J. Food. Sci*. **1991**, *56*, 1308–1312.
35. H. E. Cen, *J. Chin. Agric. Chem. Soc*. **1998**, *26(1)*, 69–75.
36. J. L. Cilliers, V. L. Singleton, *J. Agric. Food. Chem*. **1989**, *37*, 890–895.
37. N. V. Meurs, A. M. F. Dahmen, *Anal. Chim. Acta*. **1959**a, *21*, 10–16.
38. C. A. Streuli, *Anal. Chem*. **1960**, *32(3)*, 407–410.
39. L. W Marple, J. S. Fritz, *Anal. Chem*. **1962**, *34(7)*, 796–800.
40. T. Gündüz, E. Kılıç, G. Özkan, *Analyst* **1988**, *13*, 1017–1020.
41. G. Franchini, A. Marchetti, C. Preti, L. Tassi, G. Tosi, *Anal. Chem*. **1989**, *61*, 177–184.
42. T. Gündüz, E. Kılıç, G. Özkan, M. F. Awaad, M. Taştekin, *Anal. Chim. Acta*. **1990**, *234*, 339–344.
43. A. H. Aktaş, G. Yaşar, G. Alsancak, Ş. Demirci, *Turk. J. Chem*. **2001**, *25*, 501–507.

Povzetek

V opisani raziskavi smo opravili potenciometrično titracijo dveh hidroksicimetovih kislin (4-hidroksicimetove kisline in 4-hidroksi-3,5-dimetoksicimetove kisline) in dveh hidroksiliranih benzojskih kislin (3,4-dihidroksibenzojske kisline in 3-hidroksi-4-metoksi benzojske kisline) s tatrabutilamonijevim hidroksidom v 2-propanolu v dušikovi atmosferi pri 25 °C. Za multikomponentno kalibracijo pri nevtralizacijski titraciji smo uporabili umetne nevronske mreže (UNM). UNM z vzvratnim širjenjem napake smo uporabili za modeliranje kompleksnih nelinearnih povezav med koncentracijo kislin in potencialom raztopine v mV pri različnih dodatkih titrne raztopine. Kot vhodne podatke za UNM smo uporabili glavne komponente matrike potencialov v mV. Optimizirana UNM je uspešno napovedovala koncentracije posameznih kislin v sintetični mešanici. Predstavljen model UNM lahko napove koncentacije prisotnih kislin iz podatkov titracijske krivulje s povprečno realtivno napako manj kot 4,18%.