ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2010, *55*(2), 85-97

STACKED NEURAL NETWORK MODELING APPLIED TO THE PROCESS OF METAL ION SORPTION BY THE FUNCTIONAL SILICA XEROGEL

Ciprian George PIULEAC[a], Silvia CURTEANU[a] and Maria CAZACU[b]

[a]"Gh. Asachi" Technical University Iaşi, Faculty of Chemical Engineering,
Bd. D. Mangeron, No. 71A, 700050, IAŞI, Roumania
[b]"Petru Poni" Institute of Macromolecular Chemistry, Aleea Gr. Ghica Voda 41A, IAŞI

Received May 18, 2009

Silica xerogels functionalized with hydroxy-azomethine groups have been prepared by using a one-pot procedure involving both the generating of the silica and their chemical modification within the same sol-gel system, in methanol. These compounds were tested as sorbents for copper (II) ions from aqueous solutions in different conditions. The acquired data have been applied for process modeling using simple and stacked neural networks based on multilayer perceptron, generalized feed-forward and Jordan - Elman neural networks which provided good predictions in validation phases.

INTRODUCTION

The removing or recovery and recycling of the metals from wastewaters constitute an important issue both from the environmental and economical point of view.[1,2] The metals need to be removed and/or recovered from the wastewater due to either their toxicity or value, as in the case of certain catalysts or precious metals.[3]

Silica gel represents an amorphous inorganic polymer containing internal siloxane groups (Si–O–Si) and silanol groups (Si–OH) distributed on the surface. It is commonly used as a rigid matrix for ligand immobilization.[4, 5] The silica gel surface can be chemically modified with various donor atoms such as N, S, O and P with the aim of improving its adsorption and exchange properties. The incorporation of the groups selective towards certain metal ions into a modified silica gel phases is also possible. The immobilization of chelating groups containing donor atoms on the silica gel surface can occur *via* chemical bond formation between organic modifiers and silica gel phases or through simple physical adsorption processes. Such a process often incorporates selectivity in synthesized materials.

Several new chemically modified silica gels were synthesized, and applied as normal or selective solid phase extractors for many metal ions. Schiff bases, having multidentate coordination sites, such as those derived from salicylaldehyde are known to readily form stable complexes with transition metal ions.[6,7] The synthesis of the chemically modified silica gel covalently bonded with 4,4′- diaminodiphenylether and 4,4′-diaminodiphenylsulfone-salicylaldehyde Schiff bases has been reported.[6] Their comparative abilities to selectively retain the metals were determined by using batch techniques. The 4,4′-diaminodiphenylether derivative, which was proved to be more reactive, was also used for the preconcentration of Zn^{2+} and Cr^{3+} by column technique. Elution of the sorbed cations in column experimental studies constituted an example of using the chelating material in continuous cycle operation and recovery of the metal ions. The

* Corresponding author: silvia_curteanu@yahoo.com

method was found to be rapid, precise, simple and affordable.[6] Mercapto-functional silica xerogels were also used for removing metal contaminants from water solution.

The sol–gel process is a largely used tool for preparing silica with the desired functionality. The sol–gel synthesis can be performed at room temperature and allows obtaining of pure materials with uniform distribution of the functional groups. Furthermore, the properties of the resulting xerogel can be controlled by the choice of the synthetic conditions.[1]

By using such a procedure, in this paper we prepared silica xerogels with different contents of 2,4- dihydroxy-benzaldehydimine groups linked through a propylene moiety to silica.

The use of neural networks has become increasingly recommended for applications where the mechanistic description of the interdependence between variables is either unknown or very complex. They represent one of the most popular tools with applications in areas such as pattern recognition, classification, process control, optimization.[8,9] Generally, experimental and industrial practices use two types of models: mechanistic models based on the physical and chemical features and data-based empirical models. Each of these categories presents advantages and disadvantages. The mechanistic models have the advantage to be valid upon a large area of operating conditions and reflect the process phenomenology. For this reason, whenever it is possible, the main recommendation is to use the physical and chemical knowledge for the process. Regarding the disadvantages of these models, one can observe the difficulties concerning the specificity of the process and the problems in designing a mathematical model. The difficulties regarding the chemical process refer to many aspects as follow: the absence of on-line testing (measurements), the considerable delays at testing, the possibility of many answers determined by the different operating conditions. Concerning the design of the phenomenological mathematical model, several aspects can be mentioned: the complexity of reactions' mechanisms or the fact that the phenomenology of the processes is insufficiently known, the large number of chemical species involved into the system, the large number of model equations and the special methods needed for giving the solutions.

A review of studies on neural network modeling or control allows the observation of some advantages: parallel organization permits solutions to problems where multiple constraints must be satisfied simultaneously, the rules are rather implicit than explicit, the model can be easily developed and provides accurate results in the most cases.[10] On the other hand, the disadvantages seem to rely on the necessity to obtain a perfect neural network with the experimental or operational history data. Neural networks need large amount of good quality data for the training, which is normally difficult to obtain in practice. Data sparsity, 'overfitting' and poor generalization are other problems encountered by researchers when utilizing the basic neural network alone.[9] A special attention should paid to having an uniform distribution of data throughout the design space.[11] If neural network models properly trained and validated, they can be used to accurately predict the process behavior, hence, leading to improve process optimization and control performance.[12]

Our group has a series of approaches related to the modeling of different processes including polysiloxanes. For instance, a feed-forward neural network was developed to model the dependence of the monomer conversion and copolymer composition on the working conditions (temperature, reaction time, amount of catalyst and initial composition of monomers), in the synthesis of statistical dimethyl-methylvinylsiloxane copolymers.[13,14] In the heterogeneous cationic polymerization of octamethylcyclotetra-siloxane, the dependence between the main parameters of the process (conversion and molecular weight) and working conditions (temperature, reaction time, amounts of catalyst and co-catalyst) was modeled by using feed-forward neural networks.[15] Neural networks modeling was also applied for the kinetic of the equilibrium anionic polymerization of cyclic siloxanes in presence of 1,3-bis(aminopropyl)tetramethyldisiloxane. In this case study, the neural models have been used to evaluate the rates of disappearance of octamethyl-cyclotetrasiloxane and aminopropyl disiloxane at different catalyst concentrations. Alternatively, another neural model has been developed to estimate the amount of catalyst, which leads to an imposed final concentration of cyclosiloxane.[16]

In the present paper we prepared silica xerogels with different contents in chelating groups and used them to extract metal ions (Cu^{2+}) from aqueous solution. In order to find the conditions in which a maximum efficiency in the metal

absorption can be reached and to study the evolution in time of the absorbance as a function of others process parameters (saturation time, concentration of the solution, chelating groups content and pH), a series of experiments have been performed in different conditions. The acquired data were used for process modeling based on artificial intelligence instruments. The option of using simple neural networks or stacked neural networks with different architectures to model the wastewater treatment process based on silica xerogels functionalized with hydroxyl-azomethine groups was emphasized.

EXPERIMENTAL

Materials

Copper dichloride, $CuCl_2$ was purchased from Chimopar, Roumania.

Tetraethyl-orthosilicate (TEOS), purchased from Fluka (purity > 98 %, b.p.=163-167 ^oC, $d_4^{20} = 0.933$) was used as received.

3-Aminopropyltriethoxysilane (APT), Fluka, M=221.37, b.p.=213-216, $d_4^{20} = 0.949$.

2,4-Dihydroxybenzaldehyde (AR) was prepared and purified according to procedures described in literature[17] (yield: 33%, m. p. 135-137 °C).

Buffer solutions having pH = 4, 5 and 6 were prepared based on CH_3COOH and NaOH 1N solutions.

Measurements

Fourier transform infrared (FT-IR) spectra were obtained on a Bruker Vertex 70 FT-IR analyzer. Analyses were performed on the films in reflectance mode (ATR), in the 600 - 4000 cm^{-1} range, at room temperature with 2 cm^{-1} resolution and accumulation of 32 scans.

Electronic absorption spectra were measured using SPECORD M42 spectrophotometer with quartz cells of 1 cm thickness in water.

Procedure

Functionalized silica was prepared according to the procedure described in[18]: TEOS and APT were mixed in various ratios to obtain silica with desired functional groups content (Table 1) and stirred for 1 hour at room temperature in presence of dibutyltindilaurate (DBTDL) as a catalyst. The reaction mixture was then poured in a Teflon dish and left at room temperature for the reaction ending. The remained solid was extracted with water and methanol and dried in vacuum at 50^{o}C. The compounds resulted as orange powders.

Table 1

The functional groups content of the silica designed by synthesis

Sample	Functional group content in silica (mol/g)
FS1	4.50×10^{-4}
FS2	8.70×10^{-4}
FS3	13.04×10^{-4}

Metal uptake experiments

A series of solutions differing by the $CuCl_2$ content (5.5 and 11 g/l) and pH (4, 5 and 6) were prepared by using the proper buffer solutions. UV-VIS spectrum was registered for each solution. A certain amount of functional silica gel bringing a certain amount of chelating groups was introduced in such a solution and UV-VIS spectrum was registered at certain time intervals until the maximum absorbance remains constant. We named this time as saturation time, t_s.

NEURAL NETWORK MODELING

Neural networks have been increasingly used in the chemical process field, especially for dealing with some complex nonlinear processes where understanding is limited.[19,20] The quality of a neural model or its fitness-for-purpose depends on the amount and representativeness of the available training data, the network training method and the network type related to the process to be modeled.

In this paper, a consistent set of data, which covers uniformly the experimental domain, was taken into account in order to assure a good neural model. Different architectures were also considered in the modeling methodologies. Three types of neural networks that have as common characteristic the supervised learning control: Multilayer Perceptrons (MLP), Generalized Feed-Forward networks (GFF), and Jordan and Elman networks (JE), have been tested in order to put in evidence the complexity and the performance of simple and stacked neural networks.

The most common neural network architecture is the multilayer feed-forward neural network (often called multi-layer perceptron, MLP). Many papers apply a multilayered, feed-forward, fully connected network of perceptions because the simplicity of its theory, ease of programming and good results often obtained.[11] Multilayer perceptron can approximate any input/output map, but it trains

slowly and requires lots of training data. Generalized feed-forward network is a generalization of MLP, the difference being into the connections that can jump over one or more layers. In practice, generalized feed-forward networks often solve the problem much more efficiently. Jordan and Elman network supplies the multilayer perceptron with context units, which are processing elements that remember past activity. Context units provide the network with the ability to extract temporal information from the data.

In general, there is no assurance that any individual models have extracted all relevant information from the data set. Many researchers have shown that simply combining multiple neural networks could generate more accurate predictions than using any one of the individual networks alone.[21-25]

The idea of combining neural network models is based on the premise that different neural networks capture different aspects of process behavior. Moreover, aggregating this information should reduce uncertainty and provide more accurate predictions. It is always possible that a stacked network could outperform a single best trained network for the following reasons: 1) Optimal network architecture cannot always be guaranteed. 2) The optimization of network parameters is a problem with many local minima. Even for a given architecture, final network parameters can differ between one run of the algorithm and another. For example, variations in the initial starting conditions for network training can lead to different solutions for the network parameters. Therefore, the different networks might have independent errors on different subsets of the input space. 3) Different activation functions and learning algorithms can also lead to different generalization characteristics, and no particular activation function or learning algorithm is uniformly the best. 4) Convergence criteria used for network training can lead to different solutions for a given network architecture. Therefore, combining existing neural networks can provide a practical approach to developing a better overall model for prediction, rather than searching for a single best-trained model.[26]

In a stacked generalization model, three additional parameters need to be fixed: the method used to combine the parallel models, the number of data that are randomly removed from the initial data to form the individual learning data subsets and the number of levels in the stack.[27]

In this paper, the influence of the main parameters (functional group content in silica, $CuCl_2$ solution concentration, and pH) on the UV-VIS absorbance as a pointer of the Cu^{2+} amount remained in solution was evaluated using individual and stacked neural networks. A special type of stack – software stack – was designed and tested: individual neural networks were developed for each sample and a software assembling procedure was used to obtain and manipulate a single neural model. In other words, the output of the whole model is composed of the outputs of the individual networks included in the assembly, and no summation is performed on these outputs.

RESULTS AND DISCUSSION

To prepare the functionalized silica xerogels, tetraethyl-orthosilicate (TEOS) and 3-aminopropyl-triethoxysilane (APT) were employed in different ratios as precursors, to obtain the functionalized silica (Table 1). The aminopropyl groups were reacted in situ with carbonyl groups from 2,4-dihydroxy-benzaldehyde (added in slight excess in the reaction mixture) resulting in azomethine ones. IR spectra emphasized the presence of this function at about 1650 cm^{-1} (Figure 1), besides the bands specific for the other groups: 1253 (δ Si-CH_3 sym.), 1047 (Si-O-Si asym.) and 795 (CH_3 rocking asym. Si-C).

The azomethine and hydroxyl groups, being in proper positions to close the chelate ring, can co-ordinate the metal ions from an aqueous solution according to Scheme 1. The Si-OH groups on the surface permit the silica wetability and the metal ion access to the chelating groups.

A series of solutions having different $CuCl_2$ contents (5.5 or 11 g/l) and pH (4, 5 or 6) (Table 2) were prepared. UV spectrum for each solution was registered, maximum absorbance being A_0. A determined amount of functional silica gel was introduced in such a solution and UV spectrum was registered at certain time intervals. As a result of the copper complexation by silica, the concentration of the copper ions in solution decreases and, by consequence, the UV-VIS absorbance also decreases. The maximum absorbance at every time (A_t) was determined from the spectra as an indirect measure of the metal content.

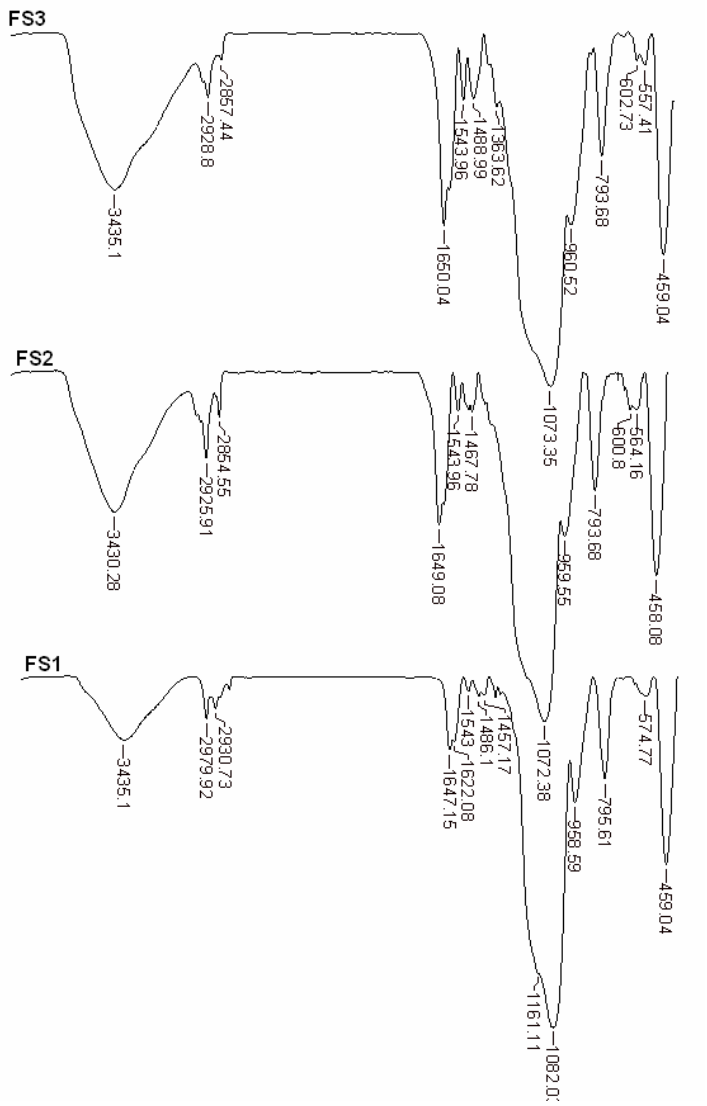
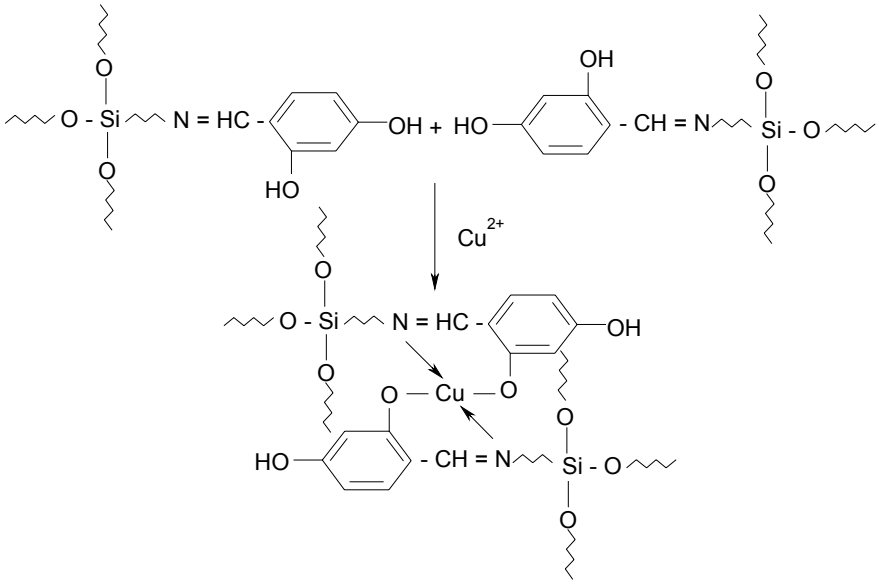


Fig. 1 – FTIR spectra of the prepared silica xerogels.



Scheme 1 – The complexing of the copper ions.

Table 2

Saturation time and relative absorbance at saturation

Sample	$CuCl_2$ solution concentration, g/l	pH	Complexing groups content in silica, mol/g	Saturation time (t_s), minute	Relative absorbance at saturation, A_{rs}
G_1	5.5	6	4.35×10^{-4}	1440	0.9093
G_2	11	6	4.35×10^{-4}	1440	0.9505
G_3	5.5	4	4.35×10^{-4}	255	0.9318
G_4	11	5	4.35×10^{-4}	120	0.9605
G_5	11	6	13.04×10^{-4}	1440	0.8602
G_6	11	6	8.70×10^{-4}	240	0.9349
G_7	5.5	6	13.04×10^{-4}	1440	0.9663
G_8	11	4	13.04×10^{-4}	1241	0.9968

For a better comparability of the results, eliminating the influence of the ratio between sorbent and solution amounts, two other parameters derived from the initial (A_0) and actual (A_t) absorbances were considered in this paper as follows:

Relative absorbance, $A_r = A_t/A_0$,

Percent relative absorbance decreasing, $A_{rd} = 100(A_0\text{-}At)/A_0$.

The time at which the absorbance remained constant was considered the saturation time (t_s). These time values, as well as the corresponding relative absorbance, A_{rs} ($A_{rs} = A_{ts}/A_0$), for the performed experiments are presented in Table 2, and the detailing of every process in time can be found in Table 3.

The main steps in neural network modeling are: collecting the experimental data sets, splitting the data in two parts, training and validation data sets, developing the neural network topology (training phase) and checking the generalization capacity of the neural model (validation phase). From the whole set of data (90 in all), about 15 % was kept for validation phase.

Table 3

The decrease of the absorbance in time for different samples

Sample	$CuCl_2$ solution concentration g/l	pH	Complexing groups content in silica, mol/g	Time, minute	Absorbance, A_t		Relative absorbance, (A_r)/relative absorbance decreasing (A_{rd})
G_1	5.5	6	4.35×10^{-4}	0	(A_0)	2.558	1.0000
				15		2.472	0.9664/3.36
				30		2.438	0.9531/4.69
				60		2.414	0.9437/5.63
				105		2.383	0.9316/6.84
				255		2.344	0.9163/8.37
				1440		2.326	0.9093/9.07
				1740		2.344	0.9163/8.36
				2880		2.366	0.9249/7.51
G_2	11	6	4.35×10^{-4}	0	(A_0)	1.2087	1.0000
				15		1.1935	0.9874/1.25
				30		1.1794	0.9757/2.42
				45		1.1669	0.9654/3.45
				1440		1.1489	0.9505/4.95
				4320		1.1663	0.9649/351
G_3	5.5	4	4.35×10^{-4}	0	(A_0)	1.0925	1.0000
				30		1.0535	0.9643/3.57
				75		1.0335	0.9460/5.40
				135		1.0305	0.9432/5.67
				195		1.0198	0.9335/6.65
				255		1.0180	0.9318/6.82
G_4	11	5	4.35×10^{-4}	0	(A_0)	2.404	1.0000
				15		2.383	0.9913/0.87
				60		2.344	0.9750/2.49
				120		2.309	0.9605/3.95
				240		2.309	0.9605/3.95
G_5	11	6	13.04×10^{-4}	0	(A_0)	2.619	1.0000
				60		2.498	0.9538/4.62
				120		2.408	0.9194/8.20

Sample	$CuCl_2$ solution concentration g/l	pH	Complexing groups content in silica, mol/g	Time, minute	Absorbance, A_t	Relative absorbance, (A_r)/relative absorbance decreasing (A_{rd})
				180	2.333	0.8908/10.92
				240	2.318	0.8851/11.49
				300	2.309	0.8816/11.83
				1440	2.253	0.8602/13.97
G_6	11	6	8.70×10^{-4}	0	(A_0) 3.010	1.0000
				60	2.975	0.9884
				120	2.903	0.9645/2.42
				180	2.846	0.9465/4.33
				240	2.814	0.9349/5.41
G_7	5.5	6	13.04×10^{-4}	0	(A_0) 1.1341	1.0000
				60	1.1223	0.9896/1.0404
				90	1.10532	0.9746/2.5395
				120	1.10318	0.9727/2.7246
				1440	1.0959	0.9663/3.3683
G_8	11	4	13.04×10^{-4}	0	(A_0) 2.210	1.0000
				1080	2.154	0.9747/2.5339
				1241	2.203	0.9968/0.3167

One important problem in the development of simple neural network models consists in the determination of the network architecture, *i.e.* the number of hidden layers and the number of neurons in each hidden layer. Firstly, potentially good topologies must be identified. Nevertheless, no good theory or rule accompanies the neural network topology that should be used, and trial-and-error is still required. This is done by testing several topologies and comparing the prediction errors. Lower errors indicate potentially good architectures, *i.e.* neural network topologies with chances to train well and to output good results.

Based on many experimental datasets, selectively exposed in Table 3, the first approach focuses on the influence of the four parameters (functional group content in silica, $CuCl_2$ solution concentration, pH, and time) upon the absorbance (A_t). Three simple neural network types were tested; MLP, GFF and JE neural networks with different topologies were trained and their performances were evaluated based on mean squared error (*MSE*), correlation between experimental data and network predictions (r) and percent error (E_p). Table 4 presents the best network for each tested type.

The best network type found by this trial and error method was MLP(4:6:1) – multilayer perceptron with 4 neurons in the input layer corresponding to the 4 input variables, one hidden layer with 6 neurons and 1 neuron in output layer for the absorbance, A_t. The transfer function applied to the hidden and output layers was TanhAxon type.

In the training phase, an average relative error of 0.9988 and a correlation between experimental and predicted data of 0.9995 were registered. Relative errors were calculated using the following formula:

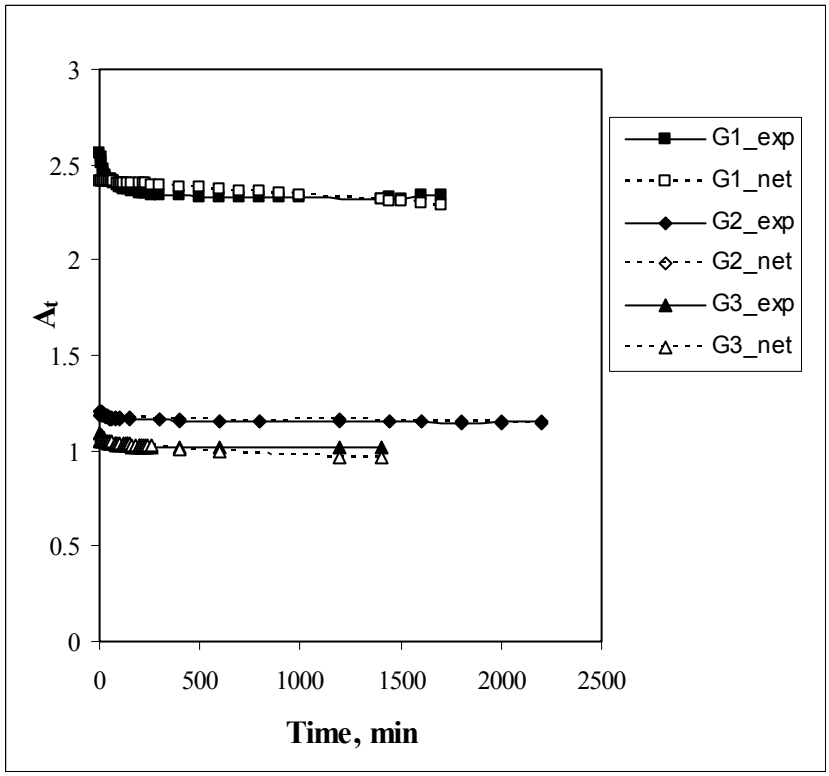
$$E_r = \frac{A_{t\_exp} - A_{t\_net}}{A_{t\_exp}} \cdot 100 \qquad (1)$$

where A_{t_exp} and A_{t_net} denote experimental and network values for the absorbance. Figure 2 (a and b) presents the training stage for different compounds belonging to the actual case study. The Figure 2 contains more experimental data than those presented in Table 3 where only a selection is given.

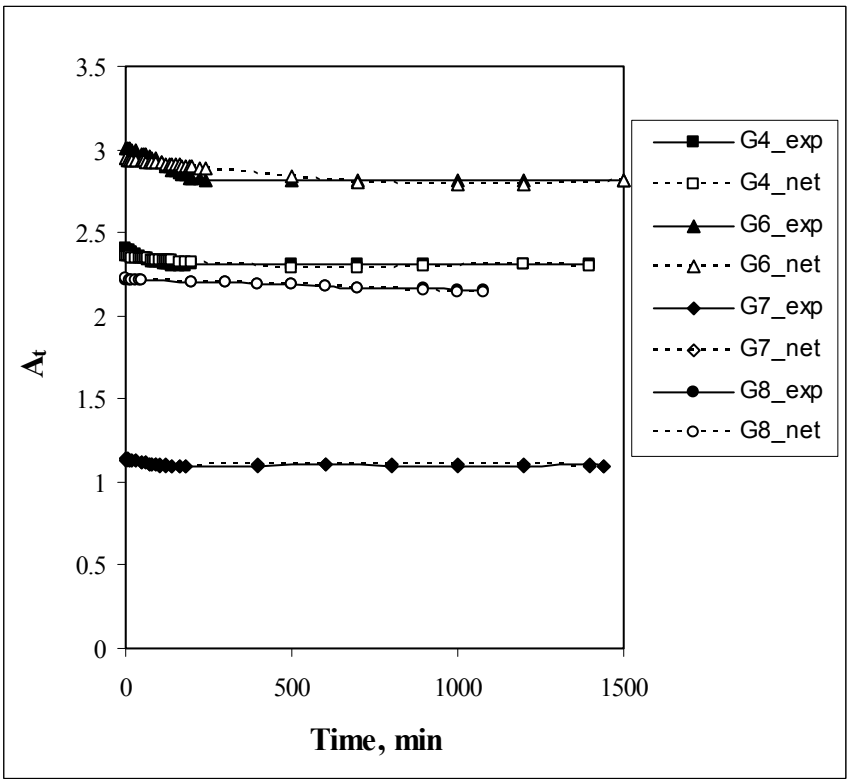
Table 4

The best neural network (for each type), individually used, developed for the modeling of absorbance, A_t

Network type	MSE	r	E_p
MLP(4:6:1)	0.000687	0.99906	0.97329
GFF(4:6:1)	0.000283	0.999613	0.791913
JE(4:6:1)	0.000164	0.999776	0.445258



(a)



(b)

Fig. 2 – The training stage for MLP(4:6:1) used for absorbance (A_t) modeling.

A key issue in neural network based process modeling is the robustness or generalization capability of the developed models, *i.e.* how well the model performs on unseen data. Thus, a serious examination of the accuracy of the neural network results requires the comparison with experimental

data, which were not used in the training phase (previously unseen data). For the validation stage, the average relative error of about 1.0664 % and the correlation value of 0.9989 reflect a very high capacity of generalization (Figure 3).

The results obtained with different types of neural networks are presented comparatively in Table 5, where E_t, r_t, E_v and r_v represent the errors and the correlations corresponding to the training and validation stages, respectively.

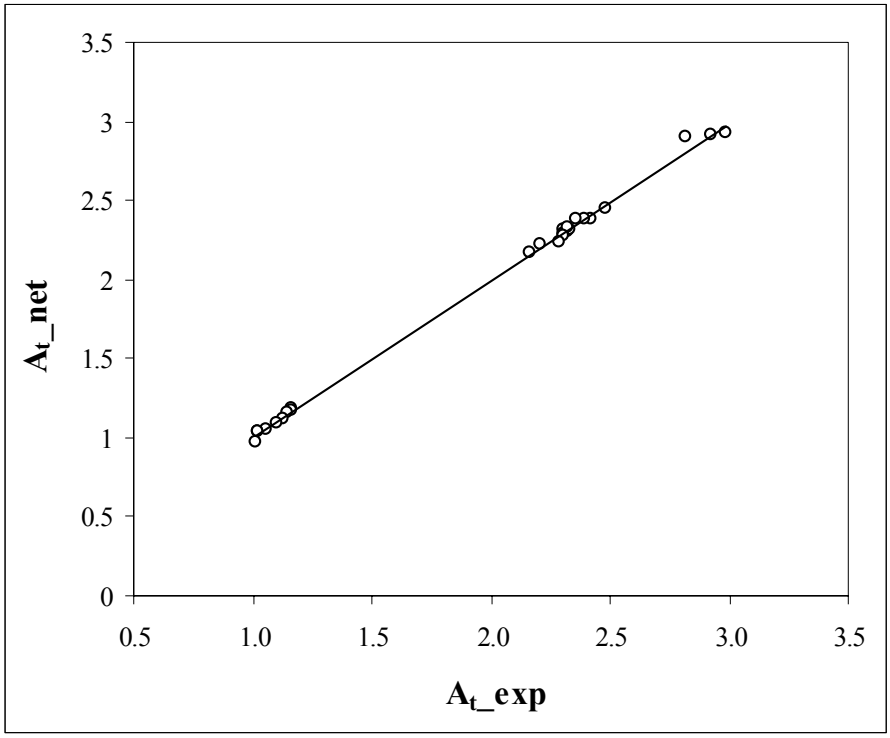


Fig. 3 – The validation stage for MLP(4:6:1) used for absorbance (A_t) modeling.

Table 5

The results obtained with individual neural networks for the absorbance (A_t) modeling

Network type	E_t %	r_t	E_v %	r_v
MLP(4:6:1)	0.973251	0.99906	1.082215	0.99927
GFF(4:6:1)	0.770185	0.9996401	1.255206	0.999305
JEN(4:6:1)	0.445542	0.999776	5.93776	0.987142

Diversity of the individual models is a key issue in the stacked generalization and it is not easy to achieve because all models are trained to do essentially similar tasks. Another variant of modeling applied in our work is based on using individual neural networks for each set of experimental data, corresponding to each sample ($G_1 \div G_8$) and then aggregating them into a stack. The problem becomes more simply because each sample is characterized by a series of initial conditions and there is only one parameter which changes when the absorbance decreases – reaction time. This type a stack is a software assembling procedure whose output is represented by the eight individual outputs. This methodology, including different neural network types, was designed with the goal to improve the generalization capability of the neural models.

The trial and error method provides the following networks: GFF(1:10:1) for G_1, JE(1:6:1) for G_2, MLP(1:10:1) for G_3, MLP(1:4:1) for G_4, GFF(1:4:1) for G_5, GFF(1:4:1) for G_6, MLP(1:10:1) for G_7 and JE(1:4:1) for G_8. Table 6 contains these individual networks with training and validation errors and correlations. A lower validation error of maximum 2.6 % was obtained for each model taken into account.

Table 6

The performances of the individual models aggregated into a software stack

Sample	Model	Training		Validation	
		E_t %	r_t	E_v %	r_v
G_1	**GFF(1:10:1)**	0.17093	0.996039	0.10586	0.998207
G_2	**JE(1:6:1)**	0.1298	0.9946	2.60999	0.9946
G_3	**MLP(1:10:1)**	0.143059	0.9954	0.10762	0.99727
G_4	**MLP(1:4:1)**	0.0858	0.996907	0.11591	0.99799
G_5	**GFF(1:4:1)**	0.22572	0.99804	1.24927	0.99675
G_6	**GFF(1:4:1)**	0.14075	0.99808	0.12564	0.9998
G_7	**MLP(1:10:1)**	0.17422	0.98486	0.31782	1
G_8	**JE(1:4:1)**	0.0.0327	0.99907	1.45834	0.99907

Figures 4 and 5 reflect a very good correlation between the experimental and predicted results, both in training and, more important, in validation phase.

The decrease of absorbance, A_{rd}, was considered as function of time, as an alternative approach to the previous case. The eight individual neural networks were: GFF(1:12:4:1) for G_1, MLP(1:12:4:1) for G_2, MLP(1:4:1) for G_3, MLP(1:4:1) for G_4, GFF(1:12:4:1) for G_5, MLP(1:12:4:1) for G_6, JE(1:12:4:1) for G_7 and MLP(1:10:1) for G_8. Figure 6 shows the stack obtained with these networks and the errors and correlations in the training and validation stages. The errors in validation phase are placed between 0.8 and 3.3 %, that means an acceptable domain of values.

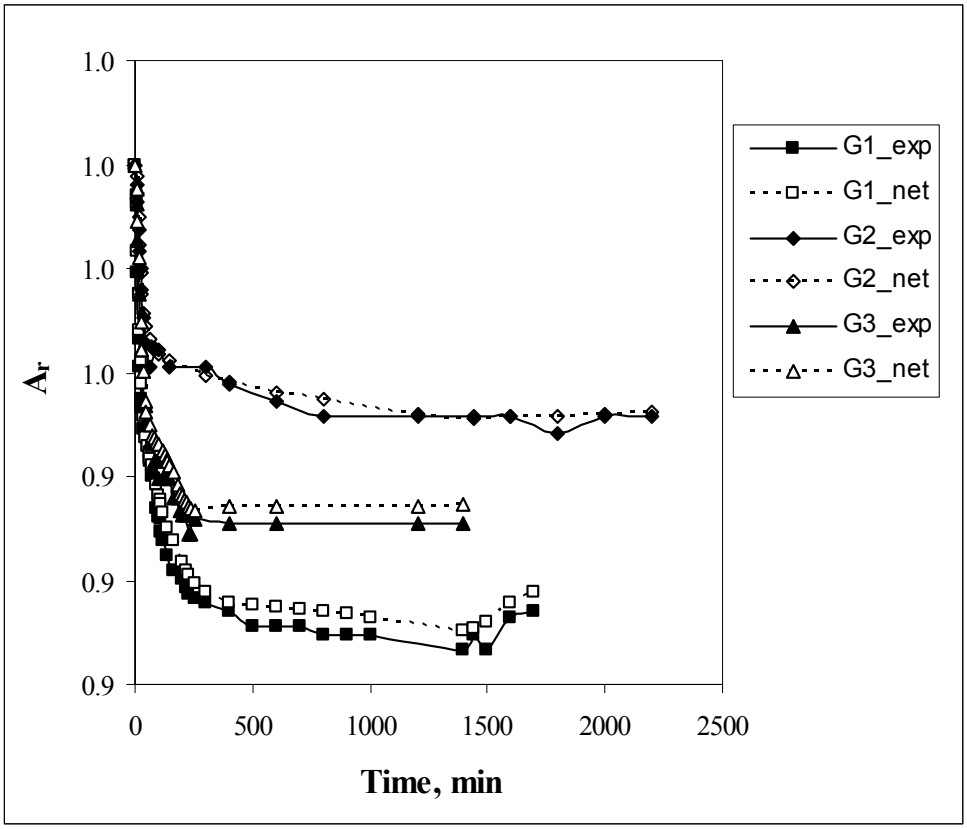


Fig. 4 – Training stage of the stacked neural networks developed for relative absorbance (A_r) modeling.

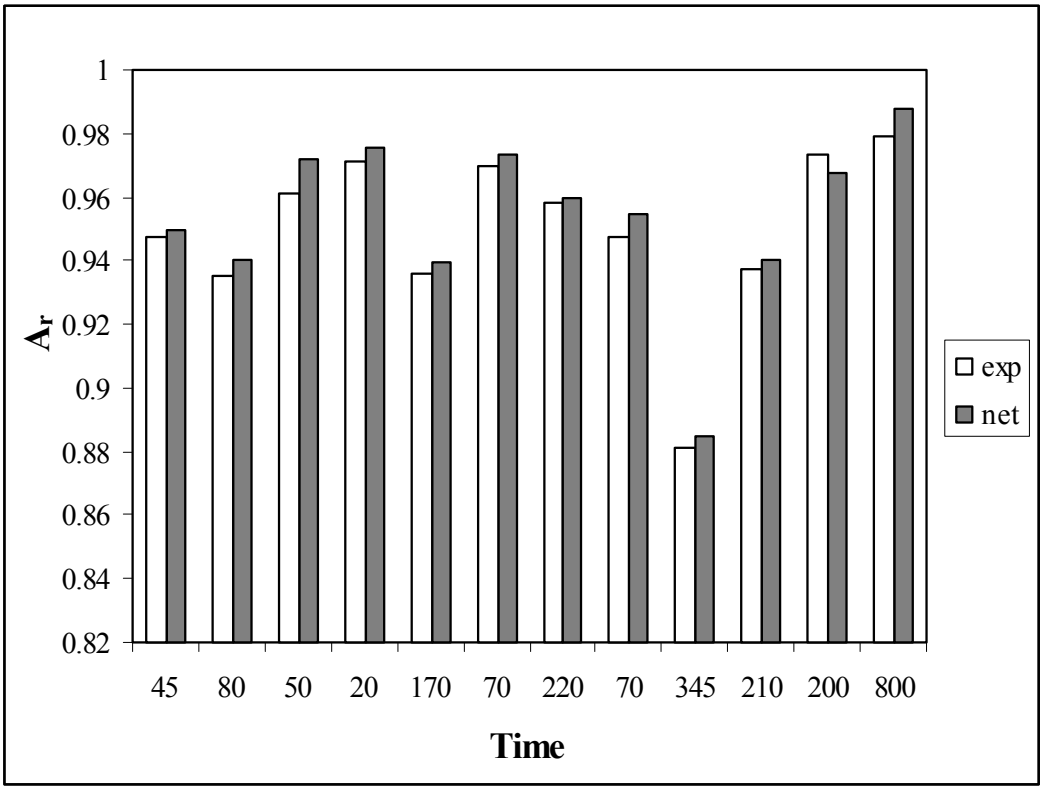


Fig. 5 – Validation stage of the stacked neural networks developed for relative absorbance (A_r) modeling.

Figure 7 shows the validation phase for the stacked neural networks designed to model the decreasing of A_{rd} in time.

Consequently, the neural network methodology provides useful and credible information for experimental practice, which can be a guide for future experiments or can substitute experiments that are time and material consuming.

CONCLUSIONS

The silica functionalized with dihydroxy-azomethine groups was proved to retain copper from aqueous solution, with equilibrium reaching between a few hours and 48 hours, depending on the silica functionality as well as the working environment (metal ion concentration or pH of the aqueous solution). The kinetic data concerning metal sorption were acquired by UV-VIS spectral analysis.

Neural networks proved to be powerful tools for modeling of the complex and nonlinear chemical systems. In this paper, the metal absorption efficiency of functionalized silica xerogels was estimated experimentally, as well as by simulation. Different modeling strategies based on simple and stacked neural networks were developed and tested for the actual case study, in order to obtain a useful model for experimental practice.

It was proved that stacked neural networks designed as a software assembly represent models with accurate capacity of generalization.

Acknoledgements: This work was done by financial support provided by Roumanian Ministry of Education and Research through National Program for Research, Development and Innovation II, project 71-006/2007.

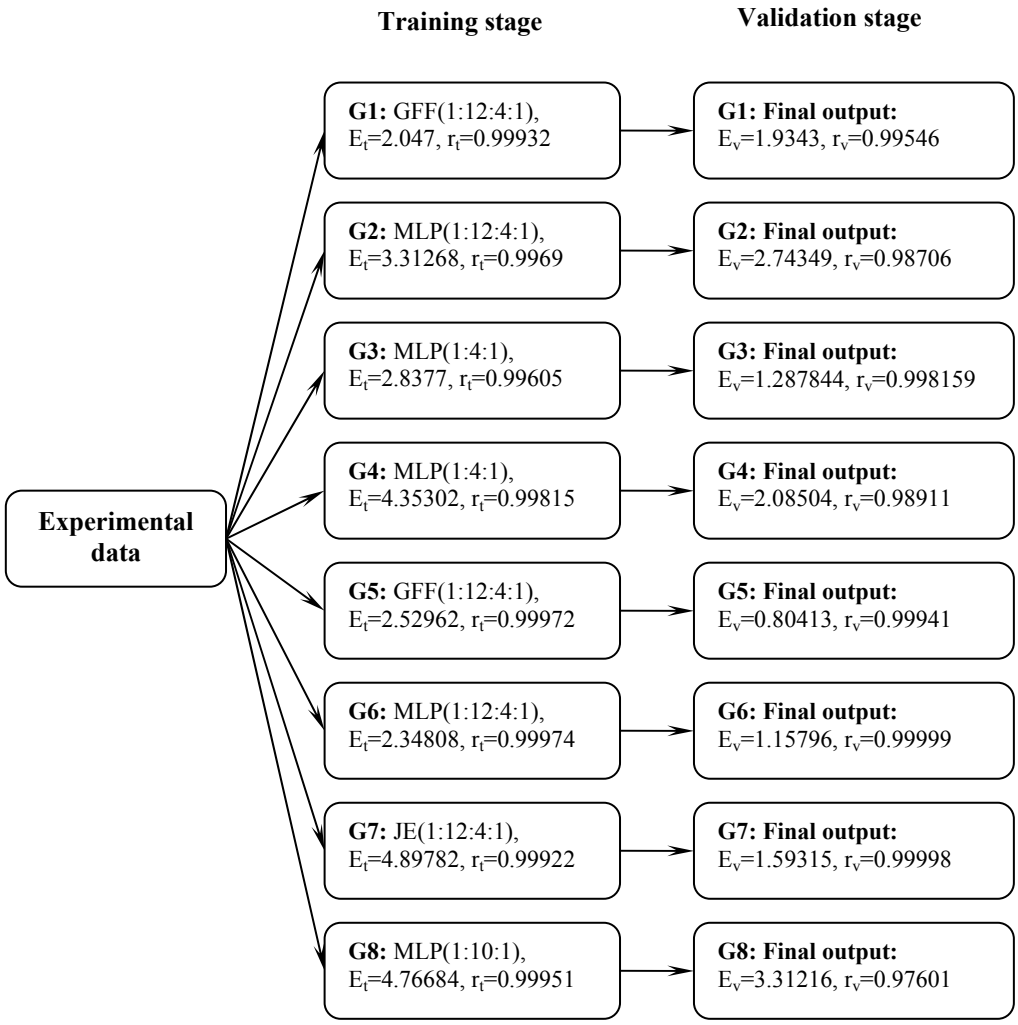


Fig. 6 – Stacked neural networks for modeling the decreasing of the absorbance, A_{rd}, with time.

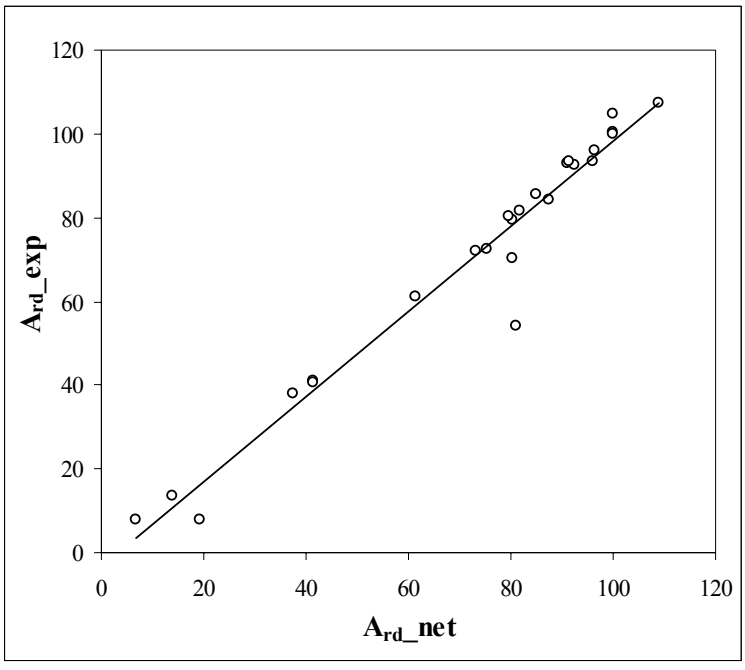


Fig. 7 – Validation phase for the stacked neural networks which model the decreasing of the absorbance (A_{rd}).

REFERENCES

1. M.C. Brasil, E.V. Benvenutti, J.R. Gregorio and A.E. Gerbase, *React. Funct. Polym.*, **2005,** *63*, 135–141.
2. B.W. Atkinson, F. Bux and H.C. Kasan, *Water SA*, **1998**, *24*, 129-135.
3. A. Duta, D. Perniu, E. Helerea, *Journal of Environment Protection and Ecology*, **2002**, *3*, 929-934.
4. M.S. Rao, I.S. Dubenko, S. Roy, N. Ali and B.C. Dave, *J. Am. Chem. Soc.*, **2001**, *123*, 1511-1512.
5. D.C. Lai, B. Dunn and J.I. Zink, *Inorg. Chem*., **1996**, *35*, 2152-2154.
6. R.K. Dey, U. Jha, A.C. Singh, S. Samal and A.R. Ray, *Anal. Sci.*, **2006**, *22*, 8.
7. M.K. Choi, H.L. Kim and D.H. Suh, *Journal of Applied Polymer Science*, **2006**, *101*, 1228-1233.
8. Z. Xiong and J. Zhang, *J. Proc. Control*, **2005**, *15*, 11-21.
9. C.W. Ng and M.A. Hussain, *Chem. Eng. Proc.*, **2004**, *43*, 559.
10. E. Fieslerand and R. Beale, "Handbook of Neural Computation", Oxford, IOP Publishing Ltd and Oxford University Press. 1997.
11. F.A.N. Fernandes and L.M.F. Lona, *Brazilian Journal of Chemical Engineering*, **2005**, *22*, 323.
12. Z. Xiong and J. Zhang, *Chemical engineering and processing*, **2005**, *44*, 477-484.
13. S. Curteanu and M. Cazacu, Journal of Macromolecular Science, **2008**, *45*, 21-34.
14. S. Curteanu and M. Cazacu, *Lecture Series on Computer and Computational Sciences*, **2006**, *6*, 120.
15. S. Curteanu, M. Cazacu, *Revue Roumaine de Chimie*, RRC 76/2008, in press.
16. A. Nistor, C.G. Piuleac, M. Cazacu and S. Curteanu, *Materiale Plastice*, **2008**, *45*, 67-73.
17. M. Bisagni, N.P. Bun-Hoiet and R. Roger, *J. Chem. Soc.*, **1955**, 3688-3693.
18. M. Cazacu, A. Vlad, A. Airinei and M. Alexandru, *Polym Int.*, **2008**, *57*, 1067-1074.
19. A.B.Bulsari, "Neural Networks for Chemical Engineers", Amsterdam, Computer-Aided Chemical Engineering, Elsevier, 1995.
20. I.M. Mujtab and M.A. Hussain, "Application of Neural Networks And Other Learning Technologies in Process Engineering", London, Imperial College Press, 2001.
21. J. Zhang, *Chemical Engineering Science*, **2008**, *63*, 1273-1281.
22. S.Q. Wang, J. Yang and K.C. Chou, *Journal of Theoretical Biology*, **2006**, *242*, 941-946.
23. E. Alpaydin, Multiple Networks for Function Learning. *Proceedings of the 1993 IEEE International Conference on Neural Networks*, Piscataway, NJ, IEEE Press, 1993.
24. L.K. Hansen and P. Salamon, *IEEE Trans. Pattern Anal. Mach. Intell.*, **1990**, *12*, 993.
25. M.P. Perrone and L.N. Copper, "When Networks Disagree: Ensemble Methods for Hybrid Neural Networks. Neural Networks for Speech and Image Processing", Chapman & Hall, London, 1993.
26. Y. Tian, J. Zhang and J. Morris, *Ind. Eng. Chem. Res.*, **2001**, *40*, 4525-435.
27. R. Lanouette, J. Thibault and J.L. Valade, *Computers and Chemical Engineering*, **1999**, *23*, 1167-1176.