Presented at 3rd International Symposium on Environmental Management, SEM – Towards Sustainable Technologies (SEM2011) at University of Zagreb, Faculty of Chemical Engineering and Technology, 26–28 September, 2011, Zagreb, Croatia

Influence of Ultrasonic and Microwave Irradiation on Cation Exchange Properties of Clay Material

A. Švarc,[a] Š. Ukić,[a] S. Radojević Lacković,[b] B. Đuričić,[a] M. Novak,[a] and T. Bolanča[a,*]

[a] University of Zagreb, Faculty of Chemical Engineering and Technology, Marulićev trg 19, 10000 Zagreb, Croatia
[b] Vinkovački vodovod i kanalizacija d.o.o., Dragutina Žanića-Karle 47a, 32100 Vinkovci, Croatia

Original scientific paper
Received: September 6, 2012
Accepted: February 1, 2013

This study deals with optimization of the clay activation process using artificial neural network models and multi-objective optimization function. Different artificial neural network models were used for description of the relation between clay sorption capacity and the activation treatment process (power and time of clay exposure to ultrasonic and/or microwave irradiation). Two methodologies (feed-forward and cascade-forward) in combination with five different training algorithms (random order incremental training with learning functions, resilient backpropagation, one-step secant backpropagation, Levenberg-Marquardt backpropagation, Bayesian regularization backpropagation) were applied in order to obtain an optimal artificial neural network model. The optimal artificial neural network model showed good predictive ability (relative error 6.02 % based on external validation data set). In-house developed multi-objective criteria function was used in combination with the developed artificial neural network model and calculated optimal activation was determined (5 minutes of ultrasonic 120 W and microwave 60 W treatment) increasing the sorption capacity by 15 %.

Key words:
Clay activation, multi-objective optimization, artificial neural networks

Introduction

In recent years, the development of pollution-control technology has significantly increased due to highly expressed public environmental awareness of the pollution impact, on both human health and Earth's ecosystem. There are few approaches for immobilization of numerous hazardous substances, but ceramic immobilization and disposal technologies are considered environmentally acceptable and the most versatile for rendering hazardous waste inert, because they destroy organic matter, immobilize regulated heavy metals in a stable matrix, and are able to convert complex chemical compositions into useful materials with the potential of market exploitation.[1–7] Clay, as the main component of ceramic, has good sorption characteristics that make it a suitable immobilization material; that is why not only ceramic materials are used for heavy metal immobilization but clay-based materials as well.[3–11] Additional exposure of clay to ultrasonic and microwave irradiation might increase the exchange capacity and thus improve the immobilization process.[12–22] However, the energy consumption can make the activation process economically infeasible. It is therefore essential to optimize the activation process by maximizing the sorption capacity and minimizing the energy consumption. This makes a multi-factorial non-linear optimization problem. Therefore, for this optimization, a superior optimization tool, such as artificial neural networks (ANNs), are desirable. ANNs are applied well in solving problems of classification and prediction; all the problems in which there are complex, nonlinear input and output relations.[23,24] To

*Corresponding author: T. Bolanča,
e-mail: Tomislav.Bolanca@fkit.hr

the authors' knowledge, no published literature can be found dealing with the application of artificial neural networks for multi-objective optimization of clay activation.

The aim of this work was to develop an artificial neural network sorption model and apply it in combination with in-house developed multi-objective criteria function for ultrasonic and microwave clay treatment process optimization in order to increase sorption capacity and reduce energy and treatment time.

Theory

ANNs have the ability to identify underlying highly complex relationships from input-output data, but they are universal function approximators that offer no audit trail from which a result can be explained. ANN models are obtained by training, i.e., the network is repeatedly presented with input/output pairs that have to be correlated. Although the training procedure can be quite time-consuming, once trained, the network predicts an answer almost instantaneously.[25] Of all the available types of artificial neural networks, multilayered perceptrons (MLPs) are the most commonly used.[26] MLP consists of several layers of neurons. The first layer is connected to the inputs, and the last one gives the network output. Between the input and output layer, there are one or more hidden layers.[27] The neurons from different layers are connected, while no connections between same-layer neurons are allowed. Each connection has its weight. Determination of optimal connection weights is called the network training process.

The signal propagates through MLP network in a forward direction, using feed-forward (FF) or cascade-forward (CF) methodology. In case of FF networks, each neuron from the subsequent layer obtains information from the previous layer only. The information is the sum of weighted signals from previous layer neurons. After obtaining the information, the neuron activates it and transfers it to all next layer neurons. CF networks are similar to feed-forward networks, with the exception that subsequent-layer neurons obtain information from all previous layers.[28]

MLPs have been applied successfully in solving various problems,[29–35] usually trained with a highly popular algorithm known as the backpropagation algorithm (BP).[36–40] The backpropagation algorithm is a gradient descent algorithm in which the network weights are moved along the negative of the gradient of the performance function.[39] There are many different variations of the basic algorithm that are based on heuristics and standard numerical

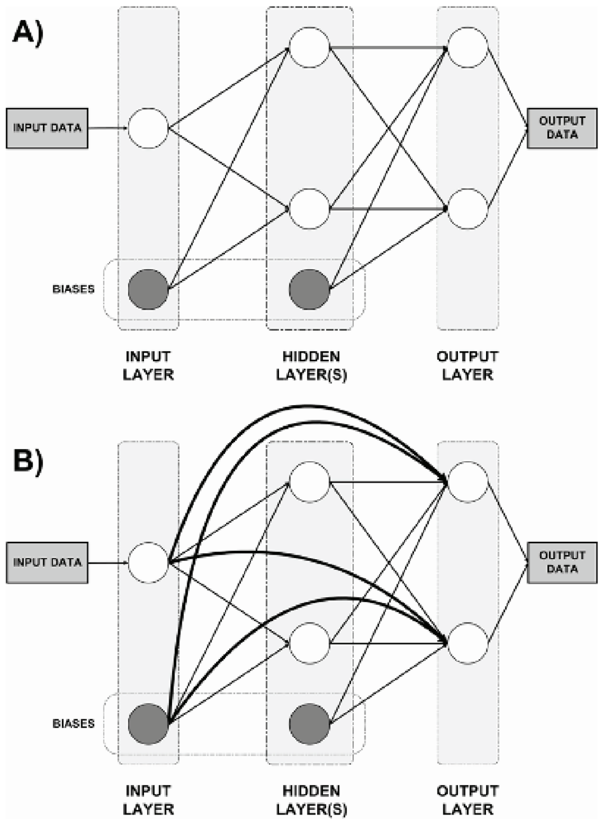


Fig. 1 – *Basic principles of feed-forward and cascade-forward artificial neural networks methodologies*

optimization techniques, such as: random order incremental training with learning functions (RI),[37] resilient backpropagation (RP)[38,41], one-step secant backpropagation (OSS),[41,42] Levenberg-Marquardt backpropagation (LM),[41,43] Bayesian regularization backpropagation (BR),[37,43] *etc.*

Multilayer networks typically use sigmoid transfer functions in hidden layers.[44] Sigmoid functions are characterized by the fact that their gradient increments tend toward zero as the absolute input value gets large. This causes a problem when BP is used to train a multilayer network with sigmoid functions, because the gradient can have a very small magnitude and, therefore, cause small changes in the weights and biases, even though the weights and biases are far from their optimal values. One of the algorithms that eliminates this harmful effect is the RP algorithm. This first-order training algorithm is the fastest algorithm on pattern recognition problems, but does not always perform well on function approximation problems; its performance also degrades as the error goal is reduced.

Furthermore, the BP algorithm is usually time-consuming in the network training phase. The quasi-Newton algorithms (second order) are one of the solutions proposed to overcome this problem.[44] Unfortunately, quasi-Newton methods in general are characterized by quadratic increasing of the storage and computational requirements with increasing of the network size.[45] The OSS algorithm, in relation to other quasi-Newton algorithms, has the requirements that are significantly lower, especially for large scale problems.[46] Nevertheless, due to the memory-less approach, the amount of second order information in OSS is also reduced (such as in comparison with probably the most popular quasi-Newton method: Broyden–Fletcher–Goldfard–Shanno method[43,47]).

However, LM algorithm trains a neural network 10 to 100 times faster than the usual gradient descent backpropagation method. On function approximation problems, for networks that contain up to a few hundred weights, LM algorithm will have the fastest convergence and is more robust (compared to above-mentioned algorithms).[48] This advantage is especially noticeable if very accurate training is required. In many cases, LM algorithm is able to obtain lower mean square errors than any of the previous algorithms. However, as the number of weights in the network increases, the advantage of LM algorithm decreases. In addition, LM algorithm performance is relatively poor on pattern recognition problems.[37]

BR is a network training function that updates the weight and bias values according to LM optimization.[37] It minimizes a combination of squared errors and weights, and then determines the correct combination to produce a network that generalizes well. This good generalization prevents underfitting and overfitting of data, because the training data set can be noisy or imprecise. The training process usually relies on some version of the least squares technique, which ideally should abstract ANN from the noise in the data. However, this feature of ANN depends on how optimal the configuration of the network is in terms of the number of layers, neurons and, ultimately, weights.

Experimental

Activation treatment

Clay samples of approximately 1.1 g each, previously sieved at 60–600 μm, were activated using microwaves (MARS-X, CEM Corporation, USA) and ultrasonic waves (Bandelin SONOREX Digital 10 P Ultrasonic bath, Sigma-Aldrich, St. Louis, USA). Four activation parameters were varied: duration of microwave treatment (0, 5, 10, 20 and 40 minutes), microwave irradiation power (0, 150, 300 and 600 W), duration of ultrasonic treatment (0, 5, 10, 20 and 40 minutes) and ultrasonic irradiation power (0, 240, 600 and 1200 W). The experiment was designed using all combinations of these four parameters' values (400 combinations). When the generally irrelevant and/or repeating experiments (*i.e.* for treatment time 0 minutes it was inessential what power was used; also under 0 W the treatment duration became irrelevant) were eliminated, a final design of 169 experiments was obtained. The clay samples were supplied from clay-pit Dren near the town of Vinkovci, Croatia.

Sorption capacity

Sorption capacity of heavy metal exchange was tested using cobalt ions (cobalt(III)-hexamine chloride, 99 %, Acros Organics, New Jersey, USA).[49,50] 1.0000 g of each sample (pretreated clay) was put in contact with 40 mL of cobalt-hexamine stock solution and then placed into the a thermostatic shaker at 200 rpm and 30 °C for 2 hours. Solutions were filtered through 0.2 μm filter, and analyzed by UV/Vis spectrophotometer (Lambda 35 UV-Vis Spectrometer, Perkin Elmer, Waltham, USA) at 475.3 nm to determine the mass of bound cobalt. Deionized water (18 MΩ, Millipore, Billerica, USA) was used in all experiments.

Artificial neural networks modeling

Each parameter in the data set (power and time of electromagnetic irradiation treatment and mass of bound cobalt) was scaled inside the interval [–1 1], to became sensitive for neural network application. The networks were three-layered with input layer (4 neurons representing activation parameters), one hidden layer, and output layer (1 neuron representing mass of bound cobalt). As an activation function for connecting neurons from input layer with hidden neurons, a logistic sigmoid function was used, whereas a linear transfer function was applied for computation of output activities. Two different neural network methodologies were applied (FF and CF) in combination with two different training approaches (incremental and batch) and five different training algorithms (RI was used for incremental training, RP as a batch first-order training, while OSS, LM and BR were used as batch second-order training algorithms). The number of neurons in the hidden layer was varied (2, 4, 8, 16, and 32) as was the number of data in the training set (4, 8, 16, 32, and 64), which led to the final number of 250 different ANNs. The developed ANN models were validated using external experimental data set.

Multi-objective optimization

In order to determine optimal power and time of exposure to microwave and ultrasonic clay treatment, the multi-objective criteria function (*CR*) was defined

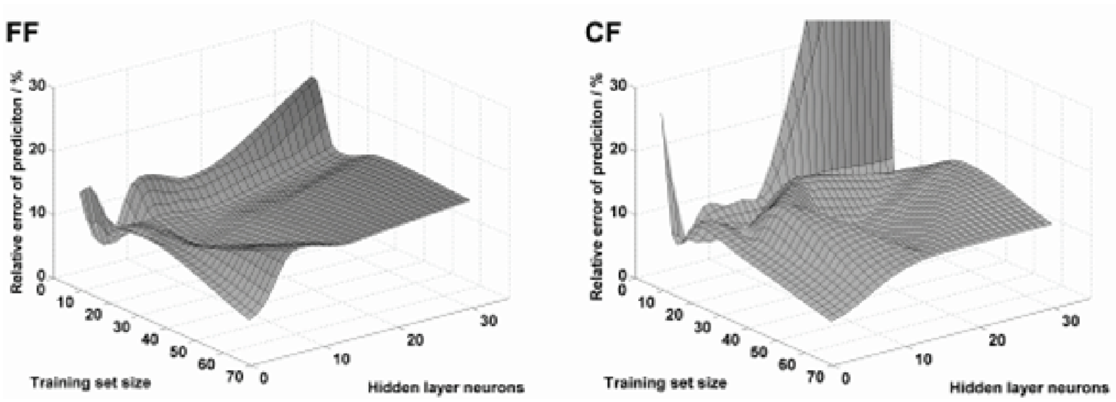
$$CR = \frac{1}{P_{\mathrm{MW}} \cdot t_{\mathrm{MW}}} + \frac{1}{P_{\mathrm{US}} \cdot t_{\mathrm{US}}} + m_{\mathrm{BND}}^{2} \qquad (1)$$

where P is power of microwave (MW) and ultrasonic (US) irradiation, t the irradiation exposure time, and m_{BND} mass of the bound cobalt ion.

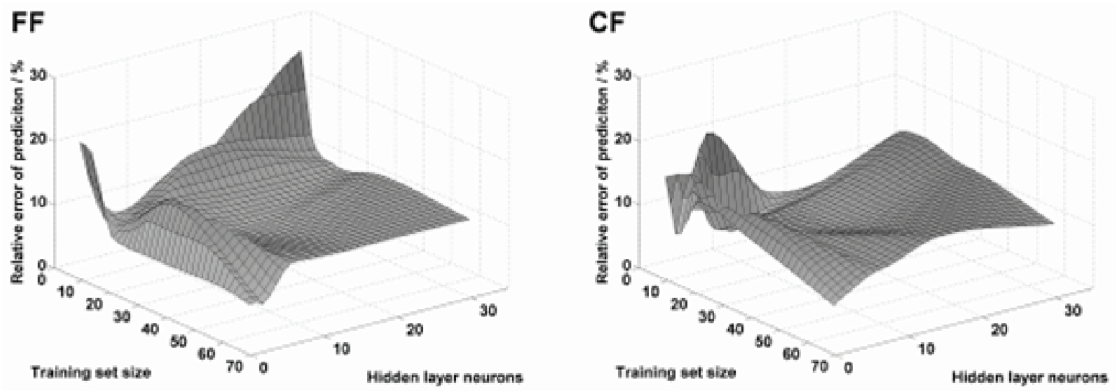
All data processing was carried out in Pentium Dual-Core CPU 2.10 GHz PC, 4.00 GB RAM, Windows 7 Home Premium (Microsoft, USA) OS using Matlab 7.8 (MathWorks, USA) environment.

Results and discussion

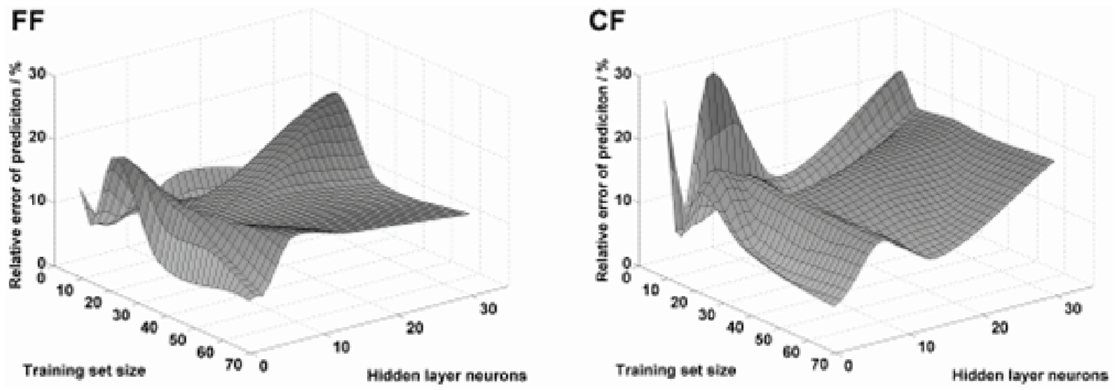
Fig. 2 to 6 illustrate the results of the developed artificial neural networks in prediction of natural clay sorption capacity using ultrasonic and microwave activation. One can observe that the models obtained using incremental type of data presenting (Fig. 2.) in combination with the RI training algorithm achieves lower predictive ability than network obtained using the batch type of data presenting in combination with the RP training algo-



F i g . 2 – *Development of artificial neural network activation model using random order incremental training with learning functions*



F i g . 3 – *Development of artificial neural network activation model using resilient backpropagation*



F i g . 4 – *Development of artificial neural network activation model using one-step secant backpropagation*

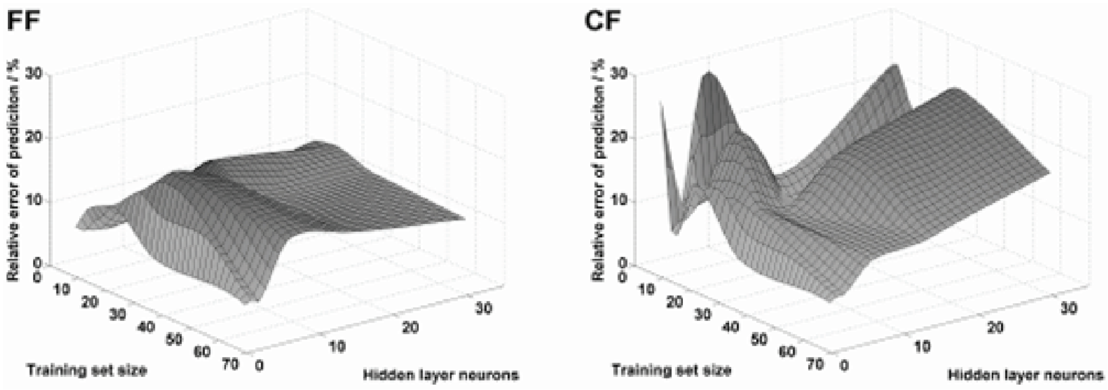


Fig. 5 – *Development of artificial neural network activation model using Levenberg-Marquardt backpropagation*

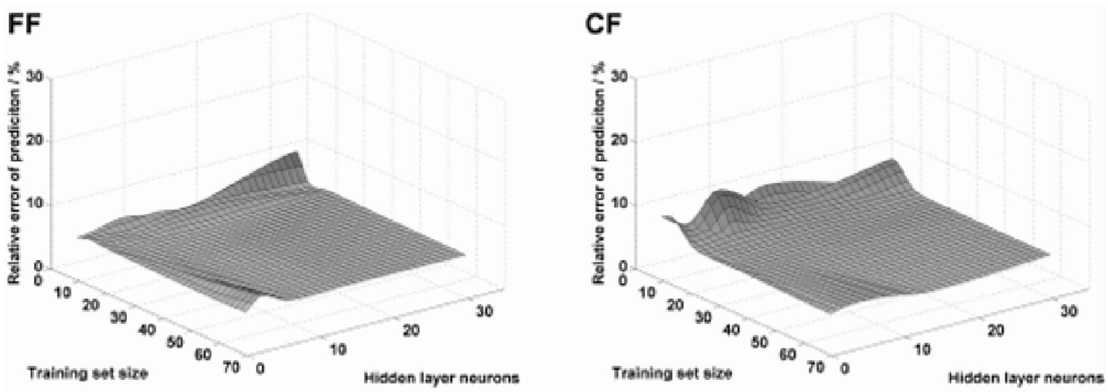


Fig. 6 – *Development of artificial neural network activation model using Bayesian regularization backpropagation*

rithm (Fig. 3). Keeping in mind that RI and RP training algorithms are based on first order information about error hyperplane, the batch type of data presenting can be clearly identified as an optimal solution. However, if the second order algorithms are considered (Figs. 4 and 5), higher predictive ability in the areas of the local minima is observed. The main drawback of the second order algorithms application is undeveloped error hyperplane with several local extremes resulting in unstable training and increased possibility of trapping in local minimums. Even the general advantage of the Levenberg Marquardt algorithm, i.e. its capability of using first-order information in combination with second-order information, does not prevent instability of the training process. On the other hand, the second order information enables significantly faster training, offering optimal solution only if local minima problems (over/undertraining) can be successfully overcome. Fig. 6 presents the upgrade of second order algorithms with the application of Bayesian regularization. It can be clearly seen that the problem with multiple minima was successfully solved yielding stable training and even higher predictive ability of developed networks (relative errors do not exceed 9.77 % in FF and 11.56 % in CF ANNs in any case). This identifies the Bayesian regularization training algorithm as the optimal solution for clay activation modeling. In addition, when the training data set consisted of less than ten experiments (Fig. 6), the percentage of the relative error was higher than when using 16 to 64 experiments in combination with 16 to 32 neurons (5.09 and 5.68 % for FF networks, and 4.27 and 6.01 % for CF networks respectively). Since the CF methodology features faster calculations, it was a reasonable selection for optimal network methodology, while optimal topology includes 16 neurons in hidden layer in order to keep the model as simple as possible (preventing possible overfitting). The optimal number of experimental data points needed for training was 32 (relative error 6.02 %) rather than 64 (relative error 5.25 %). It is obvious that predictive ability was slightly scarified in order to reduce experimental effort that, for these particular circumstances, seemed to be justified in order to promote economically feasible solutions.

The optimal artificial neural network activation model was used in combination with multi-objective criteria function in order to determine optimal activation parameters and improve sorption properties. The multi-objective criteria function (Equation 1) was designed to find the activation conditions that minimize time and power of activation (denominator) while simultaneously maximizing sorption capacity (nominator). The sum between the terms enables separate weighing of particular activation (ultrasonic and microwave) keeping their relevance reasonable, compared to sorption capacity, by avoiding their possible multiplication. In addition, squared factor is added to the sorption capacity. This way, the increase in sorption capacity was set as first priority while economic feasibility was set as second priority optimization condition. From Table 1 can be seen that overall activation lasts 10 minutes only using 120 W for ultrasonic and 60 W for microwave treatment. Application of suggested clay activation increased total clay sorption capacity by about 15 %. Having in mind relatively low energy consumption and activation duration time, the optimized electromagnetic irradiation clay treatment can be considered a promising alternative prior to application for hazardous substances removal.

Table 1 – *Mass of sorbed cobalt*

Parameter	*Optimal treatment*	*Without treatment*
P_{US} / W	120	0
t_{US} / min	5	0
P_{MW} / W	60	0
T_{MW} / min	5	0
m_S / mg g^{-1}	3.90	3.31

Conclusions

This work focuses on multi-objective optimization of natural clay sorption capacity using ultrasonic and microwave activation. For this reason, artificial neural network sorption models were developed. The results show that the application of batch data presentation in combination with second-order training algorithms and Bayesian was an optimal selection for clay activation properties ANN modeling. The cascade-forward methodology enables faster calculation while optimal topology includes 16 neurons in hidden layer. The training process was successfully conducted using 32 experimental data points, yielding good predictive ability (relative error 5.25 % based on external validation set). The developed ANN sorption methodology was applied in combination with in-house developed multi-objective criterion function. Optimal clay activation conditions include 5 minutes of ultrasonic (120 W) and microwave (60 W) treatment that results with a 15 % increase in sorption capacity. This approach can offer fast, efficient, and economically feasible approach to improve clay material properties prior to application in ecological and/or technological applications.

List of symbols

CR – multi-objective criteria function
P_{MW} – power of microwave irradiation
P_{US} – power of ultrasonic irradiation
t_{MW} – exposure time for microwave irradiation
t_{US} – exposure time for ultrasonic irradiation
m_{BND} – mass of bound cobalt ion

Abbreviations

ANN – artificial neural network
BP – backpropagation algorithm
BR – Bayesian regularization backpropagation algorithm
CF – cascade-forward methodology
FF – feed-forward methodology
LM – Levenberg-Marquardt backpropagation algorithm
MLP – multilayered perceptron
OSS – one-step secant backpropagation algorithm
RI – random order incremental training with learning functions
RP – resilient backpropagation algorithm
UV/Vis – ultraviolet/visible electromagnetic irradiation

References

1. *Pelino, M.,* Waste Manage. **20** (2000) 561
2. *Besegio, T., Berutti, F., Bernardes, A., Bergmann, C. P.,* J. Eur. Ceram. Soc. **22** (2002) 2251
3. *Theng, B.,* The Chemistry of Clay-Organic Reactions, Adam Hilger, London, 1974
4. *Churchman, G. J., Gates, W. P., Theng, B. K. G., Yuan, G.,* in Bergaya, F., Lagaly, G. and Theng, B. K. G. (Ed.), Handbook of Clay Science, pp 625–675, Elsevier, Amsterdam, 2006
5. *Tateo, F., Summa, V., Bonelli, C. G., Bentivenga, G.,* Appl. Clay Sci. **20** (2001) 97
6. *Manohar, D. M., Anoop Krishnan, K., Anirudhan, T. S.,* Water Res. **36** (2002) 1609
7. *Singh, S. P., Ma, L. Q., Harris, W. G.,* J. Environ. Qual. **130** (2001) 1961
8. *Malakul, P., Srinivasan, K. R., Wang, H. Y.,* Ind. Eng. Chem. Res. **37** (1998) 4296
9. *Ma, Q. Y., Traina, S. J., Logan, T. J., Ryan, J. A.,* Environ. Sci. Technol. **27** (1993) 1803
10. *Mench, M. V.,* Physico-chemical aspects and efficiency of trace element immobilization by soil amendments, Springer, Berlin, 1998
11. *Prost, R. Y.,* J. Soil Sci. **166** (2001) 880

12. *Wiesbrock, F., Hoogenboom, R., Schubert, U. S.,* Macromol. Rapid Commun. **25** (2004) 1739
13. *Gedye, R., Smith, F., Westaway, K., Ali, H., Baldisera, L., Laberge L., Rousell, J.,* Tetrahedron Lett. **27** (1986**)** 279
14. *Giguere, R. J., Bray, T. L., Duncan, S. C., Majetich, G.,* Tetrahedron Lett. **27** (1986) 4945
15. *Bose, A. K., Banik, B. K., Lavlinskaia, N., Jayaraman, M., Manhas, M. S.,* Chem. Tech. **27** (1997) 18
16. *Bose, A. K., Banik, B. K., Mathur, C., Wagle, D. R., Manhas, M. S.,* Tetrahedron Lett. **56** (2000) 5603
17. *Bose, A. K., Ing, Y. H., Pramanik, B. N., Bartner, P. L., Liu, Y. H., Heimark, L., Lavlinskaia, N., Sareen, C.,* J. Am. Soc. Mass Spectrom. **13** (2002) 839
18. *Bose, A. K., Manhas, M. S., Ganguly, S. N., Sharma, A. H., Banik, B. K.,* Synthesis **11** (2002) 1578
19. *Erickson, B.,* Anal. Chem. **70** (1998) 467A
20. *Richter, R. C., Link, D., Kingston, H. M. S.,* Anal. Chem. **1** (2001) 31A
21. *Lidstrom, P., Tierney, J., Wathey, B., Westman, J.,* Tetrahedron Lett. **57** (2001) 9225
22. *Larhed, M., Moberg, C., Hallberg, A.,* Acc. Chem. Res. **35** (2002) 717
23. *Melin, P. A.,* Hybrid Intelligent Systems for Pattern Recognition Using Soft Computing, Berlin-Heidelberg, Springer-Verlag, 2005
24. *Yegnanarayana, R.,* Artificial Neural Networks, Prentice-Hall of India, New Delhi, 2006
25. *Marinović, S., Bolanča, T., Ukić, Š., Rukavina, V., Jukić, A.,* Chem. Technol. Fuels Oils **48** (2012) 67
26. *Lewicki, P., Hill, T.,* Statistics: Methods and Applications, StatSoft Inc., Tulsa, USA, 2006
27. *Haykin, S.,* Neural Networks, A Comprehensive Foundation (2^{nd} ed.), Pearson Prentice Hall, Delhi, 1999
28. *Goyal, S., Kumar Goyal, G.,* Canadian Journal on Artificial Intelligence, Machine Learning and Pattern Recognition **2** (2011) 78
29. *Park, J., Diehl, F., Gales, M. J. F., Tomalin, M., Woodland, P. C.,* Comput. Speech. Lang. **25** (2011) 519
30. *Kazemi, A. R., Sobhanmanesh, F.,* Sci. Iran. **18** (2011) 1443
31. *Ghate, V. N., Dudul, S. V.,* Expert. Syst. Appl. **37** (2010) 3468
32. *Lightbody, G., O'Reilly, P., Irwin, G. W., Kelly, K., McCormick, J.,* Control Eng. Pract. **5** (1997) 1501
33. *Bolanča, T., Marinović, S., Ukić, Š., Jukić, A., Rukavina, V.,* Acta Chim. Slov. **59** (2012) 249
34. *Cerjan Stefanović, Š., Bolanča, T., Luša, M., Ukić, Š., Rogošić, M.,* Anal. Chim. Acta **716** (2012) 145
35. *Wang, D., Lu, W. Z.,* Atmos. Environ. **40** (2006) 913
36. *Drake, R., Packianather, M. S.,* Int. J. Adv. Manuf. Technol. **14** (1998) 280
37. *Damuth, H., Bale, M.,* Neural Network Toolbox, User's Guide, The MathWorks Inc., 2002
38. *Yousefizadeh, H., Zilouchian, A.,* in Zilouchian, A. and Jamshidi, M. (Ed.), Intelligent Control Systems Using Soft Computing Methodologies, pp 39–66, CRC Press, Boca Raton, 2001
39. *Bolanča, T., Cerjan-Stefanović, Š., Ukić, Š., Rogošić, M., Luša, M.,* J. Chemom. **22** (2008) 106
40. *Bolanča, T., Šipušić, J., Ukić, Š., Šiljeg, M., Ujević Bošnjak, M.,* Fresenius Environ. Bull. **21** (2012) 76
41. *Uzun, Y., Bicakci, K.,* Comput. Secur. **31** (2012) 717
42. *Battiti, R.,* Neural Comput **4** (1992) 141
43. *Bishop, C. M.,* Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995
44. *Ebrahimzadeh, A., Khazaee, A.,* Measurement **43** (2010) 103
45. *Choudhury, T. A., Hosseinzadeh, N., Berndt, C. C.,* Surf Coat Techol. **205** (2011) 4886
46. *Kostopoulos, A. E., Grapsa T. N.,* Neurocomputing **72** (2009) 3000
47. *Principe, J. C., Euliano, N. R., Lefebvre, W. C.,* Neural and Adaptive Systems: Fundamentals Through Simulations, New York, Wiley, 2000
48. *Nelles, O.,* Nonlinear System Identification, Berlin – Heidelberg – New York, Springer, 2000
49. *Renault, P., Cazevieille, P., Verdier, J., Lahlah, J., Clara, C., Favre, F.,* Geoderma **154** (2009) 101
50. *Aran, D., Maul, A., Masfaraud, J. F.,* C. R. Geosci. **340** (2008) 865.