Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 540 (2005) 455–463
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

Artificial neural network modelling of uncertainty in gamma-ray spectrometry

S. Dragović[a], A. Onjia[b], S. Stanković[a], I. Aničin[c], G. Bačić[d,*]

[a]*Institute for the Application of Nuclear Energy, INEP, Banatska 31b, 11080 Belgrade, Serbia and Montenegro*
[b]*Vinča Institute of Nuclear Sciences, P.O. Box 522, 11001 Belgrade, Serbia and Montenegro*
[c]*Faculty of Physics, University of Belgrade, Studentski trg 12-16, 11000 Belgrade, Serbia and Montenegro*
[d]*Faculty of Physical Chemistry, University of Belgrade, Studentski trg 12-16, 11000 Belgrade, Serbia and Montenegro*

Received 17 June 2004; received in revised form 12 October 2004; accepted 11 November 2004
Available online 7 January 2005

Abstract

An artificial neural network (ANN) model for the prediction of measuring uncertainties in gamma-ray spectrometry was developed and optimized. A three-layer feed-forward ANN with back-propagation learning algorithm was used to model uncertainties of measurement of activity levels of eight radionuclides (^{226}Ra, ^{238}U, ^{235}U, ^{40}K, ^{232}Th, ^{134}Cs, ^{137}Cs and ^{7}Be) in soil samples as a function of measurement time. It was shown that the neural network provides useful data even from small experimental databases. The performance of the optimized neural network was found to be very good, with correlation coefficients (R^2) between measured and predicted uncertainties ranging from 0.9050 to 0.9915. The correlation coefficients did not significantly deteriorate when the network was tested on samples with greatly different uranium-to-thorium (^{238}U/^{232}Th) ratios. The differences between measured and predicted uncertainties were not influenced by the absolute values of uncertainties of measured radionuclide activities. Once the ANN is trained, it could be employed in analyzing soil samples regardless of the ^{238}U/^{232}Th ratio. It was concluded that a considerable saving in time could be obtained using the trained neural network model for predicting the measurement times needed to attain the desired statistical accuracy.

PACS: 29.30.−h; 29.30.Kv; 29.90.+r

Keywords: ANN; Radionuclides; Uncertainty; Measurement time; Soil

1. Introduction

Artificial neural networks (ANN) are currently used in a wide variety of applications such as predictions, non-linear problems and real-time

*Corresponding author. Tel.: +381 11 630796; fax: +381 11 187133.
E-mail address: ggbacic@ffh.bg.ac.yu (G. Bačić).

0168-9002/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2004.11.045

data analysis. ANNs are data processing systems consisting of a large number of simple, highly interconnected processing elements that simulate biological neural networks. Numerous models of ANNs with different approaches both in architecture and in learning algorithms have been proposed [1–5]. There is always one input and one output layer and there should be at least one hidden layer, which enables ANNs to describe non-linear systems [6].

In gamma-ray spectrometry ANNs have been used to identify radioactive isotopes automatically from their spectra, employing pattern recognition of the entire spectrum instead of analyzing individual peaks [7,8] and for the quantitative spectrometry analysis [9–12]. Most reports deal with samples in which the radionuclides were artificially mixed, but there are few reports where ANN was utilized on environmental samples [10,12].

In the present study an ANN was employed and optimized to model the measuring uncertainties of activities of radionuclides in natural soil samples. The main aim of the work was to investigate the ability of a network trained with a small set of experimental data to provide a reliable estimate of the minimum measurement time needed to attain the desired average statistical accuracy in environmental samples of complex composition. This could be of special importance in the environmental monitoring programmes, when a large number of environmental samples have to be measured and when the time of measurement is important for the economization of experiments. We showed that the ANNs approach could determine a minimum measuring time which is requested for fulfilling quality assurance criteria for any particular application. In this study we demonstrated that an ANN which is trained on randomly selected samples of the soil is capable of predicting uncertainties for samples with different $^{238}U/^{232}Th$ ratio.

2. Experimental

2.1. Gamma-ray spectrometric measurements

The measurements were performed using a low-background HPGe gamma-ray spectrometer

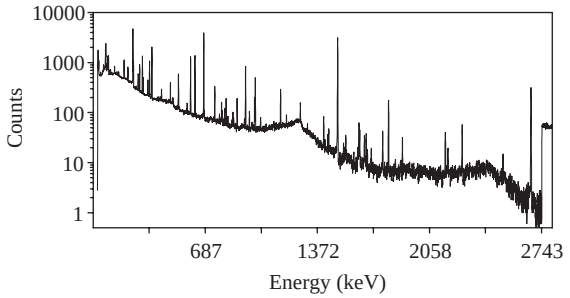

Fig. 1. A typical gamma-ray spectrum of the soil sample recorded after 17 h of measurement.

ORTEC-AMETEK (34% relative efficiency and 1.65 keV FWHM for ^{60}Co at 1.33 MeV, 8192 channels). The detector was surrounded by a shield consisting of 100 mm low-level lead (less than 50 Bq/kg), including interior lining consisting of 2 mm Cu, 1 mm Cd and 4 mm Perspex. Gamma Vision 32 software was used to process the spectra obtained [13].

The soil samples were measured in 1 L Marinelli beakers for 17 h. After this time no significant decrease of the measuring uncertainty was observed. The spectra were recorded after each hour. A typical spectrum obtained after 17 h is shown in Fig. 1. As the average difference between uncertainty data obtained for three soil samples from the same location were within 2%, a simple arithmetic mean value for each radionuclide was used. The uncertainties of activity measurements of the following radionuclides detected in soil samples were analyzed: ^{226}Ra, ^{238}U, ^{235}U, ^{40}K, ^{232}Th, ^{134}Cs, ^{137}Cs and ^{7}Be. Gamma-ray energies used in this work and activity levels obtained after 17 h of measurement with corresponding uncertainties are presented in Table 1. The efficiency calibration curve for the Marinelli beaker geometry was obtained experimentally using a commercial calibration source. According to the total summation method, the counting statistical uncertainty was calculated as the square root of the sum of the squares of the uncertainty in the gross area and the uncertainty of the background. The gross area of the peak is the sum of the contents of each channel between the background channels (including the two background channels). The

Table 1
Gamma-ray energies used in this work and average activity levels with corresponding uncertainties (for three soil samples used for ANN training) obtained after 17 h of measurement

Radionuclide	Energy (keV)	Activity level ± uncertainty ($Bq\,kg^{-1}$)
^{226}Ra	351.9; 609.3	16.3 ± 0.95
^{238}U	609.3; 1120.3; 1764.5	15.7 ± 0.63
^{235}U	143.8	0.72 ± 0.04
^{40}K	1460.8	441.5 ± 3.62
^{232}Th	338.4; 911.7; 968.9	24.9 ± 0.20
^{134}Cs	604.7; 795.8	0.09 ± 0.01
^{137}Cs	661.6	31.3 ± 0.36
^{7}Be	477.6	2.96 ± 0.41

uncertainty in the gross area is the square root of the area. The background area uncertainty is the uncertainty in the channels used to calculate the end points of the background multiplied by the ratio of the number of channels in the peak to the number of channels used to calculate the background.

2.2. Neural network algorithm

A three-layer feed-forward neural network trained with a back-propagation algorithm was used. The training process in back-propagation networks is done in two phases: feed-forward and back-propagation. In the feed-forward phase, the input layer neurons pass the input values on to the hidden layer. Each of the hidden layer neurons computes the weighted sum of its inputs, passes the sum through its activation function and presents the activation value to the output layer. After computation of the weighted sum of each neuron in the output layer, the sum is passed through its activation function, resulting in one output value for the network [14,15].

A sigmoidal function is used as the transfer function in this application:

$$f_j = \frac{1}{1 + \exp(-\sum w_{ji} o_i + b)} \quad (1)$$

where w_{ji} is the connection weight from neuron i in the lower layer to neuron j in the upper layer and an initially small random value, o_i is the output of neuron i, while b is the bias value. The bias (neuron activation threshold) is used to calculate the net input of a neuron from all neurons connected to it.

In the back-propagation phase, the error between the network output and the desired output values is calculated using the so-called generalized delta rule and weights between neurons are updated from the output layer to the input layer as follows:

$$w_{ji}^{n+1} = w_{ji}^{n} + \eta \delta_j o_j + \alpha w_{ji}^{n} \quad (2)$$

where δ_j is the error signal at neuron j, o_j is the output of neuron j, n is the number of iterations, and η and α are learning rate and momentum, respectively. The learning rate controls the rate at which the network learns. The momentum term has the effect of adding a proportion of the previous weight change during training [16]. The training process is successfully completed when the iterative process has converged.

In this work, MS-Windows-based ANN simulator, QwikNet Version 2.23 was used [17].

2.3. Training and testing the network

In order to optimize model performance, it is vital to know at what learn count a local minimum in the error surface is reached and what the magnitude of the oscillations in the forecasting error will be if training is continued. This requires three data sets: a training set, used to train the network, a test set, used to evaluate the generalization ability of the network and a validation set, used to assess the performance of the model once the training phase has been completed (the process called cross-validation) [18].

The training set consisted of uncertainties of activity measurements of radionuclides obtained after 2, 10 and 15 h. The "leave-10%-out" method was applied for cross-validation. With this method 10% of the data in the training set are not used for updating of weights, so this 10% can be used as an indication of whether or not memorization is taking place. When an ANN memorizes the training data, it produces acceptable results for the training data, but poor results when tested on

unseen data. At first the network was tested with data obtained for other times of measurement of radionuclide activities in the same sample which had not been used for network training. The network was further tested with two samples which were not included in the training process: one having a $^{238}U/^{232}Th$ ratio of 0.77, i.e. similar to that in the training sample (0.63) and one with a greatly different $^{238}U/^{232}Th$ ratio (1.72).

3. Results and discussion

3.1. ANN optimization

Since there are no theoretical principles for choosing the proper network topology, several different structures were trained. Neural networks were trained using different numbers of hidden nodes and learning epochs. At the start of the training run, all weights and all biases were initialized with random values. During training, modifications of the network weights and biases were made by back-propagation of the error. When the network was optimized, the testing data were fed into the network to evaluate the trained network.

Attempting to keep the learning speed as fast as possible, the learning rate is self-adjusted by the network to be 0.1. Also, we found that when the momentum was 0.1 the network could achieve faster convergence and avoid getting trapped in a local minimum. Low values of learning rate and momentum work well in many other applications [16,19], however, it should be noted that they tend to be application specific.

To determine the optimum number of hidden layer nodes, ANNs with different numbers of hidden layer nodes were trained. Finding the optimal number of hidden nodes is important since their function is to detect relationships between network inputs and outputs. If there is an insufficient number of hidden nodes, it may be difficult to obtain convergence during training. On the other hand, if too many hidden nodes are used, the network may lose its ability to generalize. The number of hidden nodes varied from 2 to 20 (Fig. 2) and root mean square errors (RMSE)

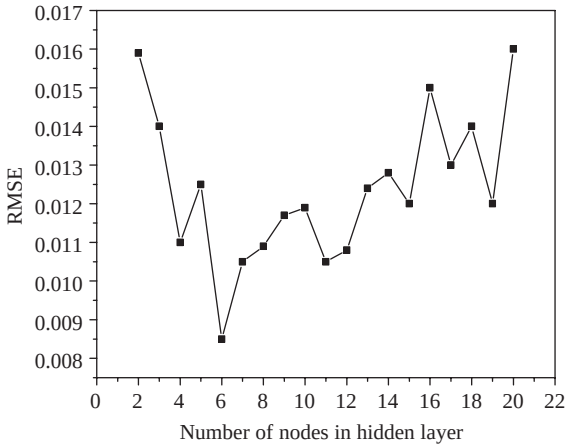

Fig. 2. The RMSE for different number of nodes in hidden layer.

were calculated:

$$\mathrm{RMSE} = \frac{\sqrt{\sum_{i=1}^{n}(d_i - o_i)^2/n}}{x} \tag{3}$$

where d_i is the desired output (experimental values) in the training or testing set, o_i the actual output (ANN predicted values) in the training or testing set, n the number of data in the training or testing set, while x is the average value of the desired output in the training or testing set [20]. Each topology was repeated five times to avoid random initialization of the weights. As shown in Fig. 2, the minimum error was obtained when the number of hidden layer nodes was equal to 6, so this value was chosen for optimization of the number of learning epochs.

To select the most favourable number of learning epochs, RMSE values vs. number of learning epochs were evaluated for both the training and testing sets (Fig. 3). In both cases, RMSE decreased rapidly at the beginning, and at 20,000 epochs the error of the testing set reached a minimum, after which the training worsened the prediction ability and the test error began to increase. This effect is called overtraining or overfitting and inspection of the cross-validation set is crucial for detection of this effect. The critical moment in training is when the error of this set starts to increase while the training error is still

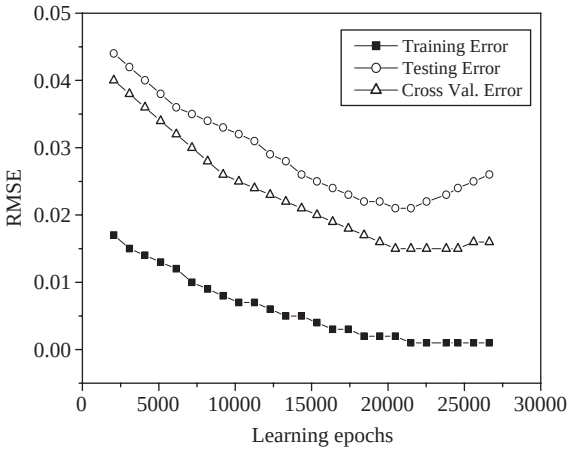

Fig. 3. The RMSE of the training, testing and cross-validation set vs. number of learning epochs.

decreasing. From that point improvement of training error is achieved by better fitting of the experimental points rather than by fitting the general trend of the data. In order to avoid overtraining, 20,000 learning epochs were selected as the optimal value.

3.2. ANN topology

The ANN optimized in this study is schematically represented in Fig. 4. The input layer consists of one node representing the time of measurement (t). The output layer consists of eight nodes representing the measuring uncertainties of the investigated radionuclides (^{226}Ra, ^{238}U, ^{235}U, ^{40}K, ^{232}Th, ^{134}Cs, ^{137}Cs and ^{7}Be). In addition, there are also biases connected to nodes in the hidden (bias 1) and output (bias 2) layers via modifiable weighted connections. The biases are responsible for accomodating nonzero offsets in the data. They act like other processing elements that have a constant output. The purpose of this is to scale the input to a useful range.

3.3. The influence of the number of training data on prediction accuracy

Reduction in the number of experimental data points used for the training set is crucial for the development of the uncertainty model without

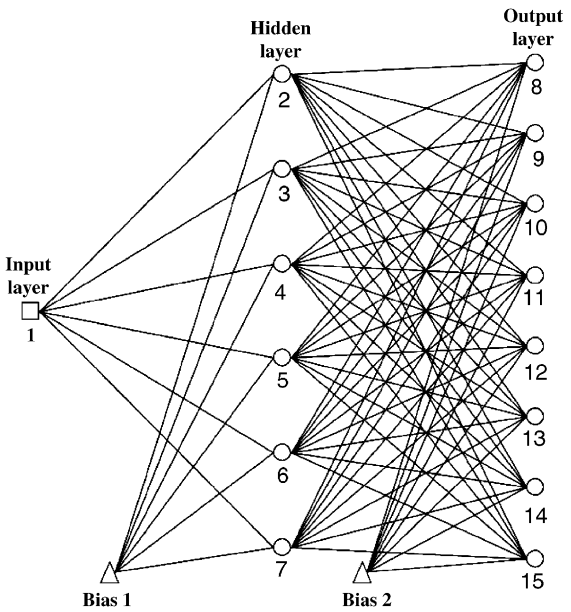

Fig. 4. Schematic representation of a three-layer feed-forward neural network used in this study. Input layer node: 1; hidden layer nodes: 2–7; output layer nodes: 8–15.

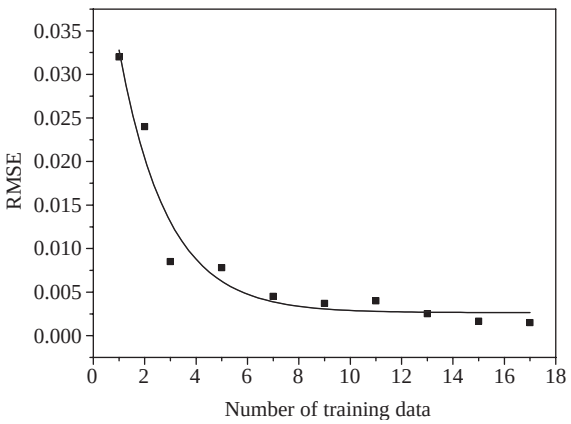

Fig. 5. Average RMSE vs. number of training points.

wasting time on unnecessary experiments. Fig. 5 shows the influence of the number of experimental data points used for the training set on the accuracy of the ANN.

The RMSE decreases with the increasing number of points in the training set, which can be explained by the fact that the modelling area is

more evenly (and in more detail) covered when more data are used in the training set [21]. However, this does not mean that further increase in the number of data in the training set will further decrease RMSE, since it has been demonstrated that the best models are not always obtained with the largest training sets [22]. Namely, when too many training data are presented to the network, it tends to learn the specific data set better than the general problem. This is undesirable, as the network should have the ability to generalize. In our application, the training set of three experimental points (three measuring times with corresponding uncertainties for all analyzed radionuclides), serving as a representative subset of the complete data set analyzed in this study, turned out to be sufficient for good prediction ability of the network.

3.4. ANN validation

The optimized neural network model was used to predict uncertainties of radionuclide activities for the remaining 14 measurement times, i.e., those which were not used for the ANN training. The predictive power was assessed by comparing experimental and predicted uncertainties (Table 2). The results show that the proposed neural network model adequately generalizes data and that it can be used for modelling of uncertainties. Correlation coefficients are shown in Table 3. What is more important is the fact that the difference between actual and ANN predicted uncertainties are not critically influenced by the absolute values of uncertainties in detecting particular radionuclides (Fig. 6) as the sign of differences in uncertainties is not a function of the measuring time. Namely, irrespective of whether we are examining radionuclides with low (^{40}K) or high (^{7}B) uncertainties, the two curves intermesh in a random fashion (Fig. 6; top row). Also, the correlation coefficient appears not to be related to the measuring uncertainties (compare Tables 2 and 3). The same pattern was found in samples with different ^{238}U/^{232}Th ratios (Fig. 6; middle and bottom row), but the correlation coefficients for all radionuclides were lower than those established for the ^{238}U/^{232}Th ratio used for training the network. The differences were

Table 2
Experimental (A) and predicted (B) standard uncertainties of measurement of activities of radionuclides in sample with ^{238}U/^{232}Th ratio of 0.63

t (h)	Uncertainty (%)															
	^{226}Ra		^{238}U		^{235}U		^{40}K		^{232}Th		^{134}Cs		^{137}Cs		^{7}Be	
	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
1	19.6	16.9	17.1	13.9	17.9	17.5	2.87	2.28	3.02	2.66	20.5	18.3	2.97	2.53	39.2	37.3
3	12.9	13.2	10.1	10.6	13.7	13.5	1.64	1.63	1.76	2.11	13.7	14.8	1.66	1.66	29.3	32.2
4	10.6	10.8	8.31	8.67	10.4	10.9	1.43	1.32	1.52	1.71	11.6	12.3	1.43	1.36	22.6	27.4
5	9.14	9.09	7.78	7.34	9.15	9.06	1.20	1.14	1.36	1.41	10.7	10.4	1.28	1.24	20.7	22.9
6	7.79	8.04	6.74	6.53	8.33	7.92	1.10	1.03	1.24	1.22	9.88	9.26	1.19	1.19	20.5	19.8
7	7.69	7.44	6.65	6.06	7.71	7.28	1.08	0.98	1.15	1.11	9.20	8.62	1.19	1.17	20.3	17.9
8	7.28	7.08	6.19	5.77	7.40	6.88	1.01	0.94	1.08	1.04	8.50	8.24	1.17	1.17	18.5	16.7
9	6.97	6.85	5.86	5.58	6.96	6.63	0.96	0.92	1.02	1.00	7.88	8.00	1.19	1.16	17.1	16.0
11	6.78	6.60	5.22	5.37	6.30	6.34	0.87	0.90	0.93	0.95	7.80	7.74	1.18	1.16	15.6	15.1
12	6.68	6.52	5.06	5.31	6.06	6.25	0.85	0.89	0.88	0.94	7.76	7.67	1.17	1.15	15.2	14.9
13	6.00	6.47	4.86	5.26	5.98	6.19	0.84	0.89	0.86	0.93	7.48	7.61	1.16	1.15	14.6	14.7
14	5.85	6.42	4.81	5.23	5.93	6.14	0.84	0.89	0.84	0.92	7.40	7.57	1.16	1.15	14.3	14.5
16	5.82	6.36	4.79	5.18	5.92	6.07	0.83	0.88	0.83	0.91	7.00	7.51	1.16	1.15	14.0	14.3
17	5.82	6.34	4.70	5.16	5.90	6.04	0.82	0.88	0.82	0.90	6.88	7.49	1.15	1.15	13.8	14.2

Table 3
Correlation coefficients between measured and predicted uncertainties for samples with different $^{238}U/^{232}Th$ ratios

$^{238}U/^{232}Th$	Correlation coefficient							
	^{226}Ra	^{238}U	^{235}U	^{40}K	^{232}Th	^{134}Cs	^{137}Cs	^{7}Be
0.63	0.9770	0.9568	0.9914	0.9765	0.9353	0.9612	0.9915	0.9291
0.77	0.9604	0.9540	0.9758	0.9625	0.9326	0.9463	0.9743	0.9221
1.72	0.9248	0.9234	0.9421	0.9229	0.9152	0.9150	0.9382	0.9050

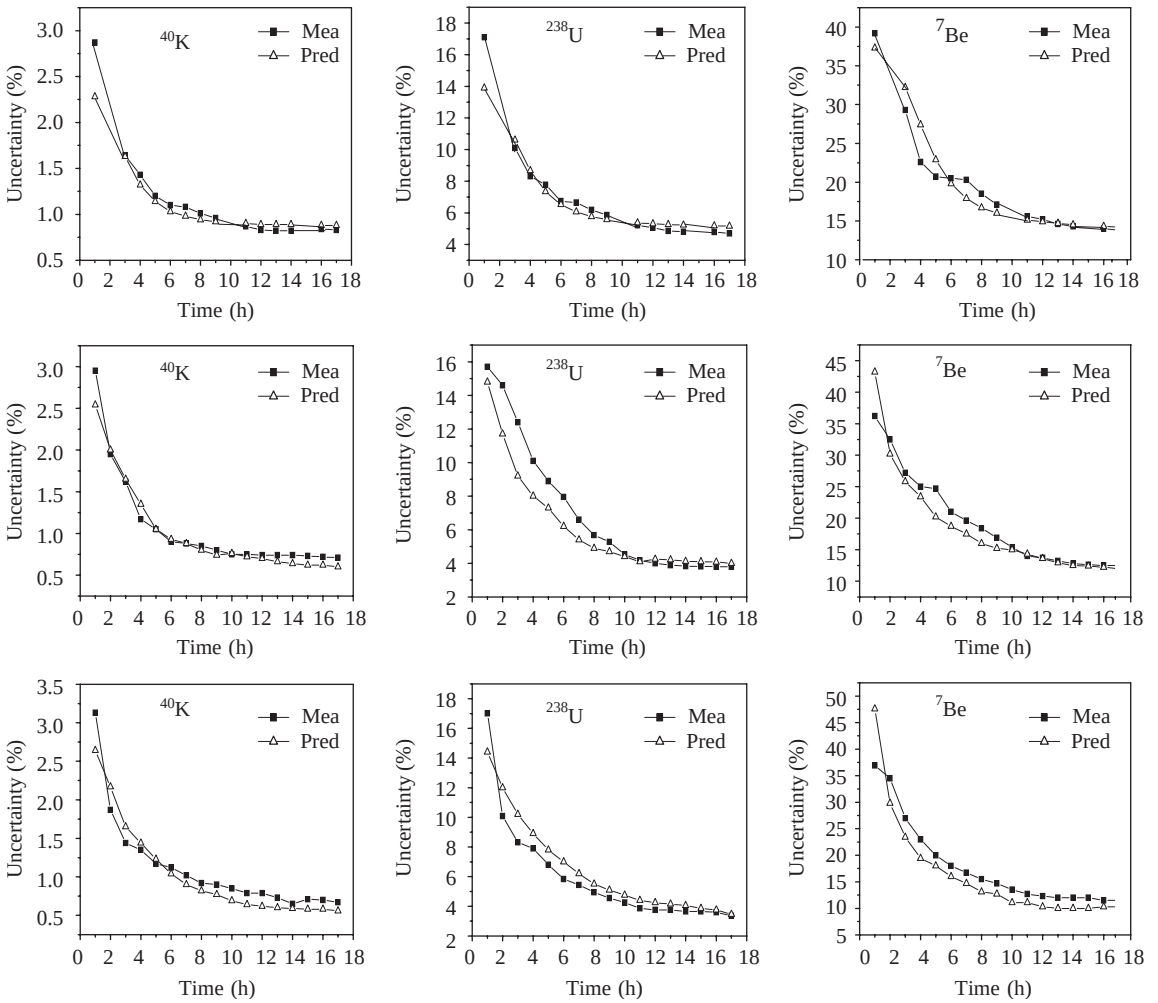

Fig. 6. Experimental and predicted uncertainties of measurement of activities of three radionuclides (^{40}K, ^{238}U and ^{7}Be) in samples with different $^{238}U/^{232}Th$ ratios (top row: 0.63; middle row: 0.77; and bottom row: 1.72) as a function of measurement time.

negligible for the sample with the $^{238}U/^{232}Th$ ratio of 0.77, but slightly increased (approximately by 4%) for the sample with a ratio of 1.72. These differences were not significant, i.e., once ANN was trained it could be successfully used for uncertainty modelling regardless of the

content of radionuclides in soil samples ($^{238}U/^{232}Th$ ratio).

There are few studies on employing the ANN approach in gamma-ray spectrometry, however, in only one paper [11] comparison between calculated and reference activities has been performed. In addition, in most studies, only samples containing artificially mixed radionuclides have been analyzed. The sole exception is the attempt to use network to analyze activities from uranium ores [9]. Here, we demonstrated that neural network could predict the uncertainties in natural soil samples.

The results obtained in our study indicate that the neural approach is a suitable alternative to the traditional approach in gamma-ray spectrometry. Moreover, because output results for a given set of input data are available in a very short time (few seconds), ANN has considerable potential for solving time-consuming problems in systematic analysis of environmental samples. We selected to test a rather simple ANN using the back-propagation algorithm for uncertainty modelling, but there are other sophisticated training algorithms available, and their suitability will be assessed in the future.

4. Conclusions

In this work ANN was used for uncertainty modelling in gamma-ray spectrometric measurements. The number of nodes in the hidden layer of ANN and the number of learning epochs were optimized. Through the above process, we found out that the optimal number of hidden layer nodes was 6, and the number of learning epochs 20,000. This study shows that ANNs are a very accurate and rapid uncertainty modelling tool with a small amount of experimental data needed for their optimization. The performance of our optimized ANN is found to be very favourable, with correlation coefficients (R^2) between measured and predicted uncertainties ranging from 0.9050 to 0.9915. The correlation coefficients were not influenced either by the absolute values of uncertainties or by the content of radionuclides in the samples ($^{238}U/^{232}Th$ ratio). It was concluded that a considerable saving in time could be obtained using this trained neural network model for predicting the measurement times needed to attain the desired statistical accuracy.

Acknowledgements

This work was supported by the Ministry of Science and Environmental Protection of the Republic of Serbia (Contract No. 2016, 1928 and 1978).

References

[1] M. Gevrey, I. Dimopoulos, S. Lek, Ecol. Model. 160 (2003) 249.
[2] K. Rajer-Kanduč, J. Zupan, N. Majcen, Chemometr. Intell. Lab. Syst. 65 (2003) 221.
[3] I.M. Raimundo, R. Narayanaswamy, Sensors Actuators B 90 (2003) 189.
[4] H. Bechtler, M.W. Browne, P.K. Bansal, V. Kecman, Appl. Thermal Eng. 21 (2001) 941.
[5] H.R. Maier, G.C. Dandy, Environ. Model. Software 15 (2000) 101.
[6] S. Agatonovic-Kustrin, M. Zecevic, Lj. Zivanovic, I.G. Tucker, Anal. Chim. Acta 364 (1998) 265.
[7] P. Olmos, J.C. Diaz, J.M. Perez, P. Gomez, V. Rodellar, P. Aguayo, A. Bru, G. Garcia Belmonte, J.L. de Pablos, IEEE Trans. Nucl. Sci. 38 (1991) 971.
[8] P. Olmos, J.C. Diaz, J.M. Perez, P. Aguago, P. Gomez, V. Rodellar, IEEE Trans. Nucl. Sci. 41 (3) (1994) 637.
[9] E. Yoshida, K. Shizuma, S. Endo, T. Oka, Nucl. Instr. and Meth. A 484 (2002) 557.
[10] V. Vigneron, J. Morel, M.C. Lepy, J.M. Martinez, Nucl. Instr. and Meth. A 369 (1996) 642.
[11] V. Pilato, F. Tola, J.M. Martinez, M. Huver, Nucl. Instr. and Meth. A 422 (1999) 423.
[12] P.E. Keller, R.T. Kouzes, IEEE NPSS Conference Proceedings, 94CH35726, 1994, p. 341.
[13] Gamma Vision 32, Gamma-Ray Spectrum Analysis and MCA Emulation, Version 5.3, ORTEC, Oak Ridge, USA, 2001.
[14] P.J.Werbos, Beyond regression: new tools for predictive and analysis in the behavioral sciences, Ph.D. Thesis, Harvard University, Cambridge, MA, 1974.
[15] D.E. Rumelhart, G.E. Hinton, R.J. William, Learning internal representations by error propagation, in: D.E. Rumelhart, J.C. McClelland (Eds.), Proceedings Parallel Distributed Processing. Foundations, vol. 1, MIT Press, Cambridge, 1986.
[16] H.R. Maier, G.C. Dandy, Environ. Model. Software 13 (1998) 193.

[17] QwikNet Version 2.23, Craig Jensen, Redmond, USA, 1999, http://www.qwiknet.home.comcast.net/.
[18] H.R. Maier, G.C. Dandy, Math. Comp. Model. 33 (2001) 669.
[19] J.R. Long, V.G. Gregoriou, P.J. Gemperline, Anal. Chim. Acta 62 (1990) 1791.
[20] R.G. Alany, S. Agatonovic-Kustrin, T. Rades, I.G. Tucker, J. Pharm. Biomed. Anal. 19 (1999) 443.
[21] T. Vasiljević, A. Onjia, Dj. Čokeša, M. Laušević, Talanta 64 (3) (2004) 785.
[22] R. Kocjančić, J. Zupan, Chemometr. Intell. Lab. Syst. 54 (2000) 21.