Neural network method for activity measuring in environmental samples.

Finkel Felix[1]; Bystrov Eugeny[2]
[1] Russian Mendeleyev Research Institute for Metrology, RUSSIAN FEDERATION
[2] Atomtex, BELARUS

Neural network method for activity measuring in environmental samples. Theses.

The spectrometric method is one of the most popular methods for activity measuring in environmental samples. It provides application of semi-conductor and scintillation spectrometers. Despite high measurement accuracy, high price and some difficulties in operation make semi-conductor detectors difficult to apply. Scintillation spectrometers, which are cheaper, are applied to determine activity of natural radionuclides. But continuous spectrum image on the monitor permits the spectrometers to measure not so many radionuclides and, thus, it restricts application of the scintillation spectrometers.

One of the ways to solve the problem is to use sophisticated mathematical analysis of decomposition of scintillation spectra. Examination of different spectrum analyses and devices, produced by means of these methods, reveals that sometimes the above solution is quite improper and has weak solution stability. Besides, some problems appear while applying these methods in field devices, which are to carry out scale analysis without a computer. As a result, simplified methods are applied, which reduce accuracy of activity determination.

The current article describes the method of decomposition of scintillation spectra into spectrum components by means of artificial neural networks. The aim of the above method is to determine activity of radionuclides in the source under measurement according to its radiation spectrum. The spectrum is acquired by means of scintillation spectrometer. If applying a laboratory spectrometer, which, as a rule, is equipped with a computer, the spectrum is transferred from the ADC to the PC, which operates with the above method. If applying field spectrometers, spectrum is to be processed by either a smart sensor or operating control unit. Activity calculation algorithm is equal in both cases.

Introduction

Nowadays industry often has bad unwholesome influence on the environment and a human being. That is why environmental friendliness is one of the quality factors of production. It causes constant development of special equipment and methods for environmental monitoring. Radioactivity stands out from other quality factors to be controlled. High radiation contamination risks, on the one hand, and strict safety regulations, on the other hand, draw special attention to metrological parameters of modern equipment of radiation monitoring.

The spectrometric method is one of the most popular methods for activity measuring in environmental samples. It provides application of semi-conductor and scintillation spectrometers. Despite high measurement accuracy, high price and some difficulties in operation make semi-conductor detectors difficult to apply. Scintillation spectrometers, which are cheaper, are applied to determine activity of natural radionuclides. But continuous spectrum image on the monitor permits the spectrometers to measure not so many radionuclides and, thus, it restricts application of the scintillation spectrometers.

One of the ways to solve the problem is to use sophisticated mathematical analysis of decomposition of scintillation spectra. Examination of different spectrum analyses and devices, produced by means of these methods, reveals that sometimes the above solution is quite improper and has weak solution stability. Besides, some problems appear while applying these methods in field devices, which are to carry out scale analysis without a computer. As a result, simplified methods are applied, which reduce accuracy of activity determination.

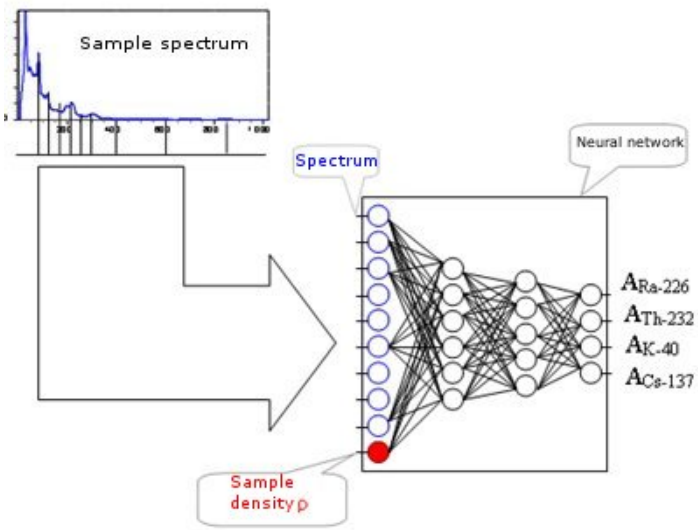
The current article describes the method of decomposition of scintillation spectra into spectrum components by means of artificial neural networks. The aim of the above method is to determine activity of radionuclides in the source under measurement according to its radiation spectrum. The spectrum is acquired by means of scintillation spectrometer. If applying a laboratory spectrometer, which, as a rule, is equipped with a computer, the spectrum is transferred from the ADC to the PC, which operates with the above method. If applying field spectrometers, spectrum is to be processed by either a smart sensor or operating control unit. Activity calculation algorithm is equal in both cases.

Activity calculation of radionuclides

Since the spectrometric method under discussion is relative, calibration of the equipment is required. According to the neural network theory such calibration is called learning.

The step-by-step algorithm of activity calculation follows. Learning (calibration) is described in part 3 of the current article.

First of all, we describe input data. We have a spectrum of a working source N collected by means of a spectrometer. The source (sample) is measured in the preset geometry with the certain density $\rho_и$ and life time. We have also a background spectrum $N_ф$ measured under the same conditions. Background is usually measured by means of a spectrometer either with an empty shielding chamber or with an empty sample dish in the chamber.



Pic.1. Diagram of activity measurement in samples by means the neural network method

To reduce the spectrum to the energy scale, we should know energy calibration. Modern spectrometers provide the relation between a spectrum channel No. and energy, which are linearly dependent as follows:

$$E_i = E_0 + k \cdot i \qquad (1),$$

where E_i – energy corresponding to the *i*-channel, keV; E_0, k – coefficients of energy calibration fixed during the device setting.

Data gained in the course of learning (calibration) is weights of neural network W, scales of each network input M_{in} and output M_{out} and energy windows where summing up R_E is carried out.

During calibration network parameters should be set up, including the number of layers and neurons in each layer and a type of activation function with parameters. The number of network inputs depends on the number of selected windows, plus one input for density. The number of network outputs corresponds to the number of radionuclides under detection.

First of all, we prepare input data for network calculation.

In the spectrum channels we define corresponding windows for each specified energy window. According to Eq.1:

$$R_{Lj} = \frac{R_{ELj} - E_0}{k} \quad , \ j=1..n \qquad (2)$$

where R_{Lj} – channel No. in the spectrum of the left edge of the *j*-window; R_{ELj} – energy of the left edge of the *j*-window, keV; n – number of windows.

The similar equation is used for the right edge:

$$R_{Rj} = \frac{R_{ERj} - E_0}{k} \quad , \ j=1..n \qquad (3)$$

where R_{Rj} – channel No. in the spectrum of the right edge of a *j*-window; R_{ERj} – energy of the right edge of a *j*-window, keV.

We define total counting rate in the specified windows for the source working spectrum and background spectrum:

$$S_j = \frac{1}{t} \cdot \sum_{i=R_{Lj}}^{R_{Rj}} N_i \quad , \text{ j=1..n} \qquad (4)$$

$$S_{\phi j} = \frac{1}{t_\phi} \cdot \sum_{i=R_{Lj}}^{R_{Rj}} N_{\phi i} \quad , \text{ j=1..n} \qquad (5)$$

where S_j and $S_{\phi j}$ – counting rates in the j-window of working and background spectra correspondingly, imp/sec; t и t_ϕ – real measurement time of working and background spectra correspondingly, sec. [1, 4].

Now we define input values of the neural network as follows:

$$X_j = \frac{S_j - S_{\phi j}}{M_{in\,j}} \quad , \text{ j=1..n} \qquad (6)$$

where X_j – value of the j-output of the neural network; $M_{in\,j}$ – scale value for the j-output. Eq. 6 reveals that the value X_j can vary from 0 to 1, which is caused by principles and architecture of neural network.

The number of network inputs m=n+1, with another special input for density, is defined as follows:

$$X_m = \frac{\rho_u}{M_{in\,m}} \qquad (7)$$

where $\rho_и$ – density of the sample under measurement, g/cm^3; $M_{in\ m}$ - scale for density input, g/cm^3.

Input data is gathered now and should be put in calculation of the neural network. Network calculation means definition of states and outputs of neurons within inner layers and the last output layer [2]. These are outputs of the last layer, which are desired unscaled activities of radionuclides. One should remember that the number of outputs is not fixed and defined according to research tasks. It means that if activity of a certain radionuclide is to be determined in the sample with a radionuclide mixture, the network can have one output corresponding to a specified radionuclide. In this case network learning is carried out by means of spectra of unmeasured radionuclides.

Network calculation is made in a layer-by-layer manner. Calculation of neuron states for the first layer is not necessary, since their outputs, or network inputs, are known. Thus, neuron states for the second layer are calculated as follows:

$$C_j^{II} = \sum_{i=1}^{P_1} X_i \cdot W_{1,i,j} \quad , \text{ j=1..P}_2 \qquad (8)$$

where P_1 - number of neurons in the first layer, similar to the number of network outputs; P_2 – number of neurons in the second layer; $W_{1,i,j}$ – weight between the i-neuron of the first layer and j-neuron of the second layer, the index *l* depicts a layer, which has a link output (the first layer for Eq. 8).

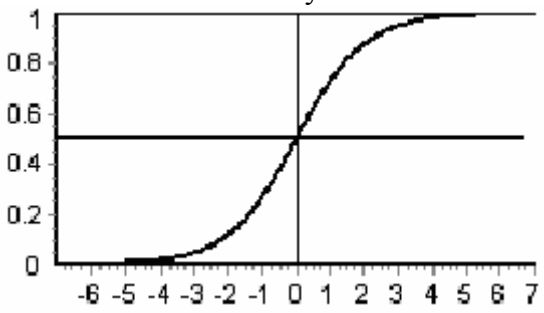
Neuron output is regarded as the function of its state:

$$y_j^{II} = f(C_j^{II}) \quad , \text{j=1..P}_2 \qquad (9)$$

where y_j^{II} - output of the j-neuron in the second layer; f – activation function, defined as:

$$f(c) = \frac{1}{1 + \exp(-\alpha \cdot c)} \quad (10)$$

where c – argument of the function or, in this case, neuron state; α – coefficient, which determines the slope of the function graph (Pic.2). The coefficient α is determined in the course of calibration and can vary from 0.5 to 3.



Pic. 2. Sigmoid activation function [3].

Despite vast variety of activation functions, the research revealed that the above type of the activation function provides particularly successful and accurate calculation.

Neuron states for the third layer are calculated as follows:

$$C_j^{III} = \sum_{i=1}^{P_2} y_i^{II} \cdot W_{2,i,j} \quad , \quad \text{j=1..P}_3 \quad (11)$$

As shown by Eq. 11, the network inputs are replaced by the outputs of the second layer. Thus, we have got the following equation, similar to Eq. 9:

$$y_j^{III} = f(C_j^{III}) \quad , \text{j=1..P}_3 \quad (12)$$

where y_j^{III} - output of the j-neuron in the third layer.

The network for the rest of layers, including the output layer [2, 3], is calculated on the same principle as illustrated above. In this case, the output layer is described as follows:

$$C_j^Z = \sum_{i=1}^{P_{Z-1}} y_i^{Z-1} \cdot W_{Z-1,i,j} \quad (13)$$

$$y_j^Z = f(C_j^Z) \quad , \text{j=1..P}_Z \quad (14)$$

where y_j^Z - output of the j-neuron in the last layer; Z – number of layers in the network.

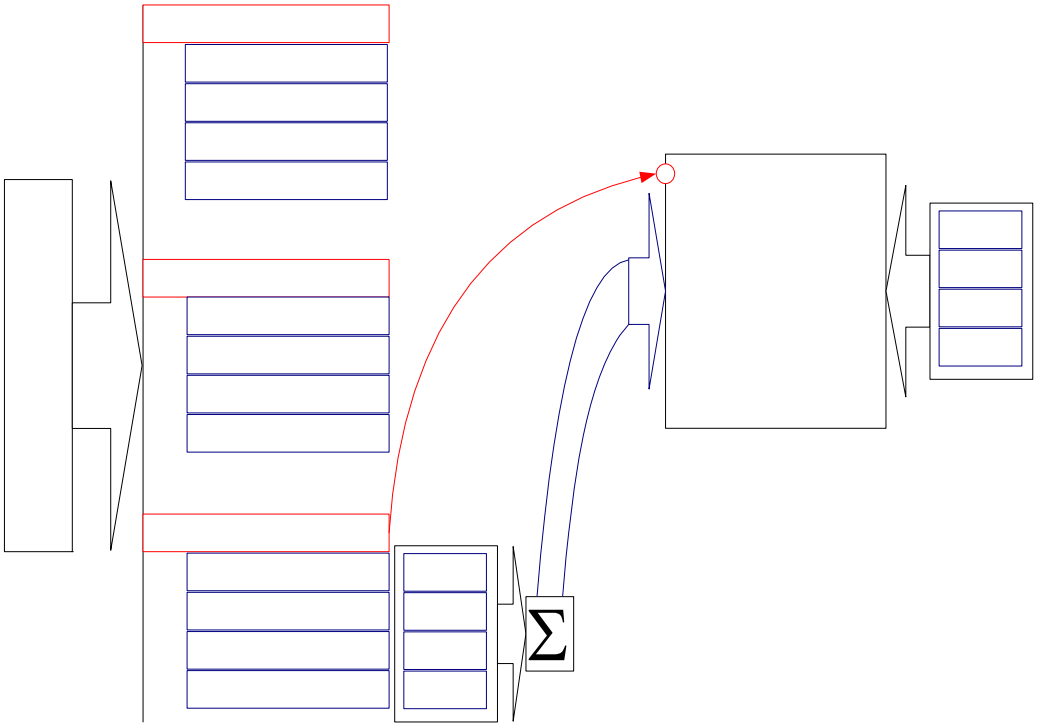
To turn directly to calculation of radionuclides activities, output scales determined during the calibration should be used.

$$A_j = (y_j^Z - k_{кр}) \cdot (\frac{M_{out\,j}}{k_{ум}}) \quad , \ j=1..P_Z \qquad (15)$$

where A_j – activity of the j-neuron in order, Bq; $M_{out\,j}$ - scale for j-output of the network; $k_{кр}$ –boundary coefficient, which enables not to use the non-linear range of a sigmoid function and is defined as 0.1; $k_{ум}$ – coefficient introduced to correct a scale disturbed by the boundary coefficient. More detailed description of the above coefficients follows.

Calibration (learning).

To train the neural network, calibration spectra are required, which should be spectra of separate radionuclides measured in calibration samples with different densities (Pic.3).



Pic. 3. Example of choosing calibration spectra for learning.

These available calibration spectra should be included into the learning sample, where each component is a spectrum with different radionuclide content and density. To get such spectra, the additivity concept of simple spectra of radionuclides is to be applied. Thus, the equation for a sample spectrum with activities A_j, where j=1..n (n – number of nuclides in a source) will be the following:

$$Q_i = Q_{\phi_i} + \sum_{j=1}^{n} q_{ij} \cdot A_j \ , \quad i=1..v \qquad (16)$$

where Q_i – counting rate in the i-channel of the mixed spectrum of the sample, imp/sec; $Q_{\phi i}$ – counting rate in the i-channel of the background spectrum, imp/sec; q_{ij} – counting rate in the i-channel for the simple spectrum of the corresponding j-radionuclide, imp; v- number of channels in a spectrum. Eq. 16 can be applied provided that energy calibration of both simple spectra and the background spectrum is similar. If the above condition is not met, the value of energy calibration is to be made equal for all the spectra by means of Eq. 1. Besides, one should take into account that simple spectra *q* are collected for samples with one and the same density.

Simple spectra *q* are calculated as follows:

$$q_{ij} = \frac{\frac{N_i}{t} - \frac{N_{\phi_i}}{t_\phi}}{A_j}, \quad i=1..v, j=1..n \qquad (17)$$

where N_i– counting in the i-channel in the real j-radionuclide spectrum, imp; $N_{\phi i}$– counting in the i-channel in the background spectrum, imp; A_j – sample activity of the j-radionuclide, Bq; t and t_ϕ – life time for measuring source spectra and background spectra correspondingly, sec [4].

As shown by Eq. 15, while changing the value A_j we can acquire the required spectrum Q for a non-existent sample with the density similar to that one, for which simple spectra q were collected. Thus, Eq. 15 can be put down as follows:

$$Q_i^k = Q_{\phi i} + \sum_{j=1}^{n} q_{ij}^k \cdot A_j^k, \quad i=1..v, \ k=1..r \qquad (18)$$

where r – number of densities; k – density No. in order.

Thus, Eq. 17 is the main one for making both learning and testing samples. The variable k, i.e. density, and variables A_j are subject to variation, or the so-called randomization. And this is the acquired spectrum Q, which is the learning element. There can be as many similar elements as possible. It means that the more information the neural network learns, the more accurate and precise results are achieved.

Now we describe the learning algorithm in details. We have simulated spectra of radionuclide mixtures Q and the activity vector $A_в$ corresponding to each of the spectra. Thus, according to the algorithm described in part 2 of the current article, initialization of the matrix weights W should be carried out by means random variables from 1 to 0.3 [2]. Then, we put different spectra from the learning sample Q in random order to the network input, calculate forward the network by means of Eqs. 1 – 15 and, as a result, get the simulated vector of radionuclide activity $A_м$. In this case we have already got activity vector $A_в$ from the learning sample for the spectrum Q.

The matrix weights are corrected in the opposite way to the network distribution. Thus, an auxiliary j-variable can be illustrated as follows:

$$g_j^Z = y_j^Z \cdot (1 - y_j^Z) \cdot (y_j^Z - y_{B\,j}), \quad j=1..P_Z \qquad (19)$$

where y_j^Z - j-neuron output of the last layer; Z – number of network layers; $y_{B\,j}$ - target value of the j-output, calculated according to the following:

$$y_{B\,j} = \frac{A_{B\,j}}{M_{out\,j}}, \quad j=1..P_Z \qquad (20)$$

where $M_{out\,j}$ - scale for j-output of the network.

Now we calculate the coefficient of weight correction:

$$\Delta W^{T}_{Z-1,i,j} = -\eta \cdot g^{Z}_{j} \cdot y^{Z-1}_{i} + \mu \cdot \Delta W^{T-1}_{Z-1,i,j}, \quad \text{i=1..P}_{\text{Z-1}}\text{, j=1..P}_{\text{Z}} \tag{21}$$

where T- number of learning iteration, if T=1 (i.e. at the beginning of learning) variables $\Delta W^{T-1}_{Z-1,i,j}$ are equal to 0; η – parameter, which determines learning rate, can vary from 0 to 1; μ – learning coefficient (Part 2) can vary from 0 to 1. The minus sign "-" in front of the first component of Eq.21 is caused by the fact that the weight changes in the opposite way to that one indicated by the derivative of error surface.

Weight correction is carried out according to the following equation:

$$W^{T}_{Z-1,i,j} = W^{T-1}_{Z-1,i,j} + \Delta W^{T}_{Z-1,i,j}, \quad \text{i=1..P}_{\text{Z-1}}\ \text{j=1..P}_{\text{Z}} \tag{22}$$

The above weight correction (Eqs. 19 – 22) is carried out in the opposite way to the network distribution for each pair of layers. This approach is called "backward error distribution".

For the next pair of inner (hidden) layers (***not for the last and next to last ones***) weights change in the following way:

$$g^{Z-1}_{j} = y^{Z-1}_{j} \cdot (1 - y^{Z-1}_{j}) \cdot \delta_{j}, \quad \text{j=1..P}_{\text{Z-1}} \tag{23}$$

where

$$\delta_{j} = \sum_{k=1}^{P_Z} g^{Z}_{k} \cdot W^{T}_{Z-1,j,k} \tag{24}$$

Eq. 23 differs from Eq. 19 on the error of a hidden layer δ_j, which does not correspond directly to target output values. That is why weight correction of the hidden layer is proportional to its "contribution" to the error of the next layer.

Then, weight correction is carried out in a similar way to Eqs. 21 and 22:

$$\Delta W^{T}_{Z-2,i,j} = -\eta \cdot g^{Z-1}_{j} \cdot y^{Z-2}_{i} + \mu \cdot \Delta W^{T-1}_{Z-2,i,j}, \quad \text{i=1..P}_{\text{Z-2}}\text{, j=1..P}_{\text{Z-1}} \tag{25}$$

$$W^{T}_{Z-2,i,j} = W^{T-1}_{Z-2,i,j} + \Delta W^{T}_{Z-2,i,j}, \quad \text{i=1..P}_{\text{Z-2}}\ \text{j=1..P}_{\text{Z-1}} \tag{26}$$

As far as next layers are concerned, weights are transformed according to Eqs. 23 – 26 if correction is carried out from the end of the network to its beginning [2, 3].

After correction completion, another spectrum Q with a density parameter is put into the input the network to start learning. The network is calculated forward and the results are compared with target values of the vector $y_{B\,j}$ specified by the spectrum Q. Then, learning is carried out in the opposite way. There can be as many learning iterates as possible (about $5 \cdot 10^6$). It stands to reason that only a computer can calculate the

above. The higher capacity computers have, the greater interest to artificial neural networks is aroused in different spheres of physics.

References

1. Vartanov, N. & Samoylov, P. (1975). Applied scintillation gamma spectrometry.
2. Kalan, R. (2003). Main concepts of neural networks.
3. Osovski, S. (2004). Neural networks for information processing.
4. Pegoev, A. (1980). Data processing methods in applied gamma spectrometry.