


# A New Method on Kerma Estimation in Mammography Screenings

Nabipour M.<sup>1</sup>, Deevband M. R.<sup>2\*</sup>, Asgharzadeh Alvar A.<sup>3</sup>, Soleimani N.<sup>4</sup>, Sadeghi S.<sup>5</sup>

## ABSTRACT

**Background:** Given the extensive use and preferred diagnostic method in common mammography tests for screening and diagnosis of breast cancer, there is concern about the increased dose absorbed by the patient due to the sensitivity of the breast tissue.

**Objective:** This study aims to evaluate the entrance surface air kerma (ESAK) before irradiation to the patient through its estimation.

**Material and Methods:** In this descriptive paper, firstly, a phantom was used to measure some data, including ESAK, Kvp, mAs, HVL, and type of filter/target. Secondly, the MultiLayer Perceptron (MLP) neural network model was trained with Levenberg-Marquardt (LM) backpropagation training algorithm and finally, ESAK was estimated.

**Results:** Based on results obtained from the program in different neuron numbers, it was found that the number of 35 neurons is the most optimal value, offering a regression coefficient of 95.7%. The Mean Squared Error (MSE) for all data was 0.437 mGy and accounting for 4.8% of the output range changes, predicting 95.2% accuracy in the present research.

**Conclusion:** Using neural networks in ESAK prediction, the method proposed in the present research leads to the possible ESAK estimation of patients before X-Ray. The results suggested that the regression coefficient represented 4.3% difference between the kerma measured by solid-state dosimeter in the radiation field and the value predicted in the research. In comparison with the Monte-Carlo simulation method, this method has better accuracy.

## Keywords

Mammography; Neural Networks, Computer; Radiation Dosimeters

## Introduction

Breast cancer is one of the most common cancers among women, resulting in half of million deaths each year, and also is the second death-causing cancer among all cancers in women [1]. Researchers believe that early diagnosis of breast cancer can decrease death by up to 30% [2].

To diagnose this disease, a diagnostic test such as the preparation of a mammography image is proposed, which causes the tumor to be diagnosed in the preliminary stages before the disease symptoms exhibited [3]. Given the extensive use of mammography for screening and

<sup>1</sup>MSc, Department of Medical Physics and Biomedical Engineering, Faculty of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>2</sup>PhD, Department of Medical Physics and Biomedical Engineering, Faculty of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>3</sup>PhD candidate, Department of Medical Physics and Biomedical Engineering, Faculty of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>4</sup>MD, Faculty of Medicine, Golestan University of Medical Sciences, Gorgan, Iran

<sup>5</sup>MD, Faculty of Medicine, Islamic Azad University, Tehran Medical Branch, Tehran, Iran

\*Corresponding author:  
M. R. Deevband  
Department of Medical Physics and Biomedical Engineering, Faculty of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran  
E-mail: mdeevband@sbm.ac.ir

early diagnosis of breast tumors, there is concern about the increased dose absorbed by the patient due to the sensitivity of breast tissue. Thus, the mean glandular dose through its estimation can help to recognize the level of absorbed dose [4].

Several methods have been presented to calculate the mean glandular dose, which most of them are based on Monte-Carlo simulation. For this purpose, ESAK is measured using a dosimeter and then the necessary conversion factor coefficients besides the calculated parameters are obtained through interpolation from tables via Monte-Carlo simulation method in different studies [5, 6]. The disadvantage of the mentioned method is calculating the mean glandular dose that it is necessary to have the result of breast tissue dosimetry. According to the statistics available, in 2017 in Iran, 722 mammography devices are available that 14 centers are equipped with suitable measurement devices. Thus, the methods based on the results of measurement for dose estimation cannot be efficient.

Another method is the simulation of a mammography device by the details of the mammography device using Monte-Carlo code. The disadvantage of Monte-Carlo is that considering the difference in devices, separate simulations should be used for any center and device, which is difficult and complex. In addition, since the devices are relatively old in our country, the function of devices will have different nominal features, which are not included in the simulation [7, 8]. In research by Mohammadi et al. on mammography examinations by ThermoLuminescent Dosimeter (TLD) and Monte-Carlo simulation methods, it was found that there is a difference in doses absorbed in the breast tissue between 7.5 and 17% for these two methods (measurement and computational) [9].

To resolve these issues and achieve an accurate and efficient estimation, the researchers decided to use a neural network for kerma estimation. The advantage of Artificial Neural

Networks (ANN) is good accuracy, as they are trained by the data of different centers [10] and in a large number and measure the air kerma automatically so that in the mentioned model, there is no need to dosimeter and trained operators and is also more user-friendly. Furthermore, it decreases the computational error and the difference between the absorbed dose measured by dosimetry and computational method of Monte-Carlo simulation method [11].

## Material and Methods

The proposed method includes collecting samples, preprocessing, training neural networks and finally, evaluation of networks. In this experimental study, 224 samples output from 32 mammography centers were collected throughout the country. To measure the data, a phantom with features similar to the breast tissue contents (fat/adipose ratio: 50-50%) was used [12]. Furthermore, the solid-state detector was used to measure the air kerma and Half-Value layer (HVL). Concurrent with recording and measuring ESAK on the phantom equivalent to the breast tissue and HVL, Kilovoltage peak (Kvp), milliAmperage-seconds (mAs), and type of filter/target were also recorded considered as the confounding factors in HVL of the device. HVL depends on the anode voltage, filter, anode material, and tube age. Thus, the type of brand and life span of any device are observed in HVL [13], and this parameter applied to the network as input. The network designed in the present research consisted of six inputs, including the device brand, Kvp value, mAs value, type and material of the filter/target, the total filter thickness, and HVL, which were coded as variables in the network input. The output was ESAK of the breast. After collection, the data were categorized into three groups: training, validation, and test sets. The training data accounted for 70% of the entire data, with the validation and test sets each capturing 15% of all data. To separate them, the simplest method, which was a selection of each set randomly, was used.

Artificial neural networks learn the general rules based on calculations on numerical data or examples. When there is a necessity to use prediction methods with the minimum error and maximum reliability, artificial neural networks are used. The advantages of ANNs over statistical methods include no limitation in the number of inputs and outputs, insensitivity to sudden changes in the data, the potential of modeling highly nonlinear behaviors, fast training, and a process to prevent overfitting [14, 15].

In calculating ESAK, a multilayer perceptron was utilized. The MLP network with the back-propagation learning method is one of the most common practical networks. Various studies have demonstrated that the MLP network is of supervised learning type, and, in case of proper selection of suitable internal structure, it can model and simulate any nonlinear system [16]. For training the neural network, a Levenberg-Marquardt back-propagation algorithm was used. This algorithm is indeed a transformation of the Newton method, designed to minimize functions which are in the form of a sum of squares of other nonlinear functions. Typically, Levenberg-Marquardt is used in multilayer networks, which have up to hundreds of weight and bias and in function approximation that the performance index is mean squared error, which is the fastest training method [17, 18].

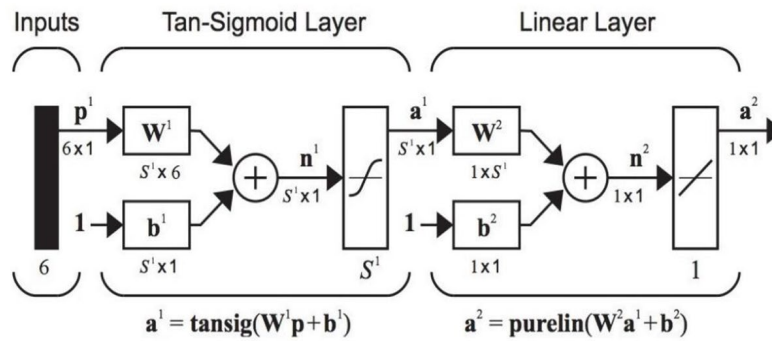
In multilayer networks, used for function approximation and pattern recognition, the number of hidden layers required in the network is not determined by the problem description; thus, any number of the hidden layer can be possible. The standard process is that the training begins with a network consisting of a hidden layer. If the performance of the bilayer network was not satisfactory, then the three-layer network can be used. The use of more than two hidden layers in ANNs is unconventional [18]. Hence, the bilayer network was also used in the present research. Furthermore, the number of neurons in each layer should also

be determined. The number of neurons in the output layer is equal to the size of the vector of the target data. Due to the presence of only one output, our network has had one output layer neuron. The number of neurons in hidden layers is determined based on the extent of the complexity of the problem, which has a minimum and maximum value. The minimum number of the first layer neurons in the bilayer neural network is experimentally obtained from Eq. (1); where  $n_1$  is the number of first layer input and  $n_0$  represents the number of output layer neurons, which are 6 and 1, respectively. Thus, the minimum number of first-layer neurons of the network is 14. The maximum number of first-layer neurons in the bilayer neural network is also obtained from Eq. (2); where  $k$  is the number of samples. Thus, the maximum number of the first-layer neurons of the network is 261 [18]. Furthermore, to train the network, the tangent sigmoid transfer function has been used (Figure 1).

$$2(n_1 + n_2) \quad (1)$$

$$\frac{k(n_1 + n_0) - n_0}{n_1 + n_0 - 1} \quad (2)$$

Sigmoid transfer functions are mostly used in multilayer networks trained by the back-propagation algorithm. One of its factors is the derivability of these functions [18]. In terms of optimization, learning in ANN is equivalent to the minimization of an error function as an index of the model performance. Among the most common and important performance indices of the model, one can mention the MSE between the output and target, and the regression index representing the correlation between the output and target [18]. Thus, to investigate the accuracy of the neural network and the accuracy of kerma estimation, in addition to the investigation of the model performance indices, the model results have been compared with the results of the reported pieces of research. The specifications of the opti-



**Figure 1:** The bilayer neural network with six inputs and tangent sigmoid transfer function and one output with purelin transfer function.

mized neural network in this study are summarized in Table 1.

### Results

After designing the network, training, and running the program, the network with different conditions and parameters was trained to achieve the best result. The error gradient, which is the function derivative in closing into zero, suggests that the function slope at that point has become zero. Thus, it shows the extremum of function with the meaning of a local minimum. Hence, it represents the convergence of the target values and output, and the program no longer needs to continue. Ac-

cordingly, the error gradient was considered  $10^{-7}$ . [19, 20].

In training the neural network in each iteration, we face a new network. For this reason and to increase the reliability of the network performance, for a certain number of neurons, we trained the network five times and the mean values of its evaluation index were regarded as the criterion [21]. As mentioned previously, the number of neurons in hidden layers is determined given the extent of the complexity of the problem, which for the present problem, this value was obtained between 14 and 261. By running the program for all of these neurons, it was found that at values above 60 neurons due to the complexity of the solution and computations, regression reached below 85% (Figure 2). Furthermore, the RMSE value for the test data reached over 1.73 at values above 60 neurons, which is around 19.05% of the range of output changes.

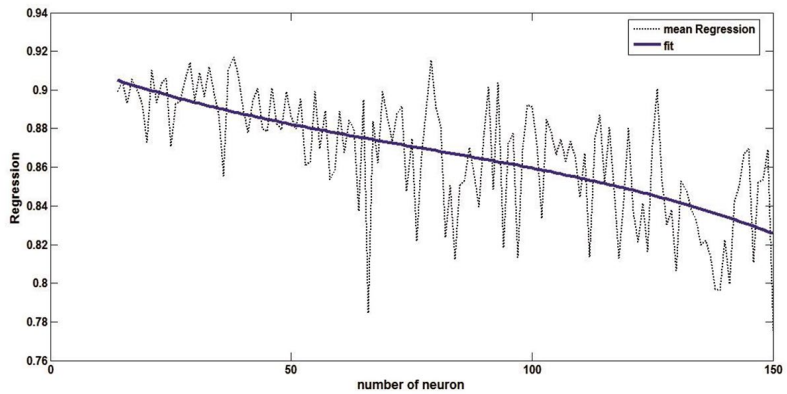
The variations of the MSE in relation to the number of neurons are presented in Figure 3. As seen in Figure 3, with the increase in the number of neurons, MSE value grows. Thus, a neural network above 60 neurons in the hidden layer is not optimal.

To achieve the most optimal number of neurons, the mean values of the network evaluation indices in relation to the number of hidden layer neurons are presented in Table 2.

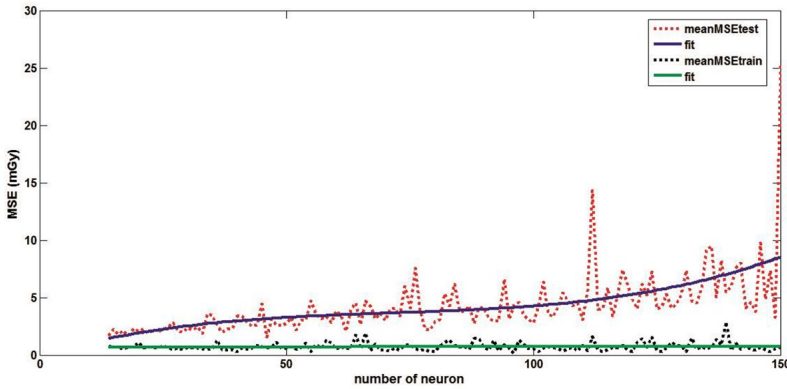
In the network with 35 neurons in the hid-

**Table 1:** Specifications of the optimized neural network.

NN architecture	MLP with 2 layers
Inputs: six inputs with tangent sigmoid transfer function	Kvp, mAs, type of filter target, total filter thickness, HVL, brand
Output: one output with linear transfer function	ESAK
train function	Levenberg-Marquardt
hidden layer size	35
divide function	Random
train, validation and test ratio	70/100, 15/100, 15/100
performance function	MSE



**Figure 2:** The output and target regression in relation to the number of neurons of the hidden layer.



**Figure 3:** The MSE in relation to the number of neurons of the hidden layer.

den layer, the network has the maximum regression between the output and target with a value of 94.9%. Furthermore, the mean MSE of the test and training data was obtained as 0.912 and 0.201 mGy, respectively, where 9.15% of the range of changes and 2.14% for the output were obtained. Thus, the number of 35 neurons is the best number of training the network. Again, the network was trained by 35 neurons in the hidden layer, whose results are presented in Figures 4 and 5. Figure 4 shows that MSE value for all data was obtained 0.437 mGy, accounting for 4.8% of the range of output variations, while the regression value between the output and target value is 95.7%. The mean and standard deviation of MSE in the histogram diagram is 0.03 and

0.66, respectively. Moreover, figure 5 shows that the MSE value for the test data is 0.503 mGy, accounting for 5.1% of the output variations, while the regression between the output and target value was obtained 95%. Also, the mean and standard deviation of MSE in the histogram diagram of test data are 0.19 and 0.69, respectively. Thus, using a trained ANN, one can estimate ESAK with a desirable accuracy (95.7%).

### Discussion

In this paper, the MLP neural network model was trained by LM training algorithm to predict air kerma based on measurable parameters. Based on Figures 4 and 5, it can be found that the regression coefficient between

**Table 2:** The mean values of the network evaluation indices in relation to the number of neurons of the hidden layer.

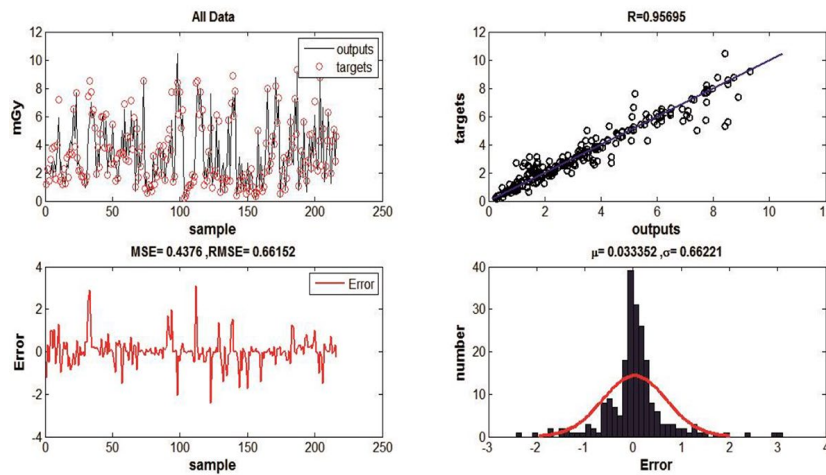
Neuron no.	MSEtrain value	MSEtest value	R (%)	Neuron no.	MSEtrain value	MSEtest value	R (%)
14	0.365	1.658	0.918	38	0.390	4.305	0.890
15	0.202	1.082	0.947	39	0.653	2.246	0.894
16	0.888	1.407	0.881	40	0.213	5.748	0.898
17	0.417	2.290	0.926	41	0.481	3.829	0.902
18	0.975	2.354	0.875	42	0.601	2.384	0.905
19	1.147	1.201	0.878	43	0.367	6.143	0.863
20	0.413	1.387	0.925	44	0.487	1.868	0.914
21	0.339	1.822	0.933	45	0.934	1.284	0.866
22	0.845	1.806	0.907	46	1.003	3.411	0.848
23	0.848	1.503	0.882	47	0.155	5.482	0.879
24	0.722	1.009	0.920	48	0.250	2.186	0.935
25	1.724	2.282	0.799	49	2.590	4.183	0.789
26	0.797	1.368	0.910	50	0.565	1.469	0.905
27	0.652	1.413	0.916	51	0.380	3.649	0.901
28	0.394	2.697	0.920	52	0.203	1.770	0.895
29	0.0820	1.583	0.824	53	0.547	1.825	0.896
30	2.270	3.540	0.816	54	0.186	4.025	0.887
31	0.149	2.919	0.937	55	0.212	1.229	0.946
32	0.363	3.181	0.908	56	0.458	5.519	0.876
33	0.304	2.438	0.904	57	0.627	2.525	0.908
34	0.419	1.947	0.933	58	0.242	1.479	0.920
35	0.201	0.912	0.949	59	0.966	4.110	0.863
36	0.497	1.578	0.925	60	0.803	0.925	0.891
37	0.290	5.239	0.897				

the output and target in the research is 95.7%, representing 4.3% difference between the air kerma measured by solid-state dosimeter and the simulation performed in this research. Furthermore, the MSE between the measured and simulated values is 0.437 mGy, accounting for 4.8% of the range of variations in the measured air kerma. The test conducted on the collected data generated positive results. In all cases, the regression correlation factor was over 94% (Figure 6).

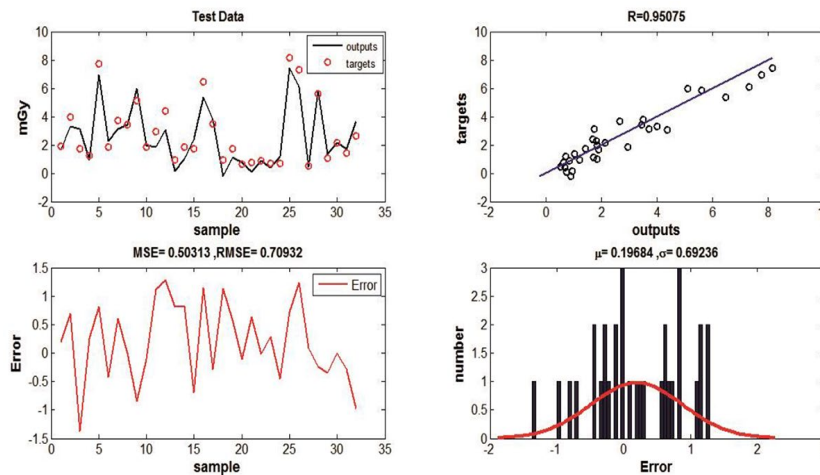
## Conclusion

There have not been more studies on this type

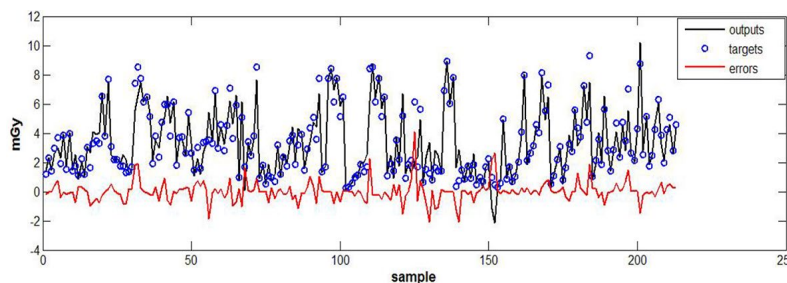
of approach in predicting the air kerma. On the other hand, extensive studies have been done by Monte-Carlo simulation method, which had 7.5-17% difference between the values measured by the dosimeter and simulated by the Monte-Carlo method. Although this approach is not originally a standard method for determining the air kerma in common mammography, the current method of using neural networks in predicting air kerma makes it possible to estimate the patient's possible air kerma before being exposed to X-ray. This approach can be a key step in the development of such neural network systems, as it can be



**Figure 4:** Top-left: Comparing the output and target for all data; bottom-left: error for all data  
 Top-right: comparing the output and target regression for all data; bottom-left: histogram diagram of the error of all data.



**Figure 5:** Top-left: Comparing the output and target for test data; bottom-left: error for test data  
 Top-right: comparing the output and target regression for test data; bottom-left: histogram diagram of the error of test data.



**Figure 6:** Comparing the neural network estimation (black), the collected data (blue), and the error value (red).

trained by more data to achieve even better results.

## Acknowledgment

This paper is taken from the project related to thesis No. 326, which has been supported by Shahid Beheshti University of Medical Sciences. The authors would like to acknowledge the Department of Biomedical Engineering and Medical Physics, Faculty of Medicine, Shahid Beheshti University of Medical Sciences.

## Conflict of Interest

None

## References

1. National Cancer Institutes. c2018. [Accessed August 21, 2018]. Available from: <https://www.cancer.gov/>.
2. Tabar L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology*. 2011;**260**:658-63. doi: 10.1148/radiol.11110469. PubMed PMID: 21712474.
3. NZ health statistics [Internet]. New Zealand National Screening Unit Website. c2018. [Accessed August 21, 2018]. Available from: <https://www.health.govt.nz/nz-health-statistics>.
4. Dance DR, Skinner CL, Carlsson GA. Breast dosimetry. *Appl Radiat Isot*. 1999;**50**:185-203. PubMed PMID: 10028637.
5. Dance DR. Monte Carlo calculation of conversion factors for the estimation of mean glandular breast dose. *Phys Med Biol*. 1990;**35**:1211-9. doi: 10.1088/0031-9155/35/9/002. PubMed PMID: 2236205.
6. Sobol WT, Wu X. Parametrization of mammography normalized average glandular dose tables. *Med Phys*. 1997;**24**:547-54. doi: 10.1118/1.597937. PubMed PMID: 9127307.
7. Nigaprucke K, Puwanich P, Phaisangittisakul N, Youngdee W. Monte Carlo simulation of average glandular dose and an investigation of influencing factors. *J Radiat Res*. 2010;**51**:441-8. doi: 10.1269/jrr.10008. PubMed PMID: 20523013.
8. Ko K, Park S, Lee J. Assessment of patient dose in mammography using Monte Carlo simulation. *J Nucl Sci Technol*. 2004;**41**:215-8.
9. Mohammadi A, Faghihi R, Mehdizadeh S, Hadad K. Total absorbed dose of critical organs in mammography, assessment and comparison of Monte-Carlo method and TLD. *Biomed Tech*. 2005;**50**:393-4.
10. Ceke D, Kunosic S, Kopric M, Lincender L. Using neural network algorithms in prediction of mean glandular dose based on the measurable parameters in mammography. *Acta Informatica Medica*. 2009;**17**:194.
11. Mohammadyari P, Faghihi R, Mosleh-Shirazi MA, Lotfi M, Hematiyan MR, Koontz C, et al. Calculation of dose distribution in compressible breast tissues using finite element modeling, Monte Carlo simulation and thermoluminescence dosimeters. *Phys Med Biol*. 2015;**60**:9185-202. doi: 10.1088/0031-9155/60/23/9185.
12. Highnam R [Internet]. Patient-Specific Radiation Dose Estimation in Breast Cancer Screening Keeping Patients Safe and Informed 2018. [Accessed 21 August 2018]. Available from: <https://www.volpara.com/assets/Uploads/VolparaDose-White-Paper.pdf>.
13. Ariga E, Ito S, Deji S, Saze T, Nishizawa K. Determination of half value layers of X-ray equipment using computed radiography imaging plates. *Phys Med*. 2012;**28**:71-5. doi:10.1016/j.ejmp.2011.01.001.
14. Haykin S. Neural networks: a comprehensive foundation. United States: Prentice Hall PTR; 1994.
15. Alvar AA, Deevband MR, Ashtiyani M. Neutron spectrum unfolding using radial basis function neural networks. *Appl Radiat Isot*. 2017;**129**:35-41. doi:10.1016/J.APRADISO.2017.07.048.
16. Anderson JA. An introduction to neural networks. Germany: MIT press; 1995.
17. Hagan MT, Menhaj MB. Training feedforward networks with the Marquardt algorithm. *IEEE transactions on Neural Networks*. 1994;**5**:989-93. doi: 10.1109/72.329697.
18. Hagan MT, Demuth HB, Beale MH, De Jess O. Neural network design (2nd Edition). Martin Hagan; 2014.
19. Iyer MS, Rhinehart RR. A method to determine the required number of neural-network training repetitions. *IEEE Transactions on Neural Networks*. 1999;**10**:427-32. doi: 10.1109/72.750573.
20. Fukumizu K, Amari S. Local minima and plateaus in multilayer neural networks. 1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470), Edinburgh, UK: IET; 1999. doi: 10.1049/cp:19991175.
21. Hamm L, Brorsen BW, Hagan MT. Comparison of stochastic global optimization methods to estimate neural network weights. *Neural Process Lett*. 2007;**26**:145-58. doi: 10.1007/s11063-007-9048-7.