Volume 20, No 1 | International Journal of Radiation Research, January 2022

Radiation dose for external exposure to gamma-ray using artificial neural network and MC simulation

K. Elhamdi*, M. Bhar, K. Belkadhi, K. Manai

Nuclear Physics and High Energy Research Unit, FST, Campus Universitaire El-Manar, 2092 El Manar Tunis

► ***Original article***

Corresponding author:
Dr. Karima Elhamdi,
E-mail:
karima.elhamdi@gmail.com

Received: July 2020
Final revised: March 2021
Accepted: April 2021

Int. J. Radiat. Res., January 2022; 20(1): 199-204

DOI: 10.52547/ijrr.20.1.30

Keywords: *Gamma ray, soil, exposure, dose, artificial neural network.*

ABSTRACT

Background: The computation of the absorbed dose in air allows the estimation of the concentrations of radionuclides in the soil and the assessment of the external exposure of the human body. The development of numerical models describing gamma ray transport in the environment provides more precise methods to analyze the pathways of external radiation dose. ***Material and Method:*** A combined method using Artificial Neural Network (ANN) and Monte Carlo Simulation (MC) has been developed to calculate the absorbed dose rate in air for photon emitters from natural radionuclides. We proposed a new class of trained ANN to GEANT4 to calculate the probability, for generated photon sources, to reach the detector. Only photons with high probability were tracked in MC Simulation. ***Results:*** A significant reduction of computation time was reached. Unscattered flux and gamma-dose-rate conversion factors were calculated and compared to previous works. ***Conclusion:*** The use of this method overcomes the problem of the long duration of computation time, obtaining a good agreement with previous works and efficient results of the dose rate conversion factor.

INTRODUCTION

Natural radioactivity is still the major source of radiation exposure that comprises cosmic radiation and the radiation arising from the primordial radioactive elements in the ground. To estimate the absorbed dose rates received at 1 m above the soil due to photon emitters derive from soil [1-2] (the natural radioactive series of ^{238}U, ^{232}Th and ^{40}K) two basic methods have been followed. The first one calculates the gamma-ray field by solving the Boltzmann transport equation [3, 4]. The second one uses Monte Carlo techniques in order to obtain information regarding the energy and angular characteristics of the radiation field in air and to calculate the absorbed dose rate in air. In the literature, several soil-detector geometries [5–8] have been proposed to achieve better accuracy for the conversion factor calculation in Monte Carlo simulation.

Saito *et al.* [5] assumed the ground and the air as infinite smooth planes contacting each other. The ground was taken up to 300m down and the air was taken up to 200m height. On the other hand, Likar *et al.* [6] generated the primary photons from a box as large as 2km. The absorbed dose rate was calculated in a box of air as large as 2km. For Clouvas *et al.* [7], radionuclides was distributed in a cylindrical of 40m radius and 1m depth. The dose was calculated at 1m above the ground. To calculate the dose rate he used two virtual detectors, a point detector and a sphere detector. Askri *et al.* [8] have proposed a theoretical model that optimizes the soil geometry implemented in GEANT4. The detector, used by Askri *et al.*, is a disc detector of radius 2m placed 1m above the soil level. The optimized geometry used by [8] is not easy to implement in GEANT4 and the disc detector used in this work hasn't the adequate geometry as we demonstrate.

The literature shows the CPU time of MC simulation is a limitation that has to be resolved. In fact, by choosing huge geometry for the soil a significant part of the generated photons will be tracked even though it will not contribute to the detected radiation. Thus, simulating photons which do not contribute to the detected radiation leads to an enormous loss of CPU time. On the other hand, choosing an arbitrary small geometry size for the soil eliminates photons that will actually contribute to the absorbed dose at 1m above the soil. Askri [8] solved this major shortcoming by determining optimized source geometry based on mathematical model and on physical assumptions that was later implanted in Monte Carlo simulation. In this work, we propose a combined method based on MC simulation and an Artificial Neural Network in order to calculate the dose in record time maintaining its efficiency. Our method presents an alternative convenient approach that simulates the absorbed dose rate and declines the necessity of determining the optimized source

[DOI: 10.52547/ijrr.20.1.30]

geometry before the simulation.

MATERIAL AND METHODS

In this work, we present a new approach in which an ANN is combined to MC simulation to define the soil part that its emitted radiation has a high probability to reach the detector. Indeed, in GEANT4, the *PrimaryGeneratorAction* class generates primary photons randomly in the source area and the code tracks it until it disappears. Using a trained ANN, the generation and the tracking of photons will be controlled. The ANN gives a prediction (P(x,y,z)) for the generated photon position (x,y,z) to reach the detector. To reduce the computation time and on the basis of applying trained neural networks, only photons having high probability to escape the soil and reach the detector are generated and tracked by GEANT4. Photon with low probability is killed and another photon is generated.

Monte Carlo simulation

Monte Carlo simulation was done using the free CERN code GEANT4 [9]. GEANT4 simulates interactions of particles with matter. The code uses C++ language and comes with basic concepts of the user's initialization and action classes.

The geometry (air, soil and detector) is incorporated in GEANT4 using G4DetectorConstruction class. Primary particles are implemented in G4Primary-Particle class. Where each primary particle is described by the coordinates of its position (x, y, z), by its momentum direction and by its energy. In MC simulation all primary particles created in the source are tracked until the particle loses all its kinetic energy, disappears due to an interaction or reaches the boundary of the simulation volume.

Three sets of photons are generated by GEANT4. Primary photon positions (x, y, z) are saved in a root file. The first set of photons is used for the training of the ANN and the second one is used for ANN architecture optimization. The third set of photons is used for accuracy estimation and for unscattered flux and dose rate conversion coefficients calculation.

Artificial neural network description

ANN is nonlinear computational algorithms inspired from biological neural systems. ANNs have been applied in the last years to a wide range in physics [11], electronics, economics [12]... An ANN is composed of an input and output node layers connected through a number of hidden layers in between. The input layer receives the input signals x_j. Then, the inputs are normalized and weighted by w_{kj}. Each node is occupied by a nonlinear information-processing unit called neuron. The nodes in each layer are connected with all neurons of the neighboring layer(s) and propagate information with a weight specifying the strength of the inter-neuron couplings [13]. The input signal connected to a neuron, is multiplied by a synaptic weight. The sum of the weighted input signals is used as a parameter of the activation function. This function, usually a sigmoid, allows limiting the output amplitude of the neuron. In mathematical terms, the artificial neuron calculation is described by equations (1) and (2):

$$u_k = \sum_{j=0}^{n} w_{kj}\ x_j \tag{1}$$

$$y_k = \varphi(u_{k+} b_k) = \varphi(v_k) \tag{2}$$

where w_{kj} are the connections weights and b_k is a bias and $\varphi(v) = 1/(1 + \exp(-v))$ for this work.

To obtain the desired output from the network, the neurons in the input layer receive input data, normalize them and propagate them successively to the hidden layers and then to the nodes in the output layer. The adjustment algorithm is referred as Learning or Training considering the comparison between the output of the network and the desired target corresponding to the training sample. One of the methods widely used for its performance is called "learning by epoch". It consists first in the summation of information for the whole pattern and then updates the weights. Each update minimizes the summed cross-entropy cost function error. The performance of an ANN is given by the accuracy of prediction measured by this error and the convergence of the learning process. As an output of the ANN, we have two curves of errors. The first one represents the error calculated for the training sample called training error and the second corresponds on the error calculated to another independent test sample not used in the training, called "generalization error". The purpose of the learning process is to obtain a low generalization error characterizing the performance of the ANN to model new data not used in the training.

This work is performed using a layered ANN type under supervised training scheme that contains a number of hidden units. This type of the ANN models showed a good agreement and has been cited as reliable universal approximators. ANN code has the ability to include the needed libraries for physics data analysis. And it is running on ROOT [14], a CERN free data analysis toolkit based on C++.

Hardware platform

The hardware platform used to run the GEANT4 code version 10.1 and ROOT framework version 5.34.0.0 is a free Scientific Linux (SL4.5) installed on a VMWare workstation machine with 4Gb RAM and a 2.5GHz CPU.

[DOI: 10.52547/ijrr.20.1.30]

RESULTS

Adequate parameters of the detector and soil

In the literature, several soil geometries and sizes [5–8] have been proposed to achieve better accuracy for the calculation of the conversion factor. Taking into account symmetry considerations, the adequate soil geometry for the simulation of the external exposure to photons is found to be cylinder. Thus, the soil geometry was implanted as cylindrical shape in Geant4 for the rest of this work. The dimensions of this cylindrical source need a careful investigation in order to define the optimum emission photon geometry size. For that purpose the unscattered flux ($\gamma cm^{-2}/Bqkg^{-1}$) was calculated using GEANT4 code for 3 different photon energies (E = 0.2, 1, 3M eV) for soil radii varying from 3m to 3000m and soil depths from 1cm to 10m. We would like to mention that the photon flux is composed by the unscattered flux (due to photons reaching the detector without any interaction with the soil and the air above) and the scattered flux (due to photons reaching the detector after scattering in air and mainly in soil). The study in this section is done using the unscattered photons because it contributes to accurate results from a statistical point of view.

Figure 1 presents the normalized unscattered flux as function of soil depth and figure 2 as function of the soil radius. The normalized unscattered flux is calculated by dividing the unscattered flux for each depth (radius) by the unscattered flux for the depth z = 10m (radius r = 3000m). To not discard photons, the choice of the soil geometry size should depend on the energy of the emitted photons.

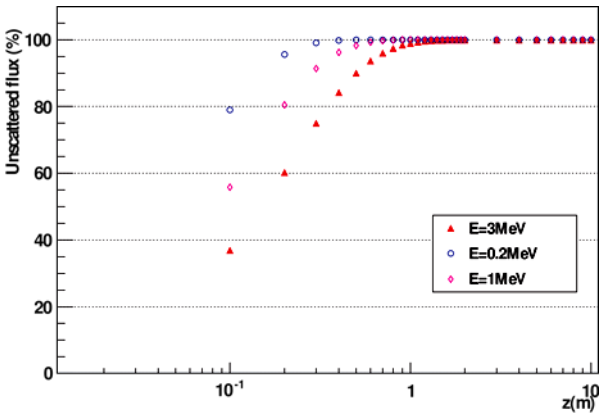

Figure 1. The normalized unscattered flux for different soil depths.

Saito, Likar and Clauvas, [5–7] have calculated the photon flux energy at 1m above the soil by different detector geometries. To study the effect of the detector geometry on the unscattered flux, we generate photons with primary energy of 1.5MeV randomly from the soil and we calculate the unscattered flux reaching two different detector geometries placed at 1m above soil: the first one is a sphere [7] of radius of 1m and the second is a disc [8] of the same radius. Figure 3 illustrates the results of the comparison between the normalized unscattered flux for the two geometries as function of the the primary emitted photon position. As seen in this figure, primary photons generated with a large distance from the detector have a low probability to reach the disc detector due to the small solid angle. The spherical shape with a high symmetry was selected in the remainder of this work.

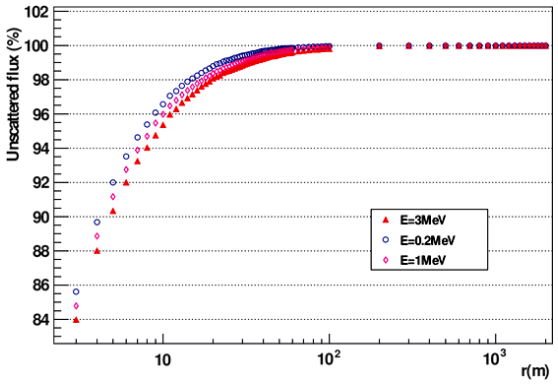

Figure 2. The normalized unscattered flux for different soil radii.

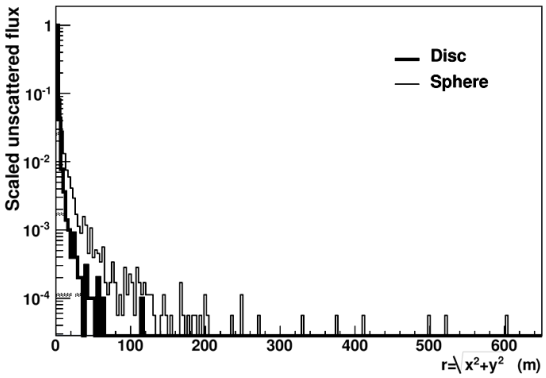

Figure 3. The unscattered flux as function of the position r of the primary emitted photon.

Training of the artificial neural network

In the training step the position (x, y, z) of the emitted photon sources was generated in the PrimaryGeneratorAction class and saved in a root file. A first set containing 10000 position values of photons reached the detector and 10000 position values (x, y, z) of photon sources not reached the detector is used for the training of the ANN. Based on (x, y, z) as inputs, the trained ANN predicts the probability P(x; y; z) that a photon reaches the detector. The ANN meets the validation requirements when it allows to reproduce a probability close or equal to 1 for photons reached the detector and close or equal to 0 for photons not reached the detector.

The architecture of the ANN (the number of layers and of nodes in each layer) is a crucial point in the construction of an Artificial Neural Network. The input layer, of our ANN, contains 3 neurons corresponding to coordinate (x, y, z) of the photon

[DOI: 10.52547/ijrr.20.1.30]

sources. The output layer of the ANN contains a single neuron assigned to the probability P(x, y, z) that the emitted photon reaches the detector.

To choose the appropriate number of hidden layers, different architectures have been tested (two, three and four hidden layers with various numbers of neurons in each hidden layer). The ANN architecture with two hidden layers with 6 and 4 neurons have validated the requirements. Training was performed for different architectures using as input to the ANN, the (x, y, z) positions of the first set of photons. A second set containing 10000 position values of photons reached the detector and 10000 (x, y, z) position values of photon sources not reached the detector is used to test the performance of the ANN. The position (x, y, z) for each photon is introduced to the trained Artificial Neural Network as input and then the ANN predict the probability for this photon to reach the detector. For each architecture the root-mean-square error (RMSE) is calculated using equation (3).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(P^i - P_{ex}^i)^2}{N}} \quad (3)$$

where N is the photons number, P^i is the output of the ANN (the probability that the emitted photon (i) reaches the detector) and P_{ex}^i expected probability (P_i = 1 for photons that have reached the detector and 0 for photons that did not reach the detector).

Figure 4 shows a histogram of the logarithmic distribution of the ANN output P(x,y,z) with two hidden layers (6 and 4 neurons) predicted for the second set of photons. As we see, there is a good separation between calculated probability by the ANN for photons reached the detector and photons not reached the detector. This separation authenticates the performance of ANN which allow us to adopt this technique.

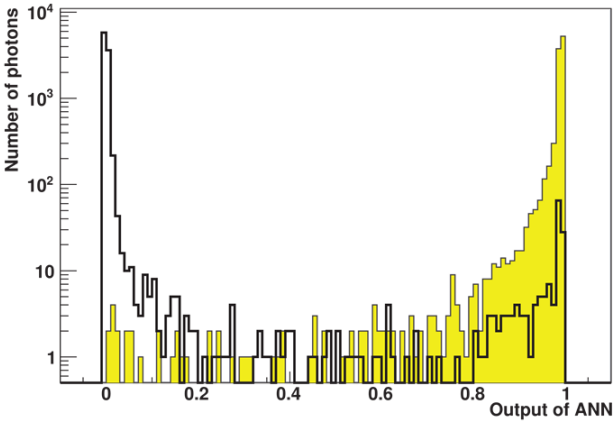

Figure 4. Logarithmic distribution of the output of the ANN. Events corresponding to photons reached the detector is presented in yellow.

ANN-MC: A combined method for dose rate conversion factor calculation

The trained Artificial Neural Network is combined with MC simulation to select and track only photons with high probability to reach the detector. A new class for the trained ANN is added to GEANT4. The position (x, y, z) of primary photon sources is randomly generated in the *PrimaryGeneratorAction* class. This position is used as input to the ANN class. The probability to reach the detector is then calculated using the trained ANN. If the probability is greater than a selected threshold, the generated photon is tracked, else, we return to the *PrimaryGeneratorAction* class to choose another (x, y, z) position. Figure 5 shows the probability to reach the detector calculated by the ANN for 108 photons generated in the *PrimaryGeneratorAction* class.

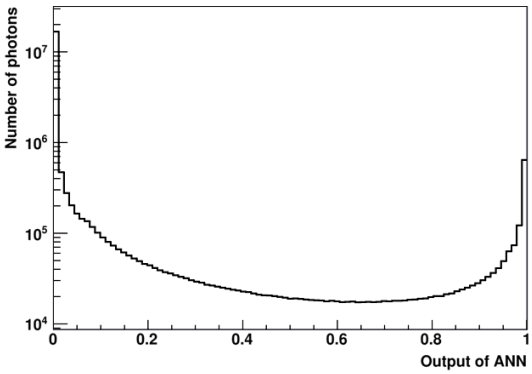

Figure 5. The output of the ANN calculated for photons generated in thePrimaryGeneratorAction class.

Figure 6. (a) shows the position z as function of x for photon sources with a probability higher than three selected thresholds (Th_1= 0.1, Th_2= 0.5, Th_3 =0.9), the number of selected photons decreases with the threshold.

Figure 5. (b) Shows the position z as function of x for Th_2 = 0.5. We draw in the same figure the surface limiting the optimized volume (red line), developed by Askri *et al.* [8], and defined by:

$$r(z) = ((\hat{\mu}h + d)^2(\frac{h-z}{\hat{\mu}h-z})^2 - (h-z)^2)^{0.5} \quad (4)$$

For primary photon energies above 100 keV, $\hat{\mu}$= $\mu_{air}/\mu_{soil} \approx 10^{-3}$. The maximum depth reached by photons for Th_2 = 0.5 is d = 1.3m.

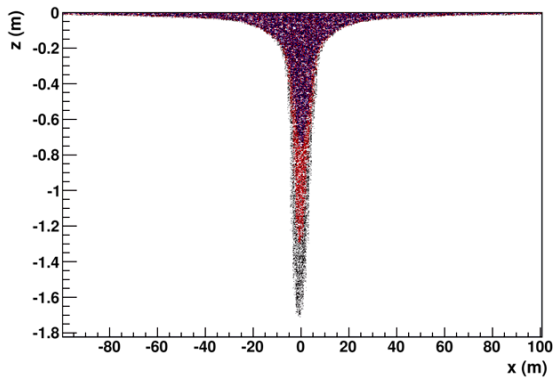

Figure 6. Top: The coordinate z as function of x for photons with a probability to reach the detector higher than three selected thresholds (Th_1 = 0.1 (black), Th_2 = 0.5 (red), Th_3 = 0.9 (blue)).

[DOI: 10.52547/ijrr.20.1.30]

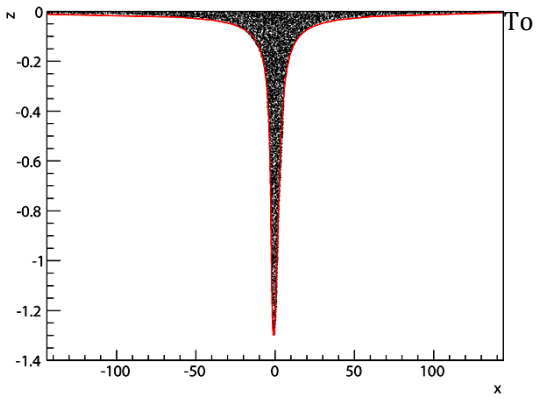

Figure 7. The position z as function of x for Th2 = 0.5 limited by the optimized volume proposed by Askri *et al*.

check our method and to choose the adequate threshold, two tests are performed. In the first one, we calculate the unscattered flux (ϕ) for 108 primary pho- tons at (3 M eV) generated randomly in a soil cylinder (r = 218m; d = 2m). All generated photons are tracked by GEANT4 from the starting point to the stopping point. In the second one, we generate the same number of the primary photons with the same energy and the same geometry. Photons are sent to the trained ANN to calculate the probability to reach the detector. Only photons with a probability higher than a fixed threshold (from 0.1 to 0.9 by a step of 0.2) are tracked by GEANT4. We note ϕ_{Th}, the unscattered flux calculated for each threshold. Table 1 shows the relative error Δϕ for five thresholds. Δϕ is calculated using formula (5):

$$\Delta\Phi = \frac{|\Phi - \Phi_{Th}|}{(\Phi + \Phi_{Th})} \times 100 \qquad (5)$$

Table 1. Presents the relative error Δφ for different threshold.

Threshold	Δφ(%)
0.1	0.9
0.3	1.4
0.5	2.6
0.7	3.2
0.9	4.5

Figure 8 shows analogy between the unscattered flux values calculated using ANN-MC and the standard values published in the literature.

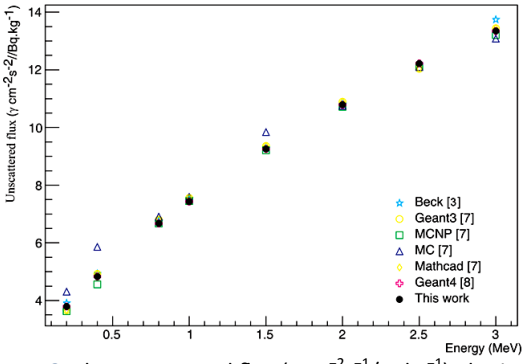

Figure 8. The unscattered flux ($\gamma cm^{-2}s^{-1}/Bqkg^{-1}$) obtained using ANN-MC method compared to previous works. Statistical discrepancy does not exceed 1%.

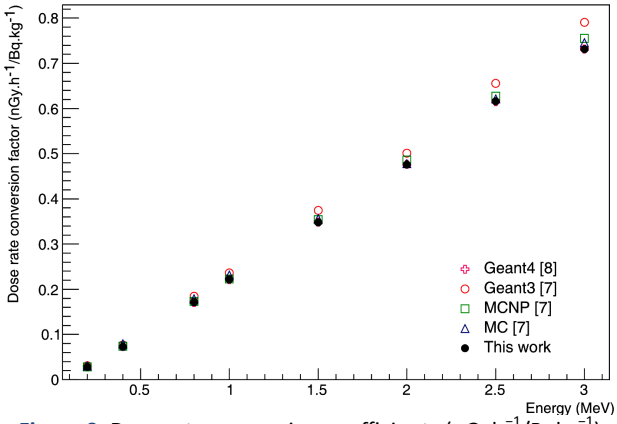

Figure 9. Dose rate conversion coefficients ($nGyh^{-1}/Bqkg^{-1}$) calculated with ANN-MC method, compared to previous works.

DISCUSSION

The results of the size source choice showed that the dimensions of soil should not be the same for all energies. However, previous studies [5,6,7,8] used the same soil size for all energies which increase the statistical error of the work. For instance, the radius and depth (r; d) of the cylinder soil, corresponding to a normalized unscattered flux= 99.9%, are (r = 280m; d = 2m) for 3 MeV energy and (r = 217m; d = 1.15m) for 1MeV energy and (r = 178m; d = 0.61m) for 0.2 MeV energy. Generally speaking our results should be adopted in Monte Carlo studies for soil parameters source in order to insure the accuracy of the results. Accordingly, in the remainder of this work, the simulation will be conducted with the proper radius and depth associated to every energy. The choice of the detector shape is another parameter that has to be studied. A comparison between disc detector and a sphere detector was made. The outcome has showed that primary photons generated with a large distance from the detector have a low probability to reach the disc detector which pinpoint that the choice of detector is another reason of the statistical error between different results.

For the rest of the work, we present a new approach in which an ANN is combined to MC simulation to determine photons that have significant probability to reach the detector. The result of the preliminary study that investigate the performance of trained ANN showed a good separation between photons that reached the detector and photons that did not which is a good indication that ANN gives a good prediction for the generated photon position (x,y,z).

Our method based on MC simulation and Artificial Neural Network has the advantage to calculate the dose in record time maintaining its efficiency along with abandoning the approach of determining the optimized source geometry before the simulation like in Askri work[8]. Indeed, Askri have made analytical calculation to define the volume that has significant contribution in the simulation,

[DOI: 10.52547/ijrr.20.1.30]

and photons were only generated in that volume. By drawing the position z as function of x of photons ANN chosen to track, we have found the same optimized shape proposed by Askri. This result is considered an affirmation that ANN prediction is correct.

A study of relative error Δϕ is necessary to point out the value of our work. The unscattered flux values calculated using ANN-MC and the standard values published in the literature showed analogy and that proves a good performance of our method.

Comparison is performed against values obtained by Askri using GEANT4 and Clauvas using GEANT3 and MCNP [8]. The comparison is also done with the theoretical calculation of Beck and the theoretical calculation of Clouvas using MathCad. The agreement with Askri (GEANT4) and Clauvas (MCNP) calculations is rather good with a discrepancy less than 1%. Values of dose rate conversion coefficient show a good agreement within the statistical error with the benefit of a significant time gaining using our method [8]. We mention for the calculation of the unscattered flux and the dose rate conversion factor, for a threshold of 0.1 corresponds ≈ 97% of gain in computation time.

CONCLUSION

We developed a new method based on ANN and MC simulation to calculate the absorbed dose rate in air. An ANN was trained and successfully implemented in GEANT4 system. ANN was used to calculate the probability for each generated primary photon, to reach the detector. Only photons with significant probability to reach the detector are tracked by GEANT4. The dose rate conversion coefficients were calculated for different energies and a good agreement with previous works was obtained. Due to reduction of the number of tracked photons in MC simulation, substantial gain in computation time was obtained (≈ 92%). The new method has solved the efficiency problem. It is important to mention that this method can be adopted by other problem of dose calculation.

ACKNOWLEDGMENTS

I would like to express my gratitude to Mr Kais Manai who guided me and offered deep insight into the project. I would also like to thank M. Bhar and K. Belkadhi for their contributions.

Ethical considerations: *The rights of research participants are protected. Scientific integrity is maintained. There is no deception or exaggeration about the aims and objectives of the research. There is no type of misleading information.*

Conflicts of interest: *None declared.*

Funding: University of Tunis El Manar

Author contributions: (K.E): Conceived and designed the simulation method, performed the analysis and wrote the paper. (M.B): Collected the data. (K.B): Contributed analysis tools. (K.M): Contributed data and analysis tools.

REFERENCES

1. Larionova NV, Panitskiy AV, Kunduzbayeva AY, Kabdyrakova AM, Ivanova AR, Aidarkhanov AO (2021) Nature of radioactive contamination in soils of the pine forest in the territory adjacent to Semipalatinsk test site. *Int J Radiat Res, **19(1)**: 113-120*.
2. Arafin AK, El-Taher A, Fazlul Hoque AKM, Ashraful Hoque M, Ferdous J, Joynal Abedi M (2020) Natural gamma radiation level detection in agriculture soil after Aila disaster and comparison with deep soil gamma activity in a specific area of Sundarban region, Satkhira, Banglsdesh. *Int J Radiat Res, **18(3)**: 397-404*.
3. Beck HL and de Planque G (1968) The radiation field in air due to distributed gamma-ray sources in the ground. Environmental Measurements Lab. Report, HASL-195.
4. Kocher DC and Sjoreen AL (1985) Dose rate conversion factors for external exposure to photon emitters in soil. *Health Phys, **48**: 193-205*.
5. Saito K and Jacob P (1995) Gamma ray fields in air due to sources in ground. *Radiat Protect Dosim, **58**: 29-45*.
6. Likar A, Vidmar T, Pucelj B (1998) Monte Carlo determination of gamma-ray dose rate with the Geant system. *Health Phys, **75**: 165-169*.
7. Clouvas A, S Xanthos, Antonopoulos-Domis M, Silva J (2000) Monte Carlo calculation of dose rate conversion factors for external exposure to photon emitters in soil. *Health Phys, **78 (3)**: 295-302*.
8. Askri B (2015) Application of optimized geometry for the Monte Carlo simulation of a gamma-ray field in air created by sources distributed in the ground. *Radiation Measurements, **72:** 1-11*.
9. http://geant4.web.cern.ch/
10. Gong C, Tang X, Fatemi S, Yu H, Shao W, Shu D, Geng C (2018) A Monte Carlo study of SPECT in boron neutron capture therapy for a heterogeneous human phantom. *Int J Radiat Res, **16(1)**: 33-43*.
11. Varley A, Tyler A, Smith L, Dale P (2015) Development of a neural network approach to characterise 226Ra contamination at legacy sites using gamma-ray spectra taken from boreholes. *J Environ Radioac, **140**: 130-140*.
12. Zakaryazad A and Duman E (2016) A profit-driven Artificial Neural Network (ANN) with applications to fraud detection and direct marketing. *Neurocomputing, **175**: 121-131*.
13. Akkurt I and Gunoglu K (2012) Estimation of bremsstrahlung photon fluence from aluminum by artificial neural network. *Int J Radiat Res, **10(1)**: 63-65*.
14. Brun R ans Rademakers F (1997) ROOT - an object oriented data analysis framework. *Nucl Instr Meth. Phys Res A, **389**: 81-8.*

[DOI: 10.52547/ijrr.20.1.30]