Environ Monit Assess (2024) 196:315
https://doi.org/10.1007/s10661-024-12459-8

RESEARCH

Using radial basis artificial neural networks to predict radiation hazard indices in geological materials

Selin Erzin

Received: 27 October 2023 / Accepted: 17 February 2024 / Published online: 28 February 2024

Abstract The estimation of exposures to humans from the various sources of radiation is important. Radiation hazard indices are computed using procedures described in the literature for evaluating the combined effects of the activity concentrations of primordial radionuclides, namely, ^{238}U, ^{232}Th, and ^{40}K. The computed indices are then compared to the allowed limits defined by International Radiation Protection Organizations to determine any radiation hazard associated with the geological materials. In this paper, four distinct radial basis function artificial neural network (RBF-ANN) models were developed to predict radiation hazard indices, namely, external gamma dose rates, annual effective dose, radium equivalent activity, and external hazard index. To make RBF-ANN models, 348 different geological materials' gamma spectrometry data were acquired from the literature. Radiation hazards indices predicted from each RBF-ANN model were compared to the radiation hazards calculated using gamma spectrum analysis. The predicted hazard indices values of each RBF-ANN model were found to precisely align with the calculated values. To validate the accuracy and the adaptability of each RBF-ANN model, statistical tests (determination coefficient (R^2), relative absolute error (RAE), root mean square error (RMSE), Nash–Sutcliffe Efficiency (NSE)), and significance tests (F-test and Student's t-test) were performed to analyze the relationship between calculated and predicted hazard indices. Low RAE and RMSE values as well as high R^2, NSE, and p-values greater than 0.95, 0.71, and 0.05, respectively, were found for RBF-ANN models. The statistical tests' results show that all RBF-ANN models created exhibit precise performance, indicating their applicability and efficiency in forecasting the radiation hazard indices of geological materials. All the RBF-ANN models can be used to predict radiation hazard indices of geological materials quite efficiently, according to the performance level attained.

Keywords Radiation hazard indices · Radial basis function · Artificial neural networks · Gamma spectrometry measurements

Introduction

Natural sources are the principle radioactive sources in the environment (Hanfi et al., 2021). Human populations are exposed to a broad spectrum of radiation in their surroundings because of the abundance of natural sources (Mehra et al., 2007; Amini Birami et al., 2019; Gaffar et al., 2021). To assess the effects of radiation exposure from both terrestrial and extra-terrestrial sources, radionuclide dispersion and radiation concentrations in the environment must be

S. Erzin (✉)
Science Faculty, Physics Department, Dokuz Eylul University, 35390 İzmir, Turkey
e-mail: selin.erzin@deu.edu.tr

understood (El-Arabi, 2007). The principal sources of extraterrestrial radiation are high-intensity cosmic rays produced from the earth's outer atmosphere (Prasad et al., 2020). Terrestrial radiation is caused by radioactive nuclides found in varying levels in rocks, construction materials, soils, water, and the atmosphere (Al-khawlany, 2017). Primordial radionuclides such as the ^{232}Th series, ^{238}U series, and ^{40}K are found in the earth's crust, which is the main root of background radiation in the environment (Haydar et al., 2021). When these radionuclides decay, their daughters emit radiations of gamma, beta, and alpha into the environment (Algattawi et al., 2019). As a result, people are constantly exposed to ionizing radiation both inside and outside their homes (Algattawi et al., 2019). These exposures may differ depending on the geology of any given place (Erzin & Yaprak, 2022).

The human body may be internally exposed to harmful radiation through three main routes: ingestion, inhalation, and skin contact (UNSCEAR, 2008). Of these, ingestion is the most prevalent and nearly inevitable route (Sarker et al., 2021). In reality, food and water intake are the primary routes via which radionuclides enter the human body (Sarker et al., 2021). Radionuclides are quickly incorporated into soft human tissues after intake. This results in complex health-threatening illnesses that may last a lifetime (Sarker et al., 2021). Therefore, individuals who wish to reduce long-term exposure should be aware of the distribution of ^{238}U, ^{232}Th, and ^{40}K (Erzin & Yaprak, 2022). Measuring radionuclide concentrations aids in the monitoring of environmental radiation (Lasheen et al., 2021). Radiation hazard indices were thus computed using procedures described in the literature for evaluating the combined effects of the activity concentrations of ^{238}U, ^{232}Th, and ^{40}K, and then compared to those obtained from other countries as well as the allowed limits defined by International Radiation Protection Organizations (UNSCEAR, 2000).

Machine learning (ML) approaches have grown in popularity in recent decades, enabling greater flexibility and less presumptuous options (Wadoux et al., 2020). Using data mining techniques, the ML approach determines if there are linear or nonlinear correlations between environmental factors and soil parameters (McBratney et al., 2003). This allows forecasts to be created without considering an unknown soil's physical field or producing a fictional dispersion from unprocessed data (Shi & Wang, 2021). The most commonly used ML models are support vector regression, artificial neural networks (ANNs), and random forest (Khaledian & Miller, 2020). Dragović (2022) outlined the use of ANN modeling in environmental radioactivity investigations. This included mapping, distributing, and detecting radionuclides, as well as anticipating their movement in the environment, improving measuring techniques, monitoring nuclear plant activities, and conducting real-time data analyses.

Broomhead and Lowe (1988) introduced the radial basis function artificial neural network (RBF-ANN) into the neural network literature as an efficient three-layer feedforward ANN based on the function approximation approach (Fu & Wang, 2003). The RBF-ANN is a typical ANN model that offers significant advantages in learning speed, parameter adjustment, and nonlinear adaptability (Dragović, 2022). RBF-ANNs are well-known for their simplicity of design, excellent tolerance to input noise, and quick and comprehensive training (Yu et al., 2011). In the presence of patterns that are not employed for learning, RBF-ANN works well (Yu et al., 2011). RBF-ANNs are widely used in a number of applications, such as noisy interpolation, function approximation, and pattern classification (Fu & Wang, 2003; Moody & Darken, 1989; Nabney, 2004; Pochmullcr et al., 1994). However, RBF-ANNs have not been extensively used in the literature to predict the activity concentrations of radionuclides except two studies carried out by Einian et al. (2015) and Erzin and Yaprak (2022). Einian et al. (2015) examined the RBF-ANN method's suitability for predicting the ^{234}U/^{238}U activity ratio. According to the findings of their investigation, the RBF-ANN technique can reliably forecast the ^{234}U/^{238}U activity ratio for alpha spectrometry. They identified three major advantages of RBF-ANN: (1) quick RBF-ANN training, (2) high accuracy in forecasting the activity ratio for spectrum data, and (3) the best network to estimate the activity ratio for alpha spectrometry. In a recent study, Erzin and Yaprak (2022) examined the suitability of RBF-ANNs in predicting the activity concentrations of primordial radionuclides, namely ^{232}Th, ^{238}U and ^{40}K. To do this, they used 126 different geological materials' gamma spectrometry measurements while developing RBF-ANN models. They concluded that the constructed RBF-ANN models are very good at predicting

the primordial radionuclides' activity concentrations in geological materials. Additionaly, RBF-ANNs have not been used in the literature to predict the hazard indices from gamma spectrometry measurements. Taking into consideration Erzin and Yaprak's (2022) findings on the appropriateness of RBF-ANNs for gamma spectrometry, the present work focus on RBF-ANNs for the prediction of radiation hazard indices from geological materials' gamma spectrometry measurements.

Using the findings of gamma spectrometry measurements, four independent RBF-ANN models were developed to predict radiation hazard indices, namely, external gamma dose rate in air (D), annual effective dose (AED), radium equivalent activity (Ra_{eq}), and external hazard index (H_{ex}). To do this, the data of 348 different geological materials' gamma spectrometry measurements acquired from Kaynar et al. (2014), Tabar et al. (2017), and Erzin and Yaprak (2022) were used. To evaluate the prediction performance of RBF-ANN models, the predicted radiation hazard indices were compared to the determined radiation hazard indices. Furthermore, the prediction accuracy of the built RBF-ANN models was assessed using a variety of performance metrics.

Experiment with gamma ray spectroscopy using a NaI(Tl) detector

To estimate the activity concentration of primordial radionuclides, gamma ray spectrometry techniques employing high resolution HPGe and NaI(Tl) detectors are used (Hofstadter, 1949). Despite their limited resolution, NaI(Tl) detectors are the most commonly employed for measuring both man-made and natural radionuclides in a wide range of materials (Eker & Çağlar, 2019; Pilakouta et al., 2018). The method of gamma ray spectroscopy for ^{232}Th, ^{238}U, and ^{40}K activity concentrations using a NaI(Tl) detector has already been widely published (Sahin & Cavas, 2008; Yaprak, 1995; Yaprak & Aslani, 2010). As a result, it will only be briefly explored here. To calculate the activity concentrations of the primordial radionuclides ^{232}Th, ^{238}U, and ^{40}K, three linear equations were solved.

$$\mathrm{eTh(Bq\ kg^{-1})} = \mathrm{C(^{232}Th)/K_1} \quad (1)$$

$$\mathrm{eU(Bq\ kg^{-1})} = [\mathrm{C(^{238}U)} - \alpha\ \mathrm{C(^{232}Th)}]/\mathrm{K_2} \quad (2)$$

$$\mathrm{K(Bq\ kg^{-1})} = [\mathrm{C(^{40}K)} - \gamma\ [\mathrm{C(^{238}U)} - \alpha\ \mathrm{C(^{232}Th)}] - \beta\ \mathrm{C(^{232}Th)}]/\mathrm{K_3} \quad (3)$$

where C(^{40}K), C(^{238}U), and C(^{232}Th) are the count rates in channels potassium, uranium, and thorium, e indicates the acceptance of series equilibrium condition, α, β, and γ are the stripping ratios, K_1, K_2, and K_3 are the sensitivity factors, while αC(^{232}Th) and βC(^{232}Th) are the count rates including thorium contributions in the windows of uranium and the potassium, γC(^{238}U) is the count rate including the uranium contribution in the window of potassium (Erzin & Yaprak, 2022). The uranium and potassium concentrations cannot be determined spectroscopically unless these count rates are taken into consideration (Erzin & Yaprak, 2022).

A variety of parameters, including detector volume, counting geometry, and the widths of the chosen energy windows, influence sensitivity factors, and stripping ratios (Yaprak, 1995; Sahin & Cavas, 2008; Yaprak & Aslani, 2010), which must be computed using the activity concentrations of ^{232}Th, ^{238}U, and ^{40}K standards for each gamma spectrometric system (Erzin & Yaprak, 2022).

Radiation hazard indices

Different radiation hazard indices, such as D, AED, Ra_{eq}, and H_{ex}, are used by national regulatory organizations and agencies worldwide to assess the harm posed by radionuclides in soil (Shabbir et al., 2023). In order to analyze the health risks associated with these gamma emissions, D, AED, Ra_{eq}, and H_{ex} are first estimated and then compared to the permissible limits (Khan et al., 2023). The D value from ^{238}U, ^{232}Th, and ^{40}K exposure at 1 m from the earth's surface depends on the activity concentrations of ^{238}U, ^{232}Th, and ^{40}K and may be calculated using the following equation (UNSCEAR, 2000);

$$D\ \mathrm{(nGy\ h^{-1})} = 0.0417\mathrm{C_K} + 0.604\mathrm{C_{Th}} + 0.462\mathrm{C_U} \quad (4)$$

where C_K, C_{Th}, and C_U are ^{40}K, ^{232}Th, and ^{238}U activity concentrations in Bq kg^{-1}, respectively.

Equation (5) can be used to calculate the AED value. In this equation, 8760 (h y^{-1}) is the exposure duration, and 0.2 and 0.7 (Sv Gy^{-1}) are the occupancy and dose conversion factors, respectively

(UNSCEAR, 2000). The equation was multiplied by10^{-3} to convert to μSv.

$$\begin{aligned}\text{AED}\ (\mu\ \text{Sv}\ \text{y}^{-1}) &= D\ (\text{nGy}\ \text{h}^{-1}) \times 8760\ (\text{h}\ \text{y}^{-1}) \\ &\times 0.2 \times 0.7\ (\text{Sv}\ \text{Gy}^{-1}) \times 10^{-3}\end{aligned} \tag{5}$$

The Ra_{eq} value is the most commonly used radiation hazard indice in environmental radioactivity investigation. The following equation can be used to calculate the Ra_{eq} of a sample having varied quantities of primordial radionuclides:

$$\begin{aligned}\text{Ra}_{\text{eq}}(\text{Bq kg}^{-1}) &= 370 \times \left(\frac{C_U}{370} + \frac{C_{\text{Th}}}{259} + \frac{C_K}{4810}\right) \\ &= C_U + 1.43C_{\text{Th}} + 0.07C_K\end{aligned} \tag{6}$$

H_{ex}, another common radiation hazard indice, is used to monitor radioactive material consumption and to limit external gamma radiation exposure to 1.5 nGy. The H_{ex} values should be less than unity to meet the safety criteria. The following calculation can be used to calculate the H_{ex} of a sample having different quantities of primordial radionuclides.

$$\text{H}_{\text{ex}} = \left(\frac{C_U}{370} + \frac{C_{\text{Th}}}{259} + \frac{C_K}{4810}\right) \tag{7}$$

Radial basis function artificial neural networks (RBF-ANNs) and RBF-ANN models

Radial basis function artificial neural networks (RBF-ANNs)

RBF-ANN has three layers: an input layer, a hidden layer, and an output layer (Fig. 1). The number of neurons in the input layer equals the number of input variables. The input layer provides data to the hidden layer, which contains neurons that compute the distance from a central point of given input values (Segal et al., 2000). The hidden layer uses RBFs to apply a nonlinear mapping to the data (Segal et al., 2000). After applying RBFs to an input vector, the hidden layer neurons' outputs are transmitted to the output layer (Segal et al., 2000). A flow chart of RBF-ANN is shown in Fig. 2.

A general formulation of RBF-ANN that expresses the output-input relationship is as follows (Robert & Howlett, 2001);

$$y_i = \sum_{k=1}^{N} w_{\text{ik}} \varnothing_k(x, c_k) = \sum_{k=1}^{N} w_{\text{ik}} \varnothing_k\left(\|x - c_k\|_2\right) \quad i = 1, 2, 3, ..m \tag{8}$$

where y is the output produced by the network, N is the hidden layer's cell number, $x \in R^{n\times1}$ is the input data, w_{ik} are weights of the output layer, $c_k \in R^{n\times1}$ are the radial-based centers selected from a subset of the input vector space, and $\varnothing_k(.)$ is the radial-based activation function (Segal et al., 2000; Szczurek & Maciejewska, 2004).

RBF-ANN models rely largely on cell centers, the weights of output layer, and the form of radial basis function (RBF). Linear, multi-quadratic, inverse multi-quadratic, Gaussian, and cubic RBFs can all be employed with RBF-ANN models. The Gauss function was chosen for this study because it can be factored. The Gauss function's mathematical structure is depicted in the following equation:

$$\varnothing_k(x) = \exp\left(-\|x - c_k\|_2{}^2 / 2\sigma^2\right) \tag{9}$$

where $\|.\|_2 \text{is the Euclidean norm}, c_k$ is the centers, x is the input vector, and σ denotes the standard deviation, commonly referred to as the spread parameter, which

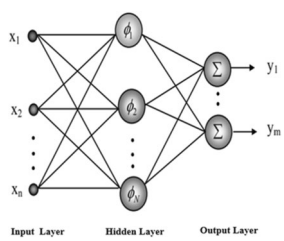

Fig. 1 The structure of RBF-ANN (Buhmann, 2003)

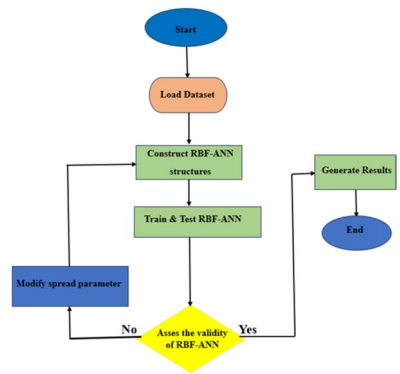

Fig. 2 Flow chart of a RBF-ANN

Table 1 The descriptive statistics of the input parameters ($C(^{232}Th)$, $C(^{238}U)$, and $C(^{40}K)$) used for all RBF-ANN models developed

Statistic	$C(^{232}Th)$	$C(^{238}U)$	$C(^{40}K)$
No of samples	348	348	348
Minimum	0	0	0
Maximum	3301	7760	58,576
Arithmetic mean ± SEM	802 ± 33	1518 ± 67	5696 ±280
Median	745	1397	6576
Skewness	0.16	0.70	5.49
Kurtosis	− 1.09	1.52	49.98
Standard deviation	623	1243	5232

SEM: standard errors of means

has a significant impact on the the RBFNN model's performance.

The training of RBF-ANN models is divided into two phases. The first step is to select centers from the training data or cluster the training data to generate centers. The second phase is the weight computation between the hidden layer and the output layer. There are numerous approaches in the literature for identifying output weights and cell centers. For finding cell centers, the Kohonen clustering and *K*-means methods are often utilized. The Moore–Penrose pseudo-inverse and least mean squares methods are used to calculate the output weights (Haykin, 2009). The spread parameter is almost completely constant. There are approximate equivalences in the literature for finding the spread parameter, but it can also be determined by trial and error (Ham & Kostanic, 2001).

RBF-ANN models

In this paper, four different RBF-ANN models designated as RBF-ANN1, RBF-ANN2, RBF-ANN3, and RBF-ANN4 were constructed for the prediction of the radiation hazard indices *D*, AED, Ra_{eq}, and H_{ex}, respectively. To achieve this, the results of the gamma spectrometry measurements acquired from Kaynar et al. (2014), Tabar et al. (2017), and Erzin and Yaprak (2022) were used. In these measurements, activity concentrations of ^{238}U, ^{232}Th, and ^{40}K of 348 geological materials of various origins, involving soil, granite, rock, fertilizer and marble samples, were determined systematically using a NaI(Tl) detector with one sigma error and then expressed relatively to dry weight. The determined activity concentrations' relative uncertainties at 68% confidence level were found to be commonly less than 10%. In all RBF-ANN models, the count rates $C(^{238}U)$, $C(^{232}Th)$, and $C(^{40}K)$ were employed as input parameters. The ouput parameter was selected as *D*, AED, Ra_{eq}, and H_{ex} in the RBF-ANN1, RBF-ANN2, RBF-ANN3, and RBF-ANN4 models, respectively. Tables 1 and 2 show the descriptive statistics for the input and output parameters, respectively. Each RBF-ANN model's input and output data are standardized into the range 0 to 1, utilizing the variable's maximum and lowest values across all data sets. The normalized data was then utilized to generate a training and testing set. The training set was used to build each RBF-ANN model. The testing set was used to evaluate the performance of each RBF-ANN model. All the RBF-ANN models used 278 training and 70 testing data from 348 data sets. The same 70 testing data was used for each RBF-ANN model.

Table 2 The descriptive statistics of the output parameters (*D*, AED, Ra_{eq}, and H_{ex}) used for RBF-ANN models developed

Statistic	Radiological hazard indices			
	D (nGy h^{-1})	AED (μSv y^{-1})	Ra_{eq} (Bq kg^{-1})	H_{ex}
No of samples	348	348	348	348
Minimum	0	0	0	0
Maximum	522	640	876	2.60
Arithmetic mean ± SEM	70 ± 3	86 ± 3	145 ± 5	0.40 ± 0.02
Median	63	78	132	0.37
Skewness	3.56	3.56	2.41	2.75
Kurtosis	24.00	24.00	12.49	15.75
Std. deviation	53	65	100	0.28

SEM: standard errors of means

Table 3 The performance indices (RAE, RMSE, and NSE) for the RBF-ANN models

Spread parameter	Data set	RBF-ANN1			RBF-ANN2			RBF-ANN3			RBF-ANN4		
		RAE (nGy h^{-1})	RMSE (nGy h^{-1})	NSE	RAE (μSv y^{-1})	RMSE (μSv y^{-1})	NSE	RAE (Bq kg^{-1})	RMSE (Bq kg^{-1})	NSE	RAE	RMSE	NSE
0.5	Training	0.11	7.04	0.98	0.11	8.63	0.98	0.11	14.64	0.98	0.11	0.04	0.98
	Testing	1.60	215.77	− 44.58	1.60	264.63	− 44.58	1.54	434.75	− 40.19	1.56	1.22	− 41.62
1	Training	0.16	9.30	0.97	0.16	11.2	0.97	0.16	18.90	0.97	0.16	0.05	0.97
	Testing	0.37	18.35	0.67	0.37	22.5	0.67	0.37	41.24	0.63	0.37	0.11	0.64
1.5	Training	0.24	12.96	0.95	0.2	13.73	0.96	**0.21**	**23.07**	**0.95**	0.20	0.06	0.95
	Testing	0.38	17.64	0.7	0.38	21.63	0.7	**0.37**	**36.52**	**0.71**	0.37	0.10	0.70
2.0	Training	0.22	12.27	0.95	0.22	15.04	0.95	0.23	25.27	0.94	0.23	0.07	0.95
	Testing	0.42	19.27	0.64	0.42	23.63	0.64	0.41	39.63	0.66	0.41	0.11	0.65
2.5	Training	0.24	12.96	0.95	0.24	15.89	0.95	0.24	26.74	0.94	0.24	0.07	0.94
	Testing	0.46	19.67	0.62	0.46	24.24	0.62	0.45	41.36	0.63	0.45	0.11	0.62
3.0	Training	0.26	14.54	0.93	0.26	17.83	0.93	0.27	30.04	0.92	0.27	0.08	0.92
	Testing	0.45	22.71	0.5	0.45	27.85	0.5	0.44	47.25	0.51	0.44	0.13	0.50
3.5	Training	0.26	14.63	0.93	0.26	17.95	0.93	0.27	30.26	0.92	0.27	0.08	0.92
	Testing	0.45	23.04	0.48	0.45	28.26	0.48	0.44	48.01	0.50	0.45	0.13	0.49
4.0	Training	0.26	14.66	0.93	0.26	17.98	0.93	0.27	30.33	0.92	0.27	0.08	0.92
	Testing	0.44	21.93	0.53	0.44	26.89	0.53	0.43	45.58	0.55	0.43	0.13	0.54
4.5	Training	0.27	15.10	0.93	0.27	18.51	0.93	0.28	31.25	0.91	0.28	0.09	0.92
	Testing	0.35	15.13	0.78	0.35	18.56	0.78	0.34	30.74	0.79	0.35	0.09	0.79
5.0	Training	**0.28**	**15.55**	**0.92**	**0.28**	**19.08**	**0.92**	0.29	32.16	0.91	**0.29**	**0.09**	**0.91**
	Testing	**0.35**	**13.31**	**0.83**	**0.35**	**16.33**	**0.83**	0.34	27.29	0.84	**0.35**	**0.08**	**0.83**
5.5	Training	0.28	15.56	0.92	0.28	19.28	0.92	0.29	32.17	0.91	0.29	0.09	0.91
	Testing	0.35	13.32	0.83	0.35	16.35	0.83	0.34	27.30	0.84	0.35	0.08	0.83
6.0	Training	0.29	15.89	0.92	0.29	19.48	0.92	0.30	32.90	0.90	0.30	0.09	0.91
	Testing	0.36	13.81	0.81	0.36	16.94	0.81	0.35	28.54	0.82	0.36	0.08	0.82
6.5	Training	0.3	16.41	0.92	0.3	20.12	0.92	0.31	33.74	0.90	0.30	0.09	0.90
	Testing	0.37	15.16	0.77	0.37	18.59	0.77	0.36	30.96	0.79	0.37	0.09	0.78
7.0	Training	0.32	17.55	0.9	0.32	21.52	0.9	0.33	36.28	0.88	0.33	0.10	0.89
	Testing	0.41	15.9	0.75	0.41	19.5	0.75	0.41	32.95	0.76	0.41	0.09	0.76
7.5	Training	0.36	19.06	0.89	0.36	23.38	0.89	0.37	39.45	0.86	0.37	0.11	0.87
	Testing	0.42	18.43	0.67	0.42	22.6	0.67	0.41	38.80	0.67	0.42	0.11	0.67

Table 3 (continued)

Spread parameter	Data set	RBF-ANN1			RBF-ANN2			RBF-ANN3			RBF-ANN4		
		RAE (nGy h^{-1})	RMSE (nGy h^{-1})	NSE	RAE (μSv y^{-1})	RMSE (μSv y^{-1})	NSE	RAE (Bq kg^{-1})	RMSE (Bq kg^{-1})	NSE	RAE	RMSE	NSE
8.0	Training	0.36	19.06	0.89	0.36	23.38	0.89	0.37	39.46	0.86	0.37	0.11	0.87
	Testing	0.42	18.47	0.67	0.42	22.66	0.67	0.42	38.88	0.67	0.42	0.11	0.67
8.5	Training	0.36	19.06	0.89	0.36	23.38	0.89	0.37	39.46	0.86	0.37	0.11	0.87
	Testing	0.42	18.51	0.66	0.42	22.7	0.66	0.42	38.95	0.67	0.42	0.11	0.67
9.0	Training	0.36	19.07	0.89	0.36	23.38	0.89	0.37	39.47	0.86	0.37	0.11	0.87
	Testing	0.42	18.55	0.66	0.42	22.74	0.66	0.42	39.01	0.67	0.42	0.11	0.67
9.5	Training	0.36	19.06	0.89	0.36	23.38	0.89	0.37	39.47	0.86	0.37	0.11	0.87
	Testing	0.42	18.57	0.66	0.42	22.78	0.66	0.42	39.05	0.67	0.42	0.11	0.66
10.0	Training	0.36	19.07	0.89	0.36	23.39	0.89	0.38	39.49	0.86	0.37	0.11	0.87
	Testing	0.42	18.6	0.66	0.42	22.81	0.66	0.42	39.09	0.67	0.42	0.11	0.66

The MATLAB7.0 neural network toolbox was used to train and evaluate each RBF-ANN model. While developing each RBF-ANN model, the Gauss function, a commonly utilized RBF (Bors & Pitas, 1996), was used. As mentioned earlier, the spread parameter has a major influence on the RBF-ANN model's performance. In this work, each RBF-ANN model's optimal spread parameter was found by gradually raising the parameter from a minimum of 0.5 to a maximum of 10, adding 0.5 to the parameter each time. To evaluate the RBF-ANN models' validity, a variety of statistical error measures can be applied. The relative absolute error (RAE), root mean square error (RMSE), and Nash–Sutcliffe efficiency (NSE) statistical error indices were employed in this study for evaluating the developed RBF-ANN models' performance, as shown in Eqs. (10), (11), and (12).

$$\mathrm{RAE} = \frac{\sum_{i=1}^{N}\left|\left(\mathrm{RHI}_i^{\mathrm{Pred}} - \mathrm{RHI}_i^{\mathrm{Exp}}\right)\right|}{\sum_{i=1}^{N}\left|\left(\mathrm{RHI}_i^{\mathrm{Exp}} - \mathrm{RHI}_i^{\mathrm{ExpMean}}\right)\right|} \tag{10}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{RHI}_i^{\mathrm{Exp}} - \mathrm{RHI}_i^{\mathrm{Pred}}\right)^2} \tag{11}$$

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(\mathrm{RHI}_i^{\mathrm{Pred}} - \mathrm{RHI}_i^{\mathrm{Exp}}\right)^2}{\sum_{i=1}^{N}\left(\mathrm{RHI}_i^{\mathrm{Exp}} - \mathrm{RHI}_i^{\mathrm{ExpMean}}\right)^2} \tag{12}$$

where $\mathrm{RHI}_i^{\mathrm{Exp}}$ and $\mathrm{RHI}_i^{\mathrm{Pred}}$ are the measured and the predicted radiation hazard indice values, $\mathrm{RHI}_i^{\mathrm{ExpMean}}$ is the mean value of the measured radiation hazard indice, N is the sample number. To normalize the error, RAE divides the total absolute error by the predictor's total absolute error. RAE values range from 0 to ∞. In a precise prediction, RAE equals 0; the numerator value increases as the model prediction error increases. The RMSE measures the difference between the actual and projected values. The smaller the difference between actual and anticipated values, the closer the RMSE is to zero. NSE is a coefficient that is particularly effective for assessing a model's ability to forecast. The NSE values vary from − ∞ to 1. A NSE score of 1 implies that the model predictions and observations agree exactly. A NSE value less than 0 indicates that the residual variance is greater than the data variance, but a NSE value of 0

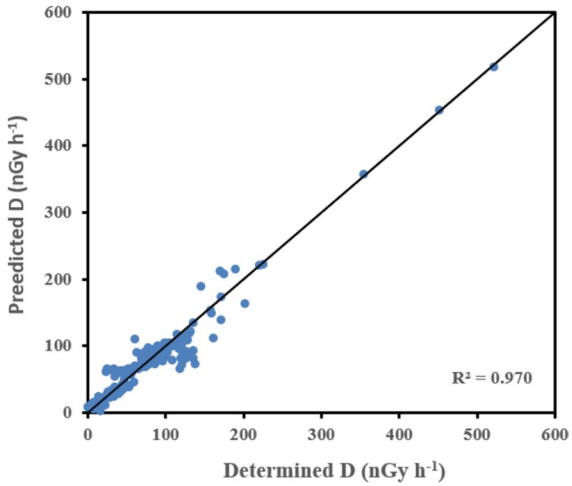

Fig. 3 Comparison of the determined D values with the predicted D values from RBF-ANN1 model for training set

shows that model predictions are as accurate as the actual data's average. In general, models with NSE between 0 and 1 are viable; otherwise, the model is considered unsuitable for application.

The details of the optimal performance of each RBF-ANN model were given in Table 3. When the performance indices (RAE, RMSE, and NSE) of each RBF-ANN model in Table 3 were compared, it was found that RBF-ANN1, RBF-ANN2, RBF-ANN3, and RBF-ANN4 models with spread parameters of 5.0, 5.0, 1.5, and 5.0, respectively, (shown in bold in

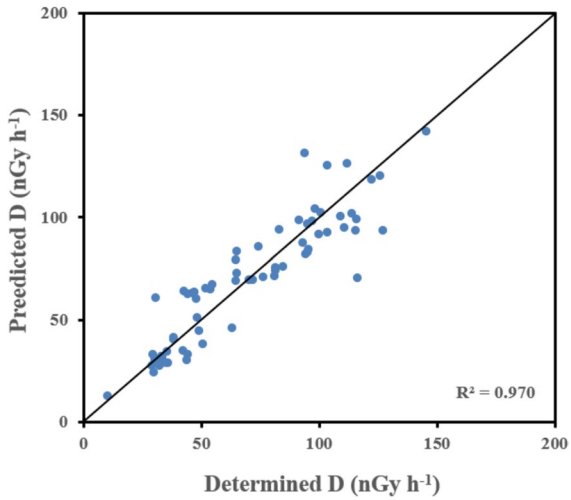

Fig. 4 Comparison of the determined D values with the predicted D values from RBF-ANN1 model for testing set

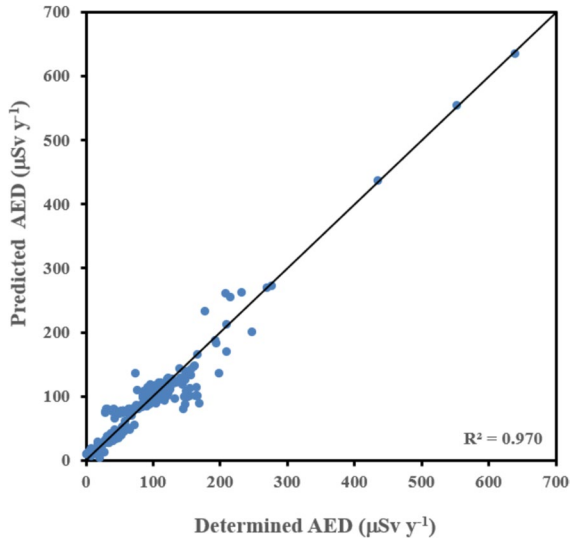

Fig. 5 Comparison of the determined AED values with the predicted AED values from RBF-ANN2 model for training set

Table 3), yielded lower RAE and RMSE values and higher NSE values and were thus chosen as the best RBF-ANN models.

In order to validate the model's fit, significance tests are conducted to the relationship between determined and predicted databases (Reyes-Téllez et al., 2020). The most widely used significance tests for testing the null hypothesis are the *F*-test, also known as the Fisher's *F*-test, and the Student's *t*-test (Verma, 2005). Student's *t*-test were used to assess the effectiveness of each RBF-ANN model's

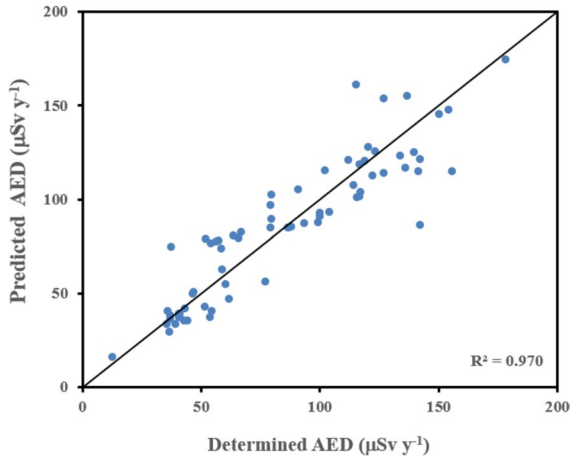

Fig. 6 Comparison of the determined AED values with the predicted AED values from RBF-ANN2 model for testing set

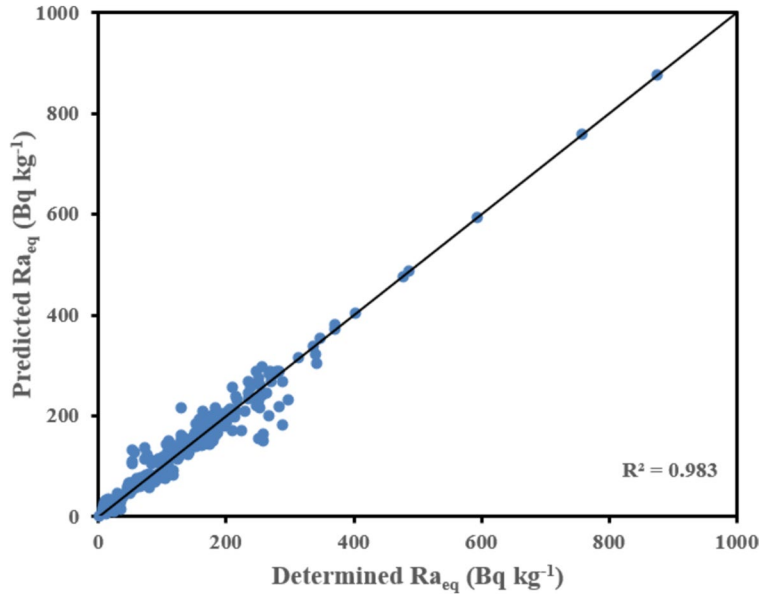

Fig. 7 Comparison of the determined Ra_{eq} values with the predicted Ra_{eq} values from RBF-ANN3 model for training set

performance. A Fisher's *F*-test was carried out by using Microsoft Excel software (Microsoft Corp., Redmond, WA, USA) prior to the Student's *t*-test to assess whether the fluctuations of the determined and anticipated radiation hazard indices were equal (Snedecor & Cochran, 1989). After conducting Fisher's *F*-tests, *p*-values for each RBF-ANN model were obtained. When the *p*-values exceed 0.05, it is not possible to reject the null hypothesis (Gupta, 2010). Stated otherwise, there exists no distinction in the variances of predicted and determined hazard indices. In addition, if the *p*-values are more than 0.05, a Student's *t*-test for each RBF-ANN model may be run, providing the variances of predicted and determined values are identical. If the *p*-values are less than 0.05, a Student's *t*-test for each RBF-ANN model may be run, providing the variances of predicted and determined values are not identical. Then, for each RBF-ANN model, a Student's *t*-test was carried out by using Microsoft Excel software (Microsoft Corp., Redmond, WA, USA) to test the null hypothesis, and *p*-values for each RBF-ANN

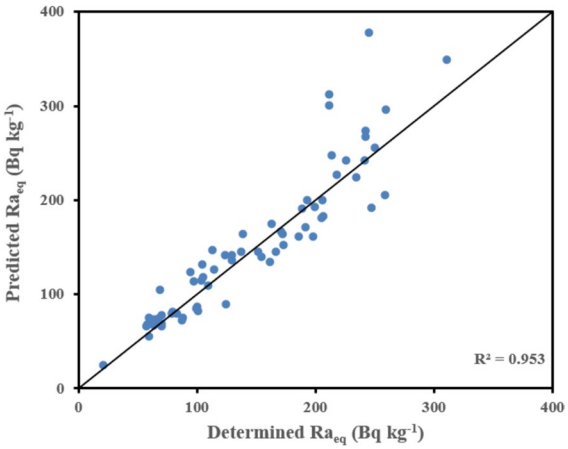

Fig. 8 Comparison of the determined Ra_{eq} values with the predicted Ra_{eq} values from RBF-ANN3 model for testing set

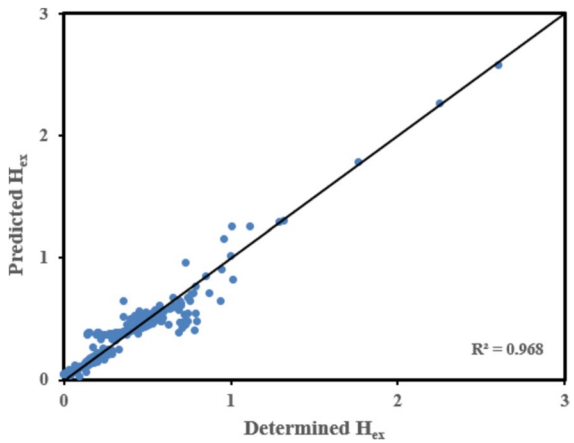

Fig. 9 Comparison of the determined H_{ex} values with the predicted H_{ex} values from RBF-ANN4 model for training set

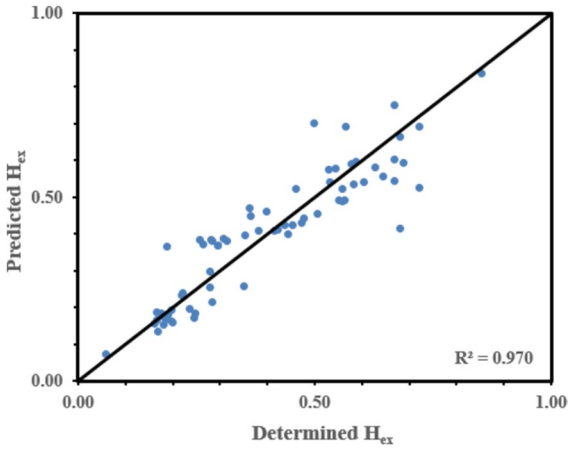

Fig. 10 Comparison of the determined H_{ex} values with the predicted H_{ex} values from RBF-ANN4 model for testing set

model were produced. The produced *p*-values exceed 0.05, indicating strong support for the null hypothesis and no significant difference in experimental and predicted data.

Results and discussion

In this study, four separate RBF-ANN models (RBF-ANN1, RBF-ANN2, RBF-ANN3, and RBF-ANN4) were developed for the prediction of radiation hazard indices of D, AED, Ra_{eq} and H_{ex}, respectively, using gamma spectrometry measurements' results acquired from the literature. Individual RBF-ANN models' prediction performance was then evaluated by comparing measured and predicted radiation hazard indices. Figures 3 and 4 compare the predicted D values from the RBF-ANN1 model with the determined D values from gamma spectrometry measurements. Figures 3 and 4 show that the RBF-ANN1 model has the minimum scatter around the line of equality between predicted and determined D values, with coefficients of determination (R^2) values of 0.970 both for training and testing samples. Figures 5 and 6 show a comparison of AED values predicted using the RBF-ANN2 model with those determined using gamma spectrometry measurements. Figures 5 and 6 illustrate that the predicted AED values were fairly similar to the determined AED values, since the R^2 values of 0.970 for both training and testing samples are very close to unity. The predicted Ra_{eq} values

Table 4 The performance indices (R^2, RAE, RMSE, and NSE) for the RBF-ANN models

	RBF-ANN1				RBF-ANN2				RBF-ANN3				RBF-ANN4			
Data	R^2	RAE (nGy h^{-1})	RMSE (nGy h^{-1})	NSE	R^2	RAE (μSv y^{-1})	RMSE (μSv y^{-1})	NSE	R^2	RAE (Bq kg^{-1})	RMSE (Bq kg^{-1})	NSE	R^2	RAE	RMSE	NSE
Training	0.970	0.28	15.55	0.92	0.970	0.28	19.08	0.92	0.983	0.21	23.07	0.95	0.968	0.29	0.090	0.91
Testing	0.970	0.35	13.31	0.83	0.970	0.35	16.33	0.83	0.953	0.37	36.52	0.71	0.970	0.35	0.080	0.83

using the RBF-ANN3 model were plotted against the determined Ra_{eq} values from gamma spectrometry measurements in Figs. 7 and 8. As shown in Figs. 7 and 8, predicted Ra_{eq} values were found to be in good agreement with determined Ra_{eq} values, with R^2 values of 0.983 and 0.952 for training and testing samples. The H_{ex} values predicted from the RBF-ANN4 model were shown against the determined H_{ex} values from gamma spectrometry measurements in Figs. 9 and 10. As demonstrated in Figs. 9 and 10, predicted H_{ex} values were very close to the determined H_{ex} values, with R^2 values of 0.968 and 0.967 for training and testing samples.

The RAE, RMSE, and NSE values for each RBF-ANN model constructed in this research were computed using Eqs. (10), (11), and (12), respectively, and were shown in Table 4. The computed performance parameters in Table 4 show that all the RBF-ANN models created exhibit precise performance, indicating their applicability and efficiency in forecasting the radiation hazard indices of geological materials.

The *p*-values for the RBF-ANN1, RBF-ANN2, RBF-ANN3, and RBF-ANN4 models were found to be 0.228, 0.228, 0.464, and 0.189, respectively, according to the results of the Fisher's *F*-tests. The *p*-value found for each RBF-ANN model is more than 0.05, making refutation of the null hypothesis unfeasible (Gupta, 2010). In layman's words, the determined and forecasted radiation hazard indices have the same variance. As a result, for each RBF-ANN model, a Student's *t*- test was performed under the premise that the variances of predicted and determined values were similar. The Student's *t*-test yielded *p*-values of 0.990, 0.990, 0.944, and 0.985 for the RBF-ANN1, RBF-ANN2, RBF-ANN3, and RBF-ANN4 models, respectively. The *p*-value for each RBF-ANN model exceed 0.05, indicating no significant difference between the determined and estimated radiation hazard indices. Given these *p*-values, it is advised that all the RBF-ANN models be used for predicting the radiation hazard indices in geological materials if the count rates C(^{40}K), C(^{238}U), and C(^{232}Th) are known.

Conclusions

In this study, four different RBF-ANN models (RBF-ANN1, RBF-ANN2, RBF-ANN3, and RBF-ANN4) have been developed for the prediction of radiation hazard indices of *D*, AED, Ra_{eq}, and H_{ex}, respectively, using gamma spectrometry measurements obtained from the literature. For this purpose, the results of 348 geological materials' the gamma spectrometry measurements acquired from Kaynar et al. (2014), Tabar et al. (2017), and Erzin and Yaprak (2022) were used. The results obtained from RBF-ANN models were compared vis-à-vis those acquired by the measurements. It is found that the radiation hazard indices predicted by RBF-ANN models match very well with those determined from the measurements, with R^2 values greater than 0.95. Performance indices (RAE, RMSE, and NSE) were computed to validate the developed RBF-ANN model's accuracy and reliability. The performance of the RBF-ANN1 model (RAE = 0.35 nGy h^{-1}, RMSE = 13.31 nGy h^{-1}, and NSE = 0.83), RBF-ANN2 model (RAE = 0.35 μSv y^{-1}, RMSE = 16.33 μSv y^{-1}, and NSE = 0.83), RBF-ANN3 model (RAE = 0.37 Bq kg^{-1}, RMSE = 36.52 Bq kg^{-1}, and NSE = 0.71), and RBF-ANN4 model (RAE = 0.36, RMSE = 0.080, and NSE = 0.83) for training samples show that all RBF-ANN model constructed in this research exhibit precise performance, highlighting their applicability and efficiency in predicting the radiation hazard indices. Furthermore, significance tests, including Fisher's *F*-test and Student's *t*-test, support the hypothesis that the variances and means of the predicted and determined values are similar, further validating the accuracy and reliability of the RBF-ANN models. Therefore, based on the obtained *p*-values, it is recommended to utilize the established RBF-ANN models for predicting radiation hazard indices of geological materials when the corresponding count rates are known.

Acknowledgements The author thanks to Professor Sermin Çam Kaynar and Associted Professor Emre Tabar for enabling the use of their datasets throughout the current investigation.

Author contributions The author confirms sole responsibility for the following: study conceptualization and design, data collection, data analysis and interpretation, and manuscript writing, including all figures and tables.

Funding This study received no funding.

Data availability All datasets created during the current investigation are available upon reasonable request from the corresponding author.

Declarations

Ethics approval The author has read, understood, and complied with the declaration on "Ethical responsibilities of Authors" given in the Instructions for Authors, and is aware that, with minimal exceptions, no changes to authorship can be made once the work is submitted.

Consent to participate This paper contains no studies involving human subjects conducted by any of the writers.

Competing interests The author declares no competing interests.

References

Algattawi, A., Fayez-Hassan, M., Khalil, E., & Elez, H. (2019). radiation effects of soil and rock samples of different Libyan regions. *Engineering, 11*, 247–259.

Al-khawlany, A. H. (2017). Assessment of natural radioactivity levels and associated radiation hazards for some environmental soil and rock samples from outskirts of Aurangabad, Maharashtra-India using gamma-ray spectrometry. *International Journal of Innovative Science, Engineering and Technology, 6*, 16592–16604.

Amini Birami, F., Moore, F., Faghihi, R., & Keshavarzi, B. (2019). Distribution of natural radionuclides and assessment of the associated radiation hazards in the rock and soil samples from a high-level natural radiation area, Northern Iran. *Journal of Radioanalytical and Nuclear Chemistry, 322*, 2091–2103.

Bors, A. G., & Pitas, I. (1996). Median radial basis function neural network. *IEEE Transactions on Neural Networks, 7*, 1351–1364.

Broomhead, D., & Lowe, D. (1988). Multivariable functional interpolation and adaptive networks. *Complex Systems, 2*, 321–355.

Buhmann, M. D. (2003). *Radial basis functions: theory and implementations*. Cambridge University Press, Cambridge.

Dragović, S. (2022). Artificial neural network modeling in environmental radioactivity studies–A review. *Science of the Total Environment, 847*, 157526.

Einian, M. R., Aghamiri, S. M. R., & Ghaderi, R. (2015). Evaluation of the suitability of neural network method for prediction of uranium activity ratio in environmental alpha spectra. *Applied Radiation Isotopes, 105*, 225–232.

Eker, G. B. C., & Çağlar, İ. (2019). A study on calculation of full energy peak efficiency of NaI(Tl) detectors using point source. *Caucasian Journal of Science, 6*, 28–36.

El-Arabi, A. M. (2007). ^{226}Ra, ^{232}Th and ^{40}K concentrations in igneous rocks from eastern desert, Egypt and its radiation implications. *Radiatation Measurements, 42*, 94–100.

Erzin, S., & Yaprak, G. (2022). Prediction of the activity concentrations of ^{232}Th, ^{238}U and ^{40}K in geological materials using radial basis function neural network. *Journal of Radioanalytical and Nuclear Chemistry, 331*, 3525–3533.

Fu, X., & Wang, L. (2003). Data dimensionality reduction with application to simplifying RBF network structure and improving classification performance. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, 33*, 399–409.

Gaafar, I., Hanfi, M., El-Ahll, L. S., & Zeidan, I. (2021). Assessment of radiation hazards from phosphate rocks, Sibaiya area, central eastern desert. *Egypt. Applied Radiation and Isotopes, 173*(April), 109734.

Gupta, A. K. (2010). Predictive modelling of turning operations using response surface methodology, artificial neural networks and support vector regression. *International Journal of Production Research, 48*, 763–778.

Ham, F., & Kostanic, I. (2001). *Principles of neurocomputing for science and engineering*. Macgraw-Hill.

Hanfi, M. Y., Emad, B. M., Sayyed, M. I., Khandaker, M. U., & Bradley, D. A. (2021). Natural radioactivity in the prospecting tunnel in Egypt: Dose rate and risk assessment. *Radiation Physics and Chemistry, 187*, 109555.

Haydar, M. A., Hasan, M. M., Jahan, I., Fatema, K., Ali, M. I., Paul, D., & Khandaker, M. U. (2021). The status of NORMs in natural environment adjacent to the Rooppur nuclear power plant of Bangladesh. *Nuclear Engineering and Technology, 53*(12), 4114–4121.

Haykin, S. (2009). *Neural networks and learning machines* (Vol. 3). Pearson.

Hofstadter, R. (1949). The detection of gamma-rays with thallium-activated sodium iodide crystals. *Physical Review, 75*, 796–810.

Kaynar, S. Ç., Saç, M. M., & Ereeş, F. S. (2014). Determination of radioactivity levels in Akhisar, Gördes, Gölmarmara and Sindirgi regions, Western Turkey. *Environmental Earth Sciences, 71*, 1581–1592.

Khaledian, Y., & Miller, B. A. (2020). Selecting appropriate machine learning methods for digital soil mapping. *Applied Mathematical Modelling, 81*, 401–418.

Khan, M., Chaudhary, M. Z., Khan, E. U., et al. (2023). Assessment of radiological indices and physiochemical characterization of sediments in Chashma Lake, Pakistan. *Environmental Monitoring and Assessment, 195*, 1219.

Lasheen, E. S. R., Rashwan, M. A., Osman, H., Alamri, S., Khandaker, M. U., & Hanfi, M. Y. (2021). Radiological hazard evaluation of some Egyptian magmatic rocks used as ornamental stone: Petrography and natural radioactivity. *Materials, 14*(23), 1–13.

McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. *Geoderma*. https://doi.org/10.1016/S0016-7061(03)00223-4

Mehra, R., Singh, S., Singh, K., et al. (2007). ^{226}Ra, ^{232}Th and ^{40}K analysis in soil samples from some areas of Malwa region, Punjab, India using gamma ray spectrometry. *Environmental Monitoring and Assessment, 134*, 333–342.

Moody, J., & Darken, C. J. (1989). Fast learning in networks of locally-tuned processing units. *Neural Computation, 1*, 281–294.

Nabney, I. T. (2004). Efficient training of RBF networks for classification. *International Journal of Neural Systems, 14*, 201–208.

Pilakouta, M., Pappa, F. K., Patiris, D. L., Tsabaris, C., & Kalfas, C. A. (2018). A methodology for expanding the use of NaI(Tl) based spectrometry in environmental radioactivity measurements. *Applied Radiation Isotopes, 139*, 159–168.

Pochmullcr, W., Halgamugc, S. K., Glcsncr, M., Schwcikcrt, P., & Pfcffcrmann, A. (1994). RBF and CBF neural network learning procedures. In: IEEE World Congress on Computational Intelligence, Orlando. pp. 407–412.

Prasad, M., Ranga, V., Kumar, G. A., & Ramola, R. C. (2020). Radiation impact assessment of soil and groundwater of Himalayan regions in Uttarakhand, India. *Journal of Radioanalytical and Nuclear Chemistry, 323*, 1269–1282.

Reyes-Téllez, E. D., Parrales, A., Ramírez-Ramos, G. E., Hernández, J. A., Urquiza, G., Heredia, M. I., & Sierra, F. Z. (2020). Analysis of transfer functions and normalizations in an ANN model that predicts the transport of energy in a parabolic trough solar collector, Desalin. *Water Treatment, 200*, 23–41.

Robert, J., & Howlett, L. C. J. (2001). Radial basis function networks 2: New advances in design.

Sahin, L., & Cavas, M. (2008). Natural radioactivity measurements in soil samples of Central Kutahya (Turkey). *Radiation Protection Dosimetry, 131*, 526–530.

Sarker, M. S. D., Rahman, R., Siraz, M. M. M., Khandaker, M. U., & Yeasmin, S. (2021). The presence of primordial radionuclides in powdered milk and estimation of the concomitant ingestion dose. *Radiation Physics and Chemistry, 188*, 109597.

Segal, R., Kothari, M. L., & Madnani, S. (2000). Radial basis function (RBF) network adaptive power system stabilizer. *IEEE Transactions on Power Systems, 15*, 722–727.

Shabbir, T., Sohail, M., Ahmed, R., et al. (2023). Radiological risk assessment from the Gamma spectrometry of soil samples in the vicinity of a research centre in Nilore, Pakistan. *Environmental Monitoring and Assessment, 195*, 851.

Shi, C., & Wang, Y. (2021). Non-parametric machine learning methods for interpolation of spatially varying non-stationary and non-Gaussian geotechnical properties. *Geoscience Frontiers, 12*, 339–350.

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods* (8th ed.). Iowa State Univ Press.

Szczurek, A., & Maciejewska, M. (2004). Recognition of benzene, toluene and xylene using TGS array integrated with linear and non-linear classifier. *Talanta, 64*, 609–617.

Tabar, E., Yakut, H., Saç, M. M., et al. (2017). Natural radioactivity levels and related risk assessment in soil samples from Sakarya, Turkey. *Journal of Radioanalytical and Nuclear Chemistry, 313*, 249–259.

UNSCEAR. (2008). *Sources and efects of ionizing radiation*. Report to General Assembly, Annex B, United Nations, New York.

UNSCEAR. (2000). *Sources and effects of ionizing radiation*. United Nations.

Verma, S. P. (2005). *Estadística Básica para el Manejo de Datos Experimentales: Aplicación en la Geoquímica (Geoquimiometría)*. Universidad Nacional Autónoma de México.

Wadoux, A. M. J. C., Minasny, B., & McBratney, A. B. (2020). Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Science Reviews, 210*, 103359.

Yaprak, G., & Aslani, M. A. A. (2010). External dose-rates for natural gamma emitters in soils from an agricultural land in West Anatolia. *Journal of Radioanalytical and Nuclear Chemistry, 283*, 279–287.

Yaprak, G. (1995). Matrix effects on gamma spectrometric analysis of radioactive materials and development a self absorption correction method. Ph.D. Thesis, Ege University, İzmir, Turkey, 81 pp (Turkish with English abstract).

Yu, H., Xie, T., Paszczynski, S., & Wilamowski, B. M. (2011). Advantages of radial basis function networks for dynamic system design. *IEEE Transactions on Industrial Electronics, 58*, 5438–5450.