scientific reports

Check for updates

OPEN

Edaphic and meteorological parameters as determinants of radon exhalation and its environmental implication in Peruvian agroecosystems

B. Pérez[1,2,3✉], L. C. Stieff[2,4], R. E. Ponce-Amanca[3,4], C. J. Guevara-Pillaca[3,4] & D. Palacios[1,4]

Radon exhalation is a natural process by which atoms of the radioactive gas radon diffuse in the soil and then exhale to an indoor and/or outdoor environment. High radon concentration levels, possibly from high radon exhalation rate levels, can generate an impact on public health and environmental safety, particularly in agricultural areas where prolonged exposure may affect nearby populations. While studies have examined radon exhalation, few have focused on modeling its behavior in agricultural settings or identifying key environmental and soil parameters that influence its variation. This study addresses this gap by applying Artificial Neural Network (ANN) models and Monte Carlo methods. Three distinct approaches were developed based on radon exhalation measurements from four Peruvian agricultural regions, incorporating meteorological and soil physicochemical data. First, the ANN model determined environmental factors affecting radon exhalation, achieving R^2 values of 0.7949 (training) and 0.7656 (validation). Second, simulations analyzed radon diffusion under varying wind conditions, assessing dispersion risks. Third, gamma radiation measurements quantified radon progeny contributions ($2.82 \times 10^{-4} \pm 1.15 \times 10^{-5}$ efficiency) for soil moisture detection. This integrated methodology advances understanding of agricultural radon dynamics, supporting improved radiological safety protocols and soil monitoring techniques.

Radon (^{222}Rn) is a radioactive noble gas that originates from the decay of ^{238}U in the Earth's crust. A key feature of the radon gas is its move through porous media such as soil, rock fractures, and building foundations. In soil, radon migrates primarily via diffusion and advection through pore spaces, with its movement influenced by soil permeability, porosity and moisture content[1,2]. Once released into the air, it diffuses mainly through diffusion but can also its movement can be affected by advection (wind) and convection (temperature-driven air movement) under certain atmospheric conditions[3]. This makes radon flux–the rate at which radon escapes from the soil–a useful measure for studying underground geological activity and assessing radiation risk.

Historically, researchers have leveraged radon flux measurements to detect uranium deposits[4], map fault zones[5], even predict indoor radon exposure risks[6,7], and tracing carbon dioxide fluxes[8]–demonstrating its versatility as a natural tracer. One notable study by Jinmin Yang in Germany revealed complex interactions between radon flux and environmental factors such as air and soil temperatures, air pressure, and especially soil moisture. Their findings showed that low moisture increases radon flux up to a threshold ($\sim$10%). At higher soil wetness (>10%), the flux rate decreases gradually[9,10].

These observations align with theoretical expectations, as moisture content modulates radon's diffusion pathways. Nonetheless, factors like grain size and porosity also play a role, and their combined effects are still not fully understood[1,2]. For this reason, further independent, site-specific studies are essential to improve our

[1]Departamento de Ciencias, Pontificia Universidad Católica del Perú, Av. Universitaria 1801, San Miguel, 15088 Lima, Peru. [2]Rad Elec Inc., Frederick, MD 21704, USA. [3]Investigación, Desarrollo e Innovación, Anphysrad SAC, 15074 Lima, Peru. [4]L. C. Stieff, R. E. Ponce-Amanca, C. J. Guevara-Pillaca and D. Palacios: These authors contributed equally to this work. ✉email: bertin.perez@pucp.edu.pe

understanding of radon dynamics across different environments, particularly in agricultural systems where both direct exposure and radon progeny influence radiological safety and measurement accuracy.

Traditionally, mathematical-physical models, including those based on Darcy's Law and diffusion equations, have been used to describe radon transport[11,12]. While these models provide valuable insights, they often rely on simplifying assumptions and may fail to capture the variability observed in real-world environments. Moreover, they do not account for the stochastic nature of the radioactive decay process[13]. Complementarily, geo-statistical methods have been employed to interpolate radon flux data spatially, yet these approaches do not directly account for the physical transport mechanisms underlying radon migration[14–16].

In recent years, more advanced approaches, such as machine learning (ML), have been adopted due to the availability of large datasets and advances in computational capacity. Algorithms such as k-Nearest Neighbors (kNN), Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forests (RF), and GBM (Gradient Boosting Machine), have been successfully applied in environmental prediction tasks. These methods can capture nonlinear and complex relationships among multiple variables directly from datasets without strict assumptions[17–19].

In parallel, Monte Carlo (MC) methods have been used to simulate the stochastic behavior of radon transport at the atomic scale, particularly the random decay processes and trajectories of radon atoms and their progeny. For example,[13] used MC methods to model radon atom pathways and lifetimes, capturing microscopic transport mechanisms beyond the reach of deterministic models. These simulations complement ML analyses by providing fundamental physical insights.

In this study, four radon flux surveys across Peruvian agricultural sites between May and August 2024 were conducted. The relationship between radon exhalation and environmental parameters, including meteorological conditions and soil physicochemical properties, was investigated using ANN. Additionally, MC methods simulated the stochastic transport of radon atoms through air, driven by experimental radon flux survey data. These simulations supported complementary studies on radiological risk from airborne radon and progeny displacement, and the biasing effects of short-lived progeny on soil moisture measurements via gamma-ray spectroscopy[20–23].

Materials and methods

This section describes the fieldwork and analytical methods used to investigate radon exhalation in Peruvian agricultural fields. First, the study area and sampling strategy are presented (Section "Study area"–"Point-selection criterion and sampling"), which provide the foundation for the subsequent measurement systems described in (Section "Measurement system"). Building on these experimental setups, Section "Artificial neural networks implementation" introduces the implementation of the ANN, analysis of influential parameters, multicollinearity assessment, and hyperparameter tuning (Section "Analysis of influential parameters"–"Hyperparameter tuning"). To complement this implementation model, section "Simulation of the transport and decay of radon and its progeny atoms in the air" details MC methods designed to model the physical transport and decay of radon and its progeny atoms in air. Finally, section "Contribution of short-lived gamma-emitting radon progeny to the proximal gamma-ray spectroscopy technique" addresses the influence of short-lived radon progeny on gamma-ray spectroscopy measurements.

Study area

Four fields were carefully chosen for this study to encompass maximum variation of parameters. The selected fields were called 'Nazca', located at approximately $14°49'52.68''$ S latitude and $74°57'31.03''$ W longitude in the Nazca region; 'Yaután', at approximately $9°30'32.72''$ S latitude and $78°00'02.32''$ W longitude in the Casma region; 'Santa Eulalia', at approximately $11°53'39.46''$ S latitude and $76°39'15.42''$ W longitude in the Huarochirí region; and 'La Agraria', at approximately $12°05'02.23''$ S latitude and $76°57'10.59''$ W longitude in the Lima region. These fields were chosen for their diverse geographical characteristics, environmental parameters and soil physiochemical properties. Fig. 1 shows a map created in QGIS of the four study fields, each classified by its geomorphology (https://metadatos.ingemmet.gob.pe:8443/geonetwork/srv/spa/catalog.search#/metadata/ae9d5935-ed4c-46a0-a826-6e0b9d5e20e2), as indicated in the legend.

Nazca is located in the Ica region of southern Peru, characterized by its hyper-arid desert environment (one of the driest regions worldwide) and the iconic Nazca Lines. Table 1 shows its subsurface characteristics. The district had a total population of 27,632, with 25,293 residents in urban areas and 2,339 in rural areas. Agricultural land covered 9,464,153.52 m^2.

Yaután lies in the Casma Valley, within the Ancash region along Peru's northern coast. This area features a coastal desert climate and fertile river valleys. Table 1 summarizes its subsurface characteristics. The total population was 8,350, with 3,309 inhabitants in urban areas and 4,496 in rural areas. Agricultural land accounted for 3,696,394.61 m^2.

Santa Eulalia is situated in the Santa Eulalia Valley (Huarochirí Province, Lima Region), within the Andean foothills of central Peru. The valley is noted for its microclimatic diversity and fertile soils. Table 1 details its subsurface characteristics. The population included 12,636 residents, predominantly urban (11,955 inhabitants). Rural areas had 681 residents, with agricultural land covering 740,155.77 m^2.

La Agraria, located at the Universidad Nacional Agraria La Molina (UNALM) in Lima, serves as a key reference site for agricultural research due to its meticulously managed fields and standardized monitoring protocols. Table 1 provides its subsurface characteristics. The district had 140,679 inhabitants, all urban. Agricultural land spanned 4,653,458.61 m^2 and was classified as an "agricultural mosaic" in the MapBiomas dataset. Notably, while agricultural areas in MapBiomas typically use a pixel value of 18, La Molina's data lacked this specific value, indicating its mosaic classification matches observed field patterns.
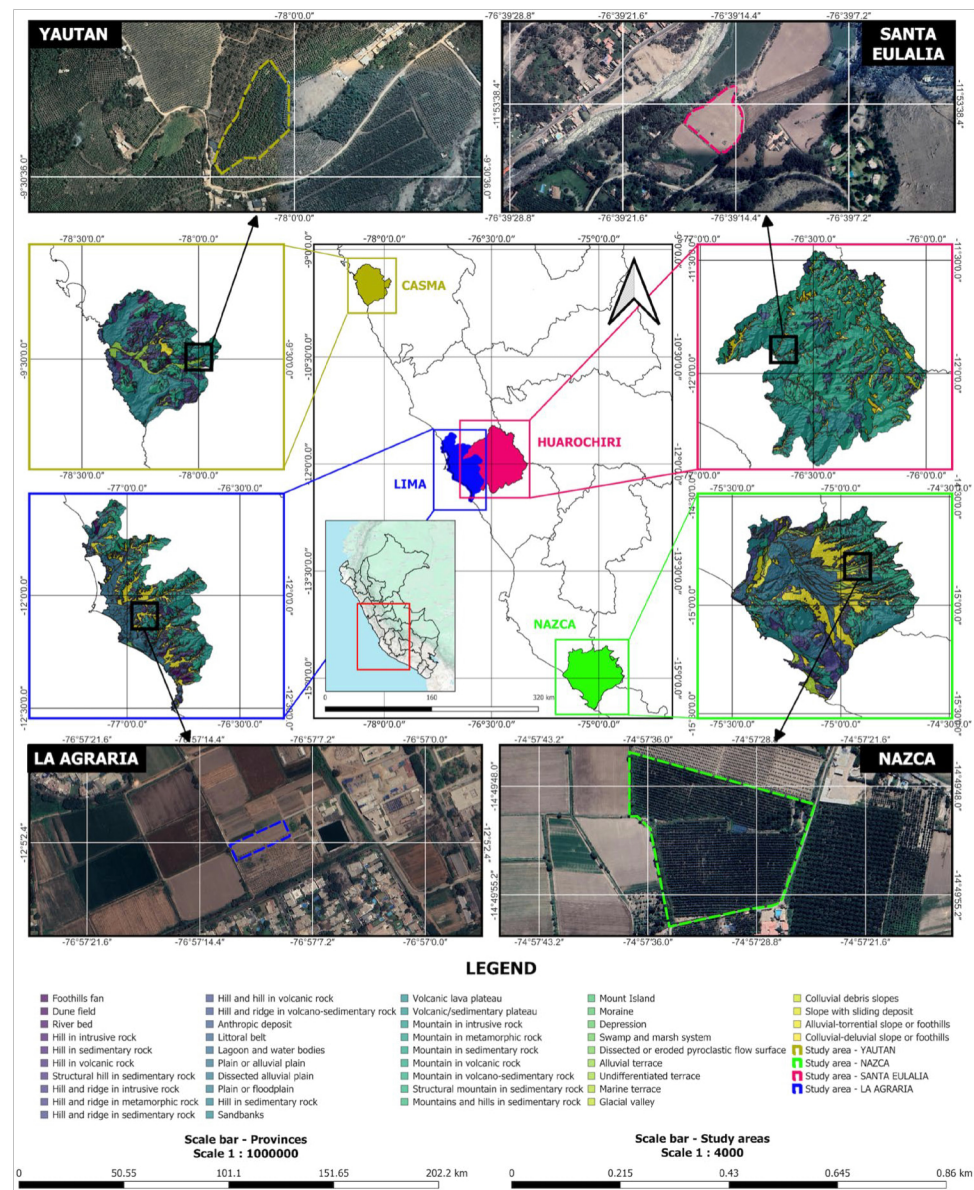
Fig. 1. Map of the four study regions, classified by geomorphology. The studied field regions and areas are outlined with dashed lines in different colors: green for Nazca, olive for Yaután, pink for Santa Eulalia, and blue for La Agraria. Created using Quantum Geographic Information System (QGIS) version 3.36.3.

Fields	Geology	Geomorphology	Source
Nazca	Alluvial deposit, composed of gravel and blocks with a sandy-silty matrix	Alluvial terrace	25
Yaután	Alluvial deposit, accumulation of gravel, sand, silt, and clays with subangular to angular clasts of varying composition	Alluvial terrace	26
Santa Eulalia	Alluvial deposit, accumulations of rock fragments (sands, gravel, boulders, etc.) deposited as ancient terraces	Hillside or alluvial-torrential piedmont	27
La Agraria	Alluvial deposit, accumulations of rock fragments (sands, gravel, boulders, etc.) deposited as ancient terraces	Hillside or alluvial-torrential piedmont	27

Table 1. Summary of geological formations, geomorphological features, and bibliographic references for the selected fields.

The demographic and agricultural data for the districts were analyzed using the 2017 National Census data[24] and land cover rasters obtained from MapBiomas Perú (https://peru.mapbiomas.org). Data processing was performed using QGIS.

To ensure representative spatial coverage across the identified study area, a point-selection strategy was applied based on both environmental variability and logistical feasibility. The next section outlines the criteria and approach used to select these sampling points.

Point-selection criterion and sampling

The sampling strategy was designed based on prior studies validating similar spatial scales for capturing soil radon variability.[28] demonstrated that 20 m spacing along transects (up to 2 km) effectively resolves radon variations linked to lithology, elevation, and edaphic properties.[29] further supported this scale, reporting coefficients of variation (15–30%) within 20×20 m grids, which informed the minimum number of sampling points required for reliable radon field characterization. This configuration optimally balances spatial heterogeneity representation with systematic coverage, minimizing data redundancy while ensuring reproducibility. Based on this evidence, each field was subdivided using a grid system, and each grid cell comprising an area approximately 400 m^2 in size. Based on this subdivision, Nazca was divided into a certain number of grid cells, resulting in approximately 199 grid nodes; Yaután into 43 grid nodes; Santa Eulalia into 18 grid nodes; and La Agraria into 8 grid nodes. Each grid node represents a measurement point.

At each measurement point, measurements of radon exhalation, meteorological parameters, soil moisture, and ground-level gamma radiation were taken. The center of each grid cell was selected for measurements of soil gas radon and thoron concentrations, as well as soil permeability. This assumes that each grid cell has a spatial physiochemical homogeneity of the soil. In other words, soil gas radon/thoron measurements and permeability were measured every 20-meter interval, where each set of four radon exhalation measurements is associated with one measurement of soil gas radon/thoron and permeability. This methodology offers higher precision than the 100-meter interval used by[30] to train the model, and offers better quality in data variability. Furthermore, 20-meter intervals ensure that individual measurements of soil gas radon/thoron and permeability do not interfere with one another, because the sphere from which the air from the soil is pumped around the buried probe has an approximate radius of 12 cm. This was described by[31] when using a device that extracts soil air through a probe.

For soil sampling, we employed a method suggested by UNALM[32]. Given the relative uniformity of the fields, a randomized sampling technique was applied to collect small portions of soil at a depth of 10 cm, avoiding fertilizers, accumulated plant material, or manure. These samples were taken from various locations throughout each studied area. These smaller portions were then combined to form a composite 1 kg sample, ensuring a representative mixture of the soil across the entire field. Care was taken to avoid sampling areas with visible biomass or fertilizer to ensure the sample was representative of the physiochemical properties of the soil. Although only a mixed soil sample was considered for each agricultural area, which can be a relevant parameter influencing radon exhalation. The largest study area, Nazca, is crossed by active geological faults[33,34], which would make possible an increase in the levels of radon exhalation in the areas adjacent to these[5,35]. The UNALM laboratory analyzed the following categories: soil physical properties, soil chemical properties, cation exchange and saturation, and soil soluble ions and saturated paste parameters.

The parameters involved in the study encompass a wide range of multiple scientific domains, as shown in Table 2.

Following the selection of sampling locations, measurements were conducted using a combination of field instruments and laboratory analyses. The following section describes the systems and protocols used to quantify radon flux and characterize relevant edaphic and meteorological parameters.

Category	Parameters
Soil Physical Properties	● Permeability (m^2) ● Bulk Density (g cc^{-1}) ● Sand (%) ● Silt (%) ● Clay (%)
Soil Moisture	● Soil Moisture (%)
Soil Chemical Properties	pH (1:1) ● CaCO$_3$ (%) ● Organic Material (%) ● Phosphorus (ppm) ● Potassium (ppm)
Cation Exchange and Saturation	Cation Exchange Capacity (meq in 100g) ● Exchangeable Ca^{2+}, Mg^{2+}, K$^+$, Na$^+$, Al^{+3} + H$^+$ ● Sum of Cations ● Sum of Bases ● Base Saturation (%) ● Saturation (%) ● Exchangeable Sodium Percentage (%)
Soil Soluble Ions and Saturated Paste	pH of Saturated Paste Extract ● EC of Saturated Extract (dS m^{-1}) ● Soluble Ca^{2+}, Mg^{2+}, K$^+$, Na$^+$ (meq L^{-1}) ● Sum of Soluble Cations ● Soluble NO$_3^-$, CO$_3^{2-}$, HCO$_3^-$, SO$_4^{2-}$, Cl$^-$ (meq L^{-1}) ● Sum of Soluble Anions ● Soluble Boron (ppm) ● Soluble Gypsum (%)
Radon and Thoron Measurements	Radon Exhalation (Bq m^{-2}day^{-1}) ● Radon Concentration (Bq m^{-3}) ● Thoron Concentration (Bq m^{-3})
Atmospheric and Weather Conditions	Air Temperature (°C) ● Relative Humidity (%) ● Absolute Humidity (g m^{-3}) ● Pressure (mmHg) ● Wind Velocity (km h^{-1}) ● Dew Point (°C)
Thermal and Comfort Indices	Thermal Sensation (°C) ● Hot Flash Index (°C)
Solar and Light-Related Variables	UV Index ● Illuminance (lx) ● Solar Irradiance (W m^{-2}) ● Solar PAR (μmol m^{-2}s^{-1})
Magnetic and Electrical Properties	Magnetic Heading (°) ● True Heading (°) ● Electrical Conductivity (dS m^{-1})
Topographical Data	Altitude (m)

Table 2. Categorization of 56 parameters (environmental, geological, and soil properties) used in radon exhalation modeling.

Measurement system

Several instruments were employed to collect different parameters. Given the extensive number of measurements points, which exceeding the available flux monitors, in each field, we developed a strategic methodology using only nine radon flux monitors. In each field, we deployed these monitors for a two-hour measurement period. During this time, we visited each of the nine points sequentially to record meteorological parameters, ground-level gamma radiation, and soil moisture. The two-hour duration allowed approximately 10 minutes at each point to capture these additional measurements before moving to the next location.

The radon exhalation rate was measured using nine flux monitors equipped with electrets from Rad Elec, USA. Each monitor has a chamber volume of 960 mL and a base area of 283 cm^2, with a stainless steel collar that is partially buried to ensure a tight seal with the soil, preventing leakage and minimizing external air interference. The flux monitors were placed on the soil surface at each point for an exposure period of two hours. This exposure time was determined based on the manufacturer guidelines, which suggest that there must be a voltage loss of at least 20 volts, and a pilot testing. The electrets inside the flux chambers are sensitive to background gamma radiation[36]; therefore, to ensure accuracy, natural gamma radiation in nGy h^{-1} was also measured at each location to correct for its influence on the electret measurements. The methodology and calculation of the radon exhalation rate are described in[37].

To measure gamma dose rate, a radiation detector from Gamma-Scout, Germany, was used. This device was placed at the soil surface near the radon flux monitor for approximately 10 minutes to obtain a reading gamma radiation levels at each measurement point. According to other studies, stable values from measurements are obtained with no fluctuations greater than ±10%[38], and 3 minutes are sufficient to obtain consistent data[39]. At the same time, meteorological data, including air temperature, relative humidity, absolute humidity, pressure, wind speed, dew point, wind chill, humidex, UV index, illuminance, solar irradiance, and solar PAR, were recorded for the same 10-min period using a wireless weather sensor with GPS from PASCO. Soil moisture levels were assessed using a Teros 10 sensor from Meter Group, USA. The Teros 10 was placed around each flux monitor at least five times at a depth of approximately 10 cm, which is a critical layer for both agricultural water management[40,41] and radon exhalation studies. This depth is particularly important because soil moisture at 10 cm (the final 'layer' before reaching the atmosphere) can directly influence the amount of radon escaping from soil into the atmosphere.

The Rad7 monitor from Durridge, USA, equipped with a stainless steel probe, was used to measure soil gas radon and thoron concentrations. The probe was inserted into the soil to a depth of approximately 80 cm at each measurement point (the center of each grid cell) for approximately 30 minutes to reach equilibrium between radon atoms and their progeny[5]. During the measurement period, the sniff mode method was applied, where the device provided readings every 5 minutes. The final result was the average of the highest stable values. Typically, soil radon concentration values initially increase and then reach a stable level during a specific time of measurement. After each measurement, a purging process with radon-free air was performed to clean the chamber monitor and remove any residual radon progeny, preventing cross-contamination between measurements. In addition, the lower detection limit is 4 Bq m^{-3}, according to the manufacturer's specifications. The Radon-Jok device from radon v.o.s., Czech Republic, was employed to measure soil permeability for radon gas at the same locations where soil gas radon and thoron concentrations were measured. Soil permeability directly influences the movement of gases within the soil, which in turn affects radon exhalation rates. In this work, values of soil permeability to radon gas has been grouped into high, medium and low[42]. These measurements were critical for understanding how the physical properties of the soil contributed to gas transport processes[43].

The collected data served as the input for data-driven modeling. Building on the experimental setup, the next section introduces the implementation of ANN, designed to model the nonlinear relationships between radon exhalation and Edaphic and meteorological parameters.

Artificial neural networks implementation

To develop model structure and input selection, an initial analysis was conducted to identify the most influential parameters affecting radon flux. This step aimed to enhance interpretability and reduce data dimensionality.

Analysis of influential parameters

The selection of factors is a critical step in constructing ANN models. This study considered six independent factors related to soil properties and environmental conditions. Pressure directly influences radon exhalation from soil. When atmospheric pressure decreases, a pressure gradient is created between the soil, where pressure is slightly higher, and the external air, driving radon migration into the atmosphere[9]. Increased temperature raises the kinetic energy of particles, accelerating diffusion processes, which means radon moves more rapidly through soil pores to the surface at higher temperatures[1,9]. Solar irradiance affects soil moisture through evaporation. Lower soil moisture increases permeability, intensifying radon exhalation. In dry soils, air-filled pores facilitate radon diffusion, resulting in high exhalation rates[44]. Permeability, which defines the ability of soil to allow fluid flow, directly influences radon migration from soil to surface. Soils with large, well-connected pores exhibit higher permeability, enhancing radon migration[45]. Bulk density significantly affects radon exhalation. Higher bulk density, typically from soil compaction, reduces pore space and connectivity, lowering soil permeability and restricting radon diffusion to the atmosphere[46].

One problem that can commonly arise is multicollinearity. This problem arises when two or more predictor variables are highly correlated. Using highly correlated or redundant variables can negatively affect model performance and interpretability[47]. This in turn can generate instability in the predictions and reduce the accuracy of the model. To avoid this, a multicollinearity analysis can be performed and correlation coefficients can be used to filter, identify and select the most relevant predictors[48]. From an initial set of 56 variables, a Pearson correlation analysis was conducted, for this study a maximum acceptable correlation threshold of 0.5

was established[19], ensuring that redundant or highly correlated factors were excluded. This step helped preserve the independence of variables and prevented potential instability in the model's performance. Additionally, variables derived from laboratory soil analyses that provided constant values for each sampling zone were excluded, as they lacked the variability needed to explain differences in radon exhalation.

Moreover, due to the inherent variability of environmental measurements, errors and uncertainties were inevitably present in the data. These variations could be attributed to natural geological factors, such as faults in the Nazca region, or occasional measurement errors (incorrect electret readings). To mitigate the influence of outliers, the 1.5 times interquartile range (IQR) rule[49] was applied. This method identified and excluded extreme values, ensuring a more robust analysis and enhancing the reliability of the results.

Thus, the variability in radon exhalation is determined by the combination of these factors, which interact nonlinearly. Traditional modeling methods struggle to capture these interactions fully, making machine learning techniques valuable tools for understanding the mechanisms governing radon exhalation.

To avoid redundancy and ensure statistical independence among predictors, multicollinearity analysis and mutual information (MI) scoring were then applied. These techniques supported the refinement of input features for robust ANN performance.

Multicollinearity analysis and mutual information (MI)
Using highly correlated or redundant predictors can negatively affect model performance[47]. To address this, multicollinearity analysis and correlation coefficients were used to identify and select the most relevant predictors. The study only considered physical characteristics of soil and environmental parameters influencing radon exhalation behavior, as shown in Table 3. In addition, values obtained from laboratory that provided a single value for each sampling area were excluded from the analysis. This decision is justified by the constant nature of these variables, which do not contribute variability data and, therefore, do not contribute to differences observed in radon exhalation.

Table 3 presents the parameters measured at the various study locations. A Pearson correlation analysis was performed to assess the relationships between soil radon exhalation values and the six influencing factors. The results revealed weak to moderate linear correlations among the factors (Fig. 2).

Moderate positive correlations were observed between solar irradiance and air temperature (0.42), pressure and bulk density (0.37), and soil moisture and pressure (0.30), indicating some degree of dependence between these variables. Additionally, weak positive correlations were found between bulk density and soil moisture (0.11) and air temperature and soil moisture (0.26). On the other hand, weak negative correlations were identified between permeability and pressure (-0.28), solar irradiance and pressure (-0.11), and bulk density and permeability (-0.14), suggesting that these variables do not exhibit a clear linear relationship. Overall, the results suggest that most of the variables are independent of one another.

As previously mentioned, a potential issue that may arise is multicollinearity, which occurs when two or more predictor variables are highly correlated. This can significantly affect the performance and interpretability of the model, leading to instability in predictions and reduced accuracy. To ensure the absence of significant multicollinearity effects among the predictors, a multicollinearity analysis was performed, which was used as a method for filtering predictors[48]. The multicollinearity analysis was performed using Tolerance (TOL) values and the Variance Inflation Factor (VIF). The criteria used to filter these predictors were VIF > 5 and TOL < 0.2[19,48,50]. This approach ensures that predictors contributing to multicollinearity are excluded, improving the robustness and reliability of the model, as shown in Table 4.

Table 4 presents the results of the multicollinearity analysis. The VIF values range from 1.113 to 1.391, and the TOL values range from 0.718 to 0.897 for all factors. Permeability exhibited the weakest linearity (VIF = 1.1137), while pressure had the highest VIF value (1.3918). It is also essential to determine the predictive capacity of the selected predictors. For this purpose, techniques such as Information Gain Ratio (IGR) can be used[19]. However, in this study, IGR is not directly applicable because it is designed for classification tasks with discrete target variables, measuring how much information a feature provides about the target class to optimize classification results through feature selection[51]. This would require prior discretization of the target dataset (radon exhalation). Nonetheless, there are alternative approaches for evaluating the informational relationship between continuous variables, as in our case. Mutual Information (MI) is a measure from information theory that quantifies the dependency between two variables (datasets), regardless of whether the relationship is linear or not. It can detect any type of relationship, including mean values, variances, or higher-order moments[52]. The calculation of MI was implemented using Python's scikit-learn library with a function that employs the kNN algorithm to estimate probability densities and calculate MI, with the number of neighbors k = 3. This approach

Category	Parameter	Data Source
	Bulk Density	Soil sample analysis
Physical Characteristics of Soil	Soil Moisture	Teros 10 sensor
	Permeability	Radon Jok
	Pressure	
Environmental Parameters	Air Temperature	Pasco weather sensor
	Solar Irradiance	

Table 3. Optimized parameters selected for ANN model construction after multicollinearity analysis and correlation-based predictor selection.
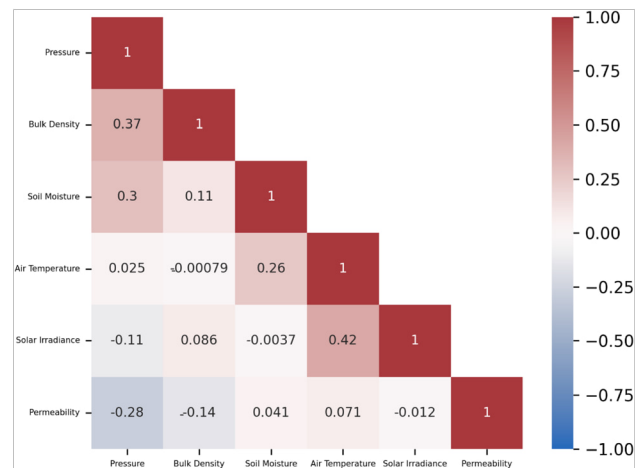
Fig. 2. Pearson correlation matrix of selected predictors influencing radon exhalation rates, with significance thresholds (p < 0.05).

Parameter	MI Score	VIF	TOL
Pressure	0.161436	1.391807	0.71849
Bulk Density	0.155884	1.187387	0.842186
Soil Moisture	0.070415	1.202375	0.831687
Air Temperature	0.063171	1.334116	0.74956
Solar Irradiance	0.029709	1.283447	0.779152
Permeability	0.003805	1.11374	0.897876

Table 4. Results of the multicollinearity analysis and contributions of the influential factors. Collinearity Statistics include VIF (Variance Inflation Factor) and TOL (Tolerance).

allows for handling continuous data without the need for manual discretization[52]. The MI results indicate that the most important factors are pressure (MI = 0.1614) and bulk density (MI = 0.1558).

With the finalized input set, the ANN architecture was optimized using hyperparameter tuning. This process aimed to improve generalization and predictive accuracy across the training and validation datasets.

Hyperparameter tuning
Once the multicollinearity relationships between predictors were identified and addressed, the next critical step was the model optimization phase. Hyperparameter tuning becomes crucial in this context, as it is essential for building highly performant deep neural networks. This is an iterative process of trial and error aimed at finding the optimal combination of values for the configurable parameters of a machine learning model[53]. These parameters, which are not learned during training, can significantly impact the model's performance. Some are related to the structural configuration of the network, while others are associated with the learning algorithms[54].

In this context, Keras Tuner was used to efficiently explore hyperparameter ranges, testing methods like Grid Search, Random Search, Hyperband, and Bayesian Optimization[55]. Bayesian Optimization was selected for its ability to model smoothness of the hyperparameter configuration space to guide the search more efficiently by building a predictive model based on previous evaluations, strategically balancing exploration and exploitation by leveraging past evaluations–unlike Random Search or Hyperband, which operate with less or no dependency between successive evaluations[56]. By automating this process, Keras Tuner facilitates the identification of architectures and model configurations that maximize both accuracy and generalization. The tuned hyperparameters in this study included the number of layers, number of units per layer, dropout rates and learning rate.

The final neural network architecture, which was optimized through this process is illustrated in Fig. 3, consists of a combination of different types of layers to address the regression task. Specifically, it included GRU (Gated Recurrent Unit) recurrent layers, dropout layers (Dropout) for regularization, and densely connected layers (Dense). The hidden layers have ReLU activation functions, their structure was determined by the search algorithm. Finally, the output layer, designed to provide the final prediction, consists of a single dense unit with linear activation function. The model was compiled using the Adam optimization algorithm with a final learning rate of 0.002 and Mean Squared Error (MSE) loss function.

For training and subsequent validation of this model, the data collected from the four measurement locations were combined into a single data set. Before splitting, preprocessing was performed to improve the quality and scale of the data. First, outliers were identified and treated using the 1.5 times interquartile range (IQR) criteria.
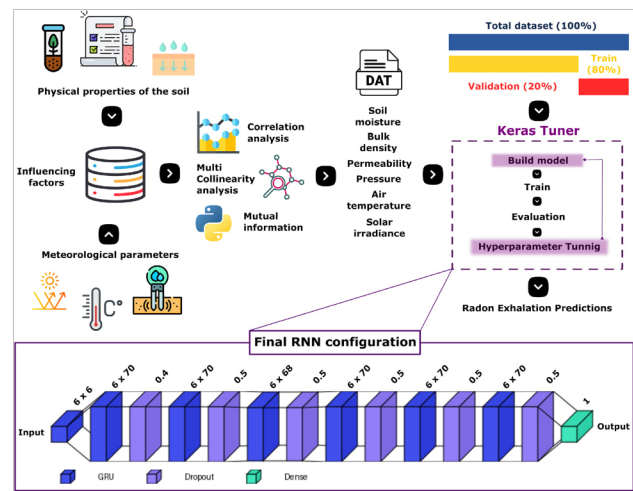
Fig. 3. Detailed workflow for ANN model construction and evaluation, including GRU layers, dropout regularization, dense layers with ReLU activation, and hyperparameter optimization via Keras Tuner.

Parameter	Set	Min	Max	Mean	SD	Q1	Q3	Skewness
Pressure (mmHg)	Train	670.42	744.43	707.86	13.02	710.38	712.22	− 0.08
	Val	670.17	744.38	705.35	19.05	693.57	712.53	− 1.12
Bulk Density (g cc^{-1})	Train	0.7	1.54	1.21	0.18	1.09	1.33	− 0.28
	Val	0.7	1.46	1.14	0.20	0.98	1.27	− 0.22
Soil Moisture (%)	Train	0	54.5	15.22	10.72	6.80	21.60	2.08
	Val	2.46	44.6	12.00	9.84	6.02	14.40	0.92
Air Temperature (°C)	Train	16.71	37.46	27.23	4.87	22.82	31.18	0.20
	Val	18.32	34.15	25.55	4.65	22.15	28.88	− 0.05
Solar Irradiance (W m^{-2})	Train	0	1060.77	195.1	304.93	14.83	225.10	1.64
	Val	0.02	941.44	204.65	327.79	11.44	206.83	1.59
Permeability	Train	0	2	2	0.51	2	2	− 3.71
	Val	0	2	2	0.46	2	2	− 2.16

Table 5. Descriptive statistics of selected predictors post-multicollinearity analysis. The permeability parameter is classified as low = 0, medium = 1, and high = 2.

Next, to prepare the input features for the neural network, they were normalized by scaling the data to the range [0, 1]. Subsequently, the combined set was split by a standard method using the *train_test_split* function of the scikit-learn library obtaining 80% of the data for training and the remaining 20% for validation. While this random split method ensures a random distribution of features across the board, it is important to note that it does not proportionally represent each of the four measurement locations in the resulting sets. To assess the representativeness and statistical comparability of these subsets, a descriptive analysis was performed. Table 5 presents descriptive statistics for each variable in the training and validation sets. The results show that, although most of the variables present very similar statistics, there are some cases in certain statistics that present more notable differences, the impact of which will be evaluated in subsequent sections. The metrics to be used for the final evaluation of the model will be Mean Absolute Error (MAE), Mean Squared Error (MSE–RMSE) and Mean Absolute Percentage Error (MAPE).

While the ANN approach captures complex empirical relationships, it does not explicitly account for underlying physical mechanisms. To address this, MC methods were applied to model the stochastic transport and decay of radon and its progeny atoms in the air.

Simulation of the transport and decay of radon and its progeny atoms in the air

In this section, we describe the simulation of Brownian diffusion of radioactive atoms using a Python script. The simulation models the transport of the radon and its progeny based on the stochastic Langevin equation, with the solution provided by[13]. Although radon transport in air involves diffusion, advection, and convection, this study considers only Brownian diffusion and advection, as convection primarily affects radon transport at higher altitudes and larger atmospheric scales, rather than in the near-surface boundary layer[57]. The solution for the one-dimensional Brownian motion of a radioactive atom, in the absence of drift, reduces to:

$$x(t + \Delta t) = x(t) + \sqrt{2D\Delta t} \cdot g \tag{1}$$

where x(t) is the atom's position along the x-axis at time t, D = 0.054 cm^2 s^{-1}[58], Δt is the time steps, and g is a Gaussian random number with zero mean and unit variance. These Gaussian random numbers were generated using the Box-Muller transform[59]. According to Box and Muller, g can be expressed as $\sqrt{-2 \cdot \ln(\xi_a)} \cdot \cos(2\pi \cdot \xi_b)$, where ξ_a and ξ_b are independent uniform random numbers between 0 and 1.

The choice of Δt, the simulation time step, is critical in accurately modeling atom diffusion. A smaller Δt allows the simulation to capture finer details of the Brownian motion and decay process, yielding a more accurate depiction of the transport of radon or its progeny from soil-air interface to air. However, too small a Δt can make the simulation computationally expensive, while a larger Δt may oversimplify atom movement and potentially miss crucial events, such as the decay of short-lived progeny. Thus, a step of 20 seconds was selected to strike an optimal balance between accuracy and computational efficiency.

In this simulation, we first estimate the number of radon atoms needed to achieve the measured radon exhalation rate. Given the measured exhalation rate in Bq m^{-2} day^{-1} we use this value to calculate the total number of radon atoms exhaled from a 1 m^2 area over a day, accounting for the radioactive decay properties of ^{222}Rn. The number of radon atoms exhaled per day from a 1 m^2 area can be estimated by using the relationship between the activity and the number of atoms present in a decaying state:

$$\text{Number of Radon Atoms} = \frac{Activity}{\lambda} \tag{2}$$

where the activity is in Bq, and λ is the radon decay constant derived from its half-life.

Next, we use random sampling to simulate the lifespan of each radon atom and its progeny. The lifespan of radon ($t_{lifespan}$) and each progeny atoms are calculated using the expression:

$$t_{lifespan} = -\tau \cdot \ln(1 - \xi) \tag{3}$$

where τ is the mean lifetime, and ξ is a uniform random number between 0 and 1[58]. This method provides a realistic estimate of each atom's total life until decay, enabling the simulation to track radon decay through its progeny until ^{210}Pb. The simulation stops at ^{210}Pb since it's the first long-lived progeny (22.3 year half-life). The preceding short-lived progeny pose the greatest radiological hazard due to their intense alpha emissions, while long-lived progeny slow decay makes it far less significant for radiation exposure assessments.

The flow diagram in Fig. 4 outlines the process, starting from the creation of a radon atom on the exhalation surface. Its transport and decay occur, as for its progeny, until the last short-lived progeny decay, ^{214}Po.

To include the influence of air movement on radon and its progeny, we introduced an additional directional component to each atom's position. This was modeled by adding the x component of the velocity that represents the wind speed in the x-direction (plane parallel to the surface soil) at each time step[60]:

$$x(t + \Delta t) = x(t) + \sqrt{2D\Delta t} \cdot g + v_x \cdot \Delta t \tag{4}$$

where v_x is the wind velocity in cm s^{-1}. This approach captures both the random diffusion of atoms and their drift due to wind, providing a more comprehensive model of radon and its progeny transport. The coordinates y and z are determined in analogous way. In addition, due to the stochastic nature of the MC method, the simulation was repeated ten times to statistically estimate the uncertainties associated with this process.

In atmospheric environments, radon progeny atoms typically exist in two states, as unattached and attached to aerosols[61]. Although unattached atoms persist in the air for a very short time (<1 seconds)[62], meaning the majority are in the attached state, this simulation simplifies the process by not accounting for these different states. This choice is made for computational efficiency, but future work will aim to incorporate these finer details to enhance the model's accuracy.

It should be noted that a constant wind speed to simplify airflow modeling and does not account for attached/unattached progeny to aerosols.
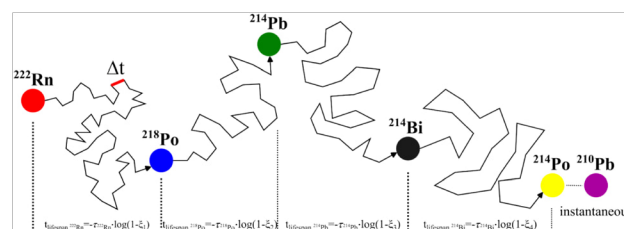

Fig. 4. The schematic flow diagram of Brownian dynamics of ^{222}Rn and its progeny, modeled with steps of duration Δt. The lifespan of each atom is calculated based on its specific mean lifetime, using uniform random numbers ξ_1, ξ_2, ξ_3, and ξ_4. ^{214}Po decays instantaneously to ^{210}Pb.

As radon progeny also contribute significantly to gamma-ray signals in environmental monitoring, their contribution must be considered. The final section evaluates the impact of short-lived gamma-emitting progeny affect on the interpretation of the proximal gamma-ray spectroscopy technique.

Contribution of short-lived gamma-emitting radon progeny to the proximal gamma-ray spectroscopy technique

This proposed simulation, provides valuable insights into the diffusion behavior of short-lived radon progeny, specifically gamma emitters such as ^{214}Pb and ^{214}Bi. These simulations enable the tracking of the final positions of these radon progeny gamma emitters just before decay. By modeling the isotropic emission of gamma rays at each decay position using MC methods implemented in Python, it is possible to estimate the total efficiency of a $3'' \times 3''$ sodium iodide (NaI) detector positioned 2.25 m above a 1 m^2 square ground, with the detector oriented perpendicular to the soil surface, for these emissions.

In the context of gamma spectrometry, the resolution of this NaI detector influences which energy peaks are most effectively analyzed. For this study, the focus is on ^{214}Bi rather than ^{214}Pb, as ^{214}Bi emits gamma radiation with distinct energies that are better suited for measurement[63]. Among its several gamma emission energies, the most probable and well-separated peak, at 609.312 keV, is ideal for minimizing interference from other peaks and it depends on the energy resolution for that peak and the ^{232}Th content due to the possible interference from the 583 keV peak of ^{208}Tl and the 609 keV of ^{214}Bi. Furthermore, if radon is exhaled from the soil, the detector will register the contribution of ^{214}Bi atoms decaying in the soil and those decaying in the air. In this case, an overestimation of ^{214}Bi concentration in the soil will be observed. Therefore, the number of net counts in the peak must be corrected for contributions from airborne ^{214}Bi atoms, ensuring that they are only due to ^{214}Bi atoms in the soil.

According to the outputs obtained from Section "Simulation of the transport and decay of radon and its progeny atoms in the air", the initial decay positions for the ^{214}Bi gamma emitter $(x_{Bi_0}, y_{Bi_0}, z_{Bi_0})$. Gamma rays are assumed to be emitted isotropically, modeled by computing the direction vector components (d_x, d_y, d_z) based on spherical coordinates. These are calculated as:

$$\begin{aligned} d_x &= \sin(\theta)\cos(\phi) \\ d_y &= \sin(\theta)\sin(\phi) \\ d_z &= \cos(\theta) \end{aligned} \tag{5}$$

where ϕ is sampled uniformly from 0 to 2π(azimuthal angle), and $\cos(\theta)$ is sampled uniformly from -1 to 1 to ensure isotropic angular distribution. This approach accurately models the random directions of emitted gamma rays.

The detector is modeled as a finite cylindrical region with radius $r_d = 3.81$ cm, height $h_d = 7.62$ cm, its height from $z_{min} = 100$ cm to $z_{max} = 100 + h_d$ cm and located at its axis perpendicular to the surface soil. For each ray, the simulation computes possible intersections with both the cylindrical surface and its bottom base. For the intersection of the gamma ray with the cylindrical surface, the quadratic equation is solved as folllows:

$$\begin{aligned} a &= d_x^2 + d_y^2 \\ b &= 2(x_{Bi_0}d_z + y_{Bi_0}d_y) \\ c &= x_{Bi_0}^2 + y_{Bi_0}^2 - r_d^2 \end{aligned} \tag{6}$$

The discriminant (Δ) determines whether the ray intersects the infinite cylinder. If the discriminant is greater than or equal to 0, the parametric values t_1 and t_2 are computed as $t_{1,2} = \frac{-b \pm \sqrt{\Delta}}{2a}$, and the corresponding z-values (z_1, z_2) are checked against the cylinder's height bounds $[z_{min}, z_{max}]$. For the bases, the intersection condition is derived from the ray's parametric equation $t_{top,bottom} = \frac{z_{max,min} - z_{Bi_0}}{d_z}$, and the corresponding x- and y-coordinates are evaluated as:

$$\begin{aligned} x_{top,bottom} &= x_{Bi_0} + t_{top,bottom} \cdot d_x \\ y_{top,bottom} &= y_{Bi_0} + t_{top,bottom} \cdot d_y \end{aligned} \tag{7}$$

These intersections (top or bottom base) are valid if the radial constraint is satisfied $(x_{top,bottom}^2 + y_{top,bottom}^2 \leq r_d^2)$. For all valid intersections, the entry and exit points of the ray within the cylinder are determined by the smallest and largest valid parametric values, t_{entry} and t_{exit}. The corresponding 3D coordinates of the entry and exit points are calculated, and the total distance traveled inside the cylinder is computed. The simulation outputs both the total number of rays that intersect the cylinder, which provides the geometrical efficiency, and the total distance traveled by these rays inside the detector. By using the mean distance traveled, the intrinsic efficiency can be calculated by simulating the fraction of gamma rays that transfer energy to the NaI crystal. This fraction is 1 - $\exp^{-\mu d}$, where μ is the total linear attenuation coefficient of NaI (excluding coherent scattering) for the ^{214}Bi energy, and d is the average distance traveled inside the detector (cylinder)[64]. This relationship quantifies the probability of gamma rays interacting with the detector material and transferring energy. In addition, the attenuation in air and in any medium between the gamma emission point and the NaI detector was not considered in this work.
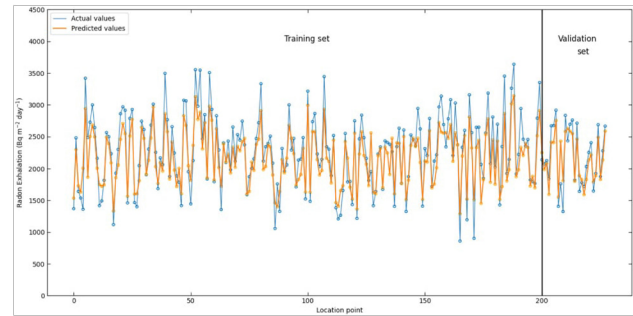
Fig. 5. Prediction results for soil radon exhalation using the ANN model, including the training and validation sets.

Tools	Training	Validation
MAE	0.3717	0.3295
MSE	0.2108	0.1700
RMSE	0.4592	0.4124
MAPE (%)	1.9904	1.2929

Table 6. Error metrics for the ANN model across training and validation datasets.

Using the total efficiency $\eta_{total} = \eta_g \cdot \eta_i$ as fraction, which combines the geometrical (η_g) and intrinsic efficiency (η_i), it is possible to optimize the gamma proximal spectrometry technique, particularly in utilizing the ^{214}Bi gamma peak for evaluating water content in soil. It is important to note that η_{total} cannot be directly combined with the neat counts of ^{214}Bi, because η_{total} uses the entire spectrum. For this reason, the peak-to-total ratio PT(E) becomes particularly relevant. According to[65], PT(E) for ^{214}Bi is approximately 0.60, and the full-energy peak efficiency for ^{214}Bi is calculated as $\eta_F = \eta_{total} \cdot 0.60$. To find the corrected count $N_{corrected}$, it should be calculated as $N_{measured} \cdot \eta_F$.

Results and discussion

The measured radon exhalation rates in agricultural soils ranged from 2040 to 2590 Bq m^{-2} day^{-1} across study sites, with values of 2200 ± 610 Bq m^{-2} day^{-1} in Nazca, 2380 ± 580 Bq m^{-2} day^{-1} in Yaután, 2590 ± 230 Bq m^{-2} day^{-1} in Santa Eulalia, and 2040 ± 260 Bq m^{-2} day^{-1} in La Agraria. These values are significantly higher than those reported for non-agricultural environments[66–69], highlighting distinct agroecosystem-specific trends. This suggests that agricultural fields may enhance radon exhalation compared to natural or urban landscapes.

Artificial neural networks implementation

Fig. 5 displays the soil radon exhalation predictions for the training (80% of the data) and validation (20%) sets, revealing a need to assess uncertainty, especially at extremes. Table 5 shows that predictor variables (pressure, soil moisture, and solar irradiance) differ in distribution (dispersion, ranges, skewness) between the two sets–a known source of model uncertainty. These variations likely contribute to the observed overestimation at low values and underestimation at high values (Fig. 5).

The values for Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE) obtained in this study are presented in Table 6. The results indicate that the model achieved an adequate level of accuracy in predicting radon gas exhalation based on the meteorological parameters and physical properties considered in this study.

The results of the metrics indicate that the model achieved an adequate level of accuracy in predicting radon gas exhalation as a function of the meteorological parameters and physical properties considered in this study. Lower values of MAE indicate good prediction accuracy[70]. On the other hand, the lower the RMSE and MSE values, the closer the prediction values are to the observation values indicating better consistency between model predictions[19,71].

The performances of the training and validation data sets of the ANN model architecture (Fig. 6a and b, respectively) revealed a good reliability of the network.

The high value of R^2 =0.7949 for training and R^2=0.7656 for validation indicate that the input factors selected for the model adequately explain the variability of the target variable around its mean. On the other hand, this represents a good model fit given that the network has learned to generalize the patterns present in the training data and applies them correctly to new data, as shown in Fig. 7.

Fig. 7 displays fitted values against residuals, revealing slight discrepancies despite an acceptable overall R^2. This suggests that the model struggles to fully capture extreme values, likely due to variability and distributional differences in environmental parameters. Similar findings in other studies support this observation, noting that radon prediction errors often arise at elevated concentrations, driven by fluctuations in key factors like
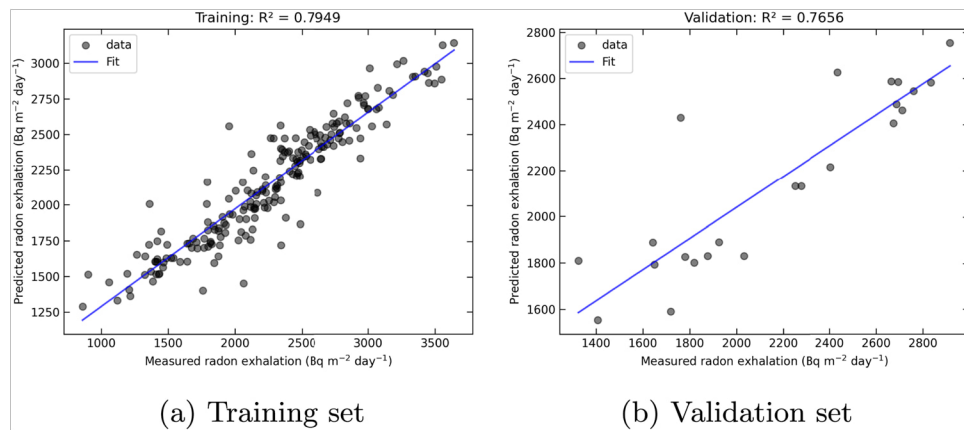
(a) Training set   (b) Validation set

Fig. 6. Prediction results for soil radon exhalation rates (Bq m^{-2}day^{-1}) using the ANN model, with R^2 performance metric.
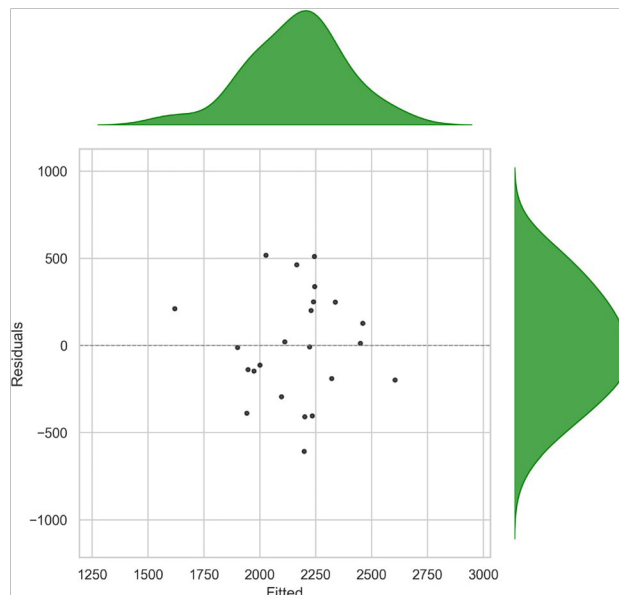

Fig. 7. Adjusted data versus residuals and density histograms.

temperature, pressure, and humidity[18]. Thus, the uncertainty in model predictions, particularly the bias observed in the residuals for low and high values, is intrinsically linked to the variability and characteristics of the distributions of the meteorological and soil property variables used as input.

Simulation of the transport and decay of radon and its progeny atoms in the air

Several studies[13,59,72]–employing similar methodologies in different contexts–have reported results comparable with experimental data, thereby validating our approach. Considering these, our simulation effectively describes radon and its progeny's diffusion dynamics, which yield essential insights into environmental dispersion and potential public health risks. This model's results emphasize two primary perspectives: the diffusion dynamics of radon and progeny atoms, particularly under varying air influence, and the gamma radiation contribution from short-lived radon progeny, particularly above the ground surface.

Diffusion dynamics of radon and progeny atoms, particularly under varying air influence
Utilizing the obtained exhalation rate values and the Equation 2, we estimate radon atom production per unit area (1 m^2), yielding around 1.05×10^9 radon atoms/day on average in Nazca, 1.13×10^9 in Yaután, 1.23×10^9 in Santa Eulalia, and 9.69×10^8 in La Agraria.

Our Monte Carlo simulation employed a 20-second time step (Δt) to model radon and progeny (^{218}Po, ^{214}Pb, ^{214}Bi) diffusion-decay dynamics. This temporal resolution accurately captures radon's decay chain evolution, though progeny modeling becomes less precise due to their shorter half-lives while remaining physically valid.

Using a simplified static-air model (Fig. 8), the spatial distributions, to understand radon and progeny behavior, were analyzed and provide complementary data for health risk assessment.

Fig. 8 shows the distributions of nuclides remaining after one day (excluding ^{210}Po, which decays fully). For radon, 16.53±0.10% of atoms decay, equivalent to $1.73{\times}10^8$ decays day^{-1} in a 256 m^3 outdoor space in Nazca (ignoring wind effects). Converting decay rates to activity for each study site yields outdoor radon concentrations of ranged from 7.24 to 9.17 Bq m^{-3}, as summarized in Table 7. These values, though low and subject to environmental fluctuations, confirming minimal outdoor health risks.

As an illustrative extension of the outdoor analysis, we modeled an hypothetical potential radon accumulation in a simplified indoor environment representative of a rural dwelling with a 4×4 m footprint and a 3 m ceiling height (48 m^3 total volume). This scenario assumed zero ventilation and a bare soil floor, representing a worst-case accumulation condition. Under these assumptions, estimated indoor radon concentrations ranged from 170 to 216 Bq m^{-3} across the four sites, with Santa Eulalia exceeding Peru's national indoor action level ([73]: 200 Bq m^{-3}), as also summarized in Table 7. Santa Eulalia, with the highest estimated indoor level (Bq m^{-3}) and a largely urban population near agricultural zones, poses particular concern.
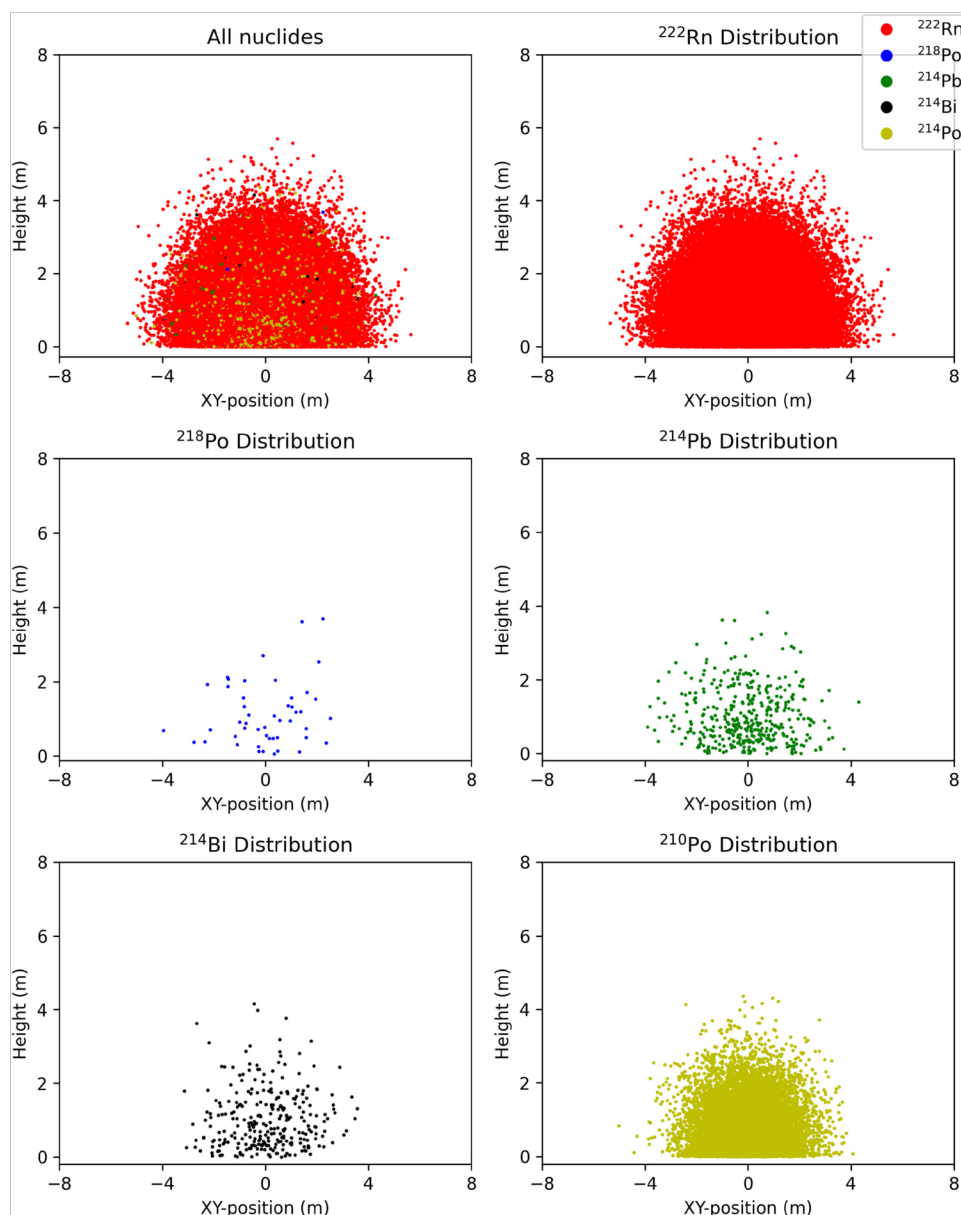
Fig. 8. Spatial distribution of radon (^{222}Rn) and its progeny (^{218}Po, ^{214}Pb, ^{214}Bi, and ^{210}Po) exhaled from a one-square-meter soil area over one day, assuming no air movement. The XY-plane represents the soil surface (in meters) with the vertical axis indicating height. The upper-left subplot displays the combined distribution of all nuclides, providing an overview of their dispersion, while the remaining subplots show individual distributions for each nuclide.

Fields	Outdoor concentration (Bq m^{-3})	Indoor concentration (Bq m^{-3})
Nazca	7.83 ± 0.01	184 ± 31
Yautan	8.43 ± 0.01	199 ± 33
Santa Eulalia	9.17 ± 0.01	216 ± 36
La Agraria	7.24 ± 0.01	170 ± 28

Table 7. Estimated radon concentrations for each study site under outdoor and hypothetical indoor scenarios. Outdoor values correspond to modeled radon concentrations in a 256 m^{-3} open environment without wind influence. Indoor values are based on a worst-case scenario assuming a 48 m^{-3} enclosed room with bare soil flooring and zero ventilation.
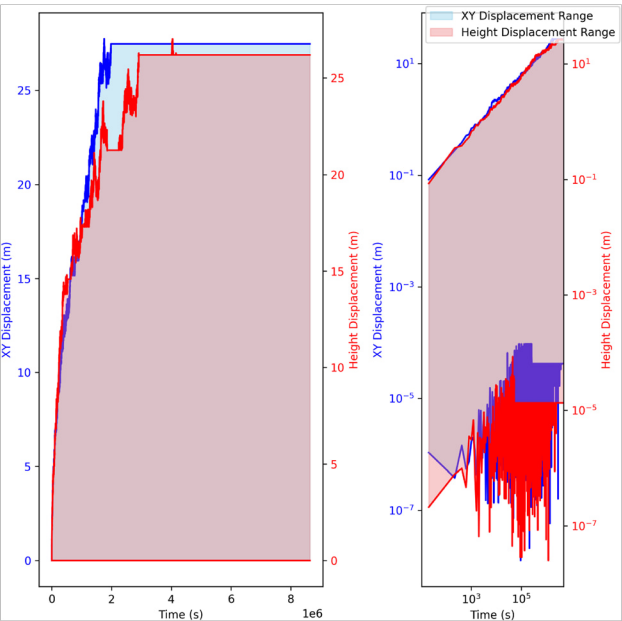
Fig. 9. Displacement of radon atoms and their progeny over time without considering air movement, with shaded regions representing the range between the lower (bottom lines) and upper (top lines) displacement values for both horizontal (XY plane) and vertical (height) movement. The left subplot shows the entire time from 0 to 8×10^6 seconds with steps of 2×10^6, while the right subplot focuses on a shorter period, from 0 to 5e^6 seconds.

The demographic profiles of the study regions provide important context for evaluating potential public health implications of both outdoor and hypothetical indoor radon exposure. Rural populations range from 681 (Santa Eulalia) to 4,496 (Yaután), with Nazca and La Molina being predominantly urban. While outdoor radon concentrations remained quite low, modeled indoor concentrations under worst-case conditions (Table 7) are within a close range of Peru's 200 Bq m^{-3} action level. Although these values are based on simplified assumptions–such as zero ventilation and full soil exposure–they offer a conservative baseline for identifying areas where further assessment may be needed. In particular, rural dwellings with limited airflow and direct soil contact in areas like Nazca and Yaután may warrant targeted indoor monitoring and implementation of low-cost mitigation strategies, such as improved ventilation or soil sealing, to minimize potential health risks.

To understand the atoms' diffusion over time and how far or high atoms move, Fig. 9 presents the total displacement of the nuclides during a day in both the XY and vertical directions, from the point of radon exhalation to the final decay of ^{214}Po.

Fig. 9 depicts the displacement of radon and its progeny over time, the shaded regions in the figure illustrate the range of displacements for both the horizontal (blue) and vertical (red) directions, highlighting its variability across time. During the initial period, fluctuations occur as expected due to Brownian motion, after which displacement stabilizes in both the vertical and horizontal directions. This stability occurs because ^{214}Po, no longer moves, as this decay happens almost instantly. We do not follow its nearest daughter (^{210}Pb) because it has a much longer half-life, which is less relevant to the public health risks associated with the inhalation of radon progeny. The vertical and horizontal displacement and the stability have important implications for radon exposure and public health, especially in confined environments. Understanding the extent and direction of radon movement aids in predicting where radon and its progeny may accumulate, which is crucial for assessing inhalation risks.

In agricultural sector and rural areas where buildings may be constructed over bare soil[74], the vertical displacement behavior illustrated here is particularly relevant. Progeny like polonium can settle in low-ventilation areas, leading to higher concentrations and, consequently, increased health risks for inhabitants. However, without accounting for the attachment of progeny to airborne particles or several surfaces, this concentration may reasonably decrease, as attachment can reduce the number of unattached radon progeny in the air[62].

Fig. 9 also illustrates the rule that nearly all atoms of a specific nuclide have decayed after approximately seven half-lives. This is evident in the plot, as around 1.93×10^8 seconds (i.e, for radon: $3.826 \times 24 \times 60 \times 60$ seconds), the displacement remains stable, indicating that the progeny no longer moves significantly due to rapid decay. As radon atoms can travel long distances along this path its progeny can be inhaled, contributing to health risks for people in the surrounding area if no external factors dilute this possibility.

Similarly, when the airflow is set to the maximum observed wind speed of 9.72 cm $_s^{-1}$, as measured by our wireless weather sensor during all measurements, the results are shown in Fig. 10. Although a constant wind speed was assumed for simplicity, this scenario represents the upper bound of radon and its progeny displacement under favorable conditions, illustrating how far they can travel. In contrast, turbulence or highly variable wind could reduce the total displacement[75].

Unlike previous distributions without airflow, Fig. 10 demonstrates a significantly wider horizontal spread of atoms due to wind-driven transport. In the ^{222}Rn distribution, undecayed radon clusters disperse over thousands of meters. While such dispersion may reduce local concentrations, it also increases the risk of radon infiltration into buildings over large areas, particularly in open regions like agricultural zones[57].

Although radon transport is primarily diffusion-limited, wind significantly enhances long-range displacement. This can lead to elevated indoor concentrations from outdoor sources, though this contribution is generally less significant than radon entering through foundation cracks or building materials. Without adequate ventilation, accumulated radon and its progeny may reach hazardous concentrations indoors. These findings underscore the critical importance of proper structural ventilation as a primary mitigation strategy, as further illustrated in Fig. 11.

Fig. 11 visually demonstrates the impact of airflow on the displacement of radon and its progeny over time. The left subplot captures the extensive horizontal (XY) displacement reaching up to 10^6 meters, while the right subplot zooms in on a shorter time scale to showcase the initial phases of displacement. The shaded regions indicate the range of movement for each nuclide, with the horizontal displacement (in blue) extending significantly more than the vertical displacement (in red) due to airflow. This pattern supports the understanding that wind direction and speed play crucial roles in determining radon's reach and concentration in open spaces. For public health, the insights from Fig. 11 underscore the necessity of considering environmental airflow when assessing radon exposure risks. In areas with strong winds, radon can disperse over long distances, potentially affecting remote locations[76,77].

Contribution of short-lived gamma-emitting radon progeny to the proximal gamma-ray spectroscopy technique
In this section, a bare crystal detector was used to estimate the total efficiency. The detector in the simulation was positioned at a height of 2.25 m above the ground, following similar approaches adopted in other studies[20–23]. This height makes the spectrometer more suitable for measuring atmospheric or dispersed radioactive contamination, such as ^{214}Bi, rather than ground-level radiation. In addition, the agricultural field selected as an example was La Agraria, due to its minimal interference from biomass at this stage, prior to sowing. This setting allows for more reliable gamma emission measurements, as the exhalation from the soil can be approximated to produce around 9.73×10^8 radon atoms per unit area over a 24-hour as results from section "Simulation of the transport and decay of radon and its progeny atoms in the air". Using these initial parameters, Table 8 presents the simulation results.

In Table 8, the second column shows the number of airborne atoms immediately before their decay and gamma emission. These gamma rays are potentially detectable by the spectrometer. The third column is derived by applying the simulation described in section "Simulation of the transport and decay of radon and its586 progeny atoms in the air", which calculates the geometrical efficiency. The simulation was run ten times to enhance the statistical results of the MC methods. The fourth column is estimated by using the total linear attenuation coefficient of 0.07818 for 609.312 keV from XCOM (https://physics.nist.gov/PhysRefData/Xcom/html/xcom1.html) and the density of the crystal, which is 3.67 g cm^{-3}. The fifth column extends the analysis to determine the total detection efficiency, integrating contributions from geometrical and intrinsic efficiencies for each nuclide. The value of 2.82×10^{-4} means that the vast majority of emitted gamma rays from ^{214}Bi, created from radon exhalation in 1 m^2 over one day, are not recorded by the detector, this implies that the contribution of ^{214}Bi from radon exhalation to the overall gamma signal is minimal. This significantly reduces the risk of overestimating the counts of ^{214}Bi in the soil. This finding is crucial in settings where there is no biomass in fields or large crops. Otherwise, the potential attachment of radon progeny to biomass (due to tall crops), can affect the movement (due to wind velocity) and distribution of ^{214}Bi. In addition, the corrected net counts from ^{214}B calculated as section "Contribution of short-lived gamma-emitting radon progeny to the proximal gamma-ray spectroscopy technique", is equal to 0.01692% of the measured net counts.

Conclusions
This study demonstrated the feasibility of using ANN to predict radon exhalation from agricultural soils based on edaphic and meteorological parameters. The model achieved strong performance (R^2 = 0.83 for validation), capturing complex, nonlinear relationships with minimal overfitting. Among the 56 variables analyzed across four fields, pressure, air temperature, solar irradiance, soil moisture, permeability, and bulk density emerged as key drivers of radon exhalation.
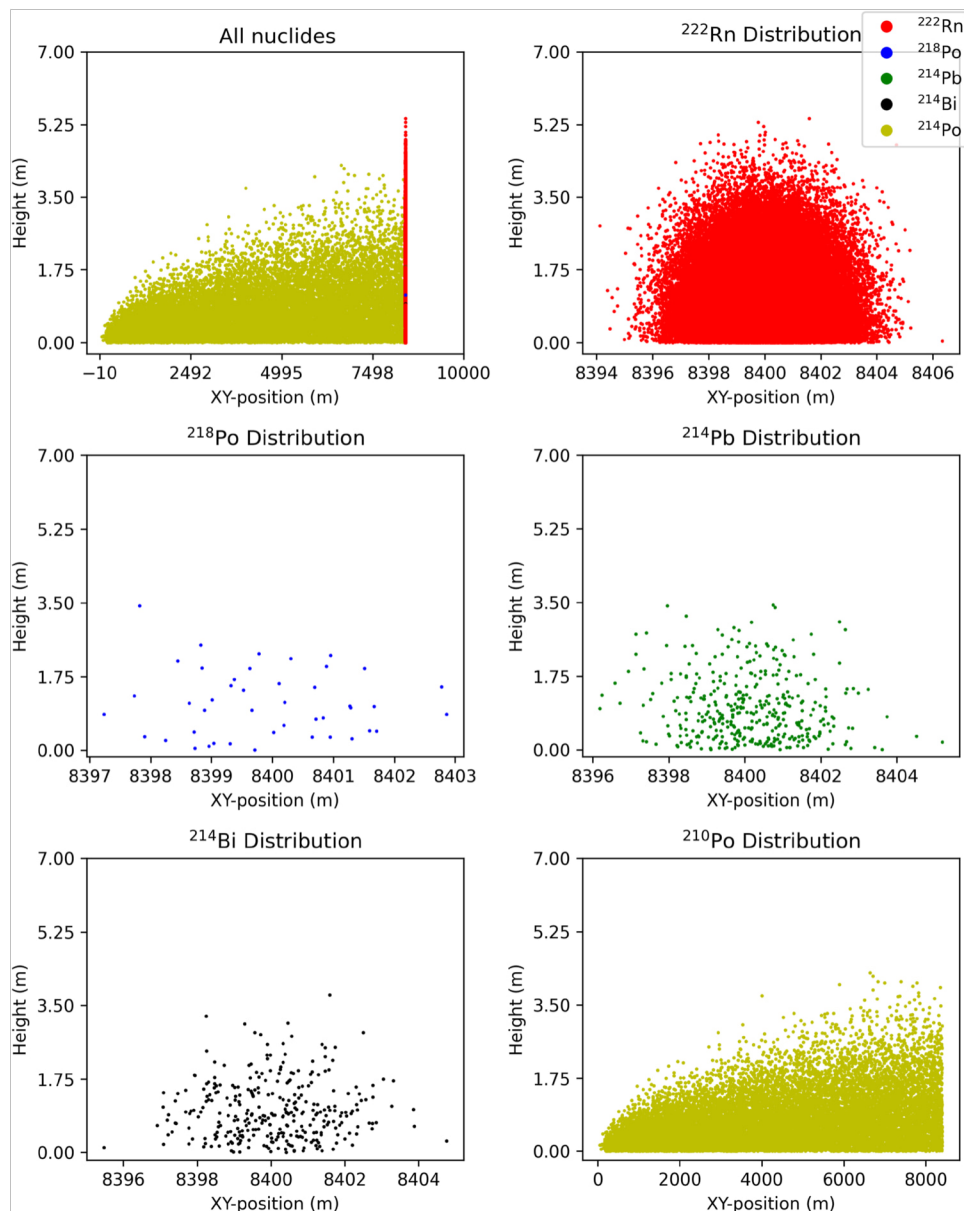
Fig. 10. Spatial distribution of radon atoms (^{222}Rn) and its progeny (^{218}Po, ^{214}Pb, ^{214}Bi, and ^{210}Po), caused by radon exhaled from the soil over a one-meter-squared area of soil during one day, considering air movement of 9.72 cm $_s^{-1}$. The top-left subplot is the combined distribution of all nuclides, while the remaining subplots are the individual subplot of each nuclide.

Complementing the ANN analysis, Monte Carlo simulations modeled the transport and decay of radon and its progeny in open environments, showing that radon atoms can migrate considerable distances under certain conditions. Although outdoor concentrations were well below health reference levels, these findings underscore the importance of spatial dispersion and support the need for radon monitoring in agricultural areas. Given the proximity of rural populations to cultivated soils in several study regions, continued monitoring is essential to ensure long-term public health protection and to guide future assessments or mitigation efforts where needed.

The study also addressed the potential influence of short-lived radon progeny on proximal gamma-ray spectroscopy for soil moisture measurement. It was found that most gamma emissions from ^{214}Bi over a 1 m^2 area were not detectable, underscoring the importance of accounting for radon-induced background effects in radiometric sensing.

Future work should focus on expanding the dataset to include more diverse field conditions and broader variability in exhalation rates, which would enhance model robustness and generalizability. Incorporating additional predictors such as radionuclide concentration, groundwater depth, and geological features may further improve prediction accuracy. In the simulation, the assumption of constant wind speed neglects near-surface turbulence and the exclusion of aerosol attachment and vegetation plate-out effects likely underestimated
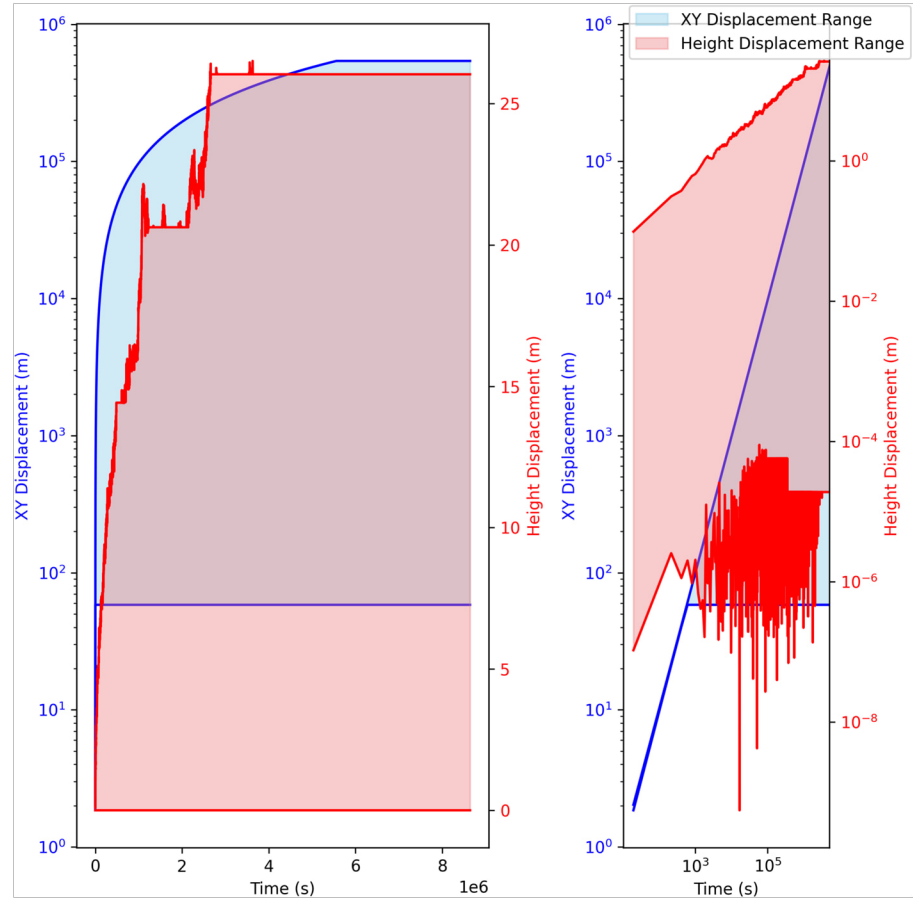
Fig. 11. Displacement of radon atoms and their progeny over time considering air movement of 9.72 cm s^{-1}, with shaded regions representing the range between the lower (bottom lines) and upper (top lines) displacement values for both horizontal (XY plane) and vertical (height) movement. The left subplot shows the entire time from 0 to 8×10^6 seconds with steps of 2×10^6, while the right subplot focuses on a shorter period, from 0 to 5e6 seconds.

Nuclide	Number of Atoms	η_g	η_i	η_{total}
^{214}Bi	557.7 ± 20.98	$3.64\times10^{-4}\pm1.37\times10^{-5}$	$0.77 \pm 5.68\times10^{-3}$	$2.82\times10^{-4}\pm1.15\times10^{-5}$

Table 8. Simulation results for the short-lived gamma-emitting radon progeny (^{214}Bi). Detection efficiencies (fractional values) were computed for a 3" × 3" NaI(Tl) gamma spectrometer.

inhalation risks. Future simulations should include these processes, as well as distinguish between attached and unattached progeny states, to improve the reliability of radiological risk assessments. Despite these simplifications, the presented models provide a strong foundation for future studies addressing environmental radon dynamics and their implications in agricultural settings.

Data availability
Data will be made available from the Corresponding Author upon reasonable request.

References
1. Sakoda, A. & Ishimori, Y. Mechanisms and modeling approaches of radon emanation for natural materials. *Hoken Butsuri (Online)* **52**(4), 296–306. https://doi.org/10.5453/jhps.52.296 (2017).
2. Sun, K., Guo, Q. & Cheng, J. The effect of some soil characteristics on soil radon concentration and radon exhalation from soil surface. *J. Nucl. Sci. Technol.* **41**(11), 1113–1117. https://doi.org/10.1080/18811248.2004.9726337 (2004).
3. Muhammad, A. & Külahcı, F. Radon transport from soil to air and Monte-Carlo simulation. *J. Atmos. Sol.-Terr. Phys.* **227**, 105803. https://doi.org/10.1016/j.jastp.2021.105803 (2022).

4. Charlton, P. & Kotrappa, P. Uranium prospecting for accurate time-efficient surveys of radon emissions in air and water, with a comparison to earlier radon and He surveys. In *Proceedings of the 2006 International Radon Symposium*, Vol. 1, 1–9 (2006).

5. Chen, Z. et al. Radon emission from soil gases in the active fault zones in the capital of China and its environmental effects. *Sci. Rep.* **8**(1), 16772. https://doi.org/10.1038/s41598-018-35262-1 (2018).

6. Čeliković, I. et al. Overview of radon flux characteristics, measurements, models and its potential use for the estimation of radon priority areas. *Atmos* **13**(12), 2022. https://doi.org/10.3390/atmos13122005 (2005).

7. Moorman, L. Radon flux measurements as predictor for indoor radon concentrations in new home residential structures. In *Proceedings of the 2005 International Radon Symposium*, 1–12 (2005).

8. Dörr, H., Kromer, B., Levin, I., Münnich, K. & Volpp, H.-J. CO_2 and radon 222 as tracers for atmospheric transport. *J. Geophys. Res. Oceans* **88**(C2), 1309–1313. https://doi.org/10.1029/JC088iC02p01309 (1983).

9. Yang, J. et al. Modeling of radon exhalation from soil influenced by environmental parameters. *Sci. Total Environ.* **656**, 1304–1311. https://doi.org/10.1016/j.scitotenv.2018.11.464 (2019).

10. Yang, J. et al. Radon exhalation from soil and its dependence from environmental parameters. *Radiat. Prot. Dosim.* **177**(1–2), 21–25. https://doi.org/10.1093/rpd/ncx165 (2017).

11. Ryzhakova, N. K. Parameters of modeling radon transfer through soil and methods of their determination. *J. Appl. Geophys.* **80**, 151–157. https://doi.org/10.1016/j.jappgeo.2012.01.010 (2012).

12. Xia, M., Ye, Y.-J. & Liu, S.-Y. Numerical simulations for radon migration and exhalation behavior during measuring radon exhalation rate with closed-loop method. *Nucl. Sci. Tech.* **35**(1), 9. https://doi.org/10.1007/s41365-024-01362-z (2024).

13. Silverman, M. P. Brownian motion of decaying particles: Transition probability, computer simulation, and first-passage times. *J. Mod. Phys.* **8**, 1809. https://doi.org/10.4236/jmp.2017.811108 (2017).

14. Astorri, F., Beaubien, S. E., Ciotoli, G. & Lombardi, S. An assessment of gas emanation hazard using a geographic information system and geostatistics. *Health Phys.* **82**(3), 358–366 (2002).

15. Buttafuoco, G., Tallarico, A. & Falcone, G. Mapping soil gas radon concentration: A comparative study of geostatistical methods. *Environ. Monit. Assess.* **131**, 135–151. https://doi.org/10.1007/s10661-006-9463-7 (2007).

16. Ciotoli, G. et al. Geographically weighted regression and geostatistical techniques to construct the geogenic radon potential map of the Lazio region: A methodological proposal for the European Atlas of Natural Radiation. *J. Environ. Radioact.* **166**, 355–375. https://doi.org/10.1016/j.jenvrad.2016.05.010 (2017).

17. Asim, K., Martínez-Álvarez, F., Basit, A. & Iqbal, T. Earthquake magnitude prediction in Hindukush region using machine learning techniques. *Nat. Hazards* **85**, 471–486. https://doi.org/10.1007/s11069-016-2579-3 (2017).

18. Elío, J., Petermann, E., Bossew, P. & Janik, M. Machine learning in environmental radon science. *Appl. Radiat. Isot.* **194**, 110684. https://doi.org/10.1016/j.apradiso.2023.110684 (2023).

19. Rezaie, F. et al. Spatial modeling of geogenic indoor radon distribution in Chungcheongnam-do, South Korea using enhanced machine learning algorithms. *Environ. Int.* **171**, 107724. https://doi.org/10.1016/j.envint.2022.107724 (2023).

20. Baldoncini, M. et al. Investigating the potentialities of Monte Carlo simulation for assessing soil water content via proximal gamma-ray spectroscopy. *J. Environ. Radioact.* **192**, 105–116. https://doi.org/10.1016/j.jenvrad.2018.06.001 (2018).

21. Baldoncini, M. et al. Biomass water content effect on soil moisture assessment via proximal gamma-ray spectroscopy. *Geoderma* **335**, 69–77. https://doi.org/10.1016/j.geoderma.2018.08.012 (2019).

22. Filippucci, P. et al. Soil moisture as a potential variable for tracking and quantifying irrigation: A case study with proximal gamma-ray spectroscopy data. *Adv. Water Resour.* **136**, 103502. https://doi.org/10.1016/j.advwatres.2019.103502 (2020).

23. Strati, V. et al. Modelling soil water content in a tomato field: Proximal gamma ray spectroscopy and soil-crop system models. *Agriculture* **8**(4), 60. https://doi.org/10.3390/agriculture8040060 (2018).

24. Censos Nacionales, I. N. E. I. XII de Población, VII de Vivienda y III de Comunidades Indígenas. *Peruvian National Census.* https://doi.org/10.1016/S0048-9697(98)00256-3 (2017).

25. Montoya Ramírez, M. A., García Márquez, W. & Caldas Vidal, J. Geología de los cuadrángulos de Lomitas, Palpa, Nasca y Puquio 30-l, 30-m, 30-n, 30-ñ–[Boletín A 53]. Technical report, INGEMMET. https://repositorio.ingemmet.gob.pe/handle/20.500.12544/174 (1994).

26. Sánchez Fernández, A. W., Molina Galdos, O. & Gutiérrez Abanto, R. Geología de los cuadrángulos de chimbote, casma y culebras. hojas: 19-f, 19-gy 20-g–[boletín a 59]. Technical report, INGEMMET (1995).

27. Palacios Moncayo, O., Caldas Vidal, J. & Vela Velásquez, C. Geología de los cuadrángulos de lima, lurín, chancay y chosica. hojas 25-i, 25-j. 24-i, 24-j-[boletín a 43]. Technical report, INGEMMET (1992).

28. Oliver, M. & Khayrat, A. A geostatistical investigation of the spatial variation of radon in soil. *Comput. Geosci.* **27**(8), 939–957. https://doi.org/10.1016/S0098-3004(00)00133-3 (2001).

29. Bunzl, K., Ruckerbauer, F. & Winkler, R. Temporal and small-scale spatial variability of 222rn gas in a soil with a high gravel content. *Sci. Total Environ.* **220**(2–3), 157–166. https://doi.org/10.1016/S0048-9697(98)00256-3 (1998).

30. Elío, J., Crowley, Q., Scanlon, R., Hodgson, J. & Long, S. Rapid radon potential classification using soil-gas radon measurements in the Cooley Peninsula, County Louth, Ireland. *Environ. Earth Sci.* **78**(12), 359. https://doi.org/10.1007/s12665-019-8339-4 (2019).

31. Alonso, H. et al. Assessment of radon risk areas in the Eastern Canary Islands using soil radon gas concentration and gas permeability of soils. *Sci. Total Environ.* **664**, 449–460. https://doi.org/10.1016/j.scitotenv.2019.01.411 (2019).

32. Okalebo, J. R., Gathua, K. W. & Woomer, P. L. Laboratory methods of soil and plant analysis: A working manual second edition. *Sacred Africa, Nairobi* **21**, 25–26 (2002).

33. Vílchez, M., Ochoa, M. & Pari, W. Peligro Geológico en la Región Ica. *Boletín Serie C: Geodinámica e Ingeniería Geológica* **67**, INGEMMET (2019).

34. Macharé, J., Benavente, C. & Audin, L. Síntesis descriptiva del mapa neotectónico 2008. *INGEMMET Boletín Serie C: Geodinámica e Ingeniería Geológica*, Vol. 40. https://hdl.handle.net/20.500.12544/245 (2009).

35. Benà, E. et al. Evaluation of tectonically enhanced radon in fault zones by quantification of the radon activity index. *Sci. Rep.* **12**(1), 21586. https://doi.org/10.1038/s41598-022-26124-y (2022).

36. Kotrappa, P. *Electret ion Chambers for Characterizing Indoor, Outdoor, Geologic and Other Sources of Radon* 2–36 (Nova Science Publishers, Inc, Radon, 2015).

37. Kotrappa, P., Stieff, F. & Steck, D. J. Radon flux monitor for in situ measurement of granite and concrete surfaces. In *Proceedings of the American Association of Radon Scientists and Technologists 2009 International Symposium, St. Louis, USA* (2009).

38. Masod Abdulqader, S., Vakanjac, B., Kovačević, J., Naunovic, Z. & Zdjelarević, N. Natural radioactivity of intrusive-metamorphic and sedimentary rocks of the Balkan Mountain range (Serbia, Stara Planina). *Minerals* **8**(1), 6. https://doi.org/10.3390/min8010006 (2017).

39. Bassey, N. E., Kaigama, U. & Oluwasegun, A. Radiometric mapping of Song area and environs, Hawal basement complex, northeast Nigeria. *Int. J. Sci. Tech.* **2**(9), 692–9 (2013).

40. Tian, H. et al. Mapping winter crops in china with multi-source satellite imagery and phenology-based algorithm. *Remote Sens.* **11**(7), 820. https://doi.org/10.1016/j.scitotenv.2019.01.411 (2019).

41. Corti, T., Wüest, M., Bresch, D. & Seneviratne, S. I. Drought-induced building damages from simulations at regional scale. *Nat. Hazards Earth Syst. Sci.* **11**(12), 3335–3342. https://doi.org/10.5194/nhess-11-3335-2011 (2011).

42. Neznal, M., Neznal, M., Matolin, M., Barnet, I. & Miksova, J. *The New Method for Assessing the Radon Risk of Building Sites* 48 (Czech Geological Survey, Prague, 2024).

43. Beltrán-Torres, S. et al. Estimated versus field measured soil gas radon concentration and soil gas permeability. *J. Environ. Radioact.* **265**, 107224. https://doi.org/10.1016/j.jenvrad.2023.107224 (2023).
44. Asare, E. O., Otoo, F., Adukpo, O. K. & Opoku-Ntim, I. Assessment of soil moisture on radon levels, radon exhalation, natural radioactivity, and radiological risks in offices and laboratories in GAEC. *J. Radiat. Res. Appl. Sci.* **17**(3), 101014. https://doi.org/10.1016/j.jrras.2024.101014 (2024).
45. Zhang, W., Zhang, Y. & Sun, Q. Analyses of influencing factors for radon emanation and exhalation in soil. *Water Air Soil Pollut.* **230**, 1–13. https://doi.org/10.1007/s11270-018-4063-z (2019).
46. Othman, N., Rashid, A. S. A., Hashim, S., Sanusi, M. S. M. & Bery, A. Radon gas migration through soil-a review. *J. Mines Met. Fuels* **69**(8), 3–8 (2021).
47. Petermann, E., Meyer, H., Nussbaum, M. & Bossew, P. Mapping the geogenic radon potential for Germany by machine learning. *Sci. Total Environ.* **754**, 142291. https://doi.org/10.1016/j.scitotenv.2020.142291 (2021).
48. Arabameri, A. et al. Performance evaluation of GIS-based novel ensemble approaches for land subsidence susceptibility mapping. *Front. Earth Sci.* **9**, 663678. https://doi.org/10.3389/feart.2021.663678 (2021).
49. Hoaglin, D. C. & John, W. Tukey and data analysis. *Stat. Sci.* **18**, 311–318 (2003).
50. Rezaie, F. et al. Application of machine learning algorithms for geogenic radon potential mapping in Danyang-Gun, South Korea. *Front. Environ. Sci.* **9**, 753028. https://doi.org/10.3389/fenvs.2021.753028 (2021).
51. Win, T. Z. & Kham, N. S. M. Information gain measured feature selection to reduce high dimensional data. PhD thesis, MERAL Portal (2019).
52. Ross, B. C. Mutual information between discrete and continuous data sets. *PLoS ONE* **9**(2), 87357. https://doi.org/10.1371/journal.pone.0087357 (2014).
53. Pon, M. Z. A. & Ka, K. K. Hyperparameter tuning of deep learning models in Keras. *Sparklinglight Trans. Artif. Intell. Quantum Comput. (STAIQC)* **1**(1), 36–40 (2021).
54. Nguyen, X. D. J. & Liu, Y. Methodology for hyperparameter tuning of deep neural networks for efficient and accurate molecular property prediction. *Comput. Chem. Eng.* **193**, 108928. https://doi.org/10.1016/j.compchemeng.2024.108928 (2025).
55. O'Malley, T. et al. Keras tuner (2019).
56. Wang, J., Xu, J. & Wang, X. Combination of hyperband and Bayesian optimization for hyperparameter optimization in deep learning. Preprint at arXiv:1801.01596v1 (2018).
57. Chen, X. et al. Responses of the atmospheric concentration of radon-222 to the vertical mixing and spatial transportation. *Boreal Environ. Res.* **21**(3–4), 299–318 (2016).
58. Nikezić, D. & Stevanović, N. Radon progeny behavior in diffusion chamber. *Nucl. Instrum. Methods Phys. Res. B* **239**(4), 399–406. https://doi.org/10.1016/j.nimb.2005.05.050 (2005).
59. Pérez, B., López, M. & Palacios, D. Theoretical and experimental study of the LR-115 detector response in a non-commercial radon monitor. *Appl. Radiat. Isot.* **160**, 109112. https://doi.org/10.1016/j.apradiso.2020.109112 (2020).
60. Marcq, P. & Naert, A. A Langevin equation for turbulent velocity increments. *Phys. Fluids* **13**(9), 2590–2595. https://doi.org/10.1063/1.1386937 (2001).
61. Mostafa, M. Y. A., Khalaf, H. N. B. & Zhukovsky, M. Radon decay products equilibrium at different aerosol concentrations. *Appl. Radiat. Isot.* **156**, 108981. https://doi.org/10.1016/j.apradiso.2019.108981 (2020).
62. Vaupotič, J. Radon and its short-lived products in indoor air: Present status and perspectives. *Sustainability* **16**(6), 2424. https://doi.org/10.3390/su16062424 (2024).
63. Knoll, G. F. *Radiation Detection and Measurement* (John & Wiley Sons Inc, Hoboken, 2010).
64. Abbas, M. I. Analytical calculations of the solid angles subtended by a well-type detector at point and extended circular sources. *Appl. Radiat. Isot.* **64**(9), 1048–1056. https://doi.org/10.1016/j.apradiso.2006.04.010 (2006).
65. Aguiar-Amado, P. P., Amado, V. A. & Aguiar, J. C. Full-energy peak determination from total efficiency and peak-to-total ratio calculations. *Nucl. Instrum. Methods Phys. Res. A.* **990**, 164980. https://doi.org/10.1016/j.nima.2020.164980 (2021).
66. Guo, Q., Sun, K. & Cheng, J. Methodology study on evaluation of radon flux from soil in China. *Radiat. Prot. Dosim.* **112**(2), 291–296. https://doi.org/10.1093/rpd/nch387 (2004).
67. Hirao, S., Yamazawa, H. & Moriizumi, J. Estimation of the global {sup 222} Rn flux density from the earth's surface. *Hoken Butsuri* https://doi.org/10.5453/jhps.45.1 (2010).
68. Hosoda, M. et al. Simultaneous measurements of radon and thoron exhalation rates and comparison with values calculated by UNSCEAR equation. *J. Radiat. Res.* **50**(4), 333–343. https://doi.org/10.1269/jrr.08121 (2009).
69. UNSCEAR. Sources and Effects of Ionizing Radiation. Report to the General Assembly. Volume I: Sources. United Nations Scientific Committee on the Effects of Atomic Radiation, New York (2000).
70. Duong, V.-H., Ly, H.-B., Trinh, D. H., Nguyen, T. S. & Pham, B. T. Development of Artificial Neural Network for prediction of radon dispersion released from Sinquyen Mine, Vietnam. *Environ. Pollut.* **282**, 116973. https://doi.org/10.1016/j.envpol.2021.116973 (2021).
71. Feng, X. et al. Groundwater radon precursor anomalies identification by EMD-LSTM model. *Water* **14**(1), 69. https://doi.org/10.3390/w14010069 (2022).
72. Fernández, D. F. P., Anaya, P. E. P. & Sajo-Bohus, L. Simulation of diffusion and decay of radon/thoron exhaled from a wall and its newly created progeny. response of a bare lr-115 detector placed on the wall.. *Appl. Radiat. Isot.* **220**, 111743. https://doi.org/10.1016/j.apradiso.2025.111743 (2025).
73. DIDP. Decreto Supremo No 009-97-EM, Reglamento De Seguridad Radiológica https://www2.congreso.gob.pe/Sicr/CenDocBib/con5_uibd.nsf/$$ViewTemplate%20for%20Documentos?OpenForm&Db=195C84C15DF1BE1C0525845F0076565B&View=yyy (2019).
74. Molina, J., Ponce, M., Horn, M. & Gómez, M. Towards a sustainable bioclimatic approach for the Peruvian high Andean rural area: Evaluation of the thermal contribution of a greenhouse attached to a dwelling. In *Proc. ISES Sol. World Congr. 2019 IEA SHC Int. Conf. Sol. Heat. Cool. Build. Ind.* 355–364 (2019).
75. Droppo, J. G. Improved formulations for air-surface exchanges related to national security needs: Dry deposition models. Technical report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States) (2006).
76. Gutiérrez-Álvarez, I. et al. Radon behavior investigation based on cluster analysis and atmospheric modelling. *Atmos. Environ.* **201**, 50–61. https://doi.org/10.1016/j.atmosenv.2018.12.010 (2019).
77. Lebel, L., John, A. & Korolevych, V. Dispersion simulations of radon discharges between neighboring buildings and their sensitivity to meteorology, discharge rate, and building geometry. *Health Phys.* **122**(3), 383–401. https://doi.org/10.1097/HP.0000000000001510 (2022).

Acknowledgements

Miguel Sanchez Delgado from the Faculty of Agriculture, for providing us the opportunity to utilize the facilities of the Centro de Investigación y Extensión de Riego (CIER), and the managers of the agricultural fields in Nazca, Yaután, and Santa Eulalia.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to B.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.