

Article

Application of Statistical Methods for the Characterization of Radon Distribution in Indoor Environments: A Case Study in Lima, Peru

Rafael Liza ¹, Félix Díaz ^{2,*}, Patrizia Pereyra ³, Daniel Palacios ³, Nhell Cerna ⁴, Luis Curo ⁵
and Max Riva ⁵

¹ Departamento Académico de Cursos Básicos, Universidad Científica del Sur, Lima 15067, Peru

² Vicerrectorado de Investigación, Universidad Autónoma del Peru, Lima 15067, Peru

³ Departamento de Ciencias, Sección Física, Pontificia Universidad Católica del Peru, Lima 15088, Peru

⁴ Facultad de Ingeniería, Universidad Tecnológica del Perú, Lima 15046, Peru

⁵ Departamento de Física, Universidad Nacional Pedro Ruiz Gallo, Lambayeque 14001, Peru

* Correspondence: felix.diaz@autonoma.pe

Abstract: This study evaluates the effectiveness of advanced statistical and geospatial methods for analyzing radon concentration distributions in indoor environments, using the district of San Martín de Porres, Lima, Peru, as a case study. Radon levels were monitored using LR-115 nuclear track detectors over three distinct measurement periods between 2015 and 2016, with 86 households participating. Detectors were randomly placed in various rooms within each household. Normality tests (Shapiro–Wilk, Anderson–Darling, and Kolmogorov–Smirnov) were applied to assess the fit of radon concentrations to a log-normal distribution. Additionally, analysis of variance (ANOVA) was used to evaluate the influence of environmental and structural factors on radon variability. Non-normally distributed data were normalized using a Box–Cox transformation to improve statistical assumptions, enabling subsequent geostatistical analyses. Geospatial interpolation methods, specifically Inverse Distance Weighting (IDW) and Kriging, were employed to map radon concentrations. The results revealed significant temporal variability in radon concentrations, with geometric means of $146.4 \text{ Bq}\cdot\text{m}^{-3}$, $162.3 \text{ Bq}\cdot\text{m}^{-3}$, and $150.8 \text{ Bq}\cdot\text{m}^{-3}$, respectively, across the three periods. Up to 9.5% of the monitored households recorded radon levels exceeding the safety threshold of $200 \text{ Bq}\cdot\text{m}^{-3}$. Among the interpolation methods, Kriging provided a more accurate spatial representation of radon concentration variability compared to IDW, allowing for the precise identification of high-risk areas. This study provides a framework for using advanced statistical and geospatial techniques in environmental risk assessment.



Academic Editor: Abdelatif Amrane

Received: 16 October 2024

Revised: 11 December 2024

Accepted: 18 December 2024

Published: 14 January 2025

Citation: Liza, R.; Díaz, F.; Pereyra, P.; Palacios, D.; Cerna, N.; Curo, L.; Riva, M. Application of Statistical Methods for the Characterization of Radon Distribution in Indoor Environments: A Case Study in Lima, Peru. *Eng* **2025**, *6*, 14. <https://doi.org/10.3390/eng6010014>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: radon; statistical methods; environmental monitoring; geostatistical mapping; public health

1. Introduction

In the past two decades, increasing attention has been directed toward investigating and measuring levels of radon gas (^{222}Rn) due to the recognized health risks associated with its accumulation in indoor environments. Radon is a naturally occurring radioactive gas, and extensive research has established a well-founded link between radon exposure and an elevated risk of certain cancers, particularly lung and gastric cancers [1]. Since 1986, the World Health Organization (WHO) has classified radon as a Group 1 carcinogen [2], highlighting the urgency of addressing this issue from a public health perspective. This

consensus is upheld by international expert bodies such as the United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) [3], the International Commission on Radiological Protection (ICRP) [4], the International Atomic Energy Agency (IAEA) [5], and the United States Environmental Protection Agency (EPA) [6], all affirming that any level of radon exposure carries a potential health risk. Consequently, these organizations have initiated actions to mitigate the risks associated with ^{222}Rn exposure, emphasizing the importance of maintaining safe indoor environments for both living and working [7].

Therefore, understanding the factors influencing indoor radon levels is important for effective risk assessment and management. The ^{222}Rn concentration in enclosed spaces is determined by a complex interaction of variables, including building characteristics, soil properties, geological formations, and environmental conditions such as temperature and humidity [8–11]. Analyzing these factors and their interrelationships makes it possible to predict indoor radon levels or estimate the likelihood of exceeding specific radon thresholds. Furthermore, while these factors operate independently, their combined effects ultimately shape the final radon concentration in any given space. This complexity suggests that indoor radon levels are better modeled using a log-normal distribution, given the skewness observed in radon data due to the random nature of emission sources and indoor accumulation patterns [12].

Numerous studies have applied various statistical methods to analyze these interactions and model radon distribution, with the log-normal distribution being one of the most widely used approaches. To confirm the suitability of this model, normality tests such as the Shapiro–Wilk and Anderson–Darling tests have been employed, verifying whether the data adequately fit a log-normal distribution [13]. Moreover, previous studies have employed parametric estimation techniques to estimate the proportion of homes surpassing a specific radon concentration threshold [14–16]. Additionally, techniques such as analysis of variance (ANOVA) have been used to determine the influence of specific factors, such as building type or ventilation, on indoor radon concentration [17–19]. ANOVA has proven particularly useful in comparing different groups, as it helps identify the factors that significantly impact the observed variability in radon levels [20,21]. Furthermore, in geospatial studies, methods such as Inverse Distance Weighting (IDW) and Kriging have been widely applied to interpolate radon concentrations in regions where direct measurement data are unavailable [22–24]. While IDW is valued for its simplicity and efficiency, Kriging is often preferred when more precise estimates are required, as it models the spatial variability of radon levels by considering both the distance and spatial correlation between measurement points [25,26]. Therefore, these methods are essential for developing risk maps that guide targeted interventions in the most affected areas.

Recent studies conducted in various countries have utilized these statistical methods to integrate multiple parameters, including housing characteristics, soil composition, geological data, and gamma dose rates [27–31]. By combining these factors, researchers have established comprehensive frameworks that enhance understanding of radon distribution patterns and identify high-risk areas. This approach has proven effective in capturing the inherent variability of radon concentrations and developing robust predictive models.

Thus, this research aims to explore and identify the most suitable statistical techniques to model indoor radon distribution in Lima, Peru. Specifically, radon concentration measurements obtained using trace detectors are analyzed alongside factors specific to the homes and lifestyles of the occupants. The log-normal distribution is applied to characterize radon concentrations, and normality tests are performed to validate this choice. Additionally, the proportion of households exceeding a specific radon concentration threshold is estimated using parametric and non-parametric estimation techniques, thereby providing a

detailed analysis of exposure risks in densely populated urban areas. Furthermore, analysis of variance (ANOVA) is employed to determine the significant influence of factors such as building type, ventilation, and geographic location on the variability of radon levels. Finally, spatial interpolation methods, including Inverse Distance Weighting (IDW) and Kriging, generate risk maps, identifying areas with potentially elevated radon concentrations.

2. Materials and Methods

2.1. Data

In the present study, conducted in San Martín de Porres, Lima, Peru, starting in mid-October 2015, three consecutive 45-day measurement periods were used to assess indoor radon activity concentrations utilizing the non-strippable kodak LR-115 Type 2 detector. This detector was attached to a plastic film slide and positioned on walls approximately 1.5 m above the floor in bare mode (1.5 cm × 1.5 cm). The first measurement period began in mid-October 2015, capturing radon levels during spring. The second period immediately started in early December, extending into late spring. The third and final period began in late January 2016, covering early summer, allowing the study to capture seasonal variations in radon levels specific to the Peruvian climate. Each detector was exposed for 45 days at the study site, covering these different measurement periods. This sequential approach allowed for observing variations in indoor radon concentrations due to environmental and seasonal changes from spring to summer. The voluntary participation of numerous households provided a broad view of radon levels across different times of the year. Radon detectors were randomly placed in any available spaces within each household, such as bedrooms, living rooms, kitchens, or other rooms. This random assignment of detectors ensured that the measurements were not influenced by subjective factors or the deliberate choice of specific spaces, thus providing a more balanced evaluation of indoor radon concentrations. After each 45-day exposure period, the LR-115 detectors in bare mode were chemically etched to reveal alpha particle tracks. The etching process involved immersing the detectors in a 2.5 N sodium hydroxide (NaOH) solution at 60 °C for 90 min, utilizing a thermostatic bath with a temperature control fluctuation of ±0.5 °C. Upon completion of the etching, the detectors were thoroughly rinsed with distilled water to stop the process and remove any residual NaOH. The background track density for unexposed detectors was measured at 1.5 tracks cm⁻², subtracted from the total track density to account for background noise. The minimum measurable radon activity concentration for this setup was calculated at 20 Bq/m³ for a 45-day exposure period. The uncertainty budget for the radon measurements was estimated at ±35%, considering calibration factor uncertainty and track counting variability. Disregarding the contribution of the Radon-220 progeny and considering the partial sensitivity of the detector and the equilibrium factor within the proposed range, the calibration factor was estimated at 0.038 ± 0.005 tracks cm⁻²·Bq⁻¹·d⁻¹·m³ (see [32] for more detail). For track counting, each etched detector was photographed at 50 distinct regions using a Leica microscope, capturing high-resolution images to ensure an accurate representation of track distribution across the surface. Subsequently, these images were analyzed in ImageJ version 1.51a [33], an open-source image analysis software, where the tracks were identified based on circularity and size criteria and track density was calculated (ρ). Following a precise calibration, this automated identification enabled the accurate counting and assessment of radon concentrations based on track density across the detector. Radon concentration was calculated using the following equation:

$$C_{Rn} = \frac{\rho}{K \cdot t} \quad (1)$$

where ρ is track density, K is the calibration factor, and t is the duration of exposure (days).

2.2. Statistical Analysis Methods

2.2.1. The Log-Normal Distribution and Considerations for Normality Tests

The log-normal distribution is commonly employed to model environmental variables that are inherently non-negative, such as radon concentrations. The probability density function for a random variable x , which follows a log-normal distribution $Ln(\mu, \sigma^2)$, is given by [12]

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x > 0 \quad (2)$$

In this formula, μ and σ denote the mean and standard deviation of the natural logarithm of x , respectively. The geometric mean (GM) and geometric standard deviation (GSD) can be calculated as $GM = \exp(\mu)$ and $GSD = \exp(\sigma)$.

Before conducting statistical analyses assuming a log-normal distribution, verifying that the log-transformed data fit a normal distribution is essential. This is because many statistical tests and models require normality of the data to ensure the validity of the results. To confirm this assumption, key normality tests are applied, including the following.

2.2.2. Anderson–Darling Test

The Anderson–Darling test is given by the following statistic:

$$A^2 = -n - \sum_{i=1}^n \frac{(2i-1)}{n} [\ln(F(y_i)) + \ln(1 - F(y_{n+1-i}))] \quad (3)$$

where n is the sample size, $F(y_i)$ is the cumulative distribution function (CDF) of the normal distribution evaluated at the i -th ordered data point y_i . And y_{n+1-i} represents the data points in descending order. The Anderson–Darling test is a statistical method used to determine how well a set of data fits a particular distribution—in this case, the normal distribution. It is especially sensitive to discrepancies in the tails of the distribution, making it well suited for datasets that include extreme values [34]. The null hypothesis is rejected if the adjusted Anderson–Darling statistic (Adj^2) exceeds the critical value corresponding to a 5% significance level (approximately 0.751). The latter indicates that the data do not follow a normal distribution, with a 5% probability of committing a Type I error.

2.2.3. Kolmogorov–Smirnov (K-S) Test

The Kolmogorov–Smirnov test is a non-parametric method used to evaluate the goodness-of-fit between an empirical sample distribution and a reference probability distribution—in this case, the normal distribution of log-transformed radon concentrations. The Kolmogorov–Smirnov test statistic (D) is defined as

$$D = \max |F_n(x) - F(x)| \quad (4)$$

where $F_n(x)$ is the empirical cumulative distribution function (ECDF) derived from the sample data and $F(x)$ is the cumulative distribution function (CDF) of the theoretical normal distribution. The calculated D value is compared against a critical value based on the chosen significance level (commonly 5%) and the sample size to interpret the test results. If D exceeds this critical value, the null hypothesis—that the log-transformed radon data follow a normal distribution—is rejected. This indicates a significant deviation from normality, suggesting that alternative statistical methods or data transformations, such as the Box–Cox transformation, may be necessary.

It is important to note that the K-S test is most sensitive around the median of the distribution and less so in the tails. Therefore, when analyzing radon data—which may

have extreme values due to localized geological factors—it is beneficial to use the K-S test in conjunction with the Anderson–Darling test, which is more sensitive to deviations in the tails.

2.2.4. Shapiro–Wilk Test

The Shapiro–Wilk test is a robust statistical method for evaluating the normality of a dataset and is particularly effective with small sample sizes (typically less than 50 observations). This makes it highly relevant in radon studies, where limited data points are common due to sampling constraints. The test statistic W is calculated as

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)}\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (5)$$

The null hypothesis is rejected if the Shapiro–Wilk test statistic (W) falls below the critical value corresponding to a 5% significance level. This test evaluates whether a sample originates from a normally distributed population and is particularly effective for small sample sizes, typically defined as fewer than 50 observations. The critical value of W , used to determine the rejection of the null hypothesis that the data follow a log-normal distribution at the 5% significance level, is contingent upon both the specific distribution and the sample size.

2.2.5. Estimating Homes Exceeding a Specific Radon Concentration Threshold

After confirming that the radon concentration data follow a log-normal distribution, one can estimate the percentage of homes exceeding a specific radon concentration threshold, denoted as NR . This estimation utilizes the geometric mean (GM) and the geometric standard deviation (GSD) derived from the dataset.

The standardized variable Z is calculated using the following equation:

$$Z = \frac{\ln(NR) - \ln(GM)}{\ln(GSD)} \quad (6)$$

Recognizing that $\ln(GM) = \mu$ and $\ln(GSD) = \sigma$, where μ and σ are the mean and the standard deviation of the log-transformed data, the equation simplifies to

$$Z = \frac{\ln(NR) - \mu}{\sigma} \quad (7)$$

This Z value represents the number of standard deviations that the logarithm of the threshold NR is away from the mean μ . To determine the probability P that a home exceeds the threshold NR , the cumulative distribution function (CDF) of the standard normal distribution is used:

$$P(Z \geq X) = 1 - P(Z \leq X) = 1 - \Phi(X) \quad (8)$$

where $\Phi(X)$ denotes the CDF of the standard normal distribution $N(0, 1)$. By transforming the log-normal distribution into a standard normal distribution, the computation of probabilities is simplified. This allows for the use of standard statistical tables or software to assess the risk and estimate the proportion of homes exceeding the specified radon concentration threshold.

2.3. Analysis of Variance (ANOVA)

Analysis of variance (ANOVA) is used to determine whether significant differences exist in mean radon concentrations across various groups. Before applying ANOVA, it

is essential to verify that the data meet the assumptions of normality and homogeneity of variances, which are critical for the validity of the test results. These assumptions are assessed through preliminary statistical tests on the indoor radon concentration data.

If these assumptions are satisfied, a one-way ANOVA is conducted, followed by post hoc multiple comparisons using the Least Significant Difference (LSD) method to identify specific group differences. The F -statistic in ANOVA is calculated using the following formula:

$$F = \frac{MSA}{MSE} = \frac{\frac{SSA}{df_A}}{\frac{SSE}{df_E}} = \frac{\frac{\sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2}{k-1}}{\frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2}{n-k}} \quad (9)$$

In this equation, k represents the number of groups, n_i is the number of observations in the i -th group, \bar{x}_i is the mean radon concentration of the i -th group, \bar{x} is the overall mean radon concentration, x_{ij} is the j -th observation in the i -th group, and n is the total number of observations across all groups. SSA (Sum of Squares Among groups) measures the variability between group means, while SSE (Sum of Squares Within groups) measures the variability within each group. The degrees of freedom associated with SSA and SSE are $df_A = k - 1$ and $df_E = n - k$, respectively. The mean square among groups is calculated as $MSA = SSA/df_A$, and the mean square within groups is $MSE = SSE/df_E$.

In situations where the assumption of homogeneity of variances is violated—which can occur due to the natural variability in radon measurements—the non-parametric Kruskal–Wallis test serves as a suitable alternative. This test does not assume normality and is effective for comparing median values across groups. The Kruskal–Wallis H -statistic is computed as

$$H = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1) \quad (10)$$

where R_i is the sum of the ranks assigned to the observations within the i -th group, and n is the total number of observations.

By applying these statistical techniques, it is possible to rigorously evaluate differences in radon concentrations among various categories such as geographic regions, building types, or building materials.

2.4. Inverse Distance Weighting (IDW) and Kriging

Inverse Distance Weighting (IDW) and Kriging are widely used geospatial techniques for generating detailed radon distribution maps from point data, aiding in the visualization of environmental risk. The choice between IDW and Ordinary Kriging (OK) significantly influences the accuracy and effectiveness of these maps. IDW is often favored for its simplicity, relying on the proximity of measured points to estimate radon concentrations, making it effective when data are sparse or unevenly distributed. However, this straightforward approach can be less precise compared to Kriging. Ordinary Kriging offers a significant advantage by considering both spatial correlation and variability between measurements, resulting in statistically optimized and more accurate estimates, particularly suitable for complex datasets.

Additionally, ArcGISPro 2022 software is commonly employed for generating these maps, using a grid typically sized at 1 km × 1 km. Each of these interpolation methods brings different strengths to radon mapping.

3. Results and Discussion

In Period 1, comprehensive normality tests conducted on the log-transformed radon concentration data, based on 34 measurements, provided clear evidence supporting the hypothesis that the data follow a normal distribution. The detailed statistical results are presented in Table 1.

Table 1. Results of normality tests in Period 1.

Test	Statistic	<i>p</i> -Value	Interpretation
Shapiro–Wilk	0.9646	0.3307	Does not reject normality
Anderson–Darling	0.3201	0.718 (5% CV)	Does not reject normality
Kolmogorov–Smirnov	0.0878	0.7283	Does not reject normality

Given that all tests produced *p*-values significantly above the common alpha level of 0.05, there is strong statistical evidence that the log-transformed data from Period 1 adheres closely to a normal distribution. Figure 1 shows histograms and Q-Q plots for the data of Period 1.

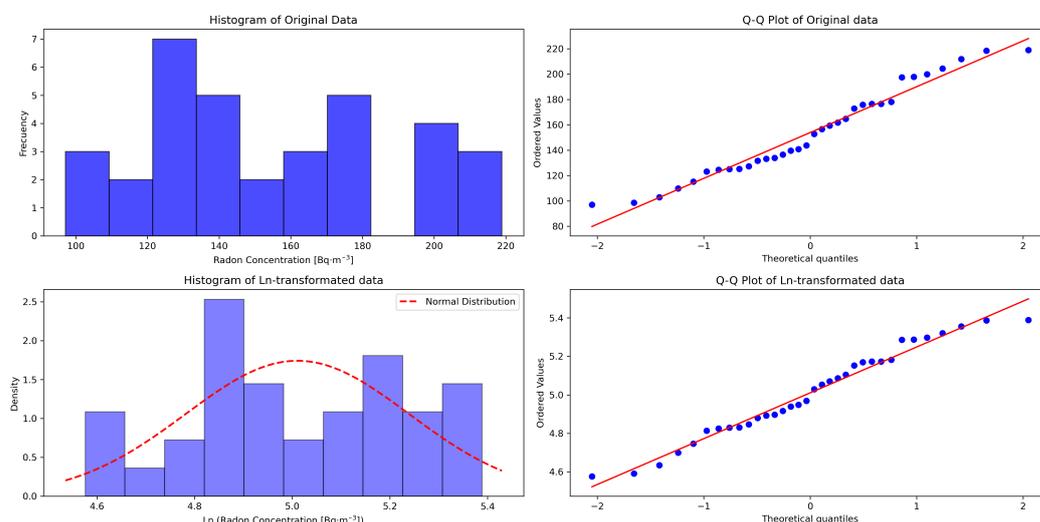


Figure 1. Graph histograms, the normal distribution fit to the log-transformed data, and the Q-Q plots for measurement in Period 1. **Top left:** Histogram of original data with log-normal fit. **Top right:** Q-Q plot of original data. **Bottom left:** Histogram of log-transformed data with normal fit. **Bottom right:** Q-Q plot of log-transformed data.

Similarly, the analysis of log-transformed data from Period 2 also supports normality, as evidenced by the *p*-values reported in Table 2.

Table 2. Results of normality tests in Period 2.

Test	Statistic	<i>p</i> -Value	Interpretation
Shapiro–Wilk	0.9790	0.6264	Does not reject normality
Anderson–Darling	0.4231	0.726 (5% CV)	Does not reject normality
Kolmogorov–Smirnov	0.1061	0.2740	Does not reject normality

The Shapiro–Wilk test confirms the normality of the data, supporting the assumption of a normal distribution. Similarly, the Anderson–Darling and Kolmogorov–Smirnov tests fail to reject the null hypothesis of normality, providing consistent evidence for this assumption. It is essential to consider that the sensitivity of normality tests varies with sample size, as in Period 1 and Period 2 with 34 and 42 observations, respectively. In contrast, even

minor deviations can lead to rejecting the null hypothesis in larger samples. In this sense, further visual analyses were performed to support the statistical results. Figure 2 shows the histogram of the data, the fit of a normal distribution to the log-transformed data, and the Q-Q plots for Period 2. This result, observed in both periods, aligns with earlier studies showing that radon concentrations tend to follow a log-normal distribution due to the multiplicative nature of the processes influencing its accumulation [35,36].

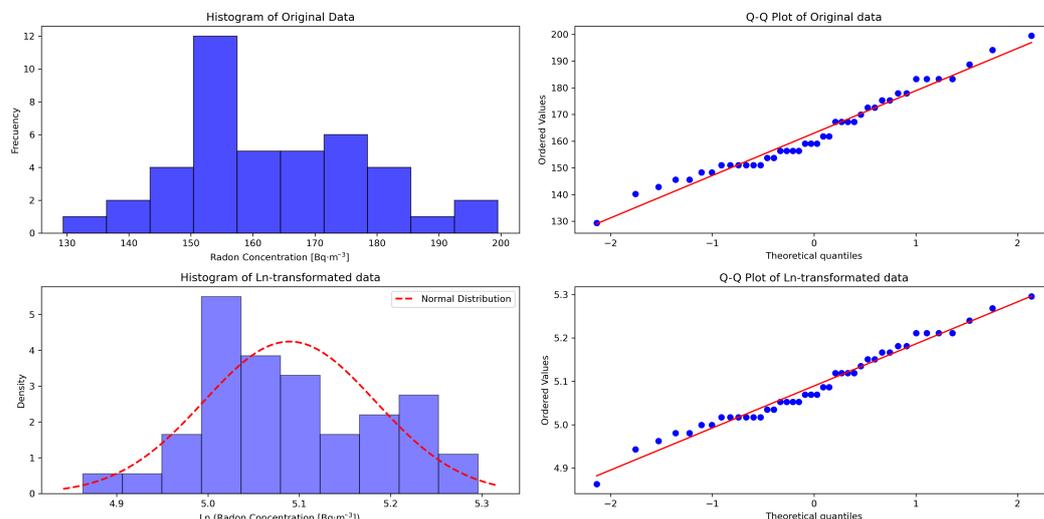


Figure 2. Graph histograms, the normal distribution fit to the log-transformed data, and the Q-Q plots for measurement in Period 2. **Top left:** Histogram of original data with log-normal fit. **Top right:** Q-Q plot of original data. **Bottom left:** Histogram of log-transformed data with normal fit. **Bottom right:** Q-Q plot of log-transformed data.

Table 3 summarizes the normality test results for Period 3. The Shapiro–Wilk and Anderson–Darling tests reject normality, while the Kolmogorov–Smirnov does not. This inconsistency reflects the varying sensitivities of normality tests to sample size and specific data characteristics. In this sense, the normality assumption for Period 3 is not strongly supported.

Table 3. Results of normality tests in Period 3.

Test	Statistic	p-Value	Interpretation
Shapiro–Wilk	0.9193	0.0019	Rejects normality
Anderson–Darling	1.3787	0.726 (5% CV)	Rejects normality
Kolmogorov–Smirnov	0.1346	0.0227	Rejects normality

The Shapiro–Wilk test, known for its sensitivity in small samples [37], was supplemented with visual inspections through histograms and Q-Q plots, as recommended in previous statistical analyses [38]. Figure 3 illustrates the graphical analysis for Period 3, including the original data, log-transformed data, and Box–Cox-transformed data. In the first row, the histogram of radon concentration and the Q-Q plot for the expected log-normal distribution reveal a poor fit to log-normality. Similarly, the second row presents the histogram and Q-Q plot for the log-transformed data, which also indicate an unsatisfactory fit to normality.

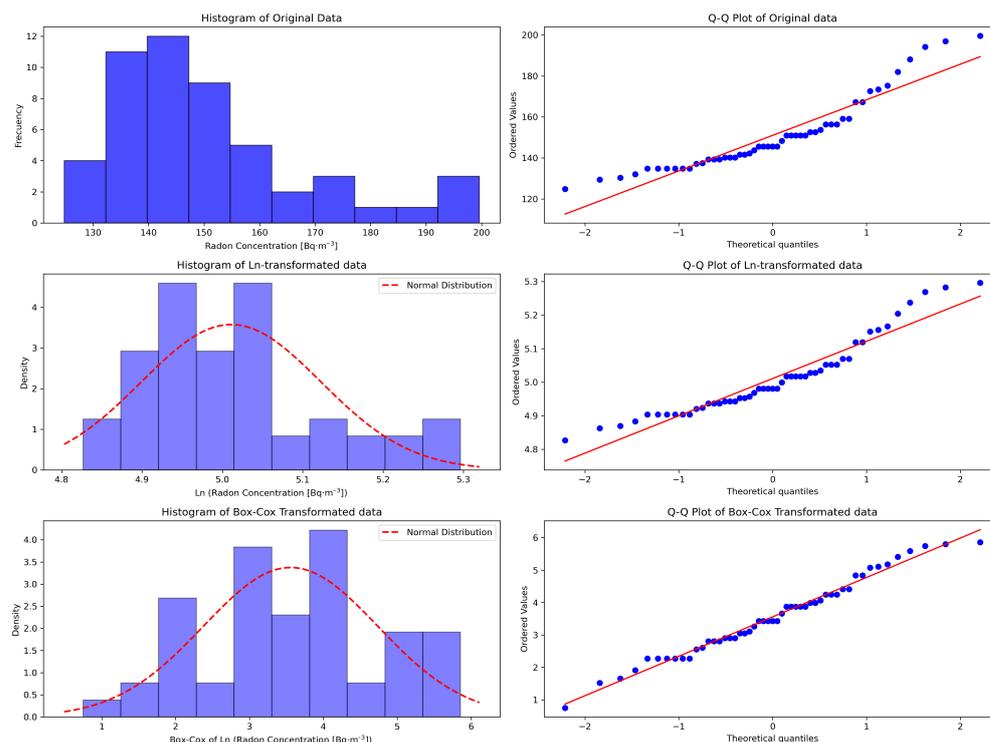


Figure 3. Graph of the fit to the normality curve for measurement in Period 3. **Top left:** Histogram of original data with log-normal fit. **Top right:** Q-Q plot of original data. **Middle left:** Histogram of log-transformed data with normal fit. **Middle right:** Q-Q plot of log-transformed data. **Bottom left:** Histogram of Box–Cox-transformed data with normal fit. **Bottom right:** Q-Q plot of Box–Cox-transformed data.

Finally, the normality tests conducted on the log-transformed radon concentration data across the three measurement periods indicated that the data do not follow a normal distribution, as shown in Table 4. A Box–Cox transformation with a lambda value of 0.5694 was applied to normalize the combined dataset from the three periods, resulting in 86 measurement points. Specifically, the first period contributed 34 data points, the second added 42, and the third included repeated measurements for most households while incorporating new measurement locations. In cases where multiple measurements were taken at the same location, the values were averaged to produce the final dataset used for the generation of the radon map. Post-transformation normality tests confirmed that the data followed a normal distribution, making them more suitable for the geostatistical methods required to generate the maps described below [39].

Table 4. Comparison of normality test results before and after Box–Cox transformation.

Test	Statistic	p-Value	Interpretation
Shapiro–Wilk	0.9706	0.0497	Rejects normality
Anderson–Darling	0.7649	0.726 (5% CV)	Rejects normality
Kolmogorov–Smirnov	0.1000	0.0486	Rejects normality
Shapiro–Wilk (Box–Cox)	0.9764	0.1212	Does not reject normality
Anderson–Darling (Box–Cox)	0.6983	0.726(5% CV)	Does not reject normality
Kolmogorov–Smirnov (Box–Cox)	0.0888	0.1325	Does not reject normality

A visual analysis of the data is shown in Figure 4.

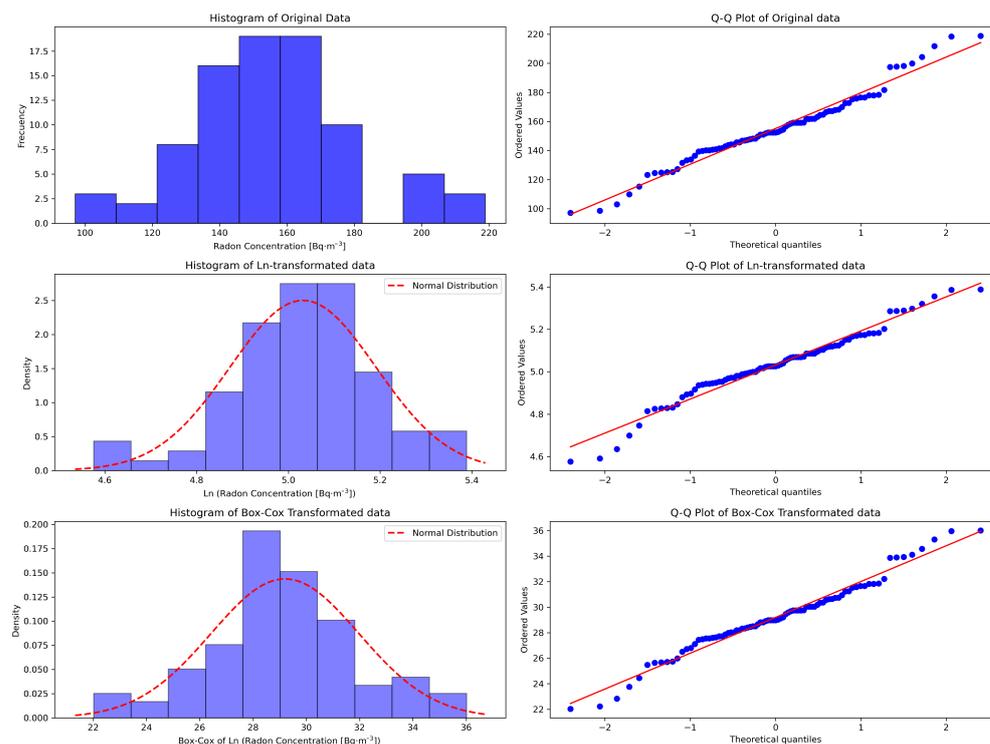


Figure 4. Graph of the fit to the normality curve for measurement in all periods. **Top left:** Histogram of original data with log-normal fit. **Top right:** Q-Q plot of original data. **Middle left:** Histogram of log-transformed data with normal fit. **Middle right:** Q-Q plot of log-transformed data. **Bottom left:** Histogram of Box–Cox-transformed data with normal fit. **Bottom right:** Q-Q plot of Box–Cox-transformed data.

An analysis of variance (ANOVA) was performed to determine whether there are statistically significant differences in radon concentrations between Period 1, Period 2, and Period 3. The results (F -statistic = 3.182, p -value = 0.0449) indicate statistically significant differences at the 0.05 significance level. These differences could be attributed to seasonal variations and changes in environmental conditions, such as temperature and humidity, which affect the emanation and accumulation of radon [40,41]. Additionally, ventilation habits and occupancy of the residents can influence indoor radon levels [42].

The boxplot in Figure 5 illustrates the comparison of indoor radon concentrations across the three periods. Regarding GM and GSD, Period 1 exhibited a GM of $146.4 \text{ Bq}\cdot\text{m}^{-3}$ with a GSD of 1.3. Period 2 recorded a GM of $162.3 \text{ Bq}\cdot\text{m}^{-3}$ and a GSD of 1.18, while Period 3 showed a GM of $150.8 \text{ Bq}\cdot\text{m}^{-3}$ with a GSD of 1.14. Finally, GM and GSD of the combined data for all measurement periods are 153.2 and 1.2, respectively.

The probability of exceeding the radon concentration threshold of $200 \text{ Bq}\cdot\text{m}^{-3}$ was calculated using Equations (7) and (9). The significance of differences between the three measurement periods was assessed based on the log-normal distribution of radon concentrations. The confirmed normality of the log-transformed data validated this calculation. The likelihood of exceeding this threshold was estimated at 11.7% for Period 1, 1.43% for Period 2, and 8.7% for Period 3. When the data from all three measurement periods were combined to generate the map, the overall probability of homes exceeding $200 \text{ Bq}\cdot\text{m}^{-3}$ was estimated at 9.5%.

Figure 6 presents two maps of radon concentration in the San Martín de Porres district, incorporating data from the three periods. The map on the left was generated using the Inverse Distance Weighted (IDW) interpolation method, while the map on the right was created using Kriging. The IDW interpolation method has a limitation in identifying regions where radon concentrations may exceed $200 \text{ Bq}\cdot\text{m}^{-3}$, as a single data point may not be

representative of the surrounding area, especially if the measurement point reflects an atypical condition. This is because this method only weights values based on distance, without considering spatial relationships between points or the natural variability of the phenomenon. As a result, areas of high concentration are generated abruptly and locally, as observed in the map, where two red points stand out due to the influence of a single measurement point. This lack of spatial continuity in the estimates makes it difficult to achieve a realistic representation and may lead to inaccurate interpretations of high-risk areas.

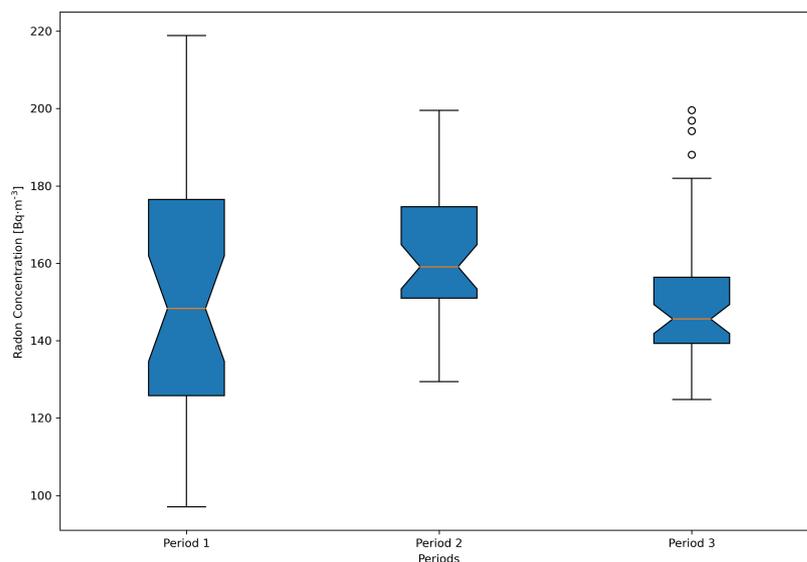


Figure 5. Box-whisker plots showing the variation in radon concentration in the three periods. The plots illustrate the range, median, and variability of radon concentrations across Period 1, Period 2, and Period 3, providing insights into the distribution and potential outliers within each period.

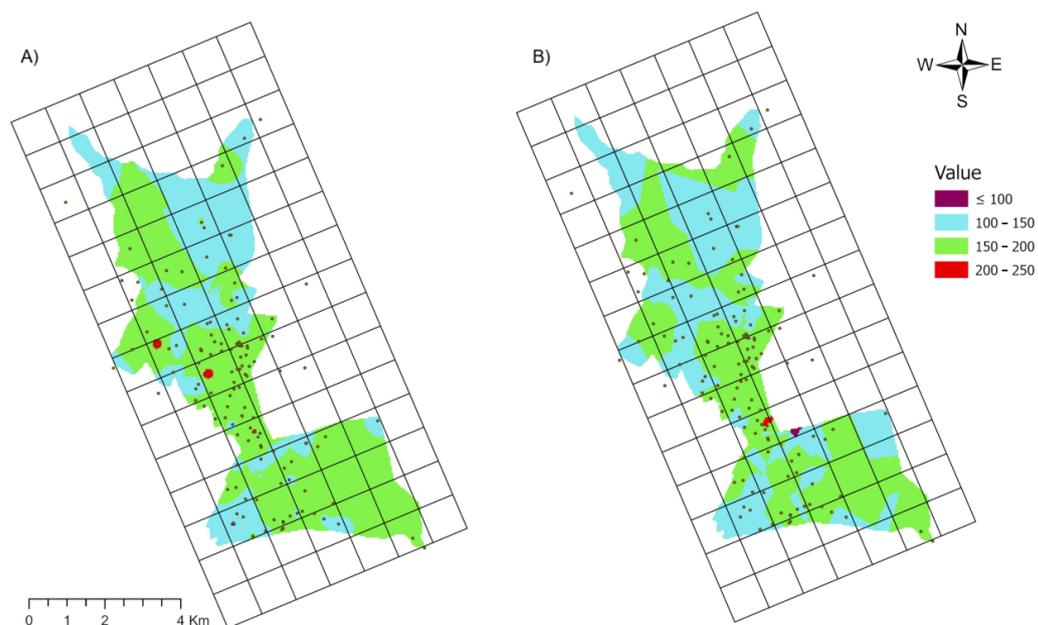


Figure 6. Predicted indoor radon map of San Martín de Porres dwellings over a grid with the dimensions of 1 km × 1 km using the (A) Inverse Distance Weighting (IDW) and (B) Ordinary Kriging interpolation methods.

In contrast, the map generated using Kriging addresses these limitations by smoothing transitions between zones and incorporating both spatial relationships and the natural

variability of radon. The red point, which indicates a high radon concentration, is located in an area where transitions to lower values in surrounding areas are smoother. This suggests that, although it represents an extreme value, Kriging incorporates it as part of a broader spatial trend by evaluating its influence based on the statistical correlation with adjacent points. Unlike IDW, which generates abrupt and highly localized “islands” around high values, Kriging integrates this red point more progressively into the overall pattern, achieving more realistic estimates that are less dependent on individual data points. Therefore, in the analyzed case, the map generated by Kriging provides improved reliability in identifying high-risk regions.

4. Conclusions

This study comprehensively analyzes the spatial and temporal distribution of indoor radon concentrations in San Martín de Porres, Lima. The significant variations observed across the three measurement periods underscore the complex interplay of environmental, structural, and human factors that influence radon accumulation in residential settings. Moreover, the application of the Box–Cox transformation significantly improved the normality of the data, thereby enhancing the precision of subsequent geostatistical analyses. This finding emphasizes the importance of transforming and validating data before applying advanced analytical techniques. The study revealed notable differences in radon concentrations across the three measurement periods, further highlighting the critical role of continuous monitoring in accurately assessing long-term exposure risks. Additionally, the results demonstrate that radon concentrations exhibit both temporal fluctuations and spatial heterogeneity, which carry important implications for risk assessment and public health strategies. The detection of statistically significant differences among the periods suggests that factors such as seasonal changes, environmental conditions, and occupant behaviors substantially impact indoor radon levels. It is essential to highlight that the radon concentrations measured during the three periods represent an estimate of the annual average behavior, considering the temporal limitation of the measurements. Future studies will aim to expand the temporal coverage, allowing for representative weighted values of yearly radon concentrations in the evaluated areas. This effort will enable a more comprehensive assessment of the associated risks and the validation of the applied methods. In summary, this study demonstrates the value of advanced statistical methods in improving data accuracy and reinforces the necessity of continuous monitoring to manage radon exposure effectively.

Author Contributions: Conceptualization, R.L.; methodology, R.L., D.P. and P.P.; software, F.D.; validation, F.D., N.C. and L.C.; formal analysis, R.L., F.D., D.P. and P.P.; investigation, R.L., F.D., D.P. and P.P.; resources, R.L. and P.P.; data curation, N.C., L.C. and M.R.; writing—original draft preparation, R.L., F.D. and N.C.; writing—review and editing, R.L., F.D. and D.P.; visualization, L.C. and M.R.; supervision, D.P. and P.P.; project administration, R.L.; funding acquisition, R.L. and F.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lorenzo-Gonzalez, M.; Torres-Duran, M.; Barbosa-Lorenzo, R.; Provencio-Pulla, M.; Barros-Dios, J.M.; Ruano-Ravina, A. Radon exposure: A major cause of lung cancer. *Expert Rev. Respir. Med.* **2019**, *13*, 839–850. [[CrossRef](#)] [[PubMed](#)]

2. IARC. *Man-Made Mineral Fibres and Radon IARC Working Group on the Evaluation of Carcinogenic Risks to Humans and International Agency for Research on Cancer*; IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals; IARC: Lyon, France, 1988; Volume 43.
3. UNSCEAR. *Sources and Effects of Ionizing Radiation*; United Nations: New York, NY, USA, 2000.
4. Valentin, J. (Ed.) *The 2007 Recommendations of the International Commission on Radiological Protection*; Elsevier: Oxford, UK, 2007; Volume 37. [[CrossRef](#)]
5. International Atomic Energy Agency. *Radiation Protection and Safety of Radiation Sources International Basic Safety Standards*; International Atomic Energy Agency: Vienna, Austria, 2011.
6. Davis, E.A.; Ou, J.Y.; Chausow, C.; Verdeja, M.A.; Divver, E.; Johnston, J.D.; Beard, J.D. Associations between school characteristics and classroom radon concentrations in Utah's public schools: A project completed by university environmental health students. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5839. [[CrossRef](#)] [[PubMed](#)]
7. Azara, A.; Dettori, M.; Castiglia, P.; Piana, A.; Durando, P.; Parodi, V.; Salis, G.; Saderi, L.; Sotgiu, G. Indoor radon exposure in Italian schools. *Int. J. Environ. Res. Public Health* **2018**, *15*, 749. [[CrossRef](#)] [[PubMed](#)]
8. Alrowaili, Z.A. Nature of radon, radium, exhalation and uranium concentration from construction materials used in Al Jouf city, Saudi Arabia. *J. Radiat. Res. Appl. Sci.* **2023**, *16*, 100579. [[CrossRef](#)]
9. Cosma, C.; Cucos-Dinu, A.; Papp, B.; Begy, R.; Sainz, C. Soil and building material as main sources of indoor radon in Băița-Ștei radon prone area (Romania). *J. Environ. Radioact.* **2013**, *116*, 174–179. [[CrossRef](#)]
10. Schubert, M.; Musolf, A.; Weiss, H. Influences of meteorological parameters on indoor radon concentrations (²²²Rn) excluding the effects of forced ventilation and radon exhalation from soil and building materials. *J. Environ. Radioact.* **2018**, *192*, 81–85. [[CrossRef](#)]
11. Nero, A.V.; Nazaroff, W. Characterising the source of radon indoors. *Radiat. Prot. Dosim.* **1984**, *7*, 23–39. [[CrossRef](#)]
12. Bossew, P. Radon: Exploring the log-normal mystery. *J. Environ. Radioact.* **2010**, *101*, 826–834. [[CrossRef](#)]
13. Branco, P.T.; Nunes, R.A.; Alvim-Ferraz, M.C.; Martins, F.G.; Sousa, S.I. Children's exposure to radon in nursery and primary schools. *Int. J. Environ. Res. Public Health* **2016**, *13*, 386. [[CrossRef](#)]
14. Lorenzo-González, M.; Ruano-Ravina, A.; Peón, J.; Piñeiro, M.; Barros-Dios, J.M. Residential radon in Galicia: A cross-sectional study in a radon-prone area. *J. Radiol. Prot.* **2017**, *37*, 728. [[CrossRef](#)]
15. García-Talavera, M.; García-Pérez, A.; Rey, C.; Ramos, L. Mapping radon-prone areas using γ -radiation dose rate and geological information. *J. Radiol. Prot.* **2013**, *33*, 605. [[CrossRef](#)] [[PubMed](#)]
16. Fernández, A.; Sainz, C.; Celaya, S.; Quindós, L.; Rábago, D.; Fuente, I. A new methodology for defining radon priority areas in Spain. *Int. J. Environ. Res. Public Health* **2021**, *18*, 1352. [[CrossRef](#)] [[PubMed](#)]
17. Park, T.H.; Kang, D.R.; Park, S.H.; Yoon, D.K.; Lee, C.M. Indoor radon concentration in Korea residential environments. *Environ. Sci. Pollut. Res.* **2018**, *25*, 12678–12685. [[CrossRef](#)] [[PubMed](#)]
18. Ivanova, K.; Stojanovska, Z.; Tsenova, M.; Badulin, V.; Kunovska, B. Measurement of indoor radon concentration in kindergartens in Sofia, Bulgaria. *Radiat. Prot. Dosim.* **2014**, *162*, 163–166. [[CrossRef](#)]
19. Appleton, J.; Miles, J. A statistical evaluation of the geogenic controls on indoor radon concentrations and radon risk. *J. Environ. Radioact.* **2010**, *101*, 799–803. [[CrossRef](#)]
20. Nikolopoulos, D.; Kottou, S.; Louizi, A.; Petraki, E.; Vogianis, E.; Yannakopoulos, P. Factors affecting indoor radon concentrations of Greek dwellings through multivariate statistics—first approach. *J. Phys. Chem. Biophys.* **2014**, *4*, 1000145. [[CrossRef](#)]
21. Yarmoshenko, I.; Vasilyev, A.; Malinovsky, G.; Bossew, P.; Žunić, Z.S.; Onischenko, A.; Zhukovsky, M. Variance of indoor radon concentration: Major influencing factors. *Sci. Total. Environ.* **2016**, *541*, 155–160. [[CrossRef](#)]
22. Bachirou, S.; Saïdou, K.; Kranrod, C.; Nkoulou, J.E.N., II; Bongue, D.; Abba, H.Y.; Hosoda, M.; Njock, M.G.K.; Tokonami, S. Mapping in a radon-prone area in Adamawa region, Cameroon, by measurement of radon activity concentration in soil. *Radiat. Environ. Biophys.* **2023**, *62*, 427–439. [[CrossRef](#)]
23. Hinojosa de la Garza, O.R.; Sanín, L.H.; Montero Cabrera, M.E.; Serrano Ramirez, K.I.; Martínez Meyer, E.; Reyes Cortés, M. Lung cancer mortality and radon concentration in a chronically exposed neighborhood in Chihuahua, Mexico: A geospatial analysis. *Sci. World J.* **2014**, *2014*, 935380. [[CrossRef](#)]
24. Akuo-ko, E.O.; Adelikhah, M.; Amponsem, E.; Csordás, A.; Kovács, T. Investigations of indoor radon levels and its mapping in the Greater Accra region, Ghana. *J. Radioanal. Nucl. Chem.* **2024**, *333*, 2975–2986. [[CrossRef](#)]
25. Munyati, C.; Sinthumule, N. Comparative suitability of ordinary kriging and Inverse Distance Weighted interpolation for indicating intactness gradients on threatened savannah woodland and forest stands. *Environ. Sustain. Indic.* **2021**, *12*, 100151. [[CrossRef](#)]
26. Elío, J.; Cinelli, G.; Bossew, P.; Gutiérrez-Villanueva, J.L.; Tollefsen, T.; De Cort, M.; Nogarotto, A.; Braga, R. The first version of the pan-European indoor radon map. *Nat. Hazards Earth Syst. Sci.* **2019**, *19*, 2451–2464. [[CrossRef](#)]
27. Friedmann, H.; Gröller, J. An approach to improve the Austrian Radon Potential Map by Bayesian statistics. *J. Environ. Radioact.* **2010**, *101*, 804–808. [[CrossRef](#)] [[PubMed](#)]

28. Andersen, C.E.; Raaschou-Nielsen, O.; Andersen, H.P.; Lind, M.; Gravesen, P.; Thomsen, B.L.; Ulbak, K. Prediction of ²²²Rn in Danish dwellings using geology and house construction information from central databases. *Radiat. Prot. Dosim.* **2007**, *123*, 83–94. [[CrossRef](#)]
29. Hauri, D.D.; Huss, A.; Zimmermann, F.; Kuehni, C.E.; Rösli, M. A prediction model for assessing residential radon concentration in Switzerland. *J. Environ. Radioact.* **2012**, *112*, 83–89. [[CrossRef](#)]
30. Mäkeläinen, I.; Arvela, H.; Voutilainen, A. Correlations between radon concentration and indoor gamma dose rate, soil permeability and dwelling substructure and ventilation. *Sci. Total. Environ.* **2001**, *272*, 283–289. [[CrossRef](#)]
31. Kemski, J.; Klingel, R.; Siehl, A.; Valdivia-Manchego, M. From radon hazard to risk prediction-based on geological maps, soil gas and indoor measurements in Germany. *Environ. Geol.* **2009**, *56*, 1269–1279. [[CrossRef](#)]
32. Pereyra, P.; Guevara-Pillaca, C.J.; Liza, R.; Pérez, B.; Rojas, J.; Vilcapoma L, L.; Gonzales, S.; Sajo-Bohus, L.; López-Herrera, M.E.; Palacios Fernández, D. Estimation of Indoor ²²²Rn Concentration in Lima, Peru Using LR-115 Nuclear Track Detectors Exposed in Different Modes. *Atmosphere* **2023**, *14*, 952. [[CrossRef](#)]
33. Ferreira, T.; Rasband, W. *ImageJ User Guide*; National Institutes of Health: Bethesda, MD, USA, 2011.
34. Laio, F. Cramer–von Mises and Anderson–Darling goodness of fit tests for extreme value distributions with unknown parameters. *Water Resour. Res.* **2004**, *40*, W09308. [[CrossRef](#)]
35. Daraktchieva, Z. Variability of indoor radon concentration in UK homes. *J. Radiol. Prot.* **2021**, *41*, 853. [[CrossRef](#)]
36. Aladeniyi, K.; Arogunjo, A.M.; Pereira, A.J.; Ajayi, O.S.; Fuwape, I.A. Radiometric evaluation of indoor radon levels with influence of building characteristics in residential homes from southwestern Nigeria. *Environ. Monit. Assess.* **2020**, *192*, 764. [[CrossRef](#)] [[PubMed](#)]
37. Sarin, A.; Kaushal, A.; Bajwa, B.S.; Sharma, N. Quantification of doses and health risks to organs and tissues corresponding to different age groups due to radon in water. *J. Radioanal. Nucl. Chem.* **2021**, *330*, 643–655. [[CrossRef](#)]
38. Rani, S.; Kundu, R.S.; Garg, V.K.; Singh, B.; Panghal, A.; Dilbaghi, N. Radon and thoron exhalation rate in the soil of Western Haryana, India. *Environ. Monit. Assess.* **2023**, *195*, 523. [[CrossRef](#)] [[PubMed](#)]
39. Kellenbenz, K.R.; Shakya, K.M. Spatial and temporal variations in indoor radon concentrations in Pennsylvania, USA from 1988 to 2018. *J. Environ. Radioact.* **2021**, *233*, 106594. [[CrossRef](#)] [[PubMed](#)]
40. Hansen, V.; Petersen, D.; Søgaard-Hansen, J.; Rigét, F.F.; Mosbech, A.; Clausen, D.S.; Mulvad, G.; Rönnqvist, T. Indoor radon survey in Greenland and dose assessment. *J. Environ. Radioact.* **2023**, *257*, 107080. [[CrossRef](#)]
41. Elmehdi, H.M.; Ramachandran, K.; Gaidi, M.; Daoudi, K. Diurnal and seasonal influence on the indoor radon levels in dwellings of Sharjah Emirate as well its estimation of annual effective dose. *Case Stud. Chem. Environ. Eng.* **2024**, *9*, 100663. [[CrossRef](#)]
42. Vienneau, D.; Boz, S.; Forlin, L.; Flückiger, B.; De Hoogh, K.; Berlin, C.; Bochud, M.; Bulliard, J.L.; Zwahlen, M.; Rösli, M. Residential radon—Comparative analysis of exposure models in Switzerland. *Environ. Pollut.* **2021**, *271*, 116356. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.