Factors Affecting Indoor Radon Concentrations of Greek Dwellings through Multivariate Statistics

Dimitrios Nikolopoulos1, Sofía Kottou2, Anna Louizi2, Ermioni Petraki1,3, Efstratios Vogiannis4 and Panayiotis H Yannakopoulos5

Abstract

A large scale nationwide radon survey was conducted in Greek dwellings between 1994 and 2000. Twelve hundred passive CR-39 detectors were distributed and collected along with 963 filled questionnaires. These were rechecked during 2012-13 to evaluate factors that affect indoor radon concentrations, such as i) area, ii) building level-floor, iii) ground type, iv) basement, v) building type, vi) construction year, vii) building walls contact, viii) wall materials, ix) floor materials. The questionnaires were prepared by the research team according to international standards. One-way and multivariate statistical methods were applied for the analysis: i) Linear Regression Analysis, ii) One way or multway ANOVA, iii) General MANOVA, iv) Stepwise Regression Analysis, v) Principal Components Analysis. Results revealed that approximately 0.1% of the dwellings exhibited outlier radon concentrations. Noteworthy statistical correlations were detected between indoor radon concentration and the factors: “building level-floor” and “wall materials”. Results of current work strengthened these considerations and provided weak evidence for the correlation of factors like “building type”, “construction year” and “floor materials” with radon concentrations. Minor association was detected with the factor “building walls contact”. Significant differences were detected in results produced by some of the applied statistical methods.

Introduction

Natural environmental radiation depends on local geology and hence, variations are addressed in human radiation exposure due to cosmic and terrestrial radiation [1]. 238U and 232Th are two natural parent isotopes which are present in soil and contribute significantly to natural terrestrial radioactivity. Radon 222Rn is a radioactive noble gas and originates from 238U. 220Rn originates from 232Th and 219Rn from 235U. 222Rn, 220Rn, 219Rn are the primary sources of radon in soil, with 222Rn being dominant in rocks, soil, building materials, underground and surface waters [2] and set to be the most hazardous radionuclide. Radon (222Rn) and its short-lived progeny (218Po, 214Po, 214Bi, 214Pb) are and surface waters [2] and set to be the most hazardous radionuclide.

Materials and Methods

A thorough investigation was performed on whether nine factors: i) area, ii) building level-floor, iii) ground type, iv) basement, v) building type, vi) construction year, vii) building walls contact, viii) wall materials, ix) floor materials, may affect indoor radon concentration independently or jointly. These factors have been recorded on 963 filled questionnaires of the greater large-scale survey in Greece [26]. A multivariate statistical analysis, based on i) Linear Regression Analysis, ii) One way or multway ANOVA, iii) General MANOVA, iv) Stepwise Regression Analysis and v) Principal Components Analysis methods, was implemented on the questionnaire data. It is noted that these data were dispersed across Greece.

Linear regression analysis

In linear regression analysis, a straight line is fitted through a set of points-observations in such a way that the sum of squared residuals is minimal [34]. In multiple linear regression the dependent variable can be written in terms of a linear combination of the independent variables. The regression equation describes the correlation of the mean value of a variable-y with specific values of x-variables used to predict y.

*Corresponding author: Dimitrios Nikolopoulos, Department of Computer Electronic Engineering, TEI of Piraeus, Greece, Petrou Ralli & Thivon 250, 122 44, Aigaleo, Athens, Greece, Tel: +0030-210-5381110, Mobile +0030-6977-208318; Fax: +0030-210-5381436; E-mail: dniko@teipir.gr, NikolopoulosDimitrios@gmail.com

Received April 2, 2014; Accepted May 22, 2014; Published May 24, 2014


Copyright: © 2014 Nikolopoulos D, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Suppose that \((x_i,y_i),(x_i,y_i),\ldots,(x_i,y_i)\) are the realisations of random variable pairs \((X_i,Y_i),(X_i,Y_i),\ldots,(X_i,Y_i)\), then the linear regression equation expressed the mean of \(Y\) as a straight-line function of \(X\) and could be represented as
\[
E(Y_i) = \beta_0 + \beta_1 X_i
\]
(2.1.1)
or
\[
E(Y_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p
\]
for \(p\) independent-predictor variables.

\(E(Y_i)\) states the mean expected value and \(i\) points the population. The estimated/fitted model is then:
\[
Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i
\]
(2.1.2)
From (2.1.2), the estimated/fitted values for each of the \(n\) observations are
\[
Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i
\]
(2.1.3)
where \(i = 1,\ldots,n\) is the consecutive number of the population. From (2.1.2) and (2.1.3) the, so called, observed error or fitted residual is calculated as:
\[
e_i = Y_i - Y_i
\]
(2.1.4)
Equation (2.1.4) calculates the estimated error of the \(i\)-th observation in the sample. From (2.1.4) the sum of squared observed errors (SSE) equals
\[
SSE = \sum (Y_i - Y_i)^2 = \sum e_i^2
\]
(2.1.5)
for all observations in a sample of size \(n\). The mean square error (MSE) equals then
\[
MSE = \frac{SSE}{n-2} = \frac{\sum e_i^2}{n-2}
\]
(2.1.6)
(“\(n-2\)” should be substituted by “\(n-p-1\)” when there are \(p\) predictor-independent variables) and this is the sample variance of error. The residual standard error is then calculated as
\[
\hat{\sigma} = \sqrt{MSE}
\]
(2.1.7)
and \(\sigma^2\) should be the constant error variance, otherwise the confidence intervals will be misleading.

As
\[
Y_i - Y = Y_i - Y + e_i
\]
(2.1.8)
then, the total sum of squares (SST) equals
\[
\sum (Y_i - Y)^2 = \sum (Y_i - Y) + \sum e_i^2
\]
(2.1.9)
with \(Y\) set to be the mean of all observed \(Y\) values.

The coefficient of determination
\[
R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}
\]
(2.1.10)
represents the proportion of variation in \(Y\) that is explained by \(X\) [35]. Parameters \(\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p\) (regression coefficients), \(\sigma^2\) (variance) and \(R^2\) (coefficient of determination) need to be estimated in order to examine if the linear regression model applies to this group of data. However, even if \(R^2\) value is close to zero, this does not mean \(X\) and \(Y\) have no nonlinear association and polynomial terms should be included to improve the fitting.

One Way or multiway ANOVA: Analysis of variance (ANOVA) approach to regression analysis is considered as the generalization of a t-test to more than two statistical groups. ANOVA is divided in two categories: i) One way, where a single factor exists and ii) two way or multiway, where two or more factors exist. ANOVA is used for distributional assumptions about a set of effects in a model, with ability to extrapolate the inference to a wider population, improve accounting for system uncertainty and the efficiency of estimation [36]. ANOVA has been implemented as the basic method for the statistical analysis of Radon concentrations in many studies [15,19,37-40]. For the analysis of factors affecting indoor radon concentrations, initially each factor was analysed independently. In such a way a first assumption for the weightiness of the effect of each factor is possible. In the next step, factors affection are no longer estimated independently; instead, factors influence each other and therefore are dependent. The aforementioned method, is called random-effects assumption of the analysis of variance [36]. ANOVA can be implemented only in sampling distributions similar to Gaussian ones, thus it was applied in the log distributions of radon measurements.

In order to construct the ANOVA table the variability in \(Y\) variable explained or not with the regression relationship of \(X\) variable was measured. This table shows additionally the Mean Square Error (MSE). The overall variation in \(Y\) is equal with the sum of regression variation and the error variation (2.1.9).

The ANOVA table (Table 1) involves the following elements:

The sum of squares for total (SST) which is the sum of squared deviations from the overall mean of \(Y\).

The sum of squared errors (SSE) which is the sum of squared

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>SSR = \sum_{i=1}^{n} (Y_i - \bar{Y})^2</td>
<td>(p)</td>
<td>MSR = SSR / (p)</td>
<td>(F = \frac{MSR}{MSE})</td>
</tr>
<tr>
<td>Error</td>
<td>SSE = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2</td>
<td>(n - p - 1)</td>
<td>MSE = SSE / (n - p - 1)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>SST = \sum_{i=1}^{n} (Y_i - \bar{Y})^2</td>
<td>(n - 1)</td>
<td>SSE / (n - 1)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The ANOVA table.
observed errors for the observed data and is a measure of the variation in $Y$ which is not explained by the regression (2.1.5).

The sum of squares of the regression (SST) defined as the difference between SST and SSE. This is a measure of the total variation of $Y$ that can be explained from the regression with $X$ variable.

The mean of square errors (2.1.6) and

The mean square of the regression (MSR) which, here, equals with the SSR [34].

The F-value, where $F(\alpha, 1, n-2)$ is the $(1-\alpha)$ quantile of the $F$ distribution and $\alpha$ is the significant level.

**General MANOVA:** Multivariate (multiple dependent variables) analysis of variance (MANOVA) is defined as a partition of the sum of squares and the sum of the cross products (SSCP) matrix

\[
SSCP = \begin{bmatrix}
SS_{11} & SCP_{12} \\
SCP_{21} & SS_{22}
\end{bmatrix}
\]  

(2.1.2.1)

into independent Wishart matrices [41]. MANOVA is applied instead of a series of one-at-a-time ANOVAs. In several situations, the power of MANOVA is inferior to ANOVA of one variable at a time, however, MANOVA takes into account the intercorrelations among the dependent variables. Hence, MANOVA is considered more efficient over ANOVA for multivariate data [42].

A MANOVA table (Table 2) includes the following elements:

The sum of squares and cross products for total ($SSCPTO$) which is the sum of squared deviations from the overall mean vector of the $Y_i$ and equals to

\[
SSCPTO = \sum \begin{bmatrix} Y_i - \bar{Y} \end{bmatrix} \begin{bmatrix} Y_i - \bar{Y} \end{bmatrix}^T
\]  

(2.1.2.2)

$SSCPTO$ is considered as a measure of the overall variation in the $Y$ vectors.

The sum of squares and cross-products of the errors ($SSCPE$)

\[
SSCPE = \sum \begin{bmatrix} Y_i - \hat{Y} \end{bmatrix} \begin{bmatrix} Y_i - \hat{Y} \end{bmatrix}^T
\]  

(2.1.2.3)

which is the sum of squared errors (residuals) for the data vectors.

$SSCPE$ is considered as a measure of the variation in $Y$ that is not explained by the multivariate regression.

The sum of squares and cross-products due to the regression ($SSCPR$) is defined as the difference between $SSCPE$ and $SSCPE$:

\[
SSCPR = SSCPTO - SSCPE
\]  

(2.1.2.4)

$SSCPR$ is a measure of the total variation in $Y$ that can be explained by the regression with the predictors [35].

**Issues in Multiple Regression:** The difficulty with model selection emerges from the fact that for $p$ predictors, there are $2^p$ different candidate models. With so many possible interactions it can be difficult to find a good model. Model selection methods try to simplify this task. A true model may only depend on a subset of $X_1, \ldots, X_p$. In other words, in model

\[
Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \varepsilon
\]  

(2.1.3.1)

some of the coefficients are zeros. The result will be the disclosure of those predictors with nonzero coefficients, i.e. the “best subset” of all predictors.

$R^2$ can be used for models with the same number of parameters/coefficients, otherwise $R_{adj}^2$ should be used. The best model has the biggest $R_{adj}^2$ value.

Selecting $p$ predictors, the Mallows’ $C_p$ criterion should be small with a value near to $p$.

**Stepwise Regression Analysis**

Stepwise methods are used in several areas of applied statistics. A statistical model can be constructed in two ways, namely (i) forward selection and (ii) backward elimination. Forward selection means that a specific number of variables exist in the beginning and gradually variables are added, one at a time, in optimal way in order to analyse the effect of each variable. Alternatively, with backward elimination, all potential variables exist in the beginning and non-effective variables are subtracted, one at a time, until a desirable stopping point is reached.

Stepwise regression forms a hybrid model between forward selection and backward elimination. More precisely, steps have a forward direction with variable addition, however if a variable is characterized as non-significant, it is removed as in backward elimination [35,43]. In literature stepwise regression has been used for the prediction of mean indoor radon concentrations [44], in the construction of radon maps based in indoor radon measurements and soil geochemical parameters [45] and in risk analysis of factors affecting lung cancer [46].

**Principal Components Analysis**

Principal components analysis (PCA) is used for the reduction of the number of possible clusters. PCA offers the ability for the identification of patterns within large sets of data [47]. Its significance rely in the occurrence of a relative redundancy in the variables, due to their correlation in the measurement of the same construct. During the analysis of the principal components, eigenvalues represent the relative participation of each factor in presenting the general variability of the sampled data [48]. PCA has several implementations in factors investigation of water quality, in drug development, in cancer detection and in health care [48-53]. PCA has been also used for the investigation of the dependence among variables and for the prediction of relationships among variables [54].

A standardisation of the various data is performed prior to analysis in order to ensure that each variable influences in the same way during the analysis. The standardised variable is the following

\[
z_i = \frac{x_i - \mu_i}{\sqrt{\sum_{i=1}^{p} \mu_i}}, i = 1, \ldots, p
\]  

(2.3.1)

with $x_i$ set to be the original variable, $\mu_i$ the mean and $Y_i$ the
variance.

Principal components $Y$ are linear combinations of $P$ random variables $X_1, X_2, \ldots, X_p$. If $S=S_p$ is a $P \times P$ sample covariance matrix with eigenvalue-eigenvector pairs

\[
\begin{pmatrix}
\lambda_1 & e_1 \\
\lambda_2 & e_2 \\
\vdots & \vdots \\
\lambda_p & e_p
\end{pmatrix}
\]

, and $x$ set to be the number of principal components, then the $i$-th principal component sample is given by

\[
Y_i = e_1^T X_1 + e_2^T X_2 + \cdots + e_p^T X_p = e_i^T X
\tag{2.3.2}
\]

where $x$ stands as any observation on the variables $X_1, X_2, \ldots, X_p$ and $i=1,2,\ldots, p$ [55].

**Results and Discussion**

Figure 1 presents characteristic residual plots calculated from the measured average concentrations of radon ($C$) and their logarithms ($\log(C)$). It is noted that the $C$ values of Figure 1 correspond to time-integration over a year, constitute representative sample for Greece, were derived in accordance to international standards and delineate the radon profile of Greece [26]. In this consensus, Figure 1 is of significance since it may show actual tendencies regarding the randomness or predictability of the employed concentration sample. Indeed, completely randomised responses to normal-distribution either of $C$ or $\log(C)$, would exhibit no deterministic normal-distribution’s residuals and, hence, be completely described by stochastic processes. The normal probability plots of Figures 1a and 1b indicate, however, that the logarithms of the measured concentration followed normal distribution up to the 95% of $C$ values, namely indicated that $C$ values followed log-normal distribution. This is also evident from the shapes of the frequency distributions of the residuals. The frequency distribution of Fig.1a was clearly log-normal, while simultaneously that of Figure 1b, clearly normal. This is also of significance because all international large-scale radon surveys reported log-normal behaviour of indoor radon concentrations. The reason is rational thus why $C$ values did not follow normal distribution as shown in the corresponding normal probability plot of Figure 1a. Under another view, the residual plots of $\log(C)$ versus values fitted to normal-distribution, showed a random pattern for fitted residuals of $\log(C)$ above 1.6. It is noted that a residual of 1.6 inlog ($C$) is consistent with uncertainty $\sigma = \frac{399 \times 1000 m^3}{\bar{C}}$ in predicted $C$ values. This, according to the recording capabilities of the employed dosemeters [56], accompany high $C$ values, namely $C$ values above the EU action limit of 200 $Bq m^{-3}$. Moreover, the majority of predicted residuals were below 1.2. This is very significant because this residual range is consistent with concentrations usually addressed, namely between 10-120 $Bq m^{-3}$. Other factors may affect concentrations in this range and surely the potential factors could not be continuous under the normal distribution. Indeed, the Versus Fits diagram of Figure 1a shows characteristic predictability different from the normal distribution for the residual $C$ range below 75 $Bq m^{-3}$. On the other hand, the residual versus observation order did not showed tendencies for concentrations up 160 $Bq m^{-3}$, either in the concentration order (Figure1a) or the order of concentration’s logarithm.

Table 3 presents the analysed factors, factor levels and level values with their corresponding description. Data of Table 3 were formulated in accordance to the contents of the 963 questionnaires which were filled during the radon survey of Greece. It is noted that these questionnaires were developed in agreement to other national surveys. The majority of factors exhibited 3-4 levels. This is noteworthy in any related analysis, since a multi-level collection of factors can distract results especially for limited number of measurements. Factor (Level-L) was 5-level marking the usual situation of apartment dwellings in big cities of Greece. Nevertheless, this 5-level factor is easily convertible to a lower-level one. Factor (Floor’s material- F) was free to fill, so a 6 level collection was finally achieved.

It is well identified that Gauss distribution offers a rigid and justified pathway for statistical analysis. Since concentrations logarithms followed the distribution of Gauss, $\log(C)$ was considered favourable. Hereafter any analysis was conducted only on $\log(C)$. Table 4 presents the unusual observations in $\log(C)$ values accounting that these followed normal distribution, viz. were treated according to the distribution of Gauss. Leverage points were considered to be those observations corresponding to extreme or outlying values of $\log(C)$ in a manner that any lack of neighbouring observations implied that the fitted Gaussian regression model passed close to the particular observation. In specific, leverage points were calculated by moving all points one-by-one up or down and calculating the proportionally constant (leverage) of the change of the corresponding Gaussian fitted value. Outliers were calculated as the observations that presented residuals above 1.5 times the interquartile range. According to Table 4, eight outlier and two leverage residual points were identified. In any case, unusual $\log(C)$ values were approximately 0.1% of the total concentration sample size. Therefore, they constituted a negligible part of measurements. Importantly, however, the latter finding indicates that only a small portion (<0.1%) of Greek dwellings presented unusual concentrations. Considering that high unusual concentration extremes may associate with high human radiation burden, this fact implies that indoor radon in Greece may not lie in the international extremes. Emphasis should be stressed also on the fact that outlier data affect any type of fit and should be removed prior to regression analysis, whereas, leverage point may or may not affect. For this reason, all outlier and leverage points were finally removed from the dataset.

Table 5 presents the combinations to define the best subsets from the nine factors of Table 3 for the regression of $\log(C)$. As in Table 4, regression was linear to the factors employed in each entry of Table 5. Mallows’ $C_p$ (MCP) was calculated for each subset of $k, k \leq p$ of explanatory variables, as $C_p = SSE_{(k)} / \bar{C}^2 + \text{msr} \cdot \text{mseo}$, where $SSE_{(k)}$ was the residual sum of squares for the subset model containing $P$ explanatory variables counting the intercept (i.e., the number of parameters in the subset model) and $n$ is the sample size. It should be emphasised that, acceptable models in the sense of minimizing the total bias of predicted values, are those models for which $C_p$ approaches the value $p$, i.e.,
those subset models that fall near the line $C_p = p$ in a plot of $C_p$ against $p$ for the collection of all subset models under consideration. Under this view, only the combination of all factors except factor G (Ground Type, Table 3) constitutes an explanatory subset for minimising total bias.

Additionally to the analysis presented so-far, one-way ANOVA was applied to single-factor data from the whole data set. Analysing factor Level-L in its full depth, an $f$ value of 2.551 was calculated whereas the critical $f$ value at the 95% confidence interval-CI is 2.03. These values imply that with $P=0.014$ results from the various different levels collected did not differ significantly. However when the full dwelling-level data were reorganised in 3-levels (ground floor, first floor and upper floors) $f$ value was found equal to 7.156 while the corresponding critical value at the 95% CI is 3.01 with corresponding value $P<0.001$. This finding is significant since it implies that lower level dwellings present higher indoor radon concentrations. Further evidence was provided by reorganising dwelling-level data in 2 levels,
namely ground floor and higher floor dwellings. Applying t-test to the average concentrations it was calculated that at \( p < 0.001 \) ground floor dwellings presented higher radon concentrations. Analysing factor “Wall Contact-Con” with one-way ANOVA, an \( f \) value of 0.893 was calculated whereas the critical \( f \) value at the 95% CI is 2.624. Namely, at \( p < 0.001 \) different Wall Contact - factor level dwellings did not present differences. Similar was the outcomes of the one-way ANOVA for factor “Floor’s material-F”. Non-significant variations were addressed, since calculated \( f \) value was 1.298 and the critical value at 95% CI, 2.119. On the contrary, the one-way analysis of factor “Wall’s material-W”, provided an \( f \) value of 4.314 with a critical value at 95% CI of 2.624 and an associated \( P \) value of 0.0051. The latter finding was associated with a tendency of higher concentrations of rock dwelling.

<table>
<thead>
<tr>
<th>Factor</th>
<th>( p ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area-A</td>
<td>0.131</td>
</tr>
<tr>
<td>Building Type-BTy</td>
<td>0.162</td>
</tr>
<tr>
<td>Floor’s material-F</td>
<td>0.185</td>
</tr>
<tr>
<td>Level-L</td>
<td>0.208</td>
</tr>
</tbody>
</table>

Table 6: General MANOVA for \( \log(C) \). Cut-off Limit for significance level < 30%

<table>
<thead>
<tr>
<th>Step</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.438</td>
<td>1.467</td>
<td>1.487</td>
</tr>
<tr>
<td>Wall Contact-Con</td>
<td>-0.070</td>
<td>-0.072</td>
<td>-0.073</td>
</tr>
<tr>
<td>T-Value</td>
<td>-1.34</td>
<td>-1.36</td>
<td>-1.40</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.162</td>
<td>0.168</td>
<td>0.162</td>
</tr>
<tr>
<td>Level-L</td>
<td>-0.948</td>
<td>-0.955</td>
<td></td>
</tr>
<tr>
<td>T-Value</td>
<td>-1.35</td>
<td>-1.52</td>
<td></td>
</tr>
<tr>
<td>P-Value</td>
<td>0.178</td>
<td>0.131</td>
<td></td>
</tr>
<tr>
<td>Wall’s material-W</td>
<td>-0.040</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-Value</td>
<td>-1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-Value</td>
<td>0.276</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.337</td>
<td>0.338</td>
<td>0.336</td>
</tr>
<tr>
<td>( R^2 ) adj</td>
<td>0.48</td>
<td>0.99</td>
<td>1.11</td>
</tr>
<tr>
<td>MCP</td>
<td>-1.2</td>
<td>-1.0</td>
<td>-0.2</td>
</tr>
</tbody>
</table>

Table 7: Stepwise Regression: \( \log (c) \) versus all factors. Value of Alpha-to-Enter was 0.5 and of Alpha-to-Remove 0.5. \( P \)-values represent calculated error probabilities and \( T \)-values, the corresponding values of t-student's test for the comparison of a \( P \)-value with \( q \). The constants of the linear fit at each step are shown in first row. Bold values represent the corresponding slopes of each factor at each step. MCP is the Mallows' \( C_p \).

Table 8: Unrotated factor loadings and communalities for 4 principal factors. The factor correlations exceeding the cut-off limit of 0.5 are marked in bold.
of contact (factor 4). A very important finding of Table 8, however, is that since the loadings of factor 1 to “Level- L”, “Building Type” and “Construction Year- Y” are negative in respect to the one of C, it is rational to accept that concentrations would increase as Level, Building Type and Year are decreased. According to Table 3, this implies that ground floor dwellings tend to present higher radon concentrations. This is rational since the lower the floor, the higher is the contribution of soil’s exhalation in indoor radon. Also detached houses tend to present higher concentrations, since other types offer pathways for radon’s interchange between dwellings in contact. Significant is also that aged dwellings, especially those of the previous century, presented higher radon concentrations. The latter finding is also reinforced by the positive loading of the “Wall’s material-W”, especially due to its rather high loading. Since higher wall values correspond to rock materials, it can be supported that higher concentrations are addressed in dwellings of the beginning of the twentieth century made of rocks. To some degree these results were supported by Table 7, since the dwelling’s “Level-L” and building “Wall’s material-W” were considered to be more significant compared to other factors.

Conclusion

Statistical analysis of data revealed that approximately 0.1% of the dwellings exhibited outlier radon concentrations. Noteworthy statistical correlations were detected between indoor radon concentration and the factors: “Level-L” and “Wall’s material-W”. Results of current work strengthen these considerations and provided weak evidence for the correlation of factors like “Building Type-Bty”, “Construction Year-Y” and “Floor’s material-F” with radon concentrations. Minor association was detected with the factor “Wall Contact-Con”. Significant differences were detected in results produced by some of the applied statistical methods.

Acknowledgement

This work was co-financed by Greece and the European Union under the European Social Fund NSRF 2007-2013 (Thales). Managing Authority: Greek Ministry of Education & Religious Affairs, Culture & Sports.

Financial Support-Funding

This work has been co-financed by Greece and the European Union, under the European Social Fund NSRF 2007-2013 (Thales). Managing Authority: Greek Ministry of Education and Religious Affairs, Culture and Sports.

References

31. Denman AR, Groves-Kirkby NP, Groves-Kirkby CJ, Crockett RGM, Phillips


