PATH MODELLING ANALYSIS OF NATURAL RADIOACTIVITY IN DRINKING WATER AND PUBLIC HEALTH IMPACTS

LISA MARTINS1*, ALCIDES PEREIRA2†, ALCINO OLIVEIRA3‡, LUÍS SANCHES FERNANDES3 & FERNANDO PACHECO1
1Department of Geology, University of Trás-os-Montes e Alto Douro, Portugal
2Department of Earth Sciences, University of Coimbra, Portugal
3Department of Civil Engineering, University of Trás-os-Montes and Alto Douro, Portugal

ABSTRACT

In 2015, the United Nations announced the Sustainable Development Goals, to achieve a better and more sustainable future for all by 2030 and promoting global attention about groundwater protection through Goal 6 (clean water and sanitation). The coefficient of determination of the structural model of total non-mineral water (n=72) is extremely low for radon contamination in groundwater ($R^2=0.083$). These results are not consistent because it is unlikely that all exogenous latent variables explain only less than 1% of radon contamination in groundwater. This discrepancy occurs because the total dataset is represented by several types of rocks with different radiological profiles. In this scenario, the complete dataset should not be singly applied in the structural equation modelling-partial least squares (SEM-PLS) model but somewhat separately in four lithological groups. From all studied lithologies, the structural model for Group V presents the highest coefficient of determination ($R^2=0.987$), mainly due to the application of hydraulic turnover time. This feature addition improves this contamination model and also confirms that the water of deeper circuits has a lower radon concentration due to its short half-life. Future studies tend to expand knowledge about other hydrogeological features that can effectively influence the radon transfer from rocks to drinking water, improving the contamination model of radon to develop a radon risk map. This study is crucial for policymakers because the radon contamination into the aquifer systems may raise concerns for the general public, requiring the continuous monitoring for the assessment of harmful health effects to humans caused by the radionuclide’s exposure.

Keywords: radon, drinking water, SEM-PLS, harmful effects.

1 INTRODUCTION

Natural radioactivity from geological materials, in particular, the radon gas ($^{222}\text{Rn}$) exposure, is considered a risk factor for human health. Regarding exposure to radon gas, health impacts are strictly related to its stable progeny. When these solid isotopes aggregate to the environmental aerosols, can be inhaled and deposited in the epithelial lung’s tissues, causing neoplasms. According to clean water goal from 2030 Agenda, several other targets were established with the purpose of water quality improvement, ensure the water-use efficiency, management, protection, and restoration of water resources. On the other hand, a study carried out by Martins et al. [1], alerts against the lack of the Portuguese Plan of Action for Radon or even the existence of any reference about the contamination of geogenic source in the Portuguese Water Plan. Radon contamination in drinking water is invisible to human sense, and as such, should be given particular attention when a high radioactive exposure in drinking water occurs. For a practical assessment of aquifer contamination risk, it is useful to establish a conceptual model that explains the production and transport of radon from rocks

*ORCID: https://orcid.org/0000-0002-7140-9036
†ORCID: https://orcid.org/0000-0002-7392-2255
‡ORCID: https://orcid.org/0000-0003-2465-2698
to drinking water [2]–[5]. For instance, when there is not mobilization of radon from rocks to water, it becomes crucial to understand which factors may be preventing this transfer process. Previous studies in fractured aquifer systems have reported that there is not a direct relationship between high-radon contents in rocks and the significant contamination of groundwater [2], [6], [7]. Then, how is it justified when in this comparison there is a high-radon production in rocks and a reduced radon concentration in groundwater? A study by Martins et al. [2] shows that a high-emanation or high-radon production in rocks may not always represent an efficient transport into the intergranular space and consequently an increased risk of groundwater contamination. For the assessment of the risk of groundwater contamination, it is crucial to conduct a survey about the connection between several factors that may increase the radon diffusion from rocks to the water such as, the weathering rate of U-bearing minerals under certain specific pH conditions [8]–[13], rock hydraulic diffusivity [14]–[16], and the emanation coefficient [17]. Several environmental studies used the structural equation modelling-partial least squares (SEM-PLS) models to establish common approaches in large flood management [18], ecological surveys [19], [20], and radiological impact in groundwater [2]. The application of cluster analysis with complex structural models was previously performed by Martins et al. [2] to study the effect of some factors on radon contamination in the aquifer system. This study aims to develop a simpler structural model, although much more robust than the previous one [2].

On the other hand, this study presents basic structural models considering only lithological grouping. The application of these structural models using the complete dataset, as a whole, was also ineffective as in Martins et al. [2]. This new approach is, therefore, crucial for the assessment of radon in aquifer systems, to prevent the radiological effects of radon exposure in drinking water consumers. The main purpose of this study is understanding the dynamics behavior of radon in aquifer systems using simple structural models that justify the presence of radioactive isotopes of geogenic source on drinking water. In addition, these radiological studies also alert policymakers about factors that enhance radiological contamination in water used for human consumption and facilitate the general public’s awareness of this environmental problem by warning on the effects of exposure to ionizing radiation.

2 METHODOLOGY

2.1 Study area

The study area is located in the north of Portugal (Fig. 1) with a contrasting topography composed by lowlands, craggy reliefs, and nearby hillslopes. The geology of the study area is extremely complex, being represented by several types of granites and metasediments. The two-mica granites were separated into two different groups: one group with high-radon production potential and another with low-radon production potential. The biotitic granites were divided from the remaining granites by their installation age. This division, based on the geochronology and radiological profile, become crucial for SEM-PLS analysis.

2.2 Conceptual model of radon contamination risk

For the development of structural equation models, the six steps depicted in Fig. 2 describes the methodology used. The study area was selected based on an excellent cartographic representation of the lithologies of the Portuguese territory (Step 1). As in Martins et al. [2], when structural models were used for the overall data (Step 2), the R² is not significant. Thus, all data collected were grouped into four clusters. Therefore, these clusters represent the main
geological units of the study area and its potential to produce radon (Step 3). After the building of the data matrix (Step 4), the structural models (Step 5) were developed according to the conceptual model depicted in Fig. 3.

Figure 1: Location, geological units, topography and drainage network of the study area. (a) Digital elevation/drainage network models; (b) Simplified geological map based on the geological map of Portugal, scale 1:500,000 and 1:50,000, available at LNEG [21] and Gomes et al. [22]. The tectonic structures were drawn from Dias et al. [23], being identified as ductile shear zones (B – Vigo-Régua) or fragile faults (II-Penacova-Régua-verín; III-Vilariça).

Figure 2: Methodology steps for the model of radon contamination risk.

Figure 3: Conceptual model of radon contamination risk in drinking water.
The geogenic features are essential to define the radiogenic profile of several lithologies for the entire region under study. Being a radioactive gas with origin in the decay series of uranium ($^{238}$U), forming from the disintegration of $^{226}$Ra, it is natural, therefore, that radon distribution is conditioned primarily to the radioactive element that initiates the chain. Uranium is widely distributed in mineralogical substrates, although it is more abundant in igneous rocks. In the subsoil, when the primary mineral-bearing uranium is subject to weathering, the chemical element is transported from the host mineral and precipitate into fractures and microfractures, on the mineral surface, or even be mobilized through groundwater flows. Depending on the specific circumstances of each study area, the radon contamination risk ($^{222}$Rn) may be influenced by the geogenic, climatic, topographic features, and hydrological properties. A primary constructor (black arrows) was developed to understand the interaction between each of these systems and to show the impact of these three crucial latent variables on the radon contamination in drinking water (Fig. 3). However, it is also essential to generate a secondary construct (dashed black arrows) to demonstrate the impact of climatic and topographic features in hydrological properties and geogenic features (Fig. 3).

For the geogenic features (Fig. 3), five measured variables that describe the radon production capacity from rocks to interconnected rock pores were used. Pereira et al. [24] widely describe the radon production potential methodology. The isotopes K$_2$O, U, and Th were analyzed by gamma spectrometry with thallium-activated sodium iodide detector (NaI(Tl)) and 3-inch diameter, branded ORTEC. A total of 72 springs and boreholes used for water supply was sampled for radon analysis and hydrological properties. Martins et al. [2] give a more detailed description of the gamma spectrometry methodology and radon analysis used. The hydrological properties define the physical-chemical conditions of drinking-water (electric conductivity, temperature, and pH) and the time of water-rock interaction (hydraulic turnover time). However, due to financial difficulties was not at all possible to evaluate the hydraulic turnover time for all the collected sites. A combination of streamflow discharge rate ($Q$, $Q$, m$^3$/s) and $a_1$ and $a_3$ constants extracted from a scatter plot of ln($\Delta Q/\Delta t$) versus ln($Q$) was developed to estimate the hydraulic turnover time (t,s) as proposed by Pacheco [25]:

$$t = \frac{1.98}{Q \sqrt{a_1a_3}}$$

(1)

The $a_1$ and $a_3$ constants were taken by intercepting y values with straight lines with slope 1 and 3, representing the lower envelope to the scatter points. The streamflow discharge rate and $a_1$, $a_3$ constants, only used for Group V structural model, were removed from Pacheco and Van der Weijden [15]. The climatic and topographic features may also contribute to radon dilution, being represented by two variables: the altitude and precipitation.

2.3 SEM-PLS

The Structural Equation Modelling-Partial Least Squares (SEM-PLS) technique was applied using the SmartPLS. SmartPLS is a prominent software application, gaining popularity since its launch in 2005 [26]. As well as being visually appealing software and freely available to research community across the globe, it has also a simple user-interface with an advanced report on the analytical depth of SEM-PLS.

After the compilation of all dataset, for each lithological group, a matrix was built, where columns are variables and each row is a water collection point. All variables were used respecting the multicollinearity rules (Pearson correlation coefficient <0.8 and VIF <5) [27].
In SEM-PLS the latent variables (LV) represented by the blue circles of Fig. 4 are considered the operators, according to the conceptual model of radon contamination risk in drinking water. Otherwise, the measured variables (MV) represented in Fig. 4 by yellow squares are the collected variables that explain the latent variables. The SmartPLS was used to build four formative models. In the interactive procedure, the algorithm attributes weights (w) for MVs and path coefficients (pc) for LVs, to achieve for each LV the highest determination coefficient ($R^2$). This algorithm also calculates a measured score for LVs, based on the MVs (eqn 2). On the other hand, for LVs that are composed by other LVs, the Smart PLS calculated a predicted score (eqn 3)

$$LV_j = \sum_{i=1}^{n}(MV_i \times W_i),$$  \hspace{1cm} (2)

$$LV_{p,c} = LV_A \times pc_{ac} + LV_B \times pc_{bc}.$$  \hspace{1cm} (3)

![Diagrammatic representation of path models for different lithological groups that are outcrops in the studied region. (a) Group I; (b) Group III; (c) Group IV; and (d) Group V. $^{222}$Rn-radon contamination in drinking water; PRn-Radon production potential in rocks.](image)

3 RESULTS AND DISCUSSION

The determination coefficient of the structural model for the total dataset (n=72) was insignificant ($R^2$=0.083). For an independent check, this low determination coefficient shows that all exogenous latent variables (geogenic, climatic and topographic features, and hydrological properties) explain only less than 1% of radon contamination in groundwater.
When the SEM-PLS is simultaneously applied to all lithologies, this same discrepancy is verified, as in Martins et al. [2]. This weak path model occurs because the radiological profile of all lithologies is quite distinct. In this scenario, the total dataset was grouped according to the geochronology and radiological profile of each lithological type. To develop more reliable path models, a multicollinearity test was carried out among the several measured variables. When R-squared for a formative factor is less than 0.80, the multicollinearity is not a problem by the criteria of VIF>5 (Fig. 4(a) and (c); Table 1 [28]). The results of this reliable model are depicted in Table 1, which shows that there is not a multicollinearity problem because the variance inflation factors (VIF) slightly exceeds 5.

Table 1: Inner VIF values for the studied models.

<table>
<thead>
<tr>
<th></th>
<th>222Rn</th>
<th>Geogenic features</th>
<th>Hydrological properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climatic and topographic features</td>
<td>5.389</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Geogenic features</td>
<td>2.723</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrological properties</td>
<td>3.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group III</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climatic and topographic features</td>
<td>2.080</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Geogenic features</td>
<td>2.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrological properties</td>
<td>1.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group IV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climatic and topographic features</td>
<td>5.509</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Geogenic features</td>
<td>2.503</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrological properties</td>
<td>3.644</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group V</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climatic and topographic features</td>
<td>2.117</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Geogenic features</td>
<td>2.869</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrological properties</td>
<td>4.495</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the first group composed by metasediments, the structural model for radon contamination in groundwater is moderate (R²=0.589; Fig. 4(a)). It means, therefore, that the combination of three latent variables is responsible for almost 60% of radon concentration in drinking water. A more comprehensive analysis of this structural model shows that the collection sites with the highest concentration of radon are located in areas with higher altitude (w×pc=0.793×0.499; Fig. 4(a)). However, the precipitation has an inverse effect, because the high precipitation causes a lower radon concentration in the water (w×pc=-0.352×0.499). This model clearly shows the dilutive behavior of rainwater in radon concentration.

On the other hand, these climatic and topographic features harm the geogenic factors and the hydrological properties (pc=-0.786 and pc=-0.810, respectively). In general, the geogenic construct plays the main role in radon contamination for the metasedimentary lithologies (Fig. 4(a)). The structural model of two-mica granites with lower radon production (PRn) presents a moderate coefficient of determination (R²=0.455; Fig. 4(b)). It means, therefore, that the exogenous latent variables explain almost 50% of this radon contamination in water used for human consumption. The thorium (Th) and the radon production potential (PRn) have a substantial impact on the geogenic features (w=1.109 and w=1.089, respectively). In the hydrological properties, the pH is highlighted from the other features (w=0.847; Fig. 4(b)). Despite the R² for this structural model becoming insignificant (R²=0.080), it has a significant negative impact on radon contamination (pc=-0.682; Fig. 4(b)). It should also
be noted that the geogenic, climatic, and topographic features almost do not influence the radon contamination in the water (pc=-0.024 and pc=-0.051; Fig. 4(b)).

Unlike the previous structural models, the algorithm for two-mica granites with high-radon production computed a weak coefficient of determination ($R^2=0.243$). It is an acceptable $R^2$, although it is difficult to provide for this work other measured variables that improve this structural model. This weak-developed path model explains only 24% of total radon contamination in groundwater. Nonetheless, it is important to highlight the great influence of hydrological properties on the radiogenic contamination of drinking water (pc=0.824; Fig. 4(c)). The $K_2O$, fracturing density, altitude and electric conductivity was the key target constructs for latent variables geogenic, climatic and topographic features, and hydrological properties, respectively (w=1.106; w=0.997; w=0.600; w=0.797; Fig. 4(c)). The model path of group V was effectively the most robust of all because it represents the major contribution of geogenic, climatic, and topographic features ($K^2=0.987$) into radon transfer for water (Fig. 4(d)). In this scenario, the radon production potential acquires a negative effect on geogenic features and consequently to the radon contamination in drinking water ($w\times pc=-1.258\times1.055$). On the other hand, for springs with extended hydraulic turnover time, there is a reduced transfer of radon ($w\times pc=0.141\times(-0.193)$; Fig. 4(d)). It should also be noted that a simulation was performed without the inclusion of the hydraulic turnover time. The results clearly show a decrease in the coefficient of determination ($R^2=0.840$). This insertion was representing, therefore, a high $R^2$ in the radon contamination, being explained by all of the exogenous constructs linked to it (Fig. 4(d)). $R^2$ is calculated as the squared correlation between specific endogenous constructs actual and predicted values. According to these results, it would be beneficial to apply the flow rate and hydraulic turnover time in these structural models with reduced $R^2$. Currently, due to lack of funds, it is not possible to do this type of analysis at all. In future studies, it is expected that these works will be accomplished, among others.

A radar mesh was developed to present the summary of all constructs generated by SmartPLS software. In Fig. 5(a) the summary constructs comprise three segments used for primary constructs (non-shaded area) and two other sections of a secondary construct (shaded area). An overview of these results shows that in the primary construct, the model path of Groups I and IV almost overlap for all coefficient paths. It means, therefore, that these two groups overlap in the primary construct of the radon contamination in water because the radon transfer to drinking water was affected equally by the same latent variables. Nevertheless, the path model of group V provides a functional overlap between predicted and measured scores ($R^2=0.987$), while the group I produce a predicted value that is almost half of measured scores ($R^2=0.589$; Fig. 5(b)). In these circumstances, for a realistic comparison between these groups, other measured variables should be included in the constructs of Group I to improve the model.

In the primary constructs of Group III, there are more reduced path coefficients than in all other groups. Moreover, in Group IV, the hydrological properties have a more significant impact on radon contamination in drinking water than in all other groups (Fig. 5(a)). For an independent check, the geogenic features have a more significant impact on groups I and V, mainly due to the high positive effects of $U$ ($w=1.227$) in group V, and $K_2O$ for Group I ($w=0.665$; Figs. 4(a), (b); Fig. 5(b)). In general, in such different geological contexts, the effect of all these measured variables on radon contamination in water may be similar or different, i.e., there is no single rule that explains radon contamination in water. This complex system that explains radon contamination in drinking water can be applied to other areas, although with other parameters applicable to the topographic, climatic, geological, and hydrogeological contexts, among others in the study area.
Figure 5: Summary of all path models computed by algorithm PLS-SEM. (a) Path coefficients with positive and negative effects for all studied lithological groups; and (b) Coefficient of determination for latent variables. $^{222}\text{Rn}$ is the radon contamination in water; GF is the climatic and topographic effect in the geogenic features; HP is the climatic and topographic effect in hydrological properties.

4 CONCLUSIONS
The entire radon transfer system from rocks to water and the processes involved are extremely complex. The application of algorithm PLS-SEM to the study area proved to be efficient and supports the knowledge about the features can influence the radon contamination in drinking water. The path model of several lithological groups under study indicates the well-predicted variables of the construct. Thus, in the path model of Group V, almost 100% of the primary construct is explained by the indicators, representing a higher level of predictive accuracy. On the other hand, for a secondary construct of groups I and IV, the climatic and topographic features present a noticeable negative effect in geogenic and hydrological processes. Therefore, when using this new methodology of path modeling in the analysis of radon contamination in water, we can anticipate that in similar geological contexts, there may be different concentrations of radon in water and vice versa. This means that the geological context alone is not preponderant for radiological contamination in water. Therefore, depending on the geological setting, the several processes of water-rock interaction, the physicochemical properties of the water, the climatic and topographic conditions together with the hydraulic turnover time have different performance in the radon transfer to the water. This work clearly shows that this innovative method should not be applied generically in several regions with similar geological context, as these areas may have different risks of contamination and consequent harmful radiological exposure to human health. Thus, this advanced tool can be useful to explain the contamination risk of regional scope in water used for human consumption. This study also alerts to the importance of hydraulic turnover time in the transfer of radon to drinking water. This concern is more pertinent because the Sustainable Development Goals of 2030 Agenda alert the policy-makers to the implementation of mitigation measures in anomalous springs, thus promoting safe water intake within the pre-established deadline.

ACKNOWLEDGEMENTS
This research was funded by the INTERACT project “Integrated Research in Environment, Agro-Chain and Technology”, no. NORTE-01-0145-FEDER-000017, in its line of research entitled BEST, co-financed by the European Regional Development Fund (ERDF) through
NORTE 2020 (North Regional Operational Program 2014/2020). For authors integrated into the CITAB research centre, it was further financed by the FEDER/COMPETE/POCI – Operational Competitiveness and Internationalization Programme, under Project POCI-01-0145-FEDER-006958, and by National Funds of FCT – Portuguese Foundation for Science and Technology, under the project UID/AGR/04033/2019. For the author integrated into the CQVR, the research was additionally supported by National Funds of FCT – Portuguese Foundation for Science and Technology, under the project UID/QUI/00616/2019. The authors would like to thank Álvaro Miranda from Trás-os-Montes and Alto Douro University and Rita Lamas for their contribution in the sampling campaigns and the laboratory analyses. The authors would also like to thank the Laboratory of Natural Radioactivity of the University of Coimbra, in the person of Alcides Pereira, for the financial support of radiological analysis in drinking water.

REFERENCES


