# RADON MAPPING IN ABRUZZO, ITALY

Sergio Palermi
*Environment Protection Regional Agency of Abruzzo (ARTA), Pescara, Italy, s.palermi@artaabruzzo.it.*
Antonio Pasculli
*Department of Geotechnology, University of Chieti, Italy, a.pasculli@unich.it*

RÉSUMÉ
La région des Abruzzes est située au centre de la péninsule italienne. L'Agence Régionale de Protection de l'Environnement (ARTA) y conduit depuis deux ans des études dont le but est d'estimer la distribution des valeurs de concentration en radon dans les habitations ainsi que la probabilité de rencontrer des zones dont les teneurs en radon seraient particulièrement élevées. Le radon est un gaz radioactif naturel exhalé du sol qui pénètre dans les habitations où il peut atteindre, selon le type de sol et les caractéristiques de construction, des concentrations particulièrement élevées et présenter ainsi un risque pour la santé. Cet article présente quelques résultats préliminaires relatifs à la distribution géographique du radon dans les Abruzzes ainsi qu'une carte de risques établie à partir de plus de 1600 mesures. Pour produire cette carte, une division du territoire étudié a été réalisée au moyen d'une grille géographique constituée de cellules carrées de 6 km de côté. Au total, une surface d'à peu près 10000 km$^2$ a été analysée à l'exclusion des régions montagneuses les plus inaccessibles qui sont aussi inhabitées. Chaque cellule a été ultérieurement subdivisée en 9 carrés de 2 km de côté. Pour chaque site étudié, les coordonnées géographiques, les caractéristiques de construction du logement ainsi que les habitudes des habitants ont été enregistrées. L'étude statistique des mesures a ainsi pu être précédée par une normalisation des mesures à partir des paramètres enregistrés. Les données utilisées en fin de compte représentent des mesures annuelles prises au rez-de-chaussée et les estimations des valeurs moyennes géométriques et les variances associées ont été obtenues en utilisant une approche Bayesienne. La carte du radon a finalement été produite en appliquant des techniques géostatistiques.

ABSTRACT
Abruzzo is a region located in the middle part of Italy. The Environment Protection Regional Agency (ARTA) since a couple of years ago has been carrying out studies aimed to assess the geographic distribution of indoor radon concentrations and the possible existence of high concentrations areas (radon prone areas). Radon in buildings can reach relatively high concentrations, depending on a number of factors (ground and housing characteristics, ventilation rates and so on). In this paper, the preliminary results regarding the geographical variations of indoor radon concentration levels are discussed and a preliminary radon map is proposed on the basis of more than 1600 measures. The selected approach has been 'the area sampling technique', through the territory subdivision in 6 km grid squares, covering almost the whole region, about 10000 km$^2$, excluding uninhabited mountain areas. Each grid square has been further subdivided into nine smaller 2 km grid squares. For each selected measurement site, a document including all those information which could have conditioned the measuring process like coordinates, building typologies, the living habits of the residents and so has been compiled. The statistical analysis of the radon data has been preceded by a normalization with respect to the ground floor values which has then been carried out in relation of both the 6 km grid and municipality areas. In both cases, geometric means and geometric standard deviations of the normalized concentrations have been evaluated applying a Bayesian approach (empirical method). Finally usual geostatistical mapping techniques have been applied to obtain a smoothed radon map.

## 1 INTRODUCTION

Radon-222 is a noble gas arising naturally from decay of Uranium-238 present in the earth's crust. Its half life of 3.82 days allows diffusion through soil into atmosphere before decaying by emission of an alpha particle, giving origin to a series of short lived radioactive progeny. Inhalation of this progeny exposes cells of bronchial epithelium to alpha irradiation (from isotopes of polonium), a possible cause of lung cancer (Darby *et al.*, 2005).

In Italy, the natural radiation hazard regulations are related only to professional exposed population and nothing exists for the safeguard of the population deriving from exposure to radon in dwelling houses. At European level, the Recommendation 90/143/Euratom (1990) is the reference for indoor radon exposure; for the existing buildings a reference level of annual average radon concentration, beyond which it is recommended to undertake remedial actions, is set at 400 Bq/m$^3$, while for future constructions the concentration should not exceed 200 Bq/m$^3$. In spite of the absence of specific regulations regarding surveys and monitoring strategies, a number of monitoring campaigns were performed in Italy in the last two decades in order to determine the radon levels in private and public places (mainly schools and kindergartens). A nationwide representative survey, carried out between the eighties and nineties and coordinated by the Italian National Institute of Health (Bochicchio *et al.*, 1996, 2005), has been of particular importance. All 21 Italian regions were involved. In a random sample of more than 5000 dwellings the radon concentration was measured over two consecutive semesters. The national average value was 70 Bq/m$^3$, while in Abruzzo, a region located in middle Italy, an average value of 60 Bq/m$^3$ was estimated from a 103 dwellings sample. Kindergartens were the subject of another large

*In : J. Locat, D. Perret, D. Turmel, D. Demers et S. Leroueil, (2008).*
*Comptes rendus de la 4e Conférence canadienne sur les géorisques: des causes à la gestion.*
*Proceedings of the 4th Canadian Conference on Geohazards : From Causes to Management. Presse de l'Université Laval, Québec, 594 p.*

survey carried out in nineties in the whole region (nearly 500 two consecutive semesters' measurements); the average turned out to be 66 Bq/m$^3$.

On the basis of the results obtained from these past surveys and taking into account the complex geomorphologic framework of the region, a new monitoring campaign has been carried out in the last few years, in order to get a more comprehensive picture of the territorial distribution of indoor radon concentration, and to identify possible radon prone areas.

In this paper, we present the preliminary results of statistical analyses carried out on the whole set of data, i.e. the ~600 measurements coming from the two past surveys and the ~1000 measurement from the present one. Information on the geomorphologic characteristics will also be used.

## 2 RADON AND GEOLOGY

The primary source of indoor radon is due to the ground underlying the building. As a matter of fact, geology and soil properties have a very important impact on radon production and transport (by convection and diffusion in the pore space and fractures) towards the surface and, thus, on its indoor concentration. On a geochemical point of view, the radon emanation depends on the uranium concentration (and therefore of radium) in the rock substrate. Substrate rock genesis has therefore an important impact on this kind of phenomenon. In Tables 1 and 2 specific activities for different types of rock are reported (Verdelocco *et al.*, 2000).

Table 1: Average specific Uranium and Thorium activity (Bq/kg) in igneous rocks

| Igneous rocks | $U^{238}$ (Bq/kg) | $Th^{232}$ (Bq/kg) |
|---|---|---|
| Acid (Granite) | 59 | 81 |
| Intermediate (Diorite) | 23 | 32 |
| Mafic (Basalt) | 11 | 11 |
| Ultrabasic (Durite) | 0.4 | 24 |

Table 2: Average specific Uranium and Thorium activity (Bq/kg) in sedimentary rocks

| Sedimentary rocks | $U^{238}$ (Bq/kg) | $Th^{232}$ (Bq/kg) |
|---|---|---|
| Limestone | 27 | 7 |
| Carbonates | 26 | 8 |
| Sandstone | 18 | 11 |
| Shales | 44 | 44 |
| Upper crust | 34 | 45 |
| Soil | 25 | 25 |

Permeability and thickness of superficial deposits also influence radon exhalation from the soil. Other key factors are considered to be the groundwater circulation, soil moisture and meteorological variables (temperature, atmospheric pressure, precipitations), which can generate important daily and seasonal variations in the radon flows towards the surface.

### 2.1 Abruzzo geomorphologic framework

Abruzzo's Apennines chain is the result of sea paleogeographic areas raising. Sedimentary processes, which predisposed the materials for the mountain chain formation, are related with two depositional systems: the carbonate platform/basin and the silico clastic foredeep (Crescenti *et al.*, 2003). In a morphologic point of view, the Abruzzo region is characterized by a succession of wide orographic homogenous areas. Going westward from the coast, one will find in progression hilly, piedmont and mountain areas. Hilly zones are located on Plio-Pleistocene sea deposits (clay, sands, conglomerates). The shoreline is constituted by short strips of high coast in the southern part of the region, against a prevalence of low and sandy coasts, affected by erosion processes. Piedmont area is characterized by reliefs whose heights can reach 1000 meters, separated by deep valleys, carved inside Pliocene and Messinian turbidity sediments, while the altitude of the mountain crests ranges between 1000-1500 meters up to about 2900 meters. Abruzzo ridges are characterized by limestone alternating with marls, dated within Upper Trias and Medium Miocene, while Laga Mountains (in the northern side) are formed by sandstone strata and banks intercropping with thin Messinian pelitic layers. Last but not least, the particular widespread occurrence of carbonates and evaporites lithotypes determined a remarkable development of karst phenomena.

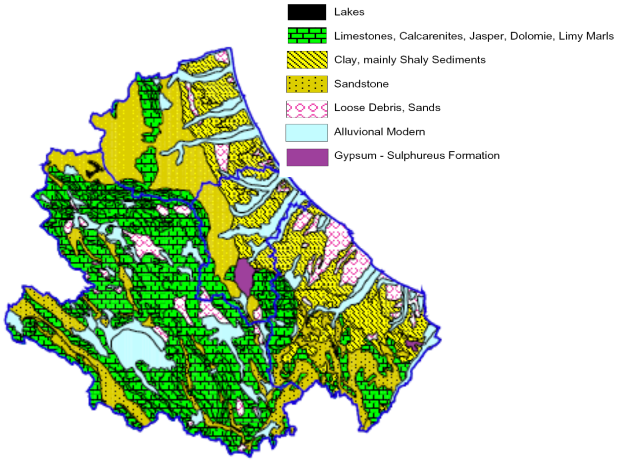

Figure 1: Lithological map of Abruzzo

The lithological characteristics of the regional territory (Figure 1), briefly mentioned above, do not seem to indicate an important indoor radon potential. On the other hand in Abruzzo, just like in the rest of Italy, beyond the lithology (moreover characterized by a high degree of heterogeneity and complexity) it is necessary to consider various occurrences like joints, faults, seismicity, volcanism, karsism and geo-thermalism (ANPA, 2000).

## 3 INDOOR RADON MEASUREMENTS

### 3.1 Measurement devices

In the current survey, radon concentrations were measured using dosimeters with CR-39 solid state track detectors, which were chemically etched after their one year exposure in dwellings and then read out by an image analyser.

### 3.2 Sampling strategy

Sampling strategies whose purpose is to identify radon prone areas often, although not necessarily, meet different criteria with respect to those aimed to estimate the population exposure to radon. Roughly speaking, the first approach meets the requirement to (more or less homogeneously) cover the entire territory (so the sampling can be carried out according to geographical criteria) while, in the second case, considerations about population distribution are required. However, with a proper design the two different approaches could be combined in order to obtain results that are both representative for the population and informative of the territorial distribution. The first among the two surveys previously cited was designed just following the latter approach, with a two-stage stratified sampling scheme whose consequence, at the regional level, was a clusterization of sampled dwellings in a few random selected municipalities (only 7 out of 305). The second survey, on the other hand, had the peculiarity to involve almost all the kindergartens in Abruzzo, with a high spreading of sampled locations over the territory. Up to now, analyzing international approaches to radon mapping (see e.g. Dubois 2005), it emerges that preference was given to monitoring strategies planned on a geographic basis. The new survey has been set up along this direction. The planning of the survey on a geographic basis stresses the problem of the territory subdivision, that is, the choice of the reference geographic unit for the sampling. Various options are available (Miles, 1998). Regarding our survey, both regular geometric meshing and administrative subdivision in municipalities have been tested. The structure of the grid is the following: 6 km side square meshes were defined and further subdivised into 2 km side squares (9 elements for each bigger square). Within each sub element (2 km side square) at least one site, (mostly dwellings, with a few workplaces, offices, shops and so on) was randomly selected. The detectors have been placed in each site, mostly in bedrooms or living rooms at ground floor.

All information (geographic coordinates, altitude, building features, floor level of the monitored room, underlying lithological characteristics and so on) related to each monitored site was collected and stored in a database (combined with a GIS software), along with data obtained from previous surveys.

### 3.3 Data analysis: normalization procedures

The set of radon concentrations available for analysis includes also the data collected in the course of the three cited campaigns. We decided to disregard, for the moment, the possible issue of homogeneity of the different sampling strategies. Almost all the measurements came from a one year exposure, but not all, thus, at first, raw data have been normalized to a full year. Furthermore, standardization to a virtual ground level condition has been performed, aiming at outlining radon emission characteristics of soil, the main source of indoor radon, reducing distortion caused by different soil-to-dwelling distances. Regarding the first procedure, analyzing the data of past measurement surveys, carried out along two consecutive semesters, the following parameter has been estimated:

$$S = \frac{\overline{R}_{winter}}{\overline{R}_{summer}} \qquad [1]$$

where $\overline{R}_{winter}$ and $\overline{R}_{summer}$ are the average values of measurements carried out respectively during autumn-winter and spring-summer semesters. The analysis has been carried out separately for the Appennines-subappennines area (practically, the entire province of L'Aquila and municipalities of the others three provinces located at an altitude above 500 m a.s.l.) and for hilly-coastal area, characterized by different climatic patterns. The following were the calculated values: $S_{App} = 1.6$, $S_{Coast} = 1.3$. The $S$ parameter shows, in some way, the influence of climate on indoor radon concentration levels as well as of living habits of the inhabitants. Starting from the $S$ parameter knowledge it is possible to standardize the measurements to an annual medium value $R_y$ whichever value $R_m$ of radon concentration, related to measurement periods shorter than 12 months, applying the following equation:

$$R_y = R_m \frac{(n_s + n_w)(S+1)}{2(n_w S + n_s)} \qquad [2]$$

where $n_w$ and $n_s$ are the months number, during which the measurement has been carried out, within, respectively, autumn-winter (w) and spring-summer (s) semesters.

The procedure of the data standardization to the ground level has requested the evaluation of the conversion coefficients $K_{i \to 0}$ of the concentrations measured at the i$^{th}$ floor respect to ground floor (0). Since no meaningful differences between the provinces have been found, the same estimate has been provided and applied for the entire territory:

$$K_{i \to 0} = \frac{GM^{(0)}}{GM^{(i)}} \qquad [3]$$

Where $GM^{(i)}$ and $GM^{(0)}$ are the geometric means of all the measured values related, respectively, to the i$^{th}$ floor and ground floor. The estimated values are (-1 means underground floor): $K_{1 \to 0} = 1.6$, $K_{-1 \to 0} = 0.8$, $K_{2 \to 0} = 1.8$, where $K_{2 \to 0}$ has been calculated including also very few data related to floors upper than the second.

In Table 3 we report the summary statistics of the normalized radon concentrations $R$, calculated from original data $R_y^{(i)}$ according to the equation:

$$R = K_{i \to 0} R_y^{(i)} \quad [4]$$

Several features are noticeable: L'Aquila, an inland and mountainous province, shows the highest mean values; the geometric standard deviations appear clearly positively correlated to the geometric means. The distribution of data shows approximately a log-normal shape. The percentages of the data exceeding reference levels of 200 Bq/m$^3$ and 400 Bq/m$^3$ are also reported.

Table 3: Summary statistics of normalized radon concentrations (Bq/m$^3$) for each province and the whole region. n= sample size, AM=arithmetic mean, GM=geometric mean, SD= standard deviation, GSD=geometric standard deviation.

| Province | Aquila | Teramo | Pescara | Chieti | Abruzzo |
|---|---|---|---|---|---|
| n | 397 | 335 | 419 | 496 | 1647 |
| AM | 86.8 | 57.2 | 47.1 | 61.5 | 63.1 |
| SD | 98.8 | 57.9 | 51.3 | 87.3 | 78.7 |
| GM | 60.0 | 42.6 | 35.5 | 43.0 | 44.3 |
| GSD | 2.3 | 2.1 | 2.0 | 2.1 | 2.2 |
| Max | 1095 | 510 | 513 | 1114 | 1114 |
| Min | 9. | 6. | 7. | 4. | 4. |
| median | 55.7 | 41.7 | 33.9 | 38.5 | 41.7 |
| R>200 | 9.3% | 3.3% | 1.9% | 2.6% | 4.2% |
| R>400 | 1.3% | 0.3% | 0.5% | 1.0% | 0.8% |

## 4 PROBABILITY DISTRIBUTION AND BAYESIAN STATISTICAL INFERENCE

On the basis of the available data set, the identification of possible radon prone areas requires the application of statistical inference techniques that opportunely take into account of a low number of sample data in many geographic units (regular meshes or municipalities areas). GM and GSD were computed for each unit, from sample data whose size is very different from case to case; as a consequence, uncertainty affecting the estimated GMs and GSDs can be very high in those units where very few data are available (even only one, preventing GSD to be computed). Therefore, a well tested Bayesian inference method was selected (Price *et al.* 1996, Giannardi *et al.* 2001, Gelman *et al.*, 2004) in order to improve the estimate of GM and GSD for each unit taking into account the measurements in the other units.

Let $R_{ij}$ be the j$^{th}$ measure of radon activity concentration (after application of normalization procedures) within the i$^{th}$ unit (geographic mesh or municipality area). The application of Bayesian inference is based on the following assumptions: the $R_{ij}$ values follow a log-normal distribution, that is, their natural logarithms $\ln(R_{ij})$ are normally distributed within each geographic unit, the mean and variance of the distribution being, respectively, $\ln(GM_j)$ and $k^2$ (assumed the same for all the units). We also assume that the observed GMi are log-normally distributed, thus ln(GM) distribution is normal, with mean and variance $\mu$, $\sigma^2$ to be determined. The log-normality hypothesis of the distribution of the indoor radon concentrations within single territorial units (and within the entire territorial dominion of study as well) is credited by the greater part of the scientific literature (apart from previous references, see Miles, 1998). Actually, experimental data match the log-normal distribution only approximately without, however, fitting it completely. Analyzing the data on a provincial basis, we have verified that log-normality fitting is satisfactory only for the province of Teramo (according to D'Agostino normality test, see Armitage and Colton, 2005) which shows, not as a case, the lowest skewness and kurtosis values (see tab. 3). The Bayesian estimate of "the true" value of the variable $\ln(GM_j)$ comes down from the application of Bayes' theorem:

$$p[\ln(GM_i) \mid \ln(R_{ij})] = \frac{p[\ln(GM_i)] \cdot p[\ln(R_{ij}) \mid \ln(GM_i)]}{p[\ln(R_{ij})]} \quad [5]$$

where the left term is the posterior probability, that is the conditional probability that the expected value of log-transformed radon concentrations, within the i$^{th}$ unit, is $\ln(GM_j)$, given the set of experimental data $\ln(R_{ij})$; in the right-hand side, $p[\ln(GM_j)]=N(\mu, \sigma^2)$ is an informative (normal) prior distribution and the remaining expression is the normalized likelihood. The development of the calculations starting from [5] (Gelman *et al.*, 2004) leads to a particularly simple result for the estimate for $\ln(GM_j)$ (known as Bayesian point or empirical estimate):

$$\ln(GM_i^{eval}) = \frac{\dfrac{\mu}{\sigma^2} + \dfrac{n_i}{k^2}\ln(GM_i^{obs})}{\dfrac{1}{\sigma^2} + \dfrac{n_i}{k^2}} \quad [6]$$

whose variance $V_i^2$ is expressible as:

$$V_i^2 = \frac{1}{\dfrac{1}{\sigma^2} + \dfrac{n_i}{k^2}} \quad [7]$$

In equations [6] and [7] $n_i$ is the sample size for the geographic unit $i$ from which the observed mean $\ln(GM_i^{obs})$ is estimated. As it turns out clearly from [6], which is actually a weighted mean between the observed local mean and $\mu$, the higher the number of measures $n_i$, the closer the estimated value $\ln(GM_i^{eval})$ to the observed value $\ln(GM_i^{obs})$. The parameters $\sigma^2$ and $k^2$ are the variance components, respectively, between geographic units and within units, and can be determined from an analysis of variance (ANOVA). The selected ANOVA method is based on a hierarchical nested model with random factor effects (Kutner *et al.* 2005), in which the geographic units are grouped in macro-unit "m" belonging to an upper level:

$$\ln(R_{mij}) = \mu + \alpha_m + \beta_{i(m)} + \varepsilon_{j(im)} \qquad [8]$$

where $\mu$ is estimated from the global mean of the data, $\alpha_m = \mu_m - \mu$ is the difference between the mean of data within macro-unit $m$ and the global mean, $\beta_{i(m)} = \ln(GM_i) - \mu_m$ is the difference between the mean in the geographic unit $i$ and the mean on macro-unit $m$ (within which the $i^{th}$ unit is nested) and finally $\varepsilon_{j(im)} = \ln(R_{mij}) - \ln(GM_i)$ is the residual difference between each observed value and the mean of the $i^{th}$ unit to which the value belongs. Formally, the variables $\alpha$, $\beta$, $\varepsilon$, follow normal distributions with zero mean and variances to be determined:

$$\alpha \cong N(0,\sigma_\alpha^2)\,;\ \beta \cong N(0,\sigma_\beta^2)\,;\ \varepsilon \cong N(0,\kappa^2) \qquad [9]$$

For which that concerns the choice of the geographic macro-units, after having explored the possibility to use the four provincial areas, we have adopted a subdivision of the regional territory in two zones with homogenous geomorphologic and climatic characters: the Appennines-subappennines area (or montane-piedmont) and the hilly-coastal area (Figure 2). In Table 4 we report the summary statistics of the radon concentrations (before and after normalization to the ground floor) separately for each macro-unit.

Table 4: Summary statistics related to the macro-units. $R$ are the annual averaged radon activity normalized to the ground floor; $R_y$ are the annual averaged measured value before normalization to ground floor.

| | | Hill-coast area | | | Apennnines-subapennines area | |
|---|---|---|---|---|---|---|
| Statistic | $ln(R)$ | $R\ (Bq/m^3)$ | $R_y\ (Bq/m^3)$ | $ln(R)$ | $R\ (Bq/m^3)$ | $R_y\ (Bq/m^3)$ |
| Nº of sites | 889 | 889 | 889 | 758 | 758 | 758 |
| Min | 1.41 | 4. | 4. | 1.93 | 7. | 5. |
| Max | 6.01 | 408 | 408 | 7.02 | 1114. | 1114. |
| Median | 3.57 | 35.4 | 32.2 | 3.90 | 49.3 | 43.9 |
| GM | | 34.5 | | | 53.8 | |
| AM | 3.62 | 47.8 | 43.9 | 3.99 | 81.0 | 75.3 |
| $\sigma/\sqrt{N}$ | 0.02 | 1.4 | 1.3 | 0.03 | 3.8 | 3.8 |
| $\sigma$ | 0.67 | 41.1 | 39.7 | 0.84 | 104.4 | 104.8 |
| $\sigma^2$ | 0.45 | 1693. | 1573. | 0.70 | 10898. | 10992. |
| GSD | | 2.0 | | | 2.3 | |

The analysis of variance according to the model [8] has been carried out two times on the set of data, with two different selection of the geographic unit: a) 6 km grid squares; b) municipality territories (in both cases, the subdivision in the territorial macro-units remaining valid).

In both ANOVA applications, the variance ratio F test shows that $\sigma_\alpha^2$ e $\sigma_\beta^2$ are not vanishing (p<0,001 related to null hypothesis), indicating a significant grouping of the variable (log-transformed ground floor indoor radon concentration) by areal units. The results of this kind of approach for a) and b) subdivisions, are

a) $k^2 = 0.512$ ; $\sigma_\alpha^2 = 0.065$ ; $\sigma_\beta^2 = 0.052$

b) $k^2 = 0.488$ ; $\sigma_\alpha^2 = 0.064$ ; $\sigma_\beta^2 = 0.079$

from the lower value of $k^2$ (unexplained or residual variance) occurred in the case b), it follows that the

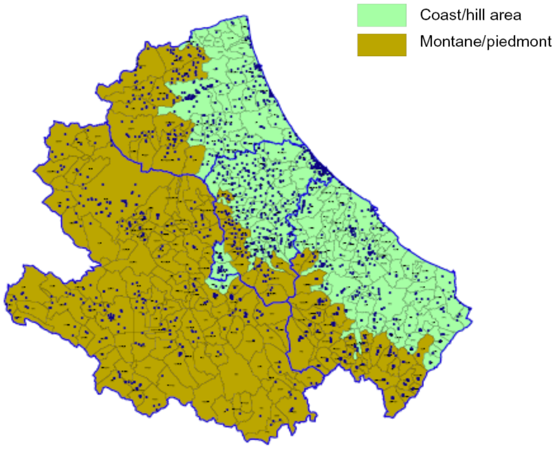

Figure 2: Abruzzo subdivision in two macro units: montane-piedmont area and hilly-coastal area.

Table 5: Analysis of variance results by model [10], for a) and b) cases for each macro-units subdivision

| | a): 6x6 km cells | | | b): municipalities | | |
|---|---|---|---|---|---|---|
| macro unit | $\mu$ | $k^2$ | $\sigma^2$ | $\mu$ | $k^2$ | $\sigma^2$ |
| hilly-coast area | 3.624 | 0.403 | 0.046 | 3.615 | 0.388 | 0.054 |
| Appennines-subappennines | 3.986 | 0.650 | 0.051 | 3.984 | 0.596 | 0.099 |

model based on administrative subdivision in municipalities is slightly more suitable to explain the global data variance.

The marked difference between means $(\mu, GM)$ and standard deviations $(\sigma, GSD)$ that characterize the two set of territorial data (Table 4) suggests to apply, to each of the two macro-units, a simplified model, respect to [8], eliminating the variable $\alpha$ :

$$\ln(R_{ij}) = \mu + \beta_i + \varepsilon_{ij} \qquad [10]$$

$$\beta \cong N(0,\sigma^2)\ ;\ \varepsilon \cong N(0,k^2) \qquad [11]$$

In Table 5 a summary prospect of the results carried out by the application of the ANOVA to the two macro-units in the cases a) and b), according to the model [10], has been reported.

## 5 IDENTIFICATION OF RADON PRONE AREAS

The application of the Bayesian inference to the available data set according to the two different areal subdivisions (grid squares or municipalities, after regional territory subdivision in two macro units) gives the estimate, for each $i^{th}$ areal unit, of the variable $\ln(GM_i^{eval})$ and of its variance $V_i^2$ (see [6] and [7]).

The variance of the data within each unit is given by $V_i^2 + k^2$ (Gelman *et al.*, 2004). Finally, a normal distribution is assumed for the variable $\ln(R_i)$ (the natural logarithm of all the measures of indoor radon at ground floor of buildings, realizable in the $i^{th}$ unit):

$$\ln(R_i) \cong N[\ln(GM_i^{eval}), \ln(GSD_i^{eval})^2] \qquad [12]$$

where $\ln(GSD_i^{eval})^2 = V_i^2 + k^2$, $GSD_i^{eval}$ is the geometric standard deviation of the measures $R_i$ after the Bayesian estimate. At this point, in order to evaluate the percentage of ground floor dwellings in which the radon concentration exceeds a selected reference level $X$ , the standard normal variable $Z$ is introduced by

$$Z = \frac{\ln(X) - \ln(GM)}{\ln(GSD)} \qquad [13]$$

and, consequently, the area $Q(Z)$ under the curve exceeding the $\ln(X)$ can be calculated.

A possible choice for *X* is 200 Bq/m$^3$, as recommended by the European Commission as upper value for future constructions, while a proportion of 20% for $Q(Z)$ (i.e. the expected percentage of ground floor dwellings with radon concentration above 200 Bq/m$^3$) could be one of the possible values to identify radon areas presenting a higher potential hazard.

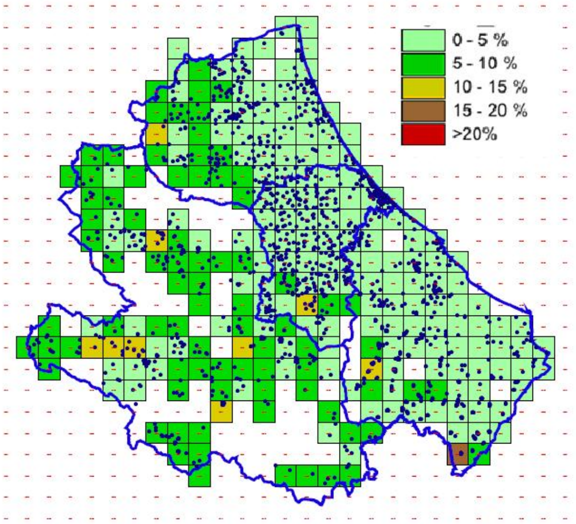

Figure 3: Probability to exceed 200 Bq/m$^3$ for radon concentration in ground floor dwellings, for each 6x6 km cell. In white, areas without data.

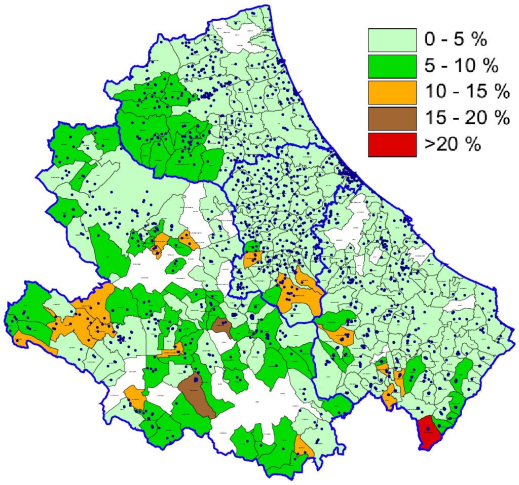

Figure 4: Same as Figure 3, but related to municipalities.

The analysis of the results (Fig. 3) shows that there is no cell in which such a value is reached. As for the municipalities, the probability map in Figure 4 seems to be less smooth, so a larger number of areas present values higher than 10% and there is one case of $Q$ exceeding 20% at the southern boundary of the region.

## 6 GEOSTATISTICAL APPROACH

Geostatistics provides interesting features for radon mapping (Zhu *et al.*, 2001, Chaouch *et al.*, 2003, Dubois *et al.*, 2007). We have applied a basic approach of this technique, intending to expand the analysis to a wider extent in the next future. From a geostatistical point of view, indoor radon concentration may be considered as a *regionalized variable*, whose peculiarity is to describe environmental phenomena as a combination of a large scale spatial structure with random local variability (see e.g. Chilès and Delphiner, 1999).

A variogram is a useful means to describe and model the spatial correlation of observed data. In a few words, it describes the dissimilarity in the variable's values as a function of the distance between two points in the spatial domain. As a first step, we have elaborated the experimental variogram of natural logarithm of our data (full year indoor radon concentration, normalized to ground floor). Note that we consider the building as the sampling unit; if two or more measurements were performed in the same building, their average was chosen as representative of the building. Figure 5 displays both the experimental and the fitted modelled variograms.

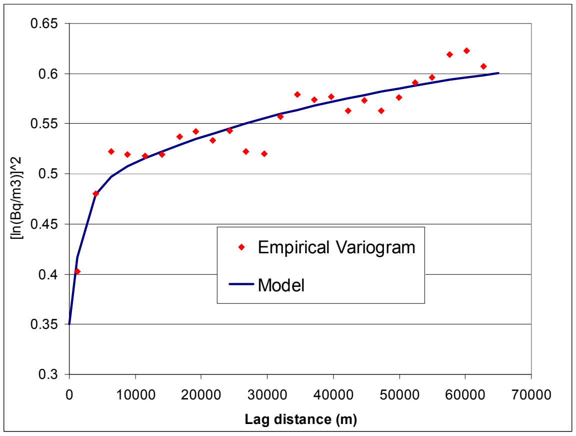

Figure 5: Empirical (dots) and modelled (solid line) variograms.

Approaching zero distance, the variogram shows a non-vanishing value. This is due to the so-called "nugget" effect, which is a characteristic occurrence in the presence of a strong variability at small scale (also due to the measurement errors), typical of the indoor radon data (Chaouch *et al.*, 2003, Dubois *et al.*, 2007).

We perform leave-one-out cross validation (see e.g. Dubois *et al.* 2007) to select a model in order to fit the empirical variogram. The best result was shown by a nested structure as follows:

$$\gamma(h)=\tau^2+\sigma_1^2\left(1-e^{-h/R_1}\right)+\sigma_2^2\left(1-e^{-h/R_2}\right) \qquad [14]$$

with a nugget $\tau^2 = 0.35$ and two exponential components (partial sills $\sigma_1^2 = 0.13$; $\sigma_2^2 = 0.16$ and ranges $R_1$ = 1900 m; $R_2$ = 49000 m). The variogram seems to point out the existence of some large scale underlying trend, probably in the NE–SW direction or from the coast toward inland. This preliminary model does not include, up to now, spatial anisotropy, ($h$ in [14] is the modulus of the separation vector) although exploratory analysis of data seem to suggest the opposite. This aspect, among other ones, will be verified when we will have a wider experimental sample data at our disposal. Another possible improvement could consist in carrying out variogram modelling separately in each of the two zones previously shown (see Fig. 2).

An ordinary kriging, based on the described model of variogram, has been performed on the log-transformed data, producing the map in Fig. 6. The same macrostructures, already made evident by the Bayesian inference (even though with coarser resolution, see Fig. 4-5), are clearly shown. The coastal area is generally characterized by pretty smooth values around or below 50 Bq/m$^3$. Besides, the radon levels in the montane-piedmont area are substantially higher, showing also a marked variability from zone to zone. On a smaller scale, some areas are conspicuous for relatively high radon level, on western and southern sides.

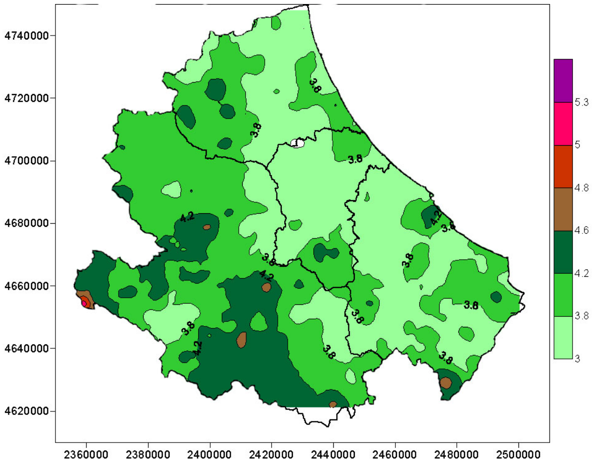

Figure 6: kriging map of the logarithm of indoor radon, normalized to the ground floor (ln(Bq/m$^3$)).

For some of the areas listed above, it is possible to infer a causal connection with local geologic features (Crescenti *et al.*, 2003). A key role, to be verified in futures studies, could be played by the diffuse presence of faults and fractures, within the remarkable areas previously outlined.

## 7 CONCLUSIONS

The paper describes the first results obtained by the Environment Protection Regional Agency (ARTA) of a regional survey made in Abruzzo. The aim of the study was the identification of high radon concentration areas. The monitoring survey was designed on the basis of the data available from similar surveys, carried out within the period 1992-2000 on the same territory. The sampling design was based on a territory subdivision through regular grid squares with two scale levels, 6 km and 2 km. We choose to randomly sample at least one dwelling (or, subordinately, places like schools, offices, shops and other workplaces) in every 2x2 km cell, in which to measure radon concentration with passive alpha-tracks detectors, in order to obtain a relatively uniform sampling of the whole territory. The set of validated data, inclusive of about 600 measurements from two previous surveys, consists of 1 647 annual mean values of indoor radon activity concentration. All these data have been georeferenced and stored in a database in combination with a Geographic Information System (GIS). The statistical analysis, based on a Bayesian inference technique, shows evidence of some critical areas in the Apennine part of the region but does not allow, up to now, a clear identification of radon risk areas. The currently available dataset, actually, lacks of a necessary number of data just in many of the potentially hazardous areas, in particular in the province of L'Aquila, the most mountainous region of the considered territory. However, a collection of a wider and more homogeneous set of data is still ongoing. Finally, further efforts will focus on improving the geostatistical analysis of data, on evaluating the representativeness of data respect to the population, and on comparing different choices of Q and reference levels.

## 8 ACKNOWLEDGEMENTS

We would like to thank the reviewers Dr. Francesco Bochicchio, Istituto Superiore di Sanità (Italy) and Dr. Gregoire Dubois, European Commission – DG Joint Research Centre, Institute for Environment and Sustainability, for their kind effort and helpful suggestions.

Many thanks to the personnel of ARTA for radon measurements and survey organization: S. D'Ostilio, D. Rancitelli, P. Ciafardone, G. Buccella, P. Pellegrini, C. Cimoroni, L. Stornelli, R. Capannolo, F. Benedetti, U. Miconi, L. Carnesale, G. Gianfelice.

## 9 REFERENCES

ANPA 2000. Sistema Informativo Territoriale per la Valutazione del Potenziale di Esalazione di Radon dal Suolo, *Serie Stato dell'Ambiente 9/2000*.

P. Armitage, T. Colton (editors) 2005. Encyclopaedia of Biostatistics, *John Wiley & Sons*.

F. Bochicchio, G. Campos Venuti, C. Nuccetelli *et al.* 1996. Results of the Representative Italian National Survey on Radon Indoors, *Health Physics, Vol. 71(5), 741-748.*

F. Bochicchio, G. Campos Venuti, S. Piermattei *et al.* 2005. Annual average and seasonal variations of residential radon concentration for all the Italian Regions, *Radiation Measurements 40, 686-694.*

Chaouch, A., M. Kanevski, M. Maignan, J. Rodriguez and G. Piller 2003. Indoor radon data mining with geostatistical tools: case study with a highly clustered and variable dataset, *Proceedings of International Association for Mathematical Geology, IAMG 2003.*

J.P. Chilès and P. Delfiner 1999. Geostatistics: Modeling Spatial Uncertainty, *John Wiley & Sons*.

U. Crescenti, E. Miccadei, A. Praturlon (editors) 2003. Guide geologiche regionali*:* Abruzzo*, Società Geologica Italiana.*

S. Darby, D. Hill, A. Auvinen, J.M. Barros-Dios, H. Baysson, F. Bochicchio *et al.* 2005*.* Radon in homes and risk of lung cancer: collaborative analysis of individual data from 13 European case–control studies. *BMJ;330:223–228.*

G. Dubois 2005. An overview of radon surveys in Europe, EUR 21892 EN, *Office for Official Publications of the European Communities*, Luxembourg.

G. Dubois, P. Bossew, H. Friedmann 2007. A Geostatistical autopsy of the Austrian indoor radon survey (1992-2002), The *Science of the Total Environment 377, 378-395.*

A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin 2004. Bayesian Data Analysis, *Chapman & Hall*.

C. Giannardi, F. Giovannini, S. Bucci *et al.* 2001. In progress identification of radon prone areas: Toscana and Veneto, *Radiation Protection Dosimetry 97*, *349-354.*

M. Kutner, C.J. Nachtsheim, J. Neter, W. Li 2005. Applied Linear Statistical Models*, McGraw-Hill*.

J. Miles 1998. Mapping Radon-Prone Areas by Lognormal Modeling of House Radon Data, *Health Physics 74(3), 370-378.*

P.N. Price, A.V. Nero, A. Gelman 1996. Bayesian Prediction of Mean Indoor Radon Concentrations for Minnesota Counties, *Health Physics 71(6), 922-936.*

S. Verdelocco, P. Turkowsky, D. Walzer 2000: "L'incidenza dei fattori geologici e delle variazioni climatiche nell'individuazione di aree ad alto potenziale di radon, *Geologia Tecnica e Ambientale*, *n. 1/2000*, *45-52.*

H.C.Zhu, J.M. Charlet, A. Poffin 2001: Radon risk mapping in southern Belgium: an application of geostatistical and GIS techniques, *The Science of the Total Environment 272, 203-210.*