16ENV10 MetroRADON

Activity 4.4.2

Radon mapping exercise

Lead organisation: Austrian Agency for Health and Food Safety (AGES)

Authors:

Valeria Gruber, Sebastian Baumann, Klara Himmelbauer, AGES, Linz, Austria

Christian Laubichler, Oliver Alber, AGES, Graz, Austria
Peter Bossew, Eric Petermann, Bundesamt für Strahlenschutz (BfS), Berlin, Germany
Giancarlo Ciotoli, Italian National Research Council, CNR-IGAG, Rome, Italy
Alcides Pereira, Filipa Domingos, University of Coimbra, Coimbra, Portugal
François Tondeur, ISIB-HE2B, Brussels, Belgium

Giorgia Cinelli, EC JRC, Ispra, Italy
Carlos Sainz, Luis Qunidos-Poncela, Alicia Fernandez, University of Cantabria (UC), Santander, Spain
Jose-Luis Gutierrez Villanueva, Radonova Laboratories AB, Uppsala, Sweden

The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States

Inhalt

1. Motivation

The purpose of the MetroRADON project, funded within the European Metrology Programme for Innovation and Research (EMPIR) is to develop reliable techniques and methodologies to enable SI traceable radon activity concentration measurements and calibrations at low radon concentrations. The need for this project has been largely motivated by the requirements of the implementation of the European Council Directive 2013/59/EURATOM (EU-BSS) (EU, 2014), one aim of which is to reduce the risk of lung cancer for European citizens due to high radon concentrations in indoor air. Furthermore, it is a goal of the project to enable uptake and exploitation of its results and experiences by all stakeholders concerned with radon, from regulators and policy makers, professionals in designing, performing, evaluating and interpreting radon surveys, radon instrument manufacturers to the construction industry and scientific community. More details about the MetroRADON project can be found at the project website (MetroRADON, 2020).

Article 103 of the EU-BSS requires, that member states identify areas where the radon concentration in a significant number of buildings is expected to exceed the relevant national reference levels. Those areas are in practice referred to radon priority areas (RPA). Definition and delineation of RPA is relevant, as specific (mandatory) measures of the radon strategy of countries depend on it (e.g. radon measurements at workplaces, preventive measures, awareness programs). Therefore, the delineation of RPA is an important tool within the transposition of EU-BSS and radon action plans in the countries, which should be implemented appropriate, accurate and reliable.

A specific work package is included in the MetroRADON project with the aim to analyse and develop methodologies for the identification of radon priority areas. As the definition of RPA in the EU-BSS allows a wide range of interpretation, different concepts and methodologies have been proposed and some already adopted. One activity of the work package is to evaluate the concept for RPA and methods of radon mapping which are already used in different countries and their usage for other countries and the harmonisation of radon priority areas across borders.

Within this framework, an activity carried out within the MetroRADON project was, to apply existing mapping methods used in different countries using harmonised data sets of various variables (e.g. indoor radon, gamma dose rate, geology, soil gas radon). The activity was focused on evaluating their comparability and their usability for other countries and is referred to as “the radon mapping exercise” and discussed in this report. The results and findings from the exercise will be included in the discussion of possibilities of harmonisation of RPAs across borders and incorporated in a guideline on the definition, estimation and uncertainty of radon priority areas for MetroRADON stakeholders.

2. Introduction

2.1 Mapping methods and radon priority areas

Radon mapping and definition of radon priority areas (RPA) are very complex topics. As discussed above, the definition of RPA in the EU-BSS allows a wide range of interpretation and therefore different concepts and methodologies have been proposed and, in some countries, already adopted. Radon mapping was also relevant already before the new EU-BSS, so in many countries, radon maps exist for many years as part of the national radon strategies. The used mapping methods and the visualisation are very different, depending on the purpose of the map and the data behind it. These different methods are based on different developments, strategies and ideas in radon protection for many years in the countries, and most of the time the basic mapping strategies and methods applied in a country remain the same, even when revised or new legal requirements apply. Consequently, a basic bottom-up harmonisation approach of mapping methods or definition of RPA will not be enforceable. Therefore comparison, evaluation and discussion for possibilities of top-down harmonisation are important.

As a starting point for this "radon mapping exercise" report, some basic information about different possible and used radon mapping methods and definitions of RPA is given, for better understanding of the situation and framework of the radon mapping exercise.

As said, radon mapping can be done (and is done in practice) with various different methodologies. The methodologies are composed of different parts, like the mapped parameter (P), the mapped unit (U) and the used display method (D). Table 1 shows an overview of possibilities for the mentioned three parts of the methodology (no guarantee to be complete!). The mapped parameter can be either the raw measurement value of e.g. indoor radon concentration or soil gas radon or an already processed (e.g. normalised) measurement value (e.g. application of seasonal correction factor, taking into account building characteristics) or a combination of different parameters (e.g. often used geogenic radon potential, defined by soil gas concentration and permeability). The units can be administrative unit, a regular grid cell or geological unit. For display methods either simple descriptive statistics like arithmetic or geometric mean of the parameter or the percentage of houses in a certain unit, which exceed a certain threshold (e.g. the reference level) can be used. Also, qualitative risk classes could be applied (e.g. parameters are classified according to risk classification scheme) or a risk index defined (e.g. classification of the different parameters and combination for an overall index to classify). The options (1-3 or 1-5) within the three different parts (U, P, D) can be combined with each other – so there are various possibilities of different overall mapping methodologies, as we can see also in practice.

Some applied examples in countries: Austria was following (and will also follow in the future for the planned new radon map and delineation of RPA) the scheme U1-P2-D1, with standardised indoor radon measurements (use of standard house), averaged for a municipality (Friedmann, 2005). Several countries follow the scheme U1/2-P1-D2/3, by using the indoor radon measurement values and displaying the percentage or probability of houses above the reference level per administrative unit or grid cell. This method is very common, as it reflects most directly the definition of the EU-BSS – areas, where the radon concentration in a significant number of buildings is expected to exceed the relevant national reference levels. But using this methodology for an accurate delineation of RPA needs a high number of representative measurements in the respective grid cells or administrative areas. A simplified display method is to use risk classes (e.g. low, medium, high risk), which can be defined by different input parameter. USA uses a risk index (multivariate classification) for radon mapping, following the scheme U1-P3-D5, taking into account indoor radon measurements, geology, air-borne gamma ray

spectrometry , soil parameters and foundation types with classification on county level. Each parameter is classified according to defined criteria and then the classification of all parameters is summed up to a risk index (EPA, 1993). The concept of a geogenic radon risk index with multivariate classification is also discussed for Europe with scheme U2-P3-D5, driven by activities by the European Commission, Joint Research Centre (JRC) (EC JRC, 2020; Bossew et al., 2016) and to further develop this (Radon Hazard Index, RHI) also a main aim of this work package of the MetroRADON project.

Table 1: Overview of possible radon mapping methodologies

Unit (U)	**Parameter (P)**	**Display method (D)**
1 - Administrative Unit (e.g. municipality)	1 - Measurement value (e.g. indoor concentration)	1 - Descriptive statistics (e.g. mean, med, max)
2 - Grid cell	2 - Modelled value (e.g. seasonal correction, reference house)	2 - % of houses/measurement values above RL
3 - Geological unit	3 - Combination of different parameters (e.g. radon conc., geology, permeability)	3 - Probability that RL is exceeded
		4 - Risk classes (qualitative)
		5 - Risk index

Besides the mapping methodology, each country has to decide about a definition of RPA. A threshold needs to be set, when an area (e.g. administrative unit, grid cell) is considered to be a radon priority area. In the EU-BSS it says “areas, where the radon concentration in a significant number of buildings is expected to exceed the relevant national reference levels defined in the EU-BSS”. What is considered to be a significant number of buildings needs to be decided by the countries and will be dependent on the radon potential of the country and on economic and political considerations, as the measures dependent on delineation of RPA need to be also manageable. Therefore also for the definition of RPA different concepts are adapted in the countries, some examples are listed in Table 2 (not complete list). Several countries define RPA with a threshold of 10% of houses above the respective RL, but also 1% (UK) and up to 30 % (Czech Republic) are used. Some countries, e.g. Austria, do not refer their RPA to the percentage of houses above RL, but use descriptive statistics in administrative units instead.

As shown, the mapping methodologies are various and so are the definition of RPA - to evaluate the situation in Europe and possibilities for harmonisation between countries and on borders was the driving factor for this work package within the MetroRADON project and also for this mapping exercise.

Table 2: Examples of radon priority areas (RPA) definitions in different European countries.

Country	**Definition of RPA**
Austria	modelled AM > 300
Belgium	Prob (C > 300) > 5%
Cyprus	AM (C) > national average
Czech Republic	Prob (C > 300) > 30%
Finland	Prob (C > 300) > 10%
Germany	Prob (C > 300) > 10% with 90% confidence
Ireland	Prob (C > 200) > 10%
Lithuania	Prob (C > 300) > 10%
Luxembourg	Prob (C > 300) > 5%
Malta	Prob (C > 200) > 1%
Spain	Prob (C, ground or 1. floor > 300) > 10 %
UK	Prob (C > 200) > 1%

2.2 The MetroRADON mapping exercise

The activity 4.4.2 within MetroRADON project, which we call “the mapping exercise” was defined in the project description as:

A4.4.2	AGES, BfS and UC will test existing mapping methods used in various countries (e.g. indoor radon, gamma dose rate, geology, soil gas radon) with different datasets accessible to JRC and BfS (e.g. national data from other countries, Austrian data set from extensive survey in 6 municipalities and the JRC database) and evaluate their comparability and their usability for other countries. Prior to the start of this exercise the methods and data which can be used for this exercise will be selected based on their applicability and ensuring sufficient data are available for a region.	**AGES**, BfS, UC

So, the idea for the exercise is, to use a provided data set and apply the individual mapping method and definition of RPA used in the country or was proposed by experts to the provided data. Afterwards the mapping and classification results for the provided data sets in the relevant areas will be compared and the usability evaluated.

The first step was to find data sets, which can be used for the exercise. The idea was to use at least two different data sets possibly different in geology, scale, co-variables, etc. to increase the scope and benefit of the exercise. As planned already from the beginning, the data set from extensive survey in six municipalities in Austria was available. As second data set Cantabria, Spain was chosen to be used for the exercise. It was important, that the data sets fulfil our needs and are available by MetroRADON partners and no data protection issue occur.

The data sets then were prepared for the exercise to reach best usability for the participants. The data were arranged in a uniform format, where necessary some anonymisation measures were taken and all data were georeferenced and either shared as table or as shp-file. All details about the data sets can be found in chapter “Exercise data”.

To find volunteers for the exercise, experts from different countries, which are known to work in the field of radon mapping, were asked, not limited to MetroRADON partners. Of course to fulfil this task – applying their mapping method to unknown, new data – needs some time and human resources, which was not applicable for all asked experts and institutions, especially if not MetroRADON partners, and no funding could be provided for it, even if a lot of interest was shown in the exercise by most of them. In the end experts from five institutes from five different countries did participate with their methods or did provide extensive data set analysis. The data set analysis and the applied methods and results are discussed in the chapters “data set analysis” and “methods and results”, the names of the experts and institution, who did the main work for each chapter are listed. Thanks to all the participants for their voluntary and important contribution!

In the end, discussed in chapter “summary and discussion”, some comparisons of methods and some interpretations of the results is done. These results and discussion will be part of the MetroRADON Deliverable D5 “Report and Guideline on the definition, estimation and uncertainty of radon priority areas (RPA)”.

In the chapters, the name of the participants who performed the major work for the chapter are listed. AGES might have added some text and explanations to the chapters and did some editing. The chapters were no names are listed were mainly written by AGES, with the help of the co-authors.

3. Exercise data

The MetroRadon exercise uses data of different radon measurement campaigns in Austria (six municipalities) and Spain (Cantabria). The data include indoor radon measurements, building characteristics of measured dwellings, soil air radon activity concentration, permeability estimation, activity concentration of soil samples, ambient dose rate and maps of geogenic parameters derived from other sources (e.g. geology, soil type, airborne radiometry). All data are georeferenced and provided in shape files (point and polygon) or TIFF raster files. Additionally, csv files are added to the data set as robust reference data for point data.

3.1 Austrian data set

The Austrian data set covers six municipalities and is separated in two distinct areas in the North and in the South of Austria (AUT North, AUT South), each consisting of three adjacent municipalities with an overall area of about 220 km² (Figure 1), about 40km² in AUT North and 180 km² in AUT South.

The area AUT North is located in the Bohemian Massif which is one of the areas in Austria with the highest geogenic Radon potential due to the predominant geology of granites and gneisses. AUT North features a homogeneous geology of a granitic pluton and interlaying migmatites (partly molten during metamorphosis). The geology of AUT South is more heterogeneous and also features a variety of felsic igneous and metamorphic rocks with a high radon potential, but also different sediments with a low radon potential.

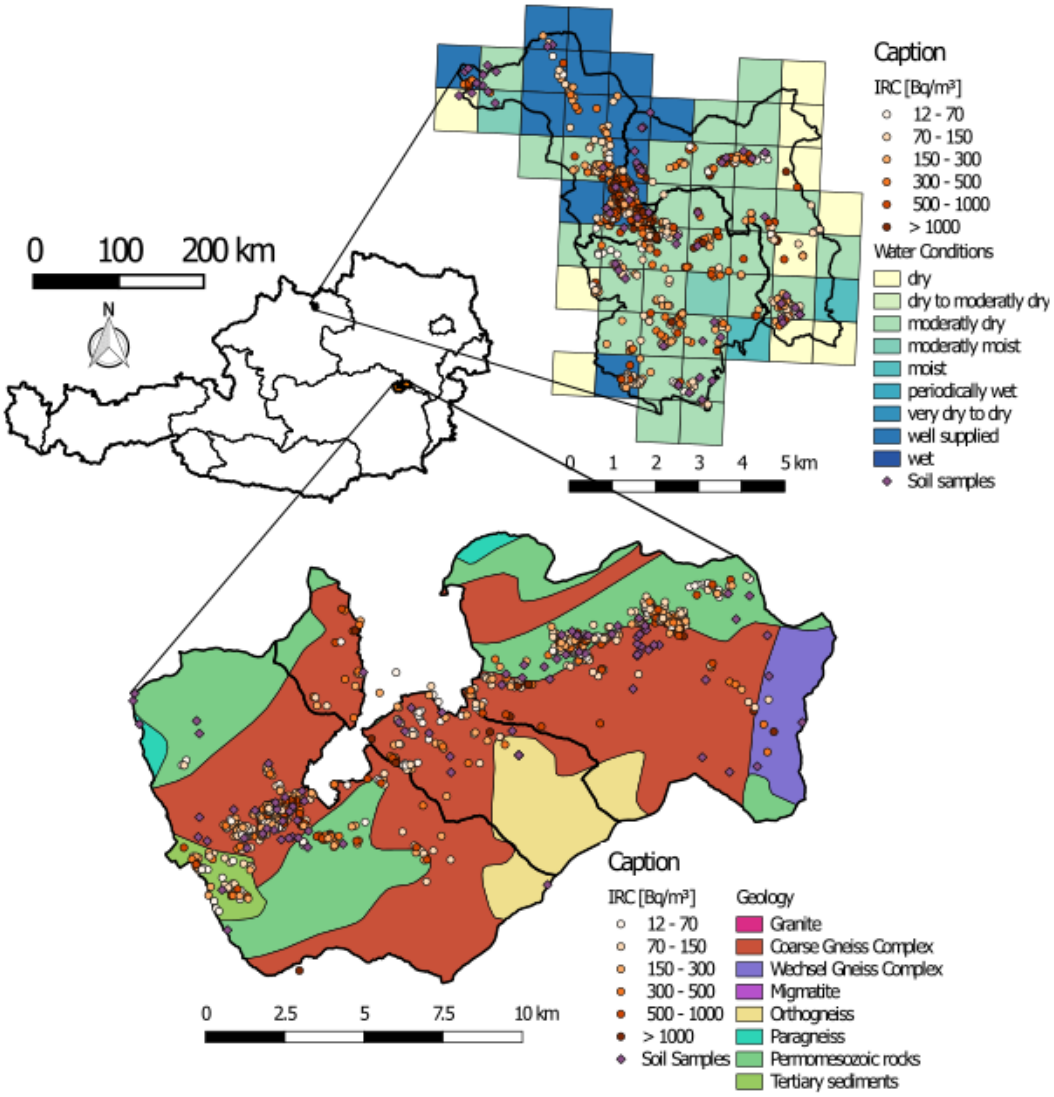

Figure 1: The Austrian data set with selected variables and a map of Austria showing the position of the areas AUT North and AUT South.

The **indoor radon concentration (IRC)** includes data of detailed indoor radon measurement campaigns in dwellings carried out 2010 and 2012 in six municipalities in Austria. Participation rate was 60 to 90 % of all households in the municipalities. IRC was measured with solid-state nuclear track detectors (SSNTD) (Raduet, RSKS Radosys). The measurement periods varied between four to six months, half winter and half summer. The detectors were located in the two most used rooms of the dwellings. Additionally, the participants completed a questionnaire of building characteristics. The data of 1.638 houses was provided, including 3.241 IRC measurements. Because of partly different questionnaires of the measurement campaigns, some data was only available for certain areas (e.g. type of heating). The data was anonymized and a small random term has been added to the original coordinates, whereby the spatial attributes of the shape files (municipality, geology and soil map) have been preserved for the new locations. Table 3 summarizes the variables stored in the IRC data set and gives examples of the attributes.

Table 3: Variable description of the indoor radon concentration data set.

Variable	Description	Unit	Examples
rn_c_r1	Radon concentration room 1	Bq/m³	120, 304, 56
rn_c_r2	Radon concentration room 2	Bq/m³	120, 304, 56
rn_c_1_err	Error radon concentration room 1	%	12, 30, 5
rn_c_2_err	Error radon concentration room 2	%	12, 30, 5
b_rn_c_AM	Arithmetic mean of radon room concentrations	Bq/m³	120, 304, 56
r1_type	Type of room 1		sleeping room, kitchen
r2_type	Type of room 2		sleeping room, kitchen
r1_floor	Floor of room 1		3, -1
r2_floor	Floor of room 2		3, -1
r1_earthb	Is room 1 earthbound?		y, n
r2_earthb	Is room 2 earthbound?		y, n
b_basement	Has the building a basement?		fully, no, partly
b_ac_units	Total number accommodation units		1, 2, ≥ 3
b_year	Building year of dwelling		1986, 2001
b_year_i	Building year of dwelling in interval		1971-2000, < 1900
b_type	Type of building		one family dwelling
b_hill	Is the property located on a slope?		y, n
b_neigh	Neighbourhood position of building		solitary, built together
b_found	Foundation of house		no foundation, strip foundation
b_floor	Floor construction in zone of foundation		screed, sand
b_walls_eb	Main material of earthbound walls		brick, concrete
b_walls	Main material of walls		brick, concrete
b_window	Air tightness of windows		low, well
b_therm	Thermal construction of building		passive, low energy
b_heating	Type of heating		central heating, electric heating
b_older_14	Number of persons in household older than 14		2, 0
b_young_14	Number of persons in household younger/equal 14		2, 0
b_rem	Remediation or extension of building		no, 1970-2000
b_m_start	Start of radon measurement		08.02.2010, 30.01.2013
b_m_end	End of radon measurement		01.07.2010, 24.06.2013
b_altitude	Altitude of building	m	373, 820

The data of **soil air radon activity concentration, permeability, soil activity concentration and ambient dose rate** originate from different measurement campaigns of radon activity concentration in soil air. 148 locations have been measured in the municipalities of the indoor radon survey (Figure 1). Additionally, permeability, soil

activity concentration and the ambient dose rate have been measured on selected locations (approximately 100).

The provided **soil air radon activity concentration** [kBq/m³] was calculated on the basis of three single measurements for every location. Steel probes of 1.6 m length and a diameter of 12 mm were used. The intended sample depth was 1.4 m. The principle of the lost tip was used to generate a cavity, which represents the effective probe volume. Soil air was vacuumed with a syringe (200 ml) attached airtight on the steel probe. The first 200 ml of vacuumed air were rejected, due to a mixture of atmospheric and soil air. The next sample of 100 ml vacuumed air were directly transferred to an Alpha Guard® for measuring the radon activity concentration. For some locations, where the intended depth of 1.4 m was not reached, the single measurements were normalized to a depth of 1.4 m. The given results are the arithmetic mean of the depth-corrected soil air radon activity concentration of the three single measurements at the sample locations.

Estimations of soil **permeability** [m²] were carried out at the same locations as the measurement of the soil air radon activity concentration. Soil air was vacuumed from the steel probe with a pump capacity of 1 litre per 60 sec (AlphaPUMP). Flow rate and pressure were measured with a flow meter. The geometry parameters (depth, length of effective probe volume, width of probe) and the flow meter results were used to calculate the permeability after Damkjaer & Korsbech (1992). The results show the arithmetic mean of three single measurements at the sample locations.

Activity concentration in soil samples [Bq/kg] of K-40, Pb-210, Ra-226, Ra-228, Th-228, U-238 were measured with gamma ray spectrometers (HPGe, LEGe). The soil samples were taken at selected locations with core samplers (2 cm diameter, 1 m profile) in the centre of the three single measurements of the soil air radon activity concentration measurements. The soil samples were dried for 24 hours in a drying cabinet (105°). Afterwards they were transferred into gas-tight loading cells. The sealed samples were stored for three weeks to ensure a radioactive equilibrium of the U-Ra radioactive series.

Note, that for the results of the radionuclide concentration, a fixed value was assigned to the limit of detection and the corresponding error was set to zero in order to avoid data loss. A conservative approach regarding the handling of detection limits is used. Thus, only numerical values for the parameters are given. This ensures data consistency.

The **ambient dose rate (ADR)** was measured in a height of 0.5 m for five minutes with a dose rate meter (Automess, 6105AD) and scintillator probe (Automess 6105AD-b/E). The built-in mean calculation of the dose rate meter was used, which ensures that the relative standard deviation is below five percent.

The following Table **4** summarizes the variables soil air radon activity concentration, permeability, soil activity concentration and ambient dose rate.

Table 4: Variable description of the soil data set.

Variable	Description	Unit
rn_sair	Radon activity concentration in soil air	kBq/m³
rn_sair_er	Radon activity concentration in soil air error	kBq/m³
permea	Permeability estimation	m²
ADR	Ambient dose rate	µSv/h
K_40	K-40 activity concentration	Bq/kg
K_40_er	K-40 activity concentration error	Bq/kg
Pb_210	Pb-210 activity concentration	Bq/kg
Pb_210_er	Pb-210 activity concentration error	Bq/kg
Ra_226	Ra-226 activity concentration	Bq/kg
Ra_226_er	Ra-226 activity concentration error	Bq/kg
Ra_228	Ra-228 activity concentration	Bq/kg
Ra_228_er	Ra-228 activity concentration error	Bq/kg
Th_228	Th-228 activity concentration	Bq/kg
Th_228_er	Th-228 activity concentration error	Bq/kg
U_238	U-238 activity concentration	Bq/kg
U_238_err	U-238 activity concentration error	Bq/kg
altitude	Altitude of measurement location	m

Data from literature include **airborne radiometry, tectonic, geological** and **soil maps.**

Airborne radiometry data derive from the geological survey of Austria (GBA) database and represents the uranium concentration of the uppermost soil layer (equivalent Uranium eU). A gamma ray spectrometer (PICO ENVIROTIC GRS410) with sodium-iodide crystals was used for the surveys. A fixed cruising altitude of 80 m, a profile distance of 200 m and a flight velocity of 125 km/h were intended. Data processing included various manipulations such as cruising altitude correction, topographic correction, vegetation correction, cosmic ray correction, radon correction and the consideration of the Compton Effect as well as the conversion of counts per peaks to concentration (IAEA, 1979). The data was only available for three municipalities of the Austrian data set (AUT North).

The data source for **geological maps and tectonic lineaments** is the geological survey of Austria with geological map of scale 1:500.000 (GBA, 2020). For the region of AUT North, an additional geological map with a finer resolution of 1:50.000 was added (GBA, 2020).

The **soil map** is a generalized data set of the Austrian soil map in a 1 x 1 km grid and includes variables as soil type, soil water content, permeability and soil depth (BFW, 2020).

3.2 Cantabrian data set

The data set in Spain covers the region Cantabria with a total area of about 5.300 km². The data set consists of different measurement campaigns of indoor radon concentration, radon concentration in soil gas, ambient dose rate and various compiled data from literature (Figure 2).

The geology of Cantabria predominantly features detritic sediments and carbonates which usually show a low to intermediate radon potential, but especially high permeability in carbonates can also locally lead to a higher radon potential. Also occurring Metasediments and volcanoclastics usually show a low radon risk and compared to the Austrian regions Cantabria has a lower geogenic radon potential.

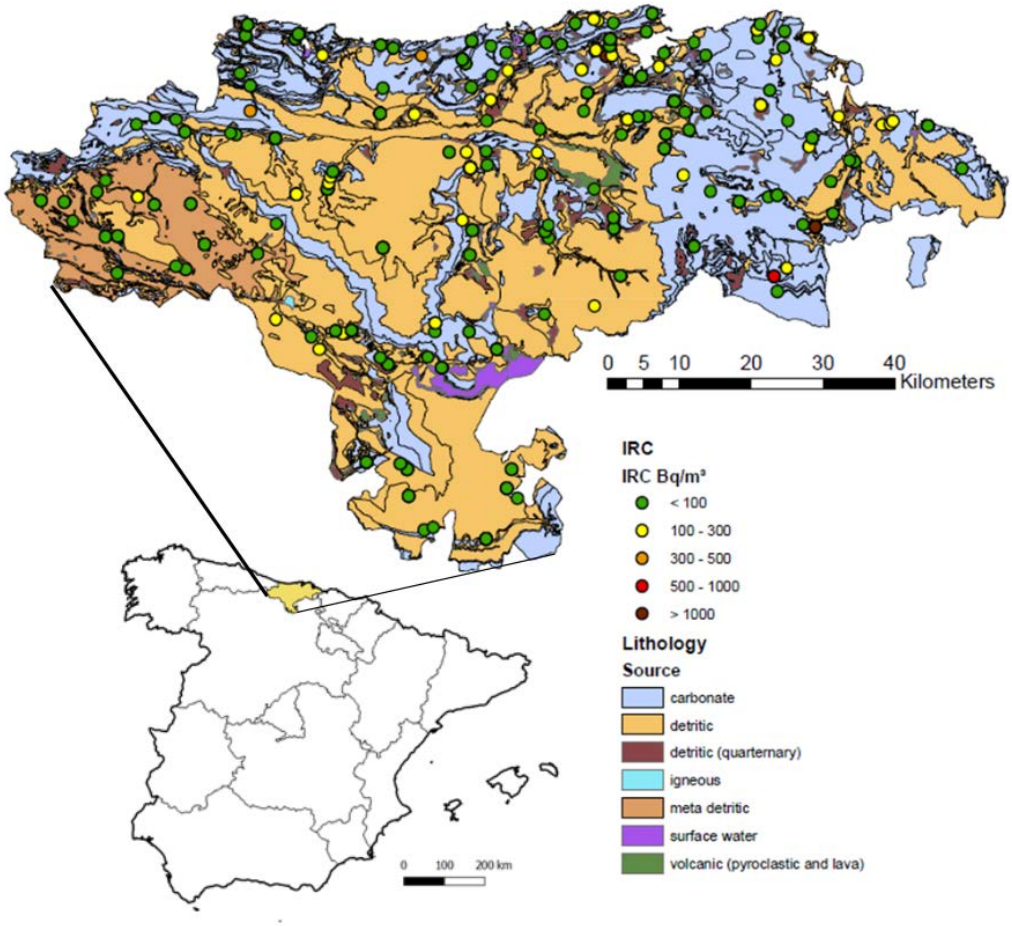

Figure 2: Map of Cantabria with selected variables and the position of Cantabria in Spain.

The data on **indoor radon concentration** in homes was obtained from the Spanish Radon Map Expansion Project. Between 2011 and 2016, the project was promoted by the Spanish Nuclear Safety Council (CSN) and the University of Cantabria (UC) with collaboration of the University of Santiago de Compostela (USC) and the Autonomous University of Barcelona (UAB). Since the 1980s, a compilation of measurements is available from the Radon Group of UC. Measurements were carried out on the ground floor of dwellings or, failing that, on the first floor, using CR-39 trace detectors following the internal location protocol of the Laboratory of Radioactivity of the University of Cantabria (LaRUC) (Sainz-Fernandez et. al., 2014). The data includes indoor radon concentration [Bq/m³] and the location of the sample. Note that this location represents the location of the main city where the measurement was performed within a 10 x 10 km grid rather than the actual measurement position.

The **radon concentration in soil gas** was measured in 260 samples and was collected from 2011 to 2016. The measurements were performed with using a sampling technique based on the collection of a soil gas sample from a depth of about one meter (Czech Method; Neznal et. al, 2004). A sampling probe was prepared with a tip at the lower end and pounded to a depth of about one meter. The punch wire into the probe was inserted and the sharp tip moved a few centimetres lower, which created a cavity at the lower end of the probe. A rubber

tube was inserted into the sampling probe and the gas was extracted with a syringe. The gas sample was introduced into a previously evacuated ionization chamber. The detection principle of the measuring system is called RM - 2 and is based on an ionization chamber operating in a current mode.

The **ambient dose rate** (ADR) in the region of Cantabria was collected in the context of the MARNA Project (Suarez et al., 1997). In 1991, the Marna project, developed in Spain, evaluated the rate of exposure to terrestrial gamma radiation at a height of one meter above the surface of the ground.

Data from both parameters, **lithology** and **permeability** originate from the geological map of Spain at a scale of 1:200,000 (IGME, 2020). Besides lithology, the map features a permeability estimation and a classification of the petrological origin. The IGME main criterion for the graphic representation was mapping of the units with significant lithostratigraphic development. It incorporates units with a high hydro-geological interest because of their lithological nature (e.g. high permeability) and because they were considered as the essential part of the definition of aquifers.

The assigned permeability values are indicative for the hydro-geological capacity (aquifer permeability) of the bedrock and do not reflect the permeability of the first meters of soil (surface formation). However, the first meters of soil are most relevant to the explanation of radon transport and furthermore to the presence or absence of radon in buildings.

Nevertheless, the map classifies lithostratigraphic units with different hydrogeological characteristics. The classification of the petrological origin groups the lithostratigraphic units in seven categories: carbonated rocks, detritic rocks, Quartenary detritic rocks, evaporite rocks, igneous rocks, volcanic rocks and meta-detritic rocks.

Data origin of **faults** is the 1:1,000,000 IGME failure map (IGME, 2020a), developed within the framework of the One Geology project.

Compiled data of the **activity concentration in soil** derive from the FOREGS (2020) and GEMAS (2020) database. The data is provided in regular grid (10 x 10 km) with the variables Potassium [%], Uranium [ppm] and Thorium [ppm]. For Potassium, the arithmetic mean was used to calculate the cell means. For Thorium and Uranium, the geometric mean was used to calculate the cell means.

Karst data is a simplified version of the IGME karst map indicating presence or absence of karst areas (IGME, 2020b).

3.3 Data set analysis

(Alcides Pereira, Filipa Domingos, University of Coimbra)

Different data sets are available for the three study areas. The data sets differ in basic characteristics as size, sample density, data extent, quality and resolution. Methods to characterize radon priority areas for the two data sets may require adequate data manipulations for different methods. Table 5 gives an overview and comparison of the Austrian and the Cantabrian data set regarding the data density, similarity of data and the origin of data (e.g. measured or derived from literature). In addition detailed analysis of the available data were carried out. In this chapter a summary of descriptive statistics is provided (Table 6, 7, 8) for the three study areas and some box-plot graphs are shown for selected characteristics. The results of a more detailed data analysis with different methods (Kruskal-Wallis test, Spearman rank correlation, variograms) can be found in the appendix. A summary of correlations is given in Chapter 3.4, Table 9.

Table 5: Overview of existing variables in the Cantabrian and the Austrian data set.

Variable	Cantabria	Austria (AUT North and AUT South)
IRC	location approx., low sample density	exact location, high sample density
Soil air Rn	*measured;* sample density similar	*measured*; sample density similar
Act. conc. in soil	European K, Th, U in soil maps (JRC) 10x10 km grid AM/GM (FOREGS, GEMAS)	^{40}K, ^{210}Pb, ^{226}Ra, ^{228}Ra, ^{228}Th, ^{238}U measurements
ADR	*measured*; sample density similar	*measured*; sample density similar
Faults	map; similar	map; similar
Geology	map; similar	map; similar
Permeability	estimates derived from lithological units	Soil permeability *measurements* + estimates derived from soil units
Karst	Binary, derived from lithological units	-
Building characteristics	-	Questionnaire; at location of IRC
Soil map	-	Soil unit, water conditions, soil depth, …
Airborne radiometry	-	eU; measured only AUT North

Austria: North Region (AUT North)

Austria North has the highest indoor radon concentrations of the data set areas. Table 6 summarizes the distributions of the numerical data. Some interesting correlations of numerical and categorical data are shown in Figure 3 – 6. The radon concentration of soil gas is positive correlated with the soil water content (Figure 3) which is in agreement with known effects in the literature (Arvela, H. et. al 2016). Although the geology of AUT north is the most homogenous, minor differences in the ADR can be observed for different rocks (Figure 4). Figure 5 shows that IRC measured in rooms that are earthbound are higher than rooms which are not earthbound. Figure 6 visualises the higher indoor radon concentrations in lower floors.

Table 6: Descriptive statistics of data from region AUT North

Variable	eU	SGR	Perm.	ADR	K-40	Pb-210	Ra-226	Ra-228	Th-228	U-238	IRC AM	IRC
Unit	**(ppm)**	**kBq/m³**	**m²**	**µSv/h**	**Bq/kg**						**Bq/m³**	
Valid N	3732	60	60	60	30	30	30	30	30	30	653	1294
Mean	1.58	104	2.83E-11	0.17	888	47	50	62	63	54	362	361
Geometric mean	n.d.	91	4.40E-12	0.17	842	46	48	57	58	53	232	223
Median	1.42	93	6.72E-12	0.16	800	46	48	57	59	52	226	216
Minimum	-0.85	13	9.00E-15	0.12	465	23	29	28	28	27	12	10
Maximum	7.65	304	1.59E-10	0.24	1630	101	115	138	139	105	2640	2765
Lower quartile	0.81	69	1.52E-12	0.15	657	39	39	43	46	45	118	112
Upper quartile	2.20	127	4.91E-11	0.19	1040	52	58	70	70	61	483	438
Standard deviation	1.01	52	4.62E-11	0.03	303	14	17	27	27	15	371	404
Skewness	0.82	1	1.89E+00	0.54	1	2	2	1	1	2	2	2
Kurtosis	0.81	3	2.52E+00	-0.38	0	7	7	1	1	4	5	7

SGR – Soil gas radon; Perm. – Permeability; ADR – Ambient dose rate; IRC AM – Arithmetic mean of indoor radon concentrations; IRC – Indoor radon concentrations of room 1 and 2 combined

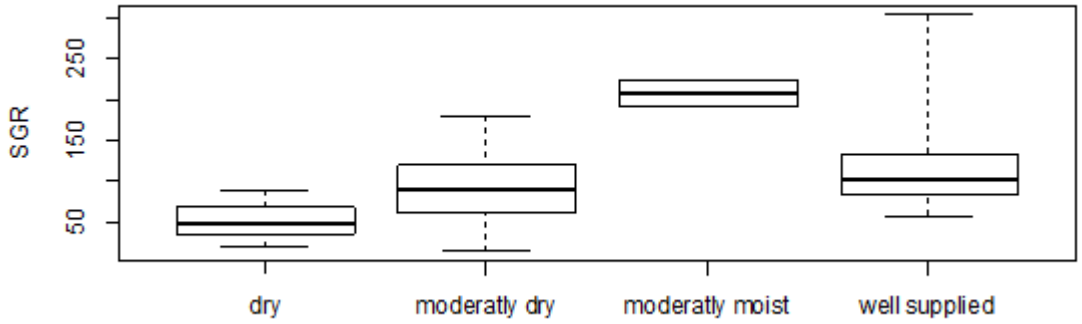

Figure 3: Box-plots of radon concentration in soil gas [kBq / m³] grouped by water content in the region AUT North.

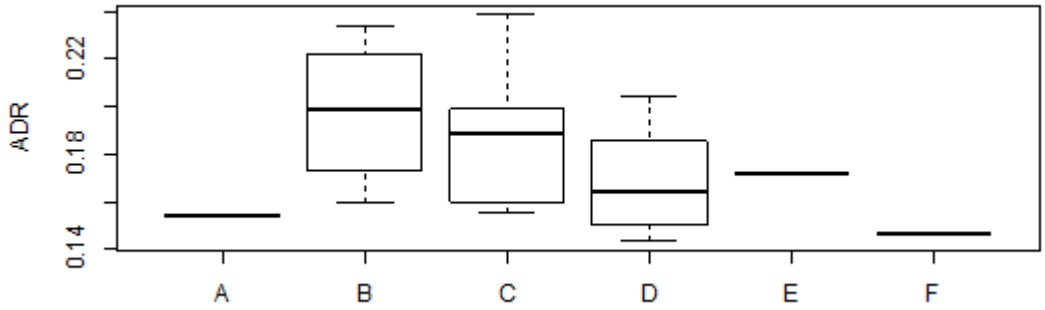

Figure 4: Box-plots of ambient dose rate [µSv/h] grouped by bedrock (fine geology) for region AUT North. A - alkaline to intermediate plutonic rock; B - coarse- to very coarsegrained biotite granite (Weinsberger); C - fine grained two mica granite (Altenberger); D – fine to intermediate grained migmatite (Meta- Diatexite), granodioritic; E – intermixing zone and fluid transition of coarse grained biotite granite and migmatite; F - valley infill.

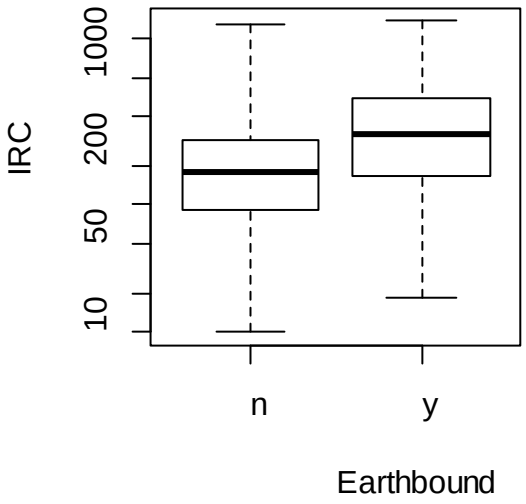

Figure 5: Box-plots of IRC grouped by earthbound (y) and non-earthbound (n) rooms (AUT North). Y-axis in logarithmic scale.

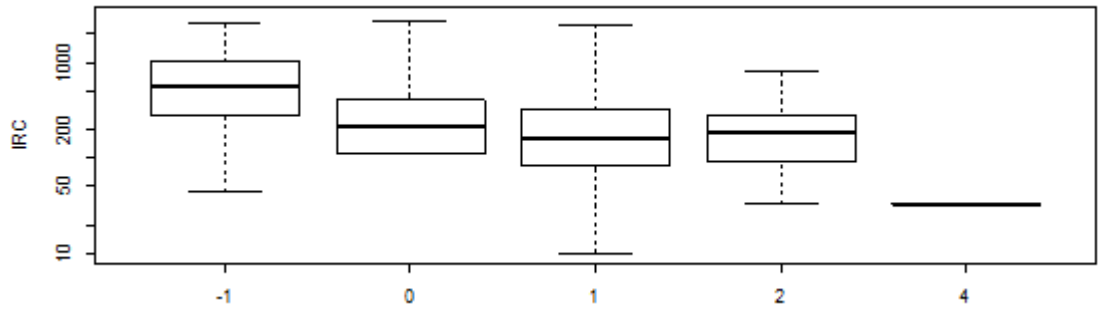

Figure 6: Box-plot of IRC grouped by floor of the measured room (AUT North). Y-axis in logarithmic scale (minimum values of zero are not displayed).

Austria: Southern Region (AUT South)

Austria South has the highest number of indoor radon measurements, but concentrations are lower than in Austria North.

Table **7** summarizes the distributions of the numerical data. AUT South is not the typical radon area in Austria and diverse in geology. The Radium concentrations are slightly lower than in AT North, Uranium concentration is clearly higher. The mean ADR is higher in Austria South and differs stronger among the bedrock types, highest in Orthogneis (see Figure 7).

Table 7: Descriptive statistics of data from region AUT South

Variable		SGR	Perm.	ADR	K-40	Pb-210	Ra-226	Ra-228	Th-228	U-238	IRC AM	IRC
Unit		**kBq/m³**	**m²**	**µSv/h**	**Bq/kg**						**Bq/m³**	
Valid N		88	8.80E+01	88	82	82	82	82	82	82	984	1933
Mean		86	1.22E-11	0.20	664	43	47	36	36	39	247	246
Geometric mean		55	7.51E-12	0.19	552	37	42	32	31	33	152	149
Median		55	1.33E-11	0.20	721	37	40	34	33	32	139	135
Minimum		1	4.20E-14	0.07	14	9	14	7	4	12	16	16
Maximum		953	2.74E-11	0.30	1190	136	168	78	81	140	4655	5218
Lower quartile		34	5.15E-12	0.18	480	25	35	24	24	22	78	76
Upper quartile		98	1.89E-11	0.23	880	48	52	45	46	44	274	275
Standard deviation		115	7.69E-12	0.04	289	27	27	17	16	26	337	351
Skewness		5	-6.22E-02	-0.31	0	2	3	1	1	2	5	6
Kurtosis		38	-1.18E+00	0.45	0	3	8	0	0	5	49	54
Kolmogorov - Smirnov	**D**	**0.2297**	0.0932	0.0951	0.0998	**0.2207**	**0.2070**	0.0681	0.0751	**0.2038**	**0.2472**	**0.2556**
	p	**< 0.01**	> 0.20	> 0.20	> 0.20	**< 0.01**	**< 0.01**	> 0.20	> 0.20	**< 0.01**	**< 0.01**	**< 0.01**
Shapiro-Wilk	**W**	**0.5125**	**0.9490**	0.9743	**0.9630**	**0.8077**	**0.7276**	**0.9573**	0.9787	**0.7758**	**0.5563**	**0.5355**
	p	**< 0.001**	**0.0017**	0.0768	**0.0183**	**< 0.001**	**< 0.001**	**0.0081**	0.1900	**< 0.001**	**< 0.001**	**< 0.001**

SGR – Soil gas radon; Perm. – Permeability; ADR – Ambient dose rate; IRC AM – Arithmetic mean of indoor radon concentrations; IRC – Indoor radon concentrations of room 1 and 2 combined

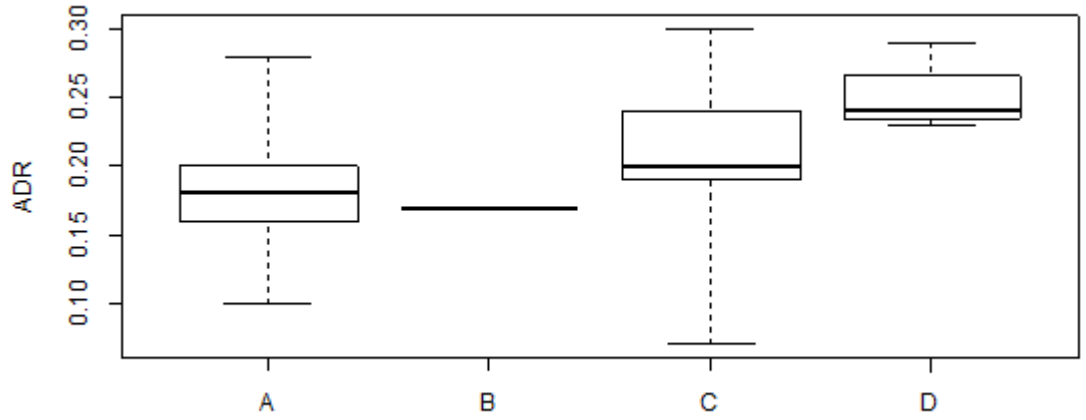

Figure 7: Box-plot of ambient dose rate according to bedrock for region AUT South. A - Carbonate rock, siliciclastics, porphyry (generally metamorphic); B - Marl, sand, gravel, limestone; C - Mica schist, paragneiss; D – Orthogneiss.

Cantabria

For Cantabria dataset a bit different parameters are available than for Austrian data sets, descriptive statistics of the numerical data is summarised in Table 8. For the much larger area, quite few IRC data are available and the IRC concentration is clearly lower than for Austria. Also the mean soil gas radon concentration is clearly below the ones in Austria. Note that the units are different for the Austrian data for K, TH, U, ADR.

Figure 8 and Figure 9 shows the ADR and the SGR by sources (carbonate, detritic, detritic quaternary, meta detritic, volcanic). It is interesting, that the ADR and the SGR have opposite distribution according to their source - e.g. carbonate origin has lowest ADR but highest SGR. Figure **10** visualises the higher mean indoor radon concentration in areas with karst.

Table 8: Descriptive statistics of data from region Cantabria

	IRC (Bq/m^3)	SGR (kBq/m^3)	K_2O* (%)	Th* (ppm)	U* (ppm)	ADR* (mR/h)
Valid N	482	238	70	70	70	77
Mean	97	23.7	1.7	7.5	1.7	7.1
Geometric mean	55	10.3	1.7	7.4	1.7	7.0
Median	54	14.0	1.7	7.2	1.7	6.9
Minimum	6	0.1	1.1	5.8	1.4	4.9
Maximum	2895	209.2	2.6	10.9	2.0	10.7
Lower quartile	29	4.5	1.4	6.8	1.6	6.4
Upper quartile	93	32.4	2.0	7.7	1.8	7.6
Standard deviation	221	29.0	0.4	1.1	0.1	0.9
Skewness	9	2.8	0.5	1.3	0.1	0.9
Kurtosis	86	10.8	-0.5	1.8	-0.7	2.6
IRC – Indoor radon concentrations; SGR – Soil gas radon; ADR – Ambient dose rate. *Zero values were excluded.						

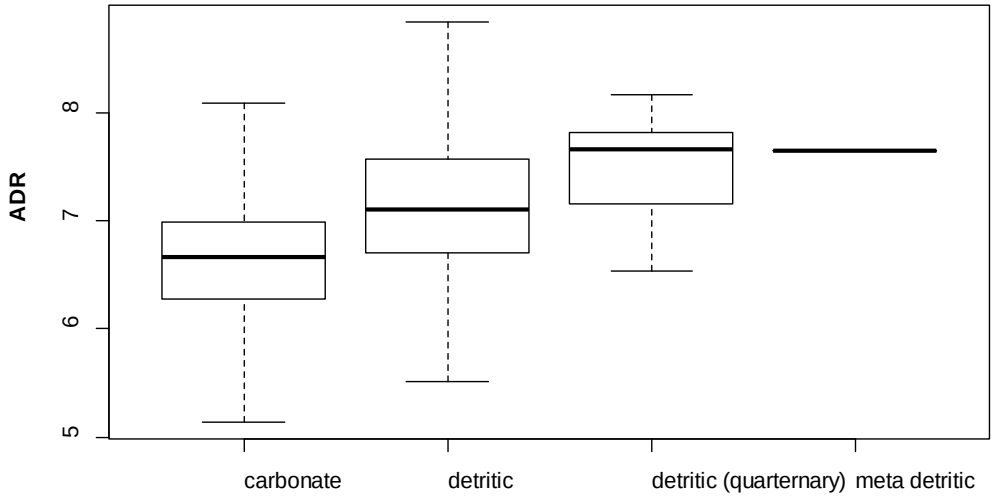

Figure 8: Box-plot of the ambient dose rate by geological source for the Cantabria region.

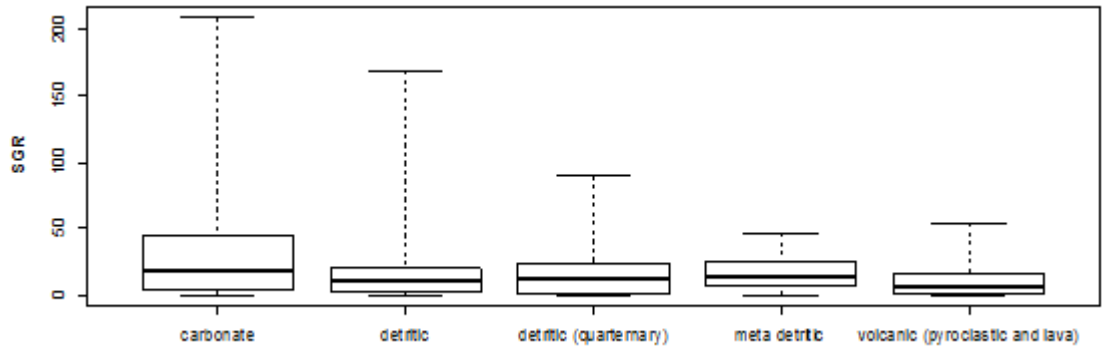

Figure 9: Box-plot of Soil gas radon by geological source for the Cantabria region.

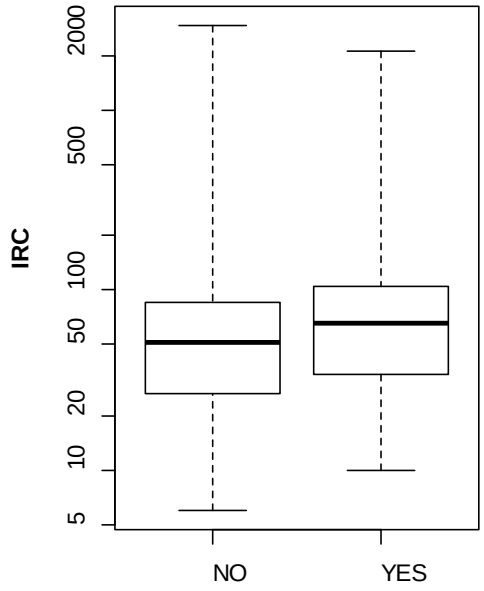

Figure 10: Box-plot of indoor radon concentration by karst for the Cantabria region (yes - karst present, no - no karst present). Y-axis in logarithmic scale.

3.4 Summary of data sets

A sound data basis is required for the delineation of radon priority areas (RPAs). The provided data sets include a variety of data from different sources and may be of interest for different concepts to identify RPAs. Radon data usually is noisy and incomplete and the same is true for the data set of this exercise. The analysis of the exercise data shows that grouping of populations, the type of correlation and the rate of spatial correlation of the same variables are not equal in different regions (Table 9).

This is especially interesting for the Austrian data set, where the measurement methods, the geogenic maps and the sampling density are comparable or even identical for both regions. For example, in the region AUT North different groups of geogenic variables show significantly different IRC, which cannot be observed in the region AUT South (see Annex, Table 27, Table 31). Furthermore, when analysing both regions together, the number of geogenic variables that show significant different IRCs increases, because the variability of IRC increases as well if considering both regions. On the other hand, both regions show similarities, such as the lack of statistical differences of the radionuclide content according to soil types or the lack of spatial correlation of almost all geogenic factors.

However, the next chapter will focus on different methods to identify RPAs and the approach how to deal with the data set. It will be interesting to see, which data are used by the different methods applied and how the data will be edited and manipulated to serve as a solid basis for the identification of RPAs.

Table 9: Data set grouped in populations, type of correlation and rate of spatial correlation in different regions.

AUT North			
	significant difference among groups	**significant correlation**	**spatial correlation**
ADR	bedrock (fine), soil map (source)	K-40, Th-228, Ra-228, TGDR	weak
eU	bedrock (fine), soil map (source)	x	strong
soil gas	soil map (type, grain size, water content)	U-238	weak
Pb-210	bedrock (coarse)	U-238, Ra-226	no
Ra-226	x	U-238, Ra-226	no
U-238	x	soil gas, Ra-226, Pb-210	no
TGDR	x	ADR, K-40, Ra-228, Th-228	no
IRC	permeability, bedrock (fine), soil map (water content), building characteristics (RT, EB, B, BT, FO,FL)	x	weak
AUT South			
ADR	bedrock	soil gas, Ra-226, TGDR	weak
soil gas	x	ADR, K-40, Pb-210, Ra-226, U-238, TGDR	no
Pb-210	x	soil gas, K-40, Ra-226, Pb-210, Ra-228, U-238, TGDR	no
Ra-226	x	soil gas, ADR, K-40, Ra-226, K-40, Pb-210, Ra-228, Th-228, U-238, TGDR	no
Ra-228	soil map (source, grain size)	K-40, Ra-226, Th-228, U-238, TGDR	no
Th-228	soil map (source, grain size)	Ra-226, Ra-228, U-238, TGDR	no
U-238	x	soil gas, K-40, Pb-210, Ra-226, Ra-228, Th-228, TGDR	no
TGDR	x	soil gas, ADR, K-40, Ra-226, K-40, Pb-210, Ra-228, Th-228, U-238, TGDR	weak
IRC	building characteristics (RT, EB, B, BT, FO,FL)	x	no
Cantabria			
ADR	lithology, source, permeability	soil gas (-), Th, K	strong
Soil gas	lithology, source, permeability	IRC, ADR (-), U (-)	no
IRC	lithology, karst	soil gas, U (-)	no

4. Exercise methods and results

The interplay of available data and the intended method usually defines the applied strategy for the delineation of RPAs. This is different in this exercise, because the available data is predefined. The comprehensive radon data sets provided in this exercise aim to be a solid basis for different strategies to identify RPAs. However, as already mentioned in the previous chapter, there is room for improvement regarding data quality. This chapter presents the methods applied for the identification of RPAs with the current data sets.

4.1 Basic analysis based on indoor radon data

(Sebastian Baumann, AGES)

The definition of RPA utilising IRC data commonly follows two basic concepts: a) The average IRC (e.g. AM, GM) of the area is compared to a threshold (e.g. 300 Bq/m³) and b) the percentage of measurements exceeding a threshold in an area is compared to a percentage threshold (e.g. 10 %). Common approaches to define radon priority areas use IRC thresholds of 100 to 300 Bq/m³ and percentage thresholds of 1 to 30 percent (see chapter 2.1. and Table **2**).

The IRC distributions differ in the regions of the exercise data sets and the concentrations are considerable higher in Austria than Cantabria (Figure 11). This is of course also true for the aggregates of the distributions that might be used for basic radon risk prediction.

Table 10, Table 11 and Table 12 give examples of descriptive statistics of the IRC in different regions and geological units, along with the number and density of the radon measurements. The spatial distribution of the IRC and geology for the different regions is shown in Figure 2, Figure 12 and Figure 13. For the Austrian data set a majority of households in the municipalities have been tested and therefore the spatial pattern of measurements mimics the distribution of overall households. The high measurement density in the Austrian regions show that low and high IRC can occur virtually in neighbouring houses, which seems – at a first glance - to contradict the concept of geogenic radon potential. However, one also needs to take into account that beside the geogenic factors also the type of building or the location of the measurement in the building (floor) largely influences the IRC. Therefore, using subsets of data e.g. by floors is a common approach to achieve comparable IRC and often is applied for radon risk prediction. The measurement density of the Cantabrian data set is much lower compared to the Austrian regions and radon risk prediction solely on the basis of IRC might show high uncertainties.

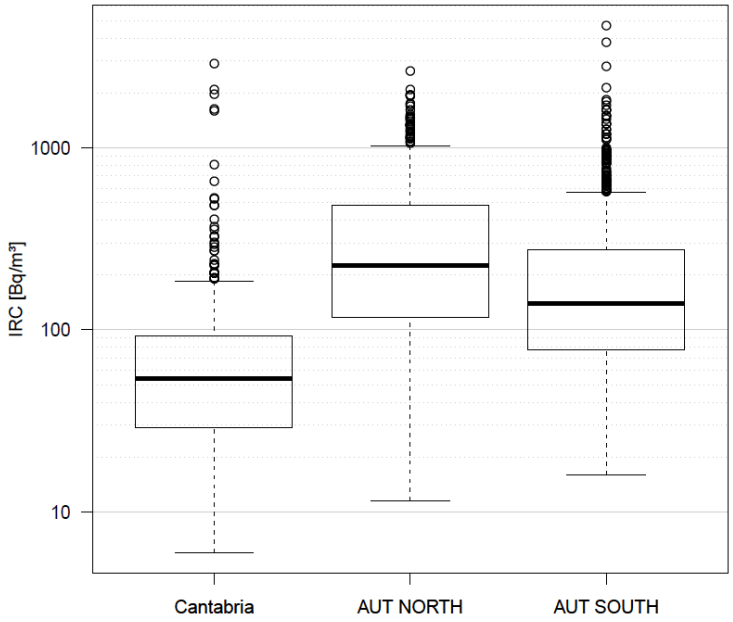

Figure 11: Boxplot shows IRC distribution in log scale for the different regions of the exercise data.

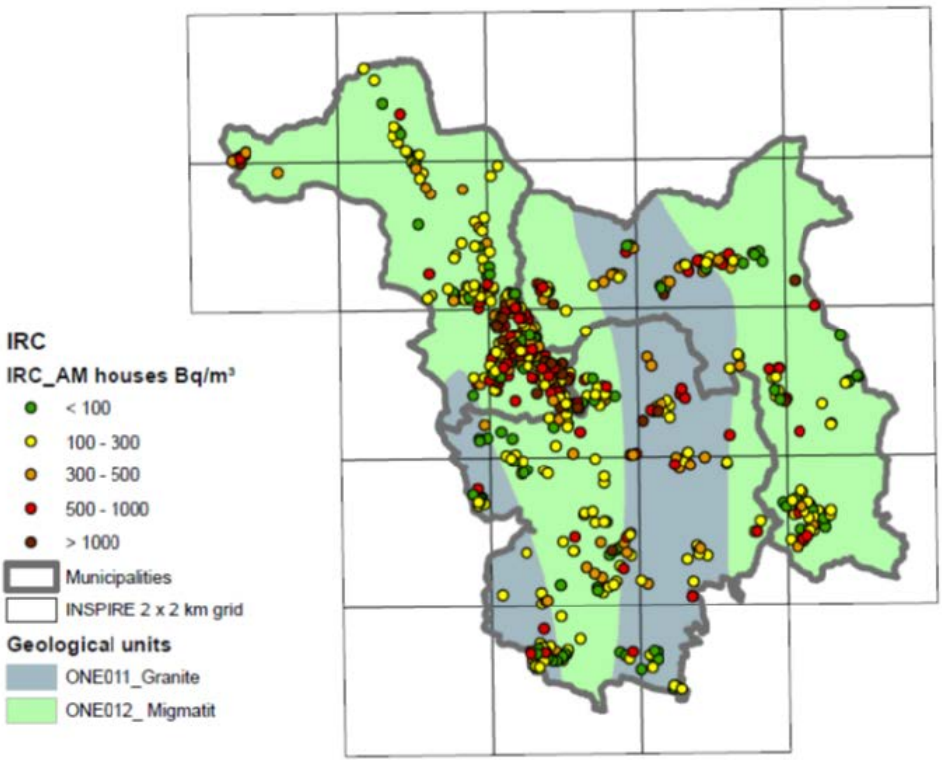

Figure 12: Indoor radon concentration (IRC) measurements and geological units in AUT North.

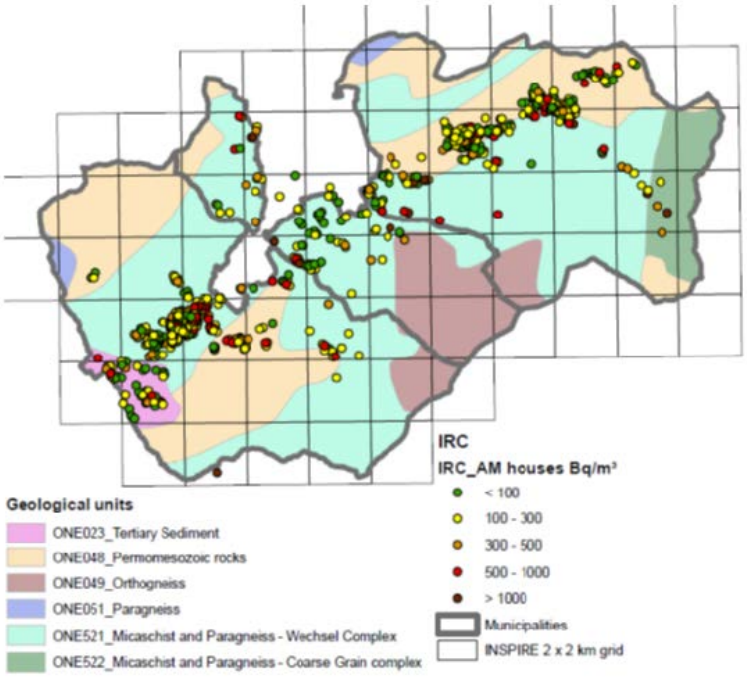

Figure 13: Indoor radon concentration (IRC) measurements and geological units in AUT South.

Table 10: Descriptive statistics per municipalities in Austria and the study region in Spain.

Country	**Austria**	**Austria**	**Austria**	**Austria**	**Austria**	**Austria**	**Spain**
Region	North	North	North	South	South	South	Cantabria
Municipality ID	1	2	3	4	5	6	-
Area [km²]	15	13	10	32	76	73	5328
Number of measurements (dwellings)	200	138	315	88	469	421	482
IRC samples per km²	13.7	10.5	32.9	2.7	6.2	5.8	0.09
Arithmetic mean (Bq/m³)	289	313	429	289	251	234	97
Geometric Mean (Bq/m³)	196	207	273	165	157	146	55
Median (Bq/m³)	197	213	266	168	144	130	54
% > 100 Bq/m³	76	77	83	64	65	61	22
% >200 Bq/m³	49	52	60	45	37	32	7
% > 300 Bq/m³	31	36	45	28	22	21	3

Table 11: Descriptive Statistics per geological unit in Austria (according to Figure 12 and Figure 13)

Geology unit	Granite	Migmatite	Tertiary sediment	Permo-mesozoic rocks	Ortho-gneiss	Micaschist and Paragneiss (coarse grain complex)	Micaschist and Paragneiss (Wechsel complex)
Area [km²]	12	26	4	52	17	105	9
number of measurements (dwellings)	143	510	56	356	1	561	4
IRC samples per km²	12	20	14	7	0.1	5	0.4
Arithmetic mean (Bq/m³)	352	364	210	210	455	266	1159
Geometric Mean (Bq/m³)	229	233	136	140	455	161	494
Median (Bq/m³)	229	225	134	130	455	145	355
% > 300 Bq/m³	41	38	20	21	-	24	-
% > 200 Bq/m³	58	54	32	33	-	37	-
% > 100 Bq/m³	78	81	63	60	-	66	-

Table 12: Descriptive Statistics per geological unit in Cantabria. Only units with more or equal to 20 IRC summarized.

Geology unit	Reef lime-stones	Gravel, sand, silt	shale, sandstone, conglomerate and sandy limestone	slate, lutite, sandstone, coal and limestone	motley clay and gypsum	sandstone lutite, marl	marl, limestone and loamy limestone
Area [km²]	**6915**	**2554**	**13625**	**3649**	**1036**	**3060**	**2337**
number of measurements (dwellings)	88	81	74	29	22	22	20
IRC samples per km²	0.012	0.032	0.005	0.008	0.021	0.007	0.009
Arithmetic mean (Bq/m³)	167	75	52	60	83	197	72
Geometric Mean (Bq/m³)	63	55	40	43	59	59	48
Median (Bq/m³)	47	62	44	47	59	62	40
% > 300 Bq/m³	8	1	0	0	0	5	5
% >200 Bq/m³	14	4	1	7	14	5	10
% > 100 Bq/m³	32	21	9	7	18	23	20

4.2 Generalized additive mixed model (GAMM)

(Christian Laubichler, Oliver Alber, AGES Graz)

The method applied in this chapter is based on the methodology used in Austria for the delineation of radon areas but taking into account more available variables than in the methodology for Austria. The method is applied in this exercise for the Austrian and Cantabria data sets.

Based on IRC measured in Austria and Cantabria, the goal is to

- identify relevant explanatory variables,
- predict the expected indoor radon concentration for a specified grid,
- assess the variability of predictions;

The IRC in dependency of explanatory variables will be estimated with the generalized additive mixed model (GAMM). The results of the final model will be used to predict the expected IRC and to calculate confidence intervals.

Subsequent modelling and analyses are carried out with the statistical programming language R version 3.5.1, using the packages gamm4 and mgcv.

Statistical Models

The additive mixed model

$$\log(IRC_{ij}) = \beta_0 + \beta_1 z_{ij} + \cdots + \beta_m z_{ij} + s(x_j, y_j) + u_j + \epsilon_{ij} \qquad (Equation\ 1)$$

$$u_j \sim N(0, \sigma^2_{house}), \quad j = 1, \ldots, n_{house}$$

$$\epsilon_{ij} \sim N(0, \sigma^2_{\epsilon})$$

is fitted to the Austrian data set, whereby the living unit u_j is taken as a random effect, thus introducing a positive correlation of measurements within the same living unit. A slightly different model, an additive model without random effects, is used for the Cantabrian data set. The Cantabrian data set qualifies as multilevel data, as the data contains multiple measurements within each location. However, introducing a random effect for location, thereby assuming a positive correlation within a location is not feasible in this case. In Cantabria measurements from a relatively large area are assigned to a particular location. Influencing factors, such as geology, etc., in such an area could be inherently different, which would contradict the positive correlation induced by the random effect.

Resulting in the additive model:

$$\log(IRC_{ij}) = \beta_0 + \beta_1 z_{ij} + \cdots + \beta_m z_{ij} + s(x_j, y_j) + \epsilon_{ij} \qquad (Equation\ 2)$$

$$\epsilon_{ij} \sim N(0, \sigma^2_{\epsilon}), \quad j = 1, \ldots, n_{location}$$

In both cases, the smooth functions $s(\,.\,)$ pertain to the class of thin plate regression splines. The z_{ij} terms represent explanatory variables and the pair (x_j, y_j) represents the coordinates of a living unit or location j.

The final model should only contain variables that show a significant influence on log(IRC). To identify these variables, a stepwise forward selection using 5-fold cross validation is applied. All available variables were used in the stepwise forward selection.

Stepwise Forward Selection

Starting with the simplest models

$$\log(IRC_{ij}) = \beta_0 + s(x_j, y_j) + u_j + \epsilon_{ij} \quad (Austria) \qquad (Equation\ 3)$$

and

$$\log(IRC_{ij}) = \beta_0 + s(x_j, y_j) + \epsilon_{ij} \quad (Cantabria) \qquad (Equation\ 4)$$

variables are added one after the other. Those variables with the highest explanatory power are chosen for the final model.

The explanatory power is determined by a 5-fold cross validation (CV) for each step, splitting the data into five blocks, whereby four blocks are used to fit the model and the fifth block serves as testing data. For measuring the error of the fitted values compared to the test block, the mean squared error of actual IRC and fitted IRC is used.

Non-relevant variables result in non-significant improvements in cross validation error. The following figures (Figure 14, Figure 15) show the differences in cross validation errors by adding variables.

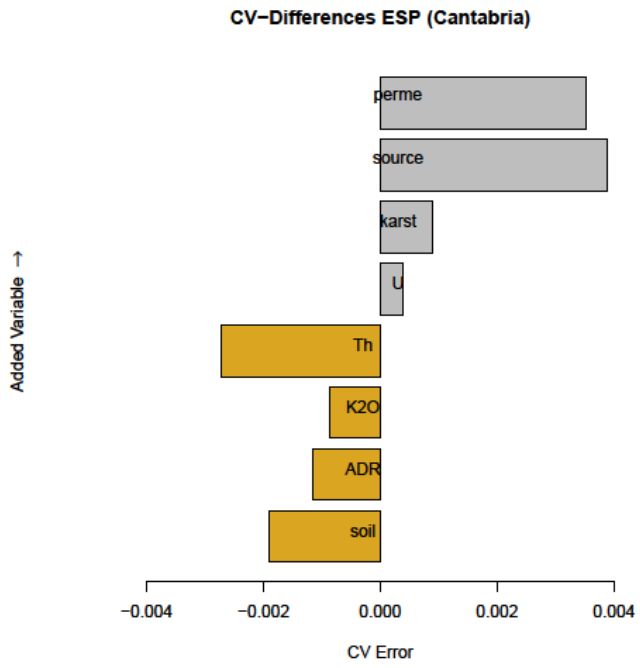

Figure 14: Difference in cross validation errors in Cantabria. Gray colored variables are not included in the model.

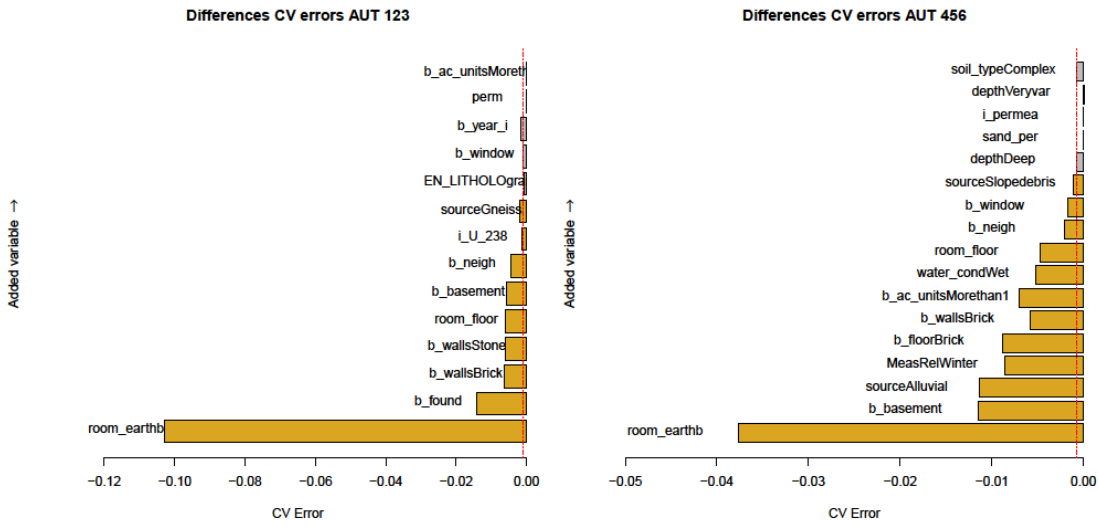

Figure 15: Differences in cross validation errors in Austria (left: AT North, right AT South). Gray colored variables are not included in the model.

Variables are only added to the final model, when the reduced cross validation error is deemed to be significant (Figure 14 and Figure 15). For Cantabria (Figure 14) the variable "Th" (Thorium activity concentration in soil) reduces the CV error significantly and the variables following "Th" increase the CV error, thus all variables up to "Th" are included in the final model (see Table 13). In Austria, Figure 15, the red striped vertical line represents the first quartile of differences in cross validation errors. The final model contains all variables which result in a reduction in CV error that is above the red line. For the Austrian data set, these are mainly building characteristics and geology units ("source"), for AUT South also water content is relevant (see Table 14 and Table 15).

The optimal basis dimension k for the smooth functions $s(.)$ was obtained before the stepwise forward selection. By using cross validation it became apparent that the optimal basis dimension k is 90 and 100 for Cantabria and Austria, respectively.

Final Models

Table 13, Table 14 and Table 15 contain the results of the final models in both study areas, Cantabria and Austria. The final models are fitted using the variables and k from the cross validation.

Table 13: Estimated coefficients of the final model for Cantabria (Soil: radon concentration in soil gas, ADR: ambient dose rate, K2O: potassium activity concentration in soil, Th: Thorium activity concentration in soil).

	$\hat{\beta}_i$	Std.Error	p-value	
Intercept ($\hat{\beta}_0$)	2.313	1.758	0.189	
Soil	0.019	0.008	0.020	*
ADR	0.273	0.245	0.266	
K2O	-1.320	0.673	0.050	.
Th	0.240	0.143	0.094	.

*** significant at $\alpha = 0.001$; ** significant at $\alpha = 0.01$

* significant at $\alpha = 0.05$; . significant at $\alpha = 0.10$

Table 14: Estimated coefficients of the final model for AUT North (municipalities 1, 2 and 3) – the building characteristics variables (b_) are explained in Table **3**, the soil data set variables (i_) in Table 4, source: geological unit) .

	$\hat{\beta}_i$	Std.Error	p-value	
Intercept ($\hat{\beta}_0$)	4.181	0.589	0.000	***
b_room_earthb	0.255	0.079	0.001	**
b_found_foundation partly	-0.144	0.176	0.415	
b_found_no foundation	0.129	0.177	0466	
b_found_strip foundation	0.101	0.091	0.265	
b_walls_Brick	0.379	0.107	0.000	***
b_walls_Stone	0.432	0.165	0.009	**
b_room_floor0	-0.406	0.098	0.000	***
b_room_floor1	-0.614	0.129	0.000	***
b_basement_no	0.085	0.144	0.554	
b_basement_partly	0.202	0.101	0.045	*
b_neigh_solitary	0.526	0.200	0.009	**
i_U_238	0.019	0.009	0.036	*
source_Gneiss	-0.508	0.215	0.019	**

*** significant at $\alpha = 0.001$; ** significant at $\alpha = 0.01$

* significant at $\alpha = 0.05$; . significant at $\alpha = 0.10$

Table 15: Estimated coefficients of the final model for AUT South (municipalities 4, 5 and 6) – the building characteristics variables (b_) are explained in Table 3, the soil data set variables (i_) in Table 4, source: geological unit, water_: water content data from soil map).

	$\hat{\beta}_i$	Std.Error	p-value	
Intercept ($\hat{\beta}_0$)	3.820	0.555	0.000	***
b_room_earthb	0.307	0.086	0.000	***
b_basementno	0.347	0.114	0.002	**
b_basementpartly	0.591	0.094	0.000	***
Source_Alluvial	0.649	0.129	0.000	***
Meas_RelWinter	1.167	0.893	0.191	
b_floorBrick1	0.911	0.271	0.000	***
b_wallsBrick1	0.222	0.073	0.002	**
b_ac_units_Morethan11	-0.189	0.062	0.002	**
water_cond_Wet1	0.523	0.171	0.002	**
b_room_floor0	-0.383	0.122	0.002	**
b_room_floor1	-0.558	0.130	0.000	***
b_neigh_solitary	0.183	0.100	0.067	.
b_window_very tight	0.222	0.092	0.016	*
source_Slopedebris	0.247	0.171	0.151	

*** significant at $\alpha = 0.001$; ** significant at $\alpha = 0.01$

* significant at $\alpha = 0.05$; . significant at $\alpha = 0.10$

Prediction of IRC

The final model is used to predict log(IRC) for a specified grid. In Austria, it is of interest to set a particular house of reference; i.e. selecting levels of the categorical explanatory variables to represent an Austrian house. For example, a standard house can be described by selecting those levels that occur most frequently. This house of reference is then assigned to the coordinates of the midpoints of the grid-cells. In Cantabria, predictions are

based on the midpoints of the grid-cells and the specific values of the explanatory variables associated with that grid.

In the case of Austria, the focus is on municipalities when predicting IRC. Predictions for municipalities can be calculated by averaging over those grid-cells that are allocated in the certain municipality.

To obtain a final IRC prediction, either for a grid-cell or a municipality, the prediction must be converted into Bq/m³ with the following equation:

$$\widehat{IRC}_{\frac{cell}{municipality}} = \exp\left(\widehat{log(IRC)}_{\frac{cell}{municipality}} + \frac{\sigma^2_{house} + \sigma^2_{\epsilon}}{2}\right) \qquad (Equation\ 5)$$

or

$$\widehat{IRC}_{cell/municipality} = \exp\left(\widehat{log(IRC)}_{cell/municipality} + \frac{\sigma^2_{\epsilon}}{2}\right) \qquad (Equation\ 6)$$

for Cantabria.

Figure 16 shows the prediction of the IRC in grid cells (10x10 km). Figure 17 and Figure 18 show the prediction of IRC in grid cells (2x2 km) and per municipality.

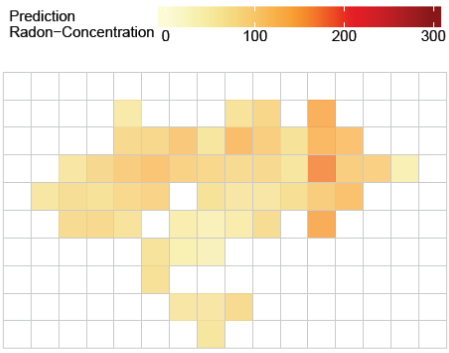

Figure 16: Prediction of radon concentration [Bq/m³] of cell midpoints in Cantabria.

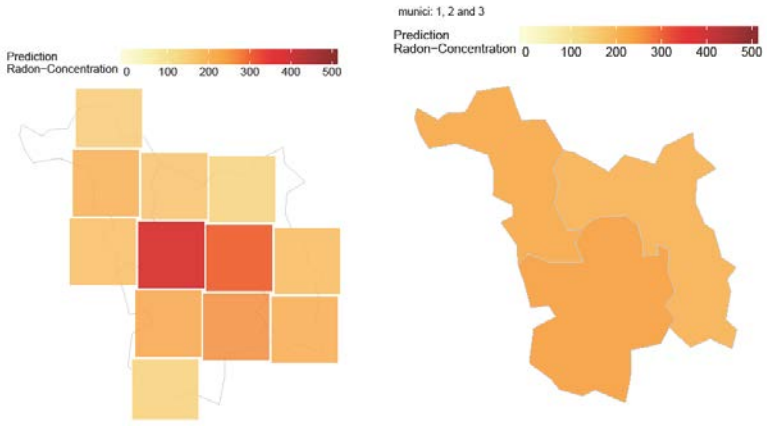

Figure 17: Prediction of radon concentration [Bq/m³] of cell midpoints and municipalities for AUT North (municipalities 1, 2 and 3).

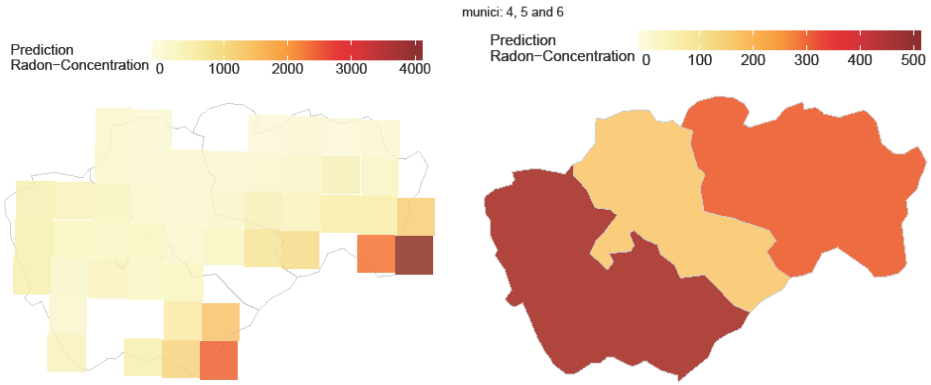

Figure 18: Prediction of radon concentration [Bq/m³] of cell midpoints and municipalities for AUT South (municipalities 4, 5 and 6).

Confidence intervals

To validate pre-defined IRC boundaries or to get an idea about the variation of predictions, confidence intervals can be of interest. Using the distributional results of estimators from the GAMM (generalized additive mixed model) theory, variances of predictions or variances of linear combinations of predictions can be obtained. Confidence intervals, at a specified level of significance ($\alpha = 0.05$), can be calculated. Figure 19 shows the confidence intervals for the grid cell prediction of the radon concentration of Cantabria, Figure 20 indicates the confidence intervals for the predicted radon concentration in the six Austrian municipalities.

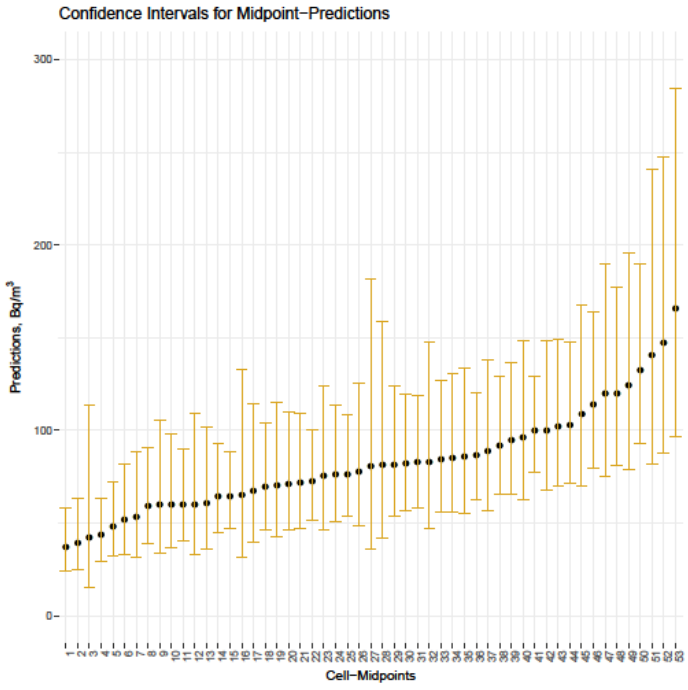

Figure 19: Confidence intervals for grid-cell midpoint predictions in Cantabria.

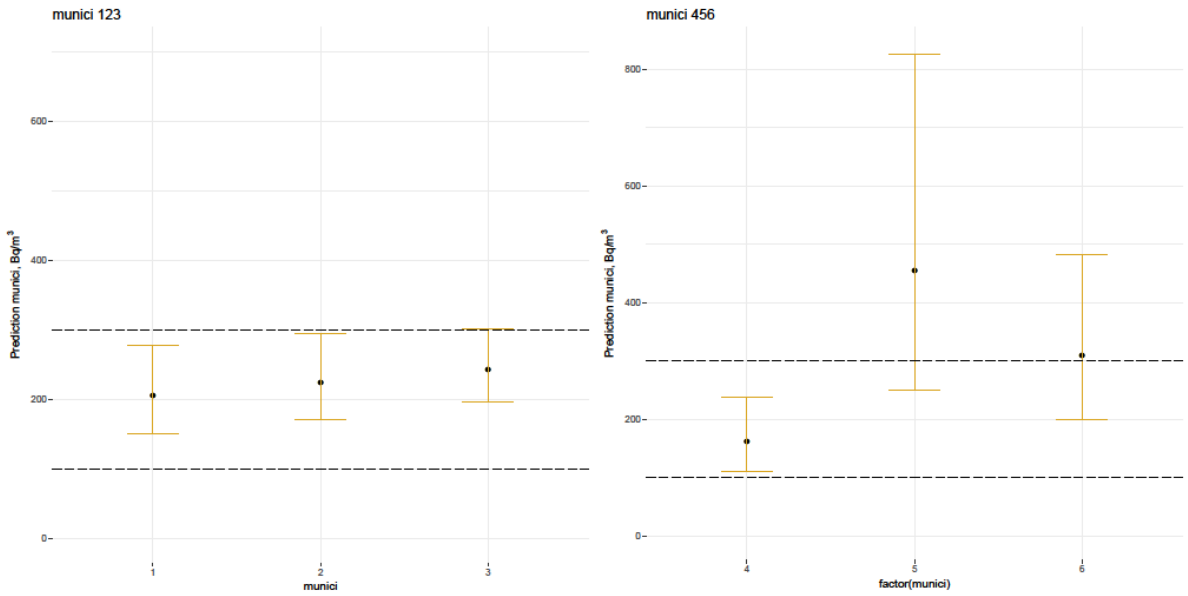

Figure 20: Confidence intervals for municipality predictions in Austria, AUT North on the left and AUT South on the right side.

4.3 Empirical Bayesian Kriging (EBK) Regression Prediction

(Giancarlo Ciotoli, Italian National Research Council)

Methodology

The Empirical Bayesian Kriging Regression Prediction is a geostatistical interpolation method that uses Empirical Bayesian Kriging (EBK) with known explanatory variable rasters to affect the value of the data that should be interpolated. This approach combines kriging with regression analysis to make predictions that are more accurate than either regression or kriging can provide on their own. More details about EBK can be found here (ESRI, 2020).

The method EBK Regression Prediction was used to generate a radon soil gas map of Cantabria. Therefore, only data from Spain (Cantabrian data set) were used in this prediction method.

The estimation by EBK Regression Prediction uses radon concentration in soil gas as response variable and raster layers of the available parameters (permeability, ambient dose rate, K-40, U-238, Th-232, fault density, presence of karst areas) as proxies (see Figure 21). The proxies were interpolated on raster cells with a resolution of 500 x 500 m by the following operations:

1. Application of the “Spatial Join” tool between the lithology layer and the geochemical soil data in order to assign average values of measured K-40, U-238, Th-232 to the lithology layer. Afterwards, the “Polygon to Raster” transformation tool was applied to obtain raster layers of the proxies.
2. Application of the kernel density algorithm to obtain a raster of the density map of faults.
3. Application of ordinary kriging to obtain a raster map of the estimated ambient dose rate.
4. Merging of the karst layer and the region boundary to obtain two polygons: karst area and no karst area. The polygons were classified by using a binary code 1 (karst area) and 0 (no karst area). The layer was transformed into a raster by using the “Polygon to Raster” transformation tool.

5. A permeability value was assigned to the lithology layer according to the qualitative description reported in the attribute table: very high (10^{-10}), high (10^{-11}), medium (10^{-12}), low (10^{-14}), very low (10^{-16}). Logarithmic values were considered to simplify calculations.
6. Application of the EBK Regression Prediction algorithm in ArcGIS Pro environment, considering the soil radon concentration as response variable and U-238, Th-232, K-40, fault, permeability, ambient dose rate and karst as independent variables.

EBK Regression Prediction results

The Root Mean Square Standardised Error indicates that EBK Regression Prediction has a good performance, though the cross-correlation graph (Figure 22) shows an underestimation of the highest values.

The map of the geogenic radon potential (Figure 23) for the Cantabrian region, obtained by using EBK Regression Prediction, suggests that mainly faulted areas and zones of high permeability affect the radon distribution in soil air. To group the areas of the geogenic radon potential map into radon priority areas, according to their concentration, the following classification is recommended:

Low radon risk area: concentration < 15 kBq/m³

Medium radon risk area: concentration between 15 – 60 kBq/m³

High radon risk area: concentration > 60 kBq/m³

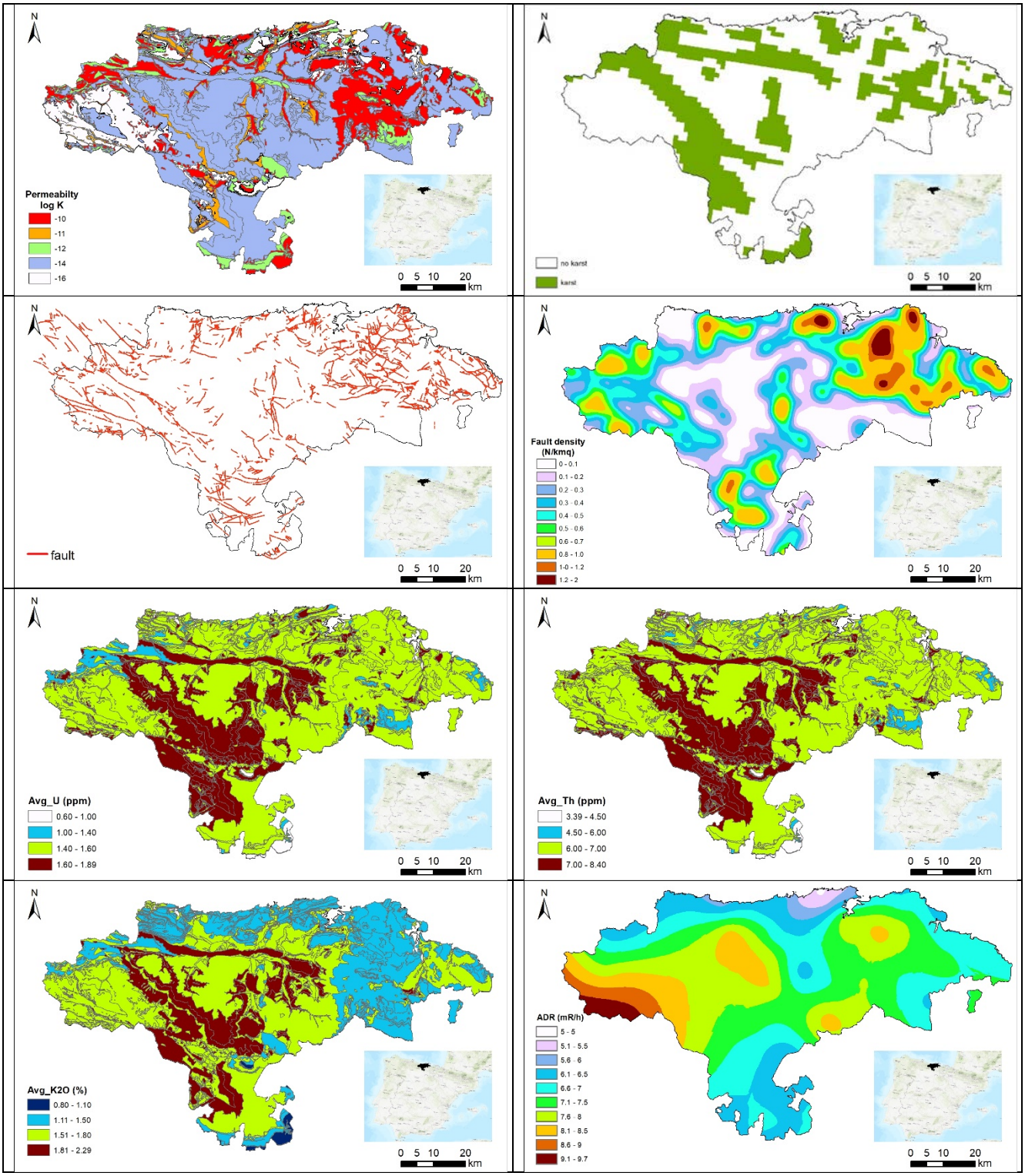

Figure 21: Maps of the available proxy variables for Cantabria

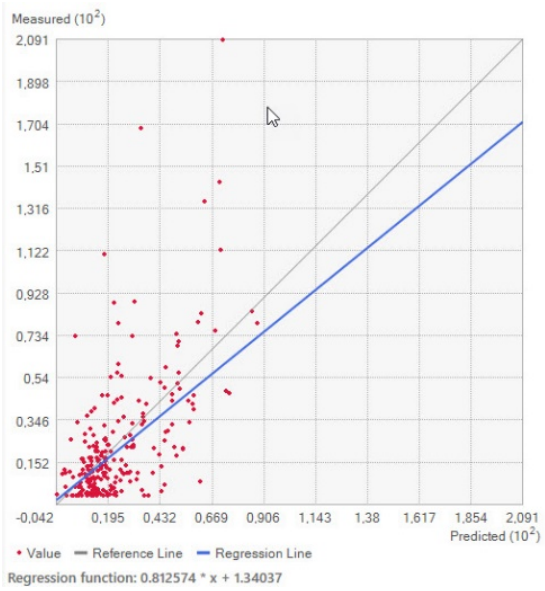

Figure 22: Cross validation graph: Measured values (y-axis) versus predicted values (x-axis)

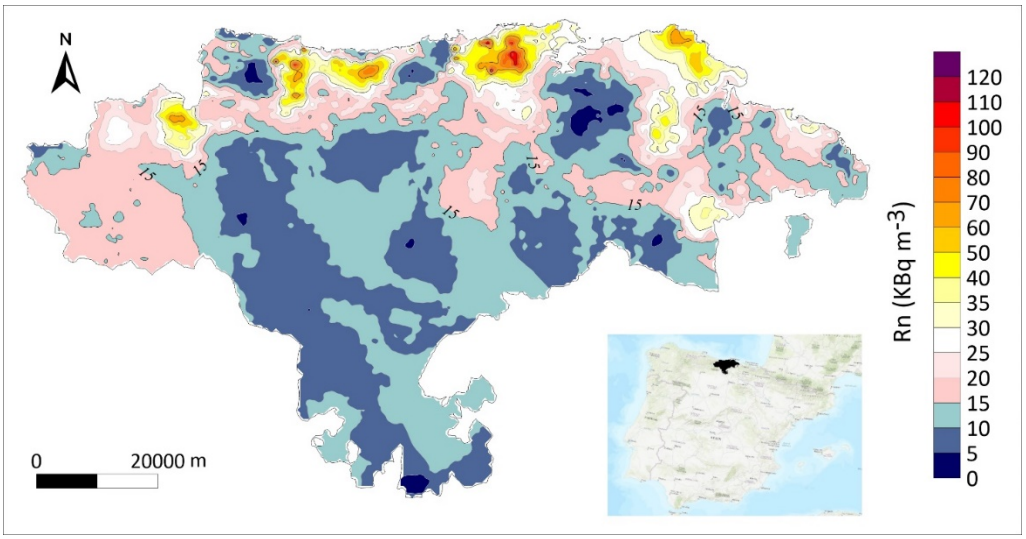

Figure 23: Geogenic radon potential map of the Cantabrian region calculated by using Empirical Bayesian Kriging Regression.

4.4 Ordinary Kriging (OK) and Indicator Kriging (IK)

(Eric Petermann, Peter Bossew, BfS - Bundesamt für Strahlenschutz)

This method describes the prediction of the indoor radon concentration in areas by using the Kriging method. This method is applied for both, the Austrian and the Cantabrian data set.

Austria

Analysis of variance (ANOVA) was performed for the following target variables: indoor radon (only ground floor measurements were considered), soil radon, soil permeability and the geogenic radon potential after Neznal et al. (2004) derived thereof (a function of soil radon and soil permeability). ANOVA revealed significant ($p < 0.05$) differences for the target variables dependent on pedological and geological characteristics such as source

material (source) and soil thickness (depth) as well as texture (EN_LEG_TEX) and lithology (EN_LITHOLO), respectively. However, differences between pedological and geological characteristics – although being significant – are not very prominent. Under consideration of the high density of indoor radon measurements in populated areas (897 measurements for AUT North; 1191 measurements for AUT South), a pure geostatistical approach using ordinary kriging and indicator kriging without any additional predictor seemed to be sufficient to estimate the radon risk for populated areas.

Spatial prediction of indoor radon concentration

The software R was used to execute ordinary kriging (Package „gstat“, function „krige“). First, the spatial autocorrelation of indoor radon concentration was tested by calculating variograms. For both Austrian areas, separate variograms were calculated (Figure 24 and Figure 25).

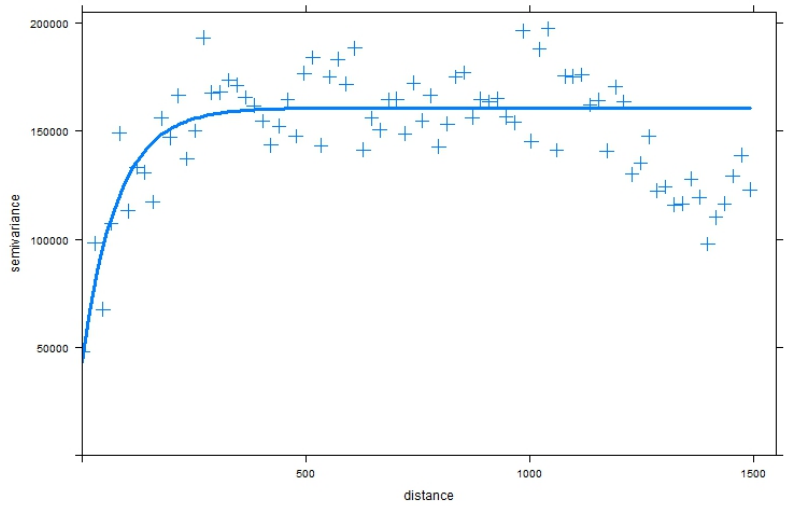

Figure 24: Variogram of ground floor indoor radon concentration for AUT North. Empirical data (crosses), fitted model (solid line).

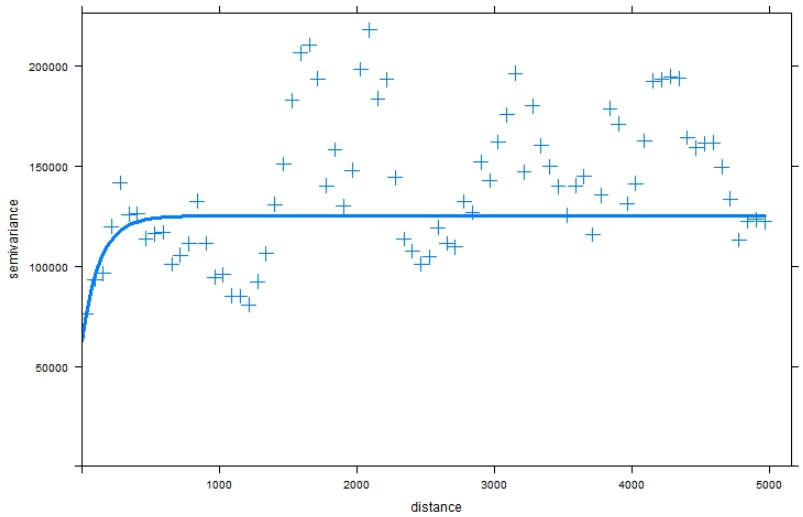

Figure 25: Variogram of ground floor indoor radon concentration for AUT South. Empirical data (crosses), fitted model (solid line).

The variograms of both Austrian areas (Figure 24 and Figure 25) reveal spatial autocorrelation over a short range of approximately 500 m. The nugget effect (local variability at distance 0) is large for both areas, but even more striking for AUT South. Both areas were tested for anisotropy of spatial autocorrelation by visual inspection of directional sample variograms for directions North, East, South and West as well as for directions North-East, South-East, South-West and North-West. Since anisotropy could not be identified, spatial autocorrelation was assumed to be isotropic. For both areas, an exponential model was fitted to the empirical variogram data (Table 16).

Table 16: Variogram model parameters for indoor radon concentration at ground floor.

	AUT North	**AUT South**
Nugget	43136	62345
Range parameter for exponential model	79	134
Partial Sill	117485	62775

Based on these variogram models and the empirical data, indoor radon concentration was kriged (Figure 26 and Figure 27) for a raster cell size of 200 m. Due to the low range of spatial autocorrelation, the estimates at large distances from the nearest observation (> 1 km) are equivalent to the mean of the whole area.

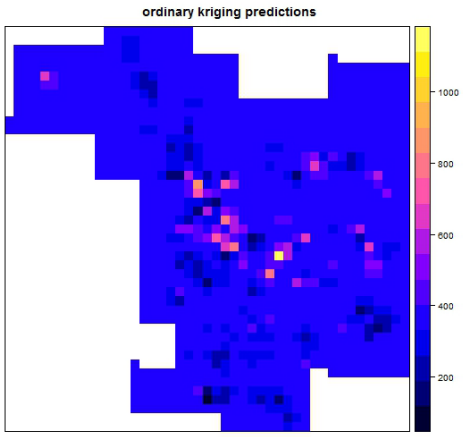

Figure 26: Estimated indoor radon concentration at ground floor based on Ordinary Kriging in AUT North.

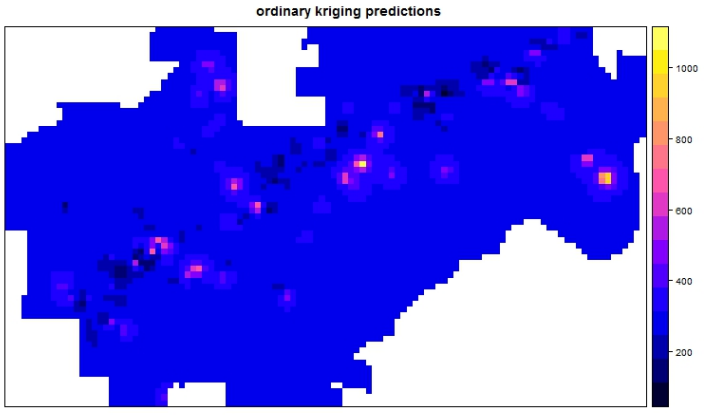

Figure 27: Estimated indoor radon concentration at ground floor based on Ordinary Kriging in AUT South.

Spatial prediction of exceeding a probability of indoor radon concentration of 300 Bq/m³ (Indicator Kriging)

The radon risk mapping was conducted using indicator kriging. Therefore, indoor radon concentration was transformed into a binary code with 0 for all observations that are smaller than 300 Bq/m³ and 1 for all observations that are greater or equal to 300 Bq/m³. Afterwards, a new variogram model was fitted to the binary coded data (see Figure 28, Figure 29 and Table 17).

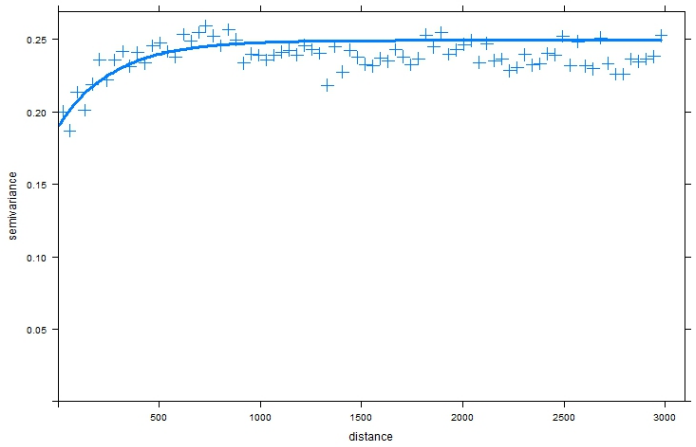

Figure 28: Variogram of binary coded ground floor indoor radon concentration for AUT North (0 = observation < 300 Bq/m³, 1 = observation ≥ 300 Bq/m³). Empirical data (crosses), fitted model (solid line)

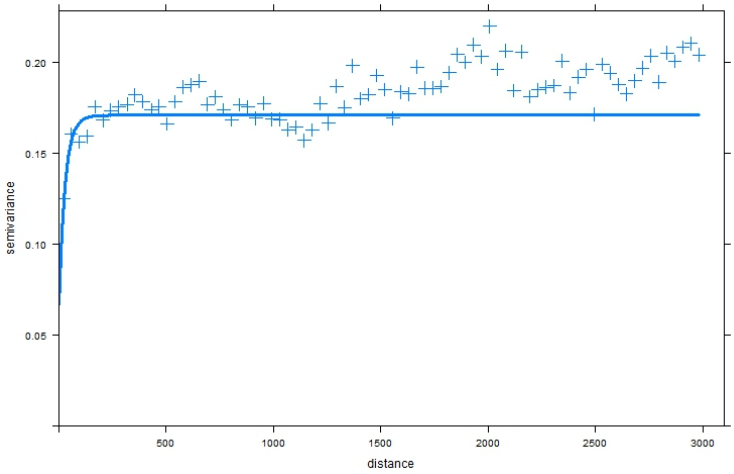

Figure 29: Variogram of binary coded ground floor indoor radon concentration for AUT South (0 = observation < 300 Bq/m³; 1 = observation ≥ 300 Bq/m³). Empirical data (crosses), fitted model (solid line);

Table 17: Variogram model parameters of binary coded indoor radon concentration on ground floor level.

	AUT North	**AUT South**
Nugget	0.19	0.07
Range parameter for exponential model	275	31
Partial Sill	0.06	0.10

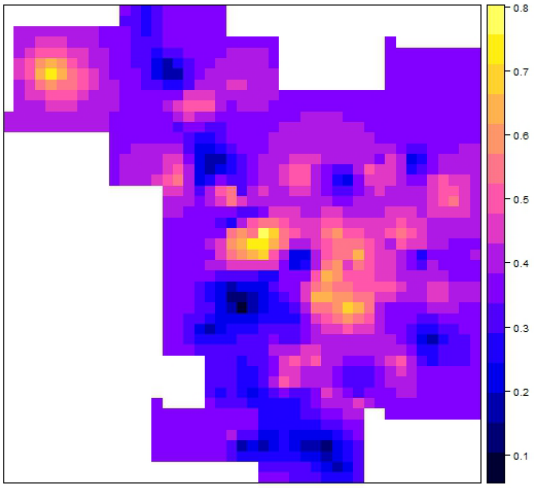

Figure 30: Estimated local probability to exceed an indoor radon concentration of 300 Bq/m³ at ground floor level based on Ordinary Kriging in AUT North.

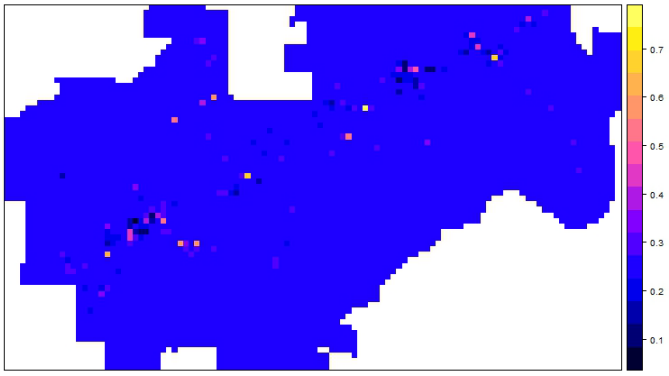

Figure 31: Estimated local probability to exceed an indoor radon concentration of 300 Bq/m³ at ground floor level based on Ordinary Kriging in AUT South.

In summary, all cells with observed values that are greater or equal to 0.1 are assigned as radon priority areas. Both Austrian regions are therefore mostly radon priority areas (except for three individual cells in the AUT South area), shown in Figure 30 and Figure 31.

Spain

Regarding data quality, the Cantabrian data set differs from the Austrian data set. For the Cantabrian data set, it is not clear whether the indoor radon measurement was conducted on the ground floor or on the first floor. This indicates that using data only from ground floor measurements, as it is the case for the Austrian data set, is not feasible.

Another characteristic of the Cantabrian data set is that coordinates, which are attributed to the indoor radon measurements, are not as accurate as desirable. All measurements from one municipality have the same coordinates. This lack of knowledge regarding radon measurement and assigned floor level and exact coordinates causes a loss of valuable information.

In order to make the data ready for kriging, all measurements from one municipality were merged into one value by calculating the arithmetic mean. Thus, each unique location is assigned to one value for indoor radon concentration.

Spatial autocorrelation of indoor radon concentration was tested but not detected, i.e. the empirical data could not be fitted in a meaningful way to the variogram model (Figure 32). Hence, kriging was not a feasible option for the delineation of radon risk areas in Cantabria.

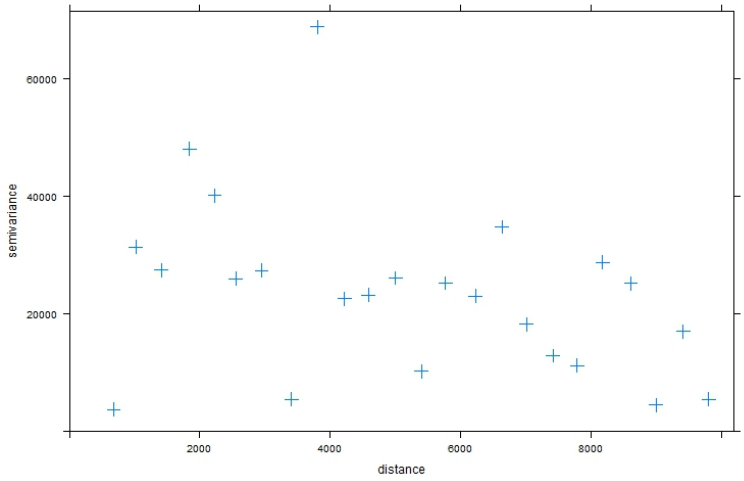

Figure 32: Variogram of indoor radon concentration for the Cantabrian data set. Empirical data (crosses); a variogram model could not be fitted due to the lack of spatial autocorrelation.

The lack of spatial autocorrelation may be a result of:

- measurement data was merged per municipality,
- distances between individual samples are larger than the range of spatial autocorrelation
- not much differentiation between ground floor and first floor measurements was given (in general, higher radon concentration would be expected on ground floor level)

Instead of geostatistical analysis of indoor radon data, the geogenic radon potential as a function of soil gas radon and soil permeability was calculated.

Kriging, based on an exponential variogram model with parameters, was conducted with the radon concentration of the soil gas (Figure 33 and Figure 34).

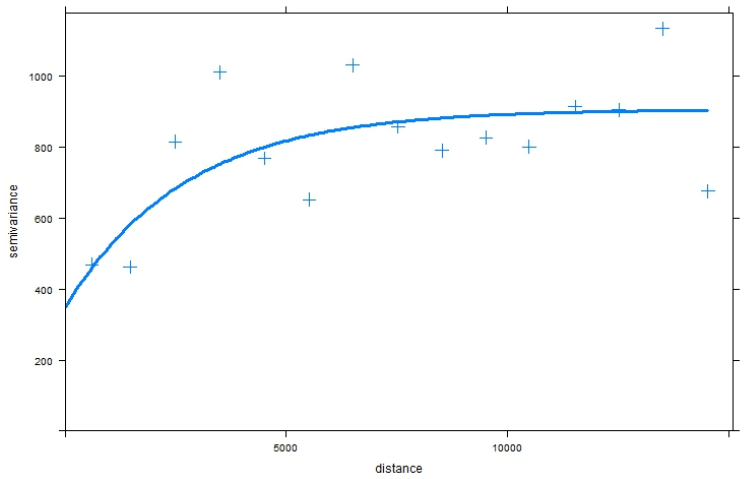

Figure 33: Variogram of radon concentration in soil gas for Cantabria. Empirical data (crosses), fitted model (solid line). Parameters (nugget: 346, partial still: 559, exponential range parameter: 2718).

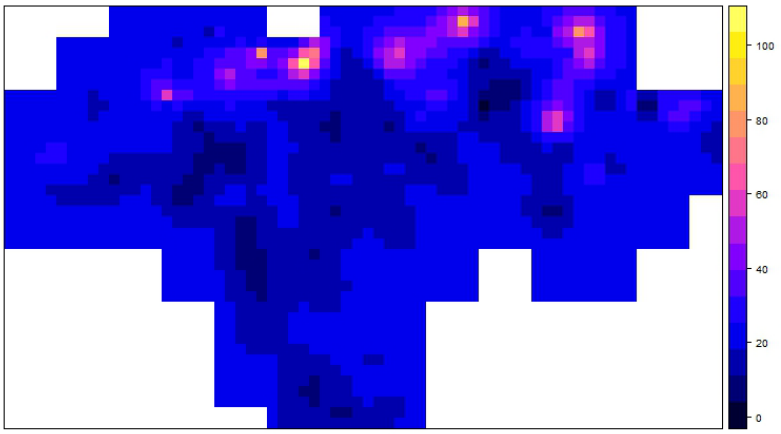

Figure 34: Estimated radon concentration in soil gas for Cantabria based on Ordinary Kriging.

Data on soil permeability was assigned to five permeability classes, depending on the lithological type. Consequently, it was not possible to use the Neznal geogenic radon potential (GRP, Neznal et al., 2004). According to the Cantabrian data set, the GRP was defined as

$$GRP = Soil\ Rn * Permeability^2 \qquad (Equation\ 7)$$

The five permeability classes ("very low", "low", "medium", "high" and "very high") were converted into numerical values (1 = "very low", 2 = "low", 3 = "medium", 4 = "high" and 5 = "very high"). Permeability data was provided as vector data, which required a transformation into raster data.

Data on soil gas radon and data on permeability are both given as raster data with the same spatial resolution (grid cell size of 2000 m). The geogenic radon potential was therefore calculated according to equation 8. Results are displayed in Figure 35.

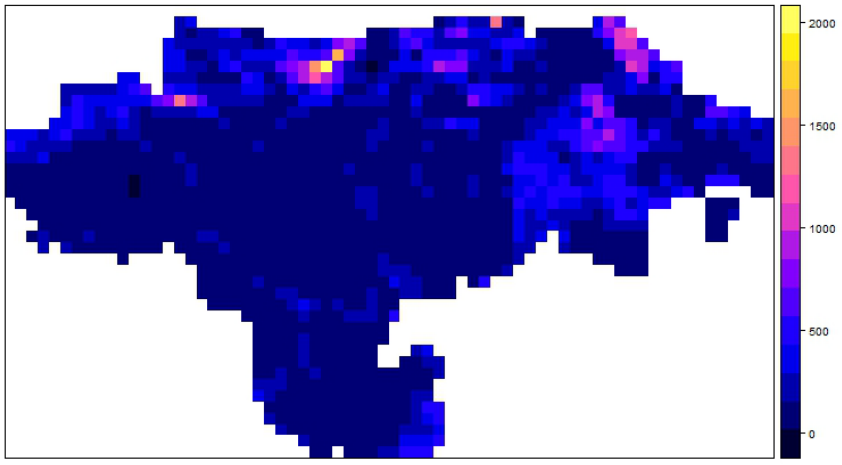

Figure 35: Geogenic radon potential (GRP) calculated from soil gas radon measurements and soil gas permeability.

In the next step, correlation between the calculated GRP and indoor radon concentration was tested. The goal was to calculate a threshold GRP value that coincides with the 10 % exceedance probability of 300 Bq/m³ indoor radon concentration at a tolerated error rate. However, there is only a weak correlation between both quantities with a Pearson correlation coefficient of 0.12 ($p<0.05$) and a spearman rank correlation coefficient of 0.08 (Figure 36). The determination of radon risk areas (radon priority areas), based on the GRP/Indoor Radon Relationship, is therefore not meaningful for the Cantabrian data set.

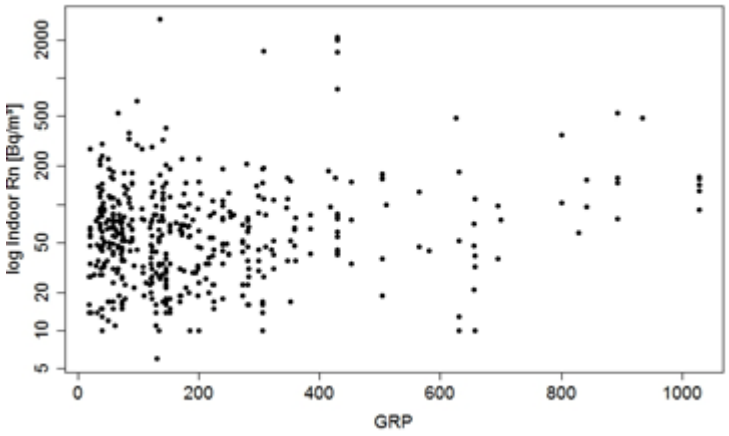

Figure 36: Correlation between geogenic radon potential (GRP) and indoor radon concentration (log).

4.5 Belgian radon risk mapping software (BRRMS)

(François Tondeur, ISIB-HE2B)

Only the Austrian data set was used for the following Belgian radon risk mapping method.

The Belgian radon mapping software

The Belgian radon risk mapping method is similar to the British one. Cinelli et al. (2011) developed the method and the corresponding software is described in Tondeur & Cinelli (2014).

The principle is to map the variations of the radon risk within geological units with the moving average method, while geological units with significantly different levels of risk are considered separately. When contiguous geological units have similar mean radon levels, they are treated as a single unit. Within a given unit, the moving average of the nearest 20 data is calculated (more precisely, the log mean, or the log median) for any chosen coordinate set, e.g. the nodes of a square grid. The percentage of data locally bypassing a chosen threshold is also predicted, assuming a lognormal distribution. The threshold used here is the European reference level of 300 Bq/m³ and the lognormal distribution is only fitted to data above the median (Cinelli & Tondeur, 2015).

The method does not include a classification of the nodes. A classification in five risk classes is used in the Belgium method for municipalities (AFCN, 2018) but was not included in the software.

Data selection

Only the highest concentration, measured on the ground floor, is kept.

Geological context

The Austrian data provided for the exercise come from two distinct rather small radon-affected areas. Each area includes different geological formations. However, the radon statistics give rather similar values for the geometrical mean indoor radon concentration in the different geological units of each area, why they were considered as a single mapping unit (Table 18).

Table 18: Geometrical mean indoor concentration in different geological units of AUT South and AUT North.

Geological unit	Number of data	Geometrical mean indoor Rn
AREA AUT North		
Granite	123	254
Migmatite	455	248
AREA AUT South		
Coarse Gneiss Complex	460	186
Permomesozoic rocks	266	161
Tertiary sediments	47	174
Other	9	233

Maps on a 500 x 500 m grid

The Belgian software evaluates the radon risk at given coordinates, e.g. at the nodes of a square grid. It does not give an average value for each square of the grid. Because of the good sampling density, the local sampling of the nearest 20 measurement data often covers a surface much smaller than the squares of the suggested 2 x 2 km grid, with the consequence that a significant part of the data might not be taken into account. Therefore, a finer grid was chosen (500 x 500 m), defined by dividing the 2 x 2 km grid initially provided, but excluding mesh points too far from any data (see Figure 37 and Figure 38).

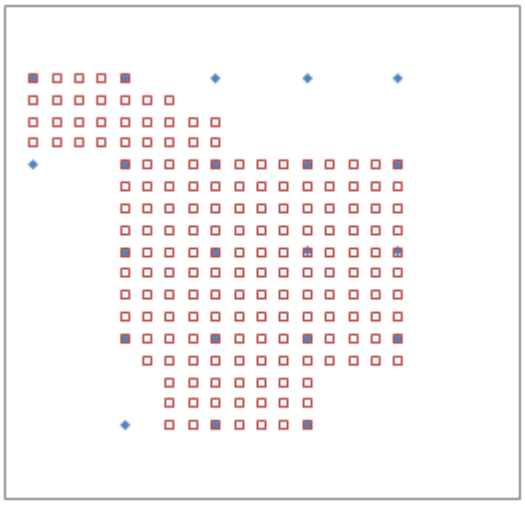

Figure 37: Grid for AUT North superimposed to the initial grid.

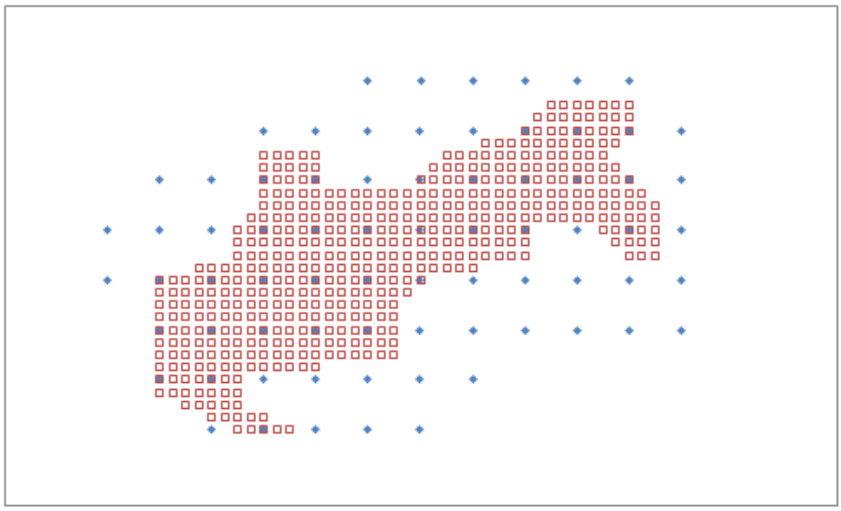

Figure 38: Grid for area AUT South superimposed to the initial grid.

Colour scales and map appearance

The colour scales used in the maps are adopted for Belgium and were chosen in order to display the contrast between unaffected areas and affected areas.

The areas considered here are affected in all their parts. Therefore, only the few colours appear in the map that correspond to radon concentration of too high and very high radon risk. The maps are given as square pixels. Note, that each pixel represents the prediction at the centre of the square, not the mean value within the square (Figure 39, Figure 40, Figure 41, Figure 42).

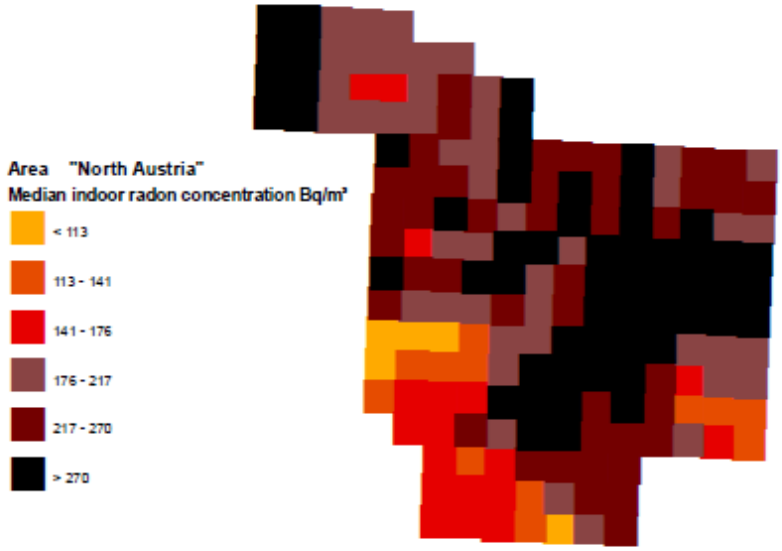

Figure 39: Map of the median indoor radon concentration in area AUT North.

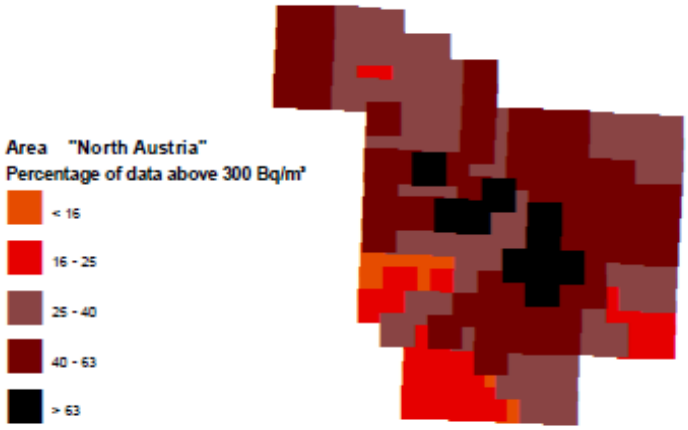

Figure 40: Map of the percentage of data above 300 Bq/m³ in area AUT North.

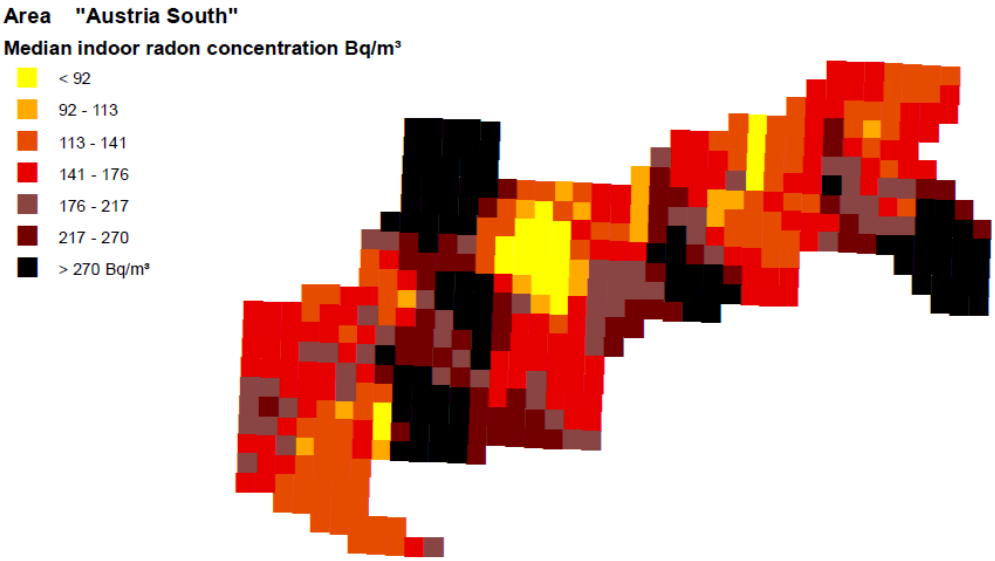

Figure 41: Map of the median indoor radon concentration in area AUT South.

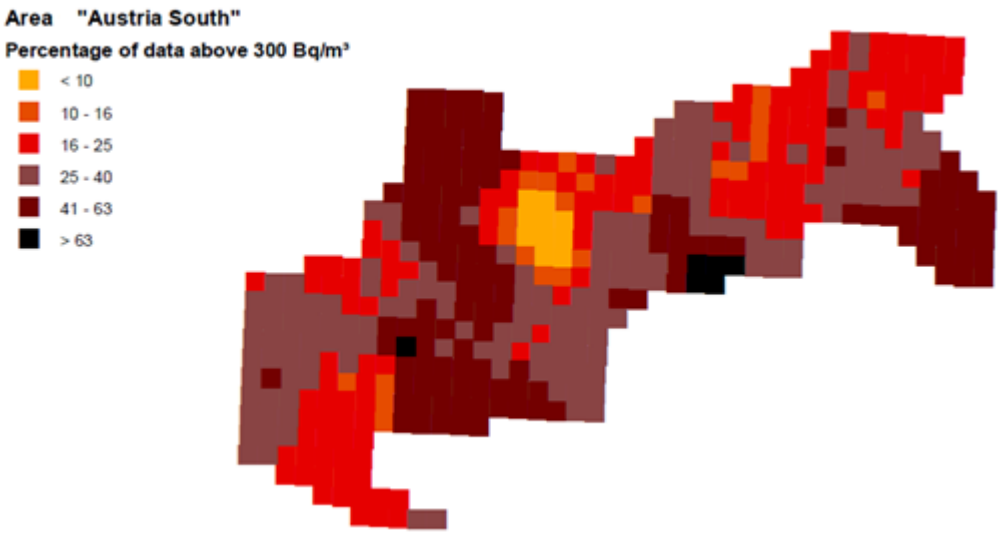

Figure 42: Map of the percentage of data above 300 Bq/m³ in area AUT South.

Conclusion

Despite the weakness of the variability related to geology, the two areas show an important variability of the percentage of dwellings above the European reference level of 300 Bq/m³. In AUT North, this percentage ranges from 7% to 67%, whereas AUT South shows a range from 11% to 78%. According to the classification used in Belgium, almost all nodes of the two areas from Austria (AUT South, AUT North) would belong to the upper risk class. Therefore, all six municipalities in the exercise would be considered as radon priority areas.

5. Summary

In the previous chapters the idea, the data sets and the applied methodologies of the mapping exercise were discussed. The main purpose of the exercise was to apply a radon classification scheme (following the concept of RPA) by using given data sets and applying different mapping methods.

Two different **data sets** were used for the exercise, Cantabria and six municipalities in Austria. The data sets differ in basic characteristics as size, sample density, data extent, quality and resolution, as shown in Table 5. In chapter 3.3 the data sets were analysed in detail, regarding significant differences among groups (e.g. soil type, geology unit), correlations among the variables and spatial dependencies for all variables, summarized in Table 9. Table 5The analysis of the exercise data shows that grouping of populations, the type of correlation and the rate of spatial correlation of the same variables are not equal in different regions.

The data sets are complex and difficult to analyse and correlations were less significant than expected. The Austrian data sets represent only small areas (6 municipalities), which seems to be too small and too geological homogenous for geogenic correlations and modelling. The Cantabrian data set represents a larger area, but the data came from different surveys and literature (e.g. GEMAS and FOREGS), which seem to be not compatible. In addition, also the data set has low sampling density and no detailed coordinates for IRC, which makes the use of IRC for modelling also challenging.

The fact, that the data are inhomogeneous and not perfect in several aspects makes it a good exercise, since also in practice, most of the time the data which are available for mapping are neither perfect nor complete which would be desirable. Consequently, the exercise can show, how different mapping methods can perform also with incomplete or heterogeneous data sets, and how classification of RPA can be done with them.

To apply the different **mapping methods** the data sets may require adequate data manipulations and not all data is used for each mapping method, and also not every mapping method can be used for the data set. Table 19 gives an overview of the applied mapping methods (see chapter 0) and the data which were used for the respective method. In general, mapping methods are mostly specified to use either IRC as target variable (e.g. basic statistics methods, Kriging IRC) or geogenic variables (EBK regression, Kriging GRP). BRRMS, the Belgium mapping method, combines IRC and geogenic variables, by taking into account geological units. The methods using IRC with building characteristics could be only applied for the Austrian data sets, as no information about building characteristics is included in the Cantabrian data set. Only the GAMM method used all available variables as well for the Austria and the Cantabrian data set. Except the basic statistic methods (IRC mean over threshold and probability of IRC over threshold per municipality or geological unit) all methods used interpolations to map the radon concentration or radon potential or radon risk.

Table 19: Overview of different methods and variables used in the respective method.

Method	IRC	Building characteristics	Soil Gas	Radionuclide contents	Geogenic factors	Interpolation
IRC mean over threshold	yes	possible subset data	no	no	no	no
Probability of IRC over threshold	yes	possible subset data	no	no	no	no
GAMM	yes	yes	yes	yes	yes	yes
EBK regression	no	no	yes	yes	yes	yes
Kriging IRC (AT)	yes	subset data	no	no	no	yes
Kriging GRP (ES)	no	no	yes	yes	yes	yes
BRRMS	yes	subset data	no	no	yes	yes

A summary of the **results** for Cantabria and the six municipalities in Austria is shown in Table 20. The table gives an arithmetic/geometric mean, median value for the IRC or the percentage of measurements above 300 Bq/m³ in Cantabria and each of the six Austrian municipalities. The results of the specific methods were discussed in detail in chapter 4. The methods which delivered results for grid cells were aggregated for the basis of region Cantabria and the municipalities for Austria, as overview and for the possibility of better comparison. The table only shows results for IRC predictions and not for geogenic radon potential (GRP). The results show that the predicted radon concentration is clearly lower for all methods in Cantabria than in Austria, and also in most cases lower in the 3 municipalities in AT South compared to AT North. The GM of Cantabria data from basic statistics and the GAMM correspond very well, also for AT Mun. 2 and 4, for the other municipalities it deviates quite strong, especially for Mun. 5 and 6. The BRRMS median concentration per municipality compared to basic statistics median deviates about 10 to 30 %, a bit stronger for the values of percentages about 300 Bq/m³. The Ordinary Kriging IRC prediction per municipality delivers clearly higher values than the basic statistics and BRRMS method. The results are compared and discussed in more detail in chapter 0.

Table 20: Results for different methods and regions for IRC (Austria and Spain)

	AM (Bq/m³)	GM (Bq/m³)	Med (Bq/m³)	% > 300	Med (Bq/m³) BRRMS	% > 300 BRRMS	GM (Bq/m³) GAMM	AM (Bq/m³) OK	% > 300 IK
Cant.	97	54	54	3	-	-	54	-	-
AT North Mun.1	289	196	197	31	231	40	243	352	36
AT North Mun. 2	313	207	213	36	240	41	201	360	39
AT North Mun. 3	429	273	266	45	230	39	208	367	39
AT South Mun. 4	289	165	168	28	209	38	153	305	26
AT South Mun. 5	251	157	144	22	183	32	241	300	26
AT South Mun. 6	234	146	130	21	173	31	310	304	26

6. Discussions and Conclusions

In the previous chapters, the data and applied mapping methods were discussed and also the results summarised. In this final chapter the methods and results should be compared and discussed and conclusions from the exercise should be drawn. As discussed in the introduction, the exercise and the delineation of radon priority areas is a multiple-step process – collecting and preparing the available data or in practice, performing the measurement campaign to get the data, selecting or developing the best mapping method for the situation and applying it to the data, and classifying the results according to the definition of RPA. The definition of RPA is mostly a political decision and not only a scientific one. As shown in chapter 2.1, different definitions of RPA concepts are adapted in the individual countries, some examples are listed in Table 2. In this chapter the results of the different applied mapping methods for the three areas (AUT North, AUT South, Cantabria) are classified and characterised according to some commonly used definitions of RPA and comparability and usability is discussed.

Correlations

The methods discussed in chapter 4 provided results for either the predicted IRC or the predicted GRP per grid cells. In Table 20 the IRC results were summarised on the basis of administrative areas, which is used also for the classification discussion below. The correlations between mapping methods were also analysed in more detail. Correlation analysis is only meaningful for methods, which provide the same variable as result (IRC, GRP) and the results need to be aggregated to the same grid cells. The GAMM method used larger grid cells for predictions and therefore have only few data points for the AT areas, which makes correlation analysis more difficult. In Figure 43 and Figure 45 two examples of correlations between different methods are shown. Figure 43 compares the EBK regression (Chapter 4.3) with the Ordinary Kriging (OK) (Chapter 4.4) for the predicted GRP for Cantabria. The data were aggregated in a 5x5 km grid and the coefficient of determination (r^2) is 0.59. The correlation between the two methods for the area is acceptable good. In Figure 44 the results (GRP predictions) of the two methods are displayed in the map (5x5 km grid). The two maps show a corresponding picture, with only some higher GRP in the North of Cantabria. In general the level of GRP prediction by OK method is a bit above the one by EBK regression.

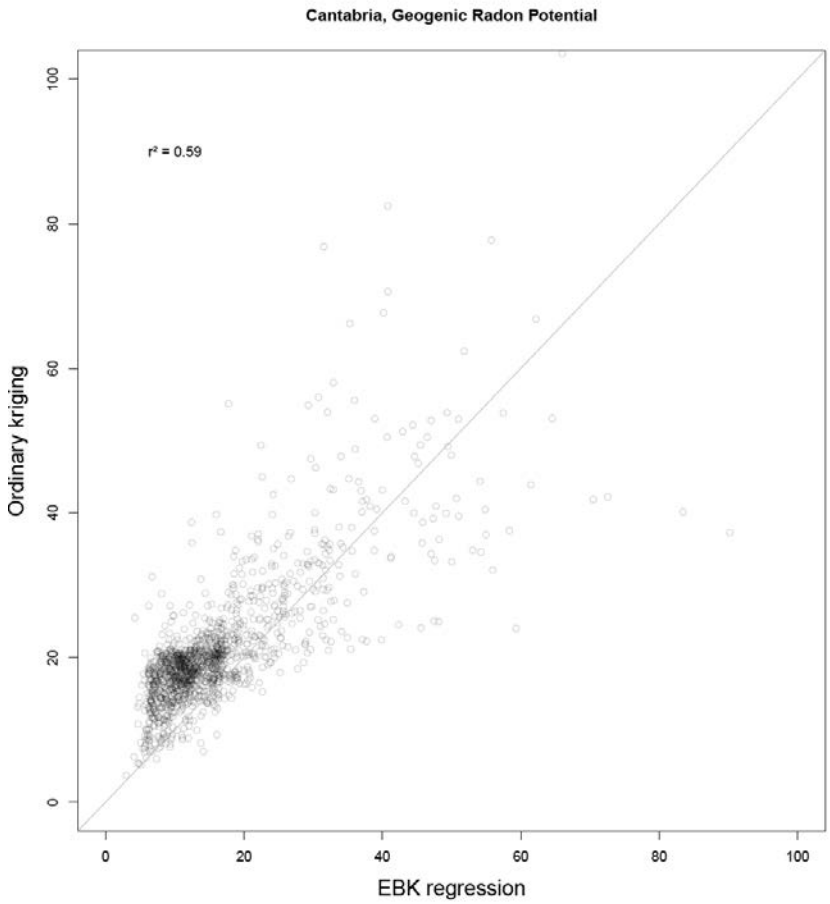

Figure 43: Correlation between 2 different mapping methods for the geogenic Radon Potential (GRP) for Cantabria data set – Ordinary Kriging (OK) (Chapter 4.4) vs. EBK regression (Chapter 4.3)

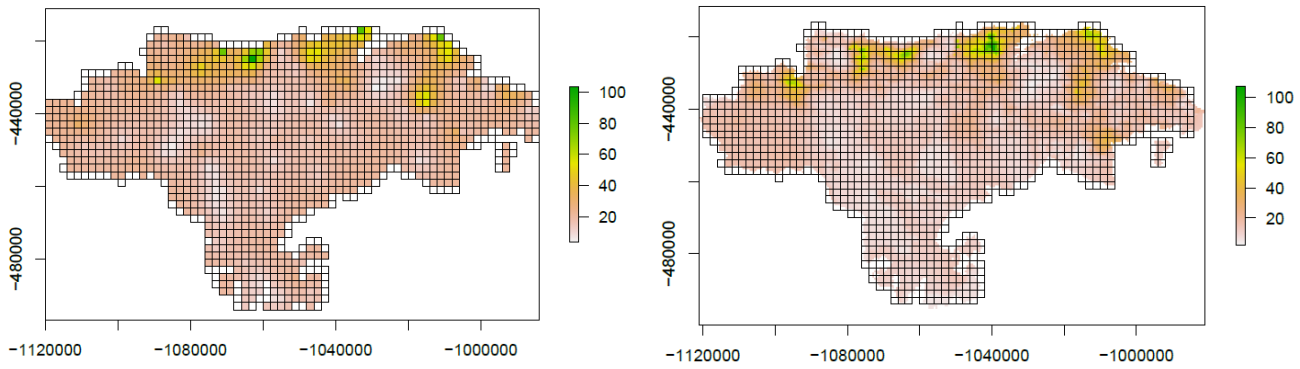

Figure 44: Mapping the GRP prediction in 5 x 5 km grid for Cantabria with Ordinary Kriging method (OK) (Chapter 4.4, left handside) and EBK regression. Predictions were aggregated for 5x5 km grids (Chapter 4.3, right handside).

Figure 45 compares the Belgian Radon Risk Mapping Software (BRRMS, Chapter 4.5) with the Indicator Kriging (IK, Chapter 4.3) for the predicted percentage of measurements above 300 Bq/m³ for the area AUT North. As basis for the comparison the coarser 500 x 500 m grid of the BRRMS was used and compared with the cell of the 200x200 m kriging raster closest to the midpoint of the BRRMS grid cell. The coefficient of determination (r^2) is 0.41, which is still a satisfying correlation. In Figure 46 the results (% of measurements over 300 Bq/m³) of the two methods are displayed in the map. The two maps show a quite similar picture, with some cells with highest radon potential in the centre. In general, the level of prediction by BRRMS method is a bit above the one by IK.

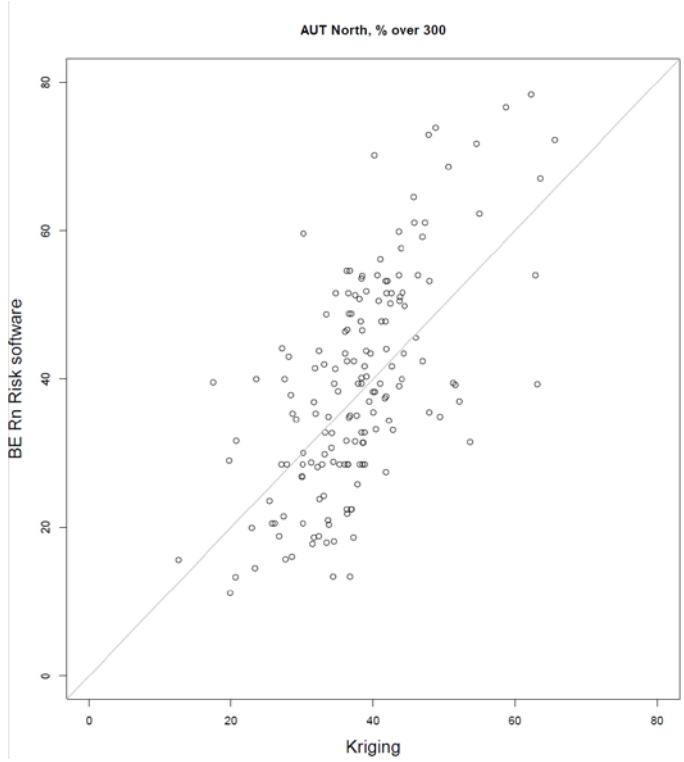

Figure 45: Correlation between 2 different mapping methods for the % above 300 Bq/m³ for the AT North data set – Belgian Radon Risk Mapping Software (BRRMS, Chapter 4.5) vs. Indicator Kriging (IK, Chapter 4.3)

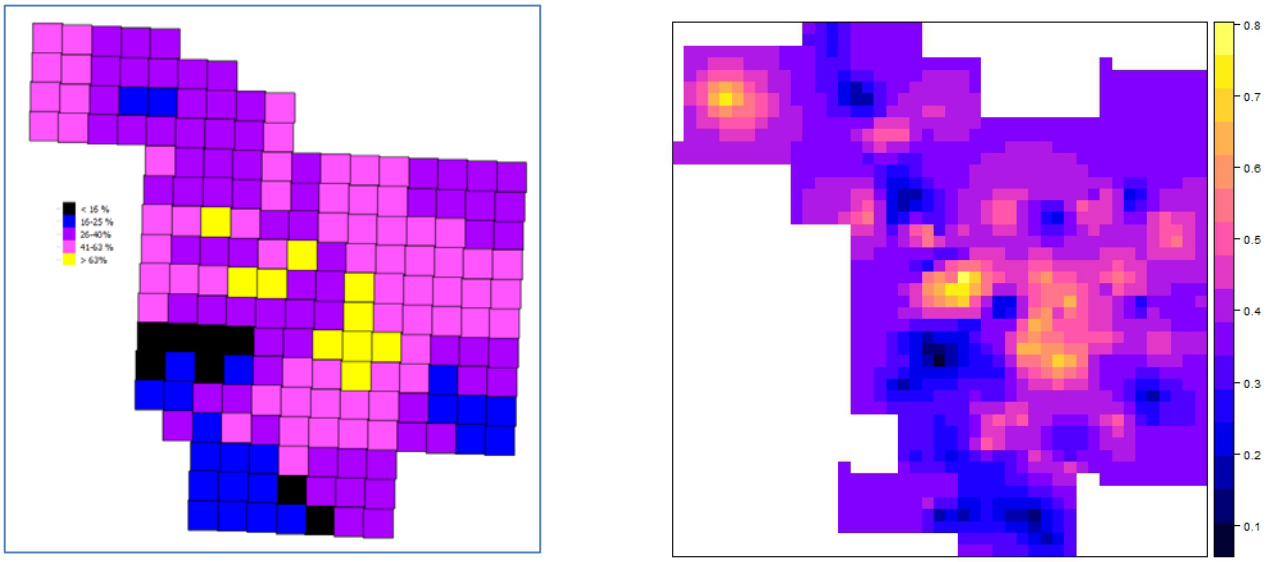

Figure 46: Mapping the prediction of % above 300 Bq/m³ for the AZ North data set with Belgian Radon Risk Mapping Software (BRRMS, Chapter 4.5, left handside) and Indicator Kriging (IK, Chapter 4.3, right handside)

The two examples show quite good correlations for the predicted cells which indicates, that they should be interchangeable for harmonisation purposes. In general, the selection of a mapping method for a certain area, will be highly depend on the available data sets. Not all mapping methods are applicable to all data and all areas as depending on data quality, sample density, heterogeneity of the area, etc. In our example the methods using building characteristics for the prediction of IRC were not possible to use for the Cantabrian data set, where this information was not available. On the other hand, methodologies based on differences between geogenic factors (e.g. EBK regression) could not be adapted to the very small, quite geogenic homogeneous areas of Austria. Also for the BRRMS, taking into account information of geological units, had problems within the AUT North area with only very few geological areas. All this information needs to be evaluated and taken into account when choosing a mapping method for a certain area or a certain available data set. If a survey for delineation of RPA (as requested in the EU-BSS) is started from scratch, the mapping method and display/classification method for the map (e.g. % above RL in administrative area) should be decided at the beginning, so that the survey (measurement density, analysed parameters, etc.) can be optimised to these requirements. For harmonisation of mapping or delineation of areas (e.g. on a European basis) a method using less parameters might be preferable, as easier to apply to different data sets.

Classification of Radon priority areas (RPA)

As discussed above, different definitions of RPA concepts are adapted in the countries. In Table 20 the results of the different methods for the administrative units (six municipalities in Austria and the region of Cantabria) were summarised. The comparison in chapter **Fehler! Verweisquelle konnte nicht gefunden werden.** showed, that some of the results correspond very well, while others not, e.g. the Ordinary Kriging IRC prediction per municipality delivers clearly higher values than the basic statistics and BRRMS method.

Here we want to evaluate, how these different results provided by different mapping methods would have an impact on the classification or delineation of RPAs. Table 21 and Table 22 show the same results as Table 20, but two common RPA classification definitions were applied to the results – mean/median/GM above a certain threshold (Table 21) and percentage of measurements/predictions above a threshold (Table 22). If the threshold of above AM/Med/GM is set to 300 Bq/m³, all six Austrian municipalities would be classified as RPA with the OK method, municipality 2 and 3 with the basic statistics method (AM) and municipality 6 with the GAMM method (marked in purple in Table 21). If, on the other hand, the threshold of above AM/Med/GM is set to 100 Bq/m³ all six Austrian municipalities with all applied methods would be classified as RPAs (marked in red and purple in Table 21). Cantabria would not be considered as RPA for all methods and classification thresholds. This shows, that the chosen threshold for the classification of RPA has a major impact, depending on the level of radon concentration in the area. For Cantabria which has a very low radon concentration, the differences in the results of the different methods do not impact the RPA classification. Whereas the Austrian municipalities show radon concentrations in the range about 150 to 400 Bq/m³, depending on municipality and mapping method. Therefore, the differences (even when small) in the radon concentration for the different methods for the same municipality can have an impact in RPA classification, when the threshold is chosen in the range of the variability of the results (e.g. 300 Bq/m³ as shown in the example). If the threshold is set with 100 Bq/m³ all municipalities are classified the same, as this threshold does not lie within the range of the measurement/prediction results and therefore the variability of the results of the different methods do not have an impact on the classification of RPAs.

If, in Table 22, the threshold of percentage of measurements/ predictions is set to 30 %, a definition which is used only in Czech Republic, all municipalities in AT North would be classified to be RPA with all three applied methods, and for all six municipalities for the BRRMS method (marked in purple in Table 22). Applying the commonly used definition of RPA in Europe (10 % of dwellings above 300 Bq/m³), all six municipalities in Austria

would be clearly considered as RPA, independent from the mapping method (marked in red and purple in Table 22). As discussed above, the variability of the results of the different methods only impact the classification of RPA when the set threshold lies within the range of the predicted/measured results.

Figure 47 displays the same results as shown in Table 21 for the three municipalities of the Austria North and Austria South area. The results (AM/GM/Med) per municipality for the respective methods is plotted and the colouring shows, for which threshold the municipality would be considered to be RPA (yellow) and Non-RPA (green).

Table 21: Results for different methods and regions related to Median, GM or AM of measured/predicted IRC with applied classification definition of RPA (purple: AM/Med/GM > 300 Bq/m³; red: AM/Med/GEM > 100 Bq/m³)

	AM (Bq/m³)	GM (Bq/m³)	Med (Bq/m³)	Med (Bq/m³) BRRMS	GM (Bq/m³) GAMM	AM (Bq/m³) OK
Cant.	97	54	54	-	54	-
AT North Mun.1	289	196	197	231	243	352
AT North Mun. 2	313	207	213	240	201	360
AT North Mun. 3	429	273	266	230	208	367
AT South Mun. 4	289	165	168	209	153	305
AT South Mun. 5	251	157	144	183	241	300
AT South Mun. 6	234	146	130	173	310	304

Table 22: Results for different methods and regions related to % of measurements/predictions above 300 Bq/m³ with applied classification definition of RPA (purple: > 10 % above 300 Bq/m³)

	% > 300	% > 300 BRRMS	% > 300 IK
Cant.	3	-	-
AT North Mun.1	31	40	36
AT North Mun. 2	36	41	39
AT North Mun. 3	45	39	39
AT South Mun. 4	28	38	26
AT South Mun. 5	22	32	26
AT South Mun. 6	21	31	26

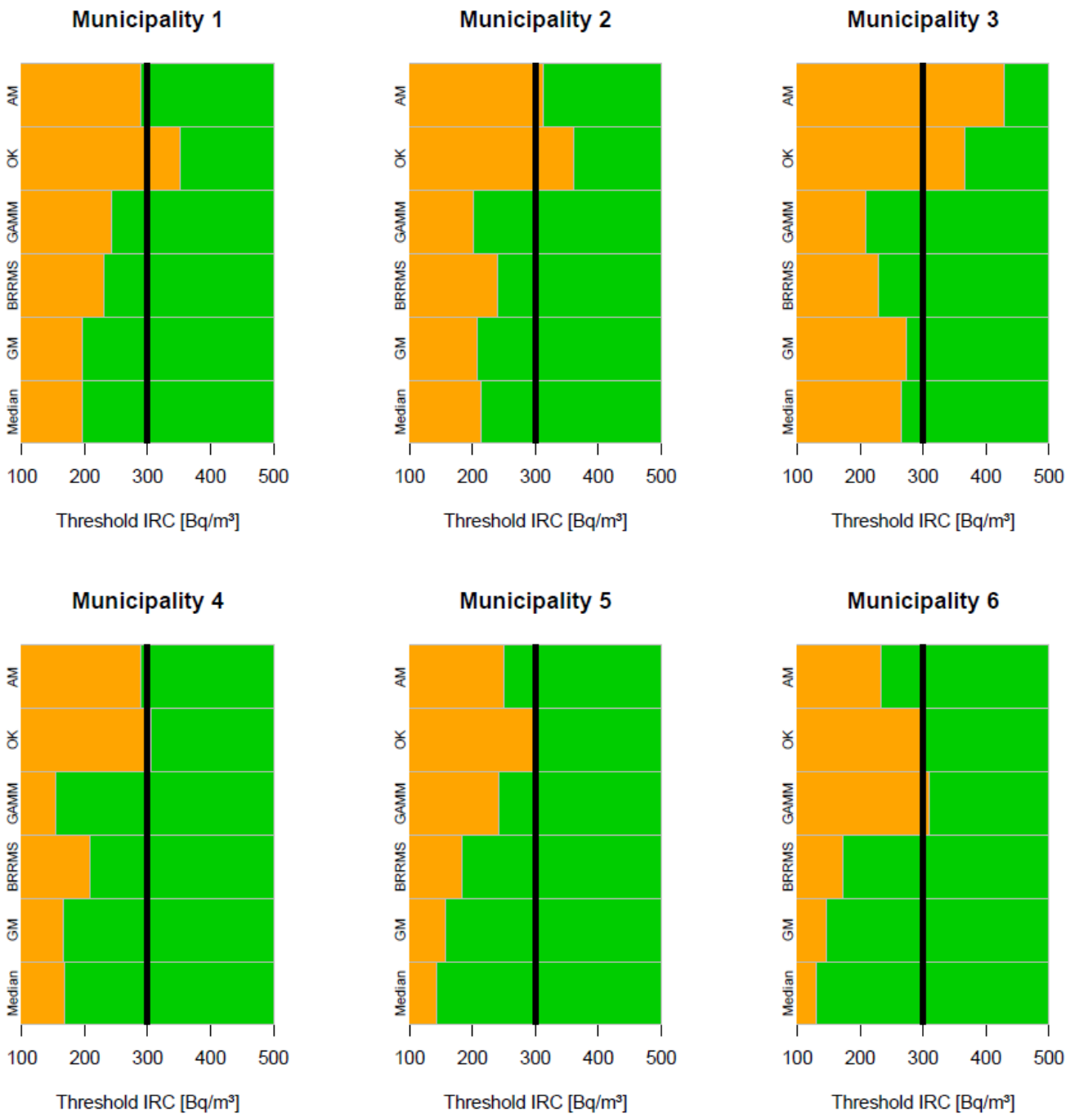

Figure 47: Classification of RPA for the 6 municipalities in Austria with different methods and for different thresholds (green: no RPA, orange: RPA – further explanation in the text)

As known and shown also within this exercise and this report, mapping methodologies are various and so are the definitions of RPAs. To evaluate the situation in Europe and possibilities for harmonisation between countries and on borders was the driving factor for the work package within the MetroRADON project, where this mapping exercise is part of. As a general conclusion from this exercise, it can be said, that applying a mapping method using data sets, which were not designed for the specific requirements of the mapping method, is challenging. Usually, data sets always have specific characteristics and are rarely comparable, even not for the same variable. Therefore, harmonisation is always a challenge. But some examples in this exercise show quite good correlations for the predicted cells which indicates, that they should in principle be interchangeable for

harmonisation purposes. In general, the selection of a mapping method for a certain area, will be highly depend on the available data sets. Not all mapping methods are usable for all data sets or areas, depending especially on data quality, sampling density, or heterogeneity of the mapping area. For harmonisation of mapping (e.g. on a European basis) a method using less parameters might be preferable, as it would be easier to apply to different data sets.

Usually the final goal of mapping is the delineation of RPA, as this is requested in the EU-BSS. It was shown in this exercise, that independent of the applied method for large intervals of classification threshold the same RPA classification is predicted. Different methods often deliver same results in RPA classification, depending on the definition of RPAs. So, the definition of thresholds is a very important factor in the process of delineation of RPA and might be as relevant as harmonising mapping methods.

7. References

AFCN, 2018: https://afcn.be/fr/dossiers-dinformation/radon-et-radioactivite-dans-votre-habitation/radon#Taux de radon dans votre commune

Arvela, H., Holmgren, O., & Hänninen, P., 2016: Effect of soil moisture on seasonal variation in indoor radon concentration: modelling and measurements in 326 Finnish houses. *Radiation protection dosimetry*, *168*(2), 277–290. https://doi.org/10.1093/rpd/ncv182

P. Bossew, G. Cinelli, T. Tollefsen, M. De Cort, 2016: Multivariate estimation of the geogenic radon potential and prediction of indoor Rn concentration. Oral presentation during the 13th International Workshop on the Geological Aspects of Radon Risk Mapping, Prague 15-16 September 2016. (available at: https://remon.jrc.ec.europa.eu/About/Atlas-of-Natural-Radiation/Geogenic-radon/Geogenic-radon)

Bundesforschungszentrum für Wald (BFW), 2020: Bodenkarte Österreich, https://bodenkarte.at

Cinelli, G., Tondeur, F., Dehandschutter, B, 2011: Development of an indoor radon risk map of the Walloon region of Belgium, integrating geological information, Environmental Earth Sciences 62(4):809-819

G. Cinelli, G. and Tondeur, F., 2015: Log-normality of indoor radon data in the Walloon region of Belgium, Journal of Environmental Radioactivity 143:100

Damkjaer, A.; Korsbech, U., 1992*:* A Small-Diameter Probe for In Situ Measurements of Gas Permeability of Soils; Radiat. Prot. Dosim. Vol. 45, p. 85-89.

Council of the European Union (EU), 2014: Council directive 2013/59/Euratom of 5 December 2013 laying down basic safety standards for protection against the dangers arising from exposure to ionising radiation, and repealing directives 89/618/Euratom, 96/29/Euratom and 2003/122/Euratom. Official J. Eur. Union, L 13, 57, (1–73)

European Commission, Joint Research Centre (EC JRC), 2020: The digital version of the European Atlas of Natural Radiation, Geogenic radon, https://remon.jrc.ec.europa.eu/About/Atlas-of-Natural-Radiation/Geogenic-radon/Geogenic-radon

EPA (United States Environmental Protection Agency), 1993: EPA's Map of Radon Zones – National Summary, EPA-402-R-93-071

Esri, ArcGIS, 2020: What is EBK regression prediction? http://pro.arcgis.com/de/pro-app/help/analysis/geostatistical-analyst/what-is-ebk-regression-prediction-.htm)

FOREGS – EuroGeo Surveys, 2020: Geochemical Atlas of Europe, http://weppi.gtk.fi/publ/foregsatlas/index.php

Friedmann, H., 2005: Final Results of the Austrian Radon Project, Health Physics, Vol. 89 (4), pp. 339-348

GEMAS- Geochemical Mapping of Agricultural and Grazing Land Soil, 2020**:** http://gemas.geolba.ac.at/

Geological Survey of Austria (GBA), 2020: Geological Map of Austria, 1:500.000, https://www.data.gv.at/katalog/dataset/onegeology-gba

Geological Survey of Austria (GBA), 2020: Geological Map of Austria, 1:50.000, https://www.geologie.ac.at/produkte-shop/karten/geologische-karte-der-republik-sterreich-150000/

INTERNATIONAL ATOMIC ENERGY AGENCY (IAEA) (1979): Gamma Ray Surveys in Uranium Exploration. - Techn. Rep. Series No. 186; Wien.

IGME, Geological and Mining Institute of Spain, 2020: Lithostratigraphic Map of Spain, 1:200.000, http://mapas.igme.es/Servicios/default.aspx#IGME_Litoestratigrafico200.

IGME Geological and Mining Institute of Spain, 2020a: One Geology Map of Spain, 1:1M. http://mapas.igme.es/gis/rest/services/oneGeology/ESP_1M_IGME_1GEGeology_EN/MapServer

IGME Geological and Mining Institute of Spain, 2020b: Karstic Map of Spain, 1:1M. http://mapas.igme.es/gis/rest/services/Cartografia_Tematica/IGME_Karst_1M/MapServer

MetroRADON, 2020: Metrology for Radon Monitoring – MetroRADON, project website, http://metroradon.eu

Neznal M., Neznal M., Matolín M., Barnet I., Mikšová J., 2004: The new method for assessing the radon risk of building sites, Czech Geological Survey Special Papers 16, 48 pages, Published by Czech Geological Survey, Prague; https://www.radon-vos.cz/pdf/metodika.pdf

Puckett, W. E., Dane, J. H., Hajer, B. F., 1985: Physical and Mineralogical Data to Determine Soil Hydraulic Properties. Soil Science Society of America Journal, 49, 931-836

Sainz-Fernandez, C., Fernandez-Villar, A., , Fuente-Merino, I., Gutierrez-Villanueva, J. L., Martin-Matarranz J. L., Garcia-Talavera, M., Casal-Ordas, S. Quindos-Poncela L. S., 2014: The Spanish indoor radon mapping strategy, Radiat. Prot. Dosim. Vol. 162 (1-2), p. 58-62

Suarez Mahou, E., Fernandez, J.A., 1997: Project MARNA: Natural Gamma Radiation Map.,

Tondeur, F. and Cinelli, G. (2014): A software for indoor radon risk mapping based on geology, Nuclear Technology and Radiation Protection XXIX:S59-S63

UNSCEAR (United Nations Scientific Committee on the Effects of Atomic Radiation), 2010: Sources and effects of ionizing radiation. UNSCEAR 2008 Report to the General Assembly with Scientific Annexes, vol. I, United Nations, pp. 683.

Annex to Chapter 3.3 – Data set analysis

(Alcides Pereira, Filipa Domingos, University of Coimbra)

Austria: Northern Region (AUT North)

Analysis of soil data (acquired by physical sampling and airborne spectrometry)

According to soil type, soil grain size, soil source and bedrock, the data analysis shows that the content of Ra-226, U-238, K-40, Ra-228 and Th-228 is not statistically different (significance level of 0.05) between the various types of soil and bedrock. This also applies for the different soil source types (Table 23). The ambient dose rate, however, is statistically different among different bedrock types and soil source types. The ADR is higher on gneissic and silicate sources compared to granitic soil sources. The ADR is higher in the Weinsberger biotite granite and the two-mica Altenberger granite, followed by migmatite rocks, alkaline to intermediate plutonic rocks and valley infills. As the ADR is significantly influenced by soil source and/or bedrock type, the terrestrial gamma dose rate (TGDR) was computed for the purpose of comparison of the combined content of radionuclides according to soil type, soil source and bedrock type. The TGDR was computed from U-238, Th-232 (assuming secular equilibrium between Th-228 and Th-232) and K-40 activity concentration [Bq/kg] according to the following equation:

$$TGDR = 0.0417 \times {}^{40}K + 0.462 \times {}^{238}U + 0.604 \times {}^{232}Th \qquad (Equation\ 8)$$

The dose conversion factors [Bq/kg] of 0.0417, 0.462 and 0.604 were retrieved from UNSCEAR (2010). The results of the Kruskal-Wallis test show the lack of statistically significant differences of the TGDR among groups.

The analysis of the soil gas radon and permeability data according to soil type, soil grain size, soil source and bedrock shows that the soil gas radon concentration is statistically different at a 0.05 significance level between the different soil types, soil grain size and soil water content (Table 24). Radon concentration in soil gas is higher in sediment brown earth compared to rock brown earth and silt compared to loamy sand. Radon concentration in soil gas is also higher in moderately moist soils compared to other moisture contents (dry, well supplied and moderately dry).

The eU concentration, determined by airborne gamma spectrometry, shows statistically significant differences among different soil sources, bedrock and soil water content (Table 24). Colluvium soils present higher eU contents, followed by soils derived from gneiss, granite and silicates. Alkaline to intermediate rocks present higher eU contents, followed by the fine grained two-mica granite (Altenberger) and the intermixing zone and fluid transition of coarse-grained biotite granite and migmatite. The fine to intermediate grained migmatite (Meta-Diatexite), valley infill sediments and the coarse to very coarse grained biotite granite (Weinsberger) present the lowest average eU contents. The permeability is not statistically significant between different soil and bedrock units, soil sources and soil water content (Table 24).

Table 23: Analysis of radionuclides concentration data, terrestrial gamma dose rate (TGDR) and ambient dose rate (ADR) by soil and bedrock type (statistically significant differences are marked in bold).

Variable	Soil type	Soil grain size	Soil source	Bedrock (geology_fine)	Bedrock (geology_coarse)
Ra-226	H(1;28) = 0.3899; p = 0.5323	H(1;28) = 0.3899; p = 0.5323	H(2;28) = 2.0819; p = 0.3531	H(2;14) = 4.5; p = 0.1054	H(1;27) = 0.1102; p = 0.7399
U-238	H(1;28) = 0.1989; p = 0.6556	H(1;28) = 0.1989; p = 0.6556	H(2;28) = 3.5252; p = 0.1716	H(2;14) = 4.5; p = 0.1054	H(1;27) = 0.1959; p = 0.6580

K-40	H(1;28) = 0.1274; p = 0.7212	H(1;28) = 0.1274; p = 0.7212	H(2;28) = 0.68; p = 0.7118	H(2;14) = 1.6236; p = 0.4441	**H(1;27) = 6.3394; p = 0.0118**
Ra-228	H(1;28) = 0.9629; p = 0.3265	H(1;28) = 0.9629; p = 0.3265	H(2;28) = 2.0862; p = 0.3524	H(2;14) = 0.5914; p = 0.7440	H(1;27) = 0.1102; p = 0.7399
Th-228	H(1;28) = 0.7958; p = 0.3724	H(1;28) = 0.7958; p = 0.3724	H(2;28) = 1.4113; p = 0.4938	H(2;14) = 0.2714; p = 0.8731	H(1;27) = 0.1294; p = 0.7191
Pb-210	H(1;28) = 0.2865; p = 0.5925	H(1;28) = 0.2865; p = 0.5925	H(2;28) = 1.1626; p = 0.5592	**H(2;14) = 8.2229; p = 0.0164**	H(1;27) = 0.0122; p = 0.9119
TGDR (calc)	H(1;28) = 0.5093; p = 0.4754	H(1;28) = 0.5093; p = 0.4754	H(2;28) = 3.2833; p = 0.1937	H(2;14) = 2.1429; p = 0.3425	TGDR: H(1;27) = 2.0694; p = 0.1503
ADR	H(1;57) = 2.2415; p = 0.1344	H(1;57) = 2.2415; p = 0.1344	**H(2;57) = 6.7742; p = 0.0338**	**H(2;41) = 8.9202; p = 0.0116**	H(1;56) = 3.1127; p = 0.0777
H – Kruskal-Wallis H test; the degrees of freedom and number of data are indicated within brackets, respectively.					

Table 24: Analysis of soil gas radon, eU content and permeability data by water content, soil and bedrock type (statistically significant differences are marked in bold).

Variable	Soil type	Soil grain size	Soil source	Bedrock (g_fine)	Bedrock (g_coarse)	Soil water content
Soil gas radon	**H(1;57) = 5.0859; p = 0.0241**	**H(1;57) = 5.0859; p = 0.0241**	H(2;57) = 1.8905; p = 0.3886	H(2;41) = 1.4992; p = 0.4726	H(1;56) = 0.4773; p = 0.4896	**H(3;57) = 10.3603; p = 0.0157**
eU (ppm)	**H(3;3732) = 13.9654; p = 0.0030**	**H(2;3732) = 6.4935; p = 0.0389**	**H(3;3732) = 80.895; p = 0.0000**	**H(3;3732) = 80.895; p = 0.0000**	**H(1;3732) = 32.5135; p = 0.00000**	**H(4;3732) = 87.2186; p = 0.0000**
Permeability	H(1;57) = 0.3179; p = 0.5729	H(1;57) = 0.3179; p = 0.5729	H(2;57) = 0.3295; p = 0.8481	H(2;41) = 0.0874; p = 0.9572	H(1;56) = 0.3101; p = 0.5776	H(3;57) = 2.9429; p = 0.4005
H – Kruskal-Wallis H test; the degrees of freedom and number of data are indicated within brackets, respectively.						

The correlation between soil gas radon (Rn-222), radionuclide concentration (K-40, Pb-210, Ra-226, Th-228, U-238), ADR, TGDR and airborne eU concentration was evaluated in Table 25. For comparing soil and airborne data, the closest value (raster cell centre) to the soil data sampling location was chosen. Soil gas radon is significantly correlated with U-238. The ADR is correlated with K-40, Ra-228/Th-228 and the TGDR. The correlation between Pb-210, Ra-226 and U-238 activity concentration is significant which indicates an equilibrium in the U decay chain. However, there is no correlation with the ADR or airborne eU. Ra-228 and Th-228 are strongly correlated, indicating an equilibrium in the Th-232 decay chain.

Table 25: Spearman rank correlation matrix. Correlation coefficients are statistically significant at a 0.05 significance level and indicated in red

	Rn-222 [kBq/m³[	**ADR**	**K-40 [Bq/kg]**	**Pb-210 [Bq/kg]**	**Ra-226 [Bq/kg]**	**Ra-228 [Bq/kg]**	**Th-228 [Bq/kg]**	**U-238 [Bq/kg]**	**TGDR**	**eU [ppm]**
Rn-222 [kBq/m³]	1.00									
ADR	-0.06	1.00								
K-40 [Bq/kg]	-0.07	0.49	1.00							
Pb-210 [Bq/kg]	0.32	-0.08	-0.05	1.00						
Ra-226 [Bq/kg]	0.11	0.11	-0.19	0.56	1.00					
Ra-228 [Bq/kg]	-0.06	0.37	0.27	0.26	0.30	1.00				

Th-228 [Bq/kg]	-0.07	0.39	0.31	0.27	0.33	0.97	1.00			
U-238 [Bq/kg]	0.37	0.09	-0.06	0.48	0.82	0.15	0.14	1.00		
TGDR	0.03	0.48	0.74	0.26	0.26	0.76	0.79	0.29	1.00	
eU [ppm]	0.10	0.02	-0.07	0.25	0.31	-0.17	-0.16	0.32	-0.06	1.00

The correlation between Ra-226, U-238 and eU data is evaluated in Figure 48. The results show poor correlation (not statistically significant) between the eU and the soil map data using both the closest eU value to the sampling location as well as the average of the closest values.

The omnidirectional variograms for the radionuclides (Pb-210, Ra-226 and U-238), soil gas radon, permeability ($x10^{10}$), ADR, the calculated TGDR and eU were computed for evaluating spatial correlations within the data set (Figure 49). Apart from the eU data, no clear spatial correlation is observed (see discussion below for more details). The eU data was mapped with the variogram presented in Figure 49. The modelled data presented in Figure 50 show a high degree of variability of the eU data, particularly within each square km of the "soilmap" layer.

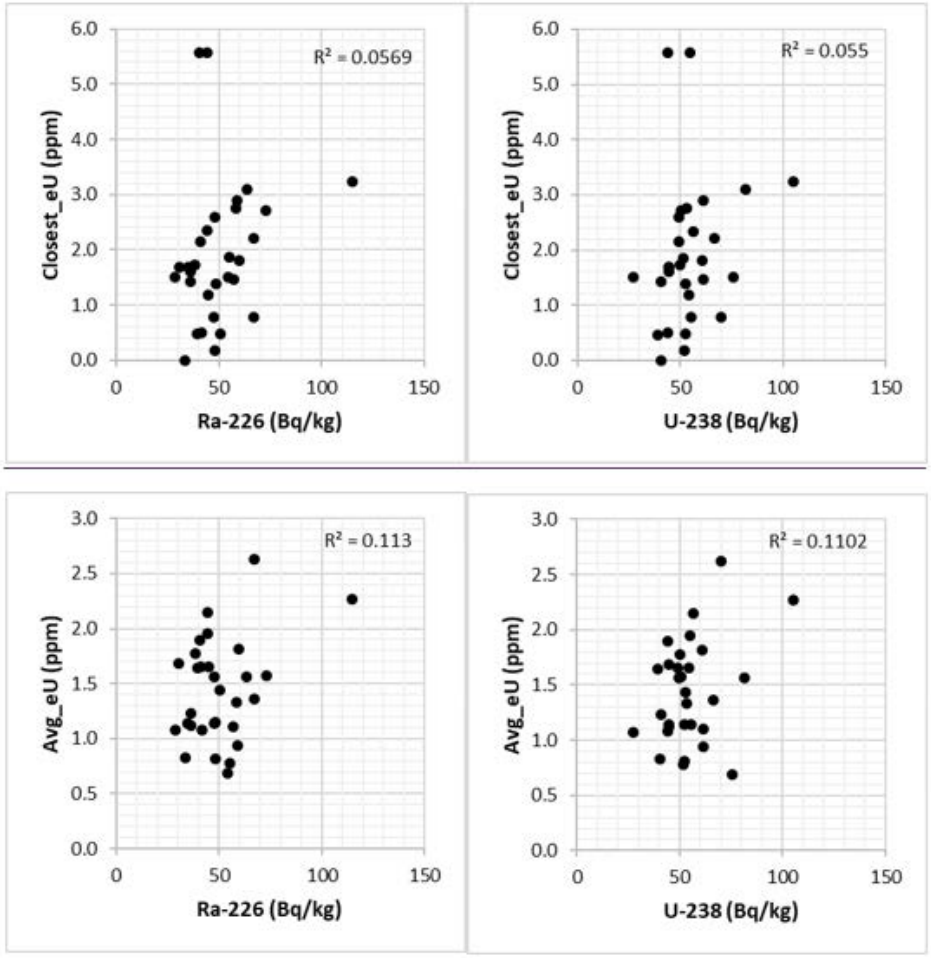

Figure 48: Correlation between Ra-226 or U-238 and airborne eU data (avg – average of eU data within a 500 m range were calculated and compared to the location of the sample; close – the closest eU value to the location of the samples were chosen to compare eU data to the radionuclide data).

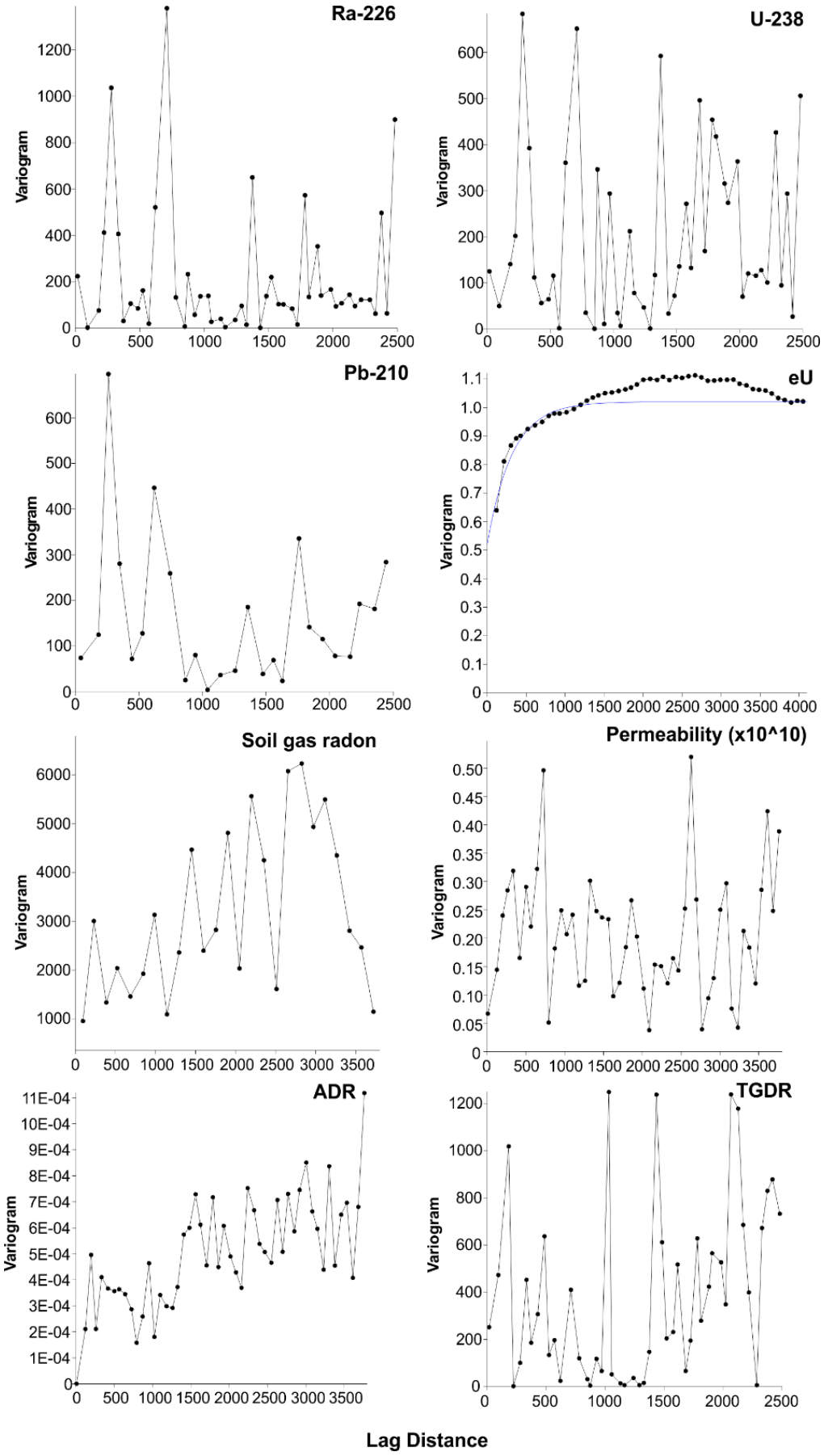

Figure 49: Omnidirectional semi-variograms of Ra-226, U-238 and Pb-210 activity concentration, airborne eU data, soil gas radon, permeability, ADR and TGDR.

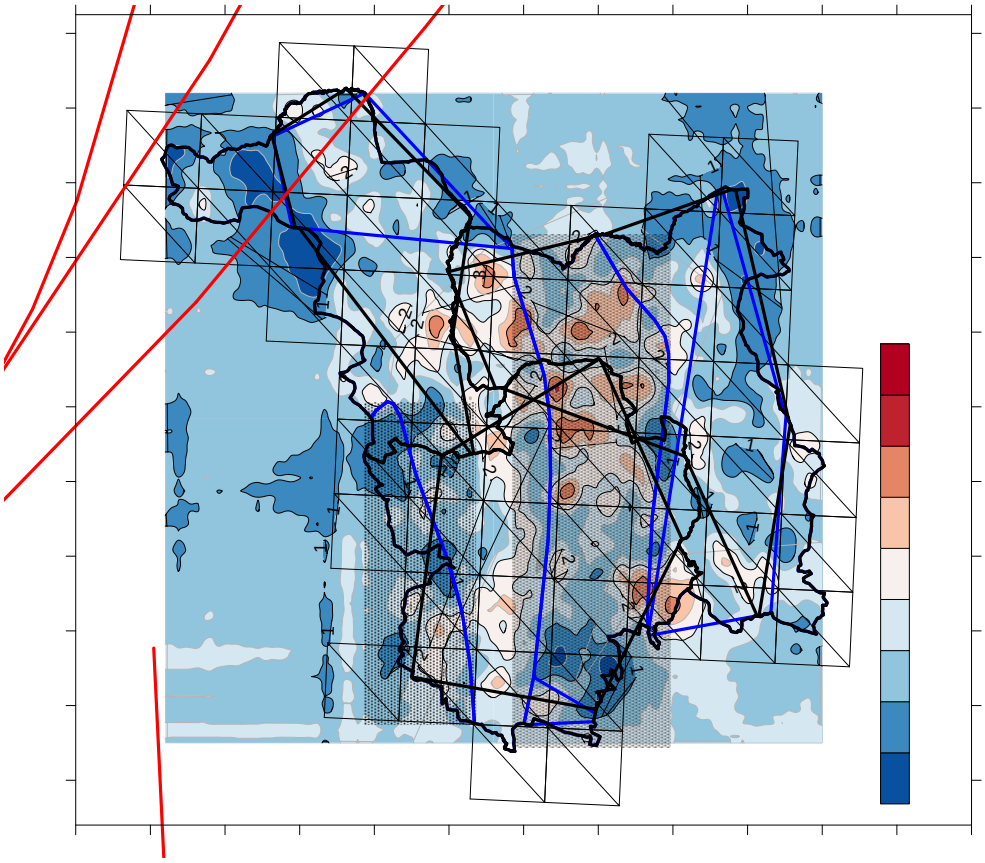

Figure 50: Map of eU data, modelled according to the variogram in Figure 49. The municipality limits and “soilmap” grid are superposed to the eU model. Shaded areas, limited by the blue line, correspond to the “granite bedrock” unit while unshaded areas correspond to the “migmatite, migmatitic paragneiss, coarse grained porphyritic granite with magmatic foliation” unit from the “geology _coarse” layer. The red lines correspond to faults (Austria GK M31 coordinate system).

Analysis of indoor radon concentration (IRC)

Data from both regions of Austria were only considered when appropriate; groups with n=1 observations were excluded from the analysis. The analysis of the indoor radon concentration data indicates that the IRC is statistically different at a 0.05 significance level between room types, earth- and non-earthbound divisions, basement, building, floor and foundation types (Table 26). The results of the multiple comparisons suggest that the data from “basement” is generally different from the data acquired in other divisions. The data from earthbound divisions is generally higher than the data acquired in non-earthbound divisions, which is also reflected in differences between the data according to basement type (“full”, “partial”, “none”). Weekend home data is different from data acquired in other types of buildings.

As a better correlation between the soil/bedrock radon exhalation rate and indoor radon concentration of earthbound divisions is expected, IRC earthbound data were analysed according to soil data properties (Table 27). The analysis of the data, considering both regions, indicates statistically significant differences among different groups of soil type, soil grain size, permeability, soil source, bedrock and water content (Table 27). The results for the AUT North region only indicate statistically significant differences according to the permeability and soil water content.

Omnidirectional variograms for the IRC data set (total and including earthbound data) were computed (Figure 51). No clear spatial correlation is observed considering the arithmetic mean of the data, however, a spatial

correlation is observed considering data from both rooms, particularly clear when only earthbound data is considered.

Table 26: Analysis of indoor radon concentration (IRC) by building characteristics (statistically significant differences are marked in bold).

Variable	IRC	Multiple comparisons
Room type	**H(7;3039) = 41.46, p < 0.001**	"Basement" different than "bed room", "kitchen", "living room", "child's room", "dining room", "home office"
Earthbound Room	**H(1;2924) = 272.15, p < 0.001**	"yes" and "no" are different
Floor	**H(6;3188)=171.90, p < 0.001**	"-1" different from "0", "1" and "2"; "1" different from "0".
Basement	**H(2,3201) = 234.39, p < 0.001**	"fully" different from "partly" and "no"
Building type	**H(4;3167) = 19.25, p = 0.0007**	"weekend home" different from "one family dwelling", "farm" and "public building";
Building foundation type	**H(3,2824) = 91.16, p < 0.001**	"Foundation fully" different from "strip foundation", "no foundation" and "foundation partly"; "strip foundation" different from "no foundation"
Building floor type	**H (4;3006) = 78.72, p < 0.001**	"brick" different from "screed" and "tural and concrete"; "screed" different from "tural (sand, soil)"
Building neighbour	**H(1;3198) = 16.52, p < 0.001**	"built together" different from "solitary"
H – Kruskal-Wallis H test; the degrees of freedom and number of data are indicated within brackets, respectively.		

Table 27: Analysis of indoor radon concentration (IRC, earth bound rooms) by soil and bedrock type, permeability and soil water content (statistically significant differences are marked in bold).

Variable	IRC, earthbound rooms (both regions)	IRC earthbound rooms (AUT North)
Soil type	**H(3;555) = 8.36, p = 0.0392**	H(2;392) = 3.14, p = 0.2080
Soil grain size	**H(3;555) = 11.50, p = 0.0093**	H(2;392) = 3.14, p = 0.2080
Permeability	**H(3;555) = 19.63, p < 0.001**	**H(2;392) = 26.28, p < 0.001**
Soil source	**H(10;555) = 23.78, p = 0.0082**	H(3;392) = 3.57, p = 0.3112
Bedrock (g_fine)	**n.d.**	H(5;393) = 10.49, p = 0.0624
Bedrock (g_coarse)	**H(4;555) = 21.57, p < 0.001**	H(1;392) = 0.69, p = 0.4045
Soil water content	**H(4;555) = 14.02, p = 0.0072**	**H(3;392) = 29.47, p < 0.001**
H – Kruskal-Wallis H test; the degrees of freedom and number of data are indicated within brackets, respectively; n.d. – not determined.		

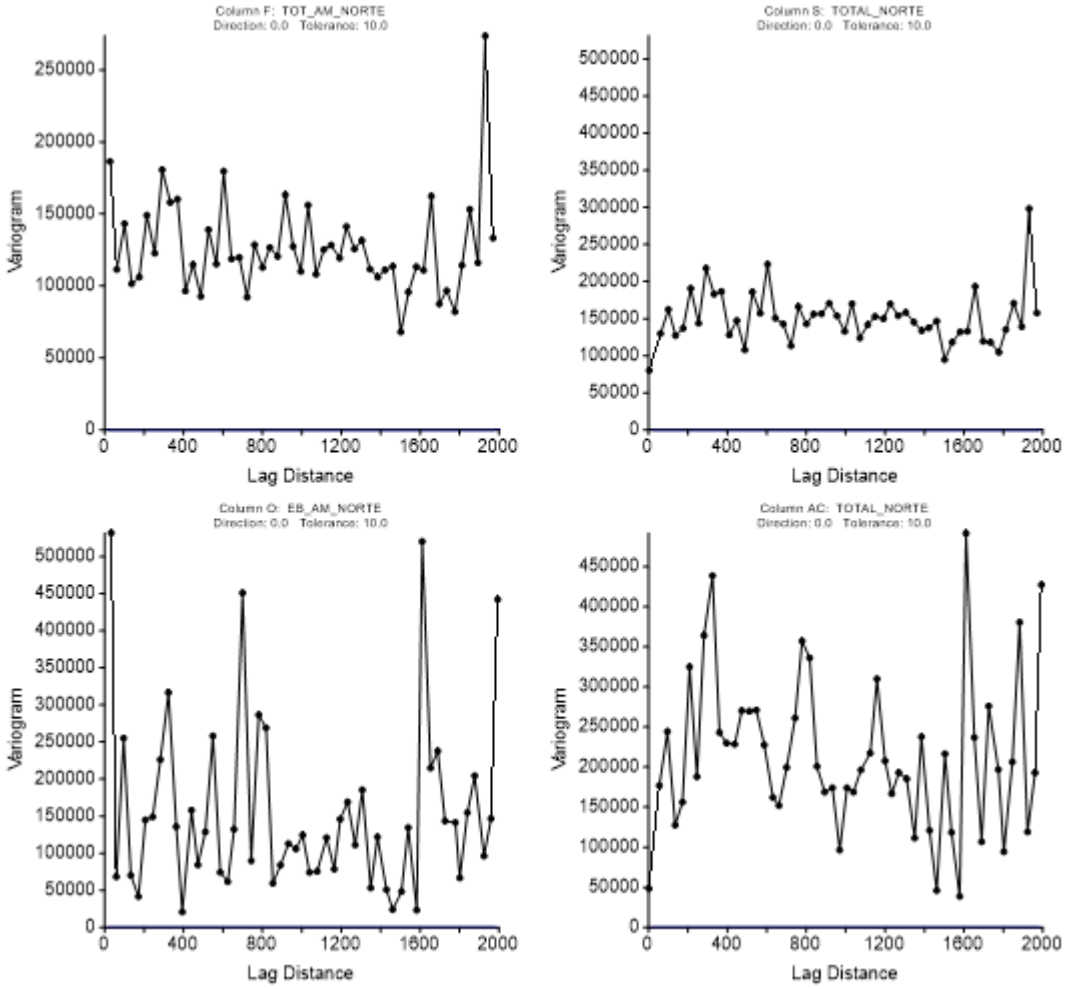

Figure 51: Omnidirectional semi-variograms of the arithmetic mean of IRC data (top left); room 1 and 2 data combined (top right), arithmetic mean of earthbound IRC data (bottom left) and room 1 and 2 data earthbound data (bottom right).

Discussion of results – Data Set Austria North

The results of **ADR** and **eU** indicate that there are statistically significant differences among different soil sources and bedrock units (of the "geology_fine" layer) while the results of **radon concentration in soil gas** indicate significant differences among soil type, grain size and water content. Apart from K-40 data comparison according to bedrock units (of the "geology_coarse" layer), the results of **permeability** and the **radionuclide content** are not statistically different among different soil types and sources, bedrock units and water content.

While ADR and eU are representative measurements of the superficial portion of the media, the radionuclide content, permeability and soil gas radon are representative of a deeper portion of the soil. As the classification of soil properties and bedrock refers to the outcropping portion of these units and because sampling took place in profiles with a depth of 1 m, the lack of statistically significant differences between the radionuclide data among different soil properties may be due to the lack of representativeness of those properties with increasing depth. This could indicate that the content of radionuclides may not be representative of the superficial portion of the soil, due to sampling of different horizons of the soil along the 1 m profile. While K-40, Ra-228 and Th-228 are correlated with the ADR thus indicating otherwise, no correlation is observed between eU and the radioisotopes of the U decay chain. This suggests that the content of U may be less representative of the superficial portion of the soil compared to other radioactive families, likely due to the higher mobility of U.

Permeability data acquired *in situ* is not statistically different among soil types, bedrock units or other soil properties. The variogram of permeability also shows a lack of spatial correlation. This implies that permeability data is site-specific, hence, difficult to model (interpolate or extrapolate).

Soil gas radon presents significant positive correlations with U-238, however, no correlation is observed with Ra-226 despite the strong correlation the later presents with U-238. The correlation observed between U-238 and Ra-226 is stronger compared to the correlation observed between Ra-226 and Pb-210. This indicates that disequilibrium in the U-238 decay chain is more intense in the last portion of the decay chain, likely due to radon migration.

The omnidirectional **variograms** for the radionuclides (Pb-210, Ra-226 and U-238), soil gas radon, permeability, ADR, the calculated TGDR and eU show that, apart from the eU data, no clear spatial correlation is observed. This is either due to the lower number of data or to the fact that the sampling interval is greater than the scale of spatial variation of the data. In fact, a high degree of variability of the concentration of eU is observed in the study area (**Fehler! Verweisquelle konnte nicht gefunden werden.**), thus, a high variability of the geogenic radon potential is expected. The eU data variogram may, however, have been altered following all correction and smoothing processes (ex. altitude, topographic, vegetation, cosmic ray and radon corrections, and the Compton Effect), leading to an increase of the spatial correlation.

Indoor radon data was evaluated according to the building and soil properties (Table 26, Table 27). The analysis shows statistically significant differences between room types (where basement is usually different), earth- and non-earthbound divisions, building, floor and foundation types. As earthbound divisions present generally higher IRC than the data acquired in non-earthbound divisions and as a better correlation between soil gas radon exhalation rate and IRC data of earthbound divisions is expected, the analysis of IRC data excluding non-earthbound data was carried out. IRC of **earthbound data** show statistically significant differences among different sources of the soil, soil water content, and/or bedrock units, permeability and soil type and soil grain size including all data. For the northern region, statistically significant differences are observed according to permeability and soil water content.

Austria: Southern Region (AUT South)

Analysis of soil data (acquired by physical sampling)

According to soil type, soil grain size, soil source and bedrock, the data analysis shows that the content of Ra-226, Pb-210, U-238 and K-40 is not statistically different (significance level of 0.05) between the various types of soil and bedrock. This also applies for soil source types (Table 28). Ra-228 and Th-228 are both statistically different between different soil grain size and soil source.

The ADR is statistically different among different bedrock types. The ADR is higher on orthogneiss followed by the remaining bedrock types (mica-schist, paragneiss, carbonate rocks, siliciclastic, porphyry, marls, sands, gravel and limestone). As the ADR is significantly influenced by the bedrock type, the TGDR was computed for comparing the content of radionuclides according to soil type, soil source and bedrock type. The TGDR was computed from U-238, Th-232 (assuming secular equilibrium between Th-228 and Th-232) and K-40 activity concentration [Bq/kg] according to the equation 8. The results of the Kruskal-Wallis test show the lack of statistically significant differences of the TGDR among groups.

According to soil type, soil grain size, soil source and bedrock, the data analysis shows that both, the radon concentration in soil gas and permeability are not statistically different (significance level of 0.05) between the various types of soil, soil grain size and soil water content (Table 29).

The correlation between soil gas radon, radionuclide concentration, ADR and TGDR was evaluated in Table 30. Soil gas radon is significantly correlated with U-238 and Ra-226. The ADR is correlated with Ra-226 and the TGDR. The correlations between Pb-210, Ra-226 and U-238 activity concentration are significant which indicates an equilibrium in the U decay chain. The correlation of U-238 with Ra-226 is higher than its correlation with Pb-210. Ra-228 and Th-228 are strongly correlated, indicating an equilibrium in the Th-232 decay chain. The TGDR is correlated with all isotopes, including soil gas radon, and with ADR.

The omnidirectional variograms for the radionuclides Ra-226, U-238, Pb-210 and K-40, soil gas radon, permeability ($x10^{10}$), ADR and the calculated TGDR were computed for evaluating of spatial correlations within the data set (Figure 52). No clear spatial correlation is observed, however, the variograms of ADR and TGDR suggest a regional trend.

Table 28: Analysis of radionuclide concentration data, terrestrial gamma dose rate (TGDR) and ambient dose rate (ADR) by soil and bedrock type (statistically significant differences are marked in bold).

Variable	Soil type	Soil grain size	Soil source	Bedrock (geology_coarse)
Ra-226	H(2;78) = 0.6355; p = 0.7278	H(1;78) = 0.1095; p = 0.7407	H(5;78) = 2.5356; p = 0.7711	H(3;78) = 2.9471; p = 0.3999
U-238	H(2;78) = 0.7682; p = 0.6811	H(1;78) = 0.0007; p = 0.9794	H(5;78) = 3.8102; p = 0.5771	H(3;78) = 3.764; p = 0.2881
K-40	H(2;78) = 2.5118; p = 0.2848	H(1;78) = 1.2586; p = 0.2619	H(5;78) = 7.8251; p = 0.1661	H(3;78) = 0.4535; p = 0.9290
Ra-228	H(2;78) = 2.1581; p = 0.3399	**H(1;78) = 5.3653; p = 0.0205**	**H(5;78) = 12.2718; p = 0.0312**	H(3;78) = 2.3585; p = 0.5014
Th-228	H(2;78) = 1.6092; p = 0.4473	**H(1;78) = 6.6286; p = 0.0100**	**H(5;78) = 12.3222; p = 0.0306**	H(3;78) = 1.9137; p = 0.5905
Pb-210	H(2;78) = 3.1064; p = 0.2116	H(1;78) = 0.0225; p = 0.8808	H(5;78) = 3.5023; p = 0.6230	H(3;78) = 4.4262; p = 0.2190
TGDR (calc)	H(2;78) = 1.2955; p = 0.5232	H(1;78) = 0.093; p = 0.7604	H(5;78) = 6.0924; p = 0.2973	H(3;78) = 0.696; p = 0.8742
ADR	H(2;84) = 0.6891; p = 0.7085	H(1;84) = 0.0017; p = 0.9670	H(5;84) = 2.7345; p = 0.7408	**H(3;84) = 12.678; p = 0.0054**
H – Kruskal-Wallis H test; the degrees of freedom and number of data are indicated within brackets, respectively.				

Table 29: Analysis of soil gas radon and permeability data by soil and bedrock type and water content (statistically significant differences are marked in bold).

Variable	Soil type	Soil grain size	Soil source	Bedrock (g_coarse)	Soil water content
Soil gas radon	H(2;84) = 0.1702; p = 0.9184	H(1;84) = 2.0023; p = 0.1571	H(5;84) = 7.7606; p = 0.1699	H(3;84) = 6.5767; p = 0.0867	H(3;84) = 5.0991; p = 0.1647
Permeability	H(2;84) = 1.7576; p = 0.4153	H(1;84) = 0.7104; p = 0.3993	H(5;84) = 9.3292; p = 0.0966	H(3;84) = 3.2349; p = 0.3568	H(3;84) = 5.9129; p = 0.1159
H – Kruskal-Wallis H test; the degrees of freedom and number of data are indicated within brackets, respectively.					

Table 30: Spearman rank correlation matrix. Correlation coefficients are statistically significant at a 0.05 significance level and indicated in red.

	Rn-222 [kBq/m^3]	ADR	K-40 [Bq/kg]	Pb-210 [Bq/kg]	Ra-226 [Bq/kg]	Ra-228 [Bq/kg]	Th-228 [Bq/kg]	U-238 [Bq/kg]	TGDR
Rn-222 [kBq/m^3]	1.00								
ADR	0.33	1.00							
K-40 [Bq/kg]	0.54	0.10	1.00						
Pb-210 [Bq/kg]	0.25	0.18	0.42	1.00					
Ra-226 [Bq/kg]	0.36	0.27	0.44	0.72	1.00				
Ra-228 [Bq/kg]	0.12	0.17	0.22	0.15	0.40	1.00			
Th-228 [Bq/kg]	0.06	0.17	0.18	0.18	0.42	0.94	1.00		
U-238 [Bq/kg]	0.24	0.14	0.48	0.60	0.72	0.42	0.46	1.00	
TGDR	0.40	0.22	0.75	0.58	0.69	0.62	0.63	0.83	1.00

Analysis of indoor radon concentration (IRC)

Data from both regions of Austria were considered when appropriate; groups with n=1 observations were excluded from the analysis. The analysis of the IRC data presented in As a better correlation between the soil/bedrock radon exhalation rate and indoor radon concentration of earthbound divisions is expected, IRC earthbound data were analysed according to soil data properties (Table 27). The analysis of the data, considering both regions, indicates statistically significant differences among different groups of soil type, soil grain size, permeability, soil source, bedrock and water content (Table 27). The results for the AUT North region only indicate statistically significant differences according to the permeability and soil water content.

Omnidirectional variograms for the IRC data set (total and including earthbound data) were computed (Figure 51). No clear spatial correlation is observed considering the arithmetic mean of the data, however, a spatial correlation is observed considering data from both rooms, particularly clear when only earthbound data is considered.

Table 26 indicate statistically significant differences between room types, earth- and non-earthbound divisions, basement, building, floor and foundation types. Thereby, similar to the analysis carried out for AUT North, IRC data were analysed according to the soil data properties after separation of earthbound data (Table 31).

The analysis of the IRC data, considering both regions, indicates statistically significant differences among different groups of soil types, soil grain size, permeability, soil source, bedrock and water content (Table 27, Table 31). The results for the AUT South region show no statistically significant difference between the various soil properties and bedrock units (Table 31).

Table 31: Analysis of indoor radon concentration (IRC, earthbound rooms) by soil and bedrock type, permeability and soil water content (statistically significant differences are marked in bold).

Variable	IRC earthbound rooms (both regions)	IRC earthbound rooms (AUT South)
Soil type	**H(3;555) = 8.36, p = 0.0392**	H(2;163) = 4.98, p = 0.0830
Soil grain size	**H(3;555) = 11.50, p = 0.0093**	H(1;163) = 0.45, p = 0.5045
Permeability	**H(3;555) = 19.63, p < 0.001**	H(2;163) = 3.05, p = 0.2175
Soil source	**H(10;555) = 23.78, p = 0.0082**	H(7;163) = 8.23, p = 0.3129
Bedrock (g_coarse)	**H(4;555) = 21.57, p < 0.001**	H(2;163) = 4.84, p = 0.0890
Soil water content	**H(4;555) = 14.02, p = 0.0072**	H(2;163) = 0.31 p = 0.8577
H – Kruskal-Wallis H test; the degrees of freedom and number of data are indicated within brackets, respectively; n.d. – not determined.		

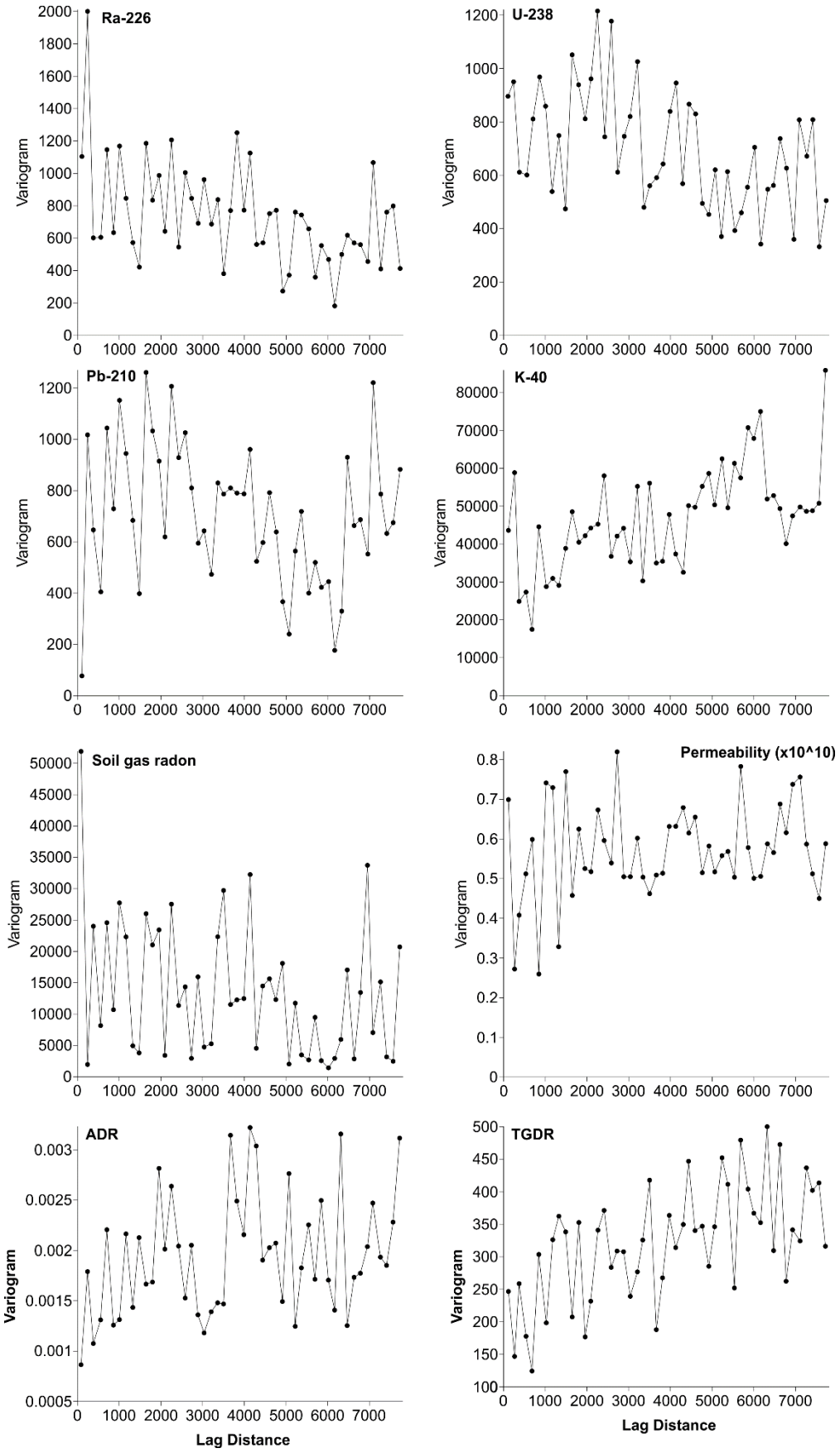

Figure 52: Omnidirectional semi-variograms of Ra-226, U-238 and Pb-210 activity concentration, airborne eU data, soil gas radon, permeability, ADR and TGDR.

The omnidirectional variograms for the IRC data set (total and including earthbound data) were computed (Figure 54). No clear spatial correlation is observed considering either the total data set (top) or only earthbound data (bottom), thus the data are spatially independent.

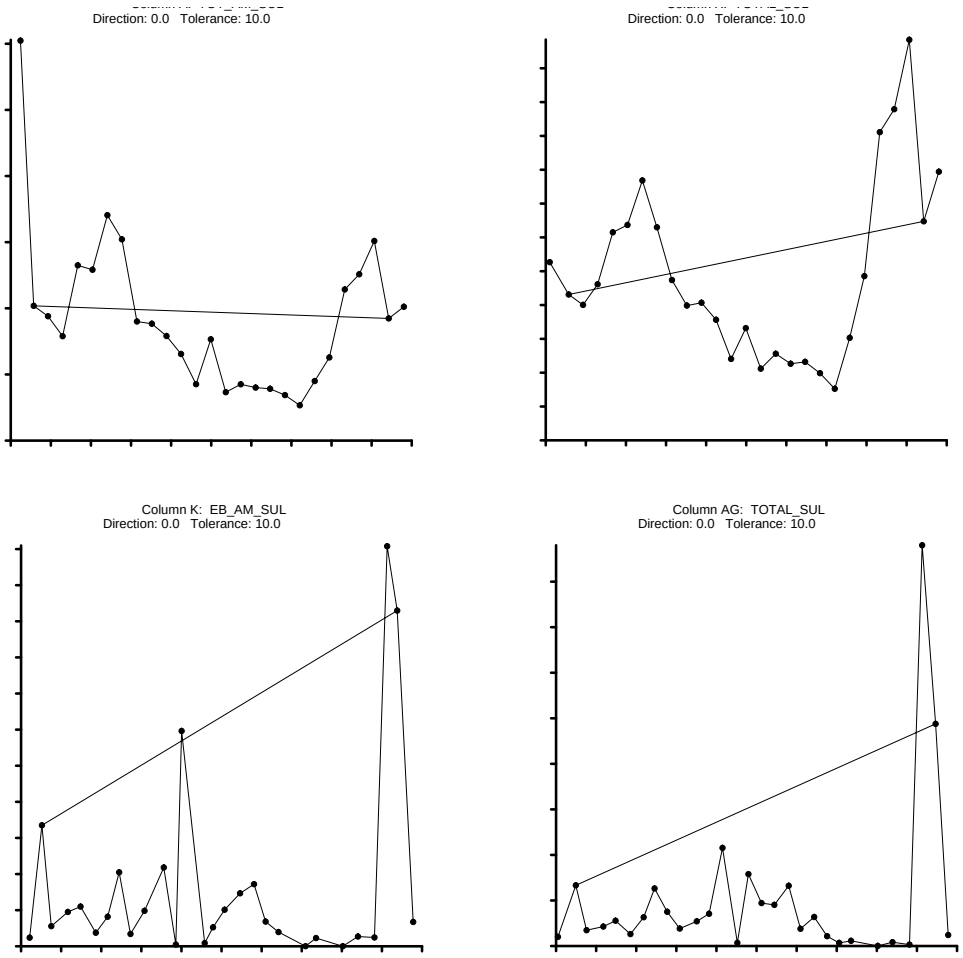

Figure 53: Omnidirectional semi-variograms of the arithmetic mean of IRC data (top left); room 1 and 2 data combined (top right), arithmetic mean of earthbound IRC data (bottom left) and room 1 and 2 data earthbound data (bottom right).
Discussion of results – AUT South

Discussion of results – Data Set Austria South

The results of **ADR** indicate that there are statistically significant differences among different bedrock units (of the "geology_coarse" layer) while the results of **soil gas radon** do not indicate significant differences among soil type, grain size or water content. Apart from Ra-228 and Th-228 data, the results of **radionuclide content** and **permeability** are not statistically different among different soil types, sources and bedrock units.

Similar to results of the discussion from the AUT North region, there is a lack of statistical differences of radionuclide composition regarding soil types. This could furthermore indicate the lack of representativeness of those properties at depth or the lack of representativeness of the radionuclide data in the superficial layer of the soil due to sampling of different horizons along the 1 m profile.

Soil gas radon presents significant positive correlations with U-238, Ra-226 and Pb-210. There is a higher correlation between U-238 and Ra-226 compared to the correlation between U-238 and Pb-210, suggesting a

higher degree of disequilibria towards the end of U-238 decay chain, likely due to radon exhalation from the ground.

Similar to results of the discussion from AUT North, the omnidirectional **variograms** for radionuclides (K-40, Pb-210, Ra-226 and U-238), soil gas radon, permeability, ADR and the calculated TGDR show no clear spatial correlation. This is either due to less data or because of the fact that the sampling interval is greater than the scale of spatial variation of the data.

The permeability data acquired *in situ* is not statistically different among soil or bedrock types and the variogram of permeability furthermore shows a lack of spatial correlation. This implies that permeability data is site-specific, hence, difficult to model (interpolate or extrapolate).

Indoor radon data of **earthbound data** for AUT South does not show statistically significant differences among different sources of soil, water content and/or bedrock units, permeability and soil type and grain size including all data, contrasting with AUT North. The incompatibility of the data and the lack of spatial dependence constrains the use of geostatistical tools to interpolate the data and predict the geogenic radon potential.

Cantabria

Analysis of soil data (acquired by physical sampling)

The analysis of ADR, soil gas radon and IRC data according to bedrock, soil source, permeability and karst is shown in Table 32. Both the ADR and radon concentration in soil gas present statistically significant differences among different bedrock units, soil sources and permeability. Indoor radon concentration behaves differently, according to bedrock type and the presence or absence of karst. The indoor radon concentration is higher when karst is present. Radon concentration in soil gas is statistically not different and therefore not influenced by the presence or absence of karst (Table 32). Glacier deposits, dolomitic rocks and the F. Bundsandstein present higher ADR than the remaining bedrock units. The radon concentration in soil gas is higher in the "dolomite, calcarenite" unit, followed by the "limestone, limestone of Picos" unit. The "Silts, clay, organic material and salt", "clay" and "dolomite, calcarenite" present the highest IRC.

Table 32: Analysis of ambient dose rate (ADR), soil gas radon and indoor radon concentration (IRC) data by karst, bedrock type and permeability (statistically significant differences are marked in bold).

Variable	Lithology	Source	Permeability	Karst
Ambient dose rate	**H(18;62) = 33.1549; p = 0.0160**	**H(3;62) = 10.0935; p = 0.0178**	**H(4;62) = 9.9015; p = 0.0421**	H(1;77) = 0.5702; p = 0.4502
Soil gas radon	**H(27;259) = 43.518; p = 0.0232**	**H(4;259) = 10.9856; p = 0.0267**	**H(4;259) = 9.7716; p = 0.0445**	H(1;260) = 0.1338; p = 0.7146
IRC	**H(25;482) = 43.172; p = 0.0134**	H(4;482) = 7.8012; p = 0.0991	H(5;482) = 5.6215; p = 0.3448	**H(1;482) = 4.9472; p = 0.0261**
H – Kruskal-Wallis H test; the degrees of freedom and number of data are indicated within brackets, respectively.				

The Spearman rank correlation coefficient between ADR, soil gas radon and IRC was computed in Table 33. A small positive correlation is observed between soil gas radon and IRC, whereas a negative correlation between soil gas radon and ADR is observed when the closest point is chosen for the comparison of the data.

Table 33: Spearman rank correlation between soil gas radon, ambient dose rate (ADR) and indoor radon concentration (IRC) (statistically significant correlations are marked in bold).

	Average of the closest points	**Closest point**
IRC x ADR	r(224) = 0.04, p = 0.5035	r(482) = -0.02, p = 0.7350
IRC x Soil gas radon	r(276) = 0.04, p = 0.5480	**r(482) = 0.11, p = 0.0099**
Soil gas radon x ADR	r(68) = -0.14, p = 0.2481	**r(260) = -0.18, p = 0.0030**
Soil gas radon x IRC	r(113) = 0.13, p = 0.1668	r(260) = 0.02, p = 0.6997
ADR x IRC	r(60) = -0.10, p = 0.4630	r(77) = -0.19, p = 0.0922
ADR x Soil gas radon	r(55) = -0.26, p = 0.0586	r(77) = -0.13, p = 0.2468

The Spearman rank correlation coefficient between ambient dose rate, soil gas radon and indoor radon concentration with the radioisotope content in soil (GEMAS and FOREGS data) was estimated. For calculating the correlation, point data (ambient dose rate, soil gas radon and indoor radon concentration) was compared to the isotope concentration of the grid cell that the point falls into. The ambient dose rate presents positive correlation with Th (r(64) = 0.25, p = 0.0499) and K (r(64) = 0.25, p = 0.0496). However, the ambient dose rate is not correlating with U (r(64) = 0.08, p = 0.5307). Soil gas radon presents a significant negative correlation with U content (r(250) = -0.23, p < 0.001). Indoor radon concentration also presents a significant negative correlation with U content (r(482) = -0.13, p = 0.0046).

The omnidirectional variograms for the ADR, soil gas radon and IRC are displayed in Figure 54. The ambient dose rate displays spatial correlation (Gaussian model with 0.5 scale and 10000 length). Soil gas radon and indoor radon concentration data are spatially independent.

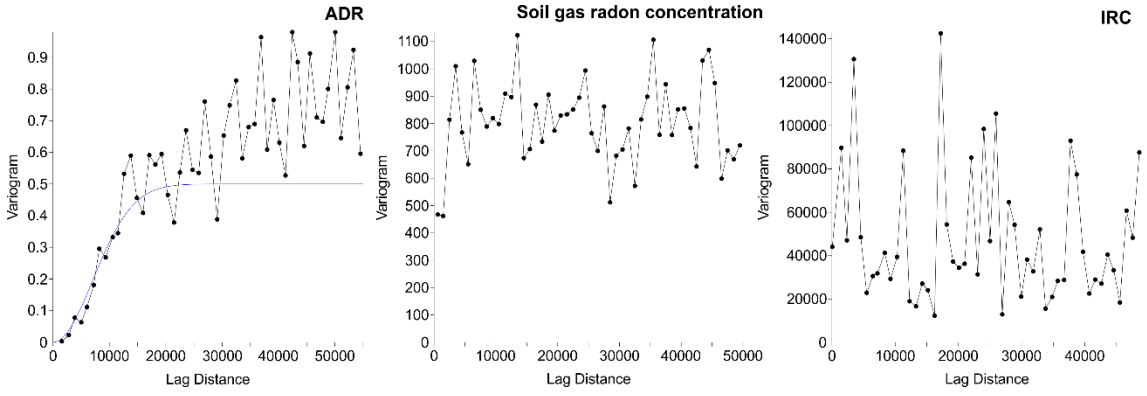

Figure 54: Omnidirectional semi-variograms of the ambient dose rate (ADR), soil gas radon and indoor radon concentration (IRC).

Discussion of results - Cantabria

The results of **ADR, radon concentration in soil gas** and **IRC** indicate that there are statistically significant differences among different bedrock units. ADR, soil gas radon and indoor radon concentration are not correlated. Soil gas radon as well as indoor radon concentration present significant negative correlations with U content, estimated from GEMAS and FOREGS data. This indicates that the data are not compatible. The ADR is correlating with Th and K, but not with U.

The omnidirectional variograms for soil gas radon and indoor radon concentration show no clear spatial correlation which is either due to less data or because of the fact that the sampling interval is greater than the scale of spatial variation. However, the variogram of indoor radon concentration is altered due to the changed location of the dwellings. The variogram of the ADR indicates a spatial dependence of the data. However, ADR is not correlating with U, soil gas radon or indoor radon data. Therefore, ADR cannot be used as a predictor of the remaining data. The incompatibility of the data and the lack of spatial dependence constrains the use of geostatistical tools to interpolate the data and predict the geogenic radon potential.