16ENV10 MetroRADON

Deliverable 4

Report on the results from the on-site comparison of indoor radon measurements and geogenic radon measurements under field conditions together with guidelines/recommendations to assist the implementation of the EU-BSS

European Commission, Joint Research Center (JRC)

Due date: March 2020
Submission: November 2020

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

Table of contents

List of authors (alphabetic order)

S. Baumann[1], P. Bossew[2], , S. Celaya[3], G. Cinelli[4], A. Fernandez[3], E. Fernandez[3], I. Fuente[3], C. Greau[5], V. Gruber[1], G. Ielsch[5], F.J. Maringer[6], E.Petermann[2], R. Pol[3], J. Quindos[3], L. Quindos[3], D. Rabago[3], C. Sainz[3], M. Stietka[6], H. Wiedner[6]

1 Österreichische Agentur für Gesundheit und Ernährungssicherheit GmbH, Austria – AGES

2 German Federal Office for Radiation Protection, Berlin, Germany – BfS

3 Universidad De Cantabria, Spain – UC

4 European Commission, Joint Research Centre, Ispra, Italy – JRC

5 Institut de Radioprotection et de Surete Nucleaire, France – IRSN

6 Physikalisch-Technischer Prüfdienst des Bundesamt für Eich- und Vermessungswesen, Austria – BEV-PTP

Introduction

This document, titled "Report on the results from the on-site comparison of indoor radon measurements and geogenic radon measurements under field conditions together with guidelines/recommendations to assist the implementation of the EU-BSS" represents the deliverable D4 of the MetroRADON project.

It reports the results of the activities developed in Task 3.3 and Task 3.4 of Work Package 3 – WP3: Comparison and harmonisation of radon measurement methodologies in Europe of the EURAMET 16ENV10 MetroRADON project. The report is structured as:

- Introduction on WP3
- Brief summary of activities results of Task 3.3
- Brief summary of activities results of Task 3.4
- Annexes reporting the full results of the activities.

Work Package 3 – WP3

One of the specific objectives of MetroRADON project is to compare existing radon measurement procedures in different European countries and to reduce inconsistency of the indoor radon measurements across the Europe. This objective is addressed within WP3 – "Comparison and harmonisation of radon measurement methodologies in Europe".

WP3 aims to:

- collect and analyse meta-information from radon surveys and existing radon databases in European countries;
- evaluate if the data and methodologies are comparable;
- identify how they could be harmonised in the event of methodical inconsistency.

Work Package 3 is divided into four Tasks:

3.1 Overview and analysis of indoor radon surveys in Europe

The aim of this Task is to analyse and evaluate indoor radon surveys in order (i) to identify the rationale and methodologies used, (ii) to identify the extent and possible sources of inconsistencies in the results of indoor radon surveys and (iii) to propose approaches to reduce inconsistencies and improve harmonisation of indoor radon data.

3.2 Overview and analysis of geogenic radon surveys in Europe

The aim of this Task is to analyse and evaluate geogenic radon surveys in order (i) to identify the rationale and methodologies used, (ii) to identify the extent and possible sources of inconsistencies in the results of outdoor geogenic radon surveys and (iii) to propose approaches to reduce inconsistencies and improve harmonisation of geogenic radon data, in analogy to indoor radon in Task 3.1.

3.3 Intercomparisons of indoor radon and geogenic radon measurements under field conditions

The aim of this Task is to organise an intercomparison of indoor radon measurements and geogenic radon measurements (including radon exhalation rate) under field conditions in order to identify physical reasons for possible inconsistencies, particularly related to sampling and measurement techniques. Three different comparisons will be performed: (i) indoor radon gas (passive and continuous monitoring devices), (ii) radon exhalation from soil and (iii) radon concentration in soil gas.

3.4 Development of options for harmonisation of indoor and geogenic radon data including practical examples

Based on the results for Tasks 3.1, 3.2 and 3.3 the aim of this Task is to develop options for harmonisation of indoor and geogenic radon data, where appropriate using practical examples.

The results of Task 3.1 and 3.2 are summarised in the MetroRADON deliverable D3 (Activity 3.2.6) – "Report on indoor and geogenic radon surveys in Europe, including strategies, the methodologies employed, inconsistencies in the results, and potential methodologies to harmonise data and reduce inconsistencies". The results for Task 3.3 and 3.4 are presented in this report/deliverable available on MetroRADON website: http://metroradon.eu/.

Task 3.3, "Intercomparisons of indoor radon and geogenic radon measurements under field conditions"

Four activities were set up to address the goals of this task (i.e. to carry out the intercomparison exercise). Below the specification of each activity according to the JRP protocol can be found:

A3.3.1: UC will organise an intercomparison based on variable indoor radon conditions along with ambient parameters that will be also subjected to change. The goal of this activity will be to compare the response of traceable and calibrated instruments in reference radon chambers to changing conditions of radon gas concentrations. This task will be carried out in UC's laboratory of natural radiation and UC will develop the protocol / documentation for the intercomparison.

A3.3.2: UC and JRC will select the participants for the comparisons of (i) indoor radon gas (passive and continuous monitoring devices), (ii) radon exhalation from soil and (iii) radon concentration in soil gas. The participants, primarily those partners involved in WP1 and WP5, will choose to either to provide their devices for testing or to carry them to the exercise's site to be tested under real conditions. UC and JRC will select the sites for the indoor and outdoor comparison. The site for the comparison of indoor radon gas will be one room inside UC's laboratory located at Saelices el Chico. The other two comparison exercises will be performed outdoors, therefore JRC and UC will visit the available outdoor areas and choose the most appropriate places. These places may be different for the comparison of radon concentration in soil gas and surface exhalation rate.

A3.3.2: UC and JRC will undertake the measurements of indoor radon gas using passive and continuous monitoring devices in the room inside UC's laboratory located at Saelices el Chico selected in A3.2.2. UC and JRC will undertake the measurements of radon exhalation ((ii) surface exhalation rate and (iii) radon concentration in soil gas) at UC's facilities including the field laboratory at Saelices el Chico selected in A3.2.2.

A3.3.3: UC and JRC will analyse and compare the results from the comparisons of (i) indoor radon gas (passive and continuous monitoring devices), (ii) radon exhalation from soil and (iii) radon concentration in soil gas obtained in A3.2.2. Using this information they will to identify physical reasons for possible inconsistencies, particularly related to sampling and measurement techniques. Based on the information from A3.2.1-A3.2.3, UC and JRC will write a report and a paper on the intercomparison of indoor radon measurements and geogenic radon measurements under field conditions.

Summary of Task 3.3 activity results

Activities A3.3.1-3.3.4 were focused on the intercomparison exercise. UC was responsible for the management of the comparisons with the support of JRC. The selection of participants and the comparison's site, and the data analysis and production of a peer-reviewed paper was undertaken by UC and JRC.

The radon intercomparison measurements were held from 5-8 November 2018 in the Laboratory of Natural Radiation (LNR) located at the facilities of the former uranium mine managed by the Spanish National Uranium Company ENUSA (Ciudad Rodrigo, Salamanca, Spain).

Twenty participants from 13 countries took part in the intercomparison "radon in air", three in the "radon exhalation from soil" and five in the "radon in soil" exercise. Figure 1 shows the countries of origin of the participants. The maximum number of participants, fixed by the organizers for operational reasons, has been reached, showing the great interest from the radon community for such kinds of activities.

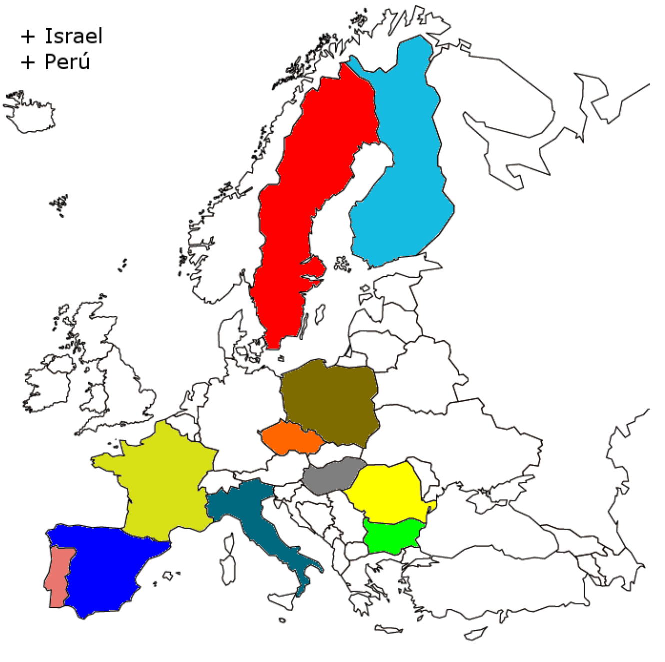

Figure 1 – The figure displays the countries from which the participants come from

Main results of the intercomparison of indoor radon and geogenic radon measurements

Regarding radon in air measurements, a total of 23 groups of passive detectors and 22 active monitors were exposed in Room1 of the LNR with variations of radon concentration from approximately 0.5 kBq m^{-3} to 30 kBq m^{-3}. Two exposures were performed with reference values of E1 = 356 ± 8 kBq m^{-3} h and E2 = 1014 ± 13 kBq m^{-3} h obtained from participant results according to ISO 13528:2015.

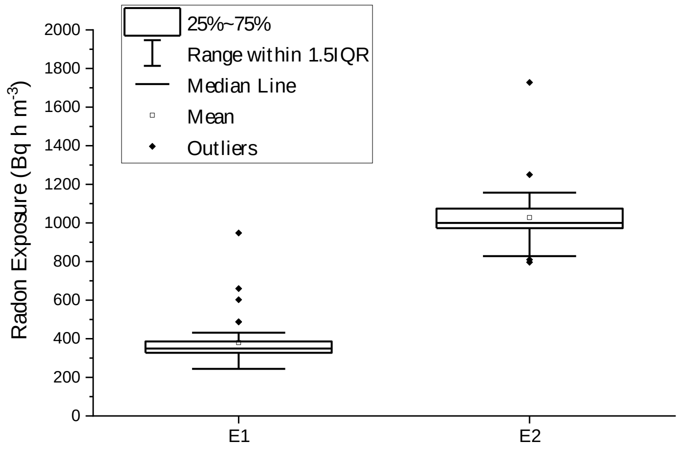

Figure 2 – Boxplot diagram of the participant's results for radon-in-air exposures E1 and E2.

Over 80 % of the results for radon-in-air exposure are within the interval defined by the reference value and the standard deviation, established as 20 % and 10 % for the first and the second exposure, respectively. The exercise was successful, taking into account the large number of different devices used. Especially in passive detectors the holder materials, diffusion chamber volume, detectors area or detection principle were diverse.

Radon exhalation and radon-in-soil measurements were carried out outside the LNR in "Green Ballesteros". Three institutes participated in the radon-exhalation measurements and the results were widespread. This could be explained, as the participants did not perform the exhalation measurements during the same time and weather conditions due to the different methodologies involved. Indeed, a participant used absorption in polycarbonate, a cumulative measurement, with a duration of two days. Another used the accumulation method with a radon monitor with a sample time of approximately one hour. The third participant used the absorption in activated charcoal canisters during 24 hours one week later due to logistics problems.

Radon-in-soil results were carried out in two different sites: "Green Ballesteros" (5 participants) and "Offices site" (3 participants). In case of "Green Ballesteros the dispersions of results is acceptable with a standard deviation of about 14%. However, in "Office site" the result performed with one instrument show a great deviation.

The detailed report on activities 3.3.1-3.3.4 is reported in **Annex 1:** "Intercomparison of indoor radon and geogenic radon measurements under field conditions", also available on the MetroRADON website.

Moreover, the findings of the intercomparison exercise of indoor radon measurements where published in a peer reviewed paper:
Rabago et al., Intercomparison of Indoor Radon Measurements Under Field Conditions In the Framework of MetroRADON European Project. Int. J. Environ. Res. Public Health 2020, 17, 1780. doi:10.2290/ijerph17051780, reported in **Annex 2**.

Task 3.4, "Development of options for harmonisation of indoor and geogenic radon data including practical examples"

Five activities were set up to develop options for harmonization of indoor and geogenic radon data. Below the specification of each activity according to the JRP protocol can be found:

A3.4.1: IRSN, BfS and UC will assess the representativeness of indoor Rn measurements over different sampling periods (different durations and seasons of the year). They will therefore explore under which circumstances and for which purposes seasonal correction factors for estimating radon exposure in dwellings are applicable and can be defined.

A3.4.2: JRC and AGES will test the comparability between existing short-term and long-term radon measurement data in order to assess the feasibility of data merging and therefore data set expansion.

A3.4.3: BfS, JRC and UC will use the results from Tasks 3.1 (A3.1.1-A3.1.5) and Task 3.3 (A3.3.1-A3.3.4) to propose instances where top-down harmonisation of indoor radon data and procedures may be reasonable. If applicable, practical cases will be demonstrated.

A3.4.4: BfS and UC will use the results of Tasks 3.2 (A3.2.1-A3.2.6) and Task 3.3 (A3.3.1-A3.3.4) to propose instances where top-down harmonisation for quantities related to geogenic radon may be reasonable. If applicable, practical cases will be demonstrated.

A3.4.5: Based on the results from A3.3.1-A3.3.4 and A3.4.1-A3.4.4, JRC, BFKH, VINS, AGES, BfS, IRSN and UC will write a report on the results from the on-site intercomparison exercises of indoor radon measurements and geogenic radon measurements under field conditions together with guidelines/recommendations to assist the implementation of the EU BSS. Once the report have been agreed by the consortium, JRC will send the report to the coordinator, who on behalf of JRC, BfS, BFKH, VINS, AGES, UC and IRSN will then submit them to EURAMET as D4 '*Report on the results from the on-site comparison of indoor radon measurements and geogenic radon measurements under field conditions together with guidelines/recommendations to assist the implementation of the EU BSS*'.

Summary of Task 3.4 activities results

In **Activity A3.4.2** the comparability between existing short-term and long-term indoor radon measurement data was tested in order to assess the feasibility of data merging and data set expansion. The exercise was designed to use existing data sets and already available analysis from literature to analyse if different radon measurements techniques in the same area can be incorporated to improve radon risk prediction. The analysis was focused to quantify if different measurement techniques with different measurement durations show similar distributions and if these measurements can be combined to predict radon risk.

In the exercise two different data sets were used and analysed. The data sets were differing in major aspects, like sample strategy or sample size. For one data set measurements were done in 47 dwellings at the same time period with different measurement methods and exposure times (track etch and electret detectors). The other data set features measurement data from different large-scale surveys (over 80.000 measurements) in the same area (Austria), done with three different sampling methods (track etch, electret and activated charcoal), different sampling strategies and at different time periods. So, with this two very differing data sets it was possible to perform interesting analysis and evaluation for the comparability between radon measurement methods, especially short-term/long-term measurements with regard of the feasibility of data merging.

Generally, it could be shown, that the analysis does not support the “out-of-the-box” merging of indoor radon concentration distributions of different measurement techniques to predict radon risk, more so if they come from surveys with different survey strategies. Of course merging distributions might still be a valid strategy if further analysis, as performed in this study, show similarity of these distributions. To extend the knowledge of the comparability of measurement methods, it would be helpful to have more data sets available for future analysis. Therefore, institutions should be encouraged to perform studies where different measurement devices at identical locations measure indoor radon concentration.

The detailed results of this activity are described in **Annex 3**: “Testing of comparability between short-term and long-term radon measurement data”.

Activity **A3.4.3 and A3.4.4** focused on proposing instances where top-down harmonisation for quantities related to indoor radon and geogenic radon may be reasonable.

Collection of radon data is the first step in the development of any kind of radon maps. Several kinds of radon maps can be created that display for example the simple arithmetic mean, % above reference level, radon potential, radon priority map, radon hazard index map. Therefore, the input radon data will strongly influence the output map. Radon data could be: indoor radon, soil gas radon, geogenic radon data (uranium concentration in soil and rock, terrestrial gamma dose rate, soil permeability, radon in water).

Difference and comparability of radon data (both indoor radon and geogenic) and their treatment have been analysed thanks to the results of the WP3 questionnaires on indoor and geogenic radon surveys. The results are available in the deliverable D3 of the MetroRADON project and have been presented during the European Radon Week (Vienna, 2020).

Different surveys – data can be harmonized in two ways as displayed in figure 3:

- Firstly, all surveys are performed methodically identically. Then homogeneity is ensured by construction; this approach is called *bottom-up*, but viable only if surveys are planned jointly from scratch.
- Realistically, surveys are conducted independently (different period, different purpose, different methods, etc.). To compare results, one has no choice but normalize results to a common standard by using models based on physical and statistical knowledge of the procedures which in each case have led to the reported results. This is called *top-down* or posterior harmonization.

In the activity A3.4.2, a practical case for harmonizing indoor radon data have been carried out (see above and **Annex 3**).

Our conclusion is that harmonization of radon data is partly possible and projects like MetroRADON and intercomparison exercises are needed to study the comparability of radon data and develop procedures for harmonization. However, it is not an easy task and further studies on comparability and harmonization should be performed in the future.

Annex 4 reports the results of A3.4.3 and A3.4.4 activities in details and indicates some need for further research.

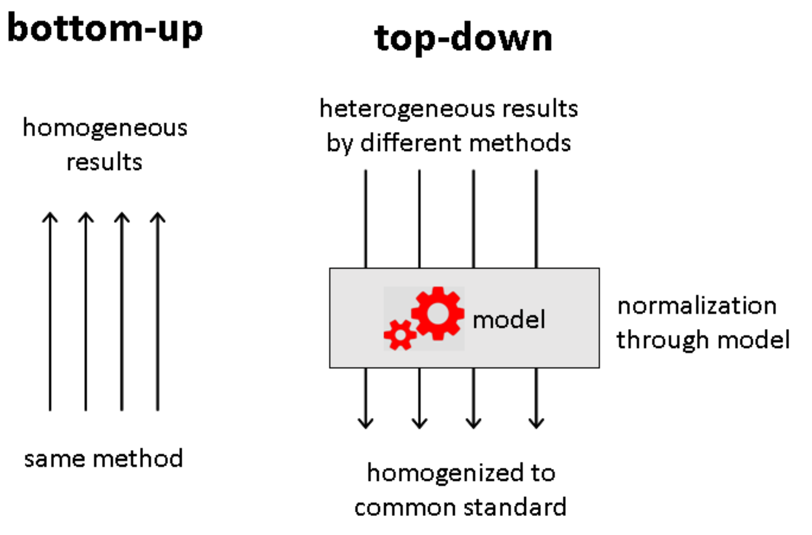

Figure 3 – Bottom-up and top-down harmonization

Guidelines/recommendations to assist the implementation of the Council Directive 2013/59/Euratom (EU-BSS)

Harmonisation and comparability

In the attempts to harmonize available data, both on indoor radon and on indoor radon related geogenic variables, the "top-down" approach has proven to be complex. Despite the usefulness of this approach to try to establish a general picture or diagnosis of the current situation related with radon exposure in Europe based on different surveys, it is necessary to develop a set of common methodologies ("bottom-up" approach) that, implemented in the different National Radon Action Plans, allow obtaining fully comparable results.

For the development of common methodologies indicated in the previous section, it is still necessary to continue the research on the validity of the use of proxies to indoor radon concentration. This research should be carried out mainly with new and extensive surveys directly designed for the task that allow increasing the available statistics and establishing multivariate models between the variables analyzed.

Requirements and design of measurements

The estimation of the annual average radon concentration from data referring to shorter measurement periods (weeks, months) continues to be a relevant research topic, taking into account the great variability and uncertainty that the seasonal factors provided in different studies in Europe show. The comparability of short-term and long-term measurements should be tested further, as if representative measurement results also would be achievable with shorter measurement duration, would be helpful, especially for fast evaluation of workplaces, building transactions etc. To extend the knowledge and analysis on this, it would help to have more data sets available. So, countries and their measurement institutes are encouraged to perform studies, where measurements with different measurement devices and different measurement durations at identical locations are done. As this data are directly comparable and do not need to take into account other factors in the evaluation, like different sampling locations or different sampling designs, which makes the analysis much more complex and uncertain.

Until this issue is not satisfactorily remedied, one-year radon concentration measurements are recommended as the best option for optimal adaptation to EU-BSS requirements. For applying shorter measurement periods, it needs to be evaluated for the specific situation (country or region specific or specific for the measurement situation (e.g. at workplaces)), if representative for the annual average radon concentration. Some countries apply shorter measurement periods in representative period of time, e.g. Austria with 6 months measurement duration, half winter and half summer, representative for the annual situation.

For new radon surveys, e.g. for mapping purposes or to evaluate the radon situation, it is recommended to carry out long-term measurements with passive track-etch detectors, as mentioned above. A uniform survey and sampling/measurement strategy should be applied for the entire survey, either for the area of interest or the specific aim of the survey (e.g. national survey, measurement in all schools, measurements in specific buildings etc.).

Data processing

If several different data sets exist in a country or for a region with different survey and/or sampling design already, the idea might occur, to merge the data sets to increase the total number of data or for a better coverage of the area - to improve radon risk prediction or for more detailed radon mapping. It can not be answered or recommended ad hoc or in general, if data merging is feasible or not, as it depends on many different factors, especially if the data come from different surveys with different survey strategies or different measurement methods (e.g. geographical sampling vs. random sampling, workplaces vs. dwellings, different measurement methods, different time period of measurement etc.). In general, before merging data sets, some analysis and tests should be carried out (e.g. comparison of similarity of distributions) and one should be aware of the problems of data merging when using and interpreting the merged data set.

Instruments and measurement techniques

Currently, there is a wide variety of instruments and techniques for measuring radon in air activity concentration. The basic classification criteria are the integration period (short period vs. long period) and the distinction between continuous measurements and measurements that provide an average value of the radon activity concentration (integrated measurements). Each of the techniques has many different purposes and capacities, so it would be necessary to establish guidelines for the use of different techniques in each National Radon Action Plan depending on the purpose of the measurements.

QA and metrological infrastructure

The great variety of instrumentation available today for the measurement of radon represents a significant challenge for the quality assurance (QA) of the provided results. QA is essential for both practical and effective prevention of radon exposure in buildings, workers protection, and also for the potential legal implications that measures may have in regulatory non-compliance situations.

It is essential in Europe to maintain metrological infrastructur related to the QA of measurements, which serve to continuously test, approve, verify and calibrate the instrumentation used in any type of radon study. These infrastructures, both those that operate under fully controlled conditions, as well as those that offer references (primary and secondary standards), comparisons and proficiency tests under field conditions, should be considered strategically and budgetary in the EU-BSS' required National Radon Action Plans.

List of Annexes:

- **Annex 1:** Report: *Intercomparison of indoor radon and geogenic radon measurements under field conditions*. Rabago, D., Quindos, L., Quindos, J., Fernandez, E., Fernandez, A., Celaya, S., Fuente, I., Sainz, C.

- **Annex 2:** Rabago, D., Fuente, I., Celaya, S., Fernandez, A., Fernandez, E., Quindos, J., Pol, R., Cinelli, G., Quindos, L., Sainz, C. Intercomparison of Indoor Radon Measurements Under Field Conditions In the Framework of MetroRADON European Project. Int. J. Environ. Res. Public Health 2020, 17, 1780. doi:10.3390/ijerph17051780

- **Annex 3:** Report: *Testing of comparability between short-term and long-term radon measurement data.* S. Baumann, V. Gruber, G. Cinelli.

- **Annex 4:** Report on activities A3.4.3 and A3.4.4

Annex 1: Report: *Intercomparison of indoor radon and geogenic radon measurements under field conditions*. Rabago, D., Quindos, L., Quindos, J., Fernandez, E., Fernandez, A., Celaya, S., Fuente, I., Sainz, C.

Intercomparison of indoor radon and geogenic radon measurements under field conditions

WP3: Comparison and harmonisation of radon measurement methodologies in Europe

Task 3.3: Intercomparisons of indoor radon and geogenic radon measurements under field conditions

Activities: A3.3.1 to A3.3.4

Daniel Rábago
Luis S. Quindós
Jorge Quindós
Enrique Fernández
Alicia Fernández
Luis Quindós
Santiago Celaya
Ismael Fuente
Carlos Sainz

Contact: daniel.rabago@unican.es
carlos.sainz@unican.es

February 2019

University of Cantabria (UC)
JRP EMPIR 16ENV10: MetroRADON
Metrology for radon monitoring

CONTENTS

1 Introduction

This report describes the activities carried out within Task 3.3 “Intercomparison of indoor radon and geogenic radon measurements under field conditions” of the MetroRADON project. The aim is to provide a direct comparison between different methodologies and to identify physical reasons for possible inconsistencies, particularly related to sampling and measurement techniques. Three different comparison exercises were performed under field conditions in order to identify physical reasons for possible inconsistencies, particularly related to sampling and measurement techniques. The main exercise was the comparison of indoor radon gas measurements performed with passive detectors, giving an integrated measurement over time, and active monitors, continuously monitoring radon concentration. A series of geogenic radon measurements, such as radon exhalation rate from soil and radon concentration in soil gas, was also conducted.

This intercomparison exercise was organized by the University of Cantabria (UC) with the support of the Joint Research Centre (JRC).

The radon intercomparison measurements were held from 5-8November 2018 in the Laboratory of Natural Radiation (LNR) located at the facilities of the former uranium mine managed by the Spanish National Uranium Company ENUSA (Address: Ctra. Ciudad Rodrigo - Lumbrales, km 7. 37592 Saelices el Chico, Salamanca, Spain).

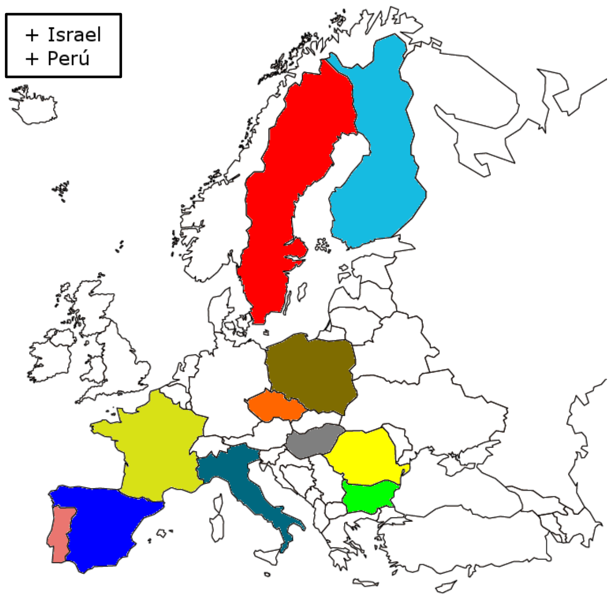

Fig. 1. Participants from the coloured European countries as well as Israel and Peru participated in the intercomparison.

1.1 Participants

The number of participants was limited to 20 due to operational reasons. All participants took part in the exercise related to radon in air measurements, whereas exhalation from soil and radon in soil measurements were performed by 3 and 5 participants, respectively. The questionnaire sent to participants to request their contact information and participation details is shown in Appendix II. The list of participants is given in Table 1. There is no correlation between this table and the code assigned to each participant in the results section. Fig. 1 shows coloured European countries of participants.

Table 1. Participants in the intercomparison sorted by alphabetical order.

Acronym	**Institution**	**Country**
CIEMAT	Centro de investigaciones energéticas, medioambientales y tecnológicas	Spain
CLOR	Central Laboratory for Radiological Protection	Poland
ENEA	ENEA Radon Service	Italy
INAIL	Italian National Institute for Insurance against Accidents at work	Italy
IRSN	Institut de Radioprotection et de Sûreté Nucléaire	France
JRC	Joint Research Centre	Italy
LaRUC (UC)	Laboratory of environmental radioactivity, University of Cantabria	Spain
LRAB - UEX	LRAB - Universidad de Extremadura	Spain
LRG	Laboratorio de Radón de Galicia	Spain
LRN-UC	Laboratorio de Radioatividade Natural - Universidade de Coimbra	Portugal
NRCN	Nuclear Research Center Negev	Israel
PUCP	Pontificia Universidad Católica Del Perú	Peru
RADONOVA	Radonova Laboratories AB	Sweden
Radosys	Radosys / Radosys Atlantic	Portugal/Hungary
RERA-CIEMAT	Centro de investigaciones energéticas, medioambientales y tecnológicas	Spain
STUK	Radiation and Nuclear Safety Authority	Finland
SUBG	Sofia University "St. Kliment Ohridski"	Bulgaria
SUJCHBO	National Institut for NBC Protection	Czech Republic
TR	TECNO RAD s.u.r.l.	Italy
UBB	Babes-Bolyai University	Romania

1.2 Site Description

The intercomparison was carried out in the former uranium mine managed by the Spanish National Uranium Company ENUSA. The reclamation of the uranium mining operations (exploited from 1972 to 2000) and the dismantling of the uranium concentrate factory started in 2001. The purpose of this reclamation is to try to restore the affected natural space and to recover it to its original state, with radiological and environmental conditions returning to those existing before the mining operations. One of the buildings was chosen to house the Laboratory of Natural Radiation (LNR) for calibration and testing of instruments and detectors for the measurement of natural radiation (see Fig. 2).

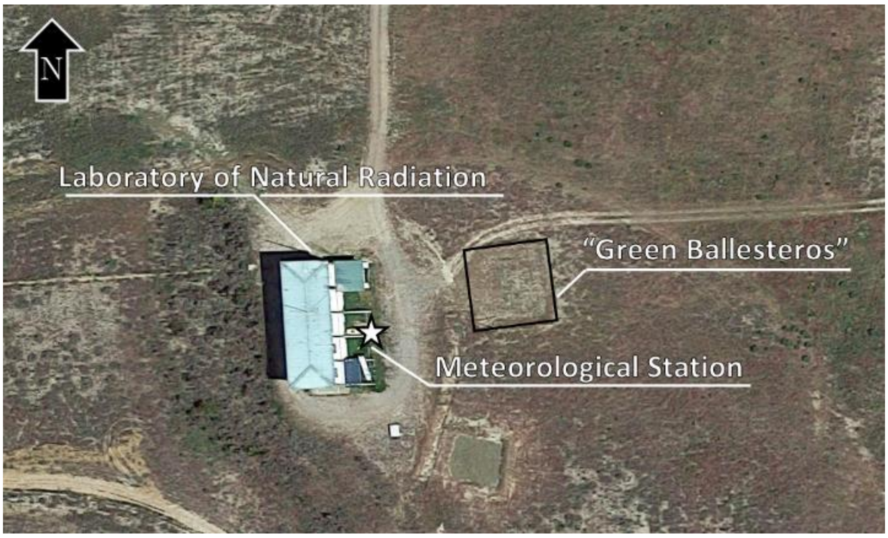

Fig. 2. Aerial view of LNR and surroundings where the intercomparison was developed.

This place has been used to carry out interlaboratory exercises both measuring radon concentration and gamma dose rate under natural environmental fluctuations. The high radioactive content in the soil along with the environmental conditions make this location a suitable place to conduct these kind of activities.

The LNR (see Fig. 3) has two floors; the first floor is used as a conference hall and multi-purpose room. The ground floor has two spaces designed as radon chambers (Room1 and Room2) with approximately 45 m^3 volume each. Room1 has no direct connection to the exterior while Room2 has an artificial ventilation system installed. The radon source is the uranium mine underground soil which has a high radium content.

In the east part of the LNR a meteorological station is set up to monitor environmental conditions. The datalogger, which has temperature and humidity sensors, is connected wirelessly to the station from the Room1.

Fig. 3. Laboratory of Natural Radiation seen from both sides.

Outside of LNR there is a place called “Green Ballesteros”, where a 5×5 m^2 and 1.5 m deep hole was dug out and filled with homogeneous soil with low radioactive content (^{226}Ra concentration about 43 ± 10 Bq/kg and gamma dose rate around 110 ± 5 nGy/h at 1 m height). The main radon source of this area is the original underlying soil.

Fig. 4. “Green Ballesteros” view and in the background the LNR.

The radon in air comparison exercise was developed inside Room1 of LNR, while the radon exhalation from soil and radon in soil activities were performed in the location called “Green Ballesteros”. Additionally, a place close to the uranium mine entrance was used to compare the radon in soil measurements, hereinafter named “Offices site”.

2 Methods

2.1 Activities

Radon in air exposure has been evaluated using passive detectors and active monitors inside Room1 of the Laboratory of Natural Radiation (LNR). Two exposures have been planned taking into account the natural radon evolution during those days. The devices were placed in Room1 on the 5 December 2018 and were taken off on 6 and 8 December 2018 for the first and second exposure, respectively, according to schedule shown in Table 2.

Table 2. Start and end dates for each exposure ((UTC+01:00) Brussels, Copenhagen, Madrid, Paris).

	Start date	End date
1st exposure E1:	05/11/2018 12:00	06/11/2018 1:00
2nd exposure E2:	05/11/2018 12:00	08/11/2018 10:00

Each participation with passive detectors requires a number of 30 units: 10 detectors for the first exposure, 10 for the second exposure and 10 transit detectors.

A total of 23 groups of passive detectors and 22 active monitors were exposed in Room1 of the Laboratory of Natural Radiation (see Fig. 5). For the second exposure E2 there are four results that are not reported due to various problems indicated by the participants.

Fig. 5. Radon devices inside Room1 (LRN).

There are several types of devices used by the participants. Passive detectors are made by different materials and/or use various technologies such as CR-39, LR 155, electret ion chambers, etc. Other procedures were implemented, e.g. using DVDs half made of polycarbonate (used as a solid state track detector) and polycarbonate foils used as a radon absorber. The features of diffusion chambers, holders, material quality and manufactures were diverse too. The overall characteristics given by participants are shown in Table 3.

Table 3. Passive detector features provided by the participants.

Detector	Diffusion chamber
CR-39 RSKS 100 mm^2 (Radosys)	Diameter 26 mm, height 55 mm 29 cm^3 volume
CR-39 24.7×36.7×1.40 (mm) (Mi-Net)	ENEA patent
CR-39 Radout 25×25×1.5 (mm) (Mi.am)	Diameter 50 mm, height 20 mm
CR-39 TASTRAK 13×37×1 (mm) (Tasl)	Diameter 58 mm, height 20 mm NRPB/SSI
CR-39 Duotrack (Radonova)	Diameter 58 mm, height 40 mm
CR-39 Radtrak2 (Radonova)	Diameter 58 mm, height 20 mm NRPB/SSI
CR-39 Rapidos (Radonova)	Diameter 58 mm, height 40 mm
ST Electret Teflon (E-PERM)	L-OO Chamber 58 mL
ST Electret Teflon (E-PERM)	S Chamber 210 mL
LR-115 type2 400 mm^2 (DOSIRAD)	Diameter 60.4 mm, height 27.6 mm Own design
LR-115 (KODAK) RAMARN device 0.012 mm film of cellulose nitrate, and coated on 0.1 mm thick polyester base	Polypropylene chamber 700 cm^3 volume
Makrofol 75.7 mm^2 STUK design "Radonpurkki"	Diameter 20 mm, height 71 mm 79 cm^3 volume
DVD half made of polycarbonate and two thin Makrofol N foils	Thin CD case

In this intercomparison different active monitors were used with various operation modes and features as shown in Table 4. This information has been obtained from the manufacturer's technical specifications.

Table 4. Active monitor features used in the intercomparison.

Monitor	**Detection technology**	**Sensitivity** (cpm at 1 kBq m^{-3})
AlphaGUARD	Ionisation chamber	50
ATMOS12 DPX	Ionisation chamber	20
SARAD EQF 3120	Silicon detector	7
Radon Scout	Silicon detector	1.8
Radon Scout Home	PIN photo diode	0.1

The organizers introduced/removed the passive detectors and active monitors from Room1. After each exposure, passive detectors were stored in a low radon concentration area. After two days, they were sealed in radon proof aluminium bags in order to allow a proper degassing. Active monitors were stopped and turned off at the end of the second exposure.

Transit detectors were stored in their original bags until the end of the second exposure. Afterwards, they were sealed in radon proof aluminium bags in order to simulate the exposed detectors conditions.

Participants have provided the exposure value and its uncertainty for each passive detector and the declared value for the first and second exposure period expressed in kBq m^{-3} h. In the case of active monitors, the overall exposure for each period was given; the individual radon concentration every hour was also included. The template for reporting results is shown in Appendix III.

For radon exhalation rate and radon in soil measurements two different points were available at the site of the intercomparison: the area “Green Ballesteros” and the “Offices site”. In the “Offices site” only radon in soil activity measurements were carried out. For these exercises, the measurements were conducted in situ and each participant used their own measuring system and sampling materials.

Participants have provided the exhalation rate value and its uncertainty for “Green Ballesteros” expressed in Bq m^{-2} h^{-1}. Radon in soil measurements were given for “Green Ballesteros” and “Offices site” expressed in kBq m^{-3}. Results template is shown in Appendix III.

Results provided by participants have been coded in order to maintain their anonymity. Such codification follows the rule:

LxxTn

where

xx is the number assigned to each participant from 01 to 20,

T is the type of measurement, **A**: radon in air with active monitor, **P**: radon in air with passive detectors, **E**: radon exhalation rate from soil, **S**: radon in soil,

n is the correlative number for more than one kind of measurements group.

2.2 Data Analysis

The determination of the assigned value and its standard uncertainty for each radon in air exposure have been obtained by using consensus value from participant results applying an iterative algorithm according to ISO 13528:2015. This algorithm considers the results of all participants and relocates the extreme values within the interval of acceptable deviation.

An outliers study has been applied in order to know the extreme values. The outlier values were found from the boxplot representation and the interquartile analysis. In this case an outlier is defined as a data point that is located 1.5 times the interquartile range (IQR) above the upper quartile and below the lower quartile. The interquartile range is defined as the difference between the third quartile (75^{th} percentile) and the first quartile (25^{th} percentile): IQR = (Q_3-Q_1).

The robust average and robust standard deviation denoted by E_{ref} and s^* have been calculated using "Algorithm A" taken from ISO 13528:2015:

There are p items of results denoted as:

$$E_i = E_1 , E_2 , E_3, \dots , E_p$$

Calculate initial values for E_{ref} and s^* as:

$$E_{ref} = \text{median of } E_i$$

$$s^* = 1.485 \text{ median of } |E_i - E_{ref}|$$

Update the values of E_{ref} and s^* as follows. Calculate:

$$\delta = 1.5\ s^*$$

$$E_i^* = \begin{cases} E_{ref} - \delta & \text{when} \quad E_i < E_{ref} - \delta \\ E_{ref} + \delta & \text{when} \quad E_i > E_{ref} + \delta \\ E_i & \text{otherwise} \end{cases}$$

Calculate the new values of E_{ref} and s^* from:

$$E_{ref} = \text{ mean of } E_i^*$$
$$s^* = 1.134 \cdot \text{SD } (E_i^*)$$

The robust estimates E_{ref} and s^* are derived by an iterative calculation, i.e. by updating the values of E_{ref} and s^* several times until the process converges.

Once the robust average and robust standard deviation have been calculated for each exposure period, the standard uncertainty of the assigned value may be estimated as:

$$u(E_{ref}) = 1.25\frac{s^*}{\sqrt{p}} \quad (1)$$

The indexes used to analyse the participants' results are the relative percentage difference $D(\%)$, the Zeta score (ζ) and the z-score (z).

The relative percentage difference $D(\%)$ has been introduced to quantify the difference between the participant's result and the reference value obtained as consensus. Therefore:

$$D_i(\%) = 100 \cdot \frac{E_i - E_{ref}}{E_{ref}} \quad (2)$$

where E_i is the exposure result i given by the participant.

The Zeta score (ζ) is a statistical index used to compare intercomparison results where the uncertainty in the measurement result is included. It is given by the following equation:

$$\zeta_i = \frac{E_i - E_{ref}}{\sqrt{u^2(E_i) + u^2(E_{ref})}} \quad (3)$$

being $u(E_i)$ the participant's own estimate the standard uncertainty of its result.

The z-score (z) index is calculated as follows:

$$z_i = \frac{E_i - E_{ref}}{\sigma_p} \quad (4)$$

where σ_p is the standard deviation for the intercomparison assessment estimated as 20% of reference value for the first exposure and 10% of reference value for the second one. This parameter should meet the following criterion: $u(E_{ref}) < 0.3\ \sigma_p$.

These indexes are interpreted as follow:

$|\ \zeta\ |$; $|\ z\ | \leq 2.0$ result is considered satisfactory

$2.0 < |\ \zeta\ |$; $|\ z\ | < 3.0$ result is considered to give a problem

$|\ \zeta\ |$; $|\ z\ | \geq 3.0$ is considered not satisfactory

The Zeta score (ζ) is used together with z-score (z) as an aid for improving the performance of participants. If a participant obtains a z-score higher than the critical value of 3.0, they may find it valuable to reassess their procedure with the subsequent uncertainty evaluation for that procedure. If the participant's ζ score also exceeds the critical value of 3.0, it implies that the participant's uncertainty evaluation does not include all significant sources of uncertainty. However, if a participant obtains a z-score ≥ 3.0 but a ζ score ≤ 2.0, this demonstrates that the participant may have assessed the uncertainty of their results accurately but that their results do not meet the performance expected for the proficiency testing scheme. The interpretation guidelines are shown in Table 5.

Table 5. Summary of guidelines to understand ζ and z scores.

ζ score	z-score	Action to take
Satisfactory	Satisfactory	Participant's result is good. No action is required.
Not satisfactory	Satisfactory	Participant's claimed uncertainty is too low, but the result fulfils the intercomparison requirements.
Satisfactory	Not Satisfactory	Participant's uncertainty assessment is accurate but the results do not fulfil the intercomparison requirements.
Not Satisfactory	Not Satisfactory	Participant's result is biased in excess. A complete revaluation should be performed.

In case of geogenic radon measurements, exhalation rate and radon in soil measurements, the methodology to assess the results is different due to the low number of results reported and the high dispersion of them. Descriptive statistics are applied and the claimed value is considered acceptable if it is within the interval defined by the mean value ± its standard deviation. In this case the analysis is not the same as radon in air activity due to the reasons mentioned.

3 Results

3.1 Environmental conditions

Below the environmental conditions in Room1 and outside during the intercomparison exercise are presented. Fig. 6 graphically shows the variation of internal and external environmental parameters, temperature and relative humidity. Table 6 gives the mean and extreme values. It is observed that the variation of temperature in Room1 is quite stable, with an absolute difference of 1 ºC, while this difference outside is about 10 ºC. Atmospheric pressure average inside Room1 was 935 ± 5 hPa with an absolute variation of 14 hPa.

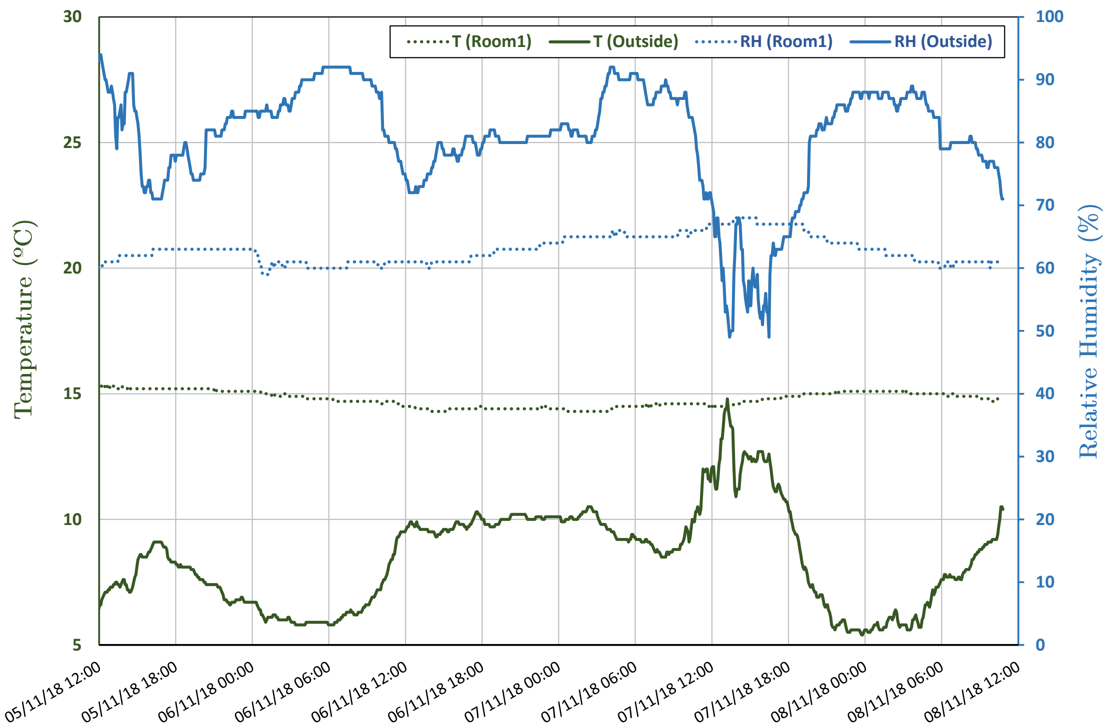

Fig. 6. Temperature (green) and relative humidity (blue) variation in Room1 (dotted line) and outside (solid line) during the intercomparison exercise. Measurements are taken every five minutes.

Table 6. Temperature and humidity conditions in Room1 and outside LNR.

	Room1	Outside
T_{mean} (ºC)	14.8 ± 0.3	8.1 ± 2.0
T_{min} (ºC)	14.3 ± 0.1	5.3 ± 0.1
T_{max} (ºC)	15.3 ± 0.1	14.8 ± 0.1
RH_{mean} (%)	63 ± 2	82 ± 9
RH_{min} (%)	57 ± 1	49 ± 1
RH_{max} (%)	68 ± 1	95 ± 1

The weather during the intercomparison exercise was rainy. On 4 November it started to rain and it continued until the 8 November. The amount of precipitation is shown in Fig. 7. Rain has a special interest in geogenic measurements because it could significantly modify the mobility of radon in the soil.

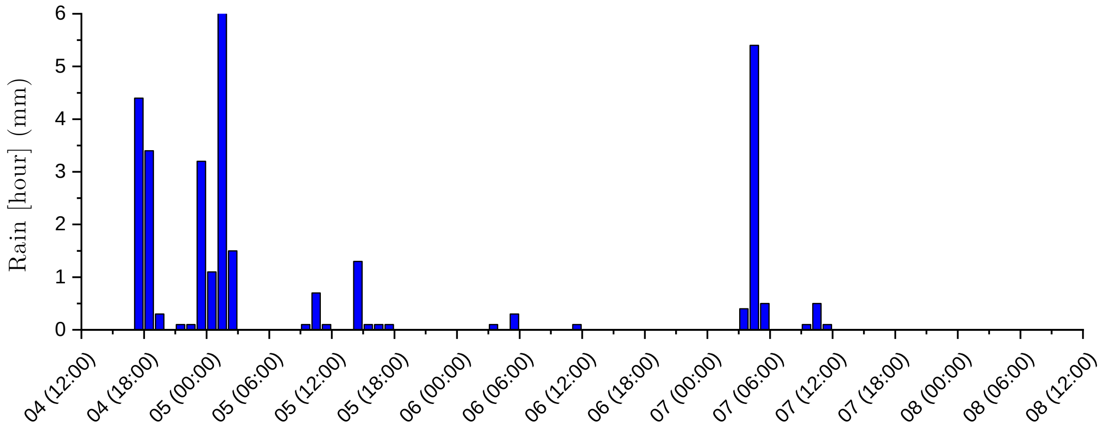

Fig. 7. Rain accumulated every hour expressed in mm from the 4 November 2018 to 8 November 2018.

3.2 Radon in air

In this subsection the radon in air exposure results are analysed. Participants submitted one exposure result together with its uncertainty per group of passive detectors and/or active monitor for the first exposure E1 and for the second exposure E2.

Appendix I contains the numerical results submitted by the participants for each exposure and the indexes used to assess their performance.

The variation of radon concentration in Room1 shows a big range of values, with levels from approximately 0.5 to 30 kBq/m^3. As an example, the radon concentration measurements of Laboratory of environmental radioactivity, University of Cantabria (LaRUC), taken by the device AlphaGUARD (S/N AG000032), are shown in Fig. 8.

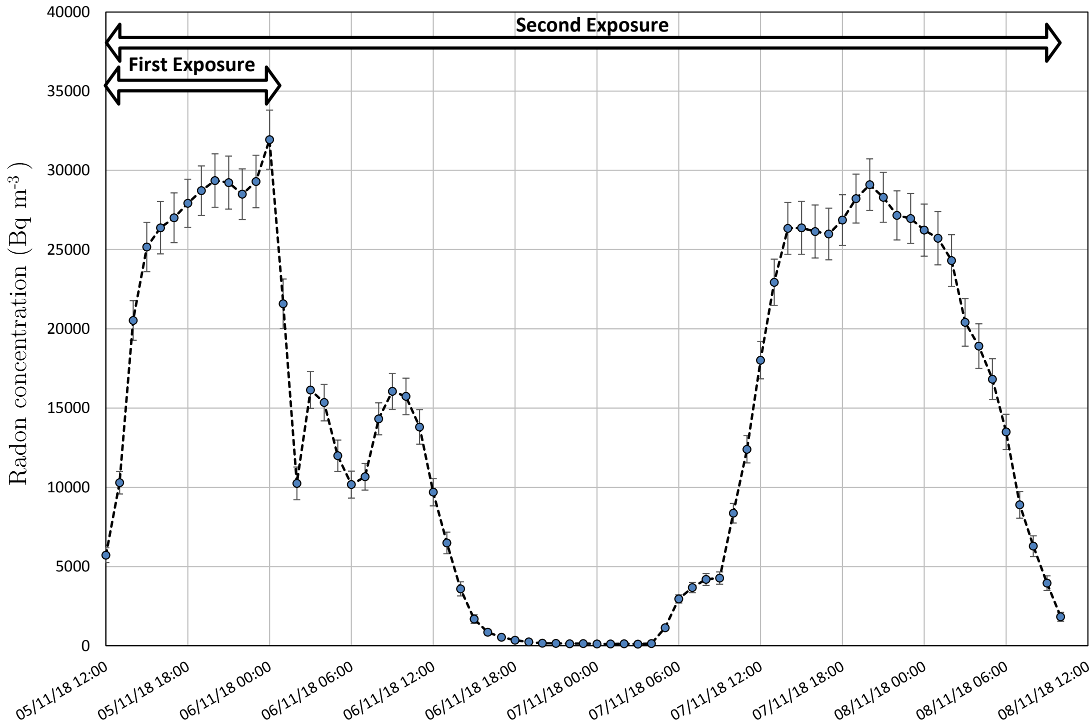

Fig. 8. Radon concentration in Room1 during the intercomparison exercise according to LaRUC. Data is displayed every hourly.

The assigned values used as reference for each exposure period are derived from a weighted average of participants' results applying the iterative algorithm described above according to ISO 13528:2015. Table 7 shows the robust average E_{ref}, the robust standard deviation s^*, the standard uncertainty $u(E_{ref})$, the number of results p and the standard deviation for the intercomparison assessment σ_p estimated as 20% of reference value for the first exposure and 10% of reference value for the second one. This parameter meets the criterion: $u(E_{ref}) < 0.3\ \sigma_p$.

Table 7. Reference parameters of the first exposure E1 and second exposure E2 expressed in kBq m^{-3} h obtained from participant results according to ISO 13528:2015. p is the dimensionless number of results.

	E_{ref}	$u(E_{ref})$	σ_p	s^*	p
1st exposure E1:	356	8	71	43	45
2nd exposure E2:	1014	13	101	68	41

As mentioned before, outliers have been identified using a bloxplot diagram (Fig. 9). The corresponding codes are displayed in Table 8. There are no statistical differences between the reference exposure value calculated taking into account the total amount of results and the one calculated without considering outliers. Therefore all the results have been considered to calculate the reference values.

Table 8. Results considered outliers from the interquartile analysis.

	Laboratory code				
1st exposure E1:	L01P2	L01P3	L02P1	L02P2	L16P1
2nd exposure E2:	L03P1	L16P1	L19P1	L20A3	

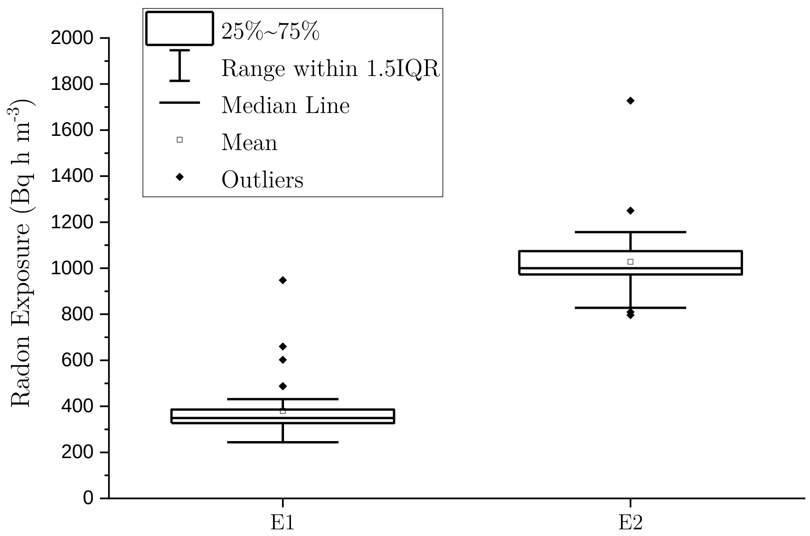

Fig. 9. Boxplot diagram of the participant's results for exposures E1 and E2.

Participant's results for the first radon in air exposure are given in Fig. 10. Each value is presented with its uncertainty (k = 1). The solid line represents the reference value obtained through consensus (356 kBq m^{-3} h) and the dashed lines denote the standard deviation for the inter-laboratory assessment estimated as 20% of the reference value. Fig. 11 shows the results for the second exposure, with the reference value of 1014 kBq m^{-3} h indicated with a solid line. In this case the dashed lines represent the 10% of that reference value.

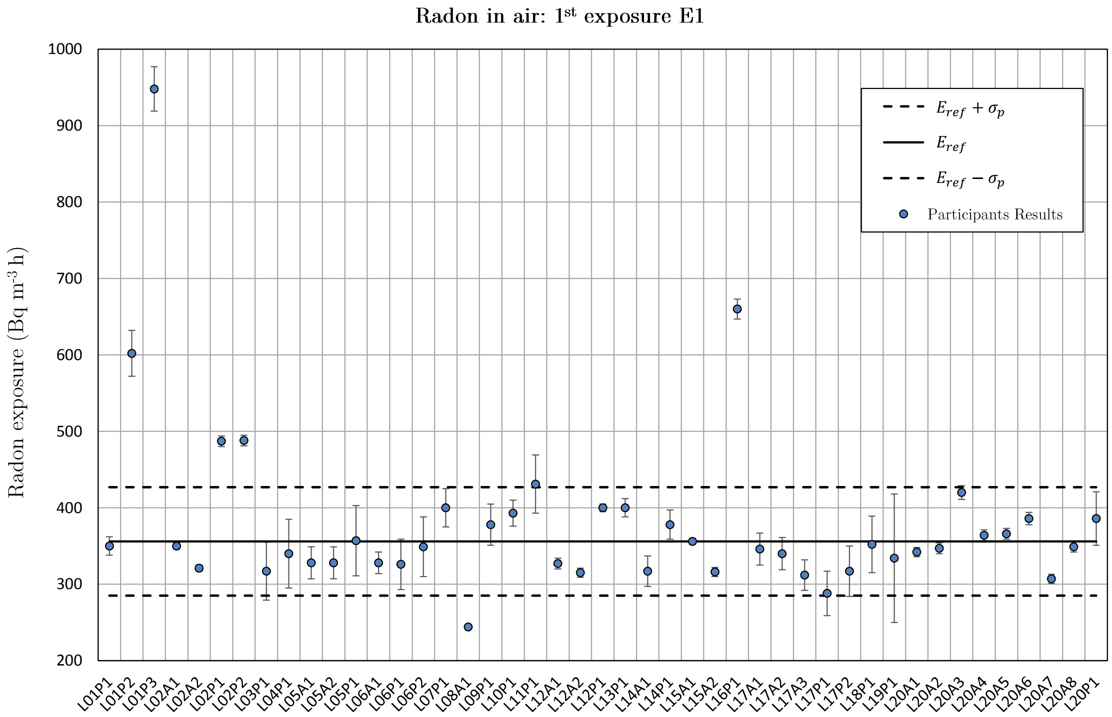

Fig. 10. Participant's results for the first exposure E1 with its associated uncertainty ($k = 1$). Exposure reference value is shown with a solid line and the standard deviation $\sigma_p = 0.2E_{ref}$ with dashed lines.

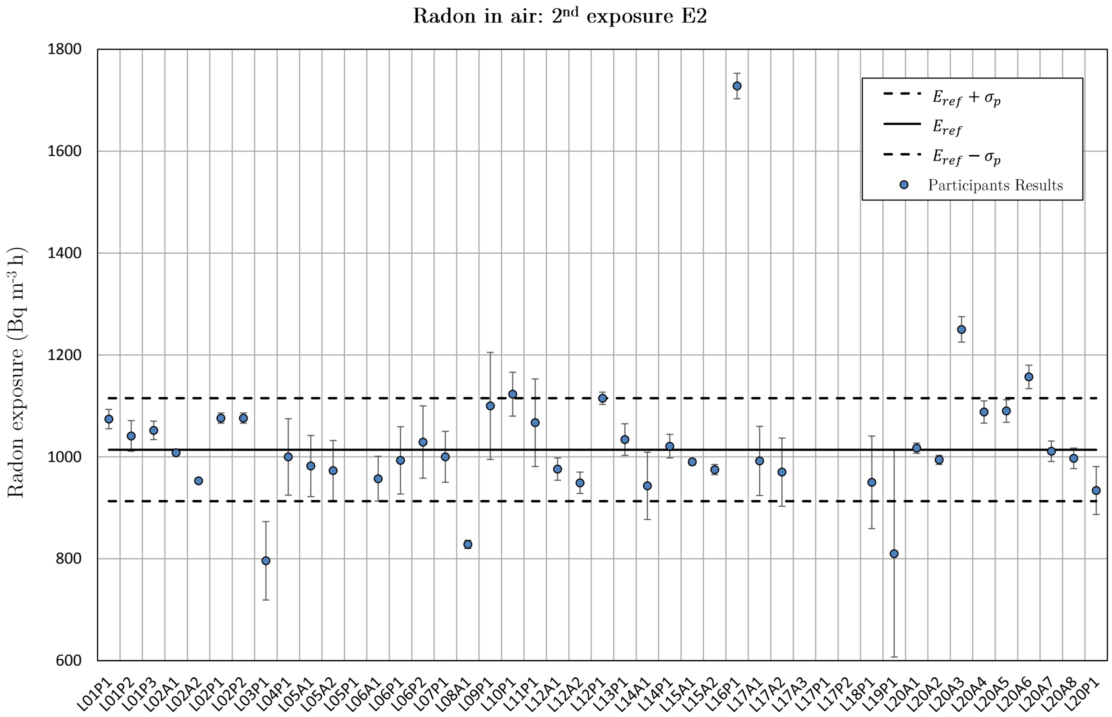

Fig. 11. Participant's results for the second exposure E2 with its associated uncertainty ($k = 1$). Exposure reference value is shown with a solid line and the standard deviation $\sigma_p = 0.1E_{ref}$ with dashed lines.

About 80% of the results presented in Fig. 10 and Fig. 11 are within the interval defined by the exposure reference value E_{ref} and the standard deviation σ_p established as 20% and 10% for the first and the second exposure, respectively. The relative difference $D(\%)$ between each single value and the reference is shown in Fig. 12. The anomalous values shown in Table 8 are clearly out of those intervals.

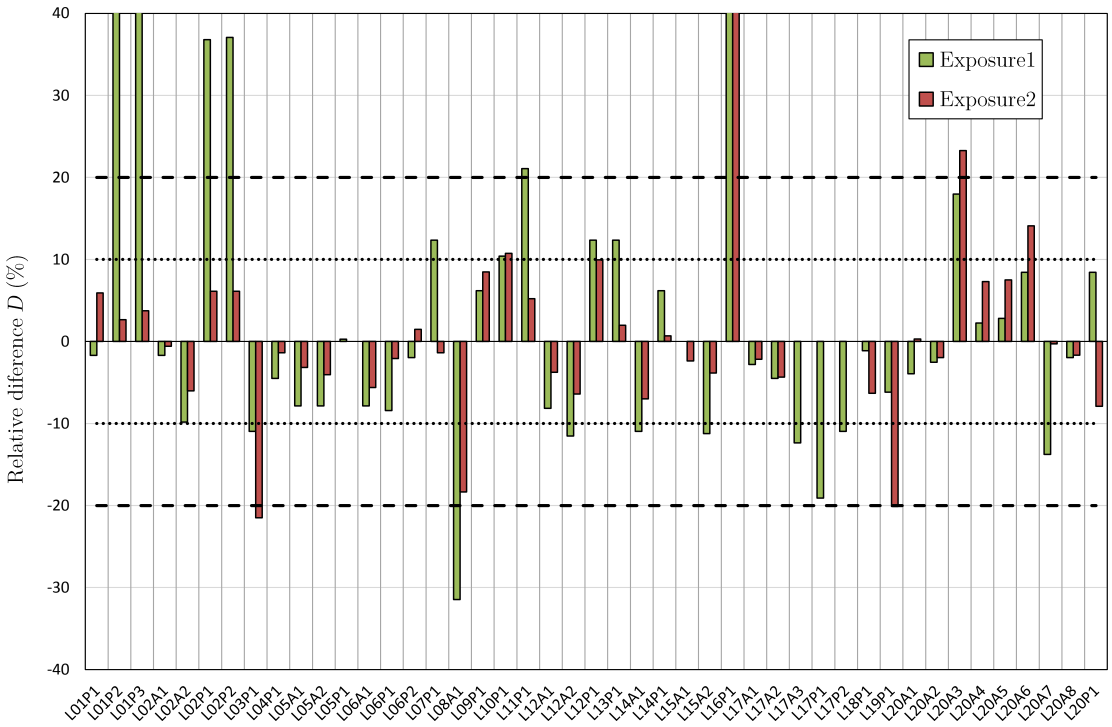

Fig. 12. Relative difference of participant's results to the mean value for the first and second exposure. Intervals established for the first exposure (± 10%) and second exposure (± 20%) are indicated.

Below it is shown the graphical representation of indexes used to assess the participant's results. In some cases the value is out of scale in order to improve the graph view. Numerical results are collected in Appendix I. In addition, Table 9 shows the percentage of results that are within the limits for each index. For the relative difference, the percentage of results within 10% and 20% of reference exposure in each case is presented.

The overall performance of results given by z-score is satisfactory, about 90% of results have a value lower than 2.0 for both exposures. Only the results of three cases have a z-score value above 3.0 for the first exposure and one result for the second exposure period. Regarding the Zeta score, about 60% of results are satisfactory ($|\zeta| \leq 2.0$), however, 29% of results for the first exposure and 20% for the second exposure period are not satisfactory, with a Zeta score $|\zeta| \geq 3.0$.

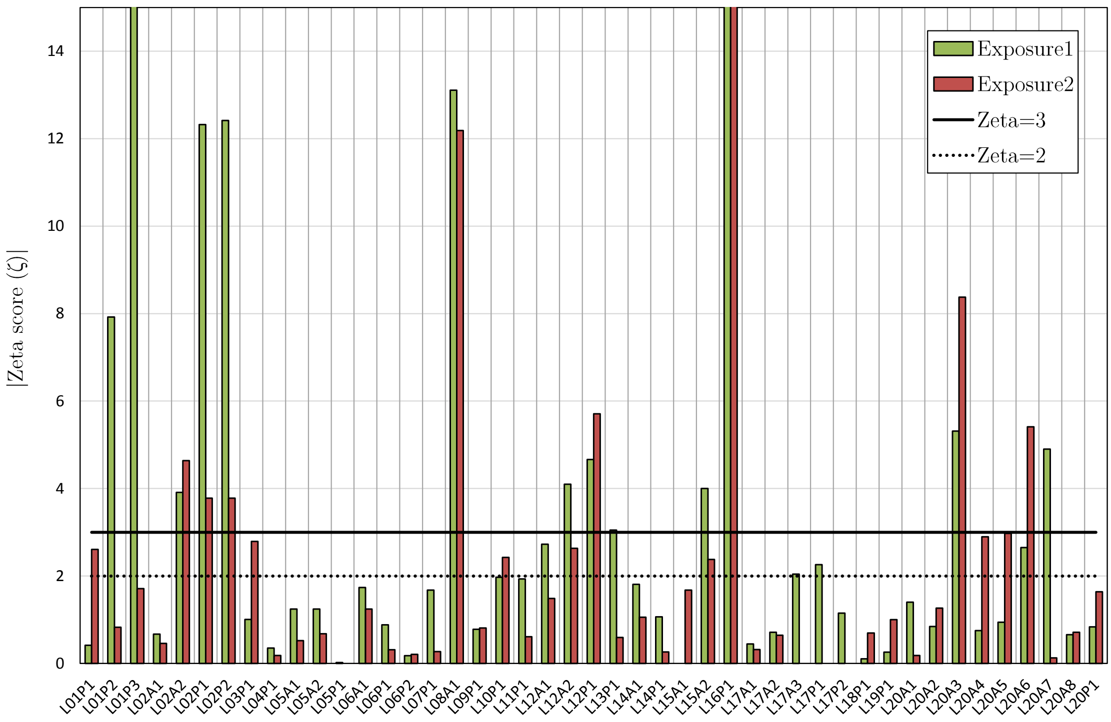

Fig. 13. Absolute values of Zeta score for the first and second exposure.

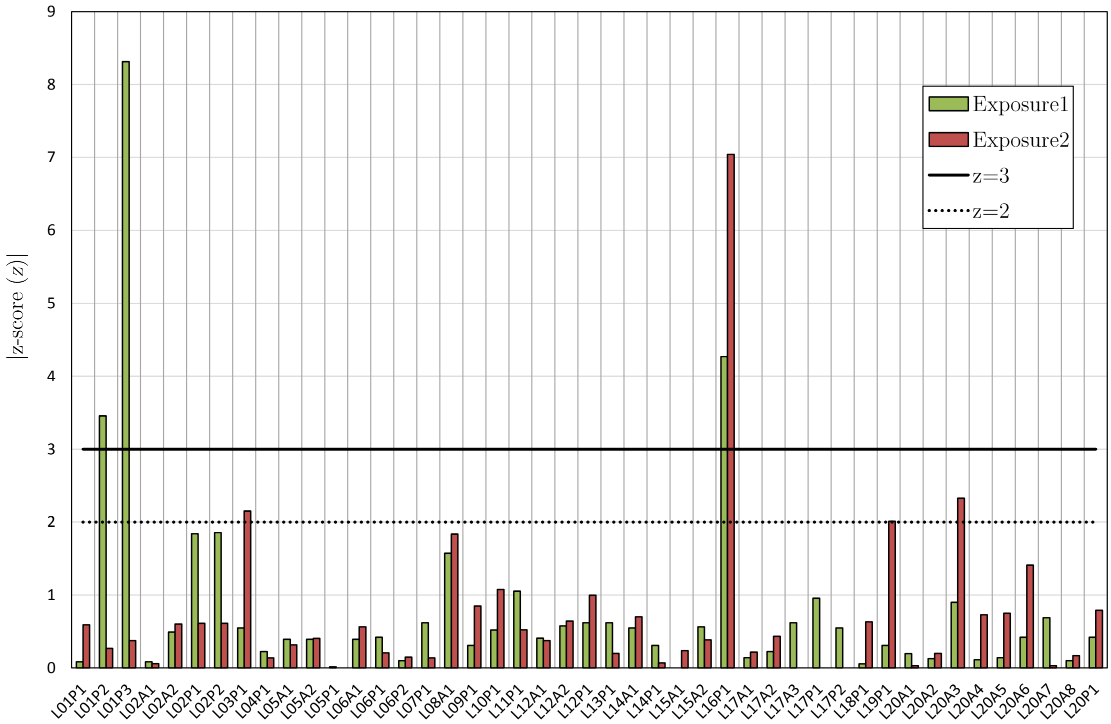

Fig. 14. Absolute values of z-score for the first and second exposure.

Table 9. Percentage of results that are within the limits for Zeta score (ζ) and z-score (z).

	Results of E1 (%)	Results of E2 (%)
$\|D(\%)\| \leq 10\%$	56	83
$\|D(\%)\| \leq 20\%$	84	90
$\| \zeta \| \leq 2.0$	62	63
$2.0 < \| \zeta \| < 3.0$	9	17
$\| \zeta \| \geq 3.0$	29	20
$\| z \| \leq 2.0$	93	90
$2.0 < \| z \| < 3.0$	0	7
$\| z \| \geq 3.0$	7	2

3.3 Exhalation from soil

In this subsection the results for exhalation rate measurements are presented. Participants submitted one single value with its uncertainty for the exhalation rate expressed in Bq m^{-2} h^{-1} in "Green Ballesteros". There were 3 participants who used their own methodology in situ, all different from each other.

Due to the different methodologies involved, the measurements were not carried out at the same time. Participant L03 used absorption in polycarbonate, which is a cumulative measurement, with a duration of two days. Participant L17 used the accumulation method with a radon monitor with a sampling time of approximately one hour. This method provides a discrete value of exhalation. In case of participant L20, the measurements were performed using the absorption in activated charcoal canisters during 24 hours one week later due to logistics problems. Results, dates and methodologies are shown in Table 10.

Table 10. Radon exhalation J and its uncertainty $u(J)$ performed in the indicated date with the methodology used by each participant to conduct the test in the "Green Ballesteros".

Code	J (Bq m^{-2} h^{-1})	$u(J)$ (Bq m^{-2} h^{-1})	Date (2018)	Methodology
L03E1	361	33	5 Nov 15:30 to 7 Nov 13:30	Gradient method with polycarbonate foils
L17E1	14719	1939	6 Nov 15:30 (approx. 1 hour.)	Accumulation method
L20E1	35100	8200	15 Nov 10:00 to 16 Nov 10:00	Absorption in activated charcoal collector

The large differences in the obtained results require some explanation. Although the measurements are very few to draw definitive conclusions, still some consideration can be made. As can be seen in Table 10, measurements L03E1 and L17E1 overlap in time, but they differ in two orders of magnitude. However, the first method is cumulative, while the second is discrete. During the period of these measurements, the weather was mostly rainy and windy and the soil was soaked, which would impede the exhalation. Due to the windy weather, there were some cloudless and sunny time windows and the discrete measurement L17E1 was carried out in such a window (see Fig. 15). The sun would dry the soil and lead to an increase in the radon exhalation, which could be a possible explanation for the observed difference between the results of the two measurements.

On the other hand, the two cumulative measurements L03E1 and L20E1 could not be compared directly, as they were carried out at different times. The weather was dry and sunny in the week of the L20E1 measurement, in contrast to the weather during the L03E1 measurement. Although no definitive conclusions can be drawn, these results indicate the significant effect of the weather on the radon exhalation rate, which deserves a more thorough study.

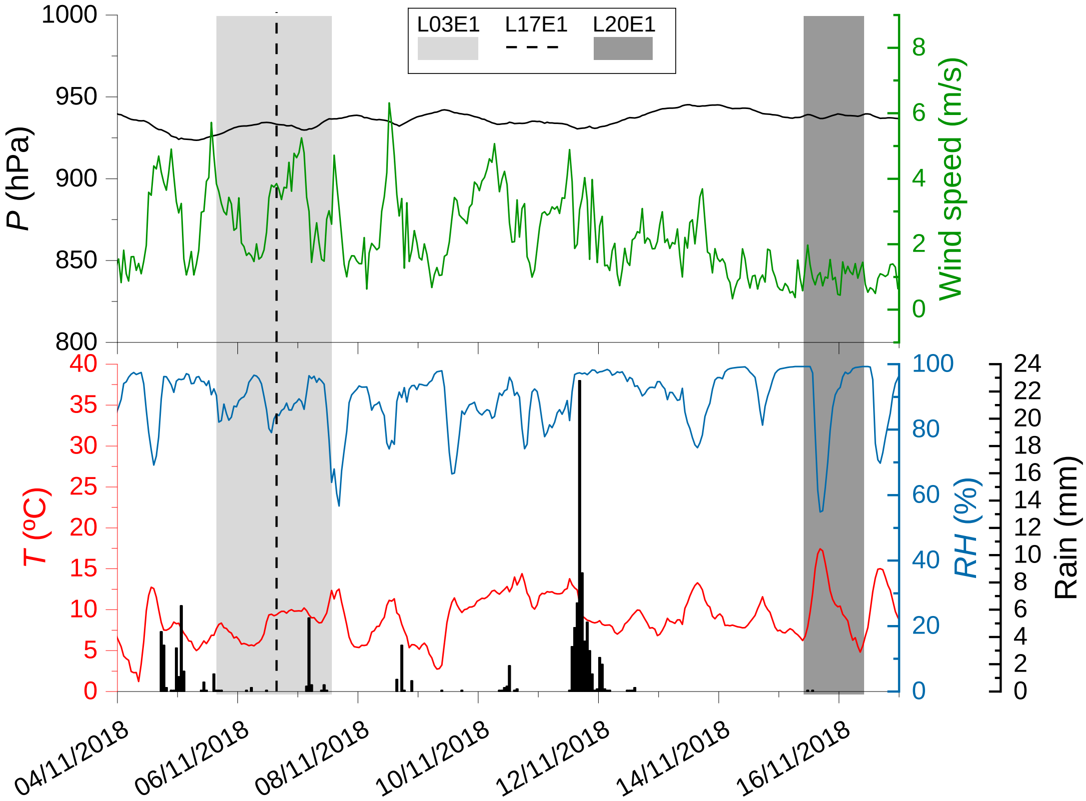

Fig. 15. Environmental conditions during the exhalation rate activity displayed every hour. Shaded areas correspond to the cumulative measurements (light grey L03E1; dark grey L20E1), and dashed line to L17E1 discrete measurement.

It is observed that there is a great difference between the results. The standard deviation (SD) is as big as the mean value (see Table 11).

Table 11. Descriptive statistics of radon exhalation results expressed in Bq m^{-2} h^{-1}.

	"Green Ballesteros"
Mean	16727
Median	14719
SD	17456
p	3

3.4 Radon in soil

Radon in soil measurements were carried out in the "Green Ballesteros" area and at "Offices site" on 6 November 2018. Table 12 gives the results and the methodology applied. Descriptive statistics for each case are shown in Table 13. It is observed that the dispersion of results is acceptable in case of "Green Ballesteros", with a standard deviation of about 14% of the mean value. However, the "Offices site" results show great differences. Graphical representation of radon in soil measurements performed in "Green Ballesteros" is shown in Fig. 16.

Table 12. Radon concentration in soil C_{soil} and its uncertainty $u(C_{soil})$ with the methodology used by each participant to conduct the measurements in "Green Ballesteros" and at the "Offices site".

	"Green Ballesteros"		"Offices site"		
Code	C_{soil} (kBq m^{-3})	$u(C_{soil})$ (kBq m^{-3})	C_{soil} (kBq m^{-3})	$u(C_{soil})$ (kBq m^{-3})	Methodology
L03S1	602	57			Absorption in polycarbonate foils
L10S1	546	143	6.3	3.3	Continuous monitoring
L13S1	789	74			Etched track detectors
L17S1	894	37	994	40	Continuous monitoring
L20S1	840	140	20	12	Grab sampling in ionization chamber plus measure with electrometer

Table 13. Descriptive statistics of radon in soil results expressed in kBq m^{-3}.

	"Green Ballesteros"	"Offices site"
Mean	734	340
Median	840	20
SD	152	566
p	5	3

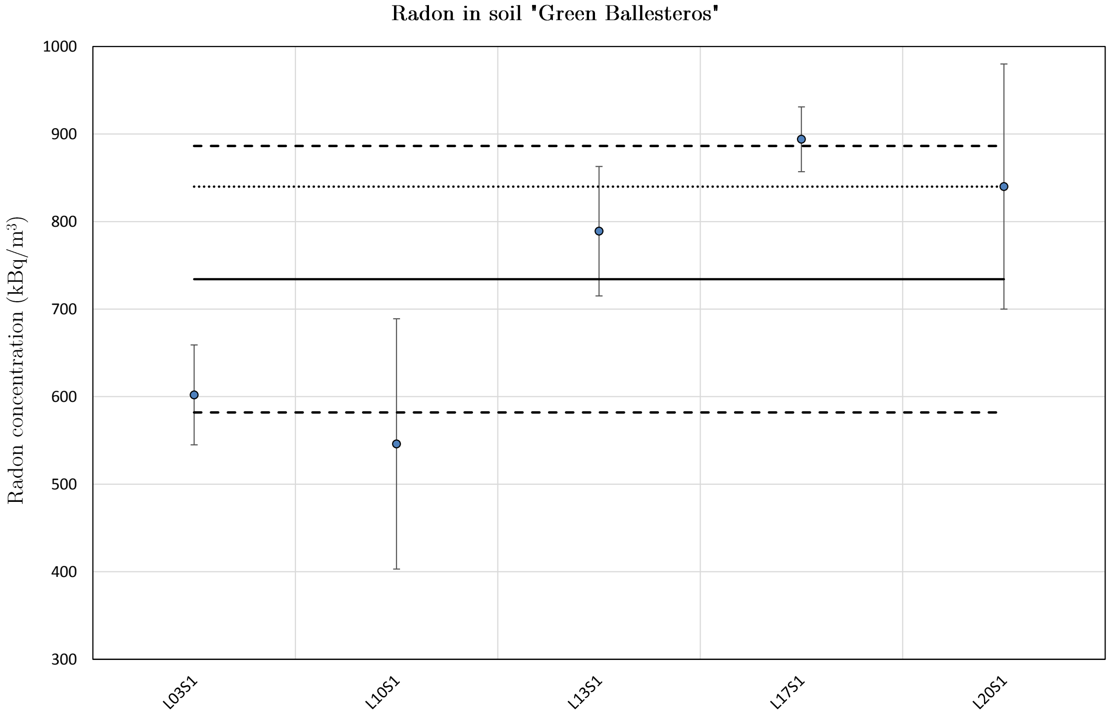

Fig. 16. Radon in soil results with its associated uncertainty ($k = 1$) in "Green Ballesteros". The solid line denotes the mean value, the dotted line the median and the dashed lines denote the standard deviation from the mean.

4 Conclusions

An inter-laboratory exercise of indoor radon and geogenic radon measurements under field conditions has been carried out in the Laboratory of Natural Radiation (LNR) between 5 and 8 December 2018. The facility is located in the former uranium mine managed by the Spanish National Uranium Company ENUSA (Saelices el Chico, Salamanca, Spain). Radon in air measurements were assessed from two exposure periods, while the geogenic radon parameters were evaluated from radon exhalation from soil and radon concentration in soil gas measurements.

Radon in air reference values for each exposure were obtained through consensus from participant's results applying an iterative algorithm according to ISO 13528:2015. The indexes used to analyse the participants results are relative percentage difference $D(\%)$, Zeta score (ζ) and z-score (z).

Over 80% of the results for radon in air exposure are within the interval defined by the reference value and the standard deviation, established as 20% and 10% for the first and the second exposure respectively. The exercise was successful, taking into account the large number of different devices used, especially in passive detectors where holder materials, diffusion chamber volume, detectors area or detection principle were diverse.

Five results of the first exposure are considered outliers. All of them are passive detectors and are overestimating the exposure from approximately 40% to 160%. Such deviations could be related with the degassing time of detector holder materials. Radon could get adsorbed in it for a long time so even after two days, when the detectors were put in radon proof bags and sealed. A further difficulty in this intercomparison exercise is that the exposures are reached in a short time period with high radon concentrations in air. At the end of the first exposure period there was a radon concentration in air around 30 kBq m^{-3} which can cause the holder degassing problem previously mentioned. In case of the second exposure, the radon concentration was under 2 kBq m^{-3} at the end of that period, therefore reducing the exposure increase due to the possible effect of adsorption and degassing.

Most z-score results are satisfactory, about 90% of the results have a value lower than 2 for both exposures. Only the results for three cases for the first exposure and one result for the second exposure period are not satisfactory, with z-score values higher than 3.0.

Regarding the Zeta score, about 60 % of results are satisfactory ($|\zeta| \leq 2.0$), however, 29% of results for the first exposure and 20% for the second exposure period are not satisfactory, with a Zeta score $|\zeta| \geq 3.0$.

Every participant has assessed their own results and revaluated their method if it necessary according to the indexes obtained.

Radon exhalation measurements were carried out in “Green Ballesteros” only by three participants. The participants did not perform the measurements during the same time due to the different methodologies involved. The weather in each case was different which could explain the widespread of results.

L03E1 and L17E1 measurements overlapped in time, but they differ in two orders of magnitude. Participant L03 used absorption in polycarbonate, which is a cumulative measurement, with a duration of two days. Participant L17 used an accumulation method with a radon monitor with a sampling time of approximately one hour. This method provides a discrete value of exhalation. On the other hand, participant L20 performed the measurements using the absorption in activated charcoal canisters during 24 hours, however the measurements were performed one week later.

During the first period of measurements, L03E1 and L17E1, the weather was mostly rainy and the soil was soaked, which would impede the exhalation. The discrete measurement L17E1 was carried out in a sunny period, the temperature raised and the relative humidity decreased. The sun would dry the soil and lead to an increase in the radon exhalation. This effect could be a possible explanation for the observed difference between the results of the two measurements. In the week of the L20E1 measurement, which is the highest value of exhalation rate, the weather was dry and sunny, in contrast to the weather during the other measurements.

Therefore, a consensus value for radon exhalation rate cannot be obtained due to the dispersion of the results. Results indicate a significant effect of the weather on the radon exhalation rate, which deserves a more thorough study.

Radon in soil measurements were carried out in “Green Ballesteros” by 5 participants and, additionally, in “Offices site” by 3 participants. Results in “Green Ballesteros” are approx. between 550 and 900 kBq m^{-3}. All participants provided acceptable values taking into account the arithmetic mean value and its standard deviation. This fact could be explained due to the homogeneity of the area. The observed differences agree with the typical spatial variability of radon in soil measurements. On the other hand, the lack of homogeneity in the “Offices site” provides an extremely high dispersion of results. Homogeneity and a historical data collection are necessary to carry out an intercomparison of geogenic radon measurements with such a low number of participants.

Appendix I: Radon in air exposure results

This appendix contains the results submitted by the participants for each exposure and the indexes used to assess their performance.

Table 14. Participant's results and their statistical indexes, relative percentage difference *D*(%), Zeta score (ζ) and z-score for the **first exposure E1**.

Code	E1 (kBq m^{-3} h)	*u* (E1) (*k*=1) (kBq m^{-3} h)	*D*(%)	ζ score	z-score
L01P1	350	12	-1.7	-0.4	-0.1
L01P2	602	30	69.1	7.9	3.5
L01P3	948	29	166.3	19.7	8.3
L02A1	350	4	-1.7	-0.7	-0.1
L02A2	321	4	-9.8	-3.9	-0.5
L02P1	487	7	36.8	12.3	1.8
L02P2	488	7	37.1	12.4	1.9
L03P1	317	38	-11.0	-1.0	-0.5
L04P1	340	45	-4.5	-0.4	-0.2
L05A1	328	21	-7.9	-1.2	-0.4
L05A2	328	21	-7.9	-1.2	-0.4
L05P1	357	46	0.3	0.0	0.0
L06A1	328	14	-7.9	-1.7	-0.4
L06P1	326	33	-8.4	-0.9	-0.4
L06P2	349	39	-2.0	-0.2	-0.1
L07P1	400	25	12.4	1.7	0.6
L08A1	244	3	-31.5	-13.1	-1.6
L09P1	378	27	6.2	0.8	0.3
L10P1	393	17	10.4	2.0	0.5
L11P1	431	38	21.1	1.9	1.1
L12A1	327	7	-8.1	-2.7	-0.4
L12A2	315	6	-11.5	-4.1	-0.6
L12P1	400	5	12.4	4.7	0.6
L13P1	400	12	12.4	3.1	0.6
L14A1	317	20	-11.0	-1.8	-0.5
L14P1	378	19	6.2	1.1	0.3
L15A1	356	4	0.0	0.0	0.0
L15A2	316	6	-11.2	-4.0	-0.6
L16P1	660	13	85.4	19.9	4.3
L17A1	346	21	-2.8	-0.4	-0.1
L17A2	340	21	-4.5	-0.7	-0.2
L17A3	312	20	-12.4	-2.0	-0.6

Code	E1 (kBq m^{-3} h)	u (E1) (k=1) (kBq m^{-3} h)	D(%)	ζ score	z-score
L17P1	288	29	-19.1	-2.3	-1.0
L17P2	317	33	-11.0	-1.1	-0.5
L18P1	352	37	-1.1	-0.1	-0.1
L19P1	334	84	-6.2	-0.3	-0.3
L20A1	342	6	-3.9	-1.4	-0.2
L20A2	347	7	-2.5	-0.8	-0.1
L20A3	420	9	18.0	5.3	0.9
L20A4	364	7	2.2	0.8	0.1
L20A5	366	7	2.8	0.9	0.1
L20A6	386	8	8.4	2.7	0.4
L20A7	307	6	-13.8	-4.9	-0.7
L20A8	349	7	-2.0	-0.7	-0.1
L20P1	386	35	8.4	0.8	0.4

Table 15. Participant's results and their statistical indexes, relative percentage difference D(%), Zeta score (ζ) and z-score for the **second exposure E2**.

Code	E2 (kBq m^{-3} h)	u (E2) (k=1) (kBq m^{-3} h)	D(%)	ζ score	z-score
L01P1	1074	19	5.9	2.6	0.6
L01P2	1041	30	2.7	0.8	0.3
L01P3	1052	18	3.7	1.7	0.4
L02A1	1008	2	-0.6	-0.5	-0.1
L02A2	953	2	-6.0	-4.6	-0.6
L02P1	1076	10	6.1	3.8	0.6
L02P2	1076	10	6.1	3.8	0.6
L03P1	796	77	-21.5	-2.8	-2.1
L04P1	1000	75	-1.4	-0.2	-0.1
L05A1	982	60	-3.2	-0.5	-0.3
L05A2	973	59	-4.0	-0.7	-0.4
L06A1	957	44	-5.6	-1.2	-0.6
L06P1	993	66	-2.1	-0.3	-0.2
L06P2	1029	71	1.5	0.2	0.1
L07P1	1000	50	-1.4	-0.3	-0.1
L08A1	828	8	-18.3	-12.2	-1.8
L09P1	1100	105	8.5	0.8	0.8
L10P1	1123	43	10.7	2.4	1.1
L11P1	1067	86	5.2	0.6	0.5

Code	E2 (kBq m^{-3} h)	u (E2) (k=1) (kBq m^{-3} h)	D(%)	ζ score	z-score
L12A1	976	22	-3.7	-1.5	-0.4
L12A2	949	21	-6.4	-2.6	-0.6
L12P1	1115	12	10.0	5.7	1.0
L13P1	1034	31	2.0	0.6	0.2
L14A1	943	66	-7.0	-1.1	-0.7
L14P1	1021	23	0.7	0.3	0.1
L15A1	990	6	-2.4	-1.7	-0.2
L15A2	975	10	-3.8	-2.4	-0.4
L16P1	1728	25	70.4	25.3	7.0
L17A1	992	68	-2.2	-0.3	-0.2
L17A2	970	67	-4.3	-0.6	-0.4
L18P1	950	91	-6.3	-0.7	-0.6
L19P1	810	203	-20.1	-1.0	-2.0
L20A1	1017	10	0.3	0.2	0.0
L20A2	994	9	-2.0	-1.3	-0.2
L20A3	1250	25	23.3	8.4	2.3
L20A4	1088	22	7.3	2.9	0.7
L20A5	1090	22	7.5	3.0	0.7
L20A6	1157	23	14.1	5.4	1.4
L20A7	1011	20	-0.3	-0.1	0.0
L20A8	997	20	-1.7	-0.7	-0.2
L20P1	934	47	-7.9	-1.6	-0.8

Appendix II: Questionnaire sent to participants

LaRUC
Laboratorio de Radiactividad Ambiental

MetroRADON

MetroRADON: Intercomparison on indoor radon at LNR

November 5th - 8th, 2018

Saelices el Chico (Salamanca, Spain)

► Participant information

Organization name: Acronym:

Address:

Country:

Contact person:

E-mail:

Phone number:

► Activities

Indicate in what activities you are going to participate ("Yes" or "No")

Radon in air with PASSIVE detectors:

Radon in air with ACTIVE monitors:

Radon in soil:

Radon Exhalation from soil:

► Additional information:

How many people are going to attend the intercomparison?[1] :

Indicate the name of the attendees[2]:

Are you going to send the Passive detectors and Active monitors by mail? [3]:

[1] Attendance is mandatory for radon in soil and exhalation activities

[2] Send by E-mail the ID card scanning of every attendee to manage the access to the facilities

[3] If so, they have to be in LaRUC facilities before October, 31th.

More information and shipping address:

LaRUC, Facultad de Medicina
C/Cardenal Herrera Oria s/n 39011 Santander,
University of Cantabria
Spain
E-mail: daniel.rabago@unican.es / laruc@unican.es
(Phone: +34 942 20 22 07)

Appendix III: Results Template

Device	S/N	1st Exposure (kBq h m^{-3})		2nd Exposure (kBq h m^{-3})	
		Value	Uncertainty	Value	Uncertainty

1st Exposure: 05/11/2018 12:00 – 06/11/2018 1:00

2nd Exposure: 05/11/2018 12:00 – 08/11/2018 10:00

Date	C (Bq m^{-3})	u(C) (Bq m^{-3})
05/11/2018 12:00		
05/11/2018 13:00		
05/11/2018 14:00		
05/11/2018 15:00		
05/11/2018 16:00		
05/11/2018 17:00		
05/11/2018 18:00		
05/11/2018 19:00		
05/11/2018 20:00		
05/11/2018 21:00		
05/11/2018 22:00		
05/11/2018 23:00		
06/11/2018 0:00		
06/11/2018 1:00		
06/11/2018 2:00		
06/11/2018 3:00		
06/11/2018 4:00		
06/11/2018 5:00		
06/11/2018 6:00		
06/11/2018 7:00		
06/11/2018 8:00		
06/11/2018 9:00		
06/11/2018 10:00		
06/11/2018 11:00		
06/11/2018 12:00		
06/11/2018 13:00		
06/11/2018 14:00		
06/11/2018 15:00		
06/11/2018 16:00		
06/11/2018 17:00		
06/11/2018 18:00		
06/11/2018 19:00		
06/11/2018 20:00		
06/11/2018 21:00		
06/11/2018 22:00		
06/11/2018 23:00		
07/11/2018 0:00		
07/11/2018 1:00		
07/11/2018 2:00		
07/11/2018 3:00		
07/11/2018 4:00		
07/11/2018 5:00		
07/11/2018 6:00		
07/11/2018 7:00		
07/11/2018 8:00		
07/11/2018 9:00		
07/11/2018 10:00		
07/11/2018 11:00		
07/11/2018 12:00		
07/11/2018 13:00		
07/11/2018 14:00		
07/11/2018 15:00		
07/11/2018 16:00		
07/11/2018 17:00		
07/11/2018 18:00		
07/11/2018 19:00		
07/11/2018 20:00		
07/11/2018 21:00		
07/11/2018 22:00		
07/11/2018 23:00		
08/11/2018 0:00		
08/11/2018 1:00		
08/11/2018 2:00		
08/11/2018 3:00		
08/11/2018 4:00		
08/11/2018 5:00		
08/11/2018 6:00		
08/11/2018 7:00		
08/11/2018 8:00		
08/11/2018 9:00		
08/11/2018 10:00		

1st Exposure (kBq h m^{-3})			2nd Exposure (kBq h m^{-3})		
Detector code	Value	Uncertainty	Detector code	Value	Uncertainty

	Overall Results (kBq h m^{-3})	
	Value	Uncertainty
1st Exposure		
2nd Exposure		

Brief description of detector: (kind of detector, chamber volume)	

	Radon in soil concentration (kBq m^{-3})	
Location	Value	Uncertainty
"Green Ballesteros" (next to the LNR)		
Offices site		

	Radon Exhalation (Bq m^{-2} h^{-1})	
Location	Value	Uncertainty
"Green Ballesteros" (next to the LNR)		

References

Cardellini, F., Chiaberto, E., Garlati, L., Giuffrida, D., Leonardi, F., Magnoni, M., Minchillo, G.,Prandstatter, A., Serena, E., Trevisi, R., Tripodi, R., & Veschetti M. (2016). Main results of the international intercomparison of passive radon detectors under field conditions in Marie Curie's tunnel in Lurisia (Italy). *Nukleonika*, *61*(3), 251-256.

ENUSA (2019). Reclamation of Former Mining Installations. Retrieved February 1, 2019, from http://www.enusa.es/en/areas-de-negocio/medioambiental/restauracion-de-antiguas-instalaciones-mineras/

Fuente, M., Rabago, D., Herrera, S., Quindos, L., Fuente, I., Foley, M., & Sainz, C. (2018). Performance of radon monitors in a purpose-built radon chamber. *Journal of Radiological Protection*, *38*(3), 1111.

Gammadata Instrument AB (2015) Atmos 12: User's Guide www.gammadata.se/products/radiationdetection/radon/radon-struments/atmos-12/

George., A. C. (1996) State-of-the-art instruments for measuring radon/thoron and their progeny in dwellings—a review *Health Phys.* 70 451–63

Gutierrez-Villanueva, J. L., Sainz-Fernández, C., Fuente-Merino, I., Sáez-Vergara, J. C., Correa-Garcés, E., & Quindos-Poncela, L. S. (2013). Intercomparison exercise on external gamma dose rate under field conditions at the laboratory of natural radiation (Saelices el Chico, Spain). *Radiation protection dosimetry*, *155*(4), 459-466.

Gutiérrez-Villanueva, J. L., Sainz, C., Fuente, I., Quindós, L., Quindós, J., Villar, A., Casal, E., Lopez, D., Arteche, D., Fernández, E., & Quindós, L. S. (2011). International Intercomparison Exercise on Natural Radiation Measurements under Field Conditions. Ed. Universidad de Cantabria.

Gutiérrez-Villanueva, J. L., Sainz, C., Fuente, I., Celaya, S., Quindós, L., Quindós, J., Villar, A., Casal, E., Lopez, D., Arteche, D., Fernández, E. & Quindós, L. S. (2015). *Inter-Laboratory Comparison on Indoor Radon Measurements under Field Conditions* (Vol. 201). Ed. Universidad de Cantabria.

Gutiérrez-Villanueva, J. L., Sainz, C., Fuente, I., Celaya, S., Quindós, L., Quindós, J., Villar, A., Lopez, D., Fernández, E. & Quindós, L. S.

(2016). *Intercomparación de monitores de gas radón en condiciones de campo* (Vol. 216). Ed. Universidad de Cantabria.

International Organization for Standardization. (2010). Conformity assessment – General requirements for proficiency testing. ISO/IEC 17043:2010.

International Organization for Standardization. (2015). Statistical methods for use in proficiency testing by interlaboratory comparison. ISO 13528:2015.

Mitev, K., Dutsov, C., Georgiev, S., Boshkova, T., & Pressyanov, D. (2019). Unperturbed, high spatial resolution measurement of Radon-222 in soil-gas depth profile. *Journal of environmental radioactivity*, *196*, 253-258.

Quindós, L. S., Sainz, C., Gutiérrez-Villanueva, J. L., Fuente, I., Celaya, S., Quindós, L., Quindós, J., Fernández, E., & Villar, A. F. (2016). The Laboratory of Natural Radiation (LNR)–a place to test radon instruments under variable conditions of radon concentration and climatic variables. *Nukleonika*, *61*(3), 275-280.

R Core Team. (2018). R: A language and environment for statistical computing [computer software]. http://www.R-project.org/

Sarad GmbH (2017) Radon Scout Home Monitor. Retrieved February 1, 2019, from www.sarad.de/product-detail.php?p_ID=67&cat_ID=2

Sarad GmbH (2019) Radon Scout Monitor. Retrieved February 1, 2019, from www.sarad.de/product-detail.php?p_ID=37&cat_ID=2

Saphymo GmbH (2012) AlphaGUARD Portable Radon Monitor.

Annex 2: Rabago, D., Fuente, I., Celaya, S., Fernandez, A., Fernandez, E., Quindos, J., Pol, R., Cinelli, G., Quindos, L., Sainz, C. Intercomparison of Indoor Radon Measurements Under Field Conditions In the Framework of MetroRADON European Project. Int. J. Environ. Res. Public Health 2020, 17, 1780. doi:10.3390/ijerph17051780

International Journal of
Environmental Research and Public Health

MDPI

Article

Intercomparison of Indoor Radon Measurements Under Field Conditions in the Framework of MetroRADON European Project

Daniel Rabago [1], Ismael Fuente [1], Santiago Celaya [1], Alicia Fernandez [1], Enrique Fernandez [1], Jorge Quindos [1], Ricardo Pol [1], Giorgia Cinelli [2,*], Luis Quindos [1] and Carlos Sainz [1]

1 Radon Group, University of Cantabria, Santander, 39011 Cantabria, Spain; daniel.rabago@unican.es (D.R.); fuentei@unican.es (I.S.); santiago.celaya@unican.es (S.C.); alicia.fernandezv@unican.es (A.F.); enrique.fernandez@unican.es (E.F.); jorge.quindos@unican.es (G.Q.); ricpol@interface.es (R.P.); quindosl@unican.es (L.Q.); sainzc@unican.es (C.S.)

2 European Commission, Joint Research Centre (JRC), I-21027 Ispra, Italy

* Correspondence: giorgia.cinelli@ec.europa.eu

Received: 30 January 2020; Accepted: 3 March 2020; Published: 9 March 2020

Abstract: Interlaboratory comparisons are a basic part of the regular quality controls of laboratories to warranty the adequate performance of test and measurements. The exercise presented in this article is the comparison of indoor radon gas measurements under field conditions performed with passive detectors and active monitors carried out in the Laboratory of Natural Radiation (LNR). The aim is to provide a direct comparison between different methodologies and to identify physical reasons for possible inconsistencies, particularly related to sampling and measurement techniques. The variation of radon concentration during the comparison showed a big range of values, with levels from approximately 0.5 to 30 kBq/m^3. The reference values for the two exposure periods have been derived from a weighted average of participants' results applying an iterative algorithm. The indexes used to analyze the participants' results were the relative percentage difference $D(\%)$, the Zeta score (ζ), and the z-score (z). Over 80% of the results for radon in air exposure are within the interval defined by the reference value and 20% and 10% for the first and the second exposure, respectively. Most deviations were detected with the overestimating of the exposure using passive detectors due to the related degassing time of detector holder materials.

Keywords: radon; proficiency test; quality assurance; metrology; interlaboratory comparison

1. Introduction

Quality assurance is essential within the internal management of laboratories to perform tests and measurements. One of the main tools to carry out such quality control is the periodic participation in interlaboratory comparisons. On one hand, this tool affects the capacity of laboratories to carry out a specific test, and the external information obtained ensures, as far as possible, that the validation of its procedure and its internal quality control strategy are sufficiently effective with a certain degree of confidence. On the other hand, such participation includes a high potential for improvement by forcing the laboratory, given unsatisfactory results, to test its ability to detect the possible source of the error, which could be from the inadequate qualification of the staff, the incomplete validation of the test procedure, or a punctual error in the operation of the equipment, etc. [1,2].

In the field of radon, the European Council Directive 2013/59/EURATOM (EU-BSS) lays down legal limits for radon concentrations in indoor air. Therefore, to fulfill the EU-BSS requirements, it is necessary to improve the radon metrological infrastructure. Quality assurance and the traceability of in situ and laboratory measurements of radon are required. The typical situation where an intercomparison of the radon devices is developed is in a radon chamber under a radon reference atmosphere and stable environmental conditions [3]. However, the real scenarios, dwellings, and working places, where the radon monitors or passive detectors are laid can differ from the stable conditions usually performed to test them. Accordingly, some intercomparisons have been carried out in field environments in order reproduce such situations, which should be periodically conducted [4–7].

The exercise presented in this article is the comparison of indoor radon gas measurements under field conditions performed with passive detectors, giving an integrated measurement over time, and active monitors, continuously monitoring radon concentration, within the European MetroRADON project (http://metroradon.eu/). This project aims to develop reliable techniques and methodologies to enable international traceable radon activity concentration measurements and calibrations at low radon concentrations and contribute to the creation of a coordinated metrological infrastructure for radon monitoring in Europe. The specific objective of the intercomparison is to provide a direct comparison between different methodologies and to identify physical reasons for possible inconsistencies, which are particularly related to sampling and measurement techniques.

2. Materials and Methods

2.1. Site Description

The intercomparison was carried out in the former uranium mine managed by the Spanish National Uranium Company ENUSA (Saelices el Chico, Salamanca, Spain). The reclamation of the uranium mining operations (exploited from 1972 to 2000) and the dismantling of the uranium concentrate factory started in 2001. Currently, the decommissioning activities continue restoring the affected natural space with the purpose of returning to the geological, radiological, and environmental conditions that the areas had before their exploitation began. Nowadays, the activities are mainly related with water treatment and the stability of tailings. There is acid drainage in the mine, due to the presence of pyrite in rocks of the site, which pollutes the water [8]. One of the buildings was chosen to house the Laboratory of Natural Radiation (LNR) for the calibration and testing of instruments and detectors for the measurement of natural radiation (see Figure 1) [9,10]. This place has been used to carry out interlaboratory exercises both measuring radon concentration and gamma dose rate under natural environmental fluctuations [4,5,11]. The high radioactive content in the soil along with the environmental conditions make this location a suitable place to conduct these kinds of activities.

(**a**) (**b**)

Figure 1. (a) Aerial view of the Laboratory of Natural Radiation (LNR) and surroundings where the intercomparison was developed; (b) Picture of LNR building taken from the south part.

In the LNR ground floor, there are two rooms designed as radon chambers (Room1 and Room2) with approximately 45 m^3 volume each. Room1 has no direct connection to the exterior or any mixing system settled, while Room2 has an artificial ventilation system installed. The radon concentration homogeneity in the measurement area of Room1 is checked periodically providing statistical suitable values [9]. The radon source is the uranium mine underground soil, which has high radium content. In the east part of the LNR, a meteorological station is set up to monitor environmental conditions.

Outside of LNR, there is a place called "Green Ballesteros", where a 5×5 m^2 and 1.5 m deep hole was dug out and filled with homogeneous soil with low radioactive content (^{226}Ra concentration about 43 ± 10 Bq kg^{-1} (k=2) and a gamma dose rate around 110 ± 5 nGy h^{-1} (k=2) at 1 m height, k means the coverage factor) with the purpose of carrying out other kinds of intercomparisons, related with external gamma dose rate and radon in soil measurements. ^{226}Ra concentration in soil was determined from 10 samples distributed in a homogenous way in the entire volume of the Green Ballesteros by the Laboratory of Environmental Radioactivity, University of Cantabria (LaRUC), which is accredited according to UNE-EN ISO/IEC 17025:2005 for activity concentration measurements of soil by gamma spectrometry using a high-purity germanium detector. The gamma dose rate was measured in situ drawn on two Reuter-Stokes devices calibrated at the Dosimetry Standards Laboratory of CIEMAT (Centre for Energy-Related, Environmental and Technological Research) [11].

2.2. Participants

The number of participants was 20 from 14 countries giving priority to MetroRADON partners. The list of participants is given in Table 1. There is no correlation between this table and the code assigned to each participant in the results section.

Table 1. Participants in the intercomparison sorted by alphabetical order.

Acronym	Institution	Country
CIEMAT	Centro de investigaciones energéticas, medioambientales y tecnológicas	Spain
CLOR	Central Laboratory for Radiological Protection	Poland
ENEA	ENEA Radon Service	Italy
INAIL	Italian National Institute for Insurance against Accidents at work	Italy
IRSN	Institut de Radioprotection et de Sûreté Nucléaire	France
JRC	Joint Research Centre	Italy
LaRUC (UC)	Laboratory of environmental radioactivity, University of Cantabria	Spain
LRAB - UEX	LRAB – Universidad de Extremadura	Spain
LRG	Laboratorio de Radón de Galicia	Spain
LRN-UC	Laboratorio de Radioatividade Natural - Universidade de Coimbra	Portugal
NRCN	Nuclear Research Center Negev	Israel
PUCP	Pontificia Universidad Católica Del Perú	Peru
RADONOVA	Radonova Laboratories AB	Sweden
Radosys	Radosys/Radosys Atlantic	Portugal/Hungary
RERA-CIEMAT	Centro de investigaciones energéticas, medioambientales y tecnológicas	Spain
STUK	Radiation and Nuclear Safety Authority	Finland
SUBG	Sofia University "St. Kliment Ohridski"	Bulgaria
SUJCHBO	National Institut for NBC Protection	Czech Republic
TR	TECNO RAD s.u.r.l.	Italy
UBB	Babes-Bolyai University	Romania

2.3. Radon exposures, devices, and logistic arragements

Radon in air exposure has been evaluated using passive detectors and active monitors inside Room1 of the LNR for two periods. The radon concentration inside Room1 is monitored remotely in order to decide when to remove the passive detectors of every exposure period. The devices were placed in Room1 on 5 December 2018 and were taken off on 6 and 8 December 2018 for the first and second exposure respectively, according to the schedule shown in Table 2.

Table 2. Start and end dates for each exposure ((UTC+01:00) Brussels, Copenhagen, Madrid, Paris).

	Start date	End date
First exposure E1:	05/11/2018 12:00	06/11/2018 1:00
Second exposure E2:	05/11/2018 12:00	08/11/2018 10:00

Each participation with passive detectors required 30 units: 10 detectors for the first exposure, 10 for the second exposure, and 10 transit detectors. A total of 23 groups of passive detectors and 22 active monitors were exposed in Room1 (see Figure 2). The statistical criteria employed to select the number of exposed and transit detectors were based on the usual international intercomparison requirements [6,7,12] and with enough of a number that assured obtaining a representative average with a low uncertainty.

Figure 2. Radon devices inside Room1 (LNR: Laboratory of Natural Radiation).

There were several types of passive devices used by the participants. Most of them were solid-state nuclear track detectors (SSNTD), CR-39 and LR 115, in which the detection principle is based on the microscopic defects in the detection material caused by the alpha particles of radon and its daughters that can be revealed by etching treatment [13,14]. In addition, electret ion chambers (EICs), consisting of a stable electret mounted inside an electrically conducting plastic chamber, were used for measuring radon. The negative ions produced inside the diffusion chamber by the alpha particles emitted reduce the electret's charged surface [15]. Other procedures were implemented, e.g., using DVDs (Digital Versatile Disc) half made of polycarbonate (used as a solid state track detector) and polycarbonate foils used as a radon absorber [16,17]. The features of diffusion chambers, holders, material quality, and manufactures were diverse, too. The overall characteristics given by participants are shown in Table 3.

Table 3. Passive detector features provided by the participants.

Detector	Diffusion chamber
CR-39 RSKS 100 mm^2 (Radosys)	Diameter 26 mm, height 55 mm 29 cm^3 volume
CR-39 24.7×36.7×1.40 (mm) (Mi-Net)	ENEA patent
CR-39 Radout 25×25×1.5 (mm) (Mi.am)	Diameter 50 mm, height 20 mm
CR-39 TASTRAK 13×37×1 (mm) (Tasl)	Diameter 58 mm, height 20 mm NRPB/SSI
CR-39 Duotrack (Radonova)	Diameter 58 mm, height 40 mm

CR-39 Radtrak2 (Radonova)	Diameter 58 mm, height 20 mm NRPB/SSI
CR-39 Rapidos (Radonova)	Diameter 58 mm, height 40 mm
ST Electret Teflon (E-PERM)	L-OO Chamber 58 mL
ST Electret Teflon (E-PERM)	S Chamber 210 mL
LR-115 type2 400 mm^2 (DOSIRAD)	Diameter 60.4 mm, height 27.6 mm Own design
LR-115 (KODAK) RAMARN device 0.012 mm film of cellulose nitrate, and coated on 0.1 mm thick polyester base	Polypropylene chamber 700 cm^3 volume
Makrofol 75.7 mm^2 STUK design "Radonpurkki"	Diameter 20 mm, height 71 mm 79 cm^3 volume
DVD half made of polycarbonate and two thin Makrofol N foils	Thin CD case

In this intercomparison, different active monitors were used. These kinds of electronic devices provide a time-series of radon concentration in which radon detection is based on detect alpha radiation (by ionization chamber, gross alpha counting, or alpha spectrometry) [18–20]. The operation modes and features are shown in Table 4. This information has been obtained from the manufacturer's technical specifications.

Table 4. Active monitor features used in the intercomparison.

Monitor	Detection technology	Sensitivity (cpm at 1 kBq m^{-3})
AlphaGUARD	Ionization chamber	50
ATMOS12 DPX	Ionization chamber	20
SARAD EQF 3120	Silicon detector	7
Radon Scout	Silicon detector	1.8
Radon Scout Home	PIN photo diode	0.1

The organizers placed/removed the passive detectors and active monitors from Room1. After each exposure, passive detectors were stored in a low radon concentration area that was less than 10 Bq m^{-3} determined with an AlphaGUARD (S/N EF 1763), which assures that the contribution to the total exposure is negligible. After two days, they were sealed in radon-proof aluminium bags in order to allow a proper degassing. Active monitors were turned off at the end of the second exposure. Transit detectors were stored in their original bags until the end of the second exposure. Afterwards, they were sealed in radon proof aluminium bags in order to simulate the exposed detectors conditions.

Participants have provided the exposure value and its uncertainty (k = 1) for each passive detector and the declared value for the first and second exposure period expressed in kBq m^{-3} h. In the case of active monitors, the overall exposure for each period was given; the individual radon concentration every hour was also included.

The results provided by participants have been coded in order to maintain their anonymity. Such codification follows the rule:

LxxTn

where xx is the number assigned to each participant from 01 to 20, T is the type of measurement: A for active monitor and P for passive detectors, and n is the correlative number for more than one kind of measurement group.

2.4. Data Analysis

The determination of the assigned value and its standard uncertainty for each radon in air exposure have been obtained by using the consensus values from participant results applying an iterative algorithm according to ISO 13528:2015 [21]. This algorithm considers the results of all participants and relocates the extreme values within the interval of acceptable deviation.

An outlier study has been applied in order to know the extreme values. The outlier values were found from the boxplot representation and the interquartile analysis. In this case, an outlier is defined as a data point that is located 1.5 times the interquartile range (IQR) above the upper quartile and below the lower quartile. The interquartile range is defined as the difference between the third quartile (75th percentile) and the first quartile (25th percentile): IQR = (Q_3–Q_1).

The robust average and robust standard deviation denoted by E_{ref} and s^* for each radon exposure have been calculated using Algorithm 1.

1: There are p items of results denoted as:
$E_i = E_1, E_2, E_3, \ldots, E_p$
2: Calculate initial values for E_{ref} and s^* as:
E_{ref} = median of E_i
$s^* = 1.483$ median of $|E_i - E_{ref}|$
3: Update the values of E_{ref} and s^* as follows. Calculate:
$\delta = 1.5\ s^*$

$$E_i^* = \begin{cases} E_{ref} - \delta & \text{when} \quad E_i < E_{ref} - \delta \\ E_{ref} + \delta & \text{when} \quad E_i > E_{ref} + \delta \\ E_i & \text{otherwise} \end{cases}$$

4: Calculate the new values of E_{ref} and s^* from:
E_{ref}= mean of E_i^*
s^*= 1.134·SD (E_i^*)

Algorithm 1. Algorithm used to calculate the robust average E_{ref} and robust standard deviation s^* taken from ISO 13528:2015 denoted as Algorithm A. The parameters are updated following the steps 3 and 4 iteratively until the process converges.

Once the robust average and robust standard deviation have been calculated for each exposure period, the standard uncertainty of the assigned value is estimated as:

$$u(E_{ref}) = 1.25 \frac{s^*}{\sqrt{p}}. \quad (1)$$

The indexes used to analyze the participants' results are the relative percentage difference D(%), the Zeta score (ζ), and the z-score (z).

The relative percentage difference D(%) has been introduced to quantify the difference between the participant's result and the reference value obtained as consensus. Therefore:

$$D_i(\%) = 100 \cdot \frac{E_i - E_{ref}}{E_{ref}} \quad (2)$$

where E_i is the exposure result i given by the participant.

The Zeta score (ζ) is a statistical index used to compare intercomparison results where the uncertainty in the measurement result is included. It is given by the following equation:

$$\zeta_i = \frac{E_i - E_{ref}}{\sqrt{u^2(E_i) + u^2(E_{ref})}} \quad (3)$$

in which $u(E_i)$ is the participant's own estimate of the standard uncertainty of its result.

The Z-score (z) index is calculated as follows:

$$z_i = \frac{E_i - E_{ref}}{\sigma_p} \quad (4)$$

where σ_p is the standard deviation for the intercomparison assessment estimated as 20% of the reference value for the first exposure and 10% of the reference value for the second one. This parameter should meet the following criterion: $u(E_{ref}) < 0.3\ \sigma_p$.

These indexes are interpreted as follows:

$|\zeta|; |z| \leq 2.0$ result is considered satisfactory
$2.0 < |\zeta|; |z| < 3.0$ result is considered to give a problem
$|\zeta|; |z| \geq 3.0$ is considered not satisfactory

The Zeta score (ζ) is used together with the z-score (z) as an aid for improving the performance of participants. If a participant obtains a z-score higher than the critical value of 3.0, they may find it valuable to reassess their procedure with the subsequent uncertainty evaluation for that procedure. If the participant's ζ score also exceeds the critical value of 3.0, it implies that the participant's uncertainty evaluation does not include all significant sources of uncertainty. However, if a participant obtains a z-score ≥ 3.0 but a ζ score ≤ 2.0, this demonstrates that the participant may have assessed the uncertainty of their results accurately but that their results do not meet the performance expected for the proficiency testing scheme. The interpretation guidelines are shown in Table 5.

Table 5. Summary of guidelines to understand the Zeta score (ζ) and z-score (z).

ζ score	z-score	Action to take
Satisfactory	Satisfactory	Participant's result is good. No action is required.
Not Satisfactory	Satisfactory	Participant's claimed uncertainty is too low, but the result fulfills the intercomparison requirements.
Satisfactory	Not Satisfactory	Participant's uncertainty assessment is accurate but the results do not fulfill the intercomparison requirements.
Not Satisfactory	Not Satisfactory	Participant's result is biased in excess. A complete revaluation should be performed.

3. Results

Participants submitted one exposure result together with its uncertainty per group of passive detectors and/or actively monitor for the first exposure E1 and for the second exposure E2.

The variation of radon concentration in Room1 shows a big range of values, with levels from approximately 0.5 to 30 kBq m^{-3}. As an example, the radon concentration measurements of the Laboratory of Environmental Radioactivity, University of Cantabria (LaRUC), taken by the calibrated device AlphaGUARD (S/N AG000032), are shown in Figure 3. The environmental conditions in Room1 during the exercise are the following: it is observed that the variation of temperature is quite stable, with an absolute difference of 1 °C, the atmospheric pressure average was (935 ± 5) hPa with an absolute variation of 14 hPa, and the mean relative humidity was (63 ± 2)% with a variation range of 11%.

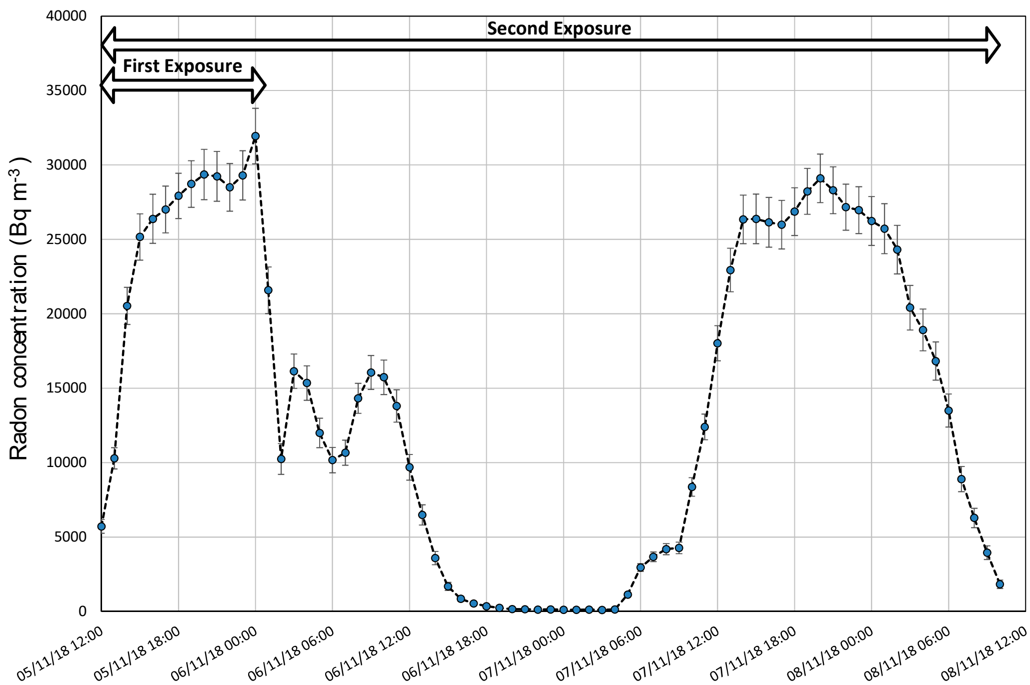

Figure 3. Radon concentration in Room1 during the intercomparison exercise according to LaRUC. Data is displayed every hour.

The assigned values used as a reference for each exposure period are derived from a weighted average of participants' results applying the iterative algorithm described above according to ISO 13528:2015. Table 6 shows the robust average E_{ref}, the robust standard deviation s^*, the standard uncertainty $u(E_{ref})$, the number of results p and the standard deviation for the intercomparison assessment σ_p estimated as 20% of reference value for the first exposure and 10% of the reference value for the second one. This parameter meets the criterion: $u\left(E_{ref}\right) < 0.3\ \sigma_p$.

Table 6. Reference parameters of the first exposure E1 and second exposure E2 expressed in kBq m^{-3} h obtained from participant results according to ISO 13528:2015. p is the dimensionless number of results.

Exposure	E_{ref}	$u(E_{ref})$	σ_p	s^*	p
E1	356	8	71	43	45
E2	1014	13	101	68	41

The boxplot diagram is shown in Figure 4, where the quartiles information and the outliers whose laboratory codes are displayed. There are no statistical differences between the reference exposure value calculated taking into account the total amount of results and the one calculated without considering outliers. Therefore, all the results have been considered to calculate the reference values.

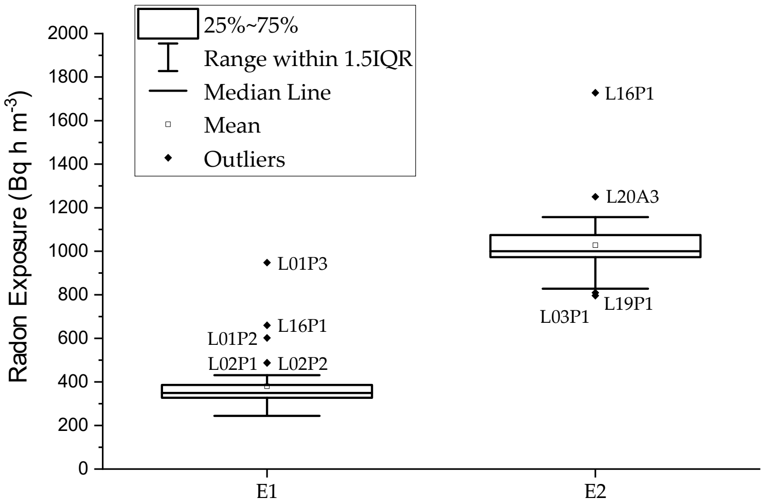

Figure 4. Boxplot diagram of the participant's results for exposures E1 and E2. Outliers are identified.

Participant's results for the first radon in air exposure are given in Figure 5. Each value is presented with its uncertainty (k = 1). The solid line represents the reference value obtained through consensus (356 kBq m^{-3} h), and the dashed lines denote the standard deviation for the interlaboratory assessment estimated as 20% of the reference value. Figure 6 shows the results for the second exposure, with the reference value of 1014 kBq m^{-3} h indicated with a solid line. In this case, the dashed lines represent 10% of that reference value.

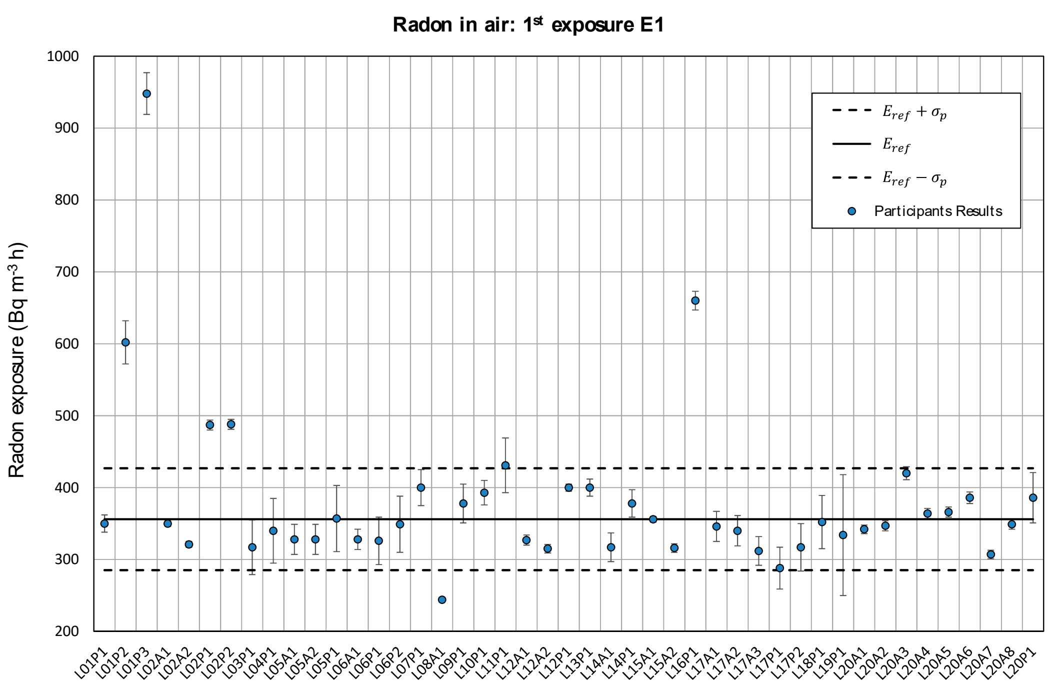

Figure 5. Participant's results for the first exposure E1 with its associated uncertainty (k = 1). The exposure reference value is shown with a solid line and the standard deviation $\sigma_p = 0.2E_{ref}$ with dashed lines.

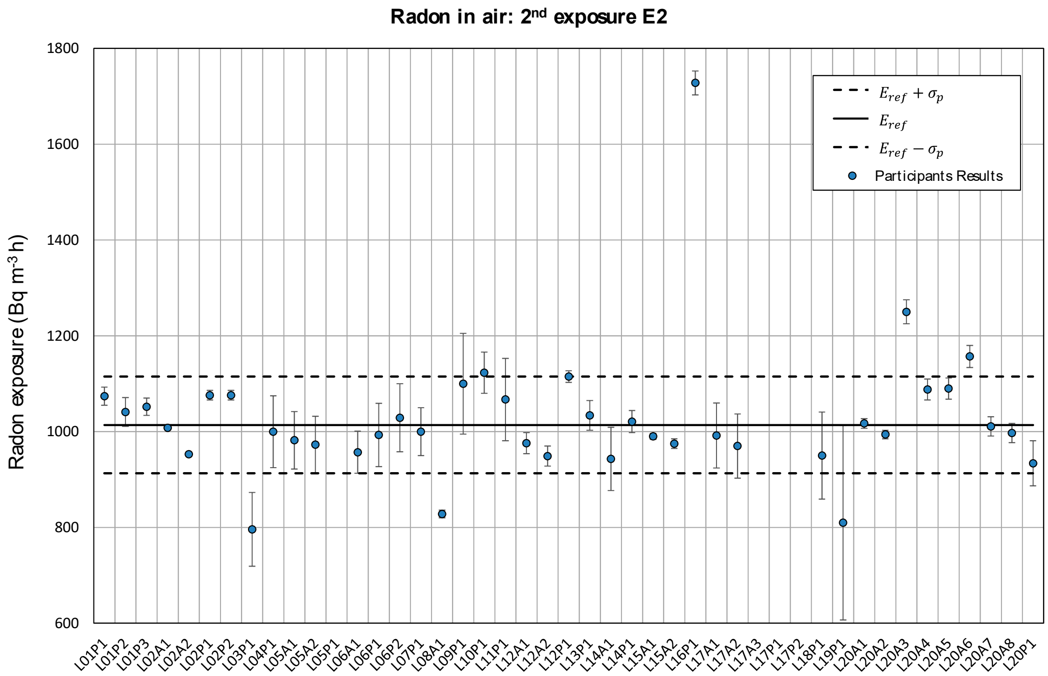

Figure 6. Participant's results for the second exposure E2 with its associated uncertainty (k = 1). Exposure reference value is shown with a solid line and the standard deviation $\sigma_p = 0.1E_{ref}$ with dashed lines.

About 80% of the results presented in Figure 5 and Figure 6 are within the interval defined by the exposure reference value E_{ref} and the standard deviation σ_p established as 20% and 10% for the first and the second exposure, respectively. The relative difference D(%) between each single value and the reference is shown in Figure 7. The anomalous values shown in Figure 4 are clearly out of those intervals.

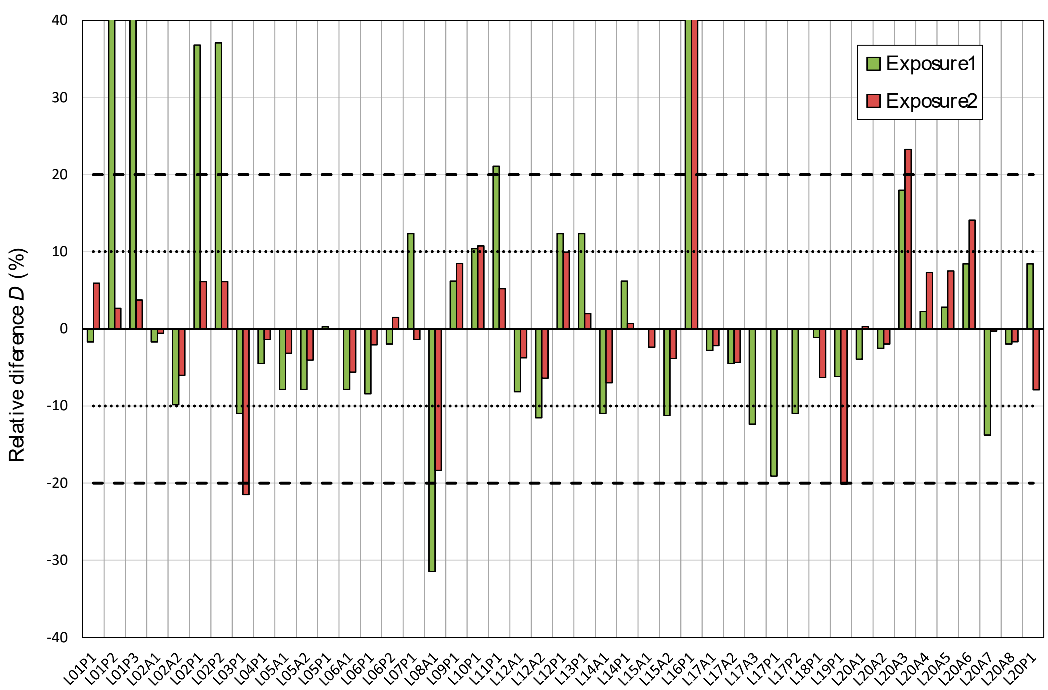

Figure 7. Relative difference of participant's results to the mean value for the first and second exposure. Intervals established for the first exposure (± 20%) and second exposure (± 10%) are indicated. In some cases, the value is out of scale in order to improve the graph view.

Figure 8 and Figure 9 show the graphical representation of indexes used to assess the participant's results. In some cases, the value is out of scale in order to improve the graph view. In addition, Table 7 shows the percentage of results that are within the limits for each index. For the relative difference, the percentage of results within 10% and 20% of the reference exposure in each case is presented.

The overall performance of results given by z-score is satisfactory; about 90% of the results have a value lower than 2.0 for both exposures. Only the results of three cases have a z-score value above 3.0 for the first exposure and one result for the second exposure period. Regarding the Zeta score, about 60% of results are satisfactory ($|\zeta| \leq 2.0$); however, 29% of results for the first exposure and 20% for the second exposure period are not satisfactory, with a Zeta score $|\zeta| \geq 3.0$.

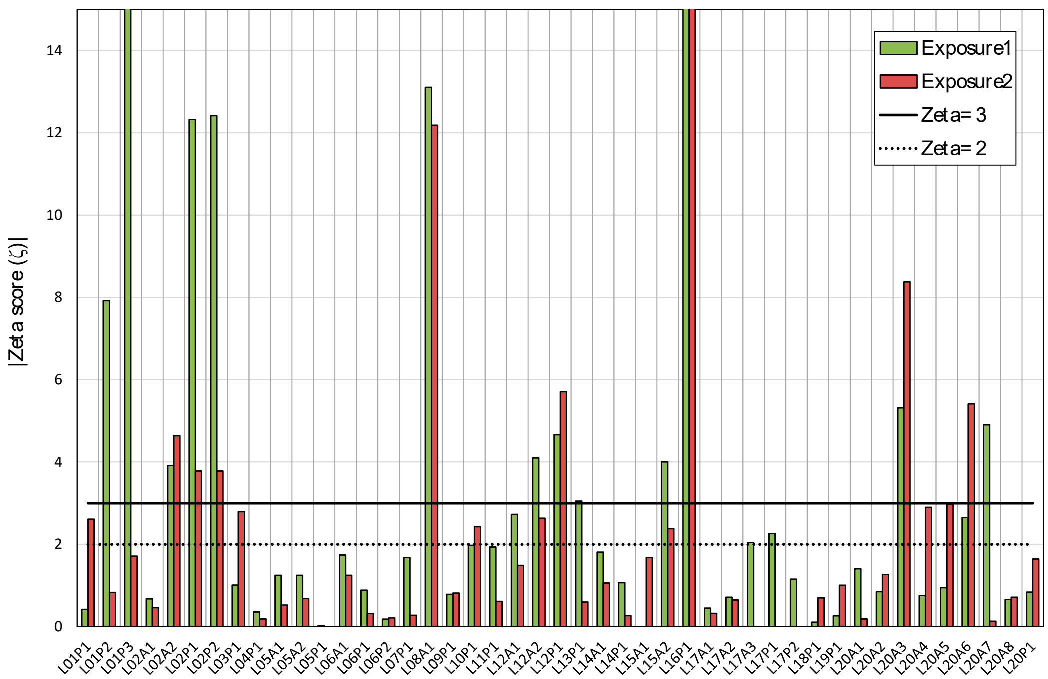

Figure 8. Absolute values of Zeta score for the first and second exposure.

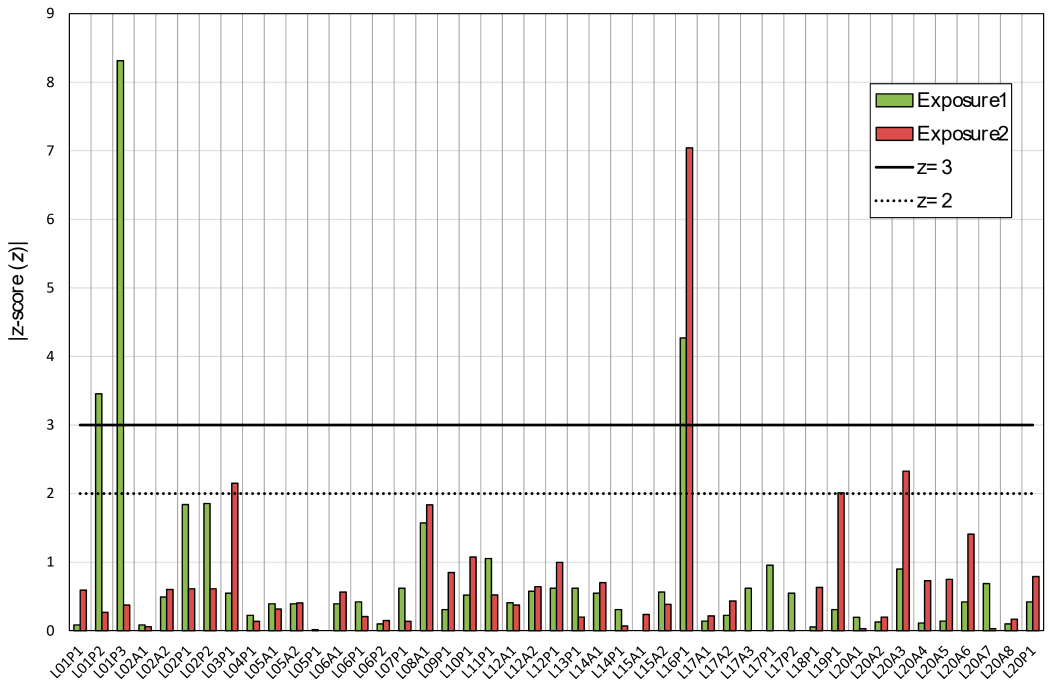

Figure 9. Absolute values of z-score for the first and second exposure.

Table 7. Percentage of results that are within the limits for relative difference *D*(%), Zeta score (ζ), and z-score (z) for each exposure period classified by passive detectors and active monitors.

	Passive detectors		**Active monitors**	
	Results E1 (%)	Results E2 (%)	Results E1 (%)	Results E2 (%)
$\lvert D(\%) \rvert \leq 10\%$	43	80	68	86
$\lvert D(\%) \rvert \leq 20\%$	74	85	95	95
$\lvert \zeta \rvert \leq 2.0$	65	65	59	62
$2.0 < \lvert \zeta \rvert < 3.0$	4	15	14	19
$\lvert \zeta \rvert \geq 3.0$	31	20	27	19
$\lvert z \rvert \leq 2.0$	87	85	100	95
$2.0 < \lvert z \rvert < 3.0$	0	10	0	5
$\lvert z \rvert \geq 3.0$	3	5	0	0

4. Conclusions

An interlaboratory exercise of indoor radon under field conditions has been carried out in the Laboratory of Natural Radiation (LNR) between 5 and 8 December 2018. The facility is located in the former uranium mine managed by the Spanish National Uranium Company ENUSA (Saelices el Chico, Salamanca, Spain). Radon in air measurements were assessed from two exposure periods.

Radon in air reference values for each exposure were obtained through consensus from participant's results applying an iterative algorithm according to ISO 13528:2015. The indexes used to analyze the participants results are the relative percentage difference *D*(%), Zeta score (ζ), and z-score (z).

Most z-score results are satisfactory; about 90% of the results have a value lower than 2 for both exposures. Only the results for three cases for the first exposure and one result for the second exposure period are not satisfactory, with z-score values higher than 3.0.

Regarding the Zeta score, about 60% of results are satisfactory ($|\zeta| \leq 2.0$); however, 29% of the results for the first exposure and 20% for the second exposure period are not satisfactory, with a Zeta score $|\zeta| \geq 3.0$.

Over 80% of the results for radon in air exposure are within the interval defined by the reference value and the standard deviation, which were established as 20% and 10% for the first and the second exposure, respectively.

Five results of the first exposure are considered outliers. All of them are passive detectors and are overestimating the exposure from approximately 40% to 160%. Such deviations could be related with the degassing time of detector holder materials. Radon could get adsorbed in it for a long time, so even after two days, when the detectors were put in radon proof bags and sealed. A further difficulty in this intercomparison exercise is that the exposures are reached in a short time period with high radon concentrations in air. At the end of the first exposure period, there was a radon concentration in air around 30 kBq m^{-3} which can cause the holder degassing problem previously mentioned. In case of the second exposure, the radon concentration was under 2 kBq m^{-3} at the end of that period, therefore reducing the exposure increase due to the possible effect of adsorption and degassing.

In general, active monitors provide better results taking into account the z-score and the relative difference parameters. However, in case of the Zeta score, worse results are obtained, which can mean that the estimated uncertainty does not include all significant sources. The exercise was successful, taking into account the large number of different devices used, especially in passive detectors where the holder materials, diffusion chamber volume, detectors area, or detection principle were diverse.

Author Contributions: Conceptualization, C.S., L.Q., D.R., and G.C.; methodology, D.R., I.F., S.C., A.F., E.F., J.Q., R.P., L.Q., and C.S.; software, D.R.; formal analysis, C.S., L.Q., and D.R.; data curation, D.R.; writing—original draft preparation, D.R. and C.S.; writing—review and editing, C.S., L.Q., D.R., G.C., I.F., and S.C.; funding acquisition, C.S. and L.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Metrology Programme for Innovation and Research (EMPIR), JRP-Contract 16ENV10 MetroRADON (http://www.euramet.org). The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States.

Acknowledgments: Authors would like to express their gratitude to the National Uranium Company of Spain (ENUSA) for the collaboration given during all these years. The cooperation of the staff of Saelices el Chico center was essential to carry out this research.

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

References

1. ISO. *ISO/IEC 17043:2010 General Requirements for Proficiency Testing*; ISO: Geneva, Switzerland, 2010.
2. ISO. *ISO/IEC 17025: General Requirements for the Competence of Testing and Calibration Laboratories*; ISO: Geneva, Switzerland, 2017.
3. Janik, M. Intercomparisons Exercises of Radon and Thoron Monitors Provided by Four Laboratories: A Review. *Jpn. J. Health Phys.* **2017**, *52*, 114–121, doi:10.5453/jhps.52.114
4. Gutiérrez-Villanueva, J.L.; Sainz, C.; Fuente, I.; Quindós, L.; Quindós, J.; Fernández, A.; Casal, E.; Lopez, D.; Arteche, D.; Fernández, E.; et al. *International Intercomparison Exercise on Natural Radiation Measurements under Field Conditions*; PUbliCan: Santander, Spain, 2012; pp. 3-7.
5. Gutiérrez-Villanueva, J.L.; Sainz, C.; Fuente, I.; Celaya, S.; Quindós, L.; Quindós, J.; Villar, A.; Casal, E.; Lopez, D.; Arteche, D.; et al. *Inter-Laboratory Comparison on Indoor Radon Measurements under Field Conditions*; PUbliCan: Santander, Spain, 2015; pp. 9-14.
6. Cardellini, F.; Chiaberto, E.; Garlati, L.; Giuffrida, D.; Leonardi, F.; Magnoni, M.; Minchillo, G.; Prandstatter, A.; Serena, E.; Trevisi, R.; et al. Main results of the international intercomparison of passive radon detectors under field conditions in Marie Curie's tunnel in Lurisia (Italy). *Nukleonika* **2016**, *61*, 251–256, doi:10.1515/nuka-2016-0042.
7. Berlier, F.; Cardellini, F.; Chiaberto, E.; Garlati, L.; Giuffrida, D.; Faure, M.; Leonardi, F.; Magnoni, M.; Minchillo, G.; Prandstatter, A.; et al. Main results of the second AIRP international radon-in-field

intercomparison for passive measurement devices. *Radiat. Meas.* **2019**, *128*, 106177, doi:10.1016/j.radmeas.2019.106177.

8. ENUSA Industrias Avanzadas S.A. *Annual Report 2018*; A.F. Diseño y Comunicación: Salamanca, Spain, 2019; pp. 50–51.
9. Quindós, L.S.; Sainz, C.; Gutiérrez-Villanueva, J.L.; Fuente, I.; Celaya, S.; Quindós, L.; Quindós, J.; Fernández, E.; Fernández, A. The Laboratory of Natural Radiation (LNR)–a place to test radon instruments under variable conditions of radon concentration and climatic variables. *Nukleonika* **2016**, *61*, 275–280, doi:10.1515/nuka-2016-0046.
10. Pressyanov, D.; Quindos Poncela, L.S.; Georgiev, S.; Dimitrova, I.; Mitev, K.; Sainz, C. ; Rabago, D. Testing and Calibration of CDs as Radon Detectors at Highly Variable Radon Concentrations and Temperatures. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3038, doi:10.3390/ijerph16173038.
11. Gutierrez-Villanueva, J.L.; Sainz-Fernández, C.; Fuente-Merino, I.; Sáez-Vergara, J.C.; Correa-Garcés, E.; Quindos-Poncela, L.S. Intercomparison exercise on external gamma dose rate under field conditions at the laboratory of natural radiation (Saelices el Chico, Spain). *Radiat. Prot. Dosim.* **2013**, *155*, 459–466, doi:10.1093/rpd/nct016.
12. Wu, Y.; Cui, H.; Zhang, Q.; Shang, B. Intercomparisons for integrating the radon–thoron detector of NIRP, China with NIRS, Japan. *Radiat. Prot. Dosim.* **2014**, *164*, 398–401, doi:10.1093/rpd/ncu286.
13. Wasikiewicz, J.M.; Daraktchieva, Z.; Howarth, C.B. Passive etched track detectors application in outdoor radon monitoring. *Perspect. Sci.* **2019**, *12*, 100416, doi:10.1016/j.pisc.2019.100416.
14. Gericke, C.; Ghose, R.; Jönsson, G.; Freyer, K.; Treutler, H.C.; Enge, W. The calibration of the solid state nuclear track detector LR 115 for radon measurements. *Radiat. Meas.* **1999**, *31*, 363–366, doi:10.1016/S1350-4487(99)00175-4.
15. Kotrappa, P. Long term stability of electrets used in electret ion chambers. *J. Electrost.* **2008**, *66*, 407–409, doi:10.1016/j.elstat.2008.04.008.
16. Pressyanov, D.; Buysse, J.; Van Deynse, A.; Poffijn, A.; Meesen, G. Indoor radon detected by compact discs. *Nucl. Instrum. Methods Phys. Res.* **2001**, *457*, 665–666, doi:10.1016/S0168-9002(00)00794-4.
17. Pressyanov, D. Use of Polycarbonate Materials of High Radon Absorption Ability for Measuring Radon. *Rom. J. Phys.* **2013**, *58*, S221–S229.
18. George, A.C. State-of-the-art instruments for measuring radon/thoron and their progeny in dwellings-a review. *Health Phys.* **1996**, *70*, 451–463, doi:10.1097/00004032-199604000-00001.
19. El-taher, A. An Overview of Instrumentation for Measuring Radon in Environmental Studies. *J. Radioanal. Nucl. Appl.* **2018**, *3*, 135–141, doi:10.18576/jrna/030302.
20. Fuente, M.; Rabago, D.; Herrera, S.; Quindos, L.; Fuente, I.; Foley, M.; Sainz, C. Performance of radon monitors in a purpose-built radon chamber. *J. Radiol. Prot.* **2018**, *38*, 1111, doi:10.1088/1361-6498/aad969.
21. ISO. *ISO 13528:2015 Statistical Methods for Use in Proficiency Testing by Interlaboratory Comparisons*; ISO: Geneva, Switzerland, 2015; pp. 53-54.

Annex 3: Report: *Testing of comparability between short-term and long-term radon measurement data.* S. Baumann, V. Gruber, G. Cinelli.

16ENV10 MetroRADON

Activity 3.4.2

Testing of comparability between short-term and long-term radon measurement data

Lead organization: Joint Research Centre, JRC

Authors:

Sebastian Baumann, Valeria Gruber (AGES)
Giorgia Cinelli (JRC)

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

Table of contents

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, 2013), 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 the end-users (e.g. companies providing radon measurement, construction industry) and the scientific community. More details about the MetroRADON project can be found at the project website (MetroRADON, 2020).

Radon surveys and radon measurements are the basis in the radon strategies in all countries. Radon surveys and radon measurements are carried out quite differently in European countries, dependent on political decisions, aim of the survey, availability of resources and infrastructure. The use of different methodologies may lead to inconsistencies, followed by communication problems within stakeholders.

Within the MetroRADON project a specific work package is included to compare and harmonise radon measurement methodologies in Europe. It is not realistic that harmonization of European radon data can be done from scratch (bottom-up approach), as a lot of surveys and data exist already and established methodologies are under use. For this data a “top-down” harmonization to make results comparable and jointly interpretable would be necessary. Therefore, it is necessary to have a picture of the whole radon situation in Europe to harmonise the radon protection standards of the public. One task of the respective MetroRADON work package is the development of options for harmonization of indoor and geogenic radon data. Among others, one activity to fulfill the goals of this task is the testing of comparability between existing short-term and long-term radon measurement data in order to assess the feasibility of data merging and data set expansion. The results and conclusions of this exercise are discussed in this report.

2. Introduction

Radon indoor measurements are commonly used to predict radon risk. The amount of data samples is crucial and the question arises if measurements from different surveys or different measurement layouts in the same survey can be merged to predict radon risk.

In the MetroRADON project, activity 3.4.2 was included to find answers to the above-mentioned problem:

A3.4.2	JRC and AGES will test the comparability between existing short-term and long-term radon measurement data in order to assess the feasibility of data merging and therefore data set expansion.	**JRC**, AGES

The exercise was designed to use existing data sets and already available analysis from literature to analyse if different radon measurements techniques in the same area can be incorporated to improve radon risk prediction. In the exercise, two different data sets (see chapter 3) are used. The focus of the analysis is to quantify if different measurement techniques with different measurement durations show similar distributions and if these measurements can be combined to predict radon risk. In addition, studies carried out in Belgium (Cinelli et al., 2009) are included in the analysis and overall discussion.

3. Data

For the analysis two different datasets of indoor radon concentrations (IRC) were available, which differ in major aspects, like sample strategy or sample size, and are summarized in the following.

3.1 Identical locations data set (ILD)

The identical location data set (ILD) is one of very few data sets, where at the same location during the same time period two radon measurement techniques with different exposure times have been carried out. In our analysis it serves us as a benchmark if under maximal possible constraints in a radon-survey the results of two measurement methods with different exposure times can be incorporated to predict radon risk.

The ILD data have been collected during a pilot education project titled “Radon in schools” carried out by the JRC at the European school of Varese between January and June 2019. The main objective of the project was to give the possibility to the students to get familiar with JRC activities on ionizing radiation through a project based on radon measurements. The project got the students involved in a scientific experimental activity, focusing on both data collection with the proper instrumentation and data analysis. The radon track-etch and electret detectors have been assembled by the students at school and then taken to their homes to be exposed.

The ILD features Rn-measurements in 47 rooms in different houses in the same geological area, where every room was measured with two measurement devices, track-etch and electret. Exposure duration for track etch measurements was 2 months (mid-March 2019 – mid-May 2019) and 1 week (mid-March 2019) for electret measurements (starting with the two detectors the same day). In 24 locations (rooms) both measurement techniques had results and for the further analysis we used only the data where results for both measurements methods in one room were available. Figure **1** illustrates the IRC of the sample locations for the measurement devices. The differences of IRC at the same location for track etch and electret ranges from close to 0 up to 70 Bq/m³ and show a high variability. More frequently track etch shows higher IRC than electret but interestingly both cases occur commonly. Considering maximum values of IRC of about 120 Bq/m³ a difference of up to 70 Bq/m³ is very high. This already indicates that for a single location it might not be sufficient to rely on short-term measurements to identify average long-term concentrations and doing so could lead to underestimation as well as to overestimation of the average IRC. These differences for single locations are relevant for decisions concerning that location, like radon remediation, but do not generally indicate that populations could not be merged for tasks as predicting radon risk based on multiple measurements. We further investigate in the Results section how radon risk prediction for the ILD is only minor influenced by the different measurement techniques.

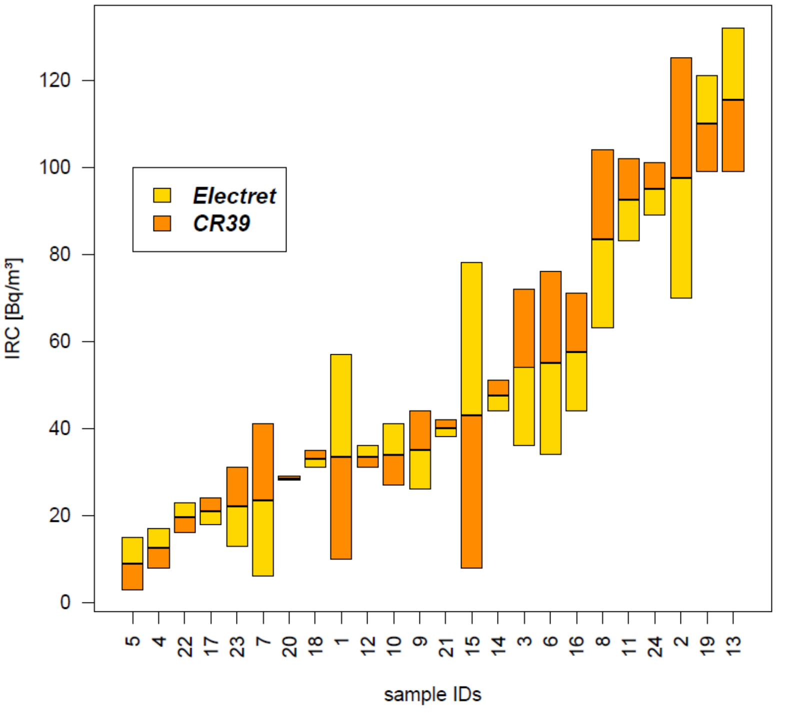

Figure 1: Indoor-Radon concentrations for the sample locations of the ILD measured with track etch (CR39) and electret. Sample IDs are ordered by increasing mean.

3.2 Different locations data set (DLD)

The different location data set (DLD) features over 80.000 samples of IRC carried out with three different measurement techniques: track etch, electret and activated charcoal. The samples were taken all over Austria and collected during two different measurements campaigns: The ONRAP - survey (Friedmann, 2005) and the recently carried out MAP – survey (Gruber et. al, 2019). The underlying sampling strategy was not intentionally designed for the purpose of our analysis and the conditions of sampling considerable differ from the ILD. Most prominently the samples of the DLD have not been taken on the same location for different measurement techniques, but on different locations. This already limits the comparability of the distributions of the different measurement techniques. Furthermore also the sampling design of the ONRAP- and MAP – survey are not fully comparable and major differences can be found in Table 1. These differences in the setup are as well crucial for our task of comparison and merging the distributions of different measurement techniques and need to be taken into account for our analysis. For example, the different goal of representativeness for the surveys of the DLD already takes samples from different populations of IRC. Therefore, in our analysis we investigate how this influences the outcome of our analysis.

Table 1: Comparison of ONRP - and MAP – survey of the DLD. T: track etch, E: electret, A: activated charcoal.

	ONRAP	MAP
Total number of Samples	~ 25.000	~ 56.000
Devices (number of samples)	T (4.400) E (3.200) A (17.400)	T
Exposure time	days to months	~ six months
Representativeness	population	spatial

Sampled area	regional for different devices	whole area, one device
Preferred indoor sampling location	no	ground floor

Another issue of the ONRAP - survey sampling design is, that the used sample devices are not equally distributed over Austria, but certain methods are more prominently used for certain regions. Figure **2** portraits the sample locations of the different measurement devices and shows that the ONRAP-survey uses different measurement techniques for different regions. This already points out the necessity to further regionalize the data set, because different regions have a different geogenic radon potential and samples from different regions have different populations even if measured exactly the same way (e.g. same device, exposure time).

One of our main goals is to compare the results of different exposure durations for different measurement devices. The DLD also features a variety of exposure times (Figure **3**), which ranges from a few days for activated charcoal up to months for track etch. Note that compared to the ILD where track etch and electret measurements feature different exposure times this is not true for the ONRAP samples, where both have exposure times around 100 days. Therefore, a comparison of track etch and electret investigates different measurement techniques and not different exposure times.

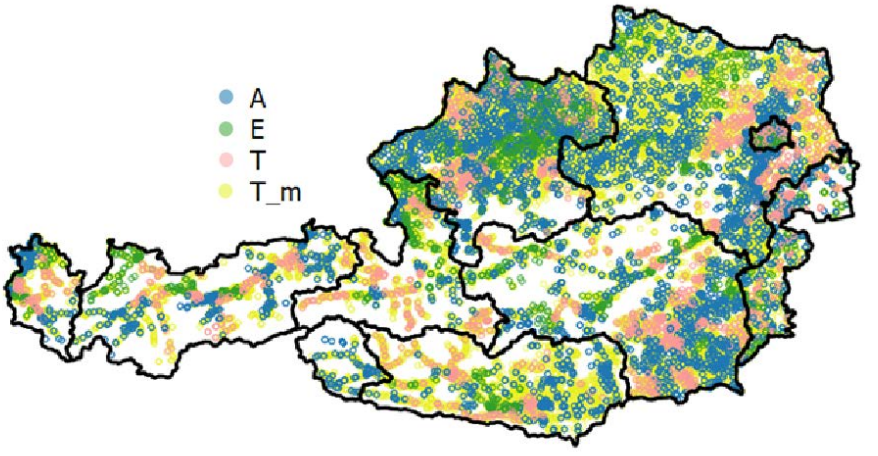

Figure 2: Sample locations for different measurement techniques in the DLD over Austria. T_m: track etch MAP-Survey, T: track etch ONRAP, E: electret ONRAP, A: activated charcoal ONRAP.

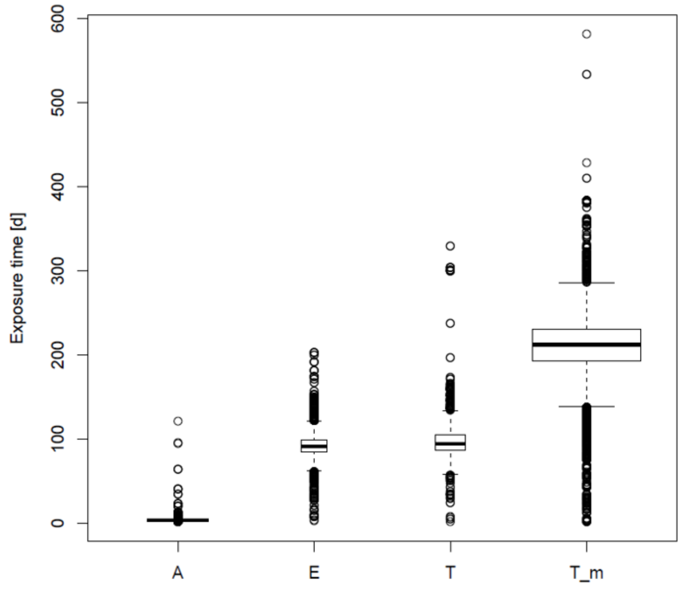

Figure 3: Comparison of exposure times of the DLD data set. T_m: track etch MAP-Survey, T: track etch ONRAP, E: electret ONRAP, A: activated charcoal ONRAP.

4. Methods

We use basic statistical visualizations of the different distributions and follow an approach suggested by Cinelli et al. (2009) to infer if populations of two different radon surveys can be merged. In the study Cinelli et al. (2009) examined short-term (activated charcoal) and long-term (track etch) data sets of IRC in the South of Belgium and found that both IRC distributions are well described by a log-normal distribution. The datasets subsampled geographically by geological units and found that these subsamples can be merged for radon risk mapping.

We follow the outlined analysis in Cinelli et al. (2009) and apply a t-test on the log-transformed distributions, where the p-value serves as proxy if two populations generally can be merged. Thereby a p-value above 0.05 suggests that merging the populations is a valid strategy, a p-value below 0.05 suggests the opposite. Further we investigate the effect of the measurement techniques on radon risk predictions. Therefore, we investigate if commonly used simple aggregations of IRC to predict radon risk show similar results for different measurement methods. As aggregations we use the arithmetic mean, median, geometric mean and the relative amount of samples exceeding different IRC thresholds.

For the ILD we use all samples for our investigation and for the DLD we further regionalize the samples as already indicated above, because of the clustered spatial sample distribution of the ONRAP-survey (Figure 2). Therefore, we subsampled the DLD by municipalities and only use municipalities in which at least two measurement techniques have 12 samples or higher. Doing so we reduce the risk of investigating subsamples with too low number of samples and follow a recommendation to use at least 12 measurements for predicting radon risk in municipalities in Austria outlined in Friedmann et al. (2005). Following this approach 478 subsamples were created in which at least two measurement techniques have a sample size of 12 or higher (Figure **4**). Beside the necessity of this regionalization because of the usage of different measurement techniques for different regions in the ONRAP – survey, another benefit of regionalized subsamples is that samples closer to each other are more similar and have a higher chance to come from the same population (e.g. same geogenic radon potential).

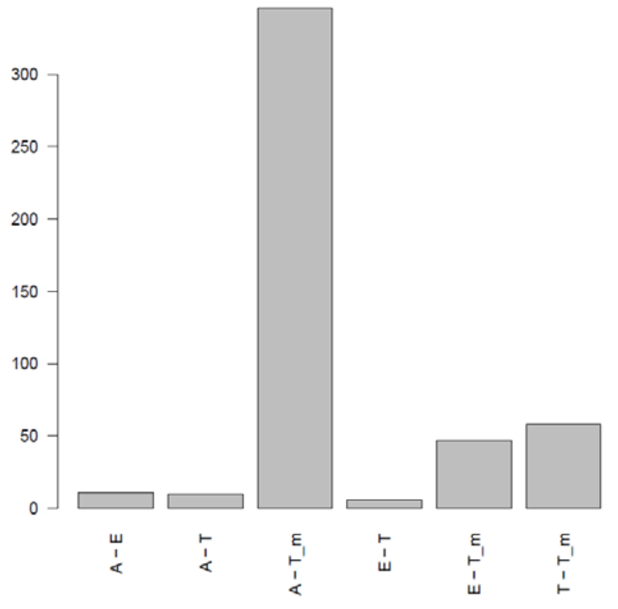

Figure 4: Number of regionalized subsamples of the DLD. T_m: track etch MAP-Survey, T: track etch ONRAP, E: electret ONRAP, A: activated charcoal ONRAP.

5. Results

We start our analysis with the ILD and then move on to the DLD where we investigate the impact of different sample survey designs.

5.1 Identical Location data (ILD) results

In the previous section we already noted that the IRC of the ILD at the same location measured with different devices can show high deviations. Nevertheless, in our analysis we are interested if multiple measurements of different measurement devices show similar distributions and can be merged for further predictions. The relative frequencies of the log-scaled distributions of the two measurement devices indicate a similar distribution (Figure **5**). This is supported by a p-value of 0.45 for t-test following the considerations in Cinelli et al. (2009) with already indicates that the two distributions can be merged for further predictions.

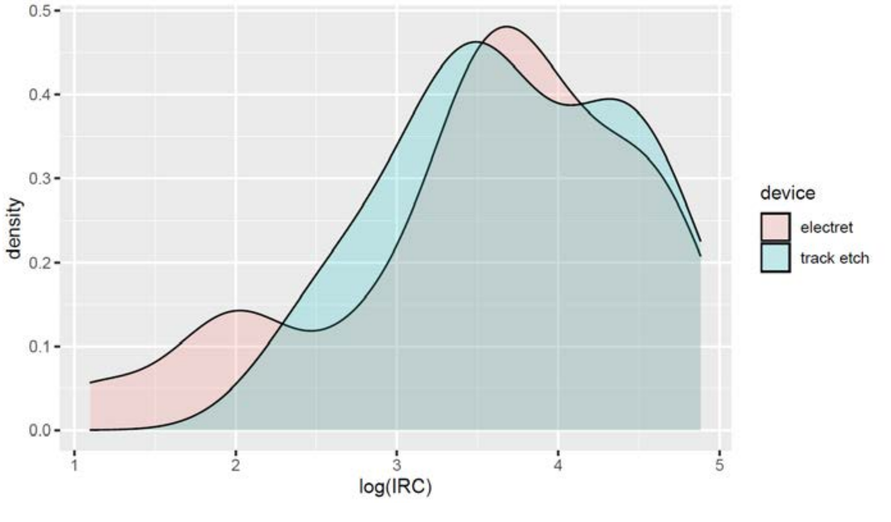

Figure 5: Relative frequencies of the log scaled distributions of the ILD. IRC in Bq/m³.

Motivated by this results we further investigate if the two distributions of the ILD show similar radon risk predictions when using appropriate aggregates (Figure **6**). The median, arithmetic and geometric mean show only minor deviations for the two distributions and only for special cases a prediction would be different. One of this cases would be if we choose the arithmetic mean to predict the radon risk and choose a threshold of 50 Bq/m³. But this is an arbitrary definition of radon risk and even if chosen both measurement techniques still produce arithmetic means of IRC around 50 Bq/m³. A look at the relative number of samples exceeding an IRC threshold also strengthens our hypothesis of merging the two distributions for radon risk classification. For example, if we take a threshold of 100 Bq/m³ and define radon risk if the relative number of samples exceeds 10 % than both measurement techniques give the same result. This is also true for many other combination of relative number of samples of thresholds and again only in very special cases they might show different classifications.

Summarized the distributions and the p-value of the t-test of 0.45 indicate that merging these for further tasks is a valid strategy and only for special arbitrary cases track etch and electret would produce different outcomes when predicting radon risk using aggregates. But the ILD is the case study, where the condition during measurements beside the exposure times are identically which is not true for the DLD in the following section.

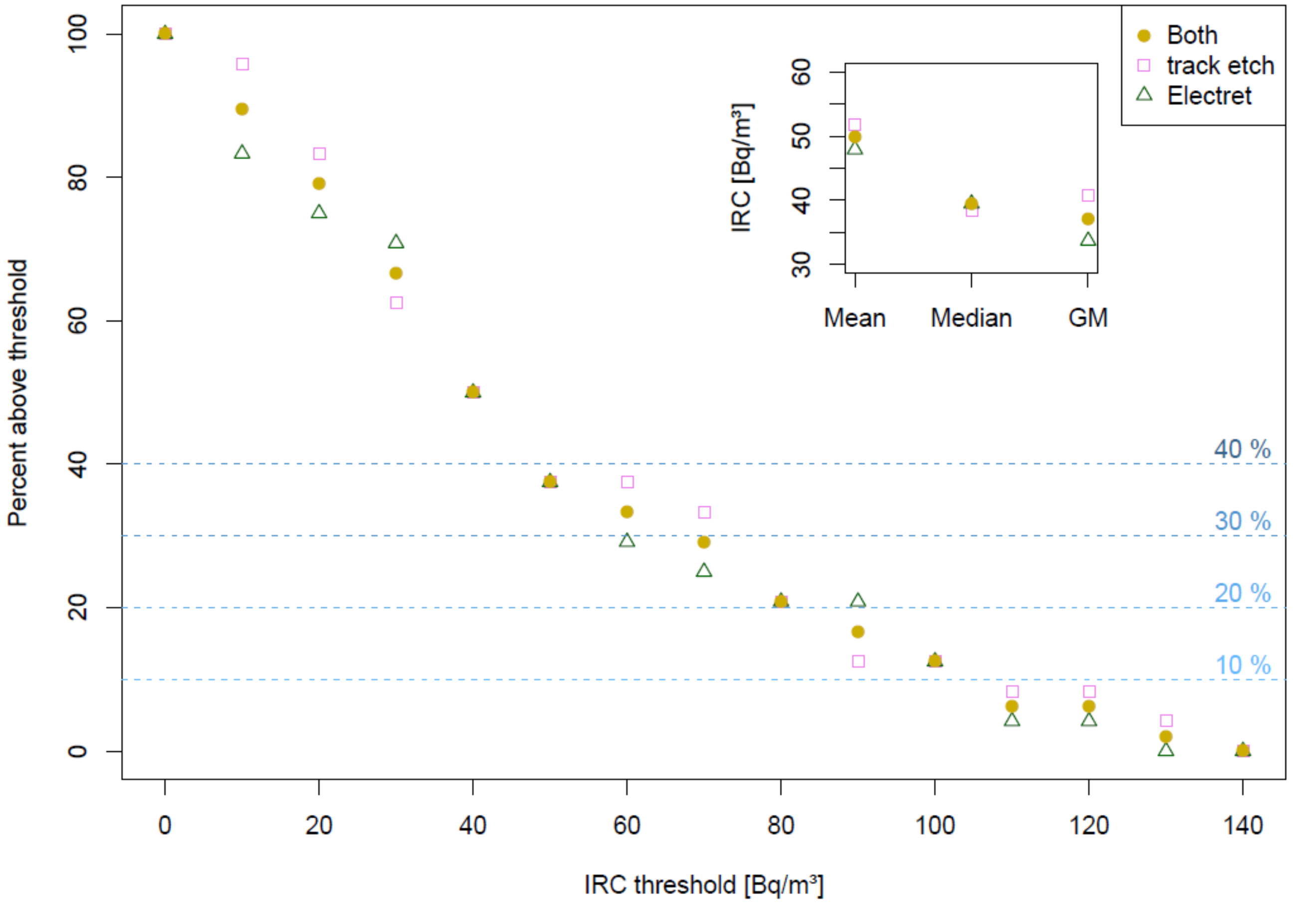

Figure 6: Aggregates of the distributions of the ILD that are commonly used for radon risk prediction. Percentages above threshold show the proportion of samples that are higher as the threshold for electret, track etch and both measurement devices. Upper right corner: arithmetic mean, median, geometric mean for electret, track etch and both measurement techniques.

5.2 Different location data results

The DLD features 478 subsamples, where the number of samples for at least two measurement techniques are equal or greater than 12 within one municipality in Austria. For these subsamples basic aggregates of the IRC were calculated and as in the previous section the p-value of the t-test for the log-transformed IRC.

Overall 179 (37.4 %) of the subsamples show p-values higher 0.05 and the percentage of p-values higher 0.05 for the different measurement combinations varies from 30 to 45 percent (Figure 7). The track etch measurements of the two different surveys (ONRAP, MAP - survey) show the highest percentage of p-values greater 0.05. Of course hereby identical measurement techniques are compared and they only differ in the survey designs. For the DLD this suggests that identical measurement techniques from different surveys are more similar than different measurement techniques within the ONRAP - survey. Electret compared to track etch for both the ONRAP- and the MAP – survey and activated charcoal show similar percentages. For some measurement combinations (A - E, A - T, E - T) the number of subsamples is low (Figure 4) and therefore these cases should be interpreted with caution.

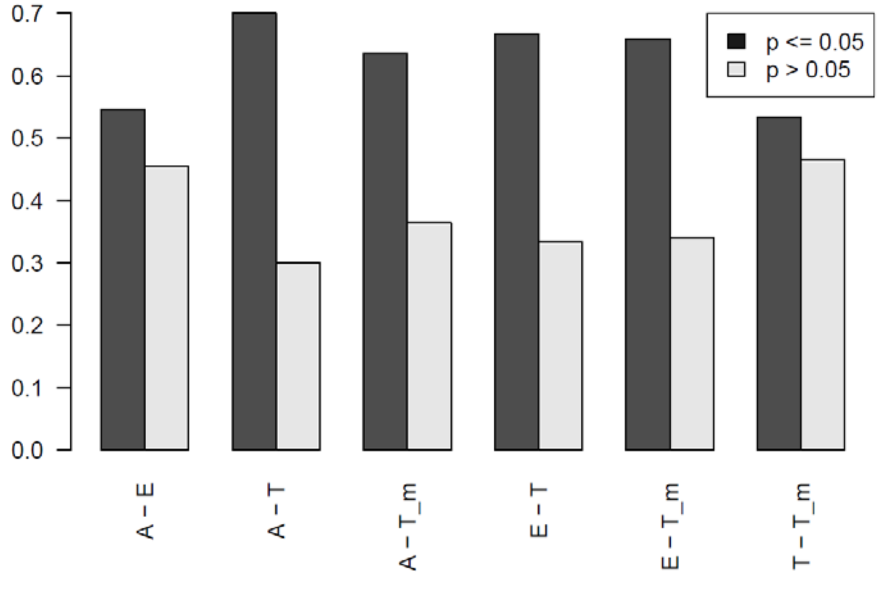

Figure 7: Relative proportions of different measurement combinations of the DLD showing a p-value above or below 0.05. T_m: track etch MAP-Survey, T: track etch ONRAP, E: electret ONRAP, A: activated charcoal ONRAP.

As outlined in the methods section we want to use the p-value as a proxy for similarity of two IRC distributions. To verify this hypothesis we calculated the absolute differences of some aggregates for IRC of the subsamples and discriminated them by their p-value (Figure 8). The arithmetic and geometric mean show that the absolute differences for cases where the p-value is greater than 0.05 are considerable lower as for p-values smaller than 0.05. For the geometric mean this is not surprisingly as the t-test was performed in the log-space. More surprisingly this is not true for the differences in percentage over 300 Bq/m³ and might indicate that the arithmetic and geometric mean are more robust aggregates to compare IRC distributions. Overall the absolute differences of aggregates support the hypothesis to use a p-value of 0.05 as a proxy for similarity of two populations.

In the following we compare the absolute IRC differences of different measurement techniques for cases with p-values above 0.05 (Figure 8). As in Figure 7 the differences of the geometric means are lower compared to the arithmetic mean for any measurement combination. All three aggregates show a similar pattern of differences for the measurement combinations, where again the comparison of track etch for the two surveys has the lowest absolute difference. As mentioned above the number of samples of some combinations (A - E, A - T, E - T) is very low, but beside some deviations of these low sample cases no major deviations for a certain measurement combination can be observed. One exception might be the comparison E – T_m for the percent over 300 Bq/m³ aggregate where the differences are much higher compared to the other measurement combinations.

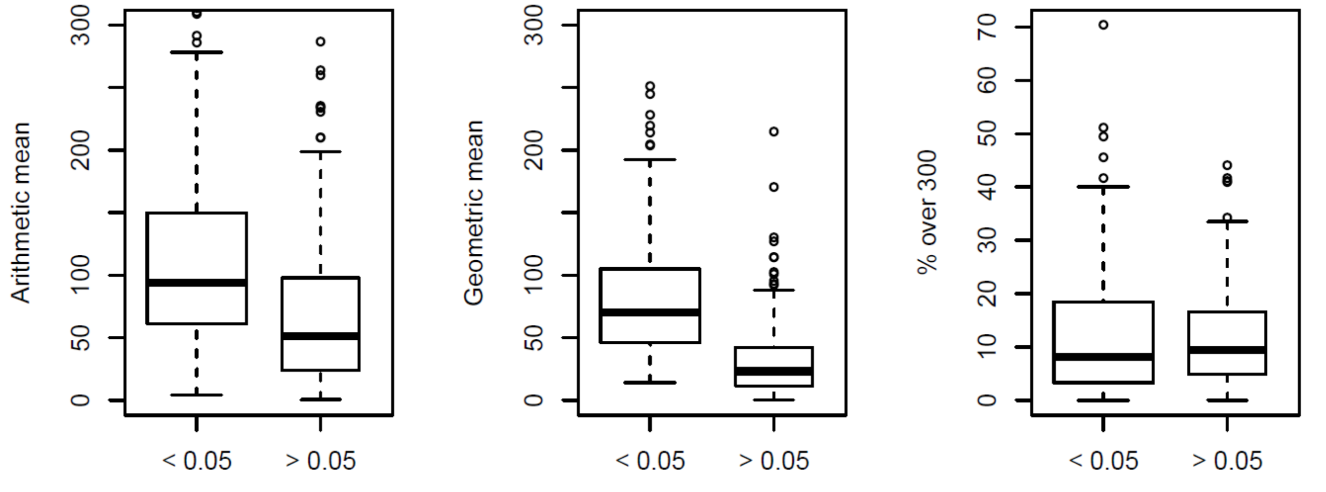

Figure 8: Boxplots of absolute differences [in Bq/m³] of aggregates for IRC of the subsamples, discriminated by their p-value

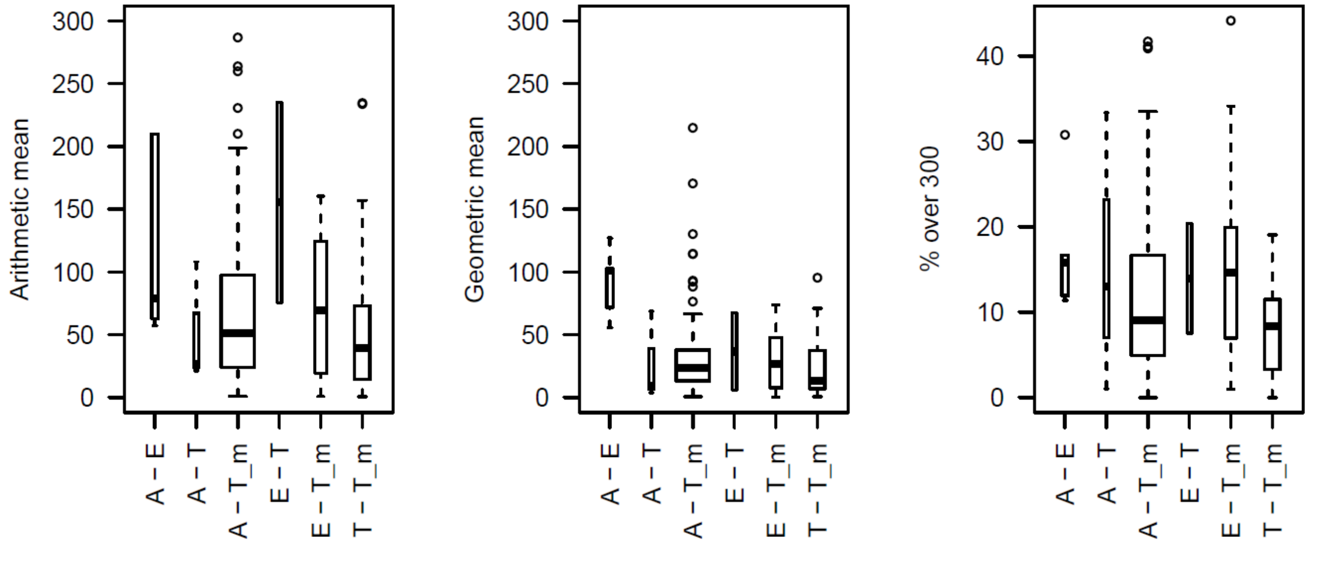

Figure 9: Boxplots of absolute differences [in Bq/m³] of aggregates for measurement pairs which show a p-value above 0.05. Width of the boxplots relative to number of samples in distribution. T_m: track etch MAP-Survey, T: track etch ONRAP, E: electret ONRAP, A: activated charcoal ONRAP.

For the ILD we examined if two measurement techniques would predict similar radon risk for basic aggregates of the IRC. For the DLD we follow the same approach but due to the high number of subsamples we compare the different measurements combinations of the DLD only for discrete aggregate predictors, which are also commonly used as basic concepts for radon risk definition: Arithmetic or geometric mean greater 100 Bq/m³, Arithmetic or geometric mean greater 300 Bq/m³, Percent of samples greater 300 Bq/m³ above 10 or 20 %. Figure 10 summarizes the results for the different measurement combinations and illustrates the percentage of cases where the two measurement techniques predict the same radon risk (both below or both above threshold) and where they differ. Again we need to consider the relative low sample size for the combinations A - E, A - T and E – T. The percent of different predictions for the combinations of measurement techniques of the DLD ranges from 0 to 100 depending on the used aggregates for predictions and combination of measurement techniques. Interestingly at least once for the chosen radon risk prediction scenarios the predictions of the two measurement techniques are different in more than 20 percent for all measurement combinations. Of course such a high percentage of different predictions could also be caused by arbitrary choices of prediction scenarios but already the most basic concept (arithmetic mean greater 300 Bq/m²) show different classifications close to 20 percent or above. This indicates that for the measurement combinations of the DLD in general the difference in prediction for two measurement techniques can be high and should be considered if merging these distributions.

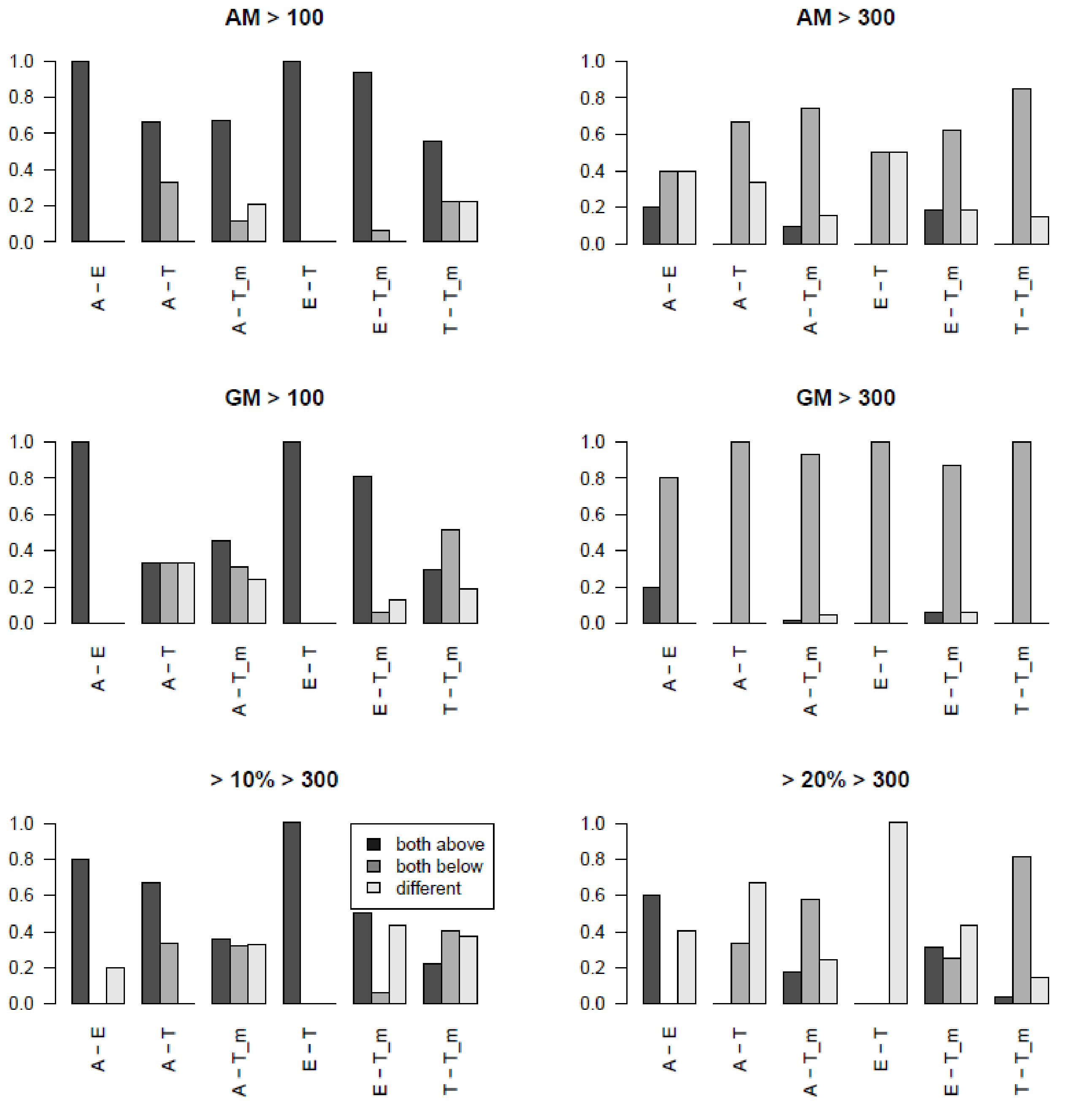

Figure 10: Relative proportions of classifications for two measurement techniques of the DLD for basic radon priority definitions. T_m: track etch MAP-Survey, T: track etch ONRAP, E: electret ONRAP, A: activated charcoal ONRAP.

6. Discussion

We analyzed different IRC data sets featuring different measurement techniques and investigated if distributions of different measurement techniques at the same location or area show similarities. Therefore, we used the p-value of the t-test for the log-transformed distribution as proxy for similarity (Cinelli et al., 2009) and analyzed the effect of the different measurement techniques on radon risk predictions. We found for the ILD that the IRC of the two measurement techniques (short term: electret, long term: track etch) show very similar distributions and a p-value of 0.45. Also the possible prediction scenarios for radon risk would perform identical prediction for the ILD for relevant cases (Figure 6). Therefore, for the ILD we suggest that the two distributions can be merged only to predict radon risk.

For the DLD we found a wide variety of p-values for the regionalized subsets. Overall 37.4 % of the regionalized subsamples showed p-values greater 0.05 for the different measurement combinations of the DLD (Figure 7, range: 30 to 45 %). In Cinelli at al. (2009) about 75 % geological subsamples of two radon surveys (short term: activated charcoal, long term: track etch) showed p-values greater 0.05. For the DLD for this measurement combination – A-T or A-T_m – only around 30 % or 38 % of the subsamples show p-values greater 0.05. One of the main reasons for this deviation might be the different subsampling techniques, where geological subsampling shows higher similarities as pure regional subsampling performed for the DLD due to clustered spatial distribution of the measurement techniques of the ONRAP – survey. But also other factors could influence the outcome, as different means or variabilities of IRC of the two sampled regions.

Summarizing, the majority of the subsamples of the DLD in contrast to previous findings show p-values lower 0.05 and therefore the distributions of the different measurement techniques should generally not be merged. Still the subsamples of the DLD with p-values greater 0.05 show less differences in aggregated IRC as the subsamples with p-values lower 0.05 (Figure 8). The performed radon risk prediction scenarios also suggest that the incorporation of distributions from different measurement techniques is not as straight forward for the DLD as for the other data sets. Already the most basic radon prediction scenario (AM > 300 Bq/m³) shows different predictions for the measurement combinations of the DLD close to or above 20 percent.

This divergent outcome of the different data sets may have various reasons. For example, different sampling designs define from which IRC populations samples are taken and therefore surveys with different designs even with the identical measurement devices show different distributions. This also shows the DLD when only comparing track etch for the ONRAP and Map – survey (T and T_m), where only around 45 % of subsamples show p-values greater 0.05 (Figure 7). Another factor which needs to be taken into account when comparing different surveys is the possible seasonal or annual variations of IRC that influence the distributions even if sample locations are identical or in the same area.

7. Conclusion

For our analysis of the comparability of different measurement devices with different exposure times we used two different data sets and the approach from Cinelli et al. (2009). The ILD featured IRC of identical sample locations carried out with different measurement techniques (short-term: electret, long-term: track etch) and we found that for the ILD the distributions of the used measurement devices can be merged to predict radon risk. For the DLD our analysis does not support this hypothesis and only for a minority of subsamples merging the distribution of two measurement devices can be performed. This is in contrast to previous findings in Cinelli et al. (2009), but one reason of the lower similarity of distributions of the DLD is the different approach of taking regionalized subsamples, where in the study of Cinelli et al (2009) also geological information was taken into account.

Generally, our analysis does not support the out-of-the-box merging of IRC distributions of different measurement techniques to predict radon risk, more so if they come from surveys with different survey strategies. Of course merging distributions might still be a valid strategy if further analysis, as performed in this study, show similarity of these distributions.

To extend the knowledge of the comparability of measurement methods, it would be good to have more data sets as the ILD available for future analysis. Therefore, institutions should be encouraged to perform studies where different measurement devices at identical locations measure IRC.

8. References

Cinelli, G., Tondeur, F. and Dehandschutter, B., 2009: Statistical analysis of indoor radon data for the Walloon region (Belgium),Radiation Effects and Defects in Solids,164:5,307 — 312.

European Union (2013). 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, 90/641/Euratom, 96/29/Euratom, 97/43/Euratom and 2003/122/Euratom. Official Journal of the European Union, OJ L13, 17.01.2014, pp. 1 – 73. https://eur-lex.europa.eu/legal-content/GA/TXT/?uri=OJ:L:2014:013:TOC

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

Gruber, V., Baumann, S., Wurm, G., Ringer, W. (2019): The new Austrian indoor radon survey – objectives, methodology, challenges and results, 9th Conference on Protection against Radon at Home and at Work, Prague, 16-20 September 2019

MetroRADON, 2020, http://metroradon.eu/

Annex 4: Report on activities A3.4.3 and A3.4.4

16ENV10 MetroRADON

Report on Activities 3.4.3 and 3.4.4

Lead organization: BfS

Authors:

Peter Bossew (BfS)

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

A3.4.3:

BfS, JRC and UC will use the results from Tasks 3.1 (A3.1.1-A3.1.5) and Task 3.3 (A3.3.1-A3.3.4) to propose instances where top-down harmonisation of indoor radon data and procedures may be reasonable. If applicable, practical cases will be demonstrated.

A3.4.4

BfS and UC will use the results of Tasks 3.2 (A3.2.1-A3.2.6) and Task 3.3 (A3.3.1-A3.3.4) to propose instances where top-down harmonisation for quantities related to geogenic radon may be reasonable. If applicable, practical cases will be demonstrated.

Activity A3.4.3 and A3.4.4 focused on proposing instances where top-down harmonisation for quantities related to indoor radon and geogenic radon may be reasonable.

Collection of radon data is the first step in the development of any kind of radon maps. Several kinds of radon maps can be created that display for example the simple arithmetic mean, % above reference level, radon potential, radon priority map, radon hazard index map. Therefore, the input radon data will strongly influence the output map. Radon data could be: indoor radon, soil gas radon, geogenic radon data (uranium concentration in soil and rock, terrestrial gamma dose rate, soil permeability, radon in water).

Difference and comparability of radon data (both indoor radon and geogenic) and their treatment have been analysed thanks to the results of the WP3 questionnaires on indoor and geogenic radon surveys. The results are available in the deliverable D3 of the MetroRADON project and have been presented during the European Radon Week (Vienna, 2020).

In the activity A3.4.2, a practical case for harmonizing indoor radon data has been demonstrated (see above and Annex 3).

Our conclusion is that harmonization of radon data is partly possible and projects like MetroRADON and intercomparison exercises are needed to study the comparability of radon data and develop procedures for harmonization. However, it is not an easy task and further studies of comparability and harmonization should be performed in the future. Below, some need for further research is indicated.

3.4.3 Top-down harmonization of indoor radon data

3.4.3.0 Bottom-up and top-down harmonization

Homogeneous results between surveys can be achieved in two ways (Figure 1). Firstly, all surveys are performed methodically identically. Then homogeneity is ensured by construction; this approach is called bottom-up, but viable only if surveys are planned jointly from scratch.

Realistically, surveys are conducted independently, following (for IRC surveys) different national Rn policies and possibly constrained differently. To compare results, one has no choice but normalize results to a common standard using models based on physical and statistical knowledge of the procedures which in each case have led to the reported results. This is called top-down or posterior harmonization.

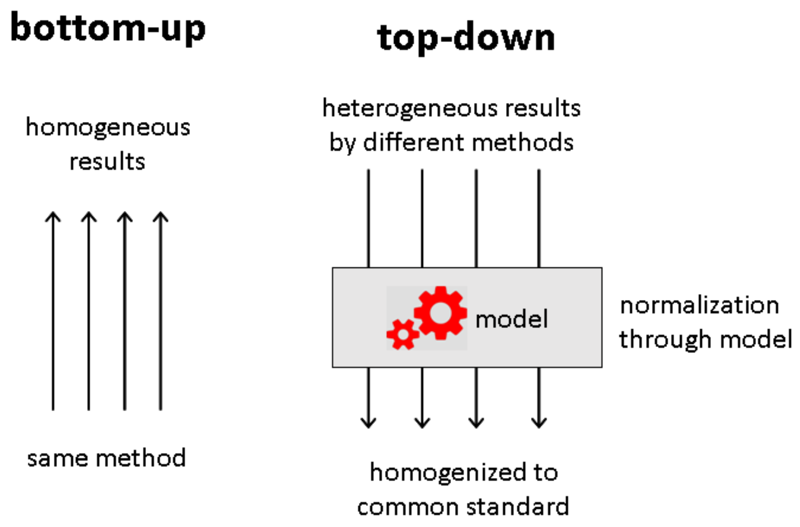

Figure 1: Bottom-up and top-down harmonization

3.4.3.1 Heterogeneity of indoor radon data

As has been shown in sections 3.1 and 3.3, different methods are used for IRC (indoor Rn concentration) assessment. The differences pertain to several levels, to be summarized in the following:

a. Survey design
b. Measurement method
c. Data aggregation

In all cases, heterogeneity causes problems in comparative data interpretation. Even if results are nominally given in Bq/m³ or a probability to exceed a reference level, they may result from a different procedure and be therefore different.

a) Survey design

IRC surveys can be based on different logic. This leads to heterogeneity in these issues:

- Selection of objects to be measured: dwellings / workplaces, e.g. schools
- Representativeness of the selection: selection criteria can be very different, from scattered voluntaries to sophisticated schemes aimed to ensure demographic representativeness. The issue is discussed in Pantelić et al. (2019) (Annex D3/2). One problem identified is that in the reports available about some European surveys, information about representativeness is scarce or missing.

- Specification of objects: All dwellings / ground floor dwellings / only buildings with basement / etc. In case of existing datasets, the specification can be carried out by filtering, i.e. only measure in rooms or building which meet the specification, or standardization to the specification through modelling. By filtering one looses data which may impair statistical quality, while by modelling, one introduces modelling uncertainty (Figure 2). For IRC surveys, when demographic representativeness is targeted, one will attempt representativeness of the sample in order to have neither to filter nor to model. This is called design-based approach. However, achieving representativeness is perhaps the most difficult task in any survey (not only radon).
- Geographical data density is very different between European countries, apart from the difference owing to different population densities: i.e., the *sampling rate* is different. Therefore, the achievable resolution of maps is different across borders (but sometimes also between regions of one country, example Germany; not only because of variable population density but because of regionally different radon policies). This can be remediated only by model-based interpolation, as shown by Elío et al. (2019).
- Selection of measurement period and duration: Short- and midterm measurements (charcoal, electret) are rare in Europe, but also long-term TE measurements can be defined differently: from 1 year to 3 months during winter or during the transition period. The CD/DVD method can even yield multi-year measurement periods.

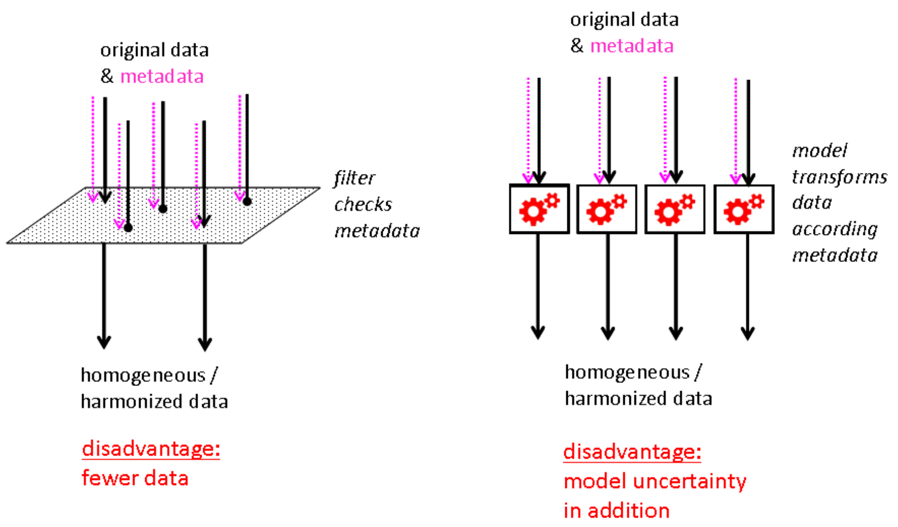

Figure 2: Harmonization by filtering and modelling.

The measurement period is relevant when measurement duration is below the length of anticipated cycles of true variability of IRC. This because in most cases (not always though) IRC is higher in winter than in summer, so measuring 3 months in winter would yield higher results than doing so in summer. For few-months measurements, the transition periods spring or autumn are therefore recommended; however, winter measurements may be chosen to get conservative results.

The IRC cycles are usually daily and annual ones. For workplaces, sometimes weekly cycles are observed related to working pattern. A spectral plot of a multi-year IRC time series from a workplace (a laboratory building in Japan, graph taken from Janik et al. 2018) is shown in Figure 3.

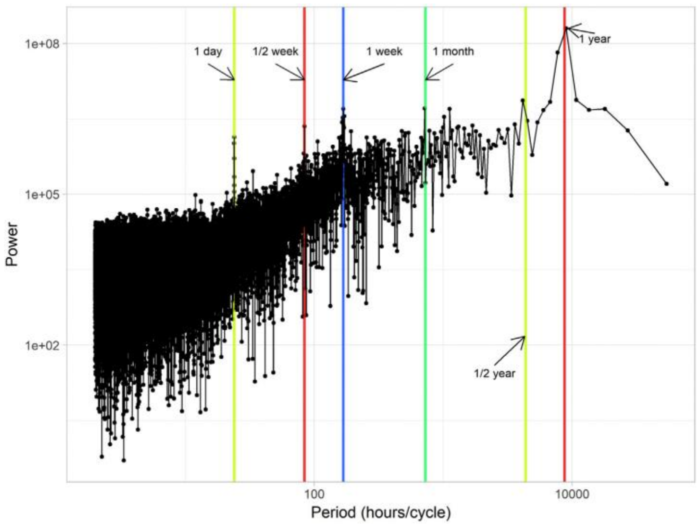

Figure 3: Fourier spectrum of IRC from a workplace (taken from Janik et al. 2018)

Apart from regular cycles (regularity referring to their length, whereas amplitude can be varying), there are irregular patterns in IRC time series. Physical reasons are weather episodes or disruptions of usual occupancy or usage patterns.

IRC time series may also be subject to trends. Physical causes may be climatic changes, restructuring of a building (e.g. upgrading of thermal insulation), or change in usage (different occupants). Note that statistics such as mean or dispersion become meaningless in the presence of trend. Whether there is an IRC trend due to climatic change, is currently a matter of debate and not yet resolved. On the other hand it has been shown that increasing thermal insulation for the sake of energy saving can lead to increased IRC (e.g., Collignan et al. 2016). How such effects should or could be included into IRC mapping has not yet been discussed in the Rn community.

b) Measurement method

It seems that these days, metrological heterogeneity proper is a minor problem. QA infrastructure is well developed in Europe and usually one can rely on metrological correctness (accuracy and precision) of IRC results within reasonable tolerance. See WPs 1, 2 and 5 for metrological QA.

One difficult issue is sensitivity to thoron, which is different between measurement devices. If contribution of Tn must be anticipated, it is still impossible to correct for, because Tn concentration is generally not known. However, this seems to be a minor problem compared to others addressed here.

c) Data aggregation

IRC data are often aggregated into geographical units such as grid cells (European IRC map), municipalities or districts. Statistics given can be arithmetic or other means or probabilities to exceed a reference level. Comparison of such data is not trivial and may require modelling.

As an example, under the *assumption* (!) that IRC in an area is log-normally distributed,

$$AM = GM \exp(SDL^2/2),$$

$$prob\ (IRC>RL)= \Phi[(AML- \ln(RL)/SDL],$$

where AML = AM(ln IRC) and SDL = SD(ln IRC).

Assuming (!) dispersion quantified as GSD (geometric standard deviation) = 2, which is reasonable for areas of characteristic linear sizes of a few to a few 10 km, hence SDL=ln(2), one finds

$$AM = GM \times 1.27,$$

for AM=300 and RL=300 Bq/m²,

$$prob(IRC>RL)=0.364,$$

that is, if the mean over an area equals 300 Bq/m³, the probability to exceed a RL of same size equals 36.4%.

Note that for making results comparable, *two assumptions were necessary* in this case, namely log-normality and known GSD. In real cases, both assumptions are fulfilled approximately, at best, leading to uncertainty when comparing values reported as aggregated differently.

A particularly complicated matter is harmonization of classified IRC, typically the RPA status (radon priority area) of an area. This is because by classification, most of statistical information about a dataset is lost. It is impossible to "back-model" a binary classification, such as RPA yes or no, into an IRC mean or exceedance probability.

3.4.3.2 Summary

Top-down homogenization is feasible. Physical knowledge and statistical tools are at hand, although in a few cases, further studies are recommended. This mainly concerns the following aspects:

- *Temporal standardization:* how to normalize IRC measurements of different duration and performed in different period of the year to an abstract long-term mean - in the framework of Metro Rn, first studies are shown in section 3.4.1.
- *Deviation from representativeness*. This is a complicated but important subject which should be studied in more depth.
- *Dwellings - workplaces.* First results show (section 4.1.1) that in general, workplaces have different Rn characteristics than dwellings. Intense work on the subject is under way, because it affects implementation of Rn regulation directly.
- *Harmonisation of RPA.* This problem could not be solved satisfyingly within Metro Rn and should be addressed on European level.

3.4.4 Top-down harmonization of data related to geogenic radon

In this section, we shall restrict discussion to the following quantities:

- Radon concentration in soil air (SRC);
- Ground permeability (k);
- Geogenic radon potential (GRP);
- Ambient dose rate (ADR).

For Rn exhalation from the surface and Rn concentration in water, information acquired during Metro Rn does not seem sufficient for further discussion. Concerning geochemical concentrations (U, Th, K), we refer, among other, to the respective methods sections of the European Atlas of Natural Radiation (Cinelli et al. 2019, EC 2019) and Domingos (2020).

3.4.4.1 Heterogeneity of SRC data

Main heterogeneity issues are:

a. Representativeness: (i) site selection for measurement; (ii) Unclear, for which area a measurement stands.

b. Temporal variability

c. Sampling and measurement method

a) Representativeness

SRC measurements may serve for site characterization, usually to decide about construction standards (whether particular provisions against Rn infiltration are required), or as data points for generating soil Rn maps. In the first case, representativeness seems to be a minor issue and protocols have been developed for this purpose (perhaps most advanced in the Czech Republic, www.radon.eu/ca.html).

For mapping, SRC measurements (the same applies for permeability and GRP) are usually located to represent variability of the geogenic factors that control geogenic Rn; mostly geological maps are used for orientation, where to locate SRC measurements. This is different from IRC, where demographic representativeness is mostly targeted.

A critical issue here - not discussed in the framework of Metro Radon - may be that geological maps may not be adequate proxies of geogenic Rn in urban environments, or ones with strongly anthropogenically altered surface geology. Additionally, SRC samples are mostly taken in the free country, not on altered ground.

So far, however, no experiences indicate heterogeneity problems of SRC or GRP maps; perhaps owed to the fact that the said factor has not yet been investigated systematically.

b) Temporal variability

SRC is measured with temporal spot measurements (grab sampling), while SRC is subject to temporal variability controlled by soil humidity, ground water level, surface soil freezing and other factors. The variability depends on sampling depth, soil type and sampling protocols which indicate situations in which it is advised against sampling, in order to avoid generating unrepresentative data.

Practitioners have collected experience with the problem and also time resolved SRC data are available, but it seems that no synthesis has been achieved, in the sense of a rule: which natural conditions (soil type, humidity, weather) and conditions related to sampling (sampling depth, air extraction technique, protocol in general) lead to which expected uncertainty of a result relative to an ideal standardized value. Uncertainty is here understood as product of the variability caused by these factors, while classical radiometric uncertainty contributes less in most cases.

c) Sampling and measurement methods

Sampling depth depends on sampling equipment, as does the volume of air extracted. In particular sampling depth influences the SRC as fraction of the theoretical equilibrium SRC. This fraction is difficult to know without more detailed investigation.

Another problem is the usage of buried Rn detectors (TE, CD-type and other) instead of grab sampling. These methods are not commonly used so far, but this may change. Continuous measurement is performed in some instances, but not for routine purposes because of the comparatively high effort. (Mostly the purpose is geophysical research.)

Surveys conducted with different methods may lead to a bias against each other which is difficult to remove. To our knowledge, such difficulty has not yet been observed between European countries - perhaps because it has not been investigated systematically. However one discrepant case is known, namely in one region of Germany where grab sampling measurements were performed together with ones based on buried detectors. The results (unpublished) proved incompatible, although no deeper analysis has yet been carried out.

Intercomparisons of SRC measurement by grab sampling between the Czech Republic, Germany and Poland showed good agreement of the results in spite of different protocols (Barnet et al. 2008; Bossew 2012).

International intercomparisons for SRC measurement are performed every two years in the Czech Republic (RIM exercise). Since SRC measurement requires practice and QA is difficult, everybody engaging in this activity is recommended to participate in this exercise (Neznal et al. 2001); some information at www.radon.eu/workshop2020/ .

3.4.4.2 Heterogeneity of permeability data

To wide extent the same as for SRC applies for permeability data. In addition, sometimes permeability is not measured but assessed from soil quality as permeability class.

We are not aware of systematic intercomparisons of permeability assessment methods.

3.4.4.3 Heterogeneity of GRP data

Again the same applies to GRP. An additional source of heterogeneity is GRP definition. As described in other sections, the GRP can be a result of measurement of SRC and permeability (typically the familiar Neznal-GRP), but also as a combination of controlling factors, leading to GRP classes.

No systematic parallel investigations exist to our knowledge which may allow to relate different conceptual definitions of the GRP. However, this is one reason, why the geogenic Rn hazard index (GRHI, section 4.3; Bossew et al. 2020) has been developed, which - at least one of its versions - is a bottom-up harmonization approach, thus avoiding top-down harmonization which does not appear feasible so far.

3.4.4.4 Heterogeneity of dose rate data

ADR is a weak predictor of the GRP; its terrestrial component a somewhat better one. Heterogeneity is mainly owed to measurement and evaluation technique. The reasons are differences in response of different measurement systems, related to internal background, cosmic response, measurement height above ground, site characteristic, and other. The physical mechanisms are known, but in many cases, the technical details of the measurement systems are not.

For the European ADR emergency warning network EURDEP, the subject has been extensively studied in the 2000s by the JRC in the AIRDOS project, but no viable rule for data harmonization has ever been achieved. See also the European Atlas of Natural Radiation (EC 2019) on this subject, and Cinelli et al. (2014) and Bossew et al. (2015a-c), where the possible usage of EURDEP stations for ADR mapping and its potential for RPA estimation is discussed.

3.4.4.5 Summary

Harmonization of SRC, permeability and GRP is more difficult than the one of IRC. The reason is that the natural conditions which determine the observed quantities, are normally poorly known. Therefore, for example, an SRC measurement in 1 m depth cannot easily be recalculated to one hypothetically performed in 70 cm depth, or v.v. Particularly difficult is normalization of GRP expressed as classes or categories.

A consequence was the development of the Geogenic Radon Hazard Index (GRHI), certain versions of which circumvent the harmonization problems; see WP 4.3.

Harmonization of SRC, permeability and GRP is mainly important for interpretation of maps of these quantities. We think that on European scale, generating a joint GRHI map, and possibly one or several (according to different objectives) GRPA (RPA derived from geogenic predictors) maps is more promising than trying to "stick together" individual maps resulting from different methodology.

Harmonization of ADR data is possible, in principle. The problem is the partial lack of metadata of the ADR monitoring stations which would be required to this end..

References

Barnet I., Pacherová P., Preusse W., Stec B. (2010): Cross-border radon index map 1:100 000 Lausitz – Jizera – Karkonosze – Region (northern part of the Bohemian Massif). J. Environm. Radioact. 101 (11), 809 - 812; doi:10.1016/j.jenvrad.2009.11.009

Bossew P. (2012): The influence of observation protocols on reported values of Rn in soil air. 11th international workshop on the geological aspects of radon risk mapping. Prague, 18 - 22 Sept. 2012. In: Barnet I., Neznal M. and Pacherová P. (eds.) Proc., 11th Intl. workshop on the geological aspects of radon risk mapping. Czech Geological Service and Radon v.o.s., Prague 2012; ISBN 978-80-7075-789-5.

Bossew P., Cinelli G., Cernohlawek N., Gruber V., Friedmann H., Dehandschutter B., Menneson F., Garcia-Talavera M., Da Silva, N., Tollefsen T., Nishev A., Hernandez-Ceballos M.A., de Cort M. (2015a): Can ambient dose rate be used for predicting the geogenic Rn potential? Presentation, Conference Radon in the Environment, Krakow, Poland, 26.-29.5.2015

Bossew P., Cinelli G., Tollefsen T., Gruber V. and De Cort M. (2015b): Estimation of the geogenic radon potential and radon prone areas from geochemical quantities and ambient dose rate. Presentation, ENVIRA 2015, Thessaloniki, Greece, 21.-25.9.2015.

Bossew P., Cinelli G., Hernández-Ceballos M., Cernohlawek N., Gruber V., Dehandschutter B., Menneson F., Bleher M., Stöhlker U., Hellmann I., Weiler F., Tollefsen T., Tognoli P.V., Nishev A., De Cort M. (2015c): The terrestrial ambient gamma dose rate. Presentation, IWEANR 2015, Verbania, Italy, 9.-13.11.2015.

Bossew P., Cinelli G. Ciotoli G., Crowley Q.G., De Cort M., Elío Medina J., Gruber V., Petermann E., Tollefsen T. (2020): Development of a Geogenic Radon Hazard Index—Concept, History, Experiences. Int. J. Environ. Res. Public Health 2020, 17(11), 4134; https://doi.org/10.3390/ijerph17114134

Cinelli G., Hernandez-Ceballos M.A., Tollefsen T., Bossew P., Sanchez I., Marin-Ferrer M., Nishev A., Bogucarskis K., Gruber, V., De Cort M. (2014): A method to estimate terrestrial gamma dose rate from EURDEP routine monitoring data to improve the European Geogenic Radon Map. Pres., 12th international workshop on the geological aspects of radon risk mapping (GARRM), 16.-19.9.2014 Prague. I. Barnet, M. Neznal and P. Pacherová (eds): 12th Intl. Workshop on the Geological Aspects of Radon Risk Mapping. Czech Geological Survey & RADON v.o.s., Prague 2014. http://www.radon.eu/workshop2014/pres.html

Cinelli, G.; Tollefsen, T.; Bossew, P.; Gruber, V.; Bogucarskis, K.; De Felice, L.; De Cort, M. (2019) Digital version of the European Atlas of natural radiation. J. Environ. Radioact. 196, 240–252, doi:10.1016/j.jenvrad.2018.02.008.

Collignan B., Le Ponner E., Mandin C. (2016): Relationships between indoor radon concentrations, thermal retrofit and dwelling characteristics. J. Environm. Radioactivity 165, 124 - 130; http://dx.doi.org/10.1016/j.jenvrad.2016.09.013

Domingos F., Cinelli G., Neves L., Pereira A., Braga R., Bossew P., Tollefsen T. (2020): Validation of a database of mean uranium, thorium and potassium concentrations in rock samples of Portuguese geological units, generated of literature data. J. Environmental Radioactivity, online https://doi.org/10.1016/j.jenvrad.2020.106338

EC (2019). European Commission, Joint Research Centre; Catalogue number KJ-02-19-425-EN-C, EUR 19425 EN; Cinelli, G., De Cort, M.; Tollefsen, T., Eds.; European Atlas of Natural Radiation, Publication Office of the European Union: Luxembourg, 2019; ISBN 978-92-76-08259-0, doi:10.2760/520053.

Janik M., Bossew P., Kurihara O. (2018): Machine learning methods as a tool to analyse incomplete or irregularly sampled radon time series data. Sci. Total Environment, https://doi.org/10.1016/j.scitotenv.2018.02.233

Neznal M., Neznal M., Matolin M. (2012): Soil-gas radon intercomparisons. In: Gutiérrez-Villanueva JL. et al. (eds): International Intercomparison Exercise on Natural Radiation Measurements under Filed Conditions, Saelices el Chico, Spain, May 2011; Santander : PUbliCan, Ediciones de la Universidad de Cantabria, D.L. 2012, ISBN 978-84-86116-64-4. www.editorial.unican.es/libro/international-intercomparison-exercise-natural-radiation-measurements-under-field-conditions ; p.40 ff.

Pantelić G., Čeliković I., Živanovića M., Vukanac I., Krneta Nikolića J., Cinelli G., Gruber V. (2019): Qualitative overview of indoor radon surveys in Europe. J. Environm. Radioact. 204, 163 - 174; doi.org/10.1016/j.jenvrad.2019.04.010