Journal of Environmental Radioactivity xxx (2016) 1–5

Contents lists available at ScienceDirect

Journal of Environmental Radioactivity

journal homepage: www.elsevier.com/locate/jenvrad

Bagged neural network model for prediction of the mean indoor radon concentration in the municipalities in Czech Republic

Jana Timkova [a,*], Ivana Fojtikova [a], Petra Pacherova [b]

[a] *National Radiation Protection Institute, Bartoskova 28, 140 00, Praha 4, Czech Republic*
[b] *Czech Geological Survey, Geologicka 6, 152 00, Praha 5, Czech Republic*

A R T I C L E I N F O

Article history:
Received 9 November 2015
Received in revised form
20 May 2016
Accepted 5 July 2016
Available online xxx

Keywords:
Indoor radon
Predictive mapping
Bagged neural network
Gamma dose rate
Lithology
Czech Republic

A B S T R A C T

The purpose of the study is to determine radon-prone areas in the Czech Republic based on the measurements of indoor radon concentration and independent predictors (rock type and permeability of the bedrock, gamma dose rate, GPS coordinates and the average age of family houses). The relationship between the mean observed indoor radon concentrations in monitored areas (~22% municipalities) and the independent predictors was modelled using a bagged neural network. Levels of mean indoor radon concentration in the unmonitored areas were predicted using the bagged neural network model fitted for the monitored areas. The propensity to increased indoor radon was determined by estimated probability of exceeding the action level of 300Bq/m^3.

1. Introduction

In the last years, international organisation involved in radiation protection and public health published new recommendations and regulations aiming at better protection from exposure to radon in homes and workplaces. One of them, the European Council Directive 2013/59/Euratom of 5 December 2013 on the basic safety standards for the protection against ionising radiation establishes inter alia the uniform basic safety standards for the health protection of individuals which are subject to occupational exposures. In the Article 54, new obligation is introduced, e.g. to regulate the exposition to radon and its progenies in workplaces situated indoors. This obligation is aimed at workplaces located on the ground or basement level within so called radon-prone areas.

This obligation constitutes a scientific challenge for Member States to delineate radon-prone areas on their territories for this purpose (Hulka (2008)). The definition given in the EC Directive 2013/59/Euratom, art. 54 (Radon-prone areas – areas where in a significant number of buildings the radon concentration is expected to exceed the relevant national reference level) gives sufficient space for appropriate concept creation. We realized to base it on the probability of exceeding national reference level in individual municipalities. When this probability is found to be high, significant number of workplaces will exceed the reference level.

We are aware that this approach is well in line with older definition given in ICRP 115/126 (Radon-prone area – a geographic area or an administrative region defined on the basis of surveys indicating a significantly higher radon concentration than in other parts of the country), but it fits quite well also with the new one.

The legislative basis for this process is in the competence of Member States. Methods used for specifying the radon-prone area are based generally on data available in the particular member State: i.e. indoor radon data or radon risk classification of foundation soils. Our approach was to base the delineation according to borders of an administrative unit, which simplifies the arrangement of measures required by the Directive.

Thus, the goal of the study was to develop a methodology for a differentiation of local administrative units according to the level of radon risk at workplaces located on the ground floor or in the basement. This methodology is based on collected data of soil and indoor radon measurements, on map sources and open data sources from the Czech Statistical Office.

In a great part of the municipalities only few or no measurements of indoor radon were collected. The average concentration of

* Corresponding author.
E-mail addresses: jana.timkova@suro.cz (J. Timkova), ivana.fojtikova@suro.cz (I. Fojtikova), petra.pacherova@geology.cz (P. Pacherova).

http://dx.doi.org/10.1016/j.jenvrad.2016.07.008
0265-931X/

indoor radon in a municipality is determined by many variables affecting the amount of radon in the ground and its entry into a house. The aim is to combine the information from indoor radon measurements in family houses with these influential variables and predict the distribution of the indoor radon concentration in insufficiently monitored areas. A similar work on the level of single houses was done by Kropat et al. (2015) (the paper also gives a detailed list of references to the various approaches and methods used in modelling and spatial interpolation of the indoor radon concentration).

The methodology of collection of the indoor radon concentration measurements in the Czech republic is described in Section 2. The statistical methods for determination of the radon-prone areas are explained in Section 3, and the obtained results are summarized in Section 4. The final section discusses the outcome of the study and used methodology.

2. Materials and methods

2.1. Indoor radon long-term measurement

Solid-state nuclear track detectors based on Kodak LR-115 foil are used for long-term radon measurements in the Czech Republic. Bare LR-115 which measures the equilibrium equivalent concentration (EEC) was used from 1985 to 2003. Results from the bare detectors depend on unknown factors (such as F). They have therefore been replaced by a system of track detectors enclosed in a diffusion chamber and the EEC were recalculated to the radon concentration using the mean equilibrium factor F. The new system is called RamaRn, and has been distributed since 2003. RamaRn detectors are traceable to PTB Braunschweig. More details on the calibration and detector parameters are given in the literature (Slezakova et al. (2013); Thinova and Burian (2008)). This measuring system has been used for long-term measurements organized by the radiation protection regulatory body for nearly 30 years. Two detectors per apartment are distributed for free among citizens interested in measurements. The national database currently contains more than 150 000 records for indoor radon measurement. For the purpose of this study, averages obtained for family houses have been used. The uncertainty of using dataset for family houses instead of a set for workplaces is expected to be acceptable, because the majority of measurements has been done in rooms on the ground level. Hence, we assume that our dataset has the same statistical characteristics as would be obtained in ordinary workplaces (libraries, offices etc.) on the ground floor.

3. Prediction of the geometric mean of indoor radon concentration

The main purpose of our work was to assign a level of risk of increased level of indoor radon concentration (denoted further as IRC) to the municipalities in the Czech Republic. We expressed the level of risk of an elevated level of the IRC by means of the radon index of a municipality (RIA, already used in previous work, Fojtikova et al. (2011)), defined as

$$\mathrm{RIA} = \log_{10}\left[P\left(\mathrm{Rn} > 300\ \mathrm{Bq}/\mathrm{m}^3\right) \times 100\right]. \tag{1}$$

It is shown that RIA is a function of the probability of the IRC exceeding the action level of 300 Bq/m^3. 10^{RIA} expresses the percentage of dwellings with IRC above the action level. In particular, RIA is equal to 1 when the probability of exceeding the action level equals 0.1, i.e. in 10% of dwellings the IRC is higher than 300 Bq/m^3.

To determine the value of RIA we need to estimate the probability of exceeding the action level. In accordance with findings in the literature (Nagda (1994); Porstendorfer (1987)) we suppose that the distribution of IRC measurements in each municipality is close to the log-normal distribution. The QQ-plots of our measurements within municipalities did not show any substantial discrepancies from the log-normal distribution. The unknown parameters of the log-normal distribution can be estimated from the observed data. The probability of exceeding the action level for each municipality is derived from the distribution function with estimated parameters plugged in. However, enough data is necessary to obtain reliable estimates of the unknown parameters. Moreover, there are areas where only few or no measurements at all were performed.

The solution is based on predicting the unknown parameters of the distribution of IRC in municipalities with no or insufficient data using the relationship between the IRC and the characteristics of every municipality. Section 3.1 gives an overview of the collected IRC measurements. In Section 3.2 we introduce the characteristics of the municipalities which serve as independent predictors in the model fitting. In Section 3.3 we explain the methods used to find the best model.

3.1. The indoor radon concentration data

A sufficient amount of IRC measurements was available for 1405 (about 22.3%) municipalities out of a total of 6298. The amount of data was considered sufficient if the measurements were taken in at least 20% of all family houses. The average number of measurements in a municipality with at least 20% was 60, the maximum value was 2339 in Trebic. The geometric means of the IRC in the municipalities with at least 20% houses measured ranged from 32 to 1174 Bq/m^3 with average equal to 215 Bq/m^3. The geometric standard deviation ranged from 1.16 to 5.65 with an average value of 1.96. In 3837 (60.9%) municipalities, some but not enough measurements were collected, while no measurements were available for 1056 (16.8%) municipalities. Hence, the geometric mean of the IRC needs to be predicted for cca 78% of municipalities.

Although we have the individual measurements of the dwellings at hand, the aim of the study is to determine the radon index for the municipalities (or higher administrative units). Hence, unlike the study of Kropat et al. (2015), where the aim was to predict the IRC on level of individual houses, here it is sufficient and more practical to work with geometric means and geometric standard deviations per municipalities. Another advantage of using the geometric mean of measurements is its lower variability in comparison to individual measurements.

3.2. The independent predictors

The following geological, radiometric, spatial and other characteristic potentially influencing the IRC were available for every municipality in the Czech Republic.

3.2.1. Gamma dose rate

The concentration of radon in the soil gas is strongly dependent on the original uranium content. The overall effect of three natural radioactive elements, U, K and Th; is shown by the gamma dose rate. Close correlation of both indoor and soil gas radon data with gamma dose rate was observed by the comparison of Czech datasets (Barnet and Fojtikova (2008)). We obtained the values of gamma dose rate (nGy/h) from a digitalised radiometric map of the Czech Republic vectorized in scale 1:500000 (Manova and Matolin (1995)). The values ranged from 25 to 180 nGy/h.

3.2.2. The type of bedrock

As already mentioned, soil gas radon concentrations are closely linked to the uranium content of the geological bedrock. The