Science of the Total Environment 599–600 (2017) 1317–1329

Contents lists available at ScienceDirect

Science of the Total Environment

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

Logistic regression model for detecting radon prone areas in Ireland

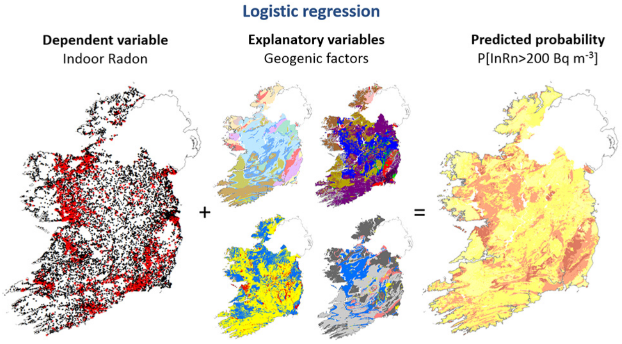
J. Elío [a], Q. Crowley [a,*,1], R. Scanlon [b], J. Hodgson [b], S. Long [c]

[a] *Geology, School of Natural Sciences, Trinity College, Dublin 2, Ireland*
[b] *Geological Survey Ireland, Ireland*
[c] *Environmental Protection Agency of Ireland, Ireland*

HIGHLIGHTS

- A new high spatial resolution radon risk map of Ireland has been developed.
- Logistic regression models were evaluated for radon mapping.
- Indoor radon measurements and relevant geological information were used.
- Probabilities of having an indoor radon concentration above 200 Bq m^{-3} were predicted.
- About 10% of the total Irish population may be affected by high indoor radon levels.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 6 March 2017
Received in revised form 2 May 2017
Accepted 8 May 2017
Available online xxxx

Editor: D. Barcelo

Keywords:
Radon
Radon mapping
Indoor radon
Geogenic factors
Logistic regression

ABSTRACT

A new high spatial resolution radon risk map of Ireland has been developed, based on a combination of indoor radon measurements ($n = 31{,}910$) and relevant geological information (i.e. Bedrock Geology, Quaternary Geology, soil permeability and aquifer type). Logistic regression was used to predict the probability of having an indoor radon concentration above the national reference level of 200 Bq m^{-3} in Ireland. The four geological datasets evaluated were found to be statistically significant, and, based on combinations of these four variables, the predicted probabilities ranged from 0.57% to 75.5%. Results show that the Republic of Ireland may be divided in three main radon risk categories: High (HR), Medium (MR) and Low (LR). The probability of having an indoor radon concentration above 200 Bq m^{-3} in each area was found to be 19%, 8% and 3%; respectively.

In the Republic of Ireland, the population affected by radon concentrations above 200 Bq m^{-3} is estimated at ca. 460 k (about 10% of the total population). Of these, 57% (265 k), 35% (160 k) and 8% (35 k) are in High, Medium and Low Risk Areas, respectively. Our results provide a high spatial resolution utility which permit customised radon-awareness information to be targeted at specific geographic areas.

1. Introduction

Radon is a natural radioactive gas present in all soils (Cothern and Smith, 1987) and represents the highest source of natural ionizing radiation to the general population (UNSCEAR, 2000a). Radon may accumulate in dwellings and may cause lung cancer when its decay products are inhaled (Field, 2015; US-EPA, 2003; WHO, 2010). In this regard, radon is

---

[*] Corresponding author.
E-mail address: crowleyq@tcd.ie (Q. Crowley).
[1] Postal address: Museum Building, Geology, School of Natural Sciences, Trinity College, Dublin 2, Ireland.

http://dx.doi.org/10.1016/j.scitotenv.2017.05.071

the second cause of lung cancer after smoking (McColl et al., 2015), and the number of lung cancer deaths linked to radon exposure may vary from 3% to 20% depending on the country (Kim et al., 2016). In the United States of America, for example, around 21,000 annual deaths from lung cancer (approximately 10% of the annual lung cancer deaths) are linked to radon exposure (Casey et al., 2015; US Cancer Statistic Working Group, 2016), while in Europe radon exposure causes around 18,000 lung cancer deaths annually (about 8% of the annual lung cancer deaths; Gray et al., 2009). Accurate prediction of radon prone areas aids regulators and local authorities to develop a proper strategy to minimise exposure to this carcinogen, resulting in an enhanced quality of life and improvement of long-term health for the general population.

Although radon is a health hazard, it can be mitigated if appropriate radon measures are implemented (e.g. Long et al., 2013; US EPA, 2001). Radon is heterogeneously distributed and certain geographic areas are more likely to have high levels of radon than others. High risk areas can be identified on radon risk maps, which are useful to target homeowners and the construction industry so that remediation work can be carried out. Such maps can also be used to identify the areas where preventive measures in new buildings should be applied (McColl et al., 2015). Therefore, radon mapping has profound economic and social implications (Gray et al., 2009), and a high-resolution, accurate and statistically robust map is crucial to inform government policy allowing development and implementation of the most cost-effective approach to reduce radon exposure in the general population.

The criteria selected to develop a radon risk map depends on national radiological protection strategies and the data available. Data on which such maps are based are conventionally indoor radon measurements (e.g. Fennell et al., 2002; Ferreira et al., 2016) or soil-gas radon measurements (e.g. Bossew, 2015a). The use of geological information, and other factors, assists to improve the accuracy of these maps (e.g. Bossew, 2015b, 2014; Ferreira et al., 2016; Miles and Appleton, 2005; Pásztor et al., 2016). Such an approach is warranted, as geology is known to be the main factor controlling indoor radon concentrations (Appleton and Miles, 2010; Watson et al., 2017).

Consequently, radon mapping can be divided in two activities, one for developing Indoor Radon Maps and another for Geogenic Radon Maps (e.g. Bossew et al., 2013; Miksova and Barnet, 2002). The principal advantage of the Geogenic Radon Maps is that it measures "*what earth delivers in terms of radon*" (Bossew et al., 2013) and, therefore, they are independent of anthropogenic factors. Recent studies suggest that this affirmation is not totally correct and some other influences exist (Bossew, 2014). On the other hand, the main advantage of Indoor Radon Maps is that radon is directly measured at the exposure point (i.e. dwellings). Nevertheless, uncertainties in the exact location of tested homes and extensive areas with little or no data may impede the generation of an accurate Indoor Radon Map. Furthermore, indoor concentrations depend on multiple factors, both natural and anthropogenic (Gunby et al., 1993; Tollefsen et al., 2014), and even some dwellings may also have been subjected to preventive/remediation measures (e.g Long et al., 2013) which may hinder the correct interpretation of data. Finally, an appropriate sample design is crucial for accurate assessments of the radiological risk (Burke and Murphy, 2011).

In this study, the use of logistic regression models has been evaluated, combining indoor radon measurements and geological information. The primary objective of this study is to estimate the probability of having an indoor radon concentration above the Irish national reference level of 200 Bq m^{-3} (Fennell et al., 2002). Previous work has been carried out in Ireland by multivariate linear regression (Hodgson et al., 2014), however, logistic regression models are a better statistical tool for studying the probability of occurrence of categorical variables (James et al., 2013). In this regard, for radon mapping purposes indoor radon can be evaluated as a binary variable taking the value 0 or 1 if it is below or above the reference level, respectively, while simultaneously reporting the probability of belonging to one of the given explanatory variable categories (in this study Bedrock Geology - BG, Quaternary Geology - QG, Subsoil permeability - SP, and aquifer type - AT). Such models are frequently used in epidemiology, for example to analyse the relationship between radon and lung cancer (Kreienbrock, 2001; Wang et al., 2002), but also to analyse natural hazards (Dong et al., 2011; Ettinger et al., 2016; Notario del Pino and Ruiz-Gallardo, 2015; Wang et al., 2016) and more recently for radon mapping (Kropat et al., 2017).

Indoor radon measurements were previously used to develop the Irish National Radon Map in grids of 100 km^2 (www.epa.ie/radiation/radonmap). Each grid was classified according to the probability of a dwelling having an indoor radon concentration above the reference level of 200 Bq m^{-3} (i.e. <1%, between 1% and 5%, between 5% and 10%, between 10% and 20% and >20%; Fennell et al., 2002), and a "High Radon Risk Area" was defined as the area in which this probability is 10% or higher. The same indoor radon data and the same criteria were used in this research, however, geological information (i.e. BG, QG, SP, and AT) has been included to add geogenic factors and develop a higher resolution radon risk map of the Republic of Ireland.

2. Material and methods

2.1. Indoor radon measurements

The indoor radon (InRn) measurements used in this study were carried out by the Environmental Protection Agency of Ireland (EPA) as part of a national survey in Irish dwellings (Fennell et al., 2002). More than 60,000 dwellings have been sampled by the EPA and local authorities across the country. However, only 32,108 measurements have been accurately geo-referenced by the Geological Survey of Ireland (Hodgson

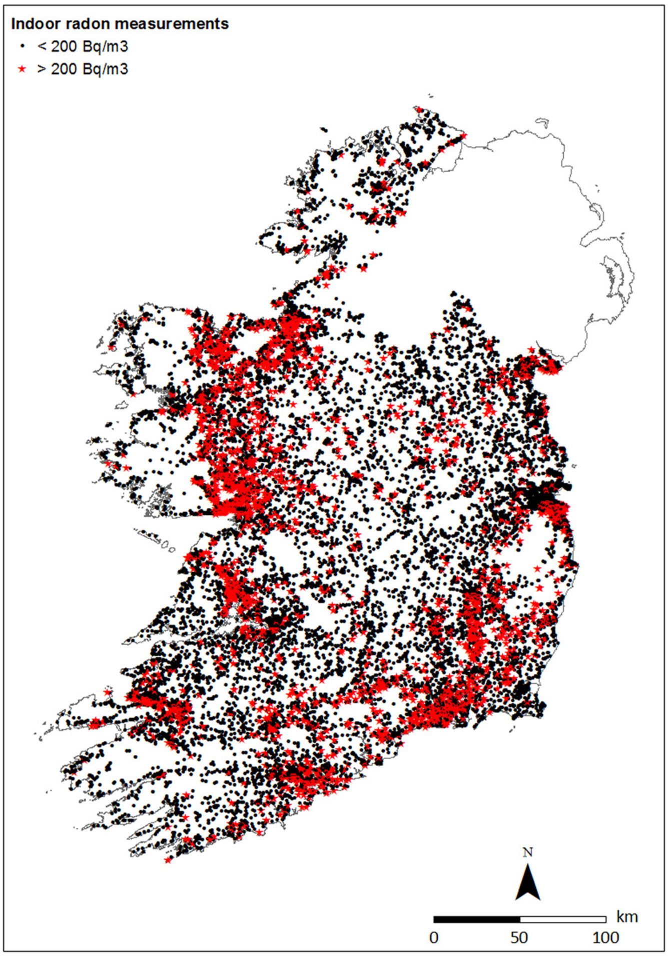


Fig. 1. Indoor radon measurements in Republic of Ireland (N = 31,910).