Construction and Building Materials 159 (2018) 755–767

Contents lists available at ScienceDirect

Construction and Building Materials

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

Virtual Special Issue
Use of NORM-containing products in construction

The NORM4Building database, a tool for radiological assessment when using by-products in building materials

Wouter Schroeyers [a,*], Zoltan Sas [a,b], Gergo Bator [c], Rosabianca Trevisi [d], Cristina Nuccetelli [e], Federica Leonardi [d], Sonja Schreurs [a], Tibor Kovacs [c]

[a] Hasselt University, CMK, NuTeC, Nuclear Technology – Faculty of Engineering Technology, Agoralaan building H, B-3590 Diepenbeek, Belgium
[b] School of Natural and Built Environment, Queen's University Belfast, David Keir Bldg., 39-123 Stranmillis Rd, Belfast BT9 5AG, United Kingdom
[c] University of Pannonia, Institute of Radiochemistry and Radioecology, H-8200, Egyetem St. 10., d Veszprém, Hungary
[d] INAIL (National Institute for Insurance against Accidents at Work)- Research Sector, DiMEILA, Via di Fontana Candida 1, 00078 Monteporzio Catone, Rome, Italy
[e] ISS (National Institute of Health), Technology and Health Department, Viale Regina Elena, 299 Rome, Italy

HIGHLIGHTS

- Datamining was used for the construction of the NORM4Building database.
- Systematic radiological evaluation of by-products for use in concrete.
- Radiological evaluation of cement, concrete, ceramics and (phospo)gypsum.
- The datamining approach enables the construction of an updated detailed database.

ARTICLE INFO

Article history:
Received 16 July 2017
Received in revised form 24 October 2017
Accepted 10 November 2017
Available online 15 November 2017

Keywords:
Natural occurring radioactive materials
Building materials
Database
Concrete
By-products
Euratom Basic Safety Standards

ABSTRACT

Scientific data on natural occurring radioactive materials (NORMs) is available in unknown quantities and the data is fragmented over several different sources. The new EU-BSS is regulating the use of NORM in building materials, however a large scale database with country specific information that can support legislators and industry in the assessment of the radiological impact of the use of by-products in construction is missing. Currently the COST Action 'NORM4BUILDING' (2014–2017) is creating such a database using a semi-automated datamining approach. In this paper radiological aspects on by-products that can find application in concrete are discussed based on the database.

1. Introduction

Europe is evolving to a more resource controlled continent. This transition is driven by the EU action plan on developing a circular economy [1] and supported by the waste framework directive [2]. The constituents of more and more construction materials, and in particular building materials, are being replaced with by-products from several industrial sectors. Upon replacing raw materials by by-products that contain increased concentrations of naturally occurring radionuclides care needs to be taken to ensure that the newly produced construction materials meet the radiological protection standards of the EU and its Member States. The new Euratom Basic Safety Standards (EU-BSS) [3] were published in 2014 and contain requirements for industrial sectors involving Naturally Occurring Radioactive Materials (NORMs) and the use of specific residues from the considered industrials sectors in building materials. The EU-BSS is expected to be transposed in national and regional legislations by February 6, 2018 for all the member states of the EU. Worldwide the safe use of NORM is also becoming increasingly important. The new IAEA-BSS (Radiation protection and safety of radiation sources: international basis safety standards) [4] was also published in 2014. Both the EU-BSS and the IAEA-BSS are based on ICRP recommendation 103 [5]. Simulated and experimental data originating from scientific

* Corresponding author.
E-mail address: Wouter.schroeyers@uhasselt.be (W. Schroeyers).

https://doi.org/10.1016/j.conbuildmat.2017.11.037
0950-0618/

papers dealing with NORMs are gathered in UNSCEAR reports [6] and these form the basis of the ICRP recommendations.

A major challenge regarding the collection of data on NORM and NORM-containing construction materials is that literature data is fragmented over a lot of different sources. It is even unclear to which extend country-specific literature data is available. Only a limited number of data sets, all of them manually collected, are available. Trevisi et al. [7] collected a lot of data on activity concentration measurements of natural radionuclides (^{226}Ra, ^{232}Th and ^{40}K) in building materials used in 26 of 27 EU Member States. Sas et al. [8] constructed a database ('By-BM database') by collecting data on raw materials, by-products and construction materials from 48 countries. Next to this database the NORM database of NIRS (National Institute of Radiological Sciences, Japan) [9] is available but it lacks data on European NORM or construction materials.

Currently the COST Action Tu1301 'NORM4BUILDING' (2014–2017) is creating an extensive database gathering radiological data on NORMs and construction materials in which NORM containing by-products are implemented. The COST Action closely collaborates with the "MetroNORM" project (MetroNORM – Metrology for processing materials with high natural radioactivity) [10], involving several European NMIs (National Metrological Institutes), to develop standardized measurement protocols for NORM and NORM-containing construction materials. In the current paper, the contents of the NORM4BUILDING database is presented and discussed.

For the collection of large amounts of data, there is a need for a systematic, as much as possible automatized, approach to data mining to investigate the different fragmented sources of data. Upon using an automatized approach of gathering scientific data, a crucial scientific challenge is the validation and selection of data for incorporation in the database, an aspect which is especially complicated when dealing with NORM. For the characterization of NORM ISO standards and standardized measurement and sampling procedures are still under development. The analytical determination of the activity concentrations of natural radionuclides in NORM samples requires specialized knowledge. Not all papers deal appropriately with the absence of secular equilibrium which in general is the rule for many industrial by-products and construction materials that are based on these by-products. The current paper describes in more detail the methodology used for the semi-automatized data mining approach that was developed to build the NORM4Building database.

Based on the information collected in the NORM4Building database, this study aims to give a realistic evaluation of the radiological properties of construction materials based on NORM by-products using the criteria set by the EU-BSS.

2. Materials & methods

2.1. Data mining for data collection

A data mining approach was used as an analytical method to extract information regarding NORM, i.e. raw materials (ores), by-products, and NORM-containing construction materials from published papers. The text mining method [11], a process of analysing collections of textual materials to capture key concepts, basic parameters, keywords and to uncover hidden relationships and trends, was applied.

The main milestones of the data mining process are the following and show the Fig. 1:

1. Automatic keyword driven selection of papers,
2. Building and applying filters based on selected keywords,
3. Handling different types of publications into the IBM SPSS Modeler™ software to extract information using natural language processing (NLP),
4. Extracting complete measurement results and origin (Country) from relevant publications available in different formats such as: Microsoft Word™, Microsoft Excel™, and Microsoft PowerPoint™, as well as Adobe PDF™, XML, HTML,
5. Applying Text Link Analysis (TLA) to define pattern rules and to compare these to relationships found in the text,
6. Structuring the collected data.

Publications from different electronic sources, such as Science Direct, Web of Science and others (IAEA and ICRP documents, National Surveys, etc.) have been processed. After applying the filter, relevant publications, which contain measurement results for activity, activity concentration, or exhalation rate were extracted. The next step was to identify the specific keywords related to these results. At the beginning of the COST project, a trial database, was built manually including 16 types of materials extracted from approximately 100 publications. To expand the list of keywords beyond these material types text link analysis (TLA) was used. TLA is a pattern-matching technology that enables to define pattern rules and compare the pattern rules to the concepts present in text of the extracted documents. Applying TLA resulted of in a lot of false-positive matches, where the concept found is not a material of interest and is not relevant for data collection. By browsing the list and selecting relevant concepts the materials list (the list of keywords) was expanded from 16 to 59 types of materials excluding general expressions. In the publications, two types of measurement information were identified: tables listing the results and grammatically complete results definitions. Both types of information require different ways of extraction because the date is structurally organized differently. Both strategies for data collection have in common that they involve library building, so the information is imported into a library. A disadvantage is that in its current version the automatic data collection approach cannot handle figures.

The geological origin of the studied materials is very relevant for the radiological properties of the studied materials. In addition, country specific circumstances (legislation, dominant industrial sectors, accessible resources, etc.) determine the properties and use of investigated materials and therefore the database aims at providing country specific information.

2.2. Validation of entries

Upon selecting the data for inclusion in a database, important aspects need to be controlled: (1) The reliability of the data used needs to be verified. (2) Another problem in data selection is that measurement results are sometimes repeated in several publications: new papers can be based on previous results and in a way measurement results can then be reported in double or even in multiple times in different sources. In this way, the data can be overrated in data analysis. (3) In numerous papers, the number of samples measured is not clearly reported.

The method used to solve these problems and validate the database data, identifying correct and useful data, was a careful reading after the final collection step. Additional analysis of the text allowed to individuate, in many cases, the number of samples. When this procedure did not succeed, the number of samples was considered as 1; that is why the total number of data in the database is certainly an underestimation. As regards to the problem of same data from different papers, a careful reading of text and references was necessary to verify if data was new or already considered from other included papers.

2.3. Database content & structure

The total number of entries in the database is 1422 and the total number of samples is 12365 (date: 01/07/2017). The database contains data on 26 (No data was found for Latvia and Malta) of the 28 Member States of the European Union and all together for 74 countries worldwide. An initial report on the start-up version of the NORM4Building database was given in [12].

The data are classified into 3 main categories:

- *Primary raw materials* – coal (coal, lignite, peat, bituminous coal), phosphate ores, ferrous ores (iron ores such as hematite, limonite, magnetite), non-ferrous ores and minerals (aluminium ore (bauxite), gold ore, manganese ore, molybdenum ore, monazite, nickelic ore, titanium ore, uranium rocks, zirconium ore, ilmenite, rutile, baddeleyite) and other natural rocks or sands (basalt, black sand, chalk, chert, clay (or clay minerals such as kaolinite), diabase rock, dolomite, gabbro, granite, gravel, marble (composed out of calcite or dolomite), marl (or marlstone), pumice, quartzite, sand, sandstone, schist, serpentinite, soil, stone, trass, tuff),
- *(By)-products* – red mud, fly ash, bottom ash, different types of slags (iron slag, steel slag, coal slag, copper slag, blast furnace slag…), and other materials (calcium carbonate, cerium oxide, copper, corundum, fertilizer material, iron oxide, mud, sludge, titanium dioxide, dross, tailing, scale…),
- *Construction materials* – cement, ceramics, concrete, phosphogypsum, gypsum and other materials (bricks, clinker, gas silicate blocks, plaster, adobe, mortar…).

The current paper limits itself to specific by-products, considered by the EU-BSS, that are likely to be used as a building material and on building materials such as concrete, ceramics and (phospo)gypsum. In total the current paper considers a selection of 460 entries (7705 samples) from the database. The references from which the extracted data (for a given country) originates are given in Table 1. From one reference several different entries linked to different types of samples can be extracted for a given country or from different countries. For most countries, the results are not statistically representative at the national level due to the low number of data available in the literature.