



Contribution ID: 21

Type: **Oral Presentation**

MACHINE LEARNING IN ENVIRONMENTAL RADON SCIENCE

Monday, 2 May 2022 11:30 (20 minutes)

Acknowledged to be a relevant hazardous agent and the most important contribution to ionizing radiation exposure, in most cases, radon (Rn –mostly referring to ^{222}Rn , but in some context also to ^{220}Rn) has therefore been subject to regulation and to intense scientific interest for many years. Temporal dynamic as well as spatial variability of environmental Rn are controlled by a number of factors such as meteorology, lithology, soil properties, hydrogeology, tectonics and seismicity. In addition, indoor radon concentration and the one of its short-lived progeny (which cause the actual radiological problem) are subject to anthropogenic factors, such as physical characteristics of a building and usage pattern.

The causal relationships between controlling factors and response variables –Rn concentration in an environmental compartment as function of time and long-term mean concentrations as function of geographical location –are mostly complicated and difficult to model, although the underlying physical processes are simple. Predictors are often intercorrelated or badly known. On the other hand, understanding and being able to predict temporal and spatial Rn dynamic are important for these reasons: (1) Rn maps are important tools of decision making in Rn policy and implementing regulation whose objective is reducing Rn exposure; (2) Analysis of Rn time series is a tool to exploit the potential of Rn as tracer of environmental processes which manifest in signals in time series.

Traditional tools are classical geostatistics for spatial analysis and prediction, most commonly methods belonging to the kriging family, and time series analysis methods such as ARMA-type approaches and multivariate linear regression, among other. However, in particular in high-dimensional multivariate settings and complicated dependence structure of covariates, these traditional methods seem to have come to an end or need to be supplemented by new ones.

New tools belong to what is commonly called machine learning (ML). Here, algorithms learn from the data, without need of defining dependence models beforehand, in general. The algorithm establishes a data-driven model which honours the (often partly hidden) dependence structure; this has the disadvantage that it does not yield a physical explanation of the effect. On the other hand, experience shows that the predictive power can be substantially higher than the one of traditional approaches.

We give a short overview on ML methods and discuss their respective merits, their application and ways of validating results. We show examples of ^{222}Rn mapping in Europe based on ML involving a number of geogenic predictors, and of analysis of time series of a long-term experiment being carried out in Chiba, Japan, involving soil and indoor Rn and thoron (Tn , ^{220}Rn) concentrations and meteorological covariates, whose ultimate target is identifying seismic signals in Rn time series, and even predicting seismicity based on detecting signals of seismic precursors in Rn dynamic.

Primary authors: Dr PETERMANN, Eric (BfS); BOSSEW, Peter (BfS Berlin); Dr ELÍO, Javier (Department of Planning, Aalborg University); Dr JANIK, Miroslaw (QST / NIRS)

Presenter: BOSSEW, Peter (BfS Berlin)

Session Classification: Applications

Track Classification: Applications