PROCEEDINGS

LANDAUER EUROPE

IUPAP

2014

RAD

Second International Conference on Radiation and Dosimetry in Various Fields of Research

CEI CENTRAL EUROPEAN INITIATIVE

www.rad2014.elfak.rs

May 27 - 30, 2014 | Faculty of Electronic Engineering | Niš | Serbia

PROCEEDINGS

2014

RAD

May 27 - 30, 2014 | Faculty of Electronic Engineering | Niš | Serbia

Second International Conference on Radiation and Dosimetry in Various Fields of Research

www.rad2014.elfak.rs

PUBLISHER: University of Niš, Faculty of Electronic Engineering
P.O.Box 73, 18000 Niš, Serbia
www.elfak.ni.ac.rs
FOR THE PUBLISHER: Prof. Dr. Dragan Tasić
EDITOR: Prof. Dr. Goran Ristić
COVER DESIGN: Vladan Nikolić, M.Sc.
TECHNICAL EDITING: Sasa Trenčić and Vladan Nikolić
PROOF-READING: Saša Trenčić, MA
ISBN 978-86-6125-101-6

The Second International Conference on Radiation and Dosymetry in Various Fields of Research (RAD 2014) and *the Second East European Radon Symposium (SEERAS)* were financially supported by:

- Central European Initiative (CEI)
- International Union of Pure and Applied Physics (IUPAP)*
- Ministry of Education, Science and Technological Development

*To secure IUPAP sponsorship, the organisers have provided assurance that RAD 2014 Conference will be conducted in accordance with IUPAP principles as stated in the IUPAP resolution passed by the General Assembly in 2008. In particular, no bona fide scientist will be excluded from participation on the grounds of national origin, nationality, or political considerations unrelated to science.

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

539.16(082)(0.034.2)

INTERNATIONAL Conference on Radiation and Dosimetry in Various Fields of Research (2nd ; 2014 ; Niš)
Procedings [Elektronski izvor] / The Second International Conference on Radiation and Dosimetry in Various Fields of Research, RAD 2014, May 27-30, 2014, Niš, Serbia ; [editor Goran Ristić]. - Niš : Faculty of Electronic Engineering, 2014 (Niš : Faculty of Electronic Engineering). - 1 elektronski optički disk (CD-ROM) ; 12 cm

Sistemski zahtevi: Nisu navedeni. - Nasl. sa naslovnog ekrana. - Nasl. sa naslovnog ekrana drugog rada: Proceedings / Second East European Radon Symposium (SEERAS), May 27-30, 2014, Niš, Serbia. - Tiraž 350. - Bibliografija uz svaki rad.

ISBN 978-86-6125-101-6

a) Јонизујуће зрачење - Дозиметрија - Зборници
COBISS.SR-ID 207467788

MULTIVARIATE ANALYSIS OF CLIMATE VARIABLES, TELECONNECTION INDICES AND ACTIVITIES OF LEAD-210 AND BERYLLIUM-7 IN SURFACE AIR IN BELGRADE, SERBIA

Dimitrije Maletić [1], Jelena Ajtić [2], Vladimir Đurđević [3], Dragana Todorović [4], Jelena Nikolić [4], Radomir Banjanac [1], and Vladimir Udovičić [1]

[1] University of Belgrade, Institute of Physics, Pregrevica 118, 11080 Belgrade, Serbia

[2] University of Belgrade, Faculty of Veterinary Medicine, Bulevar oslobodjenja 18, 11000 Belgrade, Serbia

[3] University of Belgrade, Faculty of Physics, Institute of Meteorology, Studentski trg 12, 11000 Belgrade, Serbia

[4] University of Belgrade, Vinča Institute of Nuclear Sciences, Laboratory for Environmental and Radiation Protection, 11000 Belgrade, Serbia

Abstract. *Activities of lead-210 and beryllium-7 have been monitored at the Vinča Institute of Nuclear Sciences in Belgrade, Serbia. The monthly mean activities in composite aerosol samples were determined on HPGe detectors by standard gamma spectrometry. The meteorological data, consisting of the temperature, atmospheric pressure, relative humidity, sunshine hours and cloud cover data were obtained from the European Climate Assessment & Dataset and the Republic Hydrometeorological Service of Serbia. Five teleconnection indices of large scale atmospheric circulation: North Atlantic Oscillation, East Atlantic Pattern, East Atlantic/West Russia Pattern, Scandinavia Pattern, and Polar/Eurasia Pattern were obtained from the data archive of the United States National Oceanic and Atmospheric Administration's Climate Prediction Center. The first lead-210 and beryllium-7 activities measured at the Vinča Institute date back to 1985 and 1991, respectively, and their relation with the climate variables and teleconnection indices is investigated using multivariate methods of analysis.*

The most appropriate multivariate method of analysis of these sets of measurements is selected from a wide spectrum of multivariate methods developed for data analysis in high-energy physics and implemented in the Toolkit for Multivariate Analysis software package. The evaluation ranking results based on the best signal efficiency and purity, show that the Boosted Decision Trees (BDT) multivariate method is the most suitable for the variable analysis. Further multivariate analysis results give insight into the dependence of lead-210 and beryllium-7 concentrations upon the climate variables and atmospheric circulation (via the teleconnection indices) during the time of measurements. The BDT method singles out the Scandinavia Pattern index as the variable with the highest importance for both radionuclides. Amongst the climate variables, temperature shows the strongest influence on the radionuclide concentrations, while relative humidity is the lowest ranking variable. Moreover, the multivariate regression methods give a good approximation of lead-210 and beryllium-7 concentrations for all the sets of climate variables and teleconnection indices.

Key words*: *lead-210, Be-7, air, climate variables, teleconnection indices, multivariate methods

1. INTRODUCTION

Environmental radioactivity monitoring in Belgrade, Serbia, includes continuous control of radioactivity of aerosols at meteorological station in Vinča Institute of Nuclear Sciences. The continuous control is done by measurements of radioactivity of lead-210 and beryllium-7. These two naturally occurring radionuclides have different sources: ^{210}Pb (half-life 22.3 years) in air results from ^{222}Rn decay, ground resuspension, and anthropogenic activities (mainly coal combustion), while ^{7}Be (half-life 53.28 days) is produced in the upper troposphere and lower stratosphere, and then transported into lower altitudes. However, washout is the major removal process from the atmosphere for both radionuclides, which is a likely reason for the radionuclides medium, but significant correlation in surface air [1]. The extent of influence of local climate variables on the activities of these radionuclides has been studied before [e.g., 1, 2], while the impact of large scale atmospheric transport is not as well documented [3].

The goal of this analysis is therefore to look more thoroughly into possible relations between the ^{210}Pb and ^{7}Be activities in surface air and available meteorological data and large scale atmospheric transport patterns (quantified by teleconnection indices). Multivariate methods which have been developed and used in high-energy physics were chosen for our analysis. Numerous multivariate methods and algorithms for classification and regression can be found in the analysis framework ROOT [4] and, more precisely, in the Toolkit for

Multivariate analysis (TMVA) [5]. These multivariate methods were used to test all available classifiers and regression methods implemented, in order to find a most appropriate method for the analysis of the ^{210}Pb and ^{7}Be dependence on the climate variables and teleconnection indices.

2. EXPERIMENTAL DATA

The monthly mean activities of lead-210 and beryllium-7 in composite aerosol samples are determined using HPGe detectors by standard gamma spectrometry. The measurements of activities of lead-210 started in 1985 and of beryllium-7 in 1991.[8]

The meteorological data, consisting of temperature, atmospheric pressure, relative humidity, sunshine hours and cloud cover data were obtained from European Climate Assessment & Dataset (ECA&D) [6] and the Republic Hydrometeorological Service of Serbia.

Five teleconnection indices of large scale atmospheric circulation: North Atlantic Oscillation, East Atlantic Pattern, East Atlantic/West Russia Pattern, Scandinavia Pattern, and Polar/Eurasia Pattern were obtained from the data archive of the United States National Oceanic and Atmospheric Administration's Climate Prediction Center (http://www.cpc.ncep.noaa.gov/data/teledoc/teleco ntents.shtml visited on 18 October 2013).

3. CALCULATIONS

The first step in the analysis was to calculate and rank the correlation coefficients between all the given variables and the ^{210}Pb and ^{7}Be activities (Table 1). Apart from the useful information on the connection between the variables and measured activities, the first stage calculations also help in setting up and testing the framework for running the various multivariate methods contained in the TMVA.

In the next step, the multivariate methods were used and compared in order to find the one best suited for classification (division) of the ^{210}Pb and ^{7}Be activities into two sets of data: acceptable concentrations and increased concentrations in air samples. Preferably, the result of this method comparison is a single method which, given the input climate variables and atmospheric indices, provides an output close to the observed variations in the ^{210}Pb and ^{7}Be concentrations.

In order to use the multivariate classification, a set of input events (consisting of ^{210}Pb or ^{7}Be concentration value along with meteorological variables and teleconnection indices values measured/taken simultaneously) was first split into events that correspond to the signal (i.e., events that include ^{210}Pb or ^{7}Be concentrations that are considered increased) and to the background (events that include ^{210}Pb or ^{7}Be concentration that is declared acceptable). For the purpose of this preliminary analysis, the splitting of the input events was performed at 1.1 mBq/m^{3} and 5 mBq/m^{3} for ^{210}Pb and ^{7}Be, respectively. These limits were used for the majority of the analyses. They were selected because the splitting ensured maximum employment of multivariate comparison methods. Moreover, these particular values reflected the fact that the number of background events is greater that the number of signal events. The method of multivariate regression, however, does not require preliminary splitting of input events, and is therefore more general.

4. MULTIVARIATE METHODS

The Multivariate methods in Toolkit for Multivariate Analysis (TMVA) belong to the family of "supervised learning" algorithms. They make use of training events, for which the desired output is known, to determine the mapping function that either describes a decision boundary (classification) or an approximation of the underlying functional behavior defining the target value (regression). The two most important Multivariate methods for our purposes are "Boosted Decision Trees" (BDT) and "Artificial Neural Networks" (ANN).

4.1. Boosted Decision Trees

In BDT, the selection is based on a majority vote on the result of several decision trees. Decision tree consists of successive decision nodes which are used to categorize the events in a sample as either signal or background. Typically, BDT is constructed of a forest of such decision trees. However, the advantage of the straightforward interpretation of the decision tree is lost. Detailed information on BDT can be found in the TMVA manual [5].

4.2. Artificial Neural Networks

ANN is any simulated collection of interconnected neurons, with each neuron producing a certain response at a given set of input signals [7]. ANNs in TMVA belong to the class of Multilayer Perceptrons (MLP), which are feed-forward neural networks.

5. RESULTS

A relation between the radionuclides' activities and the input variables was first investigated through correlation coefficients given in table 1. The coefficients showed a stronger correlation of the ^{7}Be activity with the climate variables (mean, minimum and maximum temperature and sunshine hours). Lead-210, on the other hand did not show correlation with any of the input variables.

The best method for MVA classification was BDT followed by MLP. Table 2 shows variable importance ranking of the top 4 variables. The method independent variable importance shows which of the input variables used contributes the most to classification of the radionuclides' activities to normal and increased ones.

Table 1 Correlation coefficients in % for the input variables and the ^{210}Pb and ^{7}Be activities

Variable	Pb-210	Be-7
Mean temperature	-8	46
Minimum temperature	-9	45
Maximum temperature	-7	46
Pressure at sea level	11	-14
Humidity	5	-24
Precipitation	-5	-2
Sunshine hours	-5	30
Cloud cover	-7	-19
North Atlantic Oscillation	6	-3
East Atlantic Pattern	-1	9
West Pacific Pattern	-2	-1
East Pacific / North Pacific Pattern	17	8
Pacific/North American Pattern	-4	-6
East Atlantic/West Russia Pattern	8	14
Scandinavia Pattern	10	-3
Polar/Eurasia Pattern	14	/

Table 2 BDT variable importance ranking of 4 most important variables

^{210}Pb	^{7}Be
Scandinavia Pattern	Scandinavia Pattern
East Pacific / North Pacific Pattern	West Pacific Pattern
Polar/Eurasia Pattern	North Atlantic Oscillation
West Pacific Pattern	East Atlantic Pattern

5.1. Regression MVA methods

Regression represents an approximation of the underlying functional behavior which defines the target value (^{210}Pb and ^{7}Be activities) in dependence on the input climate variables and atmospheric indices. Prior to commencing an investigation into the best MVA methods, we want to make method independent ranking of input variables (Table 3) by correlation with the measured concentration values of ^{210}Pb and ^{7}Be. It can be concluded from the table which of the input variables are likely to be most important for the MVA regression. Similar to the correlation coefficients table for classification methods (Table 1), the method independent table shows that for regression most important variables are the temperature, sunshine hours (and humidity) for the ^{7}Be concentration evaluations. For ^{210}Pb the most important variables were the two teleconnection indices and pressure at sea level.

In further analysis, an attempt was made to find the best regression method that gives output values (evaluated ^{210}Pb and ^{7}Be activities) as close as possible to the measured concentrations. The best multivariate regression method was found to be BDT, and the second one MLP, i.e. the same methods singled out in the case of multivariate classifiers. Figure 1 presents the comparison of frequency distributions of ^{210}Pb for measured and MLP evaluated values.

It is worth mentioning that when using the "supervised learning" MVA methods, caution in selecting resulting methods is required. This primarily means that the selected method ought to be proven good. To do that, good test algorithms are needed. Unfortunately, not all MVA regression methods have a sound test of a training process. The test process shows whether the evaluated values of concentrations are far from measured. The MLP (Figure 2) and DBT methods have good tests of the learning process, which is not the case for all MVA methods. Also, we need to mention that, if some MVA method gave distributions of evaluated ^{210}Pb or ^{7}Be concentrations with many apparent discrete values, we did not considered it as an appropriate method for our analysis, since we expect the more continual distributions to be closer to reality.

Table 3 Variable importance ranking, independent of the method for MVA regression

^{210}Pb	^{7}Be
Polar/Eurasia Pattern	Mean temperature
East Pacific / North Pacific Pattern	Maximum temperature
Pressure at sea level	Minimum temperature
Scandinavia Pattern	Sunshine hours
Minimum temperature	Humidity
North Atlantic Oscillation	Cloud cover
Mean temperature	Pressure at sea level
East Atlantic/West Russia Pattern	East Atlantic/West Russia Pattern
Cloud cover	East Atlantic Pattern
Maximum temperature	East Pacific / North Pacific Pattern
Precipitation	Pacific/North American Pattern
Sunshine hours	North Atlantic Oscillation
Humidity	Scandinavia Pattern
Pacific/North American Pattern	Precipitation
West Pacific Pattern	West Pacific Pattern
East Atlantic Pattern	Polar/Eurasia Pattern

Additionally, several MVA methods, like Function Discriminant Analysis (FDA) (Fig. 3) had negative regression coefficient for distribution of differences of the evaluated and measured values of ^{201}Pb concentrations vs ^{201}Pb concentrations. This negative regression coefficient was corrected, thus making several more MVA methods useful for our analysis. As a result, methods which could successfully be used for regression were the following MVA methods: Artificial Neural Networks, Boosted Decision Trees, Function Discriminant Analysis, Linear Discriminant, Multidimensional k-Nearest Neighbour Method, and Multidimensional

Probability Density Estimator Range-Search. All these methods gave distributions of the evaluated values similar to MLP, which encouraged us to use all the results from the MVA analysis. Since MVA methods are mathematically very diverse, getting the same results points out that they properly describe the connection between the experimental values and the meteorological variables and teleconnection indices.

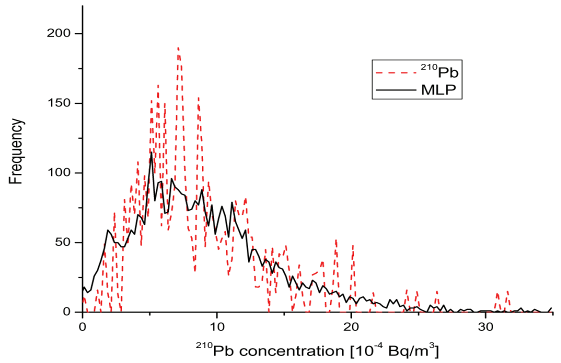

Fig. 1 Distribution of the MLP evaluated values and measured ^{210}Pb activities.

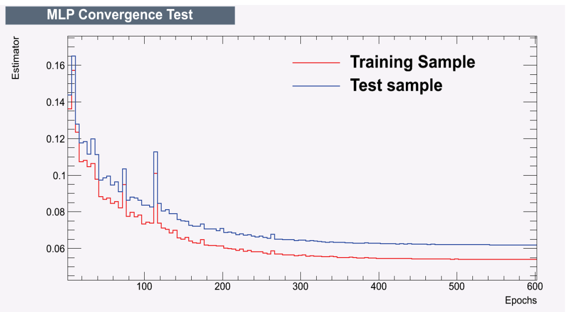

Fig. 2 Testing of MLP method. The evaluated values of ^{210}Pb concentrations were close enough to the measured values, and during a training process the difference between the evaluated and measured value was getting smaller.

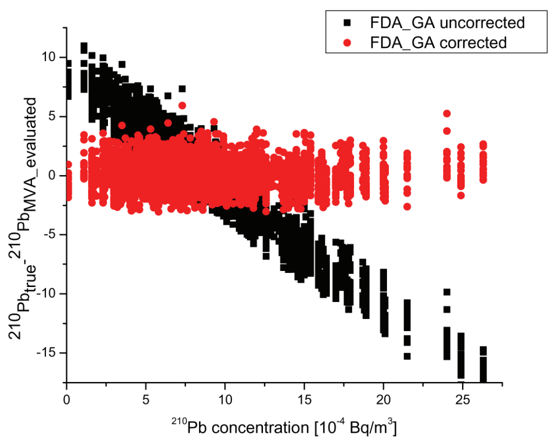

Fig. 3 Correction of evaluations for the FDA method.

6. Conclusion

The first tests of using the MVA methods in finding the relation between the ^{210}Pb and ^{7}Be activities in air, and the meteorological variables and teleconnection indices, were shown to be very promising. The MVA classification and regression methods can distinguish between inter-correlations of input variables and correlation of input variables and ^{210}Pb and ^{7}Be activities, and also to point out to the, MVA method specific, variables which contribute the most to efficient classification and evaluation of the measured ^{210}Pb and ^{7}Be activities. Successful MVA mapping of functional behavior of the ^{210}Pb and ^{7}Be activities is the most important result of this work. Having the MVA mapped functional behavior will help us with further research of activities of these isotopes in air samples.

***Acknowledgement*:**

This paper was realized within the project "Climate changes and their influence on the environment: impacts, adaptation and mitigation" (No. 43007) financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia (2011-2014).

References

1. J. Ajtić, D. Todorović, J. Nikolić and V. Djurdjević, "A multi-year study of radioactivity in surface air and its relation to climate variables in Belgrade, Serbia", Nucl. Technol. Radiat., vol. 28, pp. 381-388, December 2013.
2. M. K. Pham, M. Betti, H. Nies and P. P. Povinec, "Temporal changes of ^{7}Be, ^{137}Cs and ^{210}Pb activity concentrations in surface air at Monaco and their correlation with meteorological parameters", J. Environ. Radioact., vol. 102, pp. 1045-1054, November 2011.
3. A.-P. Leppanen et al., "Cosmogenic ^{7}Be in air: A complex mixture of production and transport", J. Atmos. Sol. Terr. Phys., vol. 72, pp. 1036-1043, August 2010.
4. R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework", Nucl. Inst. Meth. in Phys. Res. A 389 (1997), 81.
5. A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA - Toolkit for Multivariate Data Analysis", PoS ACAT 040 (2007), arXiv:physics/070303
6. A. M. G. Klein Tank et al., "Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment", Int. J. Climatol., vol. 22, pp. 1441-1453, 2002.
7. R. Rojas, "Neural Networks", Springer-Verlag, Berlin, 1996
8. D. Todorović, D. Popović, J. Nikolić, J. Ajtić, „Radioactivity monitoring in ground level air in Belgrade urban area", Radiation Protection Dosimetry, 2010, Vol.142, No. 2-4, pp. 308-313.