Communications in Physics, Vol. 29, No. 3SI (2019), pp. 411-421
DOI:10.15625/0868-3166/29/3SI/14336

STATISTICAL ANALYSIS TO EVALUATE HEAVY METAL POLLUTION IN THE AIR OBTAINED BY MOSS TECHNIQUE IN HANOI AND ITS SURROUNDING REGION

N. H. QUYET[1], L. H. KHIEM[2,3],†V. D. QUAN[2], T. T. T. MY[4], M. V. FRONTASIEVA[4],
N. T. B. MY[1,4], L. D. NAM[2], N. N. MAI[2], K. T. HONG[2], D. P. T. TIEN[5], D. V. THANG[1],
T. D. TRUNG[6], N. A. SON[7] AND T. T. THANH[8]

[1]*Institute for Nuclear Science and Technology- Vietnam Atomic Energy Institute, Hanoi, Vietnam*
[2]*Institute of Physics, Vietnam Academy of Science and Technology (VAST), Hanoi, Vietnam*
[3]*Graduate University of Science and Technology, VAST, Hanoi, Vietnam*
[4]*The Joint Institute for Nuclear Research, Dubna, Russia*
[5]*Nha Trang Institute of Technology Research and Application, VAST, Nha Trang, Vietnam*
[6]*Centre for high technology development, VAST, Hanoi, Vietnam*
[7]*Faculty of Nuclear Engineering, Dalat University, Da Lat, Lam Dong, Vietnam*
[8]*Faculty of Physics and Engineering Physics, University of Science, Vietnam National University Ho Chi Minh City*

†*E-mail:* lhkhiem@iop.vast.ac.vn

Received 22 August 2019
Accepted for publication 28 September 2019
Published 18 October 2019

Abstract. *The aim of this paper was the application of statistical analysis including principal component analysis to evaluate heavy metal pollution obtained by moss technique in the air of Hanoi and its surrounding areas and to evaluate potential pollution sources. The concentrations of 33 heavy metal elements in 27 samples of Barbula Indica moss in the investigated region collected in December of 2017 in the investigated area have been examined using multivariate statistical analysis. Five factors explaining 80% of the total variance were identified and their potential sources have been discussed.*

Keywords: statistical analysis; Hanoi; Barbula Indica moss; heavy metal elements.

Classification numbers: 82.33.Tb; 31.15.bt.

I. INTRODUCTION

Hanoi with its surrounding is one of the biggest industrial regions in Vietnam. There are many new manufacturing industries and processing plants in this region, which give significant contribution to the economic growth. Furthermore, this region is considered to have the fastest urbanization in the country. The consequence of the phenomena mentioned above is that the environment including the air is polluted. The pollution caused by heavy metals over Hanoi and its surrounding area can bring adverse effects to living organisms and human [1, 2]. Therefore, it is necessary to monitor and evaluate the level of pollution in this region for the benefit of the local residents and related communities. In addition, the pollution of heavy metal elements can also affect crops and soil content. So far, there have been no data on heavy metal pollution in the air of Hanoi region. In order to overcome this lack, a joint project for studying heavy metal pollution in the air of Hanoi region using Barbula Indica moss has been created and carried out by an international group of the Institute of Physics of Vietnam Academy for Science and Technology and the Joint Institute for Nuclear Research in Dubna, Russia. The first investigation of heavy metal air pollution was carried out in the south of Vietnam (Hue, Hoi An and Ho Chi Minh city) using *Barbula Indica* moss and neutron activation analysis [3]. Following the first investigation, another study has been carried out in Hanoi region. Twenty seven *Barbula Indica* moss samples were collected in December of 2017 at 27 different locations in Hanoi and its surrounding region. The collected moss samples were prepared and analyzed directly by neutron activation analysis using IBR-2 nuclear reactor of the Laboratory of Neutron Physics of the Joint Institute for Nuclear Research in Dubna (Russia) [4]. A table of concentration of 33 heavy metal elements including Na, Mg, Al, Cl, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Ni. Co, Zn, As, Se, Br, Cd, Sb, Ba, Cs. La. Ce, Sm, Gd, Tb, Yb, Hf, Ta, Th and U of 27 collected moss samples was obtained. The errors for analysis for almost all of the elements are less than 10%. The statistical analysis applied to this data set will be described below. In order to have some ideas about the possible sources of pollution for the selected heavy metal elements, one should apply the statistical analysis including principal component analysis to these data. This is a main objective of our study.

II. STATISTICAL ANALYSIS OF THE DATA OF HEAVY METAL ELEMENTS IN THE MOSSES SAMPLES

II.1. Descriptive Statistical Analysis

Data evaluation such as descriptive statistics, correlation analysis and factor analysis were applied to our concentration data in order to infer the spatial distribution of elements in the analyzed moss samples and to find the possible sources of the pollutants. Descriptive statistics of metal concentration data including mean, median, standard deviation, minimum, maximum, range, coefficient of variation in percent (CV=standard deviation/mean ×100%), kurtosis and skewness determined in moss samples in Hanoi were calculated using the IBM SPSS software version 20. The result of descriptive statistics for our data is listed in Table 1. In this table, skewness and kurtosis are also calculated using the Shapiro-Wilk's test (0.05 significance level) for computing the uniformity of the distribution.

From Table 1, it can be seen that in Hanoi, the descending order of the mean concentration of the elements in moss samples is: Ca $>$ K $>$ Al $>$ Mg $>$ Fe $>$ Cl $>$ Na $>$ Ti $>$ Zn $>$ Mn $>$ Ba $>$ Sr $>$ Cr $>$ V $>$ Br $>$ Ce $>$ Ni $>$ La $>$ As $>$ Co $>$ Sb $>$ Th $>$ Sc $>$ Cs $>$ Cd $>$ U $>$ Gd $>$ Hf

Table 1. Descriptive statistics

	El.	Min	Max	Mean	Median	SD	CV (%)	Kurt	Skew
1	Na	510.0	1830	922.3	816.0	339.6	36.82	0.26	0.94
2	Mg	2510	8160	4417	4080	1443	32.68	0.17	0.86
3	Al	1340	14700	6767	5560	3439	50.83	0.13	0.97
4	Cl	485.0	3620	2132	2300	926.3	43.45	-1.21	-0.18
5	K	6450	12200	9570	9780	1560	16.30	-0.87	-0.17
6	Ca	13500	62600	23544	20800	9838	41.79	9.31	2.76
7	Sc	0.20	2.46	1.16	0.97	0.61	52.42	-0.57	0.69
8	Ti	75.50	967.0	465.4	428.0	229.4	49.30	-0.16	0.74
9	V	7.54	39.10	17.28	15.50	7.50	43.32	1.53	1.24
10	Cr	7.09	46.30	20.81	17.90	10.65	51.18	0.18	1.04
11	Mn	63.30	147.8	147.8	123.0	80.75	54.63	10.00	2.68
12	Fe	1740	9060	4400	4000	2058	46.78	-0.49	0.66
13	Ni	1.90	15.50	7.08	6.13	3.18	44.87	0.51	0.85
14	Co	0.35	6.68	1.93	1.57	1.24	64.09	7.65	2.33
15	Zn	68.20	1900	457.7	306.0	432.4	94.46	4.20	2.04
16	As	1.39	7.83	3.19	3.00	1.41	44.14	3.20	1.39
17	Se	0.18	0.68	0.32	0.30	0.11	35.00	2.51	1.36
18	Br	4.29	16.10	8.46	7.99	2.87	33.82	1.27	1.08
19	Sr	21.50	131.00	49.26	38.40	28.32	57.49	2.60	1.79
20	Cd	0.29	3.21	0.88	0.70	0.62	69.71	6.85	2.13
21	Sb	0.41	3.28	1.45	1.39	0.67	46.63	0.73	0.84
22	Ba	21.20	626.0	107.8	59.50	149.2	138.4	9.45	3.19
23	Cs	0.53	2.04	1.16	1.12	0.39	33.84	0.21	0.81
24	La	1.03	8.42	4.14	3.41	2.03	49.16	0.03	0.91
25	Ce	2.15	17.30	8.13	6.45	4.01	49.31	0.06	0.92
26	Sm	0.15	1.22	0.58	0.48	0.29	49.30	-0.03	0.86
27	Gd	0.31	1.67	0.73	0.72	0.33	45.28	1.13	1.01
28	Tb	0.02	0.16	0.08	0.07	0.04	47.62	-0.30	0.76
29	Yb	0.13	0,84	0.33	0.27	0.16	46.65	3.32	1.60
30	Hf	0.27	1.48	0.68	0.62	0.35	51.92	-0.48	0.78
31	Ta	0.02	2.46	0.22	0.12	0.45	210.0	25.84	5.04
32	Th	0.35	2.74	1.35	1.09	0.70	51.75	-0.46	0.78
33	U	0.32	2.02	0.85	0.78	0.45	25.84	0.85	1.15

Min: Minimum; Max: Maximum; SD: Standard deviation; CV: Coefficient of variation (SD/Mean×100); Kurt: kurtosis; Skew: skewness

Table 2. Pearson’s correlation coefficients of heavy metal elements in the collected moss samples in Hanoi region.

	Mg	**Al**	**Cl**	**K**	**V**	**Cr**	**Mn**	**Fe**	**Ni**	**Co**	**Zn**	**As**	**Br**	**Cd**	**Sb**	**Ba**	**La**	**Ce**	**Gd**	**Tb**	**Yb**	**Hf**	**Th**
Mg	1	**0.79**	-0.39	0.08	**0.8**	0.37	**0.65**	**0.81**	**0.66**	**0.81**	-0.07	**0.69**	0.15	0.43	**0.68**	**0.56**	**0.79**	**0.79**	**0.76**	**0.81**	**0.83**	**0.85**	**0.81**
Al	**0.79**	1	-0.36	0.14	**0.84**	0.39	**0.6**	**0.87**	**0.6**	**0.79**	0.07	**0.75**	0.06	0.26	**0.57**	0.36	**0.97**	**0.97**	**0.91**	**0.97**	**0.88**	**0.88**	**0.98**
Cl	-0.39	-0.36	**1**	**0.6**	-0.37	-0.23	-0.34	-0.5	-0.43	-0.47	-0.1	-0.44	-0.01	-0.27	-0.47	-0.48	-0.39	-0.39	-0.44	-0.43	-0.37	-0.46	-0.43
K	0.08	0.14	**0.6**	**1**	0.09	-0.14	-0.03	-0.01	-0.23	0	-0.06	0	-0.02	-0.01	-0.11	-0.25	0.06	0.08	-0.07	0.06	0.05	0.14	0.08
V	**0.8**	**0.84**	-0.37	0.09	**1**	0.38	0.49	**0.84**	**0.6**	**0.69**	-0.08	**0.87**	-0.01	0.24	**0.62**	0.31	**0.79**	**0.78**	**0.83**	**0.83**	**0.72**	**0.8**	**0.83**
Cr	0.37	0.39	-0.23	-0.14	0.38	**1**	0.43	0.38	**0.8**	**0.51**	0.25	0.43	0	**0.61**	**0.62**	0.35	0.4	0.39	0.39	0.4	**0.51**	0.39	0.37
Mn	**0.65**	**0.6**	-0.34	-0.03	0.49	0.43	**1**	**0.73**	**0.52**	**0.89**	0.11	0.46	0.12	**0.64**	**0.5**	**0.76**	**0.6**	**0.6**	**0.55**	**0.64**	**0.82**	**0.64**	**0.58**
Fe	**0.81**	**0.87**	-0.5	-0.01	**0.84**	0.38	**0.73**	**1**	**0.66**	**0.88**	0.07	**0.79**	-0.11	0.38	**0.62**	**0.61**	**0.87**	**0.87**	**0.84**	**0.91**	**0.84**	**0.87**	**0.9**
Ni	**0.66**	**0.6**	-0.43	-0.23	**0.6**	**0.8**	**0.52**	**0.66**	**1**	**0.68**	0.25	**0.67**	0.03	**0.56**	**0.86**	**0.54**	**0.66**	**0.64**	**0.65**	**0.66**	**0.69**	**0.69**	**0.64**
Co	**0.81**	**0.79**	-0.47	0	**0.69**	**0.51**	**0.89**	**0.88**	**0.68**	**1**	0.12	**0.7**	0.07	**0.64**	**0.67**	**0.75**	**0.83**	**0.82**	**0.78**	**0.86**	**0.94**	**0.84**	**0.81**
Zn	-0.07	0.07	-0.1	-0.06	-0.08	0.25	0.11	0.07	0.25	0.12	**1**	0.14	-0.11	0.38	0.43	0.26	0.12	0.1	0.07	0.12	0.13	0.11	0.09
As	**0.69**	**0.75**	-0.44	0	**0.87**	0.43	0.46	**0.79**	**0.67**	**0.7**	0.14	**1**	0.05	0.37	**0.73**	0.39	**0.77**	**0.74**	**0.83**	**0.79**	**0.68**	**0.82**	**0.78**
Br	0.15	0.06	-0.01	-0.02	-0.01	0	0.12	-0.11	0.03	0.07	-0.11	0.05	**1**	0.01	0	0.03	0.07	0.08	0.16	0.04	0.05	0.05	0.04
Cd	0.43	0.26	-0.27	-0.01	0.24	**0.61**	**0.64**	0.38	**0.56**	**0.64**	0.38	0.37	0.01	**1**	**0.67**	**0.54**	0.3	0.29	0.25	0.33	**0.57**	0.46	0.28
Sb	**0.68**	**0.57**	-0.47	-0.11	**0.62**	**0.62**	**0.5**	**0.62**	**0.86**	**0.67**	0.43	**0.73**	0	**0.67**	**1**	**0.54**	**0.61**	**0.58**	**0.61**	**0.63**	**0.67**	**0.72**	**0.62**
Ba	**0.56**	0.36	-0.48	-0.25	0.31	0.35	**0.76**	**0.61**	**0.54**	**0.75**	0.26	0.39	0.03	**0.54**	**0.54**	**1**	0.43	0.4	0.41	0.48	**0.59**	0.45	0.41
La	**0.79**	**0.97**	-0.39	0.06	**0.79**	0.4	**0.6**	**0.87**	**0.66**	**0.83**	0.12	**0.77**	0.07	0.3	**0.61**	0.43	**1**	**0.99**	**0.94**	**0.99**	**0.9**	**0.89**	**0.98**
Ce	**0.79**	**0.97**	-0.39	0.08	**0.78**	0.39	**0.6**	**0.87**	**0.64**	**0.82**	0.1	**0.74**	0.08	0.29	**0.58**	0.4	**0.99**	**1**	**0.93**	**0.99**	**0.9**	**0.9**	**0.98**
Gd	**0.76**	**0.91**	-0.44	-0.07	**0.83**	0.39	**0.55**	**0.84**	**0.65**	**0.78**	0.07	**0.83**	0.16	0.25	**0.61**	0.41	**0.94**	**0.93**	**1**	**0.94**	**0.84**	**0.83**	**0.93**
Tb	**0.81**	**0.97**	-0.43	0.06	**0.83**	0.4	**0.64**	**0.91**	**0.66**	**0.86**	0.12	**0.79**	0.04	0.33	**0.63**	0.48	**0.99**	**0.99**	**0.94**	**1**	**0.91**	**0.91**	**0.99**
Yb	**0.83**	**0.88**	-0.37	0.05	**0.72**	**0.51**	**0.82**	**0.84**	**0.69**	**0.94**	0.13	**0.68**	0.05	**0.57**	**0.67**	**0.59**	**0.9**	**0.9**	**0.84**	**0.91**	**1**	**0.87**	**0.88**
Hf	**0.85**	**0.88**	-0.46	0.14	**0.8**	0.39	**0.64**	**0.87**	**0.69**	**0.84**	0.11	**0.82**	0.05	0.46	**0.72**	0.45	**0.89**	**0.9**	**0.83**	**0.91**	**0.87**	**1**	**0.91**
Th	**0.81**	**0.98**	-0.43	0.08	**0.83**	0.37	**0.58**	**0.9**	**0.64**	**0.81**	0.09	**0.78**	0.04	0.28	**0.62**	0.41	**0.98**	**0.98**	**0.93**	**0.99**	**0.88**	**0.91**	**1**

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

> Sm > Yb > Se > Ta > Tb. The highest concentrations belong to Ca, K, Al, Mg, Fe, Cl, Na and Ti. These elements are most abundant elements in the crust. The lowest concentrations belong to Ta and Tb. This reflects that the density of dust in the air is very high and air pollution in Hanoi and its surrounding region is seriously caused by windblown soil dust.

It can be seen from Table 1 that all of the heavy metals under investigation in the investigated region show strong variation in concentration, with the coefficients of variation (CV) ranging from 16.3% to 210%. High values of the coefficient of variation are likely to indicate the influence of complicated origins of these elements in mosses. Furthermore, for those elements whose value of skewness is in the range from -0.8 to 0.8 and the value of its kurtosis is in the range from -3.0 to 3.0 then its concentration can be considered to be normally distributed. According to the values of skewness and kurtosis listed in Table 1, only the concentrations of 8 elements including Hf, Th, Tb, Ti, Sc, Fe, K and Cl follow normal distribution. In addition, the coefficient of variation of these elements are smaller than 25%. It means that these elements may have the same pollution sources. For other elements, the source of pollution may be very complicated. The obtained concentration data of heavy metal elements in the moss samples were subjected to multivariate statistical analysis.

The pre-study of the obtained concentrations had been done to select out twenty three elements used for factor analysis including Mg, Al, Cl, K, V, Cr, Mn, Fe, Ni, Co, Zn, As. Br, Cd, Sb, Ba, La, Ce, Gd, Tb, Yb, Hf and Th. The elements of interest, i.e., those that join together into groups (or factors), represent possible sources of pollution in the studied areas, and the number of explained variance of each group is as high as possible. Therefore, in the further statistical analysis to be described below, only these twenty elements will be included.

II.2. Correlation coefficient analysis

In order to establish the inter-elemental relationships and trace their sources in the moss samples, Pearson's correlation coefficients were executed and presented in Table 2. It can be seen clearly from Table 2 that there are several groups in which the elements are significantly correlated with each other, which may suggest that these elements may have a common source. For illustration, we list some groups in which the elements are strongly correlated: Magnesium is correlated with Al (0.79), V (0.8), As (0.69), Sb (0.68), Ba (0.56), La (0.79), Ce (0.79), Gd (0.76), Tb (0.81), Hf (0.85) and Tb (0.81); Aluminium is correlated with V (0.84), Fe (0.87), Co(0.79), As (0.75), La (0.97), Ce (0.97) Gd (0.91), Tb (0.97), Yb (0.88), Hf (0.88) and Th (0.98); Vanadium is correlated with Fe (0.84), Ni (0.6), Co (0.69), As (0.87), Sb (0.62), La (0.79), Ce (0.78), Gd (0.83), Tb (0.83), Yb (0.72), Hf (0.8) and Th (0.8); Chromium is correlated with Ni (0.8), Co (0.51), Cd (0.61), Sb (0.62) and Yb (0.51); Chlorine is correlated only with K (0.6).

II.3. Factor analysis for source apportionment of heavy metals in the moss samples

Factor analysis is a useful tool and has been applied successfully in different areas of environmental research because of their ability to process large data sets. In our case, the main purpose of its application is to reduce the dimensionality of the concentration data of elements of the moss samples and to find the possible sources of pollution of these elements [5,6]. Factor analysis transforms the original set of inter-correlated variables into a set of uncorrelated variables, latent factors that are linear combinations of the original variables. After factor analysis, several latent factors will be obtained and each of them will be some linear combination of several original variables (in our case, the concentration of some detected elements in the moss samples). The obtained factors are ordered such that the first few significant factors retain most of the variation present in all of the original variables. Therefore, by using factor analysis, only a small number of significant factors are needed to represent most of the information in a lager set of the concentration data. The factor analysis technique and underlying mathematical theory have been described in detail in a number of publications and books.

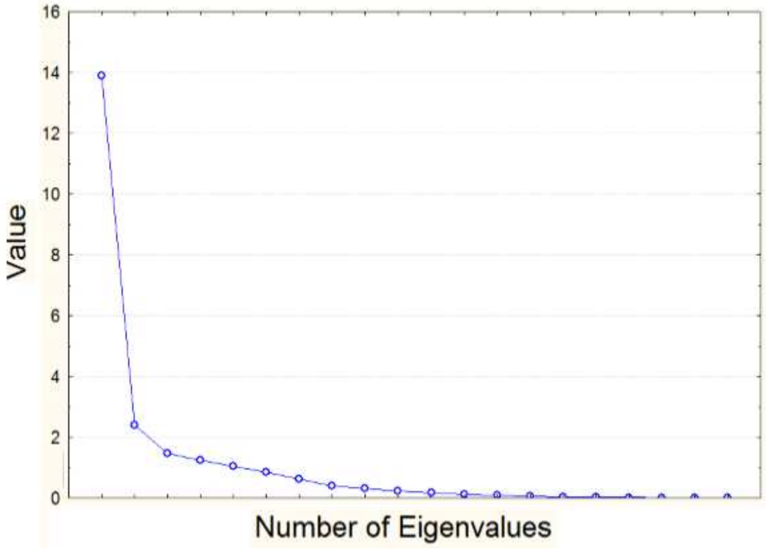

Fig. 1. Scree plot.

Factor analysis has been applied to our data set by using STATISTICA-8 software to assist in the identification of sources of the pollutants. Twenty three variables (concentration of 23 elements mentioned above) and 27 moss samples were chosen for factor analysis. In our study, the concentrations of elements vary by different orders of magnitude, therefore factor analysis was applied to the correlation matrix and each variable is normalized to unit variance. For easier interpretation of the obtained results, we applied factor analysis with varimax normalized rotation so that the variances of the factor loadings across variables for each factor will be maximized. The scree plot (see Fig. 1) confirms the choice of 5 factors. Table 3 shows the varimax-rotated components matrix for the concentration data in the moss samples in which the loadings and explained variance of the first five factors are listed. The first five factors of this study accounted for 87.16 % of the total variance in the data set namely Factor-1, Factor-2, Factor-3, Factor-4 and Factor-5 accounted for 60.36 %, 10.44 %, 6.40 %, 5.44 % and 4.51 %, respectively.

It should be noted from this table that loadings equal to or larger than 0.4 were written in bold and loadings less than 0.4 may be dominated by random errors. The factor scores of the sampling sites are listed in Table 4. Below is our interpretation of five factors.

Factor-1 explains 60.36 % of the total variance. It has high contribution to Th, Al, La, Ce, Tb, Gd, V, Hf, Fe, As, Mg, Yb, Co, moderate contribution to Ni, Sb, Mn with the factor loadings of 0.96, 0.95, 0.94, 0.94, 0.94, 0.92, 0.89, 0.88, 0.85, 0.81, 0.78, 0.78, 0.7, 0.56, 0.53 and 0.44. Aluminum, Fe, Mg, and Mn are the most common chemical elements in the crust and

Table 3. The varimax-rotated components matrix for the concentration data in the moss samples. Loadings and explained variance of the first five factors are listed.

	Element	Factor-1	Factor-2	Factor-3	Factor-4	Factor-5
1	Mg	**0.78**	0.13	0.03	0.39	0.17
2	Al	**0.95**	0.10	-0.04	0.17	0.02
3	Cl	-0.35	-0.10	**-0.80**	-0.21	0.05
4	K	0.15	-0.08	**-0.93**	-0.02	-0.05
5	V	**0.89**	0.12	0.06	0.06	0.01
6	Cr	0.27	**0.76**	0.08	0.17	0.15
7	Mn	**0.44**	0.16	0.03	**0.83**	0.08
8	Fe	**0.85**	0.10	0.14	**0.40**	-0.16
9	Ni	**0.56**	**0.67**	0.25	0.19	0.11
10	Co	**0.70**	0.23	0.07	**0.65**	0.04
11	Zn	-0.06	**0.67**	0.00	0.04	-0.30
12	As	**0.81**	0.33	0.14	0.05	0.02
13	Br	0.03	-0.05	0.01	0.04	**0.95**
14	Cd	0.11	**0.67**	-0.05	**0.61**	0.03
15	Sb	**0.53**	**0.71**	0.18	0.21	0.00
16	Ba	0.24	0.23	0.31	**0.79**	-0.04
17	La	**0.94**	0.14	0.03	0.20	0.02
18	Ce	**0.94**	0.12	0.01	0.19	0.03
19	Gd	**0.92**	0.13	0.16	0.12	0.12
20	Tb	**0.94**	0.14	0.05	0.24	-0.02
21	Yb	**0.78**	0.24	-0.02	**0.50**	0.04
22	Hf	**0.88**	0.23	0.00	0.28	0.01
23	Th	**0.96**	0.12	0.04	0.18	-0.02
	Eigenvalues	13.88	2.40	1.47	1.25	1.04
	% of variance	60.36	10.44	6.40	5.44	4.51
	Cumulative %	60.36	70.80	77.21	82.64	87.16

they are good indicator of crust-related dust. In addition, Mn, Mg, and Al are considered tracers of natural soil [7]. In Hanoi the crustal matter source contributions are likely to be a made up of windblown soil, road dust and dust generated by construction and road works. In Hanoi road works, construction of new office blocks, apartment buildings, and building refurbishments can be seen everywhere, especially in the surrounding areas of the city and, as a result, dust emission is increased quickly. The crustal matter particles can be generated by re-entrainment into the air due to wind action or by the turbulent action of vehicles passing across road surfaces [9]. These activities were carried out near the sampling sites during the period, therefore, they also should be responsible for this factor. Based on these explanations, it may be said that factor-1 represents road dust and crustal matter. This factor showed highest contributions from sites relatively close roadsides, namely M8 (Co Nhue, Tu Liem), M14 (Kieu Ky, Gia Lam), M23 (Nhu Quynh, Hung Yen), M7 (Nhan My, Tu Liem), M16 (Nhu Quynh, Hung Yen), M17 (Yen My, Hung Yen), M25 (My Hao, Hung Yen), M18 (Yen My, Hung Yen), M24 (Ngoc Lang, My Hao, Hung Yen), M1

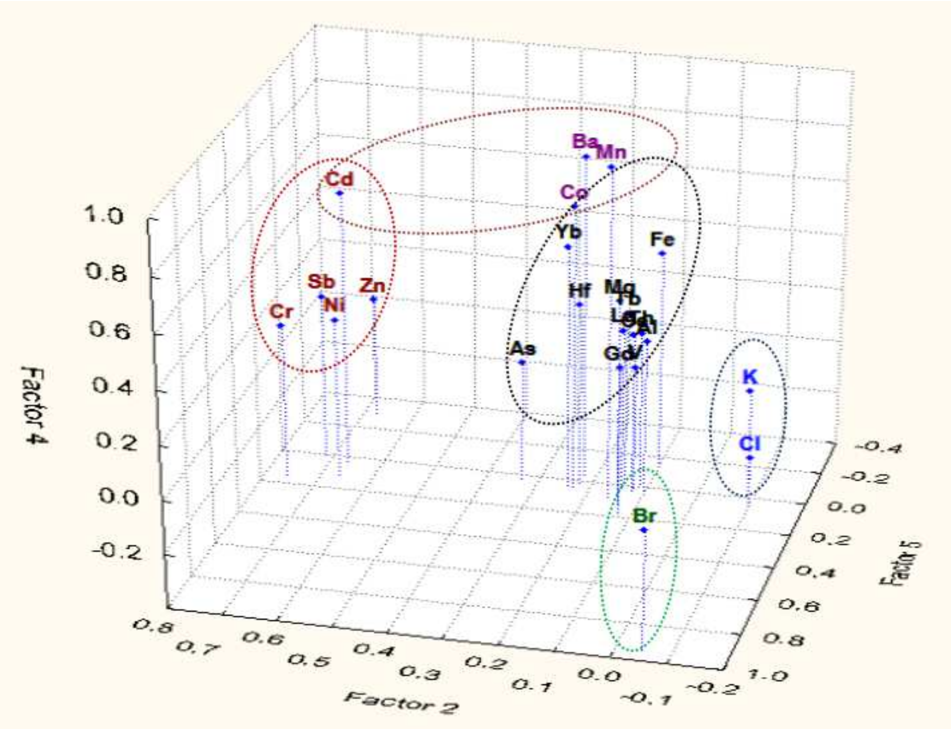

Fig. 2. 3D-plot of scores for investigated elements obtained from PCA results in the space generated by the factor-1, factor 2 and factor 5.The trace elements grouped by circle are considered to be a common source.

(Doc La, Bac Ninh), M2 (Hoan Son, Bac Ninh), M27 (Pham Kham, Van Lam, Hung Yen)), M10 (Yen Lac, Thach That, Hanoi), M22 (Xuan Dao, My Hao, Hung Yen) and M15 (Tan Quang, Van Lam, Hung Yen) with the absolute factor scores of 2.65, 1.78, 1.78, 1.46, 1.16, 1.08, 1.01, 0.95, 0.92, 0.84, 0.84, 0.75, 0.68, 0.62, and 0.61, respectively. These are suburban areas of Hanoi and in these areas, many new offices, resident buildings and new roads are under construction.

Factor-2 explains 10.44% of the total variance and it consists of Cr, Sb, Ni, Zn, Cd and As with the factor loadings of 0.76, 0.71, 0.67, 0.67, 0.67 and 0.33, respectively. These elements may be related to industrial activities such as alloy production, metal-related manufacturing, production of batteries, ... The locations having absolute high score value of factor-2 are M21(My Hao, Hung Yen), M4 (Tay Ho, Hanoi), M2 (Hoan Son, Bac Ninh), M16 (Van Lam, Hung Yen), M17 (Yen My, Hung Yen), M11 (Gia Lam, Hanoi), M26 (My Hao, Hung Yen), M24 (My Hao, Hung Yen), M27 (Van Lam, Hung Yen), M19 (Yen My, Hung Yen), M1 (Doc La, Bac Ninh), M12 (Dong Anh, Hanoi), M25 (My Hao, Hung Yen), M10 (Thach That, Hanoi), M23 (Van Lam, Hung Yen), M14 (Gia Lam, Hanoi) with the factor scores of 2.69, 1.90, 1.79, 1.25, 1.18, 1.12, 1.08, 0.99, 0.90, 0.80, 0.79, 0.78, 0.68, 0.67, 0.65 and 0.62, respectively. These areas are close to many industrial parks producing electronics, consumer electronics and there are many different kinds of metal recycling plants. Therefore, it may be concluded that this factor is partly related to industrial activities. Furthermore, this factor may be related to motor vehicles in the local roads which will

Table 4. Factor scores of the sampling sites.

	Factor-1	Factor-2	Factor-3	Factor -4	Factor - 5
M1	**0.84**	**0.79**	-0.48	**4.45**	**0.46**
M2	-0.84	**1.79**	**-1.46**	-0.58	-1.58
M3	**0.53**	-0.37	**-0.60**	0.01	0.14
M4	**0.43**	**1.90**	1.43	-0.84	**0.53**
M5	-0.56	0.06	0.18	0.09	0.34
M6	-0.46	-0.09	1.27	-0.23	0.13
M7	**1.46**	-0.34	0.35	-0.10	-0.52
M8	**2.65**	-0.36	0.43	-1.05	0.17
M9	-0.20	0.30	2.32	**1.34**	-0.82
M10	**0.68**	-0.67	-0.16	0.08	-1.29
M11	-0.05	-1.12	1.34	-0.46	**2.34**
M12	0.02	-0.78	1.72	-0.15	-1.66
M13	-0.51	-0.38	**-1.10**	-0.50	0.23
M14	**1.78**	-0.62	**-1.22**	-0.66	-1.20
M15	**0.61**	-0.01	**-1.48**	-0.02	**0.85**
M16	-1.16	-1.25	-0.27	0.01	-0.36
M17	-1.08	**1.18**	0.33	-0.35	**0.46**
M18	-0.95	0.39	-0.20	-0.27	-0.37
M19	-0.53	**0.80**	0.75	-0.44	0.24
M20	-0.40	-0.48	**-0.58**	-0.21	**1.24**
M21	-0.27	**2.69**	**-0.63**	-0.50	-0.28
M22	-0.62	-0.43	**-1.00**	0.05	**2.26**
M23	**1.78**	**0.65**	-0.08	-0.25	**0.83**
M24	-0.92	-0.99	**-0.90**	0.32	-0.41
M25	-1.01	-0.68	0.68	-0.23	0.08
M26	-0.45	-1.08	**-0.68**	-0.10	-0.90
M27	-0.75	-0.90	0.05	**0.60**	-0.92

contribute to Pb, Cr, Cu, Cd and Zn [8]. Although we do not have the concentration data of Pb and Cu because they are difficult to analyze by neutron activation analysis, but the presence of other elements which are related to motor vehicles (Cr, Cd and Zn) allows us to make the assumption that Factor-2 is partially related to motor vehicles. It is said that zinc can be emitted by vehicles through the combustion of lubricant oils as it is added during lubricant formulation [9, 10]. Cadmium is widely used as good marker of motor vehicle source category [11].

Factor-3 explains 6.40% of the total variance and it explains the highest percentage of variability for K, Cl and to some minor extent for Ba with the factor loadings of 0.93, 0.80 and 0.31, respectively. Presence of K, Cl and Ba relates this factor to the source of agriculture soil. This factor showed highest contributions from sites M9 (Yen Vien, Gia Lam), M12 (Mai Lam, Dong Anh), M15 (Van Lam, Hung Yen)), M2 (Hoan Son, Bac Ninh), M4 (Yen Phu, Tay Ho), M11 (Xuan Duc, Gia Lam), M6 (Kim Quan, Gia Lam), M14 (Gia Lam, Hanoi), M13 (Van Giang, Hung Yen), M24 (My Hao, Hung Yen), M19 (Yen My, Hung Yen), M25 (My Hao, Hung Yen), M26 (My Hao, Hung Yen) and M21 (My Hao, Hung Yen) with the absolute factor scores 2.32, 1.72, 1.48, 1.46, 1.43, 1.34, 1.27, 1.22, 1.10, 0.90, 0.75, 0.68, 0.68 and 0.63, respectively. The sampling sites were close to the rice and vegetable growing areas of the farmers.

Factor-4 explains 5.44% of the total variance and it is mainly dominated by Mn, Ba, Co and Cd with the factor loadings of 0.83, 0.79, 0.65 and 0.61, respectively. According to G. P.Vukovic [12], this factor may be related to traffic sources. In urban region, road transport is among the dominant sources of airborne trace elements. According to [12], four main traffic-related sources are: (a) the diesel or gasoline fuel combustion; (b) the lubricant oils; (c) the engine wear or abrasion of system; and (d) non-tail pipe emissions from tyre wear, barake wear and possibly from road abrasion.

The source of Mn and Ba is from diesel gasoline emissions while the source of Cd is from lubricating oil combustion and Co is from tyre wear. Therefore, we conclude that factor-4 is related to exhaust traffic sources.

Factor-4 showed highest contributions from sites M1 (Doc La, Bac Ninh), M9 (Gia Lam, Hanoi), M8 (Tu Liem, Hanoi), M4 (Tay Ho, Hanoi), M14 (Gia Lam, Hanoi) and M27 (Van Lam, Hung Yen) with the absolute factor scores of 4.45, 1.34, 1.05, 0.84, 0.66 and 0.6, respectively. These places are located near the West Lake (the biggest lake in Hanoi) and the famous Red River.

Factor-5 explains 4.51% of the total variance and it consists only of Br and some minor of Zn with the factor loadings of 0.95 and 0.30, respectively. Bromine has several strong natural and anthropogenic sources in the environment [13, 14]. The most possible natural sources of Br in the environment are from oceans, salt marshes, fungi. The anthropogenic sources of Br in the environment might be from biomass burning, automotive emissions, pesticide application, Br chemical manufacturing, coal burning, and PVC usage and disposal. Factor-5 showed highest contributions from sites M11 (Gia Lam, Hanoi), M22 (My Hao, Hung Yen), M12 (Dong Anh, Hanoi), M2 (Hoan Son, Bac Ninh), M10 (Thach That, Hanoi), M20 (Yen My, Hung Yen), M14 (Gia Lam, Hanoi), M27 (My Hao, Hung Yen), M26 (Van Lam, Hung Yen), M15 (Dong Anh, Hanoi), M23 (Yen My, Hung Yen) and M9 (Kim Quan, Gia Lam) with the absolute factor scores of 2.34, 2.26, 1.66, 1.58, 1.29, 1.24, 1.20, 0.92, 0.90, 0.85, 0.83 and 0.82, respectively. These places are located on outskirts of Hanoi having many rice and different vegetable fields, small private factories, high traffic volume and high population density. According to this list of the places having high factor scores, it may be deduced that factor-5 is related to anthropogenic sources, namely, biomass and coal burning, pesticide application and PVC usage and disposal.

III. CONCLUSIONS

In this study, statistical analysis has been applied to our obtained data of heavy elemental concentrations in the moss samples collected at different locations of Hanoi and surrounding areas. The main purpose of this study is to find the possible sources of heavy metal pollutants in the air in the investigated areas. The different statistical methods including descriptive analysis, correlation analysis and factor analysis have been used. Among 33 heavy metal elements determined by neutron activation analysis, we focused only on 23 elements including Mg, Al, Cl, K, V, Cr, Mn, Fe, Ni, Co, Zn, As. Br, Cd, Sb, Ba, La, Ce, Gd, Tb, Yb, Hf and Th. According to our obtained results, five factors have been obtained which can explain 87.16 % of the total variance in the data set. Factor-1 explains 60.36 % of the total variance. It has high contribution to Th, Al, La, Ce, Tb, Gd, V, Hf, Fe, As, Mg, Yb, Co, moderate contribution to Ni, Sb, Mn. This factor represents road dust and crustal matter. Factor-2 consists of Cr, Sb, Ni, Zn, Cd and As and explains 10.44% of the total variance. This factor is related to industrial activities. Factor-3 explains 6.40% of the total variance and it explains the highest percentage of variability for K, Cl and to some minor extent for Ba. This factor is related to the source of agriculture soil. Factor-4 explains 5.44% of the total variance and it consists of Mn, Ba, Co and Cd. This factor is related to exhaust traffic sources. Factor-5 explains 4.51 % of the total variance and it consists only of Br and some minor of Zn. This factor is related to anthropogenic sources, namely, biomass and coal burning, pesticide application and PVC usage and disposal.

ACKNOWLEDGMENT

This work was funded by the Vietnam Ministry of Science and Technology, Grant number NĐT.25.RU/17. In addition, N. N. Mai acknowledges the Institute of Physics for support under the program of supporting young researchers in 2019.

REFERENCES

[1] A.H.Tuan, N.X.Chu, and T.V.Tran, *International Journal of Scientific & Technology Research*. **6** (2017) 249–253.

[2] ICEM, 2007: http://icem.com.au/documents/envassessment/wb_cea/WB_CEA.pdf

[3] T.T.Doan Phan, T.T.M.Trinh, L.H.Khiem, M.V.Frontasyeva, and N.H.Quyet, *Asia-Pacific Journal of Atmospheric Sciences*. **55** (2019) 247-253.

[4] M. V. Frontasyeva, *Neutron News*. **16** (2005) 24–27.

[5] T. Berg, O. Royset, E. Steinne, and M. Vads, *Environmental Pollution*. **88** (1995) 67-77.

[6] C.D.Opera and Al.Mihul, *Romanian Reports in Physics*. **55** (2003) 91-110.

[7] I. Gupta, A. Salunkhe, and R. Kumar, *Indian Scientific World Journal* (2012).

[8] L.K. Boamponsem, C.R. Freitas, and D. Williams, *Atmospheric Pollution Research*. **8** (2017) 101-113.

[9] P.K. Davy, W.J. Trompetter, and A. Markwitz, *GNS Science, Wellington* (2011).

[10] B.A. Begum, S.K. Biswas, M.Nasiruddin, A.M. Showkot Hossain, and P.K. Hopke, *Environ. Eng. Sci*. **26** (3) (2009) 679-689.

[11] M.H. Sowlat, K. Naddafi, M. Yunesian, P.L. Jackson, and A. Shahsavani, *Environ. Contam. Toxicol*. **88** (5) (2012) 735-740.

[12] G.P. Vukovic, *Doctoral dissertation.University of Belgrade. Belgrade* (2015).

[13] D.C. Dowdell, G.P. Matthews, and I. Wells, *Atmospheric Environment*. **28** (1994) 1989–1999

[14] G.J. Gribble, *Environmental Science Pollution Research*. **7** (2000) 37–49.