VERSITA DOI: 10.2478/eces-2013-0006 ECOL CHEM ENG S. 2013;20(1):79-91

Celina PIESZKO[1] and Marcin SAJDAK[1*]

CHEMOMETRIC ANALYSIS OF HEAVY METALS COEXISTENCE IN DIFFERENT PLANT MATRICES IN THE UPPER SILESIAN INDUSTRIAL DISTRICT

ANALIZA CHEMOMETRYCZNA METALI CIĘŻKICH POCHODZĄCYCH Z RÓŻNYCH MATRYC ROŚLINNYCH Z TERENÓW GÓRNOŚLĄSKIEGO OKRĘGU PRZEMYSŁOWEGO

Abstract: Contents of Zn, Cd, Ni, Mn, Cu, Pb and Fe in 36 samples of Scots pine collected in 9 sites located in the Upper Silesia Province have been determined using AAS technique. The samples were digested before the analysis. On the basis of cluster analysis the existence of links between the quantitative composition, the sampling sites and the type of the matrix have been found. The coexistence of certain groups of elements was found and described.

Keywords: Scots pine, litter of conifer needles, branches, heavy metals, cluster analysis, PCA, chemometrics

Introduction

Atomic absorption spectrometry with flame atomisation (FAAS) is a commonly used technique for the determination of metals content in environmental samples [1-3]. This method is quick and accurate, but the preparation of samples for analysis is often time-consuming and tedious process. Examination of samples taken from one control point or from big area requires each time carrying out sample preparation step like mineralisation. To reduce the number of such treatments and thus to reduce the required number of tests, the dependencies between concentrations of analysed metals in the studied matrices in a given area can be established and the chemometric techniques can be applied to find links existing between them [4-9]. This will result in reducing the required processing steps thanks to the knowledge of metal - matrix A - matrix B or matrix - metal A - metal B interactions. In the present paper the results of studies conducted in the Upper Silesia Province on analysing the contents of Zn, Cd, Ni, Mn, Cu, Pb, and Fe in the different components of pine tree, such as pine needles, twigs, pine cones, litter have been described. Several chemometric and statistical techniques like analysis of the similarities, k-means

[1] Department of Analytical Chemistry, Faculty of Chemistry, The Silesian University of Technology, ul. M. Strzody 7, 44-100 Gliwice, Poland, phone +48 509 628 015
[*]Corresponding author: msajdak@ichpw.zabrze.pl

method, correlation analysis or principal component analysis have been applied for establishing the useful links.

Materials and methods

Materials for investigation: 36 samples of cones [W], needles [I], branch [G] and litter [S] of Scots pine collected from nine sites in the province of Silesia (Szczekociny, Zarnowiec, Slawniow, Podlesice, Solca, Slawkow, Dabrowa Gornicza, Jaworzno, Sosnowiec).

Collected materials were dried to an air-dried state, ground in a mill equipped with a titanium blade and then dried at 110^{o}C. After dry mineralisation at 350^{o}C the resulting ash was transferred into a beaker and dissolved in 10% HNO_3. Solutions after an appropriate preconcentration were transferred to 10 cm^3 volumetric flasks. In all collected samples of plant material the content of the following elements: Zn, Cd, Ni, Mn, Cu, Pb, and Fe was determined by flame atomic absorption spectrometry. The measurements were made using a spectrometer AAS 3 Carl Zeiss Jena. The values of recoveries for the various materials were estimated based on the reference material spiked with the standard solutions.

Cluster analysis

Cluster analysis is an exploratory multivariate method that can be used to describe the relationship among variables. Several mathematical criteria can be used to examine the similarity between variables cases. For the *hierarchical clustering analysis* (HCA), Ward's method was used to obtain a cleaner plot of clusters. This method is different from all other methods because it uses an analysis-of-variance approach to evaluate the distance between clusters. This method attempts to minimise the *sum of squares* (SS) of any two clusters that can be formed each step. This method is regarded as a very efficient one. The joining of the tree clustering method using the dissimilarity distance between objects when using the clusters. Similarities are a set of rules that serves as criteria for grouping or separating items. The Euclidean distance was chosen for the analysis.

The purpose of cluster analysis is to determine the mutual similarities between the analysed objects and attributes that describe them. In this analysis, there is no information about the homogeneity of the analysed data set. The first step in clustering objects is evaluating their similarity (or dissimilarity): the distance or correlation coefficient can be used as a measure of (dis)similarity. One way of measuring the distance between two objects *i* and *j* in HCA is to use the Euclidean distance:

$$d_{ij} = \sqrt{\sum_{k=i}^{m} \left(x_{ik} - x_{jk}\right)^2} \tag{1}$$

where: m - the number of variables, d_{ij} - Euclidean distance between objects i and j, x_{ik} - value of the variable i, x_{jk} - value of the variable j.

Using vector notation, Eq. (1) becomes:

$$d_{ij}{}^2 = \left(x_i - x_j\right)^T \cdot \left(x_i - x_j\right) \tag{2}$$

where x_i and x_j are the column vectors of the two objects and T stands for "transpose".

The smallest value is the Euclidean distance, and the largest is the similarity between the objects. The Euclidean distance can be graphically interpreted as the length of the vector starting at *i* and ending at *j* [7-9].

Principal component analysis PCA

Multivariate statistical tools, such as the *principal component analysis* (PCA) technique, have been widely applied in the treatment of high-complexity data sets. The PCA technique extracts the eigenvalues and eigenvectors from the covariance matrix of the original variables. It makes it possible to find the association between variables, thus reducing the dimensionality of the data set. The *principal components* (PCs) are the uncorrelated (orthogonal) variables obtained by multiplying the original correlated variables with the eigenvector (loading or weighing). The eigenvalues of PCs are the measure of their associated variance. The participation of the original variables in the PCs is given by the loading, and the individual transformed observations are called scores. The use of correlated variables in this analysis gives the best results. The main task is detecting the internal data structure and describing it with the other parameters resulting from the above-mentioned structure. A characteristic feature of principal components is the ability to determine the main factors (PCs) in order of decreasing volatility of the stock. The measure of this stock is the eigenvalues. In some cases, it is possible to assign the main components and groups of objects to a certain chemical or physical interpretation [7-9].

The principal component analysis technique is used primarily for:

1. The reduction of the data space,
2. The transformation of correlated input variables in the output main components, and
3. The graphical presentation of the structure of a multidimensional data set in the plane with minimal distortion of information.

Results and discussion

The contents of analysed metals in studied materials were found in the range of: 6.5-51.5 µg/g for zinc; 0.1-5.5 µg/g for lead; 5-97.2 µg/g for copper; 0.4-7.5 µg/g for nickel; 1.0-71.5 µg/g for manganese; 0.01-0.30 µg/g for cadmium and 1.0-75.2 µg/g for iron, respectively. These results were used to analyse the relationship between qualitative and quantitative composition of the materials with the aid of chemometric techniques CA, k-means and PCA.

Relationship matrix - metal content

In Figure 1 the dendrogram of cluster analysis for samples of different components of pine trees has been shown. Four main well-separated clusters can be distinguished, which are divided into smaller subgroups. The first cluster consists of two subgroups. The first one is characterized by high linear correlation between the content of cadmium in samples (0.03-0.13 µg/g) coming from different sampling sites. With such a high similarity between them, knowing the concentration of cadmium in the cones (Cd SZ) one can estimate the cadmium content in needles (Cd I) with 99% accuracy; in the litter (Cd S) with 96% accuracy; in the branches (Cd G) with 97% accuracy.

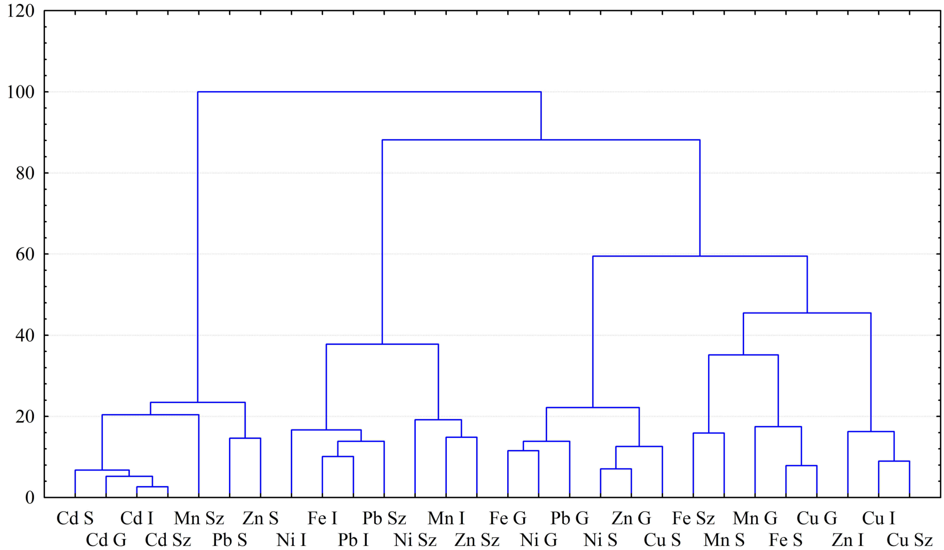

Fig. 1. Dendrogram of cluster analysis for the analysed samples
Abbreviations used:

- AA 00 XX
  - AA - studied element
  - 00 - sampling location
  - XX - type of sample

The cluster 2 consists of needles (I) and cones (SZ), in which nickel, lead and iron were analysed. The group: Pb SZ, Pb I is characterized by similar contents of lead in these components, which were found in the range of 1.6-3.5 µg/g. Moreover, the first subgroup is characterized by a high level of similarity to one another in the range of 83-86%, and describes the content of lead in the cones, pine needles in relation to the iron content in needles. Growth of lead content (Pb SZ, I) is correlated positively with iron content (Fe I) in the studied materials, according to the following formula given below and shown in Figure 1. The numerical values describing the axes are the values obtained after *standard normal variate* transformation (SNV) of raw data.

$$FeSZ = -0.1501 + 1.1018 PbI$$
$$R^2 = 0.865 \qquad (3)$$

In the cluster 3 two smaller subgroups are distinguished. The first one consists of pine branches. A characteristic feature of this subgroup is consistent nickel, lead and iron contents in the analysed materials, within the range of 1.1-2.9 µg/g, and a high correlation coefficient between iron (Fe D) and zinc (Zn G) contents. Thus knowing the iron content one can estimate zinc content with 82% accuracy. The second subgroup of cluster 3 consists of litter samples. A feature of this group is a positive correlation between nickel and copper concentration - an increase in copper content gives rise in nickel content in the examined material. This value can be calculated with 88% accuracy from the formula:

$$Cu\,S = -0.306 + 1.28 \cdot Ni\,S$$
$$R^2 = 0.881 \qquad (4)$$

In the cluster 4 the first subgroup is unspecific, while the second one has similar zinc and copper content in needles and cones in the range of 12.0-19.5 µg/g and a linear correlation between Cu (SZ) and Cu (I) of 87%. Obtained in this way relationships were used to correlate the dependence of metal content with the sampling sites.

Relationship between metal content - sample collection site

In the first cluster (Fig. 2, blue) the best correlated contents of metals were observed for Zarnowiec and Podlesice, where the correlation coefficient reached 94%. This high value is due to a convergent mean contents of metals studied in three groups designated by k-means analysis as described below and shown in Figure 3. The first group is characterized by the average content of zinc and copper in samples collected in one village but from different sampling sites (Fig. 3, blue). For Zarnowiec and Podlesice they are 23.5 µg/g, 26.9 µg/g (S 1 = 91%) for zinc and 16.4 µg/g, 6.5 µg/g (S = 39%) for copper.

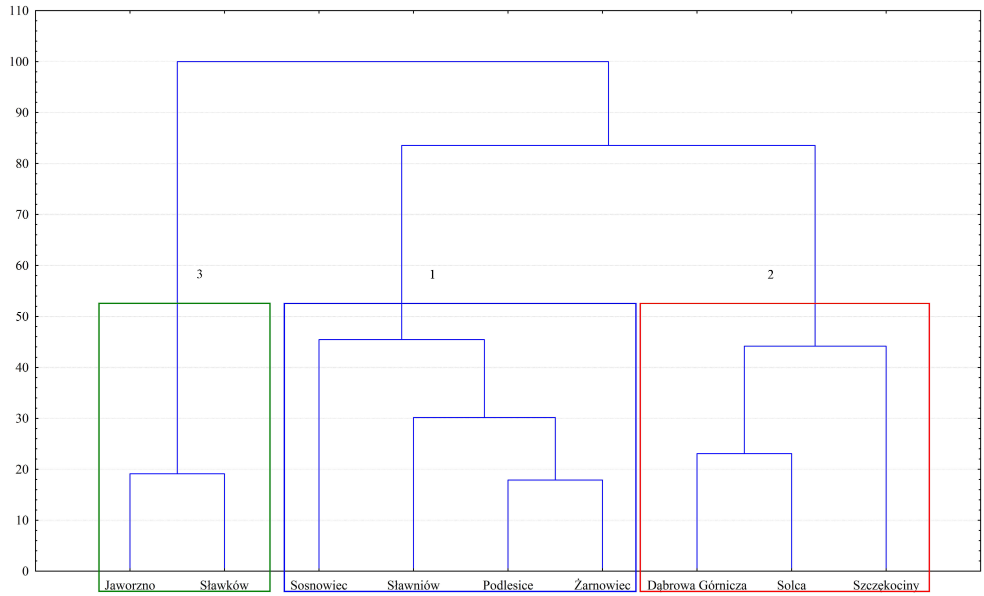

Fig. 2. Dendrogram of cluster analysis for sampling sites and analysed metals

The second cluster (Fig. 2, red) embraces Szczekociny, Dabrowa Gornicza and Solca. The best correlated are the metal contents in Solca and Dabrowa Gornicza, where the correlation coefficient reached 96%. This is due to convergent mean contents of metals studied in three groups found by the k - means analysis as shown in Figure 3. Zinc and copper contents (Fig. 3, blue) for Dabrowa Gornicza and Solca are respectively 27.2 µg/g, 25.1 µg/g (S = 94%) for zinc and 21.6 µg/g, 30.9 µgg (S = 75%) for copper. The second group describes the average content of manganese and iron (Fig. 3, red) in samples collected in one the village but from different sampling points. In the case of Dabrowa Gornicza and Solca these values are also close to each other and are 27.0 µg/g, 34.9 µg/g (S = 82%) for manganese, and 18.70 µg/g, 24.3 µg/g (S^2 = 82%) for iron, respectively. The third group was characterized by an average content of lead, nickel and cadmium (Fig. 3, green).

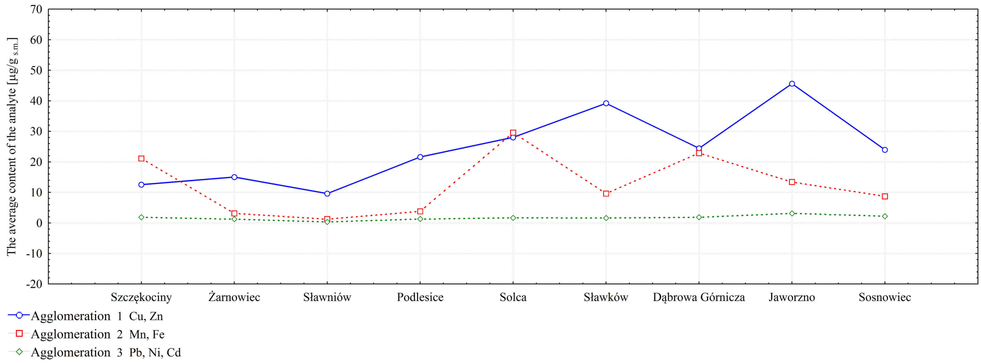

Fig. 3. Diagram of the analysis of k-means for sampling points with respect to the content of analysed metals

Mean levels of these metals in Dabrowa Gornicza and Solca were largely consistent with each other, having the similarity coefficient of 76% for lead, 96% for cadmium and 77% for nickel. These factors caused that the average contents of all the metals studied in these two cities were correlated with each other in 89%. In addition, the metal contents in Szczekociny are correlated in 91% compared with Solca and 89% compared with Dabrowa Gornicza.

The third cluster (Fig. 2, green) is the best correlated (99.9%). Taking into account the k-means analysis and sampling sites, it was found that in Jaworzno and Slawkow average contents of zinc and copper are respectively: 32.5 µg/g, 26.3 µg/g (S = 85%) - zinc and 52.1 µg/g, 58.8 µg/g. (S = 91%) - copper. The second group in this cluster is characterized by the average content of manganese and iron in the various sampling sites. In the case of Slawkow and Jaworzno these values are also similar and are respectively 15.8 µg/g, 11.4 µg/g (S = 77%) for manganese and 11.0 µg/g and 7.81 µg/g (S = 76%) for iron. The third group was characterized by an average content of lead, nickel and cadmium. The average contents of lead and cadmium in Jaworzno and Slawkow were largely consistent with each other, having the similarity coefficients of 86% for lead and 84% for cadmium. Only the average nickel content in these sites was characterized by 4% rate of convergence.

The above-mentioned types of dependencies found by chemometric analysis describe only one connection between two interesting for us data. Describing them arithmetically allowed to shorten the duration of some studies of this type. These relationships have helped in the further analyses and due to application of mentioned above chemometric techniques together it was possible to describe the new broader set of dependencies concerning present set of samples tested, which is shown below.

S^2 - coefficient of similarity between the mean values

Dependence of metal content - matrix - sample collection site

Due to the geographical locations of sampling sites (Upper Silesian Industrial District) and the size of cities and their local economy, sampling sites were initially divided into two clusters. Cities in the low density urban areas such as: Szczekociny (30), Zarnowiec (31), Slawniow (32), Podlesice (36), Solca (37), and the cities lying in a highly urbanized and industrial areas, which include: Slawkow (38), Dabrowa Gornicza (40), Jaworzno (45), Sosnowiec (46)[3]. To verify this assumption the method of k-means clustering has been applied. This method allowed to create k-clusters, differing from each other to the greatest possible extend. The X-axis (Fig. 4) refers to the concentration of tested metals, and the Y-axis shows their content in different sampling locations.

In this way, urban areas were divided into three groups:

1. Sosnowiec, Zarnowiec, Slawniow, Podlesice,
2. Szczekociny, Dabrowa Gornicza and Solca,
3. Jaworzno and Slawkow.

Qualification of Sosnowiec to the first cluster, together with villages located in the poorly urbanized and not industrialized area was caused by the fact that the average content of zinc, manganese and cadmium in this group were similar to the contents of these elements in the places considered as poorly industrialized. The same phenomenon applies to the qualification of Dabrowa Gornicza for the second cluster.

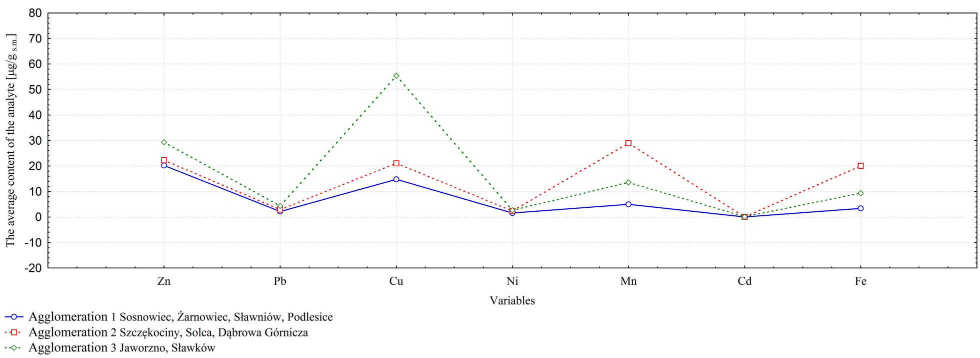

Fig. 4. Diagram of k-means analysis for metal contents with respect to sampling sites

In order to fully find out the dependencies that exist between the sampling sites, the types of pine component and the content of the elements studied, cluster analysis was executed on a set of 2268 analytical results and the dendrogram showing similarities is presented in Figure 4.

Use of cluster analysis combined with correlation analysis helped determine the relationship between the various sampling sites. The resulting dendrogram showed that the samples were clustered in seven groups, which are divided into subgroups. The first and the fifth cluster describe the presence of nickel and copper in samples of Scots pine.

The first cluster is divided into two subgroups, which in turn are divided into two more. The first subgroup describes the occurrence of nickel in urban areas (1a) and in poor urban

[3] Signs in parentheses are numbered sampling points used in the rest of this article.

areas (1b). The second subgroup describes the test material in terms of copper content, and also further divided into subgroups (1c, 1d) adequately describe the urbanized area and partially urbanized. In the first cluster, in which sampling sites are the best correlated due to the nickel and copper content, there are among others:

$$Cu\,45S = -7.9055\cdot 10^{-16} + 0.9745\cdot Cu\,36S \qquad (5)$$
$$R^2 = 0.975$$

$$Ni\,36S = -4.8816\cdot 10^{-16} + 0.9311\cdot Cu\,30SZ \qquad (6)$$
$$R^2 = 0.931$$

$$Ni\,40I = -5.5316\cdot 10^{-17} + 0.993\cdot Ni\,46SZ \qquad (7)$$
$$R^2 = 0.993$$

Cluster 5, which also describes the content of nickel and copper, is a specific grouping because it describes the presence of Ni and Cu only at two sampling points, *ie* Zarnowiec (31) and Slawniow (32) regardless of the type of samples. Correlation coefficients describing the linear dependence between the presences of copper in the area were found in the range of 0.980-0.999 and for nickel in the range of 0.966-0.999 for example:

$$Cu\,31SZ = 1.2566\cdot 10^{-16} + 0.9809\cdot Cu\,32SZ \qquad (8)$$
$$R^2 = 0.981$$

$$Ni\,32SZ = 6.7503\cdot 10^{-16} + 0.967\cdot Ni\,31G \qquad (9)$$
$$R^2 = 0.967$$

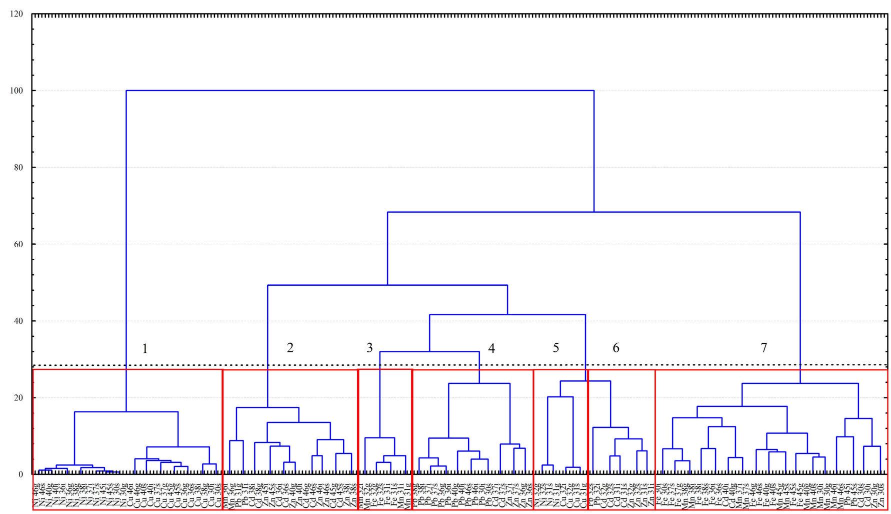

Fig. 5. Dendrogram of cluster analysis for all analysed solutions after digestion

Clusters 2 and 6 determine the presence of zinc and cadmium in the area under investigation. Cluster 2 is divided into subgroups mainly describing the presence of Zn and

Cd in the industrialized area. The population of points that are the best correlated due to the presence of these metals are such that:

$$Zn40G = -0.2051 + 1.0008 \cdot Cd36S \qquad R^2 = 0.992 \tag{10}$$

$$Zn40SZ = -0.2051 + 1.0008 \cdot Cd36SZ \qquad R^2 = 0.988 \tag{11}$$

The cluster 6 is a specific one because it identifies the presence of cadmium and zinc in different sampling points (6a $_{31.32}$, 6b $_{37}$ 6d $_{30}$), and also, like in the case of the (6c) subgroup, the presence of zinc only in the cones. Application of correlation coefficients together with the corresponding functions allowed estimation of the content of the metals studied with satisfactory accuracy of up to 85-90% for other sampling sites.

Cluster 4 describes the content of zinc and lead in the area under investigation. Unlike the previously described clusters here it is difficult to distinguish groupings describing sampling sites, the type of terrain or type of material analysed.

Clusters 3 and 7 describe the presence of iron and manganese in the analysed material. The cluster 3 is a specific one that describes the content of iron and manganese in the two measuring points: at Zarnowiec (31) and Szczekociny (32) regardless of the type of samples, which is perfectly adequate to the cluster 5. The examples of correlation coefficients describing the linear dependencies are shown below:

$$Fe32S = 0.0365 + 0.9832 \cdot Mn31SZ \qquad R^2 = 0.986 \tag{12}$$

$$Fe31SZ = 0.036 + 0.9348 \cdot Fe32G \qquad R^2 = 0.931 \tag{13}$$

Cluster 7 is divided into two subgroups, which in turn are divided into two more. The first and fourth subgroup describe the presence of iron and manganese in urban areas (7a, 7d) and the second and third subgroup describe the test material in terms of iron and manganese content in the urban poor area. Linear correlation coefficients and hence the accuracy of this mapping in this area were found in the range of 84-97%.

It was found that the content of metals in 71.5% of cones was the lowest (regardless of the sampling location). In the case of lead and nickel however the content of these metals in the cones was the highest among all components of pine tree. Lead content in cones was in accordance with the contents of this element in young branches (up to three years old) in 98%. As predicted, the highest content of the analysed elements were found in samples of litter in 86% of all cases.

The content of toxic metals - cadmium and lead in the investigated area do not exceed acceptable levels. The highest lead content (6.34 µg/g) was determined in samples from Jaworzno, while the lowest (0.19 µg/g) in samples from Slawniow (38). Cadmium contents in all tested areas were very comparable and were found in the range of 0.13-0.30 µg/g for the region from Szczekociny (30) to Jaworzno (45).

The content of heavy metals in the urban area was often up to 90% higher than in the not industrialized region. It can be seen on the example of the average content of metals such as lead, copper and manganese in the areas grouped accordingly. The average lead content in poorly industrialized area was 1.71 μg/g while in urban areas 3.89 μg/g; for copper it was 9.97 μg/g and 38.69 μg/g and for manganese it was 8.71 μg/g and 19.9 μg/g, respectively. The average zinc and nickel content in both areas did not differ significantly among themselves. They occur in a similar level in all places. Thus they cannot be used as an indicator that allows the classification of the studied area to the appropriate subgroup (the situation is distinct in the case of lead).

Analysis of principal components analysis in the description of the characteristic of industrial and non industrial areas

The PCA has been applied to explore the structure of collected measurement data. It has helped in obtaining information about the homogeneity of the data, the existence of groups and relationships between sampling sites and the type of matrix and the contents of the analytes. Visualization of the analysis carried out is shown in Figures below. In Figure 6a the cumulative percentage of variance of the data described by the main factors has been shown. Compressing the data is quite satisfactory because the first two main factors describe about 83.1% of the total variance of the data. Thanks to the PCA analysis it was possible to trace the similarities between samples collected in the province of Silesia. In Figure 6c the projection of samples on the space defined by the first two main factors has been shown. They indicate no homogenous data structure, and with affixed tags, the similarity between samples can be easily noticed.

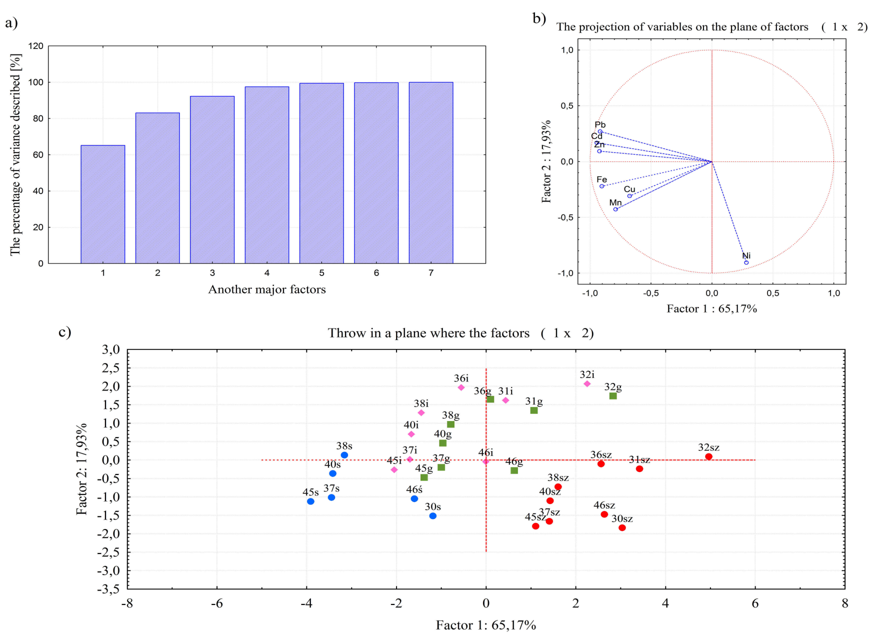

Fig. 6. The cumulative percentage of variance of the data reported in the first seven principal factors (a), projection of factorial coordinates of variables, the base currency correlation (b), projection of samples on the space defined by the first two main factors (c)

Analysing together the data from Figures 6b and 6c it can be concluded that samples of cones are described by positive values of the results along the axis of a factor 1, and the projection weights (Fig. 6b) shows a large contribution to the creation of the first principal factor by the content of zinc, lead and cadmium. Thus, for example, one could easily say that in the cone samples the level of these metals is relatively low in comparison with other samples. Between the content of lead, zinc and cadmium there are the greatest correlation which shows the projection of samples on the space defined by these three factors (Fig. 7a, b).

The content of zinc, lead and cadmium has no significant contribution to the construction of the second factor. This factor is built to the greatest extent by the nickel content in the samples. Analysing in a similar manner as described above, it can be said that the content of this element in the cones was the highest compared with other samples examined.

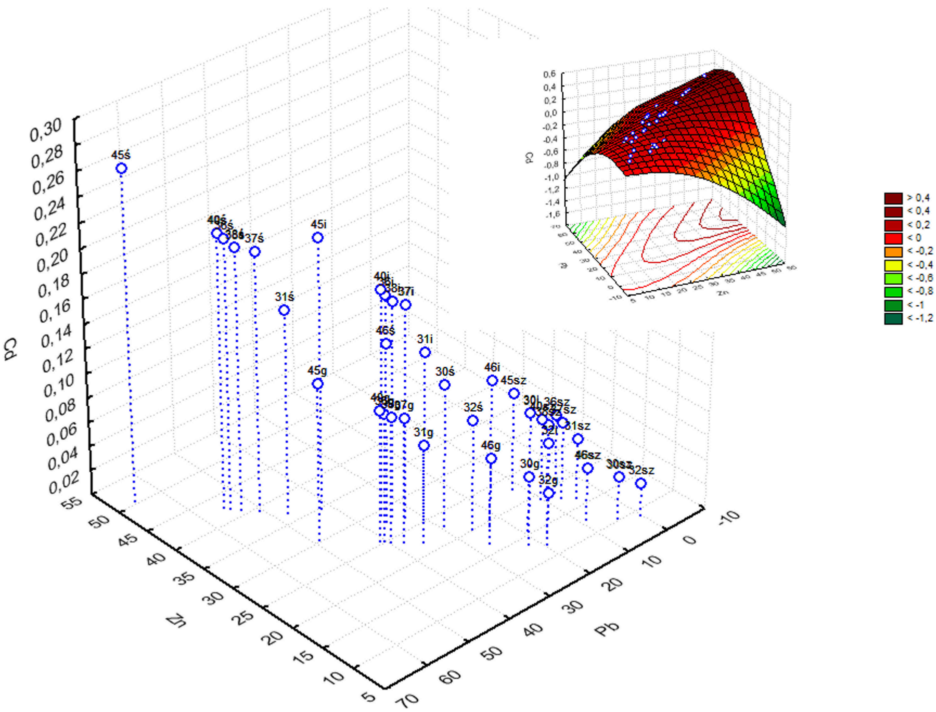

Fig. 7. Projection of samples on the space defined by the concentration of Zn, Pb and Cd: a) scatter plot, b) surface plot

By introducing to Figure 6c points corresponding to the numbering of sampling sites (Fig. 8) according to the previously described k-means method, the division of sampling sites due to the amount of the analysed metals in urban areas (red (2.3)) and poorly urbanised (blue (1)) have been illustrated.

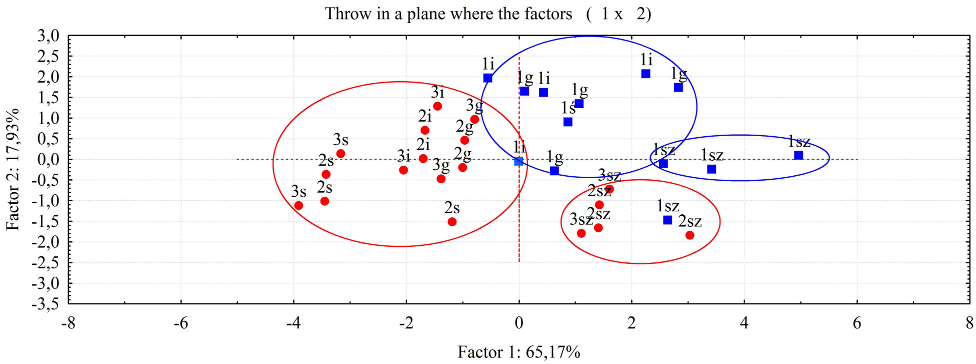

Fig. 8. The projection of samples on the space defined by the first two principal factors

Conclusions

Cluster analysis indicated the existence of significant relationships between sampling sites (site industrialized and not industrialized). It refers to a group of elements which contents are in the dependence on sampling sites ({Ni, Cu}, {Zn, Cd, Pb}, {Mn, Fe}). Principal component analysis confirms the conclusions drawn from cluster analysis and in addition allows the display of key results for easy interpretations. Application of correlation analysis helped to describe the linear dependence existing between the content of analytes in various matrices and the sampling points. With chemometric analysis used in the presented research it is possible to reduce the cost of future studies of this type conducted in this area by 30%.

References

[1] Stilwell DE, Graetz TJ. Bull. of Environ Contamin and Toxicol. 2001;67:303-308. DOI: 10.1007/s001280125.

[2] Zago C, Giblin AE, Bergamasco A. Marine Environ Res. 2001;51:389-415. DOI: 10.1016/S0141-1136(00)00135-5.

[3] Fritioff A. Greger M. Chemosphere. 2006;63:220-227. DOI: 10.1016/j.chemosphere.2005.08.018.

[4] Hida M, Sato H, Sugawara H, Mitsui T. Forensic Sci Inter. 2001;115:129-134. DOI: 10.1016/S0379-0738(00)00309-1.

[5] Szefer P, Frelek K, Szefer K, Lee ChB, Kim BS, Warzocha J, Zdrojewska I, Ciesielski T. Environ Pollut. 2002;120:423-444. http://dx.doi.org/10.1016/S0269-7491(02)00111-2.

[6] Astel A, Małek S, Makowska S. Water Air and Soil Pollut. 2008;195:137-149. DOI: 10.1007/s11270-008-9733-9.

[7] Sajdak M, Pieszko C. Cent Eur J Chem. 2012;10(5):1696-1706. DOI: 10.2478/s11532-012-0096-0.

[8] Sajdak M. Cent Eur J Chem. 2013;11(2):151-159. DOI: 10.2478/s11532-012-0145-8.

[9] Sajdak M, Pieszko C. Cent Eur J Chem. 2013;11(4):610-618. DOI: 10.2478/s11532-012-0196-x.

ANALIZA CHEMOMETRYCZNA METALI CIĘŻKICH POCHODZĄCYCH Z RÓŻNYCH MATRYC ROŚLINNYCH Z TERENÓW GÓRNOŚLĄSKIEGO OKRĘGU PRZEMYSŁOWEGO

Katedra Chemii Analitycznej, Wydział Chemiczny, Politechnika Śląska

Abstrakt: Oszacowano stężenia Zn, Cd, Ni, Mn, Cu, Pb, Fe, w 36 próbkach sosny zwyczajnej pobranych w 9 punktach umieszczonych na terenie województwa śląskiego. Mineralizaty były badane na zawartość wyżej wymienionych pierwiastków przy zastosowaniu AAS. Największe stężenie oznaczanych analitów stwierdzono w Jaworznie. Na podstawie analizy skupień stwierdzono istnienie powiązań pomiędzy składem ilościowym a miejscami pobrania próbek oraz typem matrycy, a także stwierdzono i opisano współwystępowanie obok siebie określonych grup pierwiastków

Słowa kluczowe: sosna zwyczajna, igliwie, gałęzie, metale ciężkie, analiza skupień, PCA, chemometria