Chemical speciation of metals in unpolluted soils of different types: Correlation with soil characteristics and an ANN modelling approach

J. Marković, M. Jović, I. Smičiklas, L. Pezo, M. Šljivić-Ivanović, A. Onjia, A. Popović

PII: S0375-6742(16)30063-2
DOI: doi: 10.1016/j.gexplo.2016.03.004
Reference: GEXPLO 5716

To appear in: *Journal of Geochemical Exploration*

Received date: 30 November 2015
Revised date: 23 February 2016
Accepted date: 2 March 2016

Please cite this article as: Marković, J., Jović, M., Smičiklas, I., Pezo, L., Šljivić-Ivanović, M., Onjia, A., Popović, A., Chemical speciation of metals in unpolluted soils of different types: Correlation with soil characteristics and an ANN modelling approach, *Journal of Geochemical Exploration* (2016), doi: 10.1016/j.gexplo.2016.03.004

Chemical speciation of metals in unpolluted soils of different types: correlation with soil characteristics and an ANN modelling approach

J. Marković[a], M. Jović[a]*, I. Smičiklas[a], L. Pezo[b], M. Šljivić-Ivanović[a], A. Onjia[a] and A. Popović[c]

[a] *Vinča Institute of Nuclear Sciences, University of Belgrade, P. O. Box 522, 11000 Belgrade, Serbia*
[b] *Institute of General and Physical Chemistry, University of Belgrade, 11000 Belgrade, Serbia*
[c] *The Faculty of Chemistry, University of Belgrade, Studentski trg 12–16, 11000 Belgrade, Serbia*

ABSTRACT

The distribution of elements in soil fractions affects their mobility and availability and thus their potential beneficial or harmful impact on ecosystems, biota and humans. Different mineralogical and chemical characteristics of soil influence elemental distribution. In the present study, chemical speciation of macro and micro elements (Al, Fe, Mn, K, Cd, Cr, Cu, Li, Ba, Ni, Pb and Zn) in unpolluted soils of different types, collected from the territory of the Republic of Serbia, were analysed by sequential extraction procedure. The impact of the physicochemical soil properties on the content, distribution, mobility and availability of elements was investigated. Principal component analysis was employed for the evaluation and characterization of the experimental data, understand of the relationships between soil properties and the distribution, affiliation and connection of the elements. Finally, an artificial neural network (ANN) model was developed to explore the applicability of this approach for the prediction of the elemental distribution based on soil properties. Good agreement between the model and the experimental results implied that the ANN could be considered as a useful tool for control and prediction purposes.

Keywords: Uncontaminated soil; Soil properties; Sequential extraction; Element distribution; Pattern recognition techniques; Artificial neural network.

---

*Corresponding author:

Mihajlo Jović, PhD.

University of Belgrade, Vinča Institute of Nuclear Sciences, P. O. Box 522, Belgrade, Serbia

E-mail: mjovic@vinca.rs

1. Introduction

Elements in soil are present in different levels, which reflect natural differences between soil types and consequences of soil use and pollution. The effects and functions that certain soil elements exert on living organisms range from essential to toxic. An essential element may also be a risk factor if present in inadequate amounts, i.e., concentrations higher than optimal are associated with toxicity while lower ones with nutritional deficiencies. It is generally accepted that the total content of an element in soil is not necessarily related to potential risks (Kelepertzis et al., 2015) and that the total element concentrations cannot provide necessary information on an bioavailability of element (Abollino et al., 2011; Sungur et al., 2014). Mobility and availability depend on the reactivity and binding behaviour of an element with the components of the soil matrix (Abollino et al., 2011). Several studies have investigated the relationship between soil properties and metals distribution (e.g., de Matos et al., 2001; Luz et al., 2014; Sungur et al., 2014), but mixed results were found mainly due to the different mineralogical and chemical characteristics of soils.

The sequential extraction method is a useful instrument to gain information on the bioavailability of elements. One of the most accepted and commonly used is the Tessier extraction scheme (Tessier et al., 1979). To date, numerous modifications of the basic procedure have been developed (Alvarez et al., 2006; Arcega-Cabrera et al., 2009; Lucho-Constantino et al., 2005; Riba et al., 2002; Torres and Auleda, 2013; Yu et al., 2000), whereby all of them presume a decrease in element mobility and availability along the extraction sequence. The modified Tessier methods, including the one adopted in this study, usually include alteration of the experimental conditions used for the fifth extraction phase. Namely, instead of total mineralization, which involves the use of $HClO_4$ and HF, extraction by other strong acids and their mixtures are applied. Consequently, the residual fraction becomes “pseudo-residual”. Furthermore, one additional stage has frequently been added to the Tessier scheme, to enable the determination of a water-soluble fraction of elements (Adhikari et al., 2005; Kabata-Pendias, 2001; Lucho-Constantino et al., 2005; Smičiklas et al., 2015). Using a sequential extraction method, comprehensive information can be obtained on the origin, mode of occurrence, biological and physicochemical availability, mobilization and transportation of elements (Sungur et al., 2014). In environmental impact studies, it is more important to determine element speciation patterns in soil than to assess their total pedogenic concentrations (Sungur et al., 2014).

The sequential extraction method results in a large amount of experimental data, especially when coupled to multi-element analytical techniques, such as ICP OES, which enable the simultaneous elemental analysis of a large number of samples. Moreover, the physicochemical properties of the considered soil matrix (such as particle size distribution, pH value, total carbon content, inorganic carbon content, total organic carbon, cation exchange capacity, etc.) significantly contribute to increasing the data set. Accordingly, due to simultaneous consideration of many parameters, multivariate chemometric techniques are very helpful in the visualization and interpretation of sequential extraction results (Giacomino et al., 2011).

Various multivariate statistical techniques have been employed for the evaluation and characterization of environmental data (e.g., Amiri, 2014; Giacomino et al., 2011). Principal component analysis (PCA) is a commonly used multivariate method for data reduction (the number of explanatory variables is lowered by using specific factors), but it is also used for classification and discrimination of the samples. These factors explain the major variation within the data in order to make the components more interpretable. In the last decade, PCA became accepted and used by a large number of research groups engaged in the analyses of uncontaminated, contaminated and agricultural soils (Abollino et al., 2006; Abollino et al., 2002a; Abollino et al., 2002b; Pérez and Valiente, 2005; Tokalioğlu and Kartal, 2003; Tokalioğlu et al., 2004).

The basic idea of the present research was to establish the correlations between soil physicochemical properties of soils and the distribution of a large set of elements in order to explore the predictability of their distribution based on available soil properties. In such cases, nonlinear models are more suitable due to complexity of the matrix. Artificial neural network (ANN) models are recognized as good modelling tools since they provide the solution to the problems from a set of experimental data, and are capable of handling complex systems with nonlinearities and interactions between decision variables (e.g., Almeida, 2002; Kashani et al., 2014). The developed empirical models should provide a reasonable fit to experimental data and successfully predict element mobility in different soil types.

The specific objectives of the study were: (*i*) to determine the distribution of twelve elements (minor Cd, Cr, Cu, Li, Ba, Ni, Pb and Zn, and major Al, Fe, Mn and K) in samples of different soil types having considerably different physicochemical properties (i.e., the content of sand, silt and clay fraction, pH, cation exchange capacity (CEC), total organic carbon (TOC), content of $CaCO_3$, P_2O_5 and K_2O); (*ii*) to estimate the influence of soil properties on the content, distribution, mobility and availability of the elements; (*iii*) to apply a pattern recognition technique (PCA) on the data (used as descriptors) to characterize and differentiate among the observed samples; (*iv*) to test the applicability of an ANN for the development of a mathematical model which would provide a reasonable fit of experimental data on the distribution of the elements based on the available soil properties, for control and prediction purposes.

2. Materials and methods

2.1. Soil Sampling and characterization

As the study aimed to establish correlations between soil physicochemical properties and distribution of elements, samples were selected based on their diversity in terms of their essential characteristics. The investigated samples (S_1–S_8), representing eight different soil types, were collected from various localities in the Republic of Serbia, from a depth 0–25 cm. A composite sample of each soil was prepared from five subsamples. The soils used in this study were characterized as described in detail in a previously published article (Smičiklas et al., 2015). The soil types were determined and denoted according to the World reference base for soil resources (FAO, 2006): S_1 - Humic Fluvisol, S_2 - Fluvisol, S_3 - Eutric Cambisol, S_4 - Mollic Leptosol, S_5 - Stagnosol, S_6 - Leptosol, S_7 - Dystric Cambisol and S_8 - Rendzic Leptosols. The descriptive statistics of physicochemical properties of examined soil types (Table 1) show wide ranges of soil pH, CEC, TOC, texture and nutrient content.

2.2. Sequential extraction

Soil samples were subjected to sequential extraction following a modified Tessier procedure (Tessier et al., 1979). Whereby, the examined elements were partitioned into six operationally defined fractions: water soluble (F_0), exchangeable (F_1), bound to carbonates (F_2), bound to Fe-, Mn-oxides (F_3), bound to organic matter (F_4) and residual (F_5). The sequential extraction procedure applied in this study, which was previously described in detail by Smičiklas et al. (2015), was conducted in triplicate for each sample, and the mean of the three concentration values determined for the investigated elements in each fraction were reported.

2.3. Analytical technique for the determination of the elements

The contents of the elements were determined using an inductively coupled plasma optical emission spectrometer (ICP-OES) model iCAP 6500 Duo (Thermo Scientific, United Kingdom). Radial view measurements were applied for Al, Fe and K, and axial for Zn, Ba, Cd, Cr, Cu, Li, Mn, Ni and Pb. The axial view provides better LODs (limit of detections) while the radial view is preferred for higher element concentrations. This instrument is an Echelle-type spectrometer covering the 166–847 nm range equipped with an RACID86 charge injection detector.

Determination was realized using external calibrations with matrix matched standards prepared from single stock solutions of 1000 ppm (Merck, Germany). Stock solutions were mixed into multi-standard working solutions according to their concentrations in the soil samples. Working standard solutions were matrix matched according to each extraction solution that was applied. Correlation coefficients for calibration curves were greater than 0.9999. For quality assurance, ICP multi-element standard solution VI (Merck, Germany) was used. Quality control (QC) was performed using the ACCU standard MES 21-1 as a QC standard, blank samples, standard reference material (SRM 2711) and triplicate analyses of each sample. The QC standard was measured at a frequency of every 10 analytical samples with recovery limits ±10 %. Three replicates of the SRM 2711 – Montana Soil (National Institute of Standards & Technology) were digested in a microwave accelerated reaction system model MARS 5 (CEM Corporation, USA) according to the US EPA method 3051A (US EPA, 2007) and the digestate was diluted with Milli-Q deionized water (18 MΩ). Recovery ratios between the certified and the analytical values for SRM 2711 ranged from 87.2 to 112.8 % for the investigated elements.

2.4. Statistical methods

To classify the different soil samples, PCA was applied by Eigenvalue decomposition of the correlation matrix of the obtained experimental data set in such a way that the first component contained the largest possible variance. In this manner, maximum separation between the clusters of parameters was achieved. PCA was applied in order to comprehend the relations of the obtained results, specifically the correlation between the content of the elements and the soil fractions, as well as between the soil properties and the total content of the investigated elements. In addition, linear correlation coefficients were calculated to understand the relationships between the inter-element fractions and the element–soil properties.

Data were analysed by Statistica software (Data Analysis Software System, v.10.0, StatSoft, Inc, Tulsa, OK, USA). According to the recommendation of StatSoft Statistica, the experimental database was randomly divided into three groups for the development of the ANN model: training data – 60 %, cross-validation – 20 % (used to test the performance of the network while training) and testing data – 20 % (used to examine the network generalization capability). A multi-layer perceptron model (MLP) consisting of three layers (input, hidden and output) was used in this study because it is the most common, flexible and general-purpose kind of ANN (Arsenović et al., 2013). The MLP neural network learns using an algorithm called "backpropagation". The Levenberg–Marquardt algorithm has been proved to be the fastest and particularly adapted for networks of moderate size. During this iterative process, the input data are repeatedly presented to the network (Grieu et al., 2011).

3. Results and discussion

3.1. Relations between pseudo-total metal concentrations and soil physicochemical properties

The pseudo-total amounts, the sum of the individual fractions, $\sum = F_0+F_1+F_2+F_3+F_4+F_5$, of each investigated element in the eight soil samples are given in Table 2. Of the investigated metals, concentrations of Fe and Al were commonly the highest, followed by K and Mn. Heavy metals (Cd, Cr, Cu, Ni, Pb, Zn, Ba), as well as Li were detected in a wide range of concentrations (0.0750–529 µg/g).

The PCA of pseudo-total element contents and physicochemical properties of the soil samples showed that the first two principal components (PCs) explained 57.00 % of the total variance in the original data (Fig. 1). The PC_1 contributed 29.17 %, and the PC_2 27.83 % to the total variance. The contents of Cr (which contributed 7.9 % of total variance), Cu (7.9 %), Fe (16.9 %) and Ni (11.7 %) exhibited positive scores according to PC_1, while the clay content (11.5 %) and K_2O content (9.7 %) exhibited negative influences on PC_1. The contents of Cd (8.3 %), Li (7.4 %), Zn (9.4 %) and silt fraction (9.0 %), together with pH_{KCl} (10.0 %) and pH_{H_2O} (12.8 %), exhibited positive influences on PC_2, whereas Al (7.0 %) and the sand fraction content (10.1 %) exhibited negative influences on PC_2. The Pearson's coefficients (*r*) were calculated and the obtained correlations, statistically significant at the $p<0.01$, $p<0.05$ and $p<0.10$ levels, are given in Table 3.

Soil sample of Leptosol (S_6) stands out with the highest pseudo-total concentrations of Cr, Cu, Fe and Ni, and with the lowest for Ba and K (Table 2, Fig. 1). This soil type is also characterized by the highest values of TOC (4.75 %) and the lowest values of P_2O_5 (0.01 mg $100g^{-1}$), K_2O (8.8 mg $100g^{-1}$) and clay content (22.7 %) in relation to the other observed soil types (Smičiklas et al., 2015). S_6 distinguished itself by having both the highest and the lowest deviations in terms of the obtained concentration variations (Cr, Cu, Fe, K and Ni) in the examined soil samples (Table 2). The sample of Eutric Cambisol (S_3) is characterized by a low content of $CaCO_3$ and the highest pseudo-total concentrations of Ba, K, Li and Pb. The highest pseudo-total concentrations of Al and Mn were found in soil S_5 (Stagnosol) which exhibited the smallest content of Fe in relation to the other investigated soils, and also a low content of $CaCO_3$. Dystric Cambisol (S_7) was the soil with the highest content of sand (48.8 %) and the lowest content of Cd, Cr, Li, Ni, Pb, Zn, silt, pH, TOC and $CaCO_3$. Soil sample S_1 (Humic Fluvisol) is characterized by the highest pseudo-total concentrations of Zn, pH value, P_2O_5 and K_2O contents, and the lowest content of Al and Mn; S_2 (Fluvisol) by the highest silt content and pH value, and the lowest CEC; and S_4 (Mollic Leptosol) by the highest content of Cd and the lowest values for Cu and $CaCO_3$.

The significant correlations between Cu–Cr, Ni–Cr and Ni–Cu (0.98, 1.00, 0.98, at $p<0.01$, respectively; Table 3) indicate that these metals have a common origin in the natural soils and appear together independently of soil type. The high correlation, but at a lower level of significance ($p<0.05$), was also obtained for the Fe–Cr pair (Table 3). The positive correlation between Li and Ba (0.81, $p<0.05$) suggests their natural origin and interrelationship in different types of soil. A similar conclusion could be made for the pair Al–Mn, based on the obtained significant correlation at the $p<0.05$ level (Table 3).

Considering the textural properties of soil, a negative correlation (–0.72 at $p<0.05$) was found between the sand content and the pseudo-total concentration of Cd in all soil samples. This indicates that an increased sand content is accompanied by reduced Cd concentrations, regardless of the other soil characteristics. A negative relation between total Cd and sand content was also determined for 69 sites in New Zealand (Reiser et al., 2014). The statistically significant correlations obtained in this study also indicated that an increased content of the clay fraction was associated with decreased contents of Cr and Ni (–0.72 and –0.73 at $p<0.05$, respectively) and an increased K concentration (0.78 at $p<0.05$).

Among the other relations, it was found that the soil pH decreased with increasing concentrations Al and sand contents, while the content of K_2O was negatively correlated with the concentrations of Cr and Fe (Table 3). Pearson's correlation analysis also indicated that the soil TOC was positively correlated with the amounts of Cr, Cu and Ni (0.76, 0.76 and 0.77, respectively, at $p<0.05$).

3.2. Relations between metal distribution and physicochemical properties of the soil

Mobilities of elements largely depend on the types of bonds by which they are associated with the soil components. The first two fractions of the applied sequential extraction F_0 (water soluble) and F_1 (exchangeable) constitute the available forms of elements. F_2 (bound to carbonates) and F_3 (bound to iron and manganese oxides) phases are potentially mobile, but not readily available under natural conditions. Finally, the last two fractions F_4 (bound to organic matter) and F_5 (residual) constitute unavailable pools of elements (Abollino et al., 2011). The hitherto correlations between soil characteristics and element distribution, referred to several elements and/or soil samples (Ashraf et al., 2012; Guo et al., 2005; Navas and Lindhorfer, 2005; Sipos et al., 2014) and, therefore, this study aimed at extending the database.

The different patterns of metal partitioning in soil, being the function of metal and soil properties, are demonstrated in Fig. 2. PCA was performed to analyse possible relationships between soil types and the distribution of elements in different phases. The rotation of PCs was executed by the Varimax method with Kaiser normalization. The PCA showed that the first PCs explained 64.05 % of the total variance in the original data, Fig. 3. The PC_1

contributed 44.33 %, and the PC_2 19.72 % to the total variance. The contents of Al (which contributed 14.1 % of total variance), Cr (7.9 %), Cu (7.9 %), Fe (16.9 %), K (8.2 %), Li (13.3 %), Ni (8.3 %), Zn (13.0 %), exhibited negative scores according to the PC_1. The contents of K (14.5 %), Li (8.8 %) and Zn (11.1 %) showed a positive influence on the evaluation of the PC_2, while the content of Mn (26.5 %) and Pb (19.3 %) exhibited negative influences on the PC_2 (Fig. 3). The Pearson's correlation coefficients between the content of the investigated elements in different soil fractions are presented in Table 4.

The grouping of Cd with the samples related to Phase 1 (samples No. 9–16), Fig. 3, indicated that this element was mainly bound to the exchangeable phase (F_1) in most of the examined soil samples. The obtained pseudo-total Cd concentrations in the present study were in agreement with the average concentration of Cd in the Earth’s crust (~0.2 μg/g) and with the background Cd level in surface soils (<1 mg kg^{-1}) (Gleyzes et al., 2002; Petrovic et al., 2009). Cadmium was identified as the most mobile of the investigated metals, with up to 60.5 % in the exchangeable fraction. A high mobility of Cd was detected in other uncontaminated soil samples (Dimović et al., 2013), as well as in contaminated soil around a mining area (He et al. 2013). Pueyo et al. (2003) investigated the distribution of a large number of elements (Al, Ca, Fe, Mg, Mn, As, Bi, Cd, Cu, Pb, Tl and Zn) in contaminated soil and found that Cd was the most mobile. Due to the high Cd mobility and availability, its increased concentration in soil may cause long-term risks to the ecosystem, biota and humans, and hence, Cd belongs to the group pollutants that are of major interest in bioavailability studies listed by the US Environmental Protection Agency.

The interconnection of Phase 3 (samples No. 25–32) with Mn and Pb indicate that these elements in soils are primarily linked to iron and manganese oxides. Logically, Mn as one of the main contributors in soil occurs mainly in the oxide fraction (44 to 81 %). The remaining Mn content is distributed between the other phases, in amounts which obviously depend on soil type. The mobility of Mn in soils is extremely sensitive to soil conditions, such as acidity, wetness, organic matter content, biological activity (Nadaska et al., 2012). Generally, low pH values favour the reduction of insoluble manganese oxides, resulting in increased manganese mobility, while at soil pH values above 6, Mn is efficiently bound to organic matter, oxides and silicates, resulting in decreased Mn solubility. In addition, Mn availability is high in soils with low content of organic matter (Nadaska et al., 2012). This was confirmed by the present results, in which the most mobile Mn was found in samples Dystric Cambisol (S_7), Stagnosol (S_5) and Eutric Cambisol (S_3), characterized by both low pH and low TOC content. Mn was found in the readily available phase for biogeochemistry cycles in the ecosystems (F_1) and potentially bioavailable phase (F_2), up to 16 % and 11 %, respectively, for certain soil types. A similar distribution of Mn in soil, with the highest content in F_3, was also obtained by Petrović et al. (2009), Walna et al. (2010) and Navas and Lindhorfer (2005). Pb was also mainly associated to Fe-, Mn-oxide fraction (32–66 %), followed by the oxidizable (F_4) phase (17–45 %) and the residual phase (13–27 %). Negligible Pb concentrations were found in fractions F_0 and F_1, except for Dystric Cambisol (S_7), in which the labile F_1 phase represents 11 % of the pseudo-total concentration. Abollino et al. (2002a) and Guo et al. (2005) reported that Pb was mostly associated with the residual phase, while Sarkar et al. (2014) found that the major geochemical phase for Pb was the Fe-, Mn-oxides phase, followed by the residual and the oxidizable phases, as was the case in the present study. Pb was found to have a high affinity for soil Fe-oxides. This was confirmed by Cornu et al. (2005), who reported that Pb was able to form stable hydroxide and carbonate complexes which are preferentially bound to the (slightly) positively charged Fe-oxide surfaces. Pb in the examined soil samples had a low potential mobility (0 % for F_0 and around 2 % for the F_1 and F_2 phases, on average) and, generally speaking, it could be considered as a relatively immobile element. The exception was sample S_7 with 13 % of potentially mobile

Pb (F_0+F_1). Given that S_7 was characterized by the lowest pseudo-total concentration of Pb compared with the other soil types, the higher mobility was more likely influenced by specific soil characteristics than some source of contamination.

The high correlations between K, Zn, Li and Fe (Zn–Li=0.94, Fe–Li=0.91, K–Zn=0.89, Fe–Zn=0.88, K–Li=0.81 and Fe–K=0.73 at $p<0.01$; Table 4) and their association with samples of Phase 5 (samples No. 41–48) on the bioplot PCA diagram, Fig. 3, indicate a similar distribution of these elements in the soil phases, with the highest amount in F_5. In all soil samples, Zn was the most abundant in the residual fraction (69–90 %). Negligible concentrations of this element were found in phase F_0, F_1 and F_2, indicating its low mobility. The abundance of Zn in F_4 phase of all samples ranged from 7 to 15 %. An inappreciable amount of Zn was found in phase F_3 except for in samples S_1, S_2 and S_6, in which the abundance was 15–17 %. A difference in the Zn distribution among the various soil types was evidenced by Sarkar et al. (2014). The association of Zn with the Fe-, Mn-oxide phase may be linked to the high stability constants of Zn oxides (Abdu, 2010). Fe-oxides adsorb considerable quantities of Zn and may also occlude Zn in the lattice structures (Sarkar et al., 2014). Li and K also existed mainly in the residual phase (52–74 % and 53–84 %, respectively). High content in F_5 could primarily be linked to their natural geochemical origin and their association with the soil matrix. Between 4–19 % and 2–14 % of the extracted Li and K, respectively, were associated with the exchangeable phase and thus were quite labile and easily available. Presence of Li and K in the remaining phases varies significantly depending on the soil type, except in F_0, in which negligible concentrations were found. In the available literature, it was difficult to find an adequate study to compare the distribution of these two elements. Finally, Fe is a soil macro element present mainly in the residual phase (62–83 %). In all soil types, the Fe distribution followed the order: $F_5 > F_3 > F_4$. The high percentages of Fe in the hardly accessible or inaccessible element pools confirmed its interconnection with the soil matrix.

A high correlation was also obtained between Fe and Al (Table 4). Al was primarily distributed between the F_3 (27–48 %) and F_5 (35–54 %) phases, and in the F4 phase (13–19 %). Insignificant concentrations of both Fe and Al were found in phases F_0–F_2. The obtained results for the Fe and Al distributions in the soil samples were in good agreement with those of several reported studies (e.g., Abollino et al., 2002a; Navas and Lindhorfer, 2005; Sarkar et al., 2014).

The high correlations between Cr, Ni and Cu at $p<0.01$ (Table 4) indicate their similar speciation in the investigated soils. The highest levels of Cr were found in the F_5 phase of all soils, and the highest levels of Ni and Cu were determined in this phase of the majority of the investigated soils. The dominant proportion of Ni and Cr found in the residual phase is in agreement with the results of other studies (Abollino et al., 2002a; Guo et al., 2005; Sarkar et al., 2014). Ni and Cr are, apart from the F_5, present in F_3 (6–42 % for Ni and 13–26 % for Cr) and F_4 (12–33 % for Ni and 12–25 % for Cr) phases. In soil samples S_3, S_5, S_6 and S_7, Ni was present in F_1 (~3 %), which represents its easily releasable content. Cu was distributed between the F_3, F_4 and F_5 phases, with no similar pattern, indicating that the type of soil plays an important role in the distribution and mobility of Cu. Copper has an ionic radius close to that of Fe^{3+} and thus, it could be incorporated into the structure of crystalline Fe-oxides. Furthermore, the high stability constants of Cu complexes with organic matter enable Cu binding to lipids, proteins, and carbohydrates, while the high affinity of clay minerals for Cu incorporation explains its presence in the residue (Martinez and McBride, 1998; Sarkar et al., 2014). No statistically significant correlations were found between Ba and the other examined elements. Ba was evenly distributed between phases F_1, F_2, F_3, F_4 and F_5, with a negligible concentration in phase F_0, in all the examined soil types.

As a final point, to clarify in what manner soil properties affect the mobility of examined elements, the Pearson's correlation coefficients were calculated (Table 5). The distribution of the majority of elements, in one or more soil fractions, was in correlation with the specific soil characteristics, whereas the patterns of Ba and Mn partitioning could not be connected with any of the investigated soil properties at the desired level of statistical significance.

The partitioning of the macro-elements Fe and K was significantly correlated to the soil texture, i.e., to the content of clay. With the increased percentage of clay fraction in the soil, the Fe found in Fe-, Mn-oxide phase decreased, while the K concentration increased in the most stable fractions F_4 and F_5. The total content of K in soil generally increased with increasing clay content (Peverill et al., 1999), which was confirmed by the significant positive correlation found in the present study (0.78, $p<0.05$, Table 3). Furthermore, the observed high stability of K in the clay-rich soils was in agreement with the previously reported significant positive relationship between non-exchangeable K and the illite clay content (Rezapour et al., 2009). The mobility of Al was not only strongly related to soil texture, but also to pH, two mutually connected soil properties (Table 3). The content of Al in the most mobile F_1 phase was positively correlated with the sand and negatively with the silt content. In addition, a negative relationship between soil pH and the Al content in F_2 was observed, which may be connected with negative correlation linking soil pH with the pseudo-total content of Al (–0.83 at $p<0.05$; Table 3).

Considering the distribution of the micro-elements, the effects of the soil properties were numerous and complex. The distribution of Li was mostly affected by soil pH and carbonate content. The correlations were positive, meaning a higher Li content in the F_2 fraction of carbonate-rich soils, which also exhibited higher pH values. The general weak influence of $CaCO_3$ on the distribution of the other elements may be partially ascribed to the lack of carbonates in samples S_3–S_6. A similar absence of the effect of carbonate content on the mobility of Pb, Cu, Cr, Zn, As and Sn was registered by Ashraf et al. (2012) in the ex-mining land of Bestari Jaya, Malaysia. The Cd distribution was significantly affected solely in F_2 fraction. The Cd content increased with decreasing sand content, which is in agreement with the negative correlation between the sand content and the pseudo-total concentration of Cd in various soil types (–0.72 at $p<0.05$; Table 3). In contrast, the content of Cd in F_2 was positively correlated with soil pH. The easily available content of Pb in the F_1 fraction largely increased with increasing sand fraction. On the other hand, increasing silt content is associated with stabilization of Pb through its redistribution from F_1 to F_3. The increased sand content had a similar effect on the mobility of Zn, given that the Zn concentration in F_1 increased, while simultaneously decreased in F_2, with increasing sand fraction. Soil pH had an opposing effect, i.e., with increasing pH, the Zn redistributed from F_1 to F_2.

The distributions of Cu, Cr and Ni were largely affected by the soil TOC and the clay content. The negative correlation between clay content and pseudo-total concentrations of Cr, Cu and Ni was previously observed (Table 3). In light of the results presented in Table 5, the observed trend may be interpreted as a decrease of the Cr content in F_3 and F_5, a Cu decrease in F_3, and a reduction in the Ni content in F_0, F_1, F_3 and F_5. In addition the TOC content also showed a significant correlation with the pseudo-total Cr, Cu and Ni concentrations, but a positive one (Table 3). Thus, the increasing total concentration of the mentioned elements with increasing TOC may be associated with increasing Cr content in F_2, F_3, F_4 and F_5, the Cu in F_3 and F_4, and Ni in F_0, F_1, F_3, F_4 and F_5. In addition, some stabilization of Cu with increasing TOC may be an explanation for the observed redistribution from F_2 to F_3 and $F_{4.}$ The Cu content in the carbonate fraction was also negatively correlated with the soil CEC.

3.3. ANN modelling of metal distribution in different soil types

The Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, implemented in the StatSoft Statistica evaluation routine, was used for the ANN modelling. The optimization procedures to minimize the error function between network and experimental outputs was used during the ANN training cycle (Pezo et al., 2013; Taylor, 2006) and the sum of squares (SOS) was evaluated according to the BFGS algorithm, to accelerate and stabilize the convergence of the results (Basheer and Hajmeer, 2000). The training process was repeated several times in order to obtain the best performance of the ANN, due to a high degree of variability of the parameters. It was accepted that successful training was achieved when the learning and cross-validation curves (SOS *vs*. training cycles) approached zero. Coefficient of determination (r^2) and SOS were used as parameters to check the performance (i.e., the accuracy) of the obtained ANN.

The optimum number of hidden neurons was chosen upon minimizing the difference between the predicted ANN values and the desired outputs, using *SOS* during testing as a performance indicator. The employed MLP is marked according to StatSoft Statistica's notation, "MLP" followed by the number of inputs, number of neurons in the hidden layer, and the number of outputs. According to the ANN performance (sum of r^2 and SOSs for all variables in one ANN), it was determined that the optimal number of neurons in the hidden layer is 7 (network MLP 2-4-12), Table 6.

The goodness-of-fit, between the experimental measurements and model calculated outputs, represented as ANN performance (r^2 between measured and calculated Al, Ba, Cd, Cr, Cu, Fe, K, Li, Mn, Ni, Pb and Zn) during the training, testing and validation steps, are given in Table 7.

The ANN models were used to predict the experimental variables (Al, Ba, Cd, Cr, Cu, Fe, K, Li, Mn, Ni, Pb and Zn) in the different soil types and phases. The networks were able to predict reasonably well all process outputs for a broad range of the process variables (as seen in Fig. 4, where the experimental measured and ANN model predicted values are presented).

The predicted values were very similar to the desired values in most cases, in terms of the r^2 value, for the ANN models. The SOS obtained with the ANN models were of the same order of magnitude as the experimental errors for Al, Ba, Cd, Cr, Cu, Fe, K, Li, Mn, Ni, Pb and Zn reported in the literature (Basheer and Hajmeer, 2000; Pezo et al., 2013). The values of r^2 between experimental and ANN model (MLP 2-4-12) outputs, for Al, Ba, Cd, Cr, Cu, Fe, K, Li, Mn, Ni, Pb and Zn were 0.978, 0.625, 0.910, 0.438, 0.572, 0.931, 0.815, 0.919, 0.901, 0.427, 0.920 and 0.962, respectively, during the training period. Moreover, the means and the standard deviations of the residuals were analysed. The mean of the residuals for the ANN model were 5.90, –2.88, 0.00, –4.64, –1.14, –422.99, –7.65, –0.12, –8.04, –3.66, –0.06 and –0.02 for Al, Ba, Cd, Cr, Cu, Fe, K, Li, Mn, Ni, Pb and Zn, respectively, while the standard deviations were 96.65, 18.17, 0.02, 41.01, 11.49, 1675.17, 553.68, 1.32, 225.93, 39.21, 3.06 and 3.51, respectively. These results showed a relatively good approximation to a normal distribution around zero with a probability of 95 % (2•SD), which indicates a good generalization ability of the ANN model for the ranges of the observed experimental values.

The fitting parameters imply that the ANN approach has a significant potential as a tool for the rapid assessment of metal mobility. The improvement in the prediction reliability should be tested in the future by taking into account additional soil parameters (such as mineralogy, specific surface area and porosity, anion exchange capacity (AEC), the type of organic matter, etc.). For instance, it could easily be observed from Fig. 4 that the predicted and measured K concentrations were largely scattered in the region of high concentrations, associated with fraction F_5. As clay mineralogy plays an important role in terms of the distribution of the K content (Rezapour et al., 2009), it must be taken into consideration as input data for ANN development. Furthermore, the high concentrations of Ni and Cr in sample S_6 have a pronounced impact on the fitting results (Fig. 4). Even though S_6 is an unpolluted soil, some locations in Serbia are recognized as Ni and Cr rich (Albanese et al., 2015), and such unusually high concentrations of naturally occurring elements were associated with ophiolite masses and related sedimentary rocks. Therefore, the goodness-of-fit could be enhanced by increasing the total sample size, which would, among others, also contain more extremes.

4. Conclusions

This study aimed to provide information on the speciation of Al, Fe, Mn, K, Cd, Cr, Cu, Li, Ba, Ni, Pb and Zn in different types of unpolluted soil, and establish correlations of metal contents and speciation with soil characteristics. For this purpose, a chemometric approach (PCA and ANN) was employed for the evaluation of the results of sequential extraction analysis. The patterns of metal partitioning were dependent on both metal and soil type. With the exception of Cd and Ba, these metals were commonly found in the less available fractions (F_3, F_4 and F_5). Generally, soil texture and TOC had a predominant impact on the metal contents and speciation. Samples with a higher percentage of sand exhibited increased contents of Pb, Zn and Al in phase F_1, which is readily available in the environment. An ANN model was developed in order to predict the complex distribution and mobility of elements in uncontaminated soils based on the known soil properties. High r^2 values and prediction accuracy of the observed outputs proved the ANN as useful in the prediction of metal (especially Al, Cd, Fe, K, Li, Mn, Pb and Zn) mobility in soils with a wide range of characteristics. The obtained relationships are important for a rapid assessment of metal speciation and represent a starting point for analysing mobility changes due to contamination.

Acknowledgement

This work was supported by the Ministry of Education Science and Technological Development of the Republic of Serbia (Project III43009).

References

Abdu, N., 2010. Availability, transfer and balances of heavy metals in urban agriculture of West Africa. Kassel University Press GmbH, Kassel, Germany.

Abollino, O., Aceto, M., Malandrino, M., Mentasti, E., Sarzanini, C., Barberis, R., 2002a. Distribution and mobility of metals in contaminated sites. Chemometric investigation of pollutant profiles. Environ. Pollut. 119, 177–193. doi:10.1016/S0269-7491(01)00333-5

Abollino, O., Aceto, M., Malandrino, M., Mentasti, E., Sarzanini, C., Petrella, F., 2002b. Heavy metals in agricultural soils from Piedmont, Italy. Distribution, speciation and chemometric data treatment. Chemosphere 49, 545–557. doi:10.1016/S0045-6535(02)00352-1

Abollino, O., Giacomino, A., Malandrino, M., Mentasti, E., Aceto, M., Barberis, R., 2006. Assessment of Metal Availability in a Contaminated Soil by Sequential Extraction. Water. Air. Soil Pollut. 173, 315–338. doi:10.1007/s11270-005-9006-9

Abollino, O., Malandrino, M., Giacomino, A., Mentasti, E., 2011. The role of chemometrics in single and sequential extraction assays: A review. Anal. Chim. Acta 688, 104–121. doi:10.1016/j.aca.2010.12.020

Adhikari, T., Biswas, A.K., Saha, J.K., 2005. Cadmium Phytotoxicity in Spinach with or without Spent Wash in a Vertisol. Commun. Soil Sci. Plant Anal. 36, 1499–1511. doi:10.1081/CSS-200058495

Albanese, S., Sadeghi, M., Lima, A., Cicchella, D., Dinelli, E., Valera, P., Falconi, M., Demetriades, A., De Vivo, B., 2015. GEMAS: Cobalt, Cr, Cu and Ni distribution in agricultural and grazing land soil of Europe. J. Geochemical Explor. 154, 81–93. doi:10.1016/j.gexplo.2015.01.004

Almeida, J.S., 2002. Predictive non-linear modeling of complex data by artificial neural networks. Curr. Opin. Biotechnol. 13, 72–76. doi:10.1016/S0958-1669(02)00288-4

Alvarez, J.M., Lopez-Valdivia, L.M., Novillo, J., Obrador, A., Rico, M.I., 2006. Comparison of EDTA and sequential extraction tests for phytoavailability prediction of manganese and zinc in agricultural alkaline soils. Geoderma 132, 450–463. doi:10.1016/j.geoderma.2005.06.009

Amiri, F., 2014. A nutritive value of Iranian mangrove ecosystems, northern part of the Persian Gulf. Natural Resources Research 23, 321-330.

Arcega-Cabrera, F., Armienta, M.A., Daesslé, L.W., Castillo-Blum, S.E., Talavera, O., Dótor, A., 2009. Variations of Pb in a mine-impacted tropical river, Taxco, Mexico: Use of geochemical, isotopic and statistical tools. Appl. Geochemistry 24, 162–171. doi:10.1016/j.apgeochem.2008.09.015

Arsenović, M., Radojević, Z., Stanković, S., Lalić, Ž., Pezo, L., 2013. What to expect from heavy clay? Ceram. Int. 39, 1667–1675. doi:10.1016/j.ceramint.2012.08.009

Ashraf, M.A., Maah, M.J., Yusoff, I., 2012. Chemical speciation and potential mobility of heavy metals in the soil of former tin mining catchment. ScientificWorldJournal. 2012, 125608. doi:10.1100/2012/125608

Basheer, I.., Hajmeer, M., 2000. Artificial neural networks: fundamentals, computing, design, and application. J. Microbiol. Methods 43, 3–31. doi:10.1016/S0167-7012(00)00201-3

Cornu, S., Deschatrettes, V., Salvador-Blanes, S., Clozel, B., Hardy, M., Branchut, S., Le Forestier, L., 2005. Trace element accumulation in Mn—Fe—oxide nodules of a planosolic horizon. Geoderma 125, 11–24. doi:10.1016/j.geoderma.2004.06.009

de Matos, A.T., Fontes, M.P.F., da Costa, L.M., Martinez, M.A., 2001. Mobility of heavy metals as related to soil chemical and mineralogical characteristics of Brazilian soils. Environ. Pollut. 111, 429–435. doi:10.1016/S0269-7491(00)00088-9

Dimović, S., Smičiklas, I., Šljivić-Ivanović, M., Dojčinović, B., 2013. Speciation of 90Sr and other metal cations in artificially contaminated soils: the influence of bone sorbent addition. J. Soils Sediments 13, 383–393. doi:10.1007/s11368-012-0633-7

FAO, 2006. World reference base for soil resources 2006, World Soil Resources Reports. doi:10.1017/S0014479706394902

Giacomino, A., Abollino, O., Malandrino, M., Mentasti, E., 2011. The role of chemometrics in single and sequential extraction assays: a review. Part II. Cluster analysis, multiple linear regression, mixture resolution, experimental design and other techniques. Anal. Chim. Acta 688, 122–39. doi:10.1016/j.aca.2010.12.028

Gleyzes, C., Tellier, S., Astruc, M., 2002. Methodologies for Soil and Sediment Fractionation Studies, in: Quevauviller, P. (Ed.), Methodologies for Soil and Sediment Fractionation Studies, Single and Sequential Extraction Procedures. Royal Society of Chemistry, Cambridge, p. 200. doi:10.1039/9781847551412

Grieu, S., Faugeroux, O., Traoré, A., Claudet, B., Bodnar, J.-L., 2011. Artificial intelligence tools and inverse methods for estimating the thermal diffusivity of building materials. Energy Build. 43, 543–554. doi:10.1016/j.enbuild.2010.10.020

Guo, P., Xie, Z.L., Li, J., Kang, C.L., Liu, J.H., 2005. Relationships between fractionations of Pb, Cd, Cu, Zn and Ni and soil properties in urban soils of Changchun, China. Chinese Geogr. Sci. 15, 179–185.

He, Q., Ren, Y., Mohamed, I., Ali, M., Hassan, W., Zeng, F., 2013. Assessment of trace and heavy metal distribution by four sequential extraction procedures in a contaminated soil. Soil Water Res. 8, 71–76.

Kabata-Pendias, A., 2001. Trace Elements in Soils and Plants, Third Edition, 3rd ed. CRC Press, Boca Raton, FL.

Kashani, M.H., Ghorbani, M.A., Dinpashoh, Y., Shahmorad, S., 2014. Comparison of Volterra model and artificial neural networks for rainfall–runoff simulation. Natural Resources Research 23, 341-354.

Kelepertzis, E., Paraskevopoulou, V., Argyraki, A., Fligos, G., Chalkiadaki, O., 2015. Evaluation of single extraction procedures for the assessment of heavy metal extractability in citrus agricultural soil of a typical Mediterranean environment (Argolida, Greece). J. Soils Sediments 15, 2265–2275. doi:10.1007/s11368-015-1163-x

Lucho-Constantino, C.A., Alvarez-Suárez, M., Beltrán-Hernández, R.I., Prieto-García, F., Poggi-Varaldo, H.M., 2005. A multivariate analysis of the accumulation and fractionation of major and trace elements in agricultural soils in Hidalgo State, Mexico irrigated with raw wastewater. Environ. Int. 31, 313–23. doi:10.1016/j.envint.2004.08.002

Luz, F., Mateus, A., Matos, J.X., Gonçalves, M.A., 2014. Cu- and Zn-Soil anomalies in the NE border of the South Portuguese Zone (Iberian Variscides, Portugal) identified by multifractal and geostatistical analyses, Natural Resources Research 23, 195-215.

Martinez, C.E., McBride, M.B., 1998. Coprecipitates of Cd, Cu, Pb and Zn in iron oxides:

Solid phase transformation and metal solubility after aging and thermal treatment. Clays Clay Miner. 46, 537–545.

Morais, S., Garcia e Costa, F., Pereira, M. de L., 2012. Heavy metals and human health, in: Oosthuizen, J. (Ed.), Environmental Health-Emerging Issues and Practice. InTech, Rijeka, Croatia, pp. 227–246. doi:10.5772/29869

Nadaska, G., Lesny, J., Michalik, I., 2012. Environmental aspect of manganese chemistry. Hungarian Electron. J. Sci. HEJ:ENV-10, 1–16.

Navas, A., Lindhorfer, H., 2005. Chemical Partitioning of Fe, Mn, Zn and Cr in Mountain Soils of the Iberian and Pyrenean Ranges (NE Spain). Soil Sediment Contam. An Int. J. 14, 249–259. doi:10.1080/15320380590928311

Pérez, G., Valiente, M., 2005. Determination of pollution trends in an abandoned mining site by application of a multivariate statistical analysis to heavy metals fractionation using SM&T-SES. J. Environ. Monit. 7, 29–36. doi:10.1039/b411316k

Petrovic, D., Todorovic, M., Manojlovic, D., Krsmanovic, V.D., 2009. Speciations of trace metals in the accumulation Bogovina on the Crni Timok river. Polish J. Environ. Stud. 18, 873–884.

Peverill, K., Sparrow, L., Reuter, D. (Eds.), 1999. Soil Analysis: An Interpretation Manual. CSIRO Publishing.

Pezo, L., Curcic, B., Filipovic, V., Nicetin, M., Koprivica, G., Misljenovic, N., Levic, L., 2013. Artificial neural network model of pork meat cubes osmotic dehydratation. Hem. Ind. 67, 465–475. doi:10.2298/HEMIND120529082P

Pueyo, M., Sastre, J., Hernández, E., Vidal, M., López-Sánchez, J.F., Rauret, G., 2003. Heavy Metals in the Environment: Prediction of Trace Element Mobility in

Contaminated Soils by Sequential Extraction. J. Environ. Qual. 32, 2054–2066.

Reiser, R., Simmler, M., Portmann, D., Clucas, L., Schulin, R., Robinson, B., 2014. Cadmium concentrations in new zealand pastures: relationships to soil and climate variables. J. Environ. Qual. 43, 917–25. doi:10.2134/jeq2013.09.0367

Rezapour, S., Jafarzadeh, A.A., Samadi, A., Oustan, S., 2009. Impacts of clay mineralogy and physiographic units on the distribution of potassium forms in calcareous soils in Iran. Clay Miner. 44, 327–337. doi:10.1180/claymin.2009.044.3.327

Riba, I., Delvalls, T. a., Forja, J.M., Gómez-Parra, A., 2002. Evaluating the heavy metal contamination in sediments from the Guadalquivir estuary after the Aznalcóllar mining spill (SW Spain): A multivariate analysis approach. Environ. Monit. Assess. 77, 191–207. doi:10.1023/A:1015828020313

Sarkar, S.K., Favas, P.J.C., Rakshit, D., Satpathy, K.K., 2014. Geochemical Speciation and Risk Assessment of Heavy Metals in Soils and Sediments, in: Hernandez-Soriano, M.C. (Ed.), Environmental Risk Assessment of Soil Contamination. InTech, p. 918. doi:10.5772/57295

Sipos, P., Németh, T., Choi, C., Szalai, Z., Balázs, R., 2014. Distribution, geochemical fractionation and sorption of Cu and Pb in soils characteristic of Hungary. Cent. Eur. Geol. 57, 265–285. doi:10.1556/CEuGeol.57.2014.3.3

Smičiklas, I., Jović, M., Šljivić-Ivanović, M., Mrvić, V., Čakmak, D., Dimović, S., 2015. Correlation of Sr2+ retention and distribution with properties of different soil types. Geoderma 253-254, 21–29. doi:10.1016/j.geoderma.2015.04.003

Sungur, A., Soylak, M., Ozcan, H., 2014. Investigation of heavy metal mobility and availability by the BCR sequential extraction procedure: relationship between soil properties and heavy metals availability. Chem. Speciat. Bioavailab. 26, 219–230.

doi:10.3184/095422914X14147781158674

Taylor, B.J. (Ed.), 2006. Methods and Procedures for the Verification and Validation of Artificial Neural Networks, 1st ed. Springer US. doi:10.1007/0-387-29485-6

Tessier, A., Campbell, P.G.C., Blsson, M., 1979. Sequential extraction procedure for the speciation of particulate trace metals. Anal. Chem. 51, 844–851.

Tokalioğlu, Ş., Kartal, Ş., 2003. Relationship between vegetable metal and soil-extractable metal contents by the BCR sequential extraction procedure: chemometrical interpretation of the data. Int. J. Environ. Anal. Chem. 83, 935–952. doi:10.1080/0306731031000l608740

Tokalioğlu, Ş., Kartal, Ş., Güneş, A.A., 2004. Statistical Evaluation of Bioavailability of Metals to Grapes Growing in Contaminated Vineyard Soils using Single Extractants. Int. J. Environ. Anal. Chem. 84, 691–705. doi:10.1080/03067310410001688444

Torres, E., Auleda, M., 2013. A sequential extraction procedure for sediments affected by acid mine drainage. J. Geochemical Explor. 128, 35–41. doi:10.1016/j.gexplo.2013.01.012

US EPA, 2007. Method 3051: Microwave assisted acid digestion of sediments, sludges, soils, and oils. Test Methods Eval. Solid Waste, 1–30.

Walna, B., Spychalski, W., Ibragimow, A., 2010. Fractionation of iron and manganese in the horizons of a nutrient-poor forest Soil profile using the sequential extraction method. Polish J. Environ. Stud. 19, 1029–1037.

Yu, K.C., Chang, C.Y., Tsai, L.J., Ho, S.T., 2000. Multivariate analyses on heavy metal binding fractions of river sediments in Southern Taiwan. Water Sci. Technol. 42, 193–199.

TABLES

Table 1

Descriptive statistics of physicochemical properties of the investigated soil types.

	Sand[a]	Silt[b]	Clay[c]	$CaCO_3$[d]	pH_{KCl}	pH_{H_2O}	CEC[e]	TOC[d]	P_2O_5[f]	K_2O[f]
Mean	32.2	30.8	37.1	2.90	5.46	7.05	32.03	2.28	9.69	28.19
SD[g]	8.3	9.4	8.3	5.24	1.45	0.91	8.62	1.30	10.16	11.60
Minimum	24.8	12.2	22.7	0	3.40	5.50	22.50	0.82	0.01	8.80
Maximum	48.8	43.8	50.1	11.7	6.90	8.00	47.80	4.75	30.00	43.00
Range	24.0	31.6	27.4	11.7	3.50	2.50	25.30	3.93	29.99	34.20

[a] >0.02mm, %; [b] 0.02-0.002mm, %; [c] <0.002mm, %; [d] %; [e] cmol kg^{-1}; [f] mg $100g^{-1}$; [g] standard deviation.

Table 2

Pseudo-total concentrations (µg g^{-1}) of the examined elements in different soil samples (S_1-S_8).

	Pseudo-total element concentrations in µg g^{-1} dry weight							
	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
Al	1413	1419	1853	1879	2214	1669	1851	1532
Ba	78.5	83.2	326	100	86.8	56.8	65.6	114
Cd	0.189	0.164	0.192	0.201	0.149	0.158	0.0750	0.133
Cr	43.5	40.9	62.5	24.4	45.0	501	20.7	24.6
Cu	54.8	45.9	47.4	31.3	31.7	141	40.1	33.4
Fe	11811	14461	23572	14101	11448	24354	12336	11977
K	4417	1811	5059	3480	1472	609	1760	4396
Li	13.7	15.1	22.1	12.9	12.8	15.6	10.6	18.4
Mn	216	704	1837	1247	2408	1380	217	675
Ni	69.0	73.7	62.1	22.6	61.4	529	18.4	28.1
Pb	22.6	34.6	44.4	21.7	35.6	27.5	10.0	19.7
Zn	56.6	53.2	54.7	35.6	31.5	34.9	28.3	36.9

Table 3

Correlation matrix for the pseudo-total concentrations of different elements and physicochemical parameters of the examined soil types.

	Ba	Cu	Fe	K	Li	Mn	Ni	Pb	Sand	Clay	$CaCO_3$	pH_{KCl}	pH_{H2O}	TOC	K_2O
Al	0.17	-0.22	0.01	-0.24	-0.20	0.74^{*}	-0.12	0.19	0.48	-0.15	-0.58	-0.83^{*}	-0.83^{*}	-0.33	0.04
Ba		-0.20	0.50	0.64^{x}	0.81^{*}	0.40	-0.23	0.64^{x}	-0.23	0.21	-0.11	-0.12	-0.08	-0.27	-0.30
Cd		0.06	0.32	0.44	0.40	0.33	0.07	0.58	-0.72^{*}	-0.08	-0.14	0.49	0.53	0.27	-0.06
Cr		0.98^{+}	0.71^{*}	-0.53	0.10	0.20	1.00^{+}	0.09	0.29	-0.72^{*}	-0.24	-0.05	0.12	0.76^{*}	-0.71^{*}
Cu			0.71^{x}	-0.47	0.10	0.05	0.98^{+}	0.06	0.27	-0.69^{x}	-0.23	0.02	0.19	0.76^{*}	-0.67^{x}
Fe				-0.06	0.60	0.37	0.69^{x}	0.51	0.14	-0.56	-0.28	-0.11	0.02	0.37	-0.85^{+}
K					0.57	-0.15	-0.54	0.12	-0.61	0.78^{*}	0.11	0.40	0.33	-0.09	0.35
Mn							0.18	0.72^{*}	0.00	-0.38	-0.33	-0.44	-0.32	-0.01	-0.40
Ni								0.11	0.26	-0.73^{*}	-0.21	-0.02	0.16	0.77^{*}	-0.71^{x}
Sand										-0.36	-0.51	-0.83^{*}	-0.82^{*}	-0.16	-0.14
pH_{KCl}													0.96^{+}	0.37	0.09

$^{+}$ Statistically significant at $p<0.01$ level, * statistically significant at $p<0.05$ level, ^{x} statistically significant at $p<0.10$ level; unmarked correlations are not statistically significant.

Table 4

Correlation matrix for the element contents in different phases of the soil samples (*n*=48).

	Cd	Cr	Cu	Fe	K	Li	Mn	Ni	Pb	Zn
Al	-0.42^{+}	0.37^{*}	0.51^{+}	0.79^{+}	0.53^{+}	0.66^{+}	0.56^{+}	0.39^{+}	0.67	0.66^{+}
Ba	0.26^{x}	-0.08	0.07	-0.02	0.04	0.04	0.31^{*}	-0.07	0.35^{*}	-0.05
Cd		-0.18	-0.28^{x}	-0.31^{*}	-0.16	-0.18	-0.13	-0.18	-0.34^{*}	-0.26^{x}
Cr			0.53^{+}	0.60^{+}	0.09	0.45^{+}	0.12	0.96^{+}	0.22	0.34^{*}
Cu				0.45^{+}	0.18	0.29^{*}	0.33^{*}	0.71^{+}	0.48	0.34^{*}
Fe					0.73^{+}	0.91^{+}	0.17	0.57^{+}	0.37^{*}	0.88^{+}
K						0.81^{+}	-0.11	0.04	0.07	0.89^{+}
Li							-0.08	0.39^{+}	0.18	0.94^{+}
Mn								0.20	0.79	-0.12
Ni									0.31^{*}	0.31^{*}
Pb										0.17

$^{+}$ Statistically significant at p<0.01 level; * statistically significant at p<0.05 level; ^{x} statistically significant at p<0.10 level; unmarked correlations are not statistically significant.

Table 5

Pearson's coefficients and *p*-values for the statistically significant correlations ($p<0.05$) obtained between soil properties and elements in various fractions of the soils.

Element	Phase	Property	r	p	Element	Phase	Property	r	p
Al	F1	Sand	0.846	0.008	Cd	F2	Sand	-0.731	0.039
		Silt	-0.802	0.017			pH	0.811	0.015
	F2	pH	-0.921	0.001	Ni	F0	Clay	-0.737	0.037
Cr	F2	TOC	0.770	0.026			TOC	0.724	0.042
	F3	Clay	-0.741	0.035		F1	Clay	-0.721	0.044
		TOC	0.748	0.033			TOC	0.730	0.04
	F4	TOC	0.786	0.021		F3	Clay	-0.736	0.037
	F5	Clay	-0.716	0.046			TOC	0.772	0.025
		TOC	0.760	0.029		F4	TOC	0.816	0.014
Cu	F2	TOC	-0.742	0.034		F5	Clay	-0.754	0.031
		CEC	-0.712	0.048			TOC	0.740	0.036
	F3	Clay	-0.725	0.042	Pb	F1	Sand	0.859	0.006
		TOC	0.736	0.037			Silt	-0.809	0.015
	F4	TOC	0.762	0.028		F3	Silt	0.728	0.041
Fe	F3	Clay	-0.843	0.009	Zn	F1	Sand	0.717	0.045
K	F4	Clay	0.753	0.031			pH	-0.948	0
	F5	Clay	0.761	0.028		F2	Sand	-0.786	0.021
Li	F2	$CaCO_3$	0.866	0.005			Silt	0.861	0.006
		pH	0.746	0.034			pH	0.751	0.032
	F3	CEC	0.751	0.032					

Table 6

ANN summary.

Network name	Training performance	Test performance	Validation performance	Training error	Test error	Validation error	Training algorithm	Error function	Hidden activation	Output activation
MLP 2-4-12	0.783	0.795	0.853	0.110	0.140	0.180	BFGS 73	SOS	Tanh	Logistic

Table 7

Coefficients of determination, r^2, between experimentally measured and ANN outputs, during the training, testing and validation steps.

r^2	Al	Ba	Cd	Cr	Cu	Fe	K	Li	Mn	Ni	Pb	Zn
Training	0.978	0.625	0.910	0.438	0.572	0.931	0.815	0.919	0.901	0.427	0.920	0.962
Testing	0.988	0.072	0.564	0.965	0.955	0.998	0.992	0.970	0.575	0.616	0.854	0.993
Validation	0.882	0.897	0.854	0.965	0.679	0.964	0.926	0.952	0.561	0.818	0.785	0.949

FIGURES

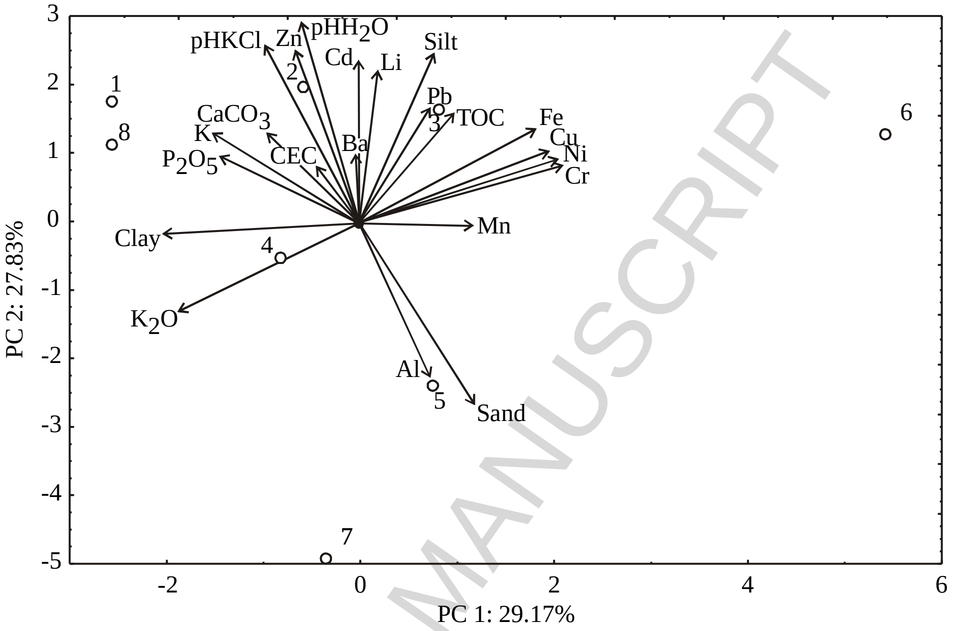

Fig.1. Biplot of correlated pseudo-total element contents and physicochemical properties of the investigated soil types.

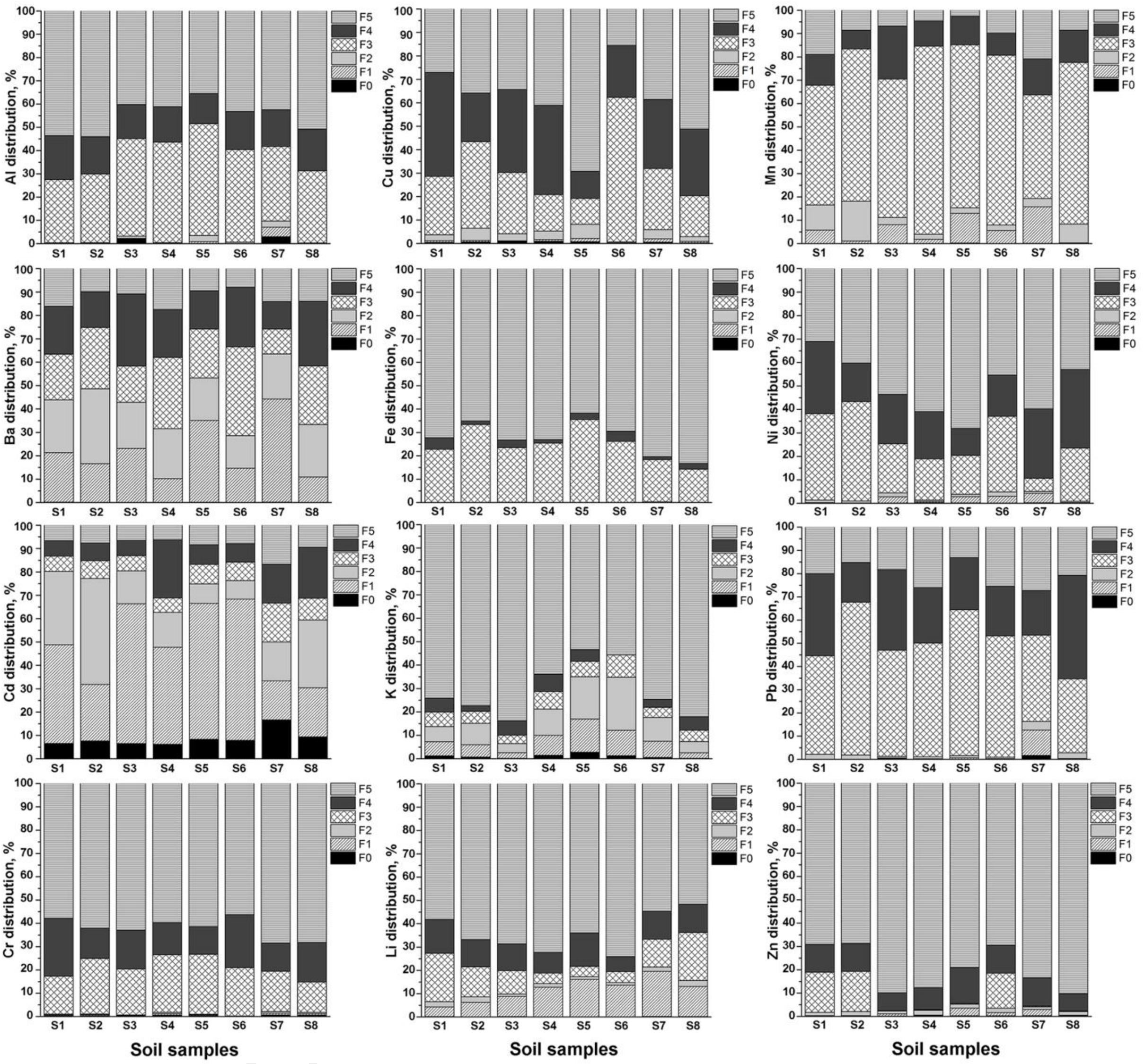

Fig.2. Distribution of Al, Ba, Cd, Cr, Cu, Fe, K, Li, Mn, Ni, Pb and Zn in the operationally defined soil fractions: F_0 - water soluble, F_1 - exchangeable, F_2 - bound to carbonates, F_3 - bound to Fe-, Mn-oxides, F_4 - bound to organic matter, F_5 - residual.

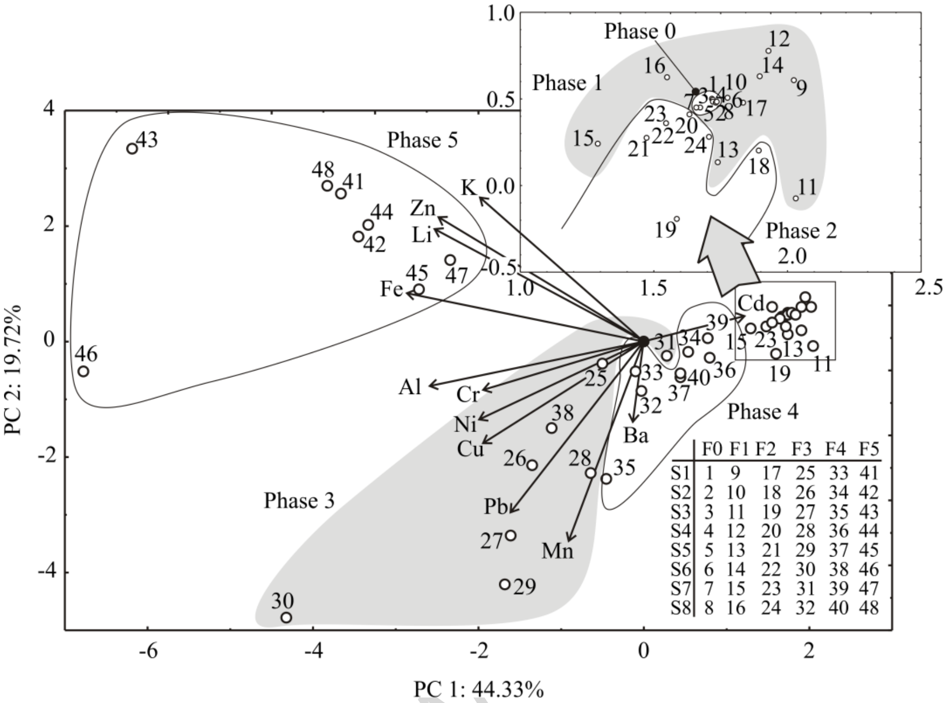

Fig. 3. Biplot of the correlated element contents in different fractions (F_0, F_1, F_2, F_3, F_4 and F_5) in the investigated soil samples (S_1–S_8).

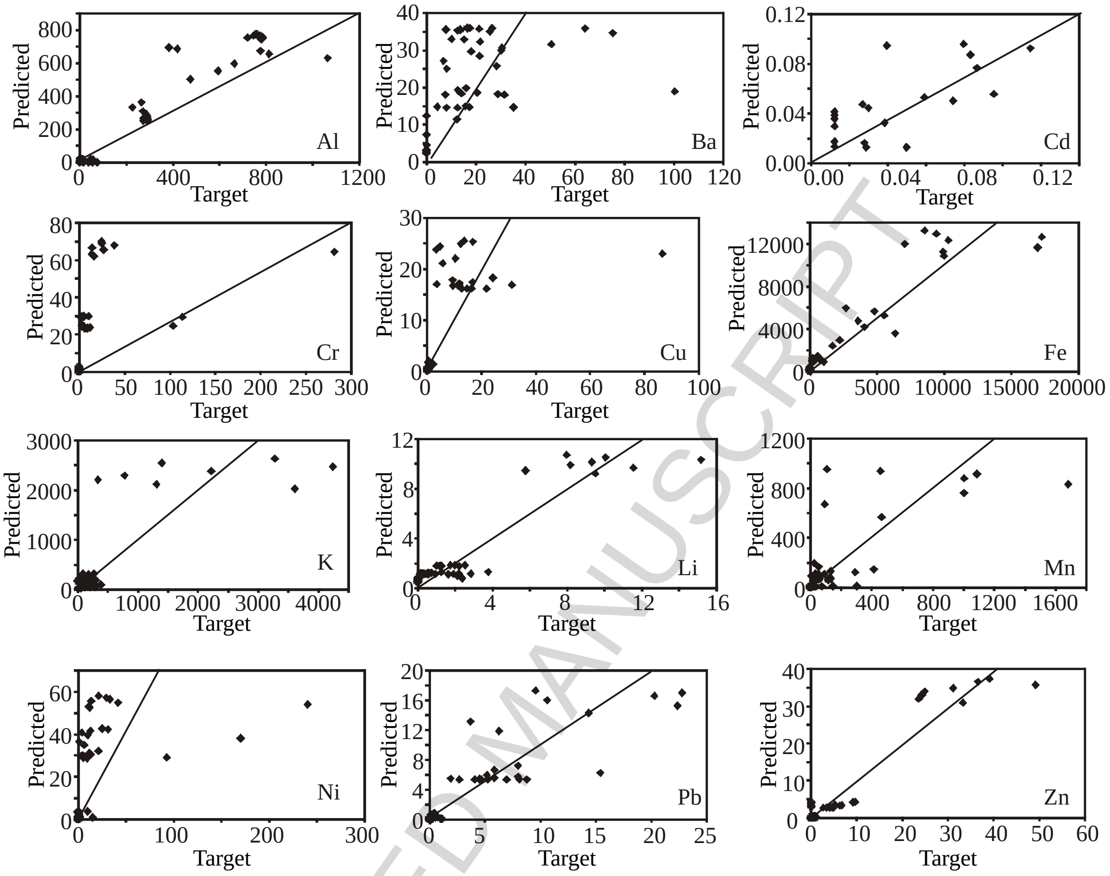

Fig. 4. Experimentally measured and ANN model predicted values of Al, Ba, Cd, Cr, Cu, Fe, K, Li, Mn, Ni, Pb, and Zn in different soil fractions.

Highlights

- Distribution of 12 metals in unpolluted soils was analyzed by sequential extraction.
- Examined metals, except Cd and Ba, were commonly found in less available fractions.
- Soil texture and TOC had a predominant effect on metal content and speciation.
- Higher content of sand fraction affects higher mobility of Pb, Zn and Al.
- ANN model accurately predicts element distribution based on soil properties.