Int. J. Environ. Sci. Technol. (2014) 11:1781–1786
DOI 10.1007/s13762-014-0496-0

SHORT COMMUNICATION

Prediction of copper and chromium concentrations in bean leaves based on an artificial neural network model

N. Hattab · R. Hambli

Received: 1 June 2012 / Revised: 19 November 2013 / Accepted: 3 January 2014 / Published online: 21 January 2014

Abstract The assessment of copper and chromium concentrations in plants requires the quantification of a large number of soil factors that affect their potential availability and subsequent toxicity and a mathematical model that predicts their relative concentrations in plants. While many soil characteristics have been implicated as altering copper and chromium availability to plants in soil, accurate, rapid and simple predictive models of metal concentrations are still lacking for soil and plant analysis. In the current study, an artificial neural network model was developed and applied to predict the exposure of bean leaves (BL) to high concentrations of copper and chromium versus some selected soil properties (pH, soil electrical conductivity and dissolved organic carbon). A series of measurements was performed on soil samples to assess the variation of copper and chromium concentrations in BL versus the soil inputs. The performance of the artificial neural network model was then evaluated using a test data set and applied to predict the exposure of the BL to the metal concentration versus the soil inputs. Correlation coefficients of 0.99981 and 0.9979 for Cu and 0.99979 and 0.9975 for Cr between the measured and artificial neural networks predicted values were found, respectively, during the testing and validation procedures. Results showed that the artificial neural network model can be successfully applied to the rapid and accurate prediction of copper and chromium concentrations in BL.

Keywords Artificial neural networks · Soil · Copper/chromium concentrations · Bean leaves

Introduction

Many soil environment factors have been implicated as potentially affecting trace metal bioavailability in plants. Some of these include partitioning within soil through cation exchange, specific adsorption, precipitation, and complexation and solid–solution partitioning factors including pH, redox potential, soil texture, clay content, organic matter content, electric conductivity, etc. (Sauve′ et al. 2000; Weng et al. 2002).

Estimating the concentration of trace elements in plants is an important step toward assessing the risk associated with their mobility and the search for the most appropriate remediation strategies. However, determining metal concentration in soil experimentally is both expensive and time-consuming.

Several attempts have been made to estimate indirectly the effect of these trace elements from easily measurable soil properties. Generally, two common statistical methods are used to develop prediction models, regression methods and artificial neural networks (ANN). Several multiple linear regression (MLR) models have been developed over

N. Hattab
ISTO, UMR 7327 - CNRS/Université d'Orléan, Campus Géosciences, 1A, rue de la Férollerie, 45071 Orleans Cedex 2, France

N. Hattab
EA 6293 GéHCO Université François-Rabelais de Tours, Parc de Grandmont, 37200 Tours, France

N. Hattab
Laboratory of Soil Science, Faculty of Agriculture, University of Damascus, Damascus, Syria

R. Hambli (✉)
Prisme Institute – MMH, 8, Rue Léonard de Vinci, 45072 Orléans Cedex 2, France
e-mail: ridha.hambli@univ-orleans.fr; ridha0007@yahoo.com

the past 20 years to predict the sorption of trace metals in soil (Tiktak et al. 1998; Schug et al. 2000). When used to model the sorption behavior of trace metals, relationships between basic soil properties and sorption characteristics have to be stated in the regression models a priori. An alternative to MLR is the application of ANN models where such relationships do not need to be formulated beforehand (Anagu et al. 2009).

It has been reported that ANNs provide superior predictive performance compared to conventional mathematical methods including MLR models (Tiktak et al. 1998; Schug et al. 2000). In regression models in many soil engineering situations, the input–output relationships are highly complex and are not well understood. The lack of physical understanding and of a powerful general tool for mathematical modeling leads to either simplifying the problem or incorporating several assumptions into mathematical models. Consequently, many mathematical models fail to simulate the complex behavior of most soil engineering problems.

In the present study, an ANN model was developed as an alternative rapid and accurate tool to estimate the effect of different soil parameters on the concentration of copper (Cu) and chromium (Cr) in bean leaves (BL) grown in 2011 on soils from an industrial site in southwestern France contaminated with Cu, Cr and As which was subjected to wood impregnation treatment and treated-wood recycling.

Table 1 Main characteristics of the measured site

Parameters	Value
Sand (%)	83.5 ± 1.1
Silt (%)	11.5 ± 0.9
Clay (%)	3.8 ± 1.2
C/N	17.2 ± 1.04
OM (g kg^{-1})	15.9 ± 0.89
CEC (cmol kg^{-1})	3.49 ± 0.27
Organic C (g kg^{-1})	9.19 ± 1.79
Total N (g kg^{-1})	0.534 ± 0.045
pH	7.0 ± 0.23
As (mg kg^{-1})	9.8 ± 0.96
Co (mg kg^{-1})	2 ± 0.12
Cu (mg kg^{-1})	674 ± 126
Cr (mg kg^{-1})	23 ± 1.67
Ni (mg kg^{-1})	5 ± 0.67
Zn (mg kg^{-1})	46 ± 3.82

Materials and methods

Sixteen soil samples (four samples at four different locations) were collected from a site located in southwestern France, Gironde County (44°43′N; 0°30′W). This site has been contaminated with high concentrations of copper (Cu) and chromium (Cr) (Mench and Bes 2009). The main characteristics of topsoils at the site are presented in Table 1. Their texture is sandy. Organic matter content is low as well as the cation exchange capacity (CEC). Total soil concentrations were in the common range of French sandy soil for Cr, As and Zn, but the total soil Cu concentration was in excess for these coarse sandy soils, i.e., 35 mg Cu kg^{-1} (Table 1).

One kilogram of each of the 16 soil samples was put in 16 different pots after sieving (2 mm). Then, four grains of dwarf beans (*Phaseolus vulgaris* L. cv vroege Limburgs) were sown in each of the 16 pots and cultivated for 18 days in controlled conditions (16-h light/8-h darkness regime). Sixty-four (16 samples × 4 grains) soil solutions were extracted from all available pots after harvesting of the dwarf beans by Rhizon soil moisture samplers from Rhizosphere research products (Wageningen, Holland). Then, dissolved organic carbon (DOC) was analyzed in the soil solution with a Shimadzu© TOC 5000A. Soil pH and electric conductivity (EC) were determined after mixing the soil with deionized water followed by pH and the EC measurements.

Mobility, availability and/or bioavailability of metals is controlled by several factors such as the soil's chemical and physical properties (pH, EC, DOC, temperature, ion charges, etc.), plant species, metal concentrations in the soil and their related factors (Kabata-Pendias and Pendias 2000; Weng et al. 2002). In the current study, a small-scale field was investigated with a measured homogeneous Cu and Cr variability (Mench and Bes 2009). Therefore, in the current preliminary study, the soil inputs were limited to the three measurable factors considered to be one of the most influential (pH, DOC and EC) on the mobility and availability of metals in the investigated soil. The measured bean leaf variables were the concentrations of Cu and Cr. At a larger-scale fields, other factors may influence the BL Cu and Cr concentrations such as the Cu and Cr concentrations in the soil. The primary aim was to illustrate the potential of the ANN method in its ability for the rapid and accurate prediction of trace element concentrations in plants rather than investigating the role of each soil factor on Cu and Cr concentrations in BL.

An ANN model was developed and trained to predict the exposure of the BL to the high concentration of Cu and Cr trace elements. The model inputs were the soil pH, the soil EC and the DOC. The outputs were the Cu and Cr concentrations in the BL. In the present work, an in-house ANN program called Neuromod written in Fortran (Hambli et al. 2006; Hambli 2009) was applied. The selected architecture is based on one hidden layer with four neurons with a learning rate factor $\eta = 0.1$ and a momentum coefficient $\alpha = 0.1$.

To prepare the training data for the ANN, 64 (16 samples $\times$ 4 grains) measurements were performed. The min/max values of the input–output variables for the training phase are given in Table 2. From these 64 measurements, respectively, 40, 16 and 8 measures which correspond to 62.5, 25 and 12.5 % of the total measurements were selected randomly for training, for testing and for validation (Hunter et al. 2000; Hambli 2009). The testing data were not used for training. The 16 testing data provided random cross-validation as suggested by (Hambli 2009) during the ANN training for verification of the network prediction accuracy.

Table 2 Range of soil parameters for soils in which bean plants were successfully grown

	Minimum	Maximum
Inputs (soil)		
pH	6.97	7.65
EC (mS cm^{-1})	111	255
DOC (mg l^{-1})	21.75	47.72
Output (plant)		
Cu concentration in the BL (mg kg^{-1})	345.17	498.91
Cr concentration in the BL (mg kg^{-1})	6.9	10.86

Results and discussion

The ANN was trained with 3×10^4 epochs (number of training cycles). The training performance was assessed by the root mean square error (RMSE). At the end of the training phase, the RMSE convergence value was 1×10^{-5}.

Correlation coefficients of 0.99981 and 0.9979 for Cu and 0.99979 and 0.9975 for Cr between the measured and ANN predicted values were found, respectively, during the testing and validation procedures (Fig. 1).

The RMSE between the prediction and experimental results was 0.091 % for the Cu and 0.084 % for the Cr. This indicates that the ANN model was able to predict the Cu and Cr concentrations in the BL rapidly and accurately.

To explore the relationships between the outputs (Cu and Cr concentrations in the BL) and the inputs (pH, EC and DOC), several response surfaces (RS) were plotted. Each consisted of a 3D graphical representation of a response plotted between two independent inputs and the outputs. The use of 3D RS plots provides insight into the behavior of the system and enables investigation of the Cu

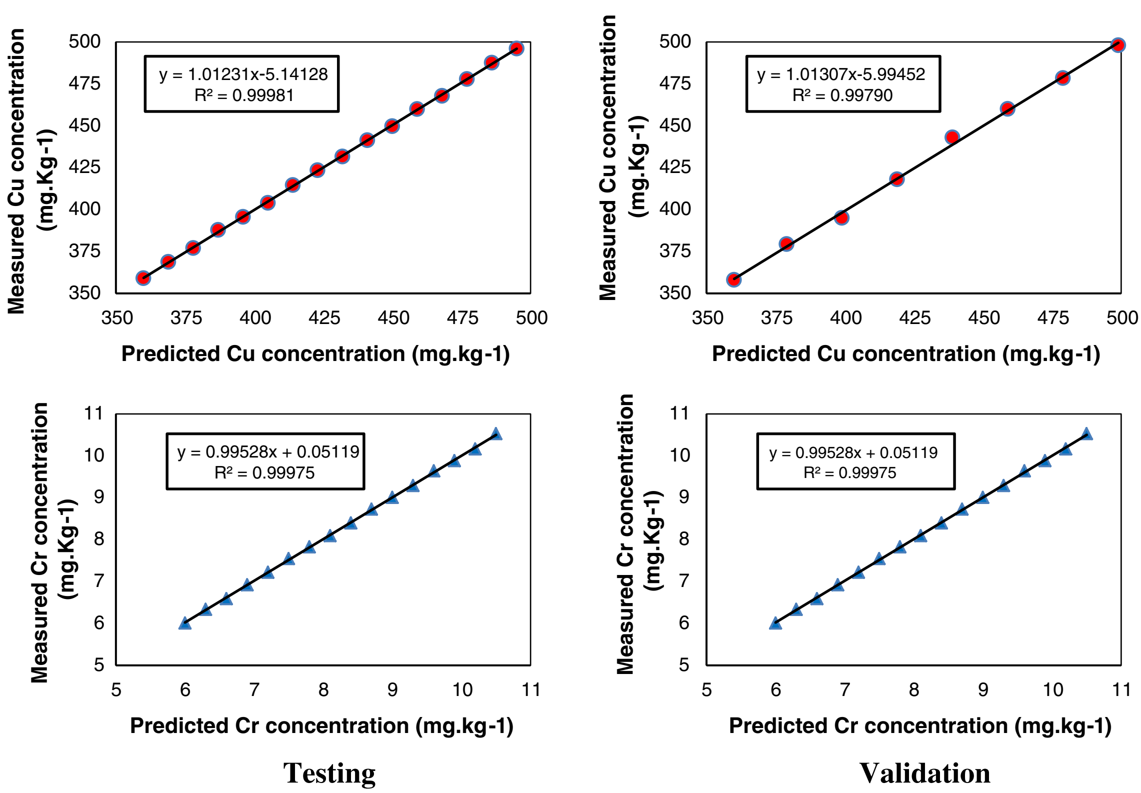

Fig. 1 Predicted (ANN) versus measured results of Cu and Cr concentrations in BL

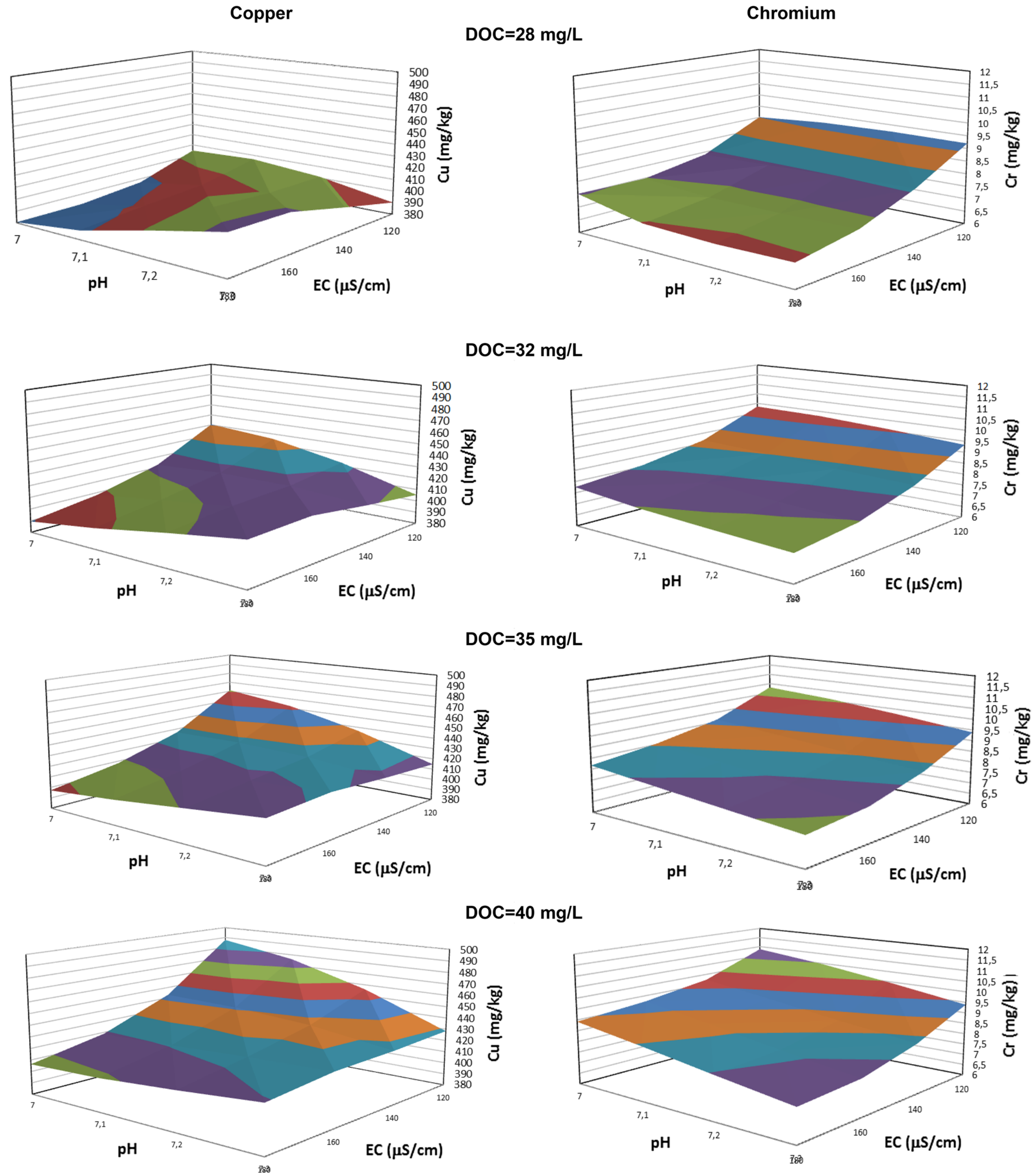

Fig. 2 Copper and chromium concentrations in BL in mg kg^{-1}: effects of soil pH and EC for four different values of DOC

and Cr concentration results versus the levels of the effecting factors predicted by the ANN model (Fig. 2).

The plotted RS indicated that Cu and Cr in BL generally fell within a narrow concentration range (360–500 mg kg^{-1}) and (6–12 mg kg^{-1}), respectively. Despite the fact that Cu exhibits a higher concentration value in the BL, there is also strong evidence that the Cr concentration depends strongly on the soil factors. The influences of pH, EC and DOC factors were found to be in agreement. It can be seen on Fig. 2 that the concentrations increase with decreasing pH, confirming that acidic soil generates an increase in the concentration of Cu and Cr.

We found that among the investigated factors, pH and DOC are the most important parameters affecting the Cu and Cr concentrations in BL. An increase in the DOC level leads to a nonlinear increase in the Cu and Cr concentrations in combination with the pH and EC of soil, indicating strong interactions between the soil inputs. Considering the entire results, one can notice that the Cr tends to evolve from linear responses to nonlinear ones with increasing DOC levels. Based on these results, controlling and monitoring Cu and Cr concentrations in BL can be performed by adjusting the soil pH, EC and DOC by adding specific soil amendments. These amendments reduce trace element mobility by promoting the formation of insoluble precipitates or by enhancing the soil's capacity to bind the trace element. The implementation of ANN modeling for robust and reliable prediction of metal concentrations in plants using additional soil data such as soil particle size, moisture levels and surface debris on soil reflectance in a practical predictive system appears very promising. Future significant work is required to achieve this goal.

Sensitivity analyses (SA) were performed by the ANN to quantify the relative importance of each soil input based on the procedure described by Hunter et al. (2000). First, the network is run using all the variables and the ANN error is recorded ($RMSE_T$). The network is run a second time without a given variable '*Vi*', and the error is recorded ($RMSE_{Vi}$). The measure of the sensitivity of a variable is given by $(RMSE_T - RMSE_{Vi})/RMSE_T$. In other words, the more sensitive a network is to an input variable, the greater the ratio. If the ratio of a variable is less than one, leaving it out either has no effect or improves the network's performance. Finally, the variables are ranked in the order of importance based on the ratio.

The sensitivity correlations (Pareto distribution) between soil inputs (pH, EC and DOC) and outputs (Cu and Cr) are given in Fig. 3.

The SA indicated that soil pH was the most important soil input. This agrees with the findings of Rieuwerts et al.

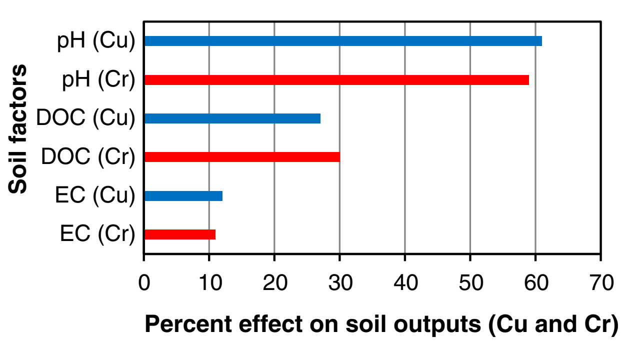

Fig. 3 Sensitivity correlations between soil inputs (pH, EC and DOC) and outputs (Cu and Cr)

(1998) that soil pH is the predominant factor among the basic soil properties influencing heavy metal sorption. Soil DOC and EC ranked, respectively, second and third.

Sauve′ et al. (2000) reported that the availability of the soil metals to the plants depends mainly on physicochemical parameters such as pH and ionic strength. The increase in soil pH leads to an increase in the binding of Cu to soil constituents (McLaren and Crawford 1973) and therefore a decrease in the mobility and availability of soil Cu. Several authors reported that the bioavailability and toxicity of Cu in the soil is increased in acidic soil due to the increase in the concentration of Cu^{+2} (free Cu) in the soil solution which generally presents the available form of Cu to the living organism (Kabata-Pendias and Pendias 2000; Tyler and Olsson 2001). González-Alcaraz et al. (2011) also found that increasing the soil pH reduced the solubility of trace elements, which can decrease the concentration of Cu by 50–60 %. These results indicate that the solubility of Cu is strongly pH dependent (Brown et al. 2005), decreasing sharply when pH > 6.

The increase in pH leads to a slight decrease in Cr concentration in plants. This result can be explained by the fact that in the case of anions [Cr(VI)], the decrease in pH promotes the release of HO ions. This then leads to a decrease in competition between anions and HO ions, which then accelerates the formation of new phases. Thus, the solubility of Cr anions decreases when pH decreases. Dzombak and Morel (1990) reported that when soil pH increases, the mobility of metals decreases, thereby decreasing the uptake of Cr by living species.

In addition, Barcelona and Holm (1991), Pantsar-Kallio et al. (2001) and Seaman et al. (2001) reported that oxidation and reduction reactions are important in the fate, transport and toxicity of Cr in the soil. These reactions are governed by many factors including organic matter, pH, aeration, soil moisture content, wetting and drying, microbial activity, clay mineral content and availability of electron donors and acceptors.

Conclusion

In many situations in soil engineering, the input–output relationships are very complex to determine and are not well understood. The lack of physical understanding and a powerful general tool for mathematical modeling lead to either simplifying the problem or incorporating several assumptions into the mathematical models. Consequently, many mathematical models fail to simulate the complex behavior of most soil engineering problems. In contrast, ANNs are based on the data alone, on which the model can be trained with input–output data patterns to generate a predictive architecture of the ANN after the training phase.

Despite their good performance in many situations, ANNs suffer from a number of limitations. First, they are not able to explain the physical relationships between the input–output data. Second, the precision of the ANN prediction depends mainly on the number of the training samples and the selected input factors affecting the soil responses. In the current primarily work, 16 soil samples were investigated using four measurements for each sample generation 64 measurements. A more general study for a large-scale field requires a higher number of samples and corresponding replicates. Nevertheless, in the current case, the investigation was limited to small contaminated field with total 64 experimental results used for the ANN modeling. Third, it should be noted that the proposed ANN approach does not take into account all the possible factors which may influence the BL Cu and Cr concentrations. The primary aim was to illustrate the potential of the ANN method in its ability for the rapid and accurate prediction of trace element concentrations in plants. The ANN model can be extended by including additional factors and their combinations to capture complex soil–plant responses.

References

Anagu I, Ingwersen J, Utermann J, Streck T (2009) Estimation of heavy metal sorption in German soils using artificial neural networks. Geoderma 152:104–112

Barcelona MJ, Holm TR (1991) Oxidation–reduction capacities of aquifer solids. Environ Sci Technol 25:1565–1571

Brown S, Christensen B, Lombi E, McLaughlin M, McGrath S, Colpaert J, Vangronsveld J (2005) An inter-laboratory study to test the ability of amendments to reduce the availability of Cd, Pb, and Zn in situ. Environ Pollut 138:34–45

Dzombak DA, Morel FMM (1990) Surface complexation modeling: hydrous ferric oxide, vol 393. Wiley, New York

González-Alcaraz MN, Conesa HM, Tercero MC, Schulinb R, Álvarez-Rogel J, Egea C (2011) The combined use of liming and Sarcocornia fruticosa development for phytomanagement of salt marsh soils polluted by mine wastes. J Hazard Mater 186:805–813

Hambli R (2009) Statistical damage analysis of extrusion processes using finite element method and neural networks simulation. Finite Elem Anal Des 45(10):640–649

Hambli R, Chamekh A, Bel Hadj Salah H (2006) Real-time deformation of structure using finite element and neural networks in virtual reality applications. Finite Elem Anal Des 42(11):985–991

Hunter A, Kennedy L, Henry J, Ferguson RI (2000) Application of neural networks and sensitivity analysis to improve prediction of trauma survival. Comput Methods Programs Biomed 62:11–19

Kabata-Pendias A, Pendias H (2000) Trace elements is soils and plants. CRC Press, Boca Raton, FL

McLaren RG, Crawford DV (1973) Studies on soil copper. II. The specific adsorption of copper by soils. J Soil Sci 24:443–452

Mench M, Bes C (2009) Assessment of ecotoxicity of topsoils from a wood treatment site. Pedosphere 19:143–155

Pantsar-Kallio M, Reinikainen SP, Oksanen M (2001) Interactions of soil components and their effects on speciation of chromium in soils. Anal Chim Acta 439:9–17

Rieuwerts JS, Thornton I, Farago ME, Ashmore MR (1998) Factors influencing metal bioavailability in soils: preliminary investigations for the development of critical loads approach for metals. Chem Speciat Bioavailab 10(2):61–75

Sauve′ S, Hendershot W, Allen HE (2000) Solid-solution partitioning of metals in contaminated soils: dependence on pH, total metal burden, and organic matter. Environ Sci Technol 34:1125–1131

Schug B, Düring RA, Gäth S (2000) Improved cadmium sorption isotherms by the determination of initial contents using the radioisotope 109 Cd. J Plant Nutr Soil 163(2):197–202

Seaman JC, Arey JS, Bertsch PM (2001) Immobilization of nickel and other metals in contaminated sediments by hydroxyapatite addition. J Environ Qual 30(2):460–469

Tiktak A, Alkemade JRM, Van Grinsven JJM, Makaske GB (1998) Modeling cadmium accumulation at a regional scale in the Netherlands. Nutr Cycl Agroecosyst 50:209–222

Tyler G, Olsson T (2001) Concentrations of 60 elements in the soil solution as related to the soil acidity. Eur J Soil Sci 52:151–165

Weng LP, Temminghoff EJM, Lofts S, Tipping E, Van Riemsdijk WH (2002) Complexation with dissolved organic matter and solubility control of heavy metals in a sandy soil. Environ Sci Technol 36:4804–4810