Journal of Hazardous Materials 174 (2010) 522–533

Contents lists available at ScienceDirect

Journal of Hazardous Materials

journal homepage: www.elsevier.com/locate/jhazmat

Influence of soil properties on the sorption and retention of cadmium, copper and lead, separately and together, by 20 soil horizons: Comparison of linear regression and tree regression analyses

F.A. Vega, M.L. Andrade, E.F. Covelo*

Departamento de Bioloxía Vexetal e Ciencia do Solo, Facultade de Bioloxía, Universidade de Vigo, Lagoas, Marcosende, 36310 Vigo, Pontevedra, Spain

ARTICLE INFO

Article history:
Received 22 June 2009
Received in revised form
15 September 2009
Accepted 16 September 2009
Available online 23 September 2009

Keywords:
Classification and regression trees
Soil
Sorption
Desorption
Heavy metal
Soil characteristics

ABSTRACT

In this paper we compare linear regression with tree regression for analysis of the influence of soil properties on the sorption and retention of added Cd, Cu and Pb by 20 soil horizons typical of cropped soils in Galicia (N.W. Spain); our measure of sorption/retention capacity was K_r, a recently introduced adimensional parameter. Sorption and retention of Cd was depressed by the presence of Cu and Pb. The soil parameters that were most associated, overall, with differences in Cd, Cu and Pb sorption and retention were cation exchange capacity (CEC), pH and Mn oxides' content. Tree regression, which can take into account variation on both global and local scales, afforded better-fitting models than linear regression, which only reflects global tendencies; but for coherent interpretability of tree regression results it is just as important to avoid overfitting as in the case of linear regression.

1. Introduction

The availability of heavy metals and other soil pollutants for uptake by plants, and the risk of their finding their way into surface or underground waters, depend on their sorption and desorption by soil components, where "sorption" encompasses adsorption, precipitation on soil particle surfaces, and fixation, and "desorption" indicates the release of sorbed species into the medium surrounding the particles by which they had been sorbed [1–3]. The sorption and desorption of metals – especially metal cations, the directly bioavailable form – predominantly involves negatively charged surfaces on organic matter, clay minerals, and metallic oxides or hydroxides, especially Fe and Mn oxides [4–8]. Depending on the relative proportions and compositions of these soil fractions, and on soil texture and total cation exchange capacity (CEC) (which itself depends mainly on soil organic matter and clay content), a soil will have more or less capacity to bind a particular heavy metal species, and will bind it more or less strongly. Accordingly, the distribution of metal cations that are added to the soil among these fractions and others depends on the properties of these fractions as well as on the nature and quantity of the metal added [7,9,10]. In fact, soil properties appear to be more determining than the characteristics of the metal in this respect [11,12]. Given the complexity of soil properties, this makes empirical characterization of their influence on sorption and retention processes essential.

Cd, Cu and Pb are among the potentially most toxic heavy metals, and are present – often together – in numerous polluting spills and in agrochemicals. In a previous work, in which we evaluated the sorption and retention of Cd, Cu and Pb, separately or in mutual competition, by 20 soil horizons, we found that the most useful measure of the capacity of a soil for sorption and retention of these metals was K_r, a novel adimensional parameter that integrates data obtained using a series of different dosage levels [13]. The influence of soil properties on a compound K_r value (a factor score afforded by a principal components analysis of K_r values obtained under various conditions) was modelled by multivariate linear regression (LR). In the present paper we compare the performance of LR of K_r values with that of tree regression [14], which for a series of native soils and different measures and different measures of sorption and retention capacity has been reported to be superior to LR [15].

Data used in previous work [15,16] were from competitive sorption and desorption experiments of Cd, Cr, Cu, Ni, Pb, and Zn. Nevertheless, in this work the models obtained are not only for competitive sorption of Cd, Cu and Pb, but also for individual sorption and retention. Moreover, results from both types of data are compared and discussed.

---

* Corresponding author. Tel.: +34 986812630; fax: +34 986812556.
E-mail address: emmaf@uvigo.es (E.F. Covelo).

0304-3894/$ – see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.jhazmat.2009.09.083

1.1. Sampling

We selected 20 soil horizons as representative of the most widely cropped soil orders in our region that commonly receive inputs containing Cd, Cu and/or Pb. Six samples of each horizon were collected using an Eijkelkamp Model A sampler and were transported in polyethylene bags to the laboratory, where they were air dried, passed through a 2 mm mesh sieve, pooled, and homogenized in a Fritsch Laborette 27 vibratory solid sample homogenizer. The homogenized sample was split into three subsamples, and each of these replicates was in turn subsampled for soil analyses and for sorption/desorption experiments.

2. Variables and methods

2.1. Soil characterization

The following soil properties were determined with a view to evaluating their influence on the capacity of the soil horizons to sorb and retain Cd, Cu and Pb.

Soil pH was determined with a pH meter in 2:1 water/soil suspensions [17]. Particle size distributions were determined following oxidation of organic matter with hydrogen peroxide; the fraction >50 mm was separated by sieving, and the sub-50 mm fraction was characterized as per Day [18]. Organic carbon was quantified by the method of Walkey and Black [19]. CEC and exchangeable cations (Ca^{2+}, Mg^{2+}, K^{+} and Na^{+}; LODs: 5.8, 1, 26, and 4.5 $\mu g\,L^{-1}$, respectively) were determined by extraction with 0.2 M ammonium chloride buffered at the soil pH [20–21], followed by quantification by inductively coupled plasma, atomic emission spectrometry (ICP-OES). Exchangeable acidity was determined using a 1 M KCl replacing solution and titration against 0.1 M NaOH, with phenolphthalein as indicator [22]. Oxides were determined using the method of Mehra and Jackson [23]: samples were shaken in a solution of 0.11 M sodium hydrogen carbonate and 0.27 M sodium citrate, and the Fe, Al and Mn contents of the extracts were determined by inductively coupled plasma, atomic emission spectrometry (ICP-OES) (LODs: 14, 11, and 0,2 $\mu g\,L^{-1}$, respectively). The measured resolution of the system is 0.006 nm at 200 nm. The 80 by 160 mm echelle grating (manufactured by Perkin-Elmer) has 79 lines per mm and a blaze angle of 63.4°.

2.2. Sorption and desorption experiments

Sorption experiments were conducted following the methods of Alberti et al. [24] and Gomes et al. [11], as modified by Harter and Naidu [25]. Non-competitive sorption was evaluated using single-metal sorption solutions of Cu, Cd or Pb nitrates at concentrations of 0.01, 0.03, 0.05, 0.08, 0.1, 0.2, 0.3, 0.4, 0.5, 1, 2 and 3 $mmol\,L^{-1}$; and competitive sorption using multi-metal solutions (Cu + Cd + Pb) in which each metal had the same concentration (again 0.01, 0.03, 0.05, 0.08, 0.1, 0.2, 0.3, 0.4, 0.5, 1, 2 or 3 $mmol\,L^{-1}$). Both single- and multi-metal solutions also contained 0.01 M $NaNO_3$ as background electrolyte. The heavy metals were used in the form of nitrates because of the high solubility of these salts, and the concentrations were chosen to range from normal values to values representative of severe pollution. Triplicate suspensions of 6 g soil samples in 100 mL of solution in polyethylene tubes were shaken in a rotary shaker for 24 h at 25 °C and then centrifuged at 5000 rpm; the pellet was set aside for use in the desorption stage of the experiment; the supernatant was filtered through Whatman 42 paper; the resulting filtrate was analysed by ICP-OES in a Perkin-Elmer Optima 4300 DV apparatus (USA); and the quantity of each metal that had been sorbed was calculated from the difference between its concentrations in solution before the addition of soil and after equilibration (shaking) with the soil.

Following Madrid and Díaz-Barrientos [26], desorption experiments were conducted using the pellets obtained in the sorption phase of the experiments. The pellets were dried at 45 °C and weighed; each pellet was shaken for 24 h in a polyethylene tube with 100 mL of 0.01 M $NaNO_3$ solution at 25 °C, which was then centrifuged at 5000 rpm; the supernatant was filtered through Whatman 42 paper; the resulting filtrate was analysed by ICP-OES; and the quantity of each metal that had been retained on the soil sample was calculated from the quantity sorbed (determined in the sorption stage of the experiment) and the concentration of the metal in solution following desorption.

For both sorption and retention data, and in both cases for competitive and non-competitive situations, the parameter K_r [13] was calculated as follows. Sorption data were fitted with equations of the forms

$$C_{s,i} = K_{r1} C_{p,i} \tag{I}$$

and

$$C_{p,i} - C_{s,i} = K_{r2} C_{p,i} \tag{II}$$

where $C_{s,i}$ is the amount of metal i that was sorbed per gram of soil and $C_{p,i}$ is the amount of metal i that was potentially sorbable, i.e. the amount in the initial sorption solution (per gram of soil). K_r was then defined as K_{r1} if the coefficient of determination of I was larger than that of Eq. (II), and $1 - K_{r2}$ otherwise. When thus calculated from sorption data, K_r varies from 0 for totally non-sorbent soils to 1 for a soil that completely eliminates metal i from solution. To measure capacity for retention of sorbed metal in desorption experiments, K_r was calculated in the same way as for sorption, except that sorbed metal was replaced by metal retained at equilibrium; in this case K_r is 0 for a soil that completely releases all sorbed metal, and for a soil that releases no metal during the desorption phase of the experiment it adopts the value obtained using the corresponding sorption data.

2.3. Regression analyses

For each metal, sorption and retention K_r data obtained as above were regressed on soil properties using both forward stepwise multiple linear regression as implemented in SPSS version 16.0 for Windows, and tree regression as implemented in STATISTICA 7. In the latter case, because of the inclusion of the three replicate samples of each horizon, the tree selected was not chosen on the basis of its cross-validated R^2 value; instead, from the tree sequence generated we chose the smallest tree with more than two terminal nodes in which the intranodal variances of all terminal nodes were no greater than 0.011. As an indication of the degree to which the relationships between the dependent and each independent variable conformed globally to the structure of each regression tree, we also calculated importance values defined by

$$I(j) = \sum_t \Delta_S(j, t)$$

where $I(j)$ is the importance of variable x_j and $\Delta_S(j,t)$ is the reduction in mean squared error S that would be achieved if node t of the tree were split using x_j [14].

3. Results and discussion

The properties of the soil horizons used have been published elsewhere, together with the corresponding K_r values obtained from the experiments described above [13]; these data are summarized in Table 1. The K_r values were regressed on soil properties using both tree regression and LR, and Table 2 compares the corresponding goodness-of-fit values (coefficients of determination R^2, calculated as usual as $\sum_s(\hat{y}_s - \bar{y})^2/\sum_s(y_s - \bar{y})^2$, where $\bar{y}$ is the