# Evaluation of single-extraction methods to estimate the oral bioaccessibility of metal(loid)s in soils

Aurélie Pelfrene, Karin Sahmer, Christophe Waterlot, Philippe Glorennec, Francis Douay, Barbara Le Bot

# Evaluation of single-extraction methods to estimate the oral bioaccessibility of metal(loid)s in soils

Aurélie Pelfrêne[1,*], Karin Sahmer[1], Christophe Waterlot[1], Philippe Glorennec[2], Francis Douay[1], Barbara Le Bot[2]

[1]Yncréa Hauts-de-France, ULR 4515, Laboratoire Génie Civil et géo-Environnement (LGCgE), F-59000 Lille, France

[2]Univ Rennes, EHESP, Inserm, Irset (Institut de recherche en santé, environnement et travail) - UMR_S 1085, F-35000 Rennes, France

*address correspondence to aurelie.pelfrene@yncrea.fr



## Abstract

Incidental ingestion of polluted soil particles exposes the population to toxic metal(loid)s. To refine the methods of exposure and risk assessment, it is relevant to use bioaccessible concentrations of metal(loid)s determined via *in vitro* digestion methods. However, some validated methods are complex and costly, involving high technical skills and numerous reagents. The objective of the present study was to evaluate the suitability of four simple chemical extractions to mimic the bioaccessible fraction of As, Cd, and Pb in the gastric (G) and gastrointestinal (GI) phases obtained using the validated UBM (unified bioaccessibility method) test. Acetic acid (0.11 M), citric acid (0.11 M), EDTA (0.16 M), and hydrochloric acid (HCl, 0.65%) were separately tested in 201 soil samples with a wide range of physicochemical parameters and metal(loid)s concentrations. Significant linear relationships were observed with HCl, EDTA, and to a lesser extent with citric acid. For the cheaper HCl method, correlations

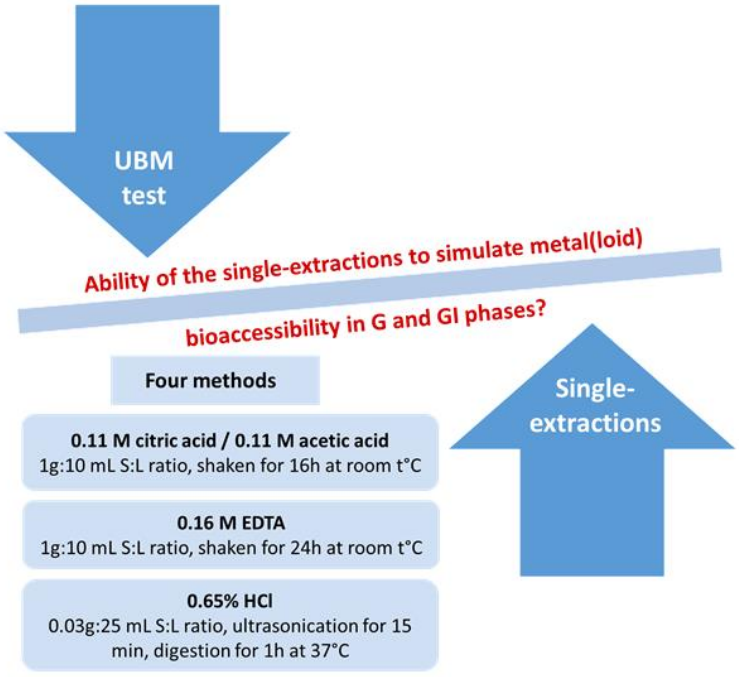
with the UBM ranged from 0.91 to 0.99 for the G phase and from 0.72 to 0.97 for the GI phase. This test can be used at least as a first-tier screening to assess the oral bioaccessibility of As, Cd, and Pb.

**Keywords:** soil pollution; metal; metalloid; human exposure; single-extraction method

## Highlights

- Single-extraction methods to mimic metal bioaccessibility in a wide range of soils
- Miming both the gastric and intestinal bioaccessibility of As, Cd and Pb in soils
- HCl (0.65%) can be used in first tier screening to assess metal bioaccessibility

## Graphical abstract



## 1. Introduction

The major sources of metal(loid)s in the environment are from mining and smelting activities, which can result in considerable soil contamination. Soils enriched with metal(loid)s such as arsenic (As), cadmium (Cd), and lead (Pb) pose a potential threat to human health if directly ingested or transferred through food. Incidental ingestion of soil particles is considered an important exposure pathway of metal(loid)s, more specifically for children through outdoor hand-to-mouth activities (Ruby et al., 1993; Paustenbach, 2000). In the context of polluted-site management, human health risk assessment is usually based on the total (or pseudo-total) metal(loid) concentrations in the soil. However, if considering that the whole amount of a contaminant is available for organisms after ingestion of soil particles, health risks are often overestimated. Therefore, to take into account the concept of bioavailability (i.e., the quantity of elements that reaches the bloodstream and human organs; Ruby et al., 1999), bioaccessibility may be used and was defined as the fraction of a contaminant that is soluble in the gastrointestinal environment and potentially available for absorption (Ruby et al., 1999; Oomen et al., 2002). Research on bioaccessibility related to contaminated-site management has attracted attention at an international level (e.g., Rodriguez and Basta, 1999; Pouschat and Zagury, 2006; Drexler and Brattin, 2007; Barsby et al., 2012; Brattin et al., 2013; Jeong et al., 2013; Pelfrêne et al., 2015; Xia et al., 2017; Tang et al., 2018) and information on this parameter can promote a more balanced assessment of sites (Alexander, 2000; Ollson et al., 2009; Latawiec et al., 2010). To mimic the availability of metal(loid)s in the human gastrointestinal tract (i.e., bioaccessibility), many in vitro extraction procedures (described in the literature as physiologically based extraction tests) were developed to be applied to soils. However, to refine the method of exposure and risk assessment, it is recommended to use bioaccessible concentrations of metal(loid)s measured according to validated and standardized protocols. In most European countries, the unified bioaccessibility method (UBM) is considered as the oral

bioaccessibility reference method because it has been validated against an in vivo model (young swine) for As, Cd, and Pb (Denys et al., 2012) and was also standardized (ISO 17924, 2016). However, this method is complex and time consuming, involves high technical skills, and requires numerous chemical and biological reagents to mimic biochemical conditions in the gastrointestinal tract. Thus, a simplified method to measure bioaccessibility might be of great interest. Single-extraction methods are used in risk assessment to dissolve metal(loid)s from soil (Bonten et al., 2008; Römkens et al., 2009; Groenenberg et al., 2010) and have given good estimates of the potential availability or geochemical reactivity as well as of the bioaccessibility of metal(loid)s (e.g., Drexler and Brattin, 2007; Madrid et al., 2008; Brattin et al., 2013; Rodrigues et al., 2013; Li and Zhang, 2013; Cruz et al., 2015; Li et al., 2015; Mendoza et al., 2017; Rodrigues et al., 2018). Because for most metallic elements, the largest part of the bioaccessible fraction is released under acidic conditions in the stomach, simplified in vitro methods in single-step extractions have been developed simulating only the action of gastric juices. Ideally, robust but easy extraction methods are needed to assess bioaccessible fractions in humans, i.e., the main available pool of metal(loid)s in soil. The reactive or adsorbed pool is often estimated by dilute acids ($HNO_3$ or HCl), chelating agents (e.g., EDTA) or other extraction solvent (e.g., glycine buffer) (Tipping et al., 2003; Rodrigues et al., 2010a, 2010b; Groenenberg et al., 2010; USEPA, 2007, 2017). To this end, the USEPA has established an analytical procedure for the validated in vitro bioaccessibility assay (IVBA) for Pb and As in soil (USEPA, 2007, 2017). Measurements of IVBA using this assay (0.4 M glycine adjusted to a pH of 1.5 using HCl) have been shown to be a reliable predictor of in vivo relative bioavailability of Pb ($R^2$ of 0.92) and As ($R^2$ of 0.87) in a wide range of soil types, anthropogenic sources and elemental forms. In recent years, owing to its simplicity, the 0.43 M $HNO_3$ extraction procedure has also been suggested as a potential in vitro method to determine bioaccessible metal(loid)s in soils. In the study of Li et al. (2015) carried out on 11 contaminated

soils, the results showed that $HNO_3$-extractable As was well correlated with bioaccessible As in the gastric phase using UBM, SBET, and PBET (physiologically based extraction test) ($R^2$ = 0.70–0.85). In 204 soils from Portugal, Brazil, and The Netherlands, Rodrigues et al. (2018) also highlighted a close 1:1 relationship between results from the single-$HNO_3$-extraction method and bioaccessible Cd and Pb in the gastric phase measured with the SBET assay. In a representative subset of samples ($n$ = 30), they showed good correlations between 0.43 M $HNO_3$ and the UBM assay in the gastric phase.

In study by Li and Zhang (2013), the bioaccessibility of Pb was measured using the PBET assay in 14 mildly acidic and alkali (pH 5.9–8.3) soils. The availability of Pb was also measured using different single-extraction methods (0.1 M $HNO_3$, 0.4 M acetic acid, 0.05 M EDTA, and 0.5 M DTPA) and it was suggested that EDTA extraction can be used to predict the bioaccessibility of Pb in these soils. These different studies reported good results for assessing the bioaccessibility in the gastric phase of metal(loid)s (mainly Pb) from soil. By contrast, the use of a single-extractant method has been scarcely studied to simulate the availability of metal(loid)s in the intestinal phase of the digestion process. In a previous work (Waterlot et al., 2017), in 27 soil samples from the north of France, we highlighted significant relationships between Cd and Pb extracted in both the gastric and gastrointestinal phases by using UBM and single-extraction methods. More specifically, the best relationships were established using acetic and citric acids for Cd, whereas for Pb, citric acid and EDTA were identified as the best extractants.

These preliminary studies yielded encouraging results but in most of the cases (except for the work of Rodrigues et al., 2018), a limited number of soil samples located in a specific environmental context were assessed. It has been clearly shown that the bioaccessibility of metal(loid)s in soils critically depends on soil type (Ruby et al., 1996; Kelley et al., 2002; Poggio et al., 2009; Pelfrêne et al., 2015), the chemical speciation of pollutants (Denys et al.,

2007), and the solid-phase distribution in soils (Beak et al., 2006; Denys et al., 2007; Wragg et al., 2007; Reis et al., 2014). Therefore, the challenge is to find single-extraction methods that could be applicable to various soil types.

The objective of the present study was to evaluate the potential suitability of four chemical single-extraction methods using acetic acid (0.11 M), citric acid (0.11 M), EDTA (0.16 M), and hydrochloric acid (HCl; 0.65%) to mimic the bioaccessible fraction of As, Cd, and Pb obtained with UBM in soil samples with a wide range of physicochemical parameters and metal(loid)s concentrations. The interest was to provide information on the applicability of simplified method(s) to study jointly both the gastric and the intestinal phases.

This paper presents a new statistical modeling approach to establish a robust model from an extended database of 201 soil samples by using: (i) a training set of 140 samples to construct both linear regression models (partial least squares regression and simple linear regression) and to select the main chemical extractants to estimate the bioaccessibility of As, Cd, and Pb in both phases, and (ii) an independent test set of 61 samples to validate the regression models and to compare both regressions according to statistical indicators.

## 2. Materials and methods

### *2.1. Soil samples and characterization of soil parameters*

A total of 201 soil samples were collected in 25 polluted sites located in France. These sites had different uses and have been classified into the following categories (Table 1): industrial activities (current and past), former smelting activities, former gas stations, urban gardens, brownfields, and agricultural activities (naturally and artificially contaminated). These samples included a wide range of physicochemical parameters and metal(loid)s concentrations.

**Table 1.** Soil parameters and pseudototal concentrations of As, Cd, and Pb ($n$ = 201) for the different soil samples according to their categories.

| Categories | No of samples | | Soil parameters | | | | | | Pseudototal concentrations | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Clay | Silt | Sand | OM | pH (in water) | Total $CaCO_3$ | As | Cd | Pb |
| | | | % | % | % | % | | g $kg^{-1}$ | mg $kg^{-1}$ | mg $kg^{-1}$ | mg $kg^{-1}$ |
| Industrial activities | 29 | Min | 13 | 32 | 6.3 | 0.1 | 7.4 | 1.4 | 1.9 | 0.1 | 9.1 |
| | | Mean | 24 | 56 | 21 | 1.9 | 8.1 | 144 | 10.6 | 13.8 | 944 |
| | | Max | 53 | 73 | 51 | 7.3 | 8.4 | 734 | 49.8 | 329 | 12330 |
| Former industrial activities | 33 | Min | 6.8 | 16 | 12 | 1.6 | 7.5 | 4.0 | 6.3 | 0.2 | 19 |
| | | Mean | 14 | 42 | 44 | 22 | 8.0 | 48 | 35.1 | 26.3 | 1636 |
| | | Max | 26 | 66 | 74 | 67 | 9.1 | 333 | 199 | 483 | 9790 |
| Former smelting activities | 84 | Min | 11 | 19 | 6.6 | 2.2 | 4.3 | 0.3 | 5.5 | 0.2 | 31 |
| | | Mean | 21 | 41 | 39 | 8.8 | 7.3 | 24 | 30.7 | 10.5 | 684 |
| | | Max | 43 | 71 | 70 | 36 | 8.3 | 162 | 228 | 257 | 8971 |
| Former gas stations | 2 | Min | 11 | 17 | 45 | 4.2 | 7.5 | 86 | 13.3 | 0.3 | 137 |
| | | Mean | 17 | 24 | 59 | 5.1 | 7.6 | 128 | 17.7 | 0.4 | 225 |
| | | Max | 24 | 31 | 72 | 5.9 | 7.8 | 169 | 22.2 | 0.5 | 312 |
| Urban gardens | 10 | Min | 10 | 20 | 38 | 2.8 | 6.6 | 1.0 | 11.2 | 0.1 | 90 |
| | | Mean | 15 | 35 | 49 | 7.2 | 7.4 | 42 | 34.4 | 1.5 | 341 |
| | | Max | 18 | 42 | 70 | 16 | 8.3 | 123 | 94.5 | 4.3 | 660 |
| Brownfields | 6 | Min | 6.6 | 24 | 34 | 1.5 | 7.8 | 36 | 7.0 | 0.2 | 268 |
| | | Mean | 9.0 | 33 | 58 | 2.9 | 8.1 | 65 | 10.8 | 0.3 | 1249 |
| | | Max | 13 | 53 | 70 | 5.7 | 8.2 | 95 | 21.7 | 0.5 | 2659 |
| Agricultural activities | 10 | Min | 8.0 | 1.2 | 10 | 0.9 | 4.9 | 0.6 | 3.3 | 0.1 | 18 |
| | | Mean | 18 | 29 | 51 | 2.6 | 6.7 | 8.9 | 14.6 | 1.0 | 71 |
| | | Max | 42 | 73 | 84 | 5.1 | 8.2 | 31 | 74.4 | 4.5 | 288 |
| Agricultural activities - spiked soils | 27 | Min | 18 | 48 | 4.0 | 2.1 | 4.7 | 1.0 | 3.4 | 0.5 | 231 |
| | | Mean | 35 | 55 | 6.3 | 2.8 | 6.5 | 35 | 13.7 | 4.6 | 540 |
| | | Max | 47 | 69 | 13 | 6.4 | 8.3 | 110 | 17.0 | 8.4 | 2848 |

OM: organic matter

The soil samples were air-dried at a temperature below 40°C and sieved at 2 mm (Sieve Shaker AS 200 Control, Retsch). Particle-size distribution (i.e., clay, silt, and sand contents) was obtained through sedimentation and sieving (NF X 31−107). Soil pH was measured in water suspension (NF ISO 10390), and organic matter content was obtained with the NF ISO 10694 standard (i.e. determination of organic carbon by dry combustion). Total carbonate content was determined by measuring the volume of $CO_2$ released after a reaction with HCl (NF ISO 10693). For each of the 201 soil samples, a representative subsample was sieved at 250 µm (Sieve Shaker As Control, Retsch). The pseudototal metal(loid) concentrations (As, Cd, and Pb) in the soil subsamples were obtained by Hot Block system-assisted digestion (Environmental Express® SC100, Charleston, SC, USA) and determined by inductively coupled plasma mass spectrometry (ICP-MS, 7900 Agilent Technologies, France). More specifically, 300 mg of soil samples was digested in a mixture of 1.5 mL $HNO_3$ (70%) and 4.5 mL HCl (37%) at 95°C for 90 min. After mineralization (Waterlot et al., 2012), digestion products were completed to 25 mL with ultra-pure water (resistivity 18.0 MΩ $cm^{-1}$) and stored at 4°C prior to analysis. All

precautions were taken regarding the protocol application and the calibration. Quality control was based on the use of a certified sample (BCR CRM 141) and an internal control and provided good recoveries for As, Cd, and Pb (98−105% for CRM 141, $n$ = 4, and 84–104% for internal control, $n$ = 5). Seven blanks were included and the limits of detection (LD) were calculated (LD of 0.2 µg $L^{-1}$ for As, 0.1 µg $L^{-1}$ for Cd, and 1 µg $L^{-1}$ for Pb).

### *2.2. Chemical extraction methods*

#### *2.2.1. In vitro oral bioaccessibility measurement*

Bioaccessibility of As, Cd, and Pb was assessed using a widely employed, validated, and standardized in vitro extraction method, the unified bioaccessible method (UBM; Denys et al., 2012; ISO 17924, 2016). The simulated digestive fluids were prepared individually 1 day before the extraction (Table 2). In the UBM protocol, 0.6 g of soil subsamples was placed in centrifuge tubes and mixed with 9 mL of saliva. After quick manual shaking (10 s), 13.5 mL of gastric solution was added and the pH of the solution was adjusted to 1.20 ± 0.05 with HCl (37%). The tubes were shaken end-over-end at 37°C for 1 h and centrifuged at 4500 *g* for 5 min (Rotanta 460 Hettich, Tuttlingen, Germany). The first stage constituted the gastric-only phase (G phase). To simulate the gastric and intestinal phases together, a duplicate gastric phase solution was produced on which 27 mL of duodenal fluid and 9 mL of bile solution were added to give a final pH ranging from 6.25 to 6.35 (adding 10 M NaOH). The tubes were shaken end-over-end at 37°C for 4 h and centrifuged at 4500 *g* for 5 min. This second phase constituted the gastrointestinal phase of the extraction technique (GI phase). The resulting supernatants were removed and bioaccessible concentrations of As, Cd, and Pb were measured by ICP-MS. In vitro metal(loid) bioaccessibility was determined in triplicate soil samples. To evaluate analytical recovery, blanks and an NIST Standard Reference Material (SRM 2710a) were used. Mean recoveries in SRM 2710a ($n$ = 22) ranged from 83 to 128%, 94 to 113%, and 83 to 129%

for As, Cd, and Pb, respectively, in the G phase, and from 83 to 130%, 71 to 126%, and 71 to 148% for As, Cd, and Pb, respectively, in the GI phase. The LDs were calculated from the blanks ($n$ = 22) (LD of 0.4 µg $L^{-1}$ for As, 0.2 µg $L^{-1}$ for Cd, and 2 µg $L^{-1}$ for Pb in both G and GI phases).

**Table 2.** Composition of the digestive solutions used during the phases of the UBM (from Pelfrêne et al., 2011).

| Saliva | Gastric juice | Intestinal juice | Bile |
|---|---|---|---|
| | *Inorganic solution (500 mL)* | | |
| KCl 896 mg | NaCl 2752 mg | NaCl 7012 mg | NaCl 5259 mg |
| $NaH_2PO_4$ 888 mg | $NaH_2PO_4$ 266 mg | $NaHCO_3$ 5607 mg | $NaHCO_3$ 5785 mg |
| KSCN 200 mg | KCl 824 mg | $KH_2PO_4$ 80 mg | KCl 376 mg |
| $Na_2SO_4$ 570 mg | $CaCl_2$ 400 mg | KCl 564 mg | 180 µL HCl 37% |
| NaCl 298 mg | $NH_4Cl$ 306 mg | $MgCl_2$ 50 mg | |
| 1.8 mL NaOH 1.0 M | 8.3 mL HCl 37% | 180 µL HCl 37% | |
| | *Organic solution (500 mL)* | | |
| Urea 200 mg | Glucose 650 mg | Urea 100 mg | Urea 250 mg |
| | Glucuronic acid 20 mg | | |
| | Urea 85 mg | | |
| | Glucosamine hydrochloride 330 mg | | |
| | *Added compounds* | | |
| Amylase 145 mg | Bovine albumine 1g | $CaCl_2$ 200 mg | $CaCl_2$ 222 mg |
| Mucin 50 mg | Mucin 3 g | Bovine albumine 1g | Bovine albumine 1.8 g |
| Uric acid 15 mg | Pepsin 1 g | Pancreatin 3 g | Porcine bile 6 g |
| | | Lipase 500 mg | |
| | *pH of solutions* | | |
| 6.5 ± 0.5 | 1.1 ± 0.1 | 7.4 ± 0.2 | 8.0 ± 0.2 |

### *2.2.2. Single-extraction methods*

Based on previous studies (Le Bot et al., 2010, 2011; Waterlot et al., 2017), 0.11 mol $L^{-1}$ acetic acid (AA), 0.11 mol $L^{-1}$ citric acid (CA), 0.16 mol $L^{-1}$ EDTA, and 0.65% HCl were selected as single-extraction solutions.

For AA, CA, and EDTA extractions (Waterlot et al., 2017), 1 g of each soil subsample was weighed in triplicate and was placed into centrifuge tubes before adding 10 mL of one single-extraction solution in order to have a 1/10 (weight/volume) soil/extractant ratio. Each mixture

was shaken on a rotor disc at 10 r $min^{-1}$ for 16 h for AA and CA extractions or 24 h for EDTA extraction at room temperature. The extract was separated from the solid residue by centrifugation at 4500 *g* for 20 min. Afterwards, the supernatant was filtered over an acetate Millipore membrane (0.45-µm porosity, Minisart, Sartorius, Göttingen, Germany), acidified by adding $HNO_3$ (70%), and stored at 4°C until analysis by ICP-MS. The EDTA solution was adjusted to pH 7.5 with ammoniac. Quality control was based on an internal control and provided good recoveries for As, Cd, and Pb (92−106% for AA, $n$ = 6, 90–112% for CA, $n$ = 6, and 94–104% for EDTA, $n$ = 6).

To improve assessment of human exposure and risk, a protocol using HCl was developed for measuring the bioaccessible (stomach accessibility) concentrations of metal(loid)s, and more specifically Pb, in wipe samples of dust and soil (Le Bot et al., 2010) and then adapted directly on dust or soil samples (Le Bot et al., 2011). For HCl extraction, 0.03 g of each soil subsample was weighed in triplicate and was placed into polyethylene tubes before adding 25 mL HCl at 0.65%. To mimic peristalsis in the stomach, the samples were ultrasonicated (Elmasonic S120, Grosseron, France) for 15 min and then digested in a graphite block (Digiprep System, SCP Science) at 37°C for 1 h. Samples were then ultrasonicated for another 15 min. After volume adjustment to 40 mL, a 20-mL aliquot was filtered (0.45-µm filter, Millipore, St Quentin en Yvelines, France), the solution was diluted with 5% $HNO_3$ and 9.35% HCl before analysis by ICP-MS. Quality control was based on two reference material samples (EnvironMAT™ from SCP Science) and the mean recoveries for As, Cd, and Pb were 89–115% for reference no. 140-025-001, $n$ = 15, and 75–110% for reference no. 140-025-002, $n$ = 9.

### *2.3. Model estimation and validation*

The 201 soil samples were randomly divided into two sets to create a training set ($n$ = 140) and a test set ($n$ = 61; i.e., 30% of the soil samples).

The first approach consisted in model estimation from the set of 140 soil samples. This training set was used to construct linear regression models and to select the best chemical extraction method(s) to estimate the metal(loid) bioaccessibility in both the G and the GI phases. To construct the best regression models, a double statistical analysis was conducted on the training set. Two types of models were estimated: (i) partial least squares (PLS) regression using the four extraction methods as explanatory variables and (ii) simple linear regression separately for each extraction method.

In a previous study (Waterlot et al., 2017) carried out on a limited number of soil samples located in a former smelting area, we showed that the best relationships for predicting bioaccessibility were established using AA and CA for Cd, whereas for Pb, CA and EDTA were identified as the best extractants. Other experiments highlighted the value of using HCl as a suitable proxy to assess the bioaccessibility of As, Cd, and Pb (Le Bot et al., 2010). According to these results, linear multiple regressions, and more specifically PLS regressions (because of multicollinearity among extractants), were performed in an initial step to determine whether combining the four extractants allowed for the best estimation to be obtained. The results of PLS regression with all four components were considered. In the second step, simple linear regressions were used for each of the four extractants. The prediction ability of each of the five models (PLS regression and four simple linear regression models) was analyzed on the test set of 61 soil samples. This set was not included in the construction of the models but was only used for validation.

The model performance was evaluated by calculating the coefficient of determination ($R^2$) in the training set and the root mean square error of prediction (RMSEP) in the test set. This allowed us to compare the accuracy of the prediction and the ability of the chemical extractions to simulate metal(loid) bioaccessibility:

$$\mathrm{RMSEP} = \left[\frac{1}{n}\sum_{i=1}^{n}(B_i - \beta_i)^2\right]^{1/2}$$

where $B_i$ is the measured bioaccessibility at location i, $\beta_i$ is the predicted bioaccessibility at location i, and $n$ is the sample size. Smaller RMSEP values indicate good predictions and that the extractant performs well.

Statistical analyses were performed using R (R Core Team, 2019) and its package pls (Mevik et al., 2019) and XLSTAT 2019.3.1 (Addinsoft). All data were log transformed because of the log-normal distribution of the variables.

## 3. RESULTS AND DISCUSSION

### *3.1. General properties of the soils*

The physicochemical characterization and the degree of contamination of As, Cd, and Pb in the 201 soil samples selected for this study are presented in Table 1. The texture of the soil samples was highly variable since the percentage of clay, silt, and sand ranged from 6.6 to 53.2%, from 1.2 to 73.4%, and from 4.0 to 83.7%, respectively. On average, soils from areas with current industrial and agricultural activities, and from brownfields, had low OM content (1.9, 2.6, and 2.9%, respectively), while soils from areas with former industrial activities had the highest content of organic matter (OM; 22%). Some soil samples of this last category were also contaminated by organic pollutants, and more specifically by polycyclic aromatic hydrocarbons (PAHs, from a former gas plant). PAHs are among the compounds taken into account in the determination of organic matter content, which explain the high percentage obtained for some samples. The pH in the database varied from 4.3 to 9.1 and wide ranges were observed in all categories of soils. The soils were characterized by a very large variation in the content of carbonates (0.3–734 g $kg^{-1}$), where soils from agricultural activities tend to have lower values (mean of 8.9 g $kg^{-1}$) compared with soils from current industrial activities (mean of 144 g $kg^{-1}$). Contamination levels of As, Cd, and Pb also varied widely within and between the different categories of soils and varied from very low levels (1.9, 0.1, and 9 mg $kg^{-1}$, respectively, for

As, Cd, and Pb), reflecting relatively undisturbed soils, to extremely high levels (228, 483, and 12,300 mg $kg^{-1}$, respectively, for As, Cd, and Pb) in polluted soils comprising mainly soils from industrial and smelting activities and from brownfields. These soils with different parameters made them good candidates for investigating the relationships between the different extractable concentrations of metal(loid)s.

The 201 soil samples were then divided into a training set of 140 soils used for initial model development, and a test set of 61 soils used to independently validate the models. The physicochemical parameters and metal(loid)s concentrations for both sets are presented in Table S1 of the Supporting Information.

*3.2. Extractable concentrations of metal(loid)s in the study soils*

The concentrations of metal(loid)s in both the G and GI phases obtained using the single-extraction methods and UBM are presented in Figure 1. Their ranges varied from very low to extremely high.

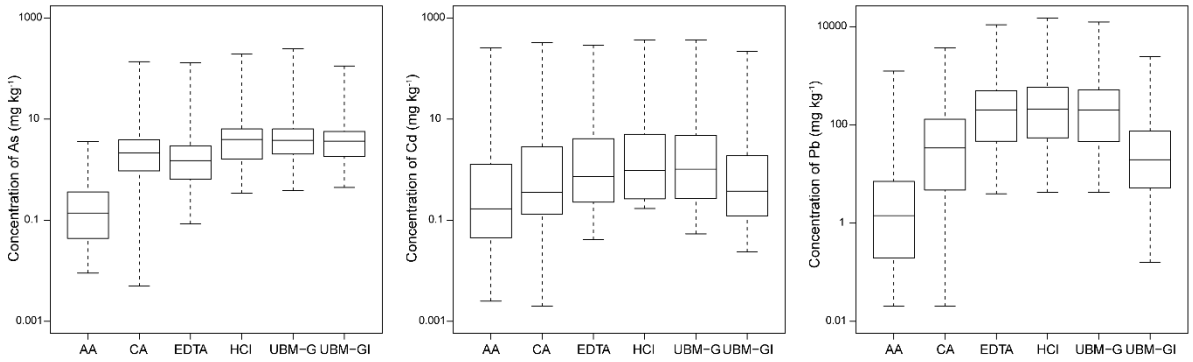



**Figure 1.** Boxplots of metal(loid)s concentrations extracted by acetic acid (AA), citric acid (CA), EDTA, HCl, and UBM in the gastric (UBM-G) and gastrointestinal (UBM-GI) phases ($n$ = 201; expressed in mg $kg^{-1}$).

For As, the median values extracted by HCl (3.91 mg $kg^{-1}$) were similar to those obtained for UBM-G (3.83 mg $kg^{-1}$) and UBM-GI (3.64 mg $kg^{-1}$). For Cd, the median values obtained with HCl, and to a lesser extent with EDTA, and UBM-G were very close (0.76, 0.64, and 0.73 mg $kg^{-1}$, respectively). It can also be noted that the median values of Cd extracted by CA (0.37 mg $kg^{-1}$) were close to those obtained for UBM-GI (0.29 mg $kg^{-1}$). For Pb, the median concentrations extracted by both EDTA and HCl (189 and 198 mg $kg^{-1}$, respectively) were similar to those measured by UBM-G (185 mg $kg^{-1}$).

Expressed in percentage of total concentration, the extraction efficiencies of the methods varied according to the soil and metal(loid)s (0.03–86.3% for As, 0.9–100% for Cd, and 0.05–100% for Pb; Table 3).

**Table 3.** Concentrations of metal(loid)s (expressed in percentage of total concentration) determined in the soil samples ($n$ = 201) with the different extraction methods.

| | | Simple extraction methods | | | | UBM | |
|---|---|---|---|---|---|---|---|
| | | AA | CA | EDTA | HCl | G phase | GI phase |
| | | % | | | | % | |
| As | min | 0.03 | 0.1 | 1.1 | 2.3 | 8.0 | 6.0 |
| | median | 0.8 | 10.9 | 9.2 | 19.8 | 21.9 | 20.1 |
| | mean | 1.5 | 15.9 | 13.2 | 25.2 | 26.0 | 23.2 |
| | max | 12.8 | 73.0 | 62.2 | 86.3 | 86.0 | 64.8 |
| Cd | min | 1.1 | 0.9 | 3.7 | 5.8 | 21.7 | 6.2 |
| | median | 19.4 | 52.9 | 73.2 | 85.4 | 82.7 | 35.7 |
| | mean | 20.2 | 46.6 | 71.7 | 83.5 | 80.5 | 37.2 |
| | max | 60.2 | 93.6 | 100.0 | 100.0 | 100.0 | 100.0 |
| Pb | min | 0.05 | 0.1 | 8.9 | 17.4 | 16.9 | 0.4 |
| | median | 0.6 | 19.0 | 76.2 | 83.5 | 76.7 | 9.1 |
| | mean | 1.4 | 19.2 | 73.9 | 80.1 | 75.1 | 12.6 |
| | max | 15.1 | 79.1 | 100.0 | 100.0 | 100.0 | 68.4 |

AA: acetic acid; CA: citric acid; EDTA: ethylenediaminetetraacetic acid; HCl: hydrochloric acid; UBM: unified bioaccessibility method in the gastric (G) and gastrointestinal (GI) phases.

More specifically, according to the bioaccessibility of metal(loid)s measured by UBM, the results showed that: (i) the bioaccessibility of As, Cd, and Pb in the G phase was 8.0−86.0% (with a mean of 26.0% and a median of 21.9%), 21.7−100% (with a mean of 80.5% and a median of 82.7%), and 16.9−100% (with a mean of 75.1% and a median of 76.7%), respectively; and (ii) in the GI phase, the bioaccessibility of these elements was 6.0−64.8%

(with a mean of 23.2% and a median of 20.1%), 6.2−100% (with a mean of 37.2% and a median of 35.7%), and 0.4−68.4% (with a mean of 12.6% and a median of 9.1%), respectively (Table 3). Notable differences were found in the behaviors of the metallic elements, which highlights that As was, on average, less bioaccessible in the G phase compared with Cd and Pb. It is well recognized that only a fraction of total As in soil is bioaccessible and that (iron) Fe oxyhydroxides were identified as key factors influencing the bioaccessibility of As (Wragg et al., 2007; Girouard and Zagury, 2009; Appleton et al., 2012; Kim et al., 2014). This influence was observed for different in vitro bioaccessibility assays. The high bioaccessibility of Cd and Pb observed in the G phase was related to the acidic pH of this environment (pH = 1.2), which induced an acidification of the organic salt functions and protonation of the Fe, manganese (Mn), and aluminum (Al) oxide/oxyhydroxide surfaces (Waterlot et al., 2017). During the GI phase, the bioaccessible fractions of Cd and Pb decreased considerably. In near-neutral pH and in the carbonate-rich environment of the intestine phase, these metals may be stabilized in solution by processes of adsorption or adsorption in remaining soil particles or other indigestible materials present in chime and/or by precipitation as relatively insoluble compounds (Davis et al., 1997; Basta and Gradwohl, 2000; Grøn and Andersen, 2003), allowing them to be less bioaccessible.

In most cases, As bioaccessibility in the G phase was slightly higher than that in the GI phase, which was also observed by Juhasz et al. (2009). Several studies have reported the importance of Fe oxides in controlling As sorption in soils (Smith et al., 1998; Smedley and Kinniburgh, 2002; Juhasz et al., 2007). Some of the As associated with Fe oxides may be dissolved in the G phase. Because the soluble Fe may be precipitated as amorphous oxides when pH increases from the gastric to the intestinal compartments (Mercer and Tobiason, 2008), dissolved As can be sorbed to these amorphous Fe oxides by surface complexation or ligand exchange to surface hydroxyl functional groups, thereby reducing the concentrations of dissolved As in the GI phase

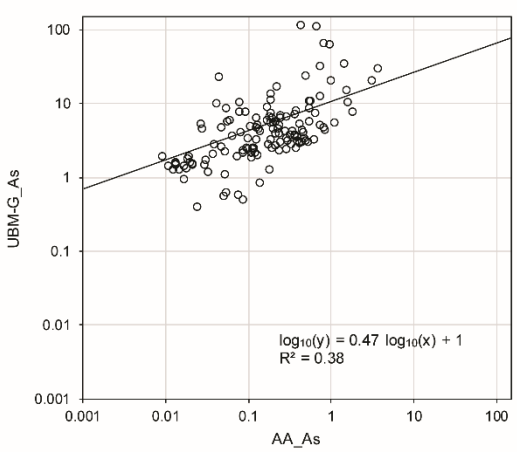
(Juhasz et al., 2009). For some soil samples (22%), bioaccessible As increased slightly from the G to the GI solutions, reflecting its desorption from solid phases at higher pH in the GI fluid than that in the G fluid (Beak et al., 2006). Moreover, Wenzel et al. (2001) clearly demonstrated the role of phosphates in the desorption of As from solid phases in digestive fluids.

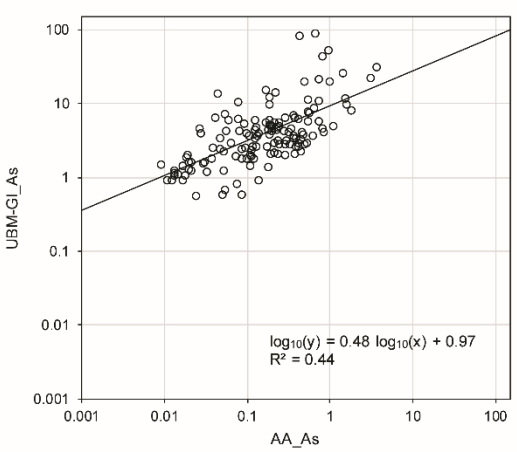
*3.3. Predicting As, Cd, and Pb bioaccessibilities*

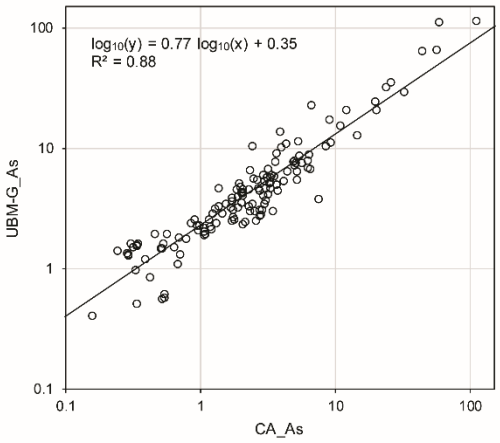
The training set of 140 soil samples was used to construct linear regression models and to select the best chemical extraction method(s) to estimate the bioaccessibility of As, Cd, and Pb in both the G and GI phases. To do this, both simple and multiple (PLS) linear regressions were conducted. Figures 2, 3, and 4 present the simple relationships between extractable concentrations obtained with UBM in both phases and those obtained with AA, CA, EDTA or HCl for As, Cd, and Pb, respectively.

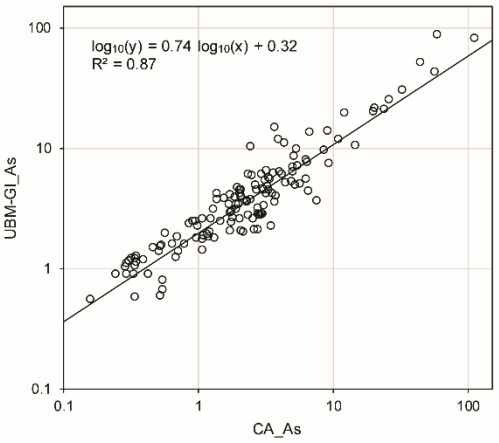
For As (Figure 2), the best relationships were established using HCl in both phases with a determination coefficient of $R^2$ = 0.91. To a lesser extent, the results showed also important correlations between As extracted in the G and GI phases and using CA ($R^2$ = 0.88 and 0.87, respectively) and EDTA ($R^2$ = 0.81 and 0.77, respectively). In all cases, according to the intercepts of the equations (from 0.13 to 0.45), either a more important or a less important underestimation of As bioaccessible values was observed by using HCl, EDTA, and CA.

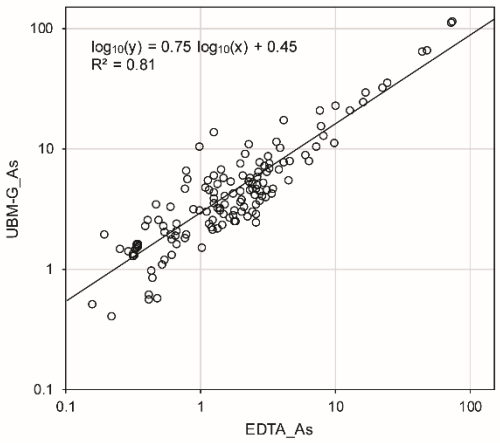
For Cd (Figure 3), the concentrations measured with all single-extraction methods generally correlated with bioaccessible Cd in the G phase, with $R^2$ = 0.81−0.99, and that in the GI phase, with $R^2$ = 0.78−0.97. However, for this element, specific behaviors were highlighted. With CA, two groups of soil samples appeared, where group 1 ($n$ = 120; indicated by black circles in Figure 3) corresponds to soils with low or moderate carbonate contents (0.3–90 g $kg^{-1}$, mean of 20 g $kg^{-1}$) while the other group ($n$ = 20; black dots) corresponds to soils that are highly carbonated (91–333 g $kg^{-1}$).

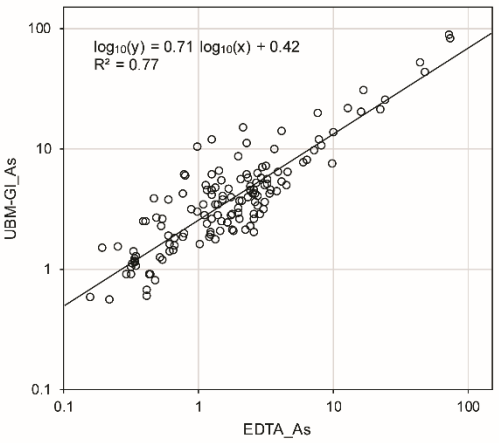

UBM-G_As
AA_As
log10(y) = 0.47 log10(x) + 1
R² = 0.38

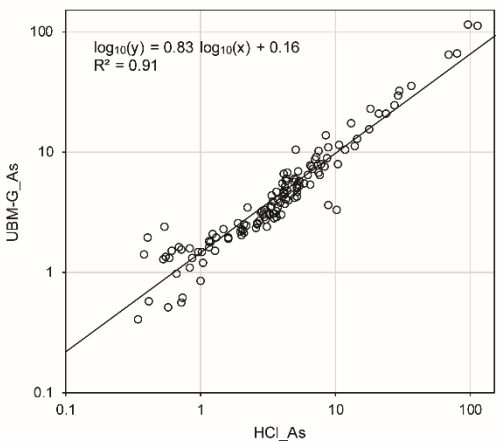

UBM-GI_As
AA_As
log10(y) = 0.48 log10(x) + 0.97
R² = 0.44


UBM-G_As
CA_As
log10(y) = 0.77 log10(x) + 0.35
R² = 0.88


UBM-GI_As
CA_As
log10(y) = 0.74 log10(x) + 0.32
R² = 0.87


UBM-G_As
EDTA_As
log10(y) = 0.75 log10(x) + 0.45
R² = 0.81


UBM-GI_As
EDTA_As
log10(y) = 0.71 log10(x) + 0.42
R² = 0.77


UBM-G_As
HCl_As
log10(y) = 0.83 log10(x) + 0.16
R² = 0.91

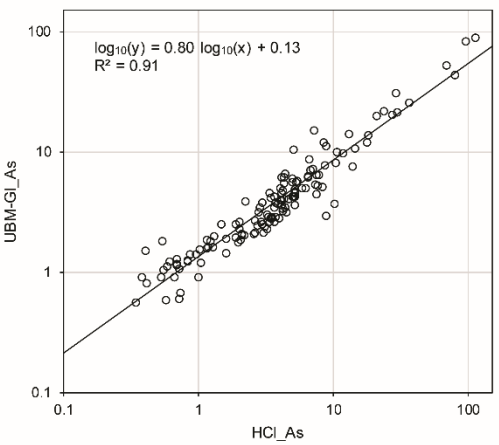
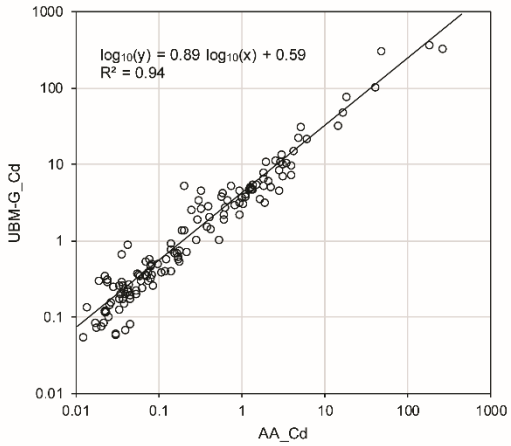
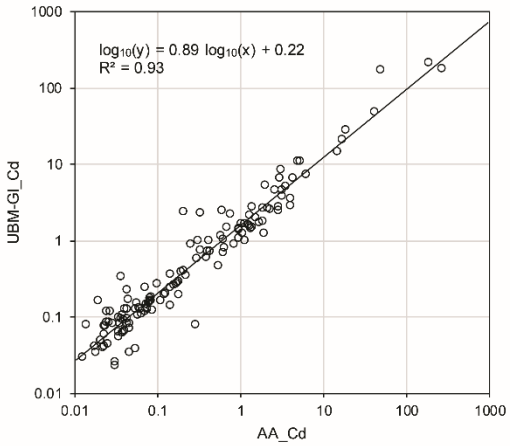
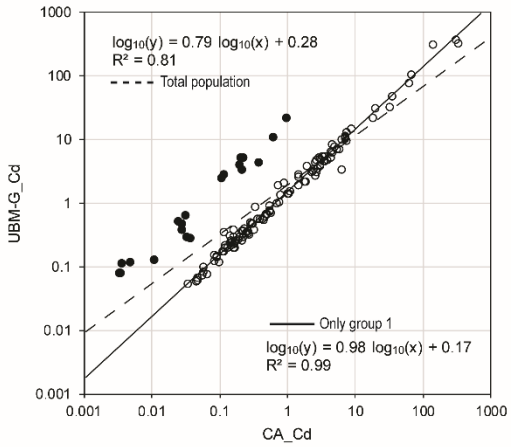
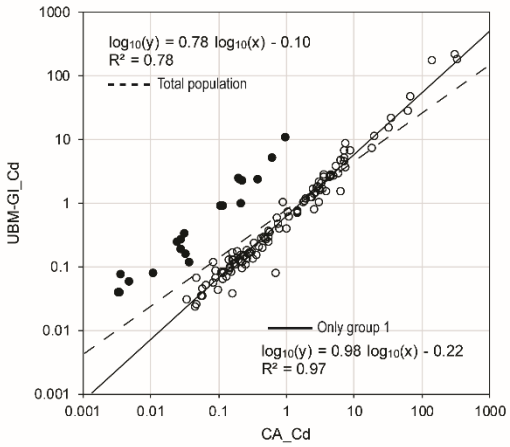

UBM-GI_As
HCl_As
log10(y) = 0.80 log10(x) + 0.13
R² = 0.91

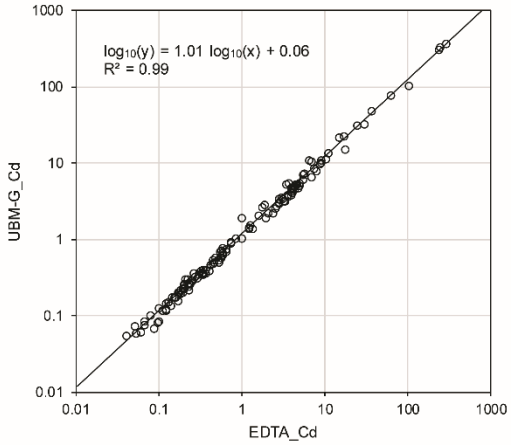
**Figure 2.** Relationships between extractable concentrations of As by UBM in gastric (UBM-G) and gastrointestinal (UBM-GI) phases and those by acetic acid (AA), citric acid (CA), EDTA, and HCl ($n$ = 140; training set). All extracted concentrations are expressed in mg $kg^{-1}$.

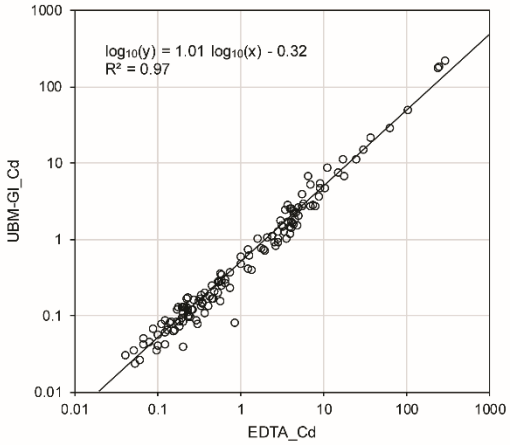

log10(y) = 0.89 log10(x) + 0.59
R² = 0.94
UBM-G_Cd
AA_Cd

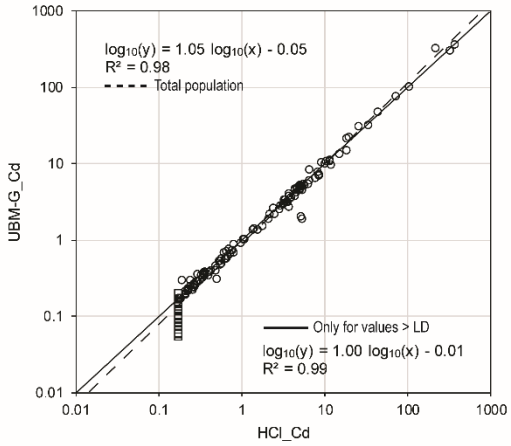

log10(y) = 0.89 log10(x) + 0.22
R² = 0.93
UBM-GI_Cd
AA_Cd


log10(y) = 0.79 log10(x) + 0.28
R² = 0.81
Total population
Only group 1
log10(y) = 0.98 log10(x) + 0.17
R² = 0.99
UBM-G_Cd
CA_Cd


log10(y) = 0.78 log10(x) - 0.10
R² = 0.78
Total population
Only group 1
log10(y) = 0.98 log10(x) - 0.22
R² = 0.97
UBM-GI_Cd
CA_Cd


log10(y) = 1.01 log10(x) + 0.06
R² = 0.99
UBM-G_Cd
EDTA_Cd


log10(y) = 1.01 log10(x) - 0.32
R² = 0.97
UBM-GI_Cd
EDTA_Cd


log10(y) = 1.05 log10(x) - 0.05
R² = 0.98
Total population
Only for values > LD
log10(y) = 1.00 log10(x) - 0.01
R² = 0.99
UBM-G_Cd
HCl_Cd

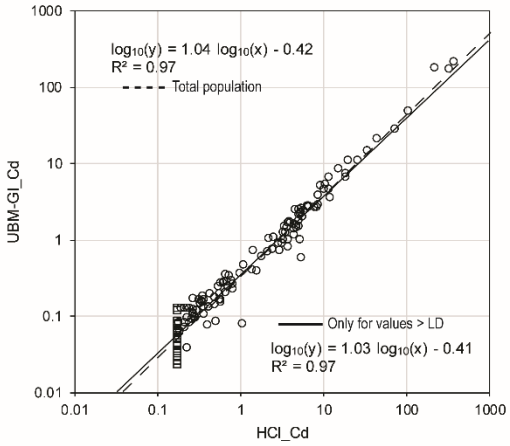
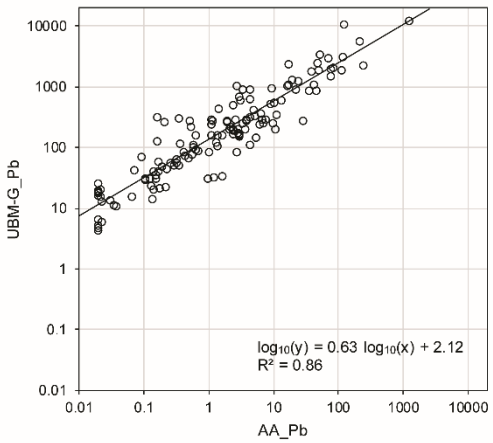
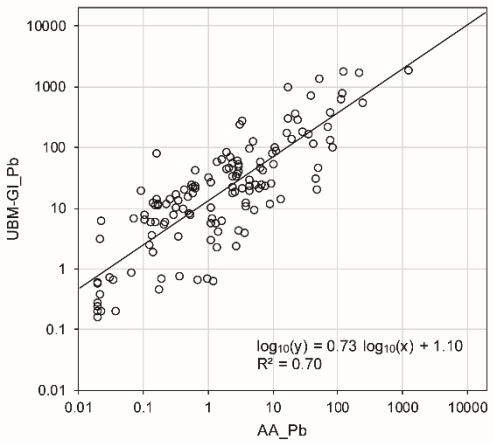
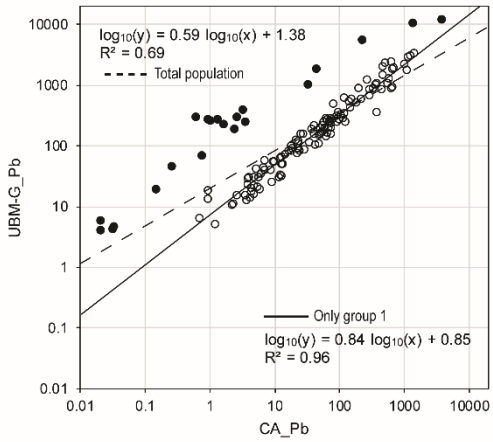
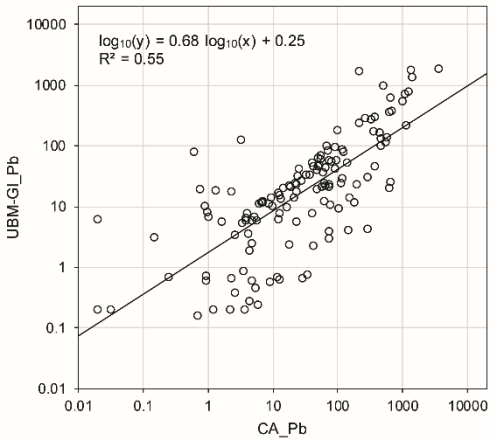

log10(y) = 1.04 log10(x) - 0.42
R² = 0.97
Total population
Only for values > LD
log10(y) = 1.03 log10(x) - 0.41
R² = 0.97
UBM-GI_Cd
HCl_Cd

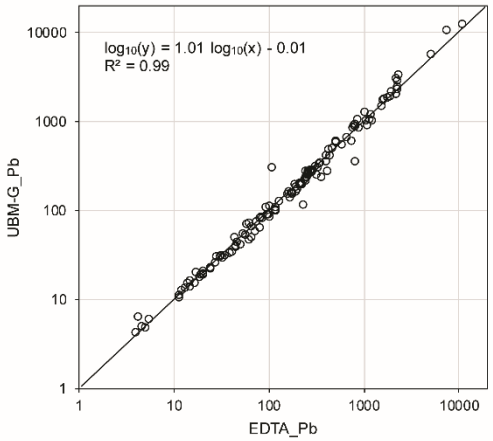
**Figure 3.** Relationships between extractable concentrations of Cd by UBM in gastric (UBM-G) and gastrointestinal (UBM-GI) phases and those by acetic acid (AA), citric acid (CA), EDTA, and HCl (*n* = 140; training set). All extracted concentrations are expressed in mg $kg^{-1}$. Group 1 (i.e., black circles) corresponds to soil samples with low or moderate carbonate content; black dots correspond to highly carbonated soils; LD: limit of detection.

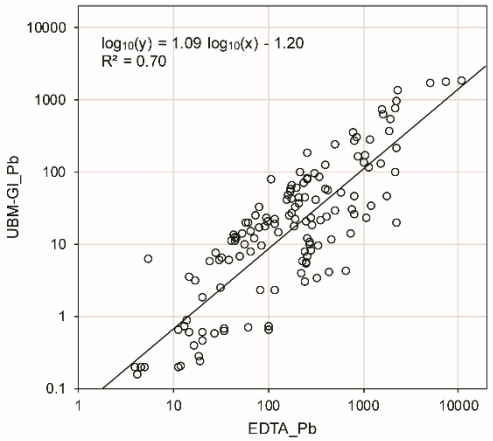

UBM-G_Pb
AA_Pb
log10(y) = 0.63 log10(x) + 2.12
R² = 0.86

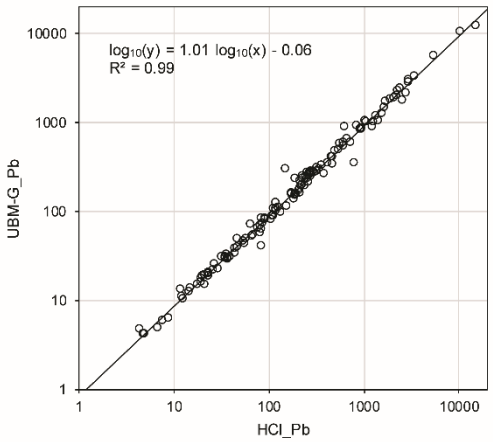

UBM-GI_Pb
AA_Pb
log10(y) = 0.73 log10(x) + 1.10
R² = 0.70


UBM-G_Pb
CA_Pb
log10(y) = 0.59 log10(x) + 1.38
R² = 0.69
Total population
Only group 1
log10(y) = 0.84 log10(x) + 0.85
R² = 0.96


UBM-GI_Pb
CA_Pb
log10(y) = 0.68 log10(x) + 0.25
R² = 0.55


UBM-G_Pb
EDTA_Pb
log10(y) = 1.01 log10(x) - 0.01
R² = 0.99


UBM-GI_Pb
EDTA_Pb
log10(y) = 1.09 log10(x) - 1.20
R² = 0.70


UBM-G_Pb
HCl_Pb
log10(y) = 1.01 log10(x) - 0.06
R² = 0.99

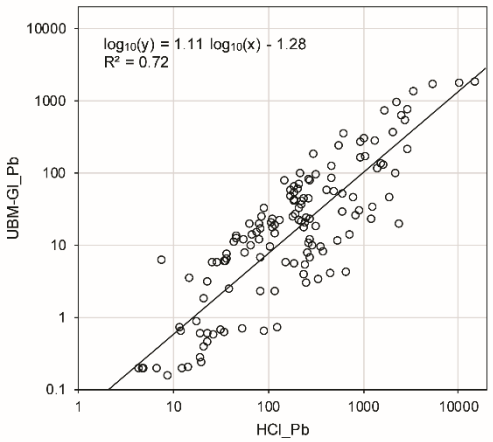

UBM-GI_Pb
HCl_Pb
log10(y) = 1.11 log10(x) - 1.28
R² = 0.72

**Figure 4.** Relationships between extractable concentrations of Pb by UBM in gastric (UBM-G) and gastrointestinal (UBM-GI) phases and those by acetic acid (AA), citric acid (CA), EDTA, and HCl ($n$ = 140; training set). All extracted concentrations are expressed in mg $kg^{-1}$. Group 1 (i.e., black circles) corresponds to soil samples with low or moderate carbonate content; black dots correspond to highly carbonated soils.

By considering only group 1 instead of the total soil population, significant relationships were observed in both the G and GI phases ($R^2$ = 0.99 and 0.97, respectively) and the regressions were improved (i.e., slopes close to 1 and very low intercepts of +0.17 and −0.22, respectively). With HCl, some soil samples ($n$ = 19) presented extractable concentrations less than the LD (< 0.2 µg $L^{-1}$). These soils were characterized by pseudo-total Cd concentrations below 0.20 mg $kg^{-1}$, which do not represent a risk for human exposure. By removing these values, significant relationships were obtained in both phases ($R^2$ = 0.99 and 0.97, respectively). In the G phase, the equation between UBM and HCl showed a slope of 1 and an intercept close to 0, while in the GI phase, an overestimation of bioaccessible values was observed (intercept of -0.41). With EDTA, important relationships were observed in both phases ($R^2$ = 0.99 and 0.97, respectively) within the range of soils studied, and the equations showed a slope of 1 (Figure 3).

For Pb (Figure 4), the best relationships were established using both EDTA and HCl but only in the G phase with a determination coefficient of $R^2$ = 0.99. For these two extractants, the equations obtained showed a slope of 1 and an intercept close to 0. With CA, the same behavior as Cd was observed but only in the G phase. Indeed, the same two soil groups appeared with a distinction between soils with low or moderate carbonate contents (group 1, $n$ = 120) and highly carbonated soils ($n$ = 20). As for Cd, when considering only group 1, important relationships were observed ($R^2$ = 0.96) and the regression was improved. However, whatever single-extraction method used, the simulation of Pb bioaccessibility in the GI phase was the least successful one.

In the present study, the four extractants were chosen for their potential ability to select the reactive pool of metal(loid)s either by competing with the metallic element for solid-phase binding sites (the case of dilute acids) or by competing with the binding sites to complex the element (the case of EDTA) (Groenenberg and Lofts, 2014). The solubility of metals that occur as cations is controlled by precipitation, co-precipitation, and sorption equilibria between the solution and the reactive surfaces in the soil solid phase (e.g., carbonate minerals, organic matter, amorphous metal oxides, and clay), all of these reactions being pH-dependent (Degryse et al., 2003; Poggio et al., 2009; Rodrigues et al., 2010b; Pelfrêne et al., 2015). With CA as single-extraction method, the chemistry of metallic elements, and more specifically Cd and Pb, was affected differently. Specific behavior was observed depending on the carbonate contents in the soil samples studied. In the case of slightly or moderately carbonated soils, all carbonate phases were dissolved by the acidic solutions (both CA and UBM in the G phase), which explains the significant relationships between the two methods. When considering only the highly carbonated soils, a significant relationship was also observed; however, the results highlighted an underestimation of Cd and Pb bioaccessible values with CA, which could be explained by the change of speciation of this extractant. Indeed, the affinity of CA toward cations depends on its speciation, which in turn depends on the pH of the mixture "soil particles and extracting solution," and on the ability of the acid to form complexes with metals (e.g., mono-, bi-, and polynuclear species as well as bi-, tri-, and multidentate complexes with metals, according to their ionic charge; Waterlot et al., 2017). To examine this hypothesis, the pH of the mixture was tested for all soil samples, and ranged between 8.0 and 10.3. At these pH values, CA is in its citrate ion form; thus, the speciation of the ion did not explain the underestimation of Cd and Pb bioaccessible concentrations. However, in the highly carbonated soils, the $Ca^{2+}$ ions could be in competition with $Cd^{2+}$ and $Pb^{2+}$ according to the complexation and precipitation reactions with citrate ions, thereby reducing the concentrations of dissolved metals such as Cd

and Pb. In both cases, the CA extraction method could be used to assess the bioaccessibility of Cd and Pb in the G phase; however, two different equations should be taken into account depending on the carbonate content in the soil.

In parallel, PLS regressions were conducted to determine whether combining the four extractants allows a better estimate to be obtained. Using the test set of 61 soil samples, both regressions were compared according to the RMSEP (Table 4) and the ratios (Figure S1 of the Supporting Information) between the measured bioaccessible concentrations by UBM and the predicted bioaccessible concentrations by either the best single extractants (EDTA or HCl) or a mix of extractants (PLS).

**Table 4.** Values of RMSEP (root mean square error of prediction) for model validation ($n$ = 61).

| Extractant | As-G | As-GI | Cd-G | Cd-GI | Pb-G | Pb-GI |
|---|---|---|---|---|---|---|
| AA | 0.36 | 0.31 | 0.27 | 0.29 | 0.25 | 0.46 |
| CA | 0.3 | 0.25 | 0.43 for all values<br>0.10 only for group 1 | 0.45 for all values<br>0.14 only for group 1 | 0.41 for all values<br>0.19 only for group 1 | 0.52 |
| EDTA | 0.17 | 0.16 | 0.16 | 0.18 | 0.089 | 0.39 |
| HCl | 0.12 | 0.11 | 0.14 for all values<br>0.13 only for values > LD | 0.19 for all values<br>0.19 only for values > LD | 0.079 | 0.37 |
| PLS | 0.17 | 0.13 | 0.13 | 0.14 | 0.075 | 0.36 |

AA: acetic acid; CA: citric acid; EDTA: ethylenediaminetetraacetic; HCl: hydrochloric acid; PLS: partial least square regression; As-, Cd-, and Pb-G: metal(loid) in the gastric phase; As-, Cd-, and Pb-GI: metal(loid) in the gastrointestinal phase. LD: limit of detection. Group 1: soils with low or moderate carbonate content.

For As, variations of RMSEP values were observed between the different extraction methods and the values decreased in the order AA > CA > EDTA > HCl (Table 4). The lowest RMSEP values were obtained for HCl to predict bioaccessibility in both the G and the GI phases. With PLS, the values were in the same order of magnitude as those calculated with EDTA or HCl alone and did not improve the prediction of As bioaccessibility (Table 4 and Figure S1 of the Supporting Information). For Cd, the RMSEP values decreased in the order CA > AA > EDTA

≥ HCl (Table 4). The lowest values were obtained for EDTA and HCl to mimic bioaccessibility in both phases. However, by considering only soil samples from group 1 (i.e., slightly or moderately carbonated), the best results were observed with CA, with RMSEP values of 0.10 and 0.14, respectively, in the G and GI phases. As previously shown for As, the PLS did not improve the prediction of Cd bioaccessibility (Table 4 and Figure S1 of the Supporting Information). For Pb, the RMSEP values decreased in the order CA > AA >> EDTA ≥ HCl when considering the G phase, while in the GI phase the values were high whatever the extractants used (0.37−0.52) (Table 4). In the G phase, the RMSEP values obtained with EDTA and HCl (0.089 and 0.079, respectively) were much smaller than those of the other extractants. The same behavior as for Cd was highlighted for Pb in the G phase, where the RMSEP was significantly decreased when taking into account only group 1 of the soil samples (from 0.41 to 0.19). With PLS, the values were in the same order of magnitude as those calculated with HCl alone and did not improve the prediction of Pb bioaccessibility (Table 4 and Figure S1 of the Supporting Information).

When considering all the soil samples studied, it is clear that the predicted bioaccessibility values of As, Cd, and Pb by CA (differentiating between soils with high and those with low or moderate carbonate content), EDTA, or HCl were in agreement with the values measured in both phases (to a lesser extent in the GI phase for Pb). Indeed, the RMSEP values were low, suggesting a satisfactory and effective calibration of the models. The results also highlighted that the RMSEP values were low with PLS. However, the differences between the two regression methods were not large enough to justify the use of four extractants to mimic the bioaccessibility of metal(loid)s.

In view of the previously obtained results and when considering all soil samples studied, the validation was carried out by taking into account the following as single-extraction methods: (i) HCl for As (Figure S2 of the Supporting Information), and (ii) CA, EDTA, and HCl for Cd

(Figure S3 of the Supporting Information) and Pb (Figure S4 of the Supporting Information) in both the G and the GI phases.

Table 5 summarizes the best regression models for As, Cd, and Pb representing bioaccessible concentrations in the G and GI phases as a function of extractable concentrations by HCl, EDTA, or CA and their conditions of use in the range of soils studied.

Regarding the range of the degree of contamination of the soils studied, the most relevant results for estimating the oral bioaccessibility of As were gained from the HCl extraction method. Given the human health hazard of this element, additional investigations were carried out to confirm this observation, by integrating 27 additional soils with significantly higher levels of As contamination. The results are presented in the Supporting Information (Figure S5).



**Table 5.** Best regression models for As, Cd, and Pb representing bioaccessible concentrations with UBM in gastric (G) and gastrointestinal (GI) phases as a function of extractable concentrations by HCl, EDTA, or citric acid (CA) ($n$ = 140, training set).

| Element | Phase | Extractant | Equation | $R^2$ | Recommendations of use |
|---|---|---|---|---|---|
| As | G | HCl | $\log_{10}[As]_{UBM} = 0.83 \times \log_{10}[As]_{HCl} + 0.16$ | 0.91 | In the range of soils studied |
| | GI | HCl | $\log_{10}[As]_{UBM} = 0.80 \times \log_{10}[As]_{HCl} + 0.13$ | 0.91 | In the range of soils studied |
| Cd | G | HCl | $\log_{10}[Cd]_{UBM} = 1.00 \times \log_{10}[Cd]_{HCl} - 0.01$ | 0.99 | In soils with $[Cd]_{total} > 0.20$ mg kg$^{-1}$ |
| | | EDTA | $\log_{10}[Cd]_{UBM} = 1.01 \times \log_{10}[Cd]_{EDTA} + 0.06$ | 0.99 | In the range of soils studied |
| | | CA | $\log_{10}[Cd]_{UBM} = 0.98 \times \log_{10}[Cd]_{CA} + 0.17$ | 0.99 | In soils with low or moderate carbonate content |
| | | | $\log_{10}[Cd]_{UBM} = 0.98 \times \log_{10}[Cd]_{CA} + 1.26$ | 0.96 | In soils highly carbonated (> 10%) |
| | GI | HCl | $\log_{10}[Cd]_{UBM} = 1.03 \times \log_{10}[Cd]_{HCl} - 0.41$ | 0.97 | In soils with $[Cd]_{total} > 0.20$ mg kg$^{-1}$ |
| | | EDTA | $\log_{10}[Cd]_{UBM} = 1.01 \times \log_{10}[Cd]_{EDTA} - 0.32$ | 0.97 | In the range of soils studied |
| | | CA | $\log_{10}[Cd]_{UBM} = 0.98 \times \log_{10}[Cd]_{CA} - 0.22$ | 0.97 | In soils with low or moderate carbonate content |
| | | | $\log_{10}[Cd]_{UBM} = 0.94 \times \log_{10}[Cd]_{CA} + 0.89$ | 0.94 | In soils highly carbonated (> 10%) |
| Pb | G | HCl | $\log_{10}[Pb]_{UBM} = 1.01 \times \log_{10}[Pb]_{HCl} - 0.06$ | 0.99 | In the range of soils studied |
| | | EDTA | $\log_{10}[Pb]_{UBM} = 1.01 \times \log_{10}[Pb]_{EDTA} - 0.01$ | 0.99 | In the range of soils studied |
| | | CA | $\log_{10}[Pb]_{UBM} = 0.84 \times \log_{10}[Pb]_{CA} + 0.85$ | 0.96 | In soils with low or moderate carbonate content |
| | | | $\log_{10}[Pb]_{UBM} = 0.69 \times \log_{10}[Pb]_{CA} + 2.09$ | 0.93 | In soils highly carbonated (> 10%) |
| | GI | HCl | $\log_{10}[Pb]_{UBM} = 1.11 \times \log_{10}[Pb]_{HCl} - 1.28$ | 0.72 | In the range of soils studied |
| | | EDTA | $\log_{10}[Pb]_{UBM} = 1.09 \times \log_{10}[Pb]_{EDTA} - 1.20$ | 0.7 | In the range of soils studied |

## 4. Conclusion

An investigation was undertaken to assess the potential suitability of chemical single-extraction methods using AA (0.11 M), CA (0.11 M), EDTA (0.16 M), and HCl (0.65%) to mimic the

bioaccessible fraction of As, Cd, and Pb in 201 soil samples with a wide range of physicochemical parameters and metal(loid)s concentrations (1.9– 228 mg $kg^{-1}$ for As, 0.1–483 mg $kg^{-1}$ for Cd, and 9–12,300 mg $kg^{-1}$ for Pb). A statistical modeling approach was followed to establish a robust model by using: (i) a training set of 140 samples to construct regressions models and (ii) an independent test set of 61 samples to validate the regression models.

We showed in 201 various soil samples that the oral bioaccessibility of As, Cd, and Pb can be predicted by HCl, EDTA, or CA. These three cheaper single-extraction methods are ease of use for analytical laboratories, fast and reproducible. Because the method with HCl is closer to physiological conditions (i.e., solid/liquid ratio, reagent, temperature, pH and residence time in the stomach) compared to others, it can be used to assess the bioaccessibility of As, Cd, and Pb at least in first-tier screening. More specifically, from the extractable concentrations of metal(loid)s by HCl, the predicted bioccessibility may be calculated by taking the equations from simple linear regressions. In a second-tier study and a complementary validation approach, it is recommended to use the UBM protocol in a few soil samples for better assessment of human exposure. The UBM bioaccessibility test has only been validated against in vivo data for As, Cd and Pb. The UBM and the simplified UBM (by using HCl) protocols could be used for other elements; however, the data obtained can only be considered as a useful line of evidence in interpreting results.



**Acknowledgments**

The authors wish to thank ADEME (French Agency for the Environment and Energy Management) for the financial support of this research (contract n°1472C0046) and the scientific partners for providing the soil samples. We think also Séverine Durand, Delphine Pelle and Julien Sinquin from EHESP for sample analyses and their contributions to this work.

## References

Alexander, M., 2000. Aging, bioavailability, and overestimation of risk from environmental pollutants. Environ. Sci. Technol. 34, 4259−4265.

Appleton, J.D., Cave, M.R., Wragg, J., 2012. Anthropogenic and geogenic impacts on arsenic bioaccessibility in UK topsoils. Sci. Total Environ. 435−436, 21−29.

Barsby, A., McKinley, J.M., Ofterdinger, U., Young, M., Cave, M.R., Wragg, J., 2012. Bioaccessibility of trace elements in soils in Northern Ireland. Sci. Total Environ. 433, 398−417.

Basta, N., Gradwohl, R., 2000. Estimation of Cd, Pb, and Zn bioavailability in smelter-contaminated soils by a sequential extraction procedure. J. Soil Contam. 9, 149−164.

Beak, D.G., Basta, N.T., Schekel, K.G., Traina, S.J., 2006. Bioaccessibility of arsenic (V) bound to ferrihydrite using a simulated gastrointestinal system. Environ. Sci. Technol. 40, 1364−1370.

Bonten, L.T.C., Groenenberg, J.E., Weng, L., Van Riemsdijk, W.H., 2008. Use of speciation and complexation models to estimate heavy metal sorption in soils. Geoderma 146, 303−310.

Brattin, W., Drexler, J., Lowney, Y., Griffin, S., Diamond, G., Woodbury, L., 2013. An in vitro method for estimation of arsenic relative bioavailability in soil. J. Toxicol. Environ. Health Part A 76, 458−478.

Cruz, N., Rodriguez, S.M., Tavares, D., Monteiro, R.J.R., Carvahlo, L., Trindade, T., Duarte, A.C., Pereira, E., Römkens, P.F.A.M., 2015. Testing single extraction methods and in vitro tests to assess the geochemical reactivity and human bioaccessibility of silver in urban soils amended with silver nanoparticles. Chemosphere 135, 304−311.

Davis, A., Ruby, M.V., Goad, P., Eberle, S., Chryssoulis, S., 1997. Mass balance on surface-bound mineralogic, and total lead concentrations as related to industrial aggregate bioaccessibility. Environ. Sci. Technol. 31, 37−44.

Degryse, F., Broos, K., Smolders, E., Merckx, R., 2003. Soil solution concentration of Cd and Zn can be predicted with a $CaCl_2$ extract. Eur. J. Soil Sci. 54, 149−157.

Denys, S., Caboche, J., Tack, K., Rychen, G., Wragg, J., Cave, M.R., Jondreville, C., Feidt, C., 2012. In vivo validation of the Unified BARGE Method to assess the bioaccessibility of arsenic, antimony, cadmium, and lead in soils. Environ. Sci. Technol. 46, 6252−6260.

Denys, S., Caboche, J., Tack, K., Delalain, P., 2007. Bioaccessibility of lead in high carbonate soils. J. Environ. Sci. Health A 42, 1331−1339.

Drexler, J.W., Brattin, W.J., 2007. An in vitro procedure for estimation of lead relative bioavailability: with validation. Hum. Ecol. Risk Assess. 13, 383−401.

Girouard, E., Zagury, G.J., 2009. Arsenic bioaccessibility in CCA-contaminated soils: Influence of soil properties, arsenic fractionation, and particle-size fraction. Sci. Total Environ. 407, 2576−2585.

Groenenberg, J.E., Lofts, S., 2014. The use of assemblage models to describe trace element partitioning, speciation, and fate: a review. Environ. Toxicol. Chem. 33, 2181−2196.

Groenenberg, J.E., Römkens, P.F.A.M., Comans, R.N.J., Luster, J., Pampura, T., Shotbolt, L., Tipping, E., De Vries, W., 2010. Transfer functions for solid solution partitioning of cadmium, copper, nickel, lead and zinc in soils: derivation of relations for free metal ion activities and validation on independent data. Eur. J. Soil Sci. 61, 58−73.

Grøn, C., Andersen, L., 2003. Human bioaccessibility of heavy metals and PAH from soil; Environmental project No. 840, Technology Programme for Soil and Groundwater Contamination. Danish Environmental Protection Agency.

ISO/DIS 17924:2016. Soil quality – Assessment of human exposure from ingestion of soil and soil material – Procedure for the estimation of the human bioaccessibility/bioavailability of metals in soil.

Jeong, S., Moon, H.S., Nam, K., 2013. Differential in vitro bioaccessibility of residual As in a field-aged former smelter site and its implication for potential risk. Sci. Total Environ. 463−464, 348−354.

Juhasz, A.L., Weber, J., Smith, E., Naidu, R., Rees, M., Rofe, A., Kuchel, T., Sansom, L., 2009. Assessment of four commonly employed in vitro arsenic bioaccessibility assays for predicting in vivo relative arsenic bioavailability in contaminated soils. Environ. Sci. Technol. 43, 9487−9494.

Juhasz, A.L., Smith, E., Weber, J., Rees, M., Rofe, A., Kuchel, T., Sansom, L., Naidu, R., 2007. In vitro assessment of arsenic bioaccessibility in contaminated (anthropogenic and geogenic) soils. Chemosphere 69, 69−78.

Kelley, M.E., Brauning, S., Schoof, R., Ruby, M.V., 2002. Assessing oral bioavailability of metals in soils. Battelle Press, Columbus, OH, USA.

Kim, E.J., Yoo, J.C., Baek, K., 2014. Arsenic speciation and bioaccessibility in arsenic-contaminated soils: sequential extraction and mineralogical investigation. Environ. Pollut. 186, 29−35.

Latawiec, A.E., Simmons, P., Reid, B.J., 2010. Decision-makers' perspectives on the use of bioaccessibility for risk-based regulation of contaminated land. Environ. Int. 36, 383−389.

Le Bot, B., Gilles, E., Durand, S., Glorennec, P., 2010. Bioaccessible and quasi-total metals in soil and indoor dust. Eur. J. Mineral. 22, 651−657.

Le Bot, B., Arcelin, C., Briand, E., Glorennec, P., 2011. Sequential digestion for measuring leachable and total lead in the same sample of dust or paint chips by ICP-MS. J. Environ. Sci. Health A 46, 63−69.

Li, S.W., Li, J., Li, H.B., Naidu, R., Ma, L.Q., 2015. Arsenic bioaccessibility in contaminated soils: coupling in vitro assays with sequential and $HNO_3$ extraction. J. Hazard. Mater. 295, 145−152.

Li, Y., Zhang, M.K., 2013. A comparison of physiologically based extraction test (PBET) and single-extraction methods for release of Cu, Zn, and Pb from mildly acidic and alkali soils. Environ. Sci. Pollut. Res. 20, 3140−3148.

Madrid, F., Biasioli, M., Ajmone-Marsan, F., 2008. Availability and bioaccessibility of metals in fine particles of some urban soils. Arch. Environ. Con. Tox. 55, 21−32.

Mendoza, C.J., Garrido, R.T., Quilodrán, R.C., Segovia, C.M., Parada, A.J., 2017. Evaluation of the bioaccessible gastric and intestinal fractions of heavy metals in contaminated soils by means of a simple bioaccessibility extraction test. Chemosphere 176, 81−88.

Mercer, K.L., Tobiason, J.E., 2008. Removal of arsenic from high ionic strength solutions: effects of ionic strength, pH, and preformed versus in situ formed HFO. Environ. Sci. Technol. 42, 3797−3802.

Mevik, B.H., Wehrens, R., Liland, K.H., 2019. pls: Partial Least Squares and Principal Component Regression. R package version 2.7-1. https://CRAN.R-project.org/package=pls

Ollson, C.A., Koch, I., Smith, P., Knopper, L.D., Hough, C., Reimer, K.J., 2009. Addressing arsenic bioaccessibility in ecological risk assessment: a novel approach to avoid overestimating risk. Environ. Toxicol. Chem. 28, 668−675.

Oomen, A.G., Hack, A., Minekus, M., Zeijdner, E., Schoeters, G., Verstraete, W., Wiele, T.V.D., Wragg, J., Rompelberg, C.J.M., Sips, A.J.A.M., Wijnen, J.H.V., 2002. Comparison of five in vitro digestion models to study the bioaccessibility of soil contaminants. Environ. Sci. Technol. 36, 3326−3334.

Paustenbach, D.J., 2000. The practice of exposure assessment: a state of the art review. J. Toxicol. Environ. Health B Crit. Rev. 3, 179−291.

Pelfrêne, A., Détriché, S., Douay, F., 2015. Combining spatial distribution with oral bioaccessibility of metals in smelter-impacted soils: implications for human health risk assessment. Environ. Geochem. Hlth. 37, 49−62.

Pelfrêne, A., Waterlot, C., Mazzuca, M., Nisse, C., Bidar, G., Douay, F., 2011. Assessing Cd, Pb, Zn human bioaccessibility in smelter-contaminated agricultural topsoils (northern France). Environ. Geochem. Hlth. 33, 477−493.

Poggio, L., Vrščaj, B., Schulin, R., Hepperle, E., Marsan, F.A., 2009. Metals pollution and human bioaccessibility of topsoils in Grugliasco (Italy). Environ. Pollut. 157, 680−689.

Pouschat, P., Zagury, G.J., 2006. In vitro gastrointestinal bioavailability of arsenic in soils collected near CCA-treated utility poles. Environ. Sci. Technol. 40, 4317−4323.

Reis, A.P., Patinha, C., Wragg, J., Dias, A.C., Cave, M., Sousa, A.J., Costa, C., Cachada, A., Ferraira da Silva, E., Rocha, F., Duarte, A., 2014. Geochemistry, mineralogy, solid-phase fractionation and oral bioaccessibility of lead in urban soils of Lisbon. Environ. Geochem. Hlth. 36, 867−881.

Rodrigues, S.M., Cruz, N., Carvalho, L., Duarte, A.C., Pereira, E., Boim, A.G.F., Alleoni, L.R.F., Römkens, P.F.A.M., 2018. Evaluation of a single extraction test to estimate the human oral bioaccessibility of potentially toxic elements in soils: towards more robust risk assessment. Sci. Total Environ. 635, 188−202.

Rodrigues, S.M., Cruz, N., Coelho, C., Henriques, B., Carvalho, L., Duarte, A., Pereira, M.E., Römkens, P.F.A.M., 2013. Risk assessment for Cd, Cu, Pb and Zn in urban soils: chemical availability as the central concept. Environ. Pollut. 183, 234−242.

Rodrigues, S.M., Henriques, B., Ferreira da Silva, E., Pereira, M.E., Duarte, A.C., Römkens, P.F.A.M., 2010a. Evaluation of an approach for the characterization of reactive and available pools of twenty potentially toxic elements in soils: part I – the role of key soil properties in the variation of contaminants reactivity. Chemosphere 81, 1549−1559.

Rodrigues, S.M., Henriques, B., Ferreira da Silva, E., Pereira, E., Duarte, A.C., Groenenberg, J.E., Römkens, P.F.A.M., 2010b. Evaluation of an approach for the characterization of reactive and available pools of twenty potentially toxic elements in soils: part II – solid-solution partition relationships and ion activity in soil solutions. Chemosphere 81, 1560−1570.

Rodriguez, R.R., Basta, N.T., 1999. An in vitro gastrointestinal method to estimate bioavailable arsenic in contaminated soils and solid media. Environ. Sci. Technol. 33, 642−649.

Römkens, P.F.A.M., Guo, H.Y., Chu, C.L., Liu, T.S., Chiang, C.F., Koopmans, G.F., 2009. Characterization of soil heavy metal pools in paddy fields in Taiwan: chemical extraction and solid-solution partitioning. J. Soils Sediments 9, 216−228.

Ruby, M.V., Schoof, R., Brattin, W., Goldade, M., Post, G., Harnois, M., Mosby, D.E., Casteel, S.W., Berti, W., Carpentier, M., Edwards, D., Cragin, D., Chappell, W., 1999. Advances in evaluating the oral bioavailability of inorganics in soil for use in human health risk assessment. Environ. Sci. Technol. 33, 3697−3705.

Ruby, M.V., Davis, A., Schoof, R., Eberle, S.S.C.M., 1996. Estimation of lead and arsenic bioavailability using a physiologically based extraction test. Environ. Sci. technol. 30, 422−430.

Ruby, M.V., Davis, A., Link, T.E., Schoof, R., Chaney, R.L., Freeman, G.B., Bergstrom, P., 1993. Development of an in vitro screening-test to evaluate the in vivo bioaccessibility of ingested mine-waste lead. Environ. Sci. Technol. 26, 1242−1248.

Smedley, P.L., Kinniburgh, D.G., 2002. A review of the source, behaviour and distribution of arsenic in natural waters. Appl. Geochem. 17, 517−568.

Smith, E., Naidu, R., Alston, A.M., 1998. Chemistry of arsenic in soils: I. Sorption of arsenate and arsenite by four Australian soils. J. Environ. Qual. 28, 1719−1726.

Tang, W., Xia, Q., Shan, B., Ng, J.C., 2018. Relationship of bioaccessibility and fractionation of cadmium in long-term spiked soils for health risk assessment based in four in vitro gastrointestinal simulation models. Sci. Total Environ. 631−632, 1582−1589.

Tipping, E., Rieuwerts, J., Pan, G., Ashmore, M.R., Lofts, S., Hills, M.T.R., Farago, M.E., Thornton, I., 2003. The solid-solution partitioning of heavy metals (Cu, Zn, Cd, Pb) in upland soils of England and Wales. Environ. Pollut. 125, 213−225.

USEPA, 2007. Estimation of relative bioavailability of lead in soil and soil-like materials using in vivo and in vitro methods. OSWER 9285.7-80.

USEPA, 2017. Validation assessment of in vitro arsenic bioaccessibility assay for predicting relative bioavailability of arsenic in soils and soil-like materials at superfund sites. OLEM 9355.4-29.

Waterlot, C., Bidar, G., Pruvot, C., Douay, F., 2012. Effects of grinding and shaking on Cd, Pb and Zn distribution in anthropogenically impacted soils. Talanta 98, 185−196.

Waterlot, C., Douay, F., Pelfrêne, A., 2017. Chemical availability of Cd, Pb and Zn in anthropogenically polluted soil: assessing the geochemical reactivity and oral bioaccessibility. Pedosphere 27, 616−629.

Wenzel, W.W., Kirchbaumer, N., Prohaska, T., Stingeder, G., Lombi, E., Adriano, D.C., 2001. Arsenic fractionation in soils using an improved sequential extraction procedure. Anal. Chim. Acta 436, 309−323.

Wragg, J., Cave, M., Nathanail, P., 2007. A study of the relationship between arsenic bioaccessibility and its solid-phase distribution in soils from Wellingborough, UK. J. Environ. Sci. Health A 42, 1303−1315.

Xia, Q., Lamb, D., Peng, C., Ng, J.C., 2017. Interaction effects of As, Cd and Pb on their respective bioaccessibility with time in co-contaminated soils assessed by the Unified BARGE method. Environ. Sci. Pollut. Res. 24, 5585−5594.

## Supporting Information

**Table S1.** Soil parameters and pseudototal concentrations of As, Cd, and Pb for both the training set and the test set of soil samples ($n$ = 140 and 61, respectively).

| Set of soil samples | No of samples | | Soil parameters | | | | | | Pseudototal concentrations | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Clay % | Silt % | Sand % | OM % | pH (in water) | Total $CaCO_3$ g $kg^{-1}$ | As mg $kg^{-1}$ | Cd mg $kg^{-1}$ | Pb mg $kg^{-1}$ |
| Training set | 140 | Min | 7 | 1 | 4.0 | 0.2 | 4.3 | 0.3 | 2.0 | 0.1 | 9.1 |
| | | Mean | 21 | 43 | 36 | 8.1 | 7.4 | 39 | 24.8 | 12.5 | 686 |
| | | Max | 47 | 73 | 84 | 66.9 | 9.1 | 333 | 199 | 483 | 12330 |
| Test set | 61 | Min | 7.0 | 19 | 4 | 0.1 | 4.7 | 1.0 | 1.9 | 0.1 | 9.6 |
| | | Mean | 22 | 46 | 31 | 9.0 | 7.5 | 73 | 25.2 | 9.0 | 1138 |
| | | Max | 53 | 72 | 74 | 60.7 | 8.3 | 734 | 228 | 257 | 9790 |

OM: organic matter

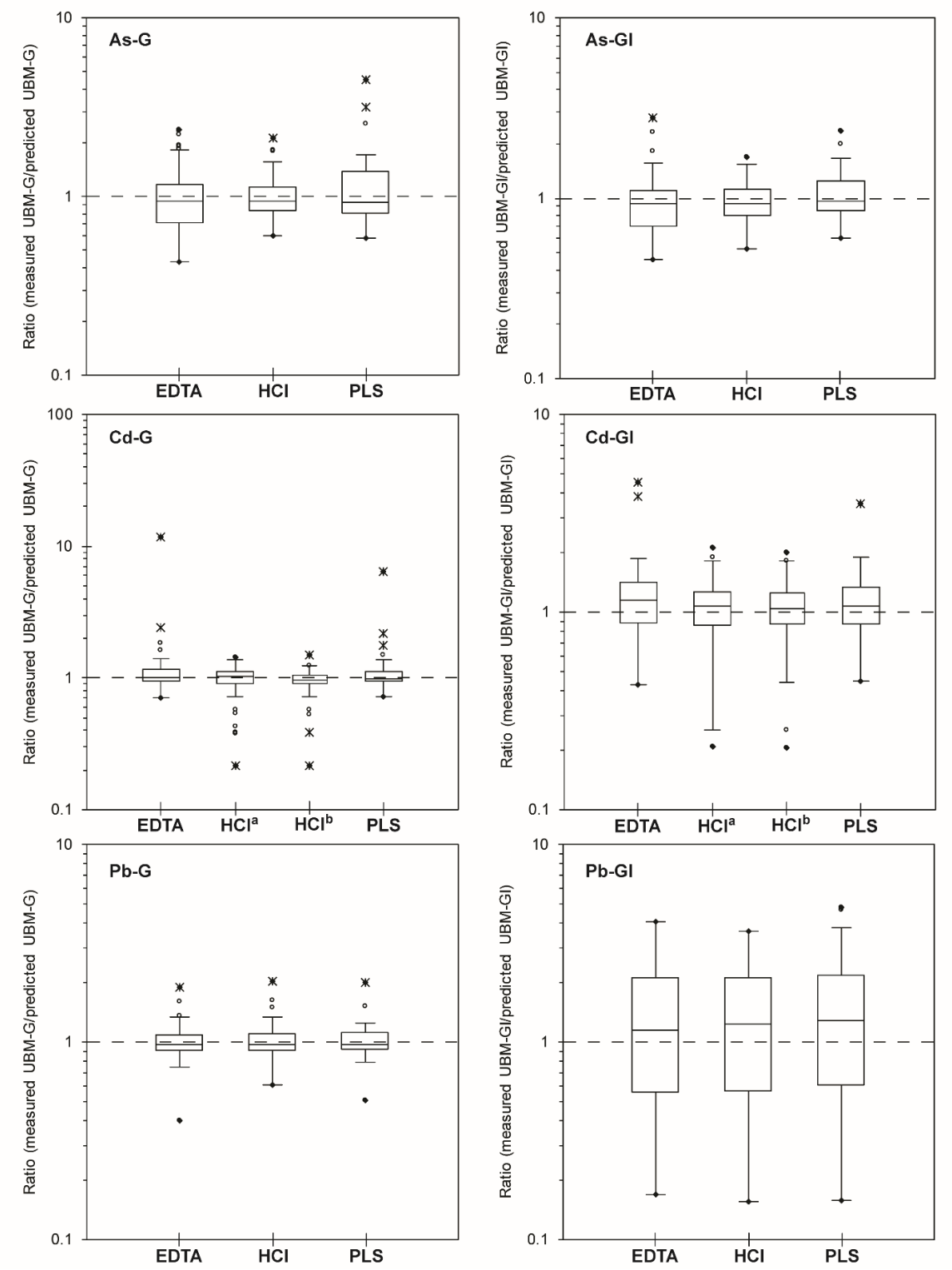



**Figure S1.** Boxplots of ratios ($n$ = 61, test set) between measured bioaccessible concentrations by UBM in gastric (G) and gastrointestinal (GI) phases and predicted bioaccessible concentrations by either single extractant (EDTA or HCl) or a mix of extractants (PLS; partial least square regression). $HCl^a$ and $HCl^b$ correspond to ratios calculated from all soil samples in the test set and from only soil samples with pseudototal concentrations of Cd greater than 0.20 mg $kg^{-1}$, respectively.

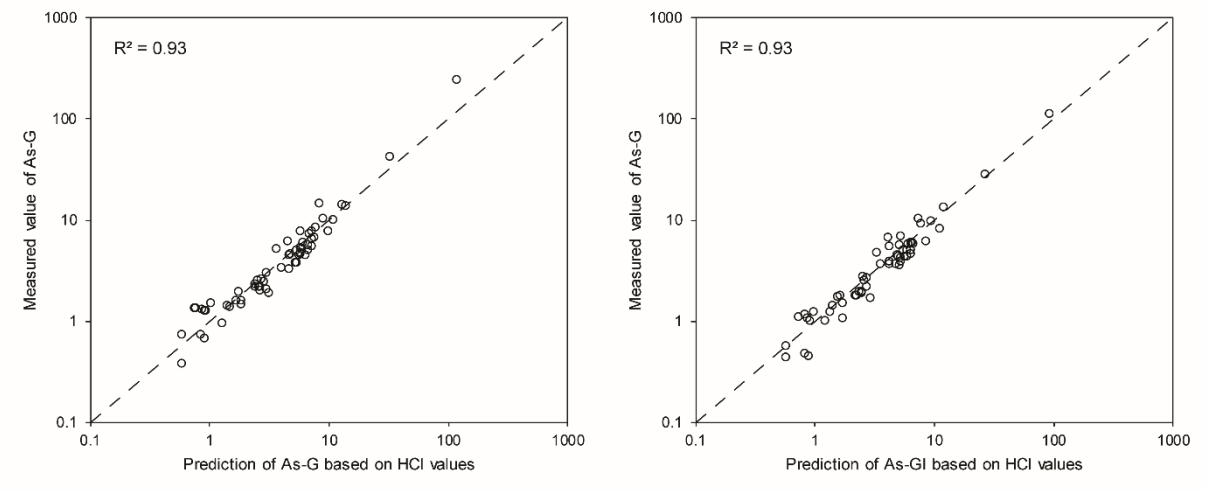



**Figure S2.** Bioaccessibility of As in gastric (As-G) and gastrointestinal (As-GI) phases measured by UBM versus predicted values by HCl for model validation ($n$ = 61, test set).

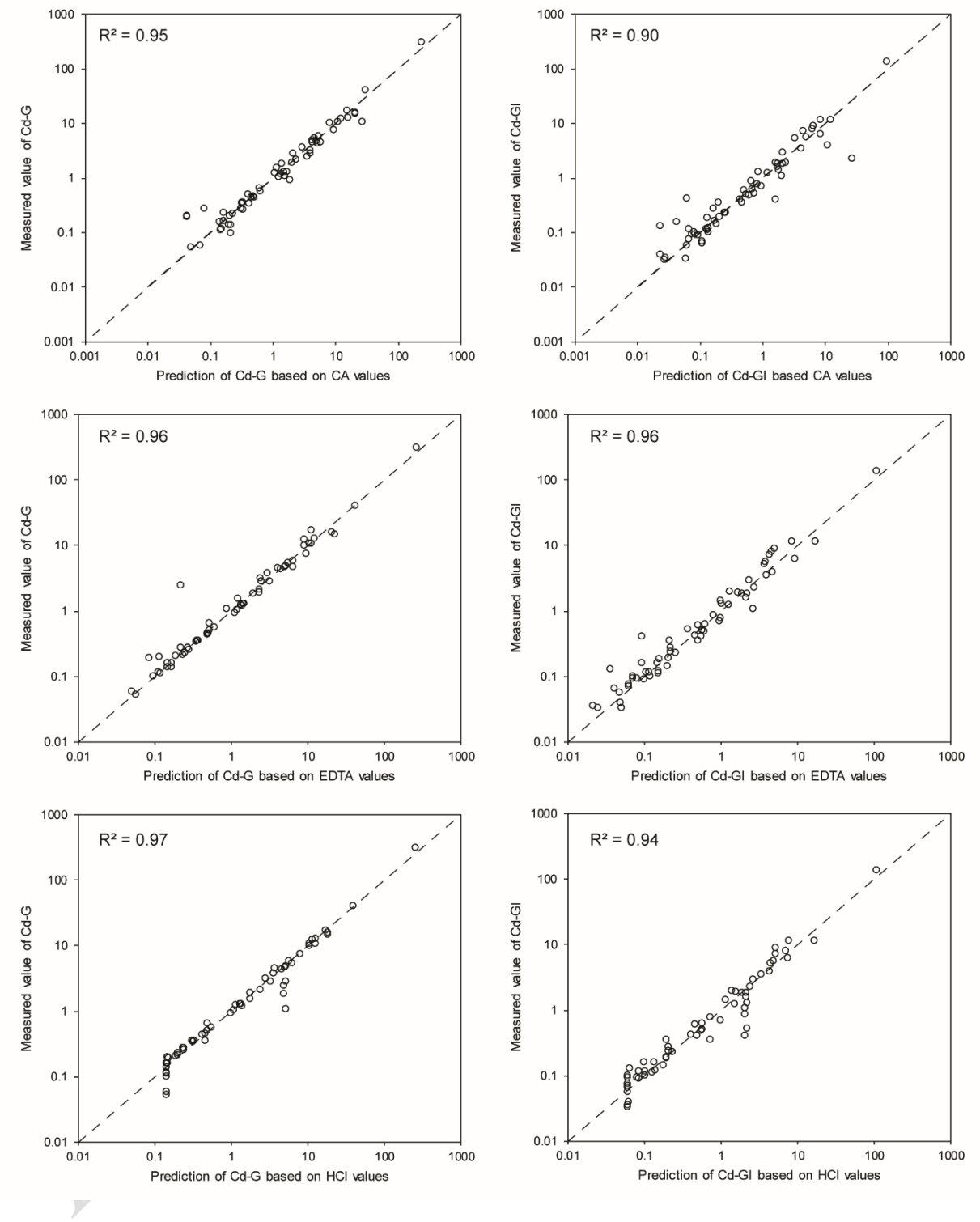



**Figure S3.** Bioaccessibility of Cd in gastric (Cd-G) and gastrointestinal (Cd-GI) phases measured by UBM versus predicted values by CA (when considering the two groups according to their carbonate contents), EDTA, and HCl for model validation ($n$ = 61, test set).

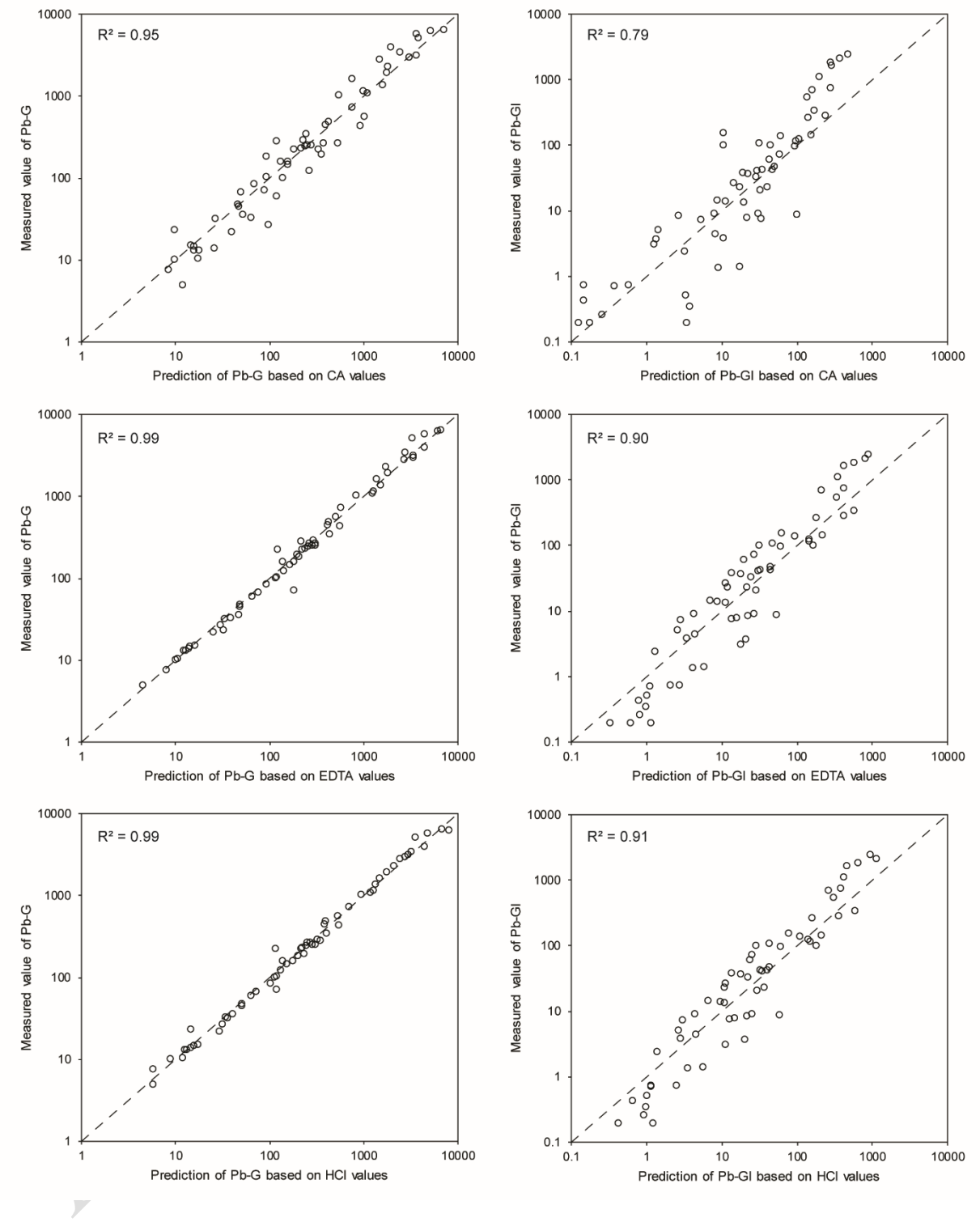



**Figure S4.** Bioaccessibility of Pb in gastric (Cd-G) and gastrointestinal (Cd-GI) phases measured by UBM versus predicted values by CA, EDTA, and HCl for model validation ($n$ = 61, test set).

**Additional investigations for As**

A total of 27 soil samples were collected from mining settings. The soil samples were air-dried at a temperature below 40°C and sieved at 2 mm and then at 250 µm. The analytical determinations consisted in measurements of: (i) the pseudototal concentrations of As; (ii) the bioaccessible concentrations of this element in the gastric (G) and gastrointestinal (GI) phases using UBM (unified bioaccessibility method); and (iii) the concentrations of As extracted by the HCl extraction method.

For this additional test set, the contamination level of As in the soil samples varied from 68 to 2641 mg $kg^{-1}$ (with a mean of 422 mg $kg^{-1}$).

When considering both the initial test set ($n$ = 61) and the additional test set ($n$ = 27), the validation was performed by taking into account the equations established with HCl to mimic the bioaccessibility of As in the G and GI phases (Figure S5). Results showed that the HCl extraction method can predict As bioaccessibility in both phases quite successfully ($R^2$ = 0.94 and 0.95, respectively, for the G and the GI phases) in soils with a wide range of As concentrations (1.9–2641 mg $kg^{-1}$).

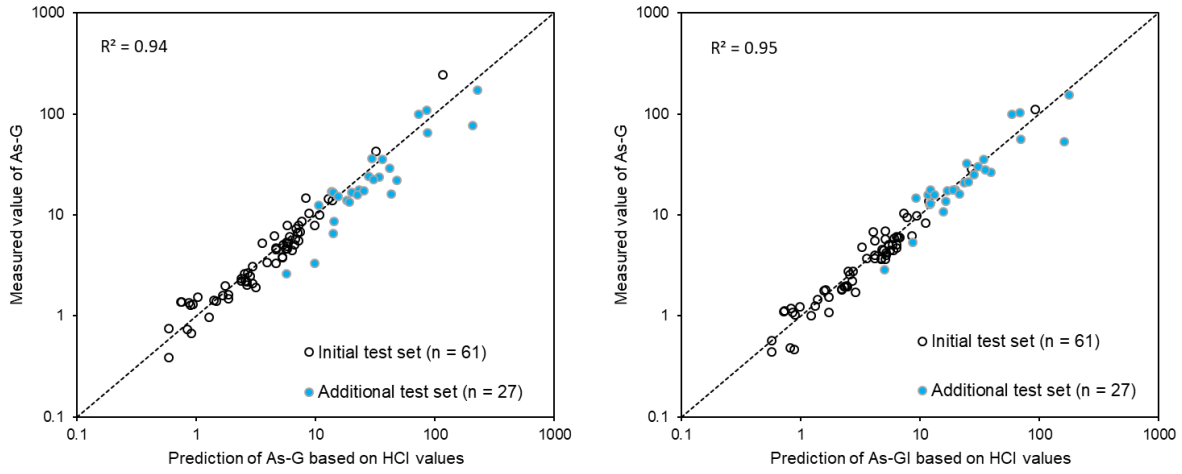



**Figure S5.** Bioaccessibility of As in gastric (As-G) and gastrointestinal (As-GI) phases measured by UBM versus predicted values by HCl for model validation from the initial test set ($n$ = 61) and from the additional test set ($n$ = 27).