JOURNAL OF SYNCHROTRON RADIATION

ISSN 1600-5775

Multi-element effects on arsenate accumulation in a geochemical matrix determined using μ-XRF, μ-XANES and spatial statistics

Aakriti Sharma,[a] Amanda Muyskens,[b]‡ Joseph Guinness,[b]§ Matthew L. Polizzotto,[a]¶ Montserrat Fuentes,[b]‡‡ Ryan V. Tappero,[c] Yu-chen K. Chen-Wiegart,[c]§§ Juergen Thieme,[c] Garth J. Williams,[c] Alvin S. Acerbo[c]¶¶ and Dean Hesterberg[a]*

[a]Department of Crop and Soil Sciences, North Carolina State University, Raleigh, NC 27695, USA, [b]Department of Statistics, North Carolina State University, Raleigh, NC 27695, USA, and [c]National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY 11973, USA. *Correspondence e-mail: dean_hesterberg@ncsu.edu

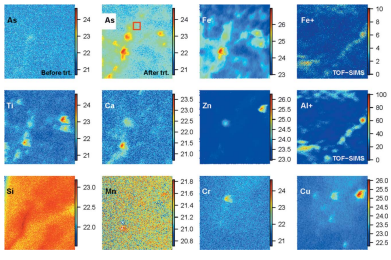
Received 31 May 2019
Accepted 13 September 2019

Edited by S. M. Heald, Argonne National Laboratory, USA

‡ Current affiliation: Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA.
§ Current affiliation: Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853, USA.
¶ Current affiliation: Department of Earth Sciences, University of Oregon, Eugene, OR 97403, USA.
‡‡ Current affiliation: Office of the Executive Vice President and Provost, University of Iowa, Iowa City, IA 52242, USA.
§§ Current affiliation: Department of Materials Science and Chemical Engineering, Stony Brook University, Stony Brook, NY 11794, USA and National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY 11973, USA.
¶¶ Current affiliation: Center for Advanced Radiation Sources (CARS), University of Chicago, Chicago, IL, 60637, USA.

Keywords: reactive microsites; multi-component complexity; arsenic; partial correlation; spatial regression.

Supporting information: this article has supporting information at journals.iucr.org/s



Soils regulate the environmental impacts of trace elements, but direct measurements of reaction mechanisms in these complex, multi-component systems can be challenging. The objective of this work was to develop approaches for assessing effects of co-localized geochemical matrix elements on the accumulation and chemical speciation of arsenate applied to a soil matrix. Synchrotron X-ray fluorescence microprobe (μ-XRF) images collected across 100 μm × 100 μm and 10 μm × 10 μm regions of a naturally weathered soil sand-grain coating before and after treatment with As(V) solution showed strong positive partial correlations (r' = 0.77 and 0.64, respectively) between accumulated As and soil Fe, with weaker partial correlations ($r' > 0.1$) between As and Ca, and As and Zn in the larger image. Spatial and non-spatial regression models revealed a dominant contribution of Fe and minor contributions of Ca and Ti in predicting accumulated As, depending on the size of the sample area analyzed. Time-of-flight secondary ion mass spectrometry analysis of an area of the sand grain showed a significant correlation (r = 0.51) between Fe and Al, so effects of Fe versus Al (hydr)oxides on accumulated As could not be separated. Fitting results from 25 As *K*-edge microscale X-ray absorption near-edge structure (μ-XANES) spectra collected across a separate 10 μm × 10 μm region showed ∼60% variation in proportions of Fe(III) and Al(III)-bound As(V) standards, and fits to μ-XANES spectra collected across the 100 μm × 100 μm region were more variable. Consistent with insights from studies on model systems, the results obtained here indicate a dominance of Fe and possibly Al (hydr)oxides in controlling As(V) accumulation within microsites of the soil matrix analyzed, but the analyses inferred minor augmentation from co-localized Ti, Ca and possibly Zn.

1. Introduction

Contamination of soils and water by toxic trace elements from natural or anthropogenic sources threatens human and ecosystem health. Soils regulate retention and mobilization of trace elements mainly by pH- and redox-affected adsorption–desorption and precipitation–dissolution reactions (Basta *et al.*, 2005; Borch *et al.*, 2010; Cullen & Reimer, 1989; Fendorf *et al.*, 2010; Inskeep *et al.*, 2001; McBride, 1989; Park *et al.*, 2016). However, directly measuring reaction mechanisms that are important for preventing adverse environmental impacts of trace elements is confounded by the association of minerals, organic matter, non-crystalline inorganic solids and biota into complex soil assemblages. Consequently, studies aimed at

determining mechanisms of trace-element binding in soils and other geochemical systems have often used model analogs of isolated soil components, including synthesized minerals or extracted humic substances (Adra *et al.*, 2013; Brown & Sturchio, 2002; Dixit & Hering, 2003; Ehlert *et al.*, 2018; Fritzsche *et al.*, 2011; Grafe *et al.*, 2001; Masue *et al.*, 2007; Mikutta & Kretzschmar, 2011; Otero-Fariña *et al.*, 2017; Raven *et al.*, 1998; Silva *et al.*, 2010; Violante & Pigna, 2002; Zeng *et al.*, 2008). The multicomponent complexity of soil makes it difficult to discriminate bonding mechanisms at the level of specificity that can be determined from model systems (Kizewski *et al.*, 2011).

Here we combined synchrotron μ-XRF and μ-XANES analyses to study the retention of arsenic in a complex soil matrix. Past research on As in natural geochemical matrices such as soils, subsurface and sediments generally showed that As retention is affected by oxidation state, with As(V) typically being sorbed to a greater extent than As(III) (Fendorf & Kocar, 2009; Fendorf *et al.*, 2010). Model-system studies showed that As(V) (arsenate) is retained by inner-sphere surface complexes on Fe-, Al- and Mn-oxides (Adra *et al.*, 2013; Arai *et al.*, 2001; Di Iorio *et al.*, 2018; Dzade & De Leeuw, 2018; Fendorf *et al.*, 1997; Foster *et al.*, 2003; Ladeira *et al.*, 2001; Manceau, 1995; Manning *et al.*, 2002; Silva *et al.*, 2010; Waychunas *et al.*, 1995; Zhu & Elzinga, 2015). Although one can envision a direct connection between model systems of individual geochemical components (minerals, *etc.*) and natural geochemical systems containing similar components, this connection is not straightforward. For example, research on simple binary mixtures of model components of Fe and Al (hydr)oxides (Antelo *et al.*, 2015; Khare *et al.*, 2005; Liu & Hesterberg, 2011; Murray & Hesterberg, 2006), minerals and organic matter (Alcacio *et al.*, 2001; Chen & Sparks, 2018; Davis, 1984; Hu *et al.*, 2018; Mikutta & Kretzschmar, 2011; Otero-Fariña *et al.*, 2017; Sowers *et al.*, 2018), and minerals and microbes (Du *et al.*, 2017; Gadd, 2010) all show non-additive effects on the binding of phosphate or arsenate. Such non-additive effects could limit the transferability of trace-element binding mechanisms from model systems to soils and other multicomponent geochemical matrices, and confound our ability to measure specific matrix components that are responsible for trace-element binding, redox transformations and other reactions.

Here we present a unique combination of analytical and statistical approaches for assessing whether retention of As analyzed at the microscale is affected by multiple chemical elements in soil solids that are co-localized at this analytical scale. For example, elements in geochemical solids that are known from model systems to affect As geochemical speciation include Al in Al (hydr)oxides that adsorb As(V) or As(III); Fe and Mn in (hydr)oxides that also adsorb As and oxidize As(III) to As(V); C(0,−IV) in organic matter that can reduce As(V) to As(III); and S(−II) in organic matter and sulfide minerals that can reduce As or form As-sulfide minerals (Anderson *et al.*, 1976; Borch *et al.*, 2010; Cullen & Reimer, 1989; Fritzsche *et al.*, 2011; Grafe *et al.*, 2001; Jiang *et al.*, 2009; Manning *et al.*, 2002; Mikutta & Kretzschmar, 2011; Violante & Pigna, 2002). Synchrotron μ-XRF imaging was used to assess As co-localization with a suite of geochemical matrix elements, and μ-XANES was used to assess As speciation within microscale volumes (microsites) of soil material (Brown & Sturchio, 2002; Burton *et al.*, 2014; Foster & Kim, 2014; Gamble *et al.*, 2018; Gräfe *et al.*, 2008; Kim *et al.*, 2013; Kopittke *et al.*, 2017; Langner *et al.*, 2013; LeMonte *et al.*, 2017; Mikkonen *et al.*, 2019; Serrano *et al.*, 2015; Strawn *et al.*, 2002; Voegelin *et al.*, 2007; Wovkulich *et al.*, 2012).

Many studies have used simple, pairwise (Pearson) correlation analysis of μ-XRF images to infer bonding mechanisms via co-localization of As with geochemical-matrix elements, including Fe, Mn, S, Cu, Cr and Zn (Fan *et al.*, 2014; Gillispie *et al.*, 2016; Gräfe *et al.*, 2008; Landrot *et al.*, 2012; Langner *et al.*, 2013; LeMonte *et al.*, 2017; Mikkonen *et al.*, 2019; Schwer & McNear, 2011; Strawn *et al.*, 2002; Voegelin *et al.*, 2007). However, simple correlations do not account for co-localization of the matrix elements themselves, which occurs, for example, at the molecular scale with Al and Fe in Al-substituted Fe (hydr)oxides and at the microscale when multiple colloidal solids are physically associated. Moreover, simple correlations do not account for spatial autocorrelation of each element to itself as a function of distance from a given point in the sample, *e.g.* moving across a concentrated 'hotspot' of an element. Autocorrelation effects are expected to increase with increasing spatial resolution of the imaging tool. In fact, applying a simple correlation to spatially correlated imaging data violates the assumption of independent data points across space and can lead to unreliable significance tests and biased parameter estimates (Beale *et al.*, 2010). Therefore, more advanced statistical approaches that account for the spatial component of μ-XRF data (Guinness *et al.*, 2014; Terres *et al.*, 2018) should better discriminate As retention in relation to individual geochemical matrix elements. Ultimately, integrating such statistical approaches into complex-system analyses using multiple spatio-analytical techniques, *e.g.* imaging, spectroscopy, diffraction (Manceau *et al.*, 2002), should increase the power of such techniques in determining trace-element binding mechanisms in complex, multi-component geochemical systems.

Our overarching research goal is to determine how interactions between multiple co-localized soil components affect trace element reactivity. The specific objective of this research was to develop synchrotron μ-XRF and μ-XANES approaches for evaluating effects of co-localized soil matrix elements on As retention within reactive microsites (Hesterberg *et al.*, 2011; Werner *et al.*, 2017). To avoid historical effects of legacy contaminants, we present a novel technique for directly reacting a non-contaminated soil matrix coating on a quartz sand grain with aqueous arsenate, and imaging the same sample area before and after the treatment. Partial correlation (Guinness *et al.*, 2014) and spatial regression (Guinness, 2018; Hoeting *et al.*, 2006; Minasny & McBratney, 2005, 2007) analyses were applied to multi-element images, and μ-XANES spectroscopy was used to determine chemical speciation.

2. Materials and methods

2.1. Sand-grain collection and preparation

A sample was collected from the B horizon (40–50 cm depth) of a forested Wagram soil (loamy, kaolinitic, thermic Arenic Kandiudults) at the Central Crops Research Station in Clayton, NC, USA. Soil material was scraped from a newly exposed face of a soil pit using a glass rod to avoid metal contamination. The sample was transported on ice and stored in a refrigerator until used. Characterization data for the sampled soil profile are reported elsewhere (National Cooperative Soil Survey Characterization Database, https://ncsslabdatamart.sc.egov.usda.gov/rptExecute.aspx?p=25542&r=1&). The bulk soil sample had a pH of 5.5 and contained 1.2 ± 0.1 g kg^{-1} of carbon in total and 0.1 cmol kg^{-1} of exchangeable Ca. Oxalate extractable Fe and Al in this soil were 107 ± 8 mg kg^{-1} (n = 2) and 188 ± 4 (n = 2) mg kg^{-1}, respectively. A 30 g sub-sample was dispersed in deionized water (1:10 soil:water) by shaking on a reciprocating shaker at ambient temperature for 0.5 h at a rate of 5.7 s^{-1}. Sand grains collected on a polypropylene mesh with 1 mm openings were repeatedly rinsed with deionized water to remove fine particles, then dried in an acrylic Petri dish under N_2. All grains were handled with plastic forceps to avoid metal contamination. For our analyses, one quartz sand grain with an ∼2 mm × 1.5 mm face containing areas of thin, reddish-colored Fe-oxide coatings on a visibly flat surface was selected from among many grains observed under a light microscope (see Fig. S1 of the supporting information). The grain was physically fixed (without glue) between two walls separated by 1.4 mm that were machined into a polypropylene sample mount configured for attachment to a kinematic sample holder for the Submicron Resolution X-ray Spectroscopy (SRX) beamline (5-ID) at the National Synchrotron Light Source II (NSLS-II), Brookhaven National Laboratory in Upton, NY, USA (Chen-Wiegart *et al.*, 2016).

2.2. Micro-X-ray fluorescence imaging

Spatial distributions of chemical elements in the sand-grain coating were imaged using μ-XRF analysis at the SRX beamline. Images were collected from the same 100 μm × 100 μm, reddish-colored region (ROI-100; Fig. S1) before and after treatment with aqueous As(V). The ∼1 μm × 1 μm beam of 13.5 keV incident energy was step scanned using a 0.5 s dwell time. For data management, 25 μ-XRF elemental images of 20 μm × 20 μm sub-areas were collected in sequence and batch-stitched into each composite image of ROI-100. However, one sub-area image was inadvertently omitted during data collection after As(V) treatment, giving a dataset with a total of 38 400 voxels for the treated sample. Also, because the images collected before and after the As(V) treatment were misaligned in the vertical (by 9.5 μm), the images collected before and after treatment were aligned statistically by shifting one vertically in 0.5 μm steps so that the absolute error of differences in Ti fluorescence signals between images was minimized. This shift also decreased the number of voxels for statistical analyses to 34 284. We also collected a μ-XRF image of a 10 μm × 10 μm subregion (ROI-10a) of ROI-100 from the As(V)-treated sample at 12.1 eV incident energy, the highest energy of a three-dimensional (x, y, energy) μ-XANES stack that we collected but have not presented here. The aim of analyzing μ-XRF data from ROI-10a was to increase the likelihood of detecting any effects of Ti, Ca, Mn, Ni, Zn and Cu on As accumulation, given that the Fe fluorescence signal was less dominant than signals from these other matrix elements. For all μ-XRF images, the full fluorescence spectrum from a multichannel analyzer (MCA) for each voxel was fit using the *PyXRF* software (Li *et al.*, 2017) to extract element-specific fluorescence intensities. Details can be found in the *Additional Methods* section of the supporting information.

Our unique sampling approach provided a means to image the effects of a chemical treatment on a thin layer of soil matrix while avoiding the use of chemical resins. *HEPHAESTUS* software (Ravel & Newville, 2005) calculations indicated that a 13.5 keV incident X-ray beam is attenuated by 95% after passing through 1 mm of the quartz core of a sand grain, so only the front-surface coating was imaged. Effectively, the grain coating is a naturally thin section, although it is likely to have microscale thickness variations across the imaged area.

2.3. As(V) treatment

After imaging ROI-100 of the non-treated sand grain, the kinematic sample holder was removed from the sample stage to treat the sand grain with As(V) solution under ambient conditions. The mounted grain was first covered with a 150 μl drop of deionized water for 30 min to hydrate, then the excess water was carefully removed via capillarity by touching the side of the drop with a Kimwipe without touching the imaged surface. The moistened grain was treated with a 150 μl drop of 0.1 mM KH_2AsO_4 in a 0.1 mM KCl background solution, which was pre-adjusted to pH 5.0 to approximately match the pH of the soil. Preliminary experiments showed that this As concentration gave contrasting fluorescence signals across such grains. After equilibrating for 35 min, the As(V)-treatment solution and four subsequent 150 μl rinses with deionized water were recovered with the pipet and combined for analysis of any dissolved or weakly bound As using inductively coupled plasma optical emission spectrometry (ICP-OES). Based on a dissolved As concentration of 3.6 μM in the recovered sub-sample of ∼600 μl of aqueous As(V) plus rinsate solutions that was measured gravimetrically to the nearest 0.0001 g, the total As accumulated on the sand grain determined based on loss from solution was ∼0.0128 μmol or ∼85% of applied As(V). The sand grain was dried in an N_2-purged glovebox chamber maintained for 1 h under a weak vacuum of −20 kPa, which was applied incrementally to avoid boiling. The kinematic holder with the dried grain was re-mounted on the beamline stage, and re-aligned by coarse mapping a small area around an Fe hotspot before collecting a post-treatment μ-XRF image across ROI-100.

2.4. Micro-XANES spectroscopy and linear combination fitting

As *K*-edge μ-XANES spectra were collected from the As(V)-treated sand grain using the NSLS-II X-ray Fluorescence Microprobe (XFM) beamline, with an ~2 μm × 2 μm spot size. The Si(111) monochromator was calibrated to 11875 eV at the white-line peak of As(V) imaged in a micrometre-sized inclusion in a topaz crystal. Details of the data collection and analysis methods are summarized in the supporting information.

Spectra were collected at each of six locations (spectra 1L–6L) across a 200 μm × 200 μm region (ROI-200) that included ROI-100 imaged at the SRX beamline (Fig. S2). The spectrum for location 6L represents an average of 25 spectra collected across a 10 μm × 10 μm area within ROI-200. Also, 25 individual spectra were separately collected from a 10 μm × 10 μm subregion (ROI-10b) within ROI-200 (Fig. S3), which showed a concentration gradient of elements, in particular As and Fe. All μ-XANES spectra were collected between 11767 eV and 12117 eV, with a minimum step size of 0.5 eV across the edge region and 10 s dwell times.

The μ-XANES spectra were analyzed using the *ATHENA* software (Ravel & Newville, 2005) and established procedures (Kelly *et al.*, 2008). The spectra were baseline subtracted using a linear function between −100 eV and −30 eV relative to the first-derivative maximum (E0), normalized using a linear function between 40 eV and 240 eV, and flattened to remove quadrature. Linear combination fitting (LCF) analyses of μ-XANES spectra were performed across an energy range from −10 eV to 25 eV relative to E0. Fitting standards that were considered to be representative model analogs of possible species found in soils included spectra from ten adsorbed As(V) or As(III) standards that were collected previously at unfocused beamlines – X11A, X11B or X18B (Lopez *et al.*, 2018) – and μ-XANES spectra from scorodite and an arsenate–ferrihydrite coprecipitate (co-ppt) collected at the XFM beamline along with our sample spectra (Table S1 and Fig. S4 of the supporting information). Spectra from the unfocused beamlines were calibrated by simultaneously collecting data from an As(0) powder reference (E0 = 11 867 eV) or an Au foil (E0 = 11 919 eV). Reported fits were obtained using a modification of the standard-elimination approach (Manceau *et al.*, 2012), with combinatoric fitting (Kelly *et al.*, 2008) performed as a consistency check for fits to spectra 1L–6L (details are given in the supporting information).

2.4.1. Geochemical-based merging of As μ-XANES spectra. For LCF analysis of the 25 μ-XANES spectra from ROI-10b, we used a unique approach of grouping μ-XANES spectra according to microsite chemistry. Because As was largely correlated with Fe accumulation across ROI-100 (discussed below), and As(V) is known to have a high affinity for Fe(III)-(hydr)oxides, we assumed that As speciation within our soil microsites would vary with Fe content. Therefore, As spectra were grouped into four quartiles based on Fe fluorescence signals intensities: 'high Fe' (75–100% of the maximum Fe signal), 'medium high Fe' (50–75% of maximum), 'medium low Fe' (25–50% of maximum) and 'low Fe' (0–25% of maximum); see Fig. S3. The six or seven As μ-XANES spectra from voxels grouped into each Fe quartile were merged before normalizing to effectively yield a signal-weighted average of diminished spectral noise, then normalized as described above. The modified standard elimination procedure described in the supporting information showed that these four merged spectra could all be fit with combinations of two standards: our arsenate-ferrihydrite coprecipitate and As(V) adsorbed on boehmite. These two standards were subsequently used to fit the 25 individual μ-XANES spectra.

2.5. TOF-SIMS analysis

Spatial distributions of Al and Fe were analyzed on the As(V)-treated sand grain by time-of-flight secondary ion mass spectrometry (TOF-SIMS) (TOF SIMS V instrument, ION TOF, Inc. Chestnut Ridge, NY, USA), with a pixel size of 0.39 μm × 0.39 μm and estimated penetration depth of 3 nm. Using an optical microscope, a 100 μm × 100 μm region was selected, which, based on physical features, was visually judged to be close to or overlapping the μ-XRF-imaged region ROI-100. Surface contamination was first removed by sputtering for 20 s with the Bi ion beam used in TOF-SIMS analysis, then mass spectral images of Al^+ and Fe^+ were collected (NCSU-AIF, 2018). Arsenic was not detected. Data were collected as a 256 × 256 voxel image and the acquisition time at each voxel was 60 μs.

2.6. Statistical analyses

The μ-XRF intensity data normalized to the incoming X-ray intensity (I_0) at the SRX beamline were transformed using a natural log to approximate a normal distribution, which is a criterion for conducting linear correlation and regression models. Natural log transformations also augmented visualization of data points in statistical plots, given the orders-of-magnitude variations in μ-XRF signals for many elements in the images that we collected. Statistical relationships were developed between natural log-transformed As accumulated in ROI-100 or ROI-10a versus soil matrix elements imaged before the As(V) treatment. We chose to analyze As accumulation in relation to μ-XRF signals of matrix elements before treatment to obtain more reliable information about the initial spatial distribution of native soil matrix elements with which the applied As(V) first reacted. We observed minor changes in spatial patterns of Ca, Cu, Cr and Zn after the As(V) treatment (discussed below). Analyses included simple (Pearson) correlation as has been conventionally used for μ-XRF data, partial correlation to remove effects of other co-localized matrix elements and multivariate regression modeling with or without spatial statistical analyses that accounted for autocorrelation of elements. These approaches were also used to evaluate the 25 μ-XANES fitting results from ROI-10b with respect to fluorescence signals of matrix elements, which were only imaged after the As(V) treatment at the XFM beamline.

Four specific multivariate regression models were evaluated for predicting μ-XRF intensities of accumulated As from intensities of soil matrix elements in ROI-100 and ROI-10a: (i) a simple multiple linear regression model (independent error model), (ii) a spatial-likelihood linear regression model (spatially correlated error model) (Guinness, 2018; Guinness *et al.*, 2014; Minasny & McBratney, 2005, 2007), (iii) a non-linear (second-order polynomial) regression model and (iv) a spatial-likelihood non-linear (second-order polynomial) regression model. All models were fit using *R* statistical software. Unlike simple ('non-spatial') linear regression models, the spatial-likelihood linear and non-linear models ('spatial models') were used to account for the spatial autocorrelation in the residuals (Beale *et al.*, 2010). These residuals were assumed to follow a Gaussian process and a Matérn isotropic covariance function (Guinness, 2018), which induces spatial correlation based on distance (Guinness, 2018; Minasny & McBratney, 2005, 2007). The Matérn covariance function describes the spatial covariance of a random variable process, *i.e.* the covariance between measurements at two points separated by a given distance. Three non-negative parameters in the covariance function were variance, range and smoothness. Variance measures the deviation of a random variable from its mean, range represents the distance to which data are correlated, and smoothness reflects the steepness of variations with distance (Minasny & McBratney, 2005). The spatial likelihood models were fit using a composite likelihood approach (Guinness, 2018). Backward variable selection was based on the Akaike Information Criteria (AIC), which estimates the relative quality of each statistical model for a given set of data (Snipes & Taylor, 2014), to parameterize linear effects of matrix elements, and these mean parameters were further tested for their significance at $\alpha = 0.05$. Once the significant matrix elements that described As accumulation were selected from the linear spatial model, only these elements were included in non-spatial linear models (with or without interaction terms) and a non-linear model for predicting As fluorescence intensities. Variable selection procedures were conducted independently for non-spatial and spatial models.

3. Results

3.1. Spatial patterns of soil matrix elements and accumulated arsenic

Micro-XRF images of ROI-100 and ROI-10a on the sand grain are shown in Figs. 1 and 2. TOF-SIMS images of Fe^{+} and

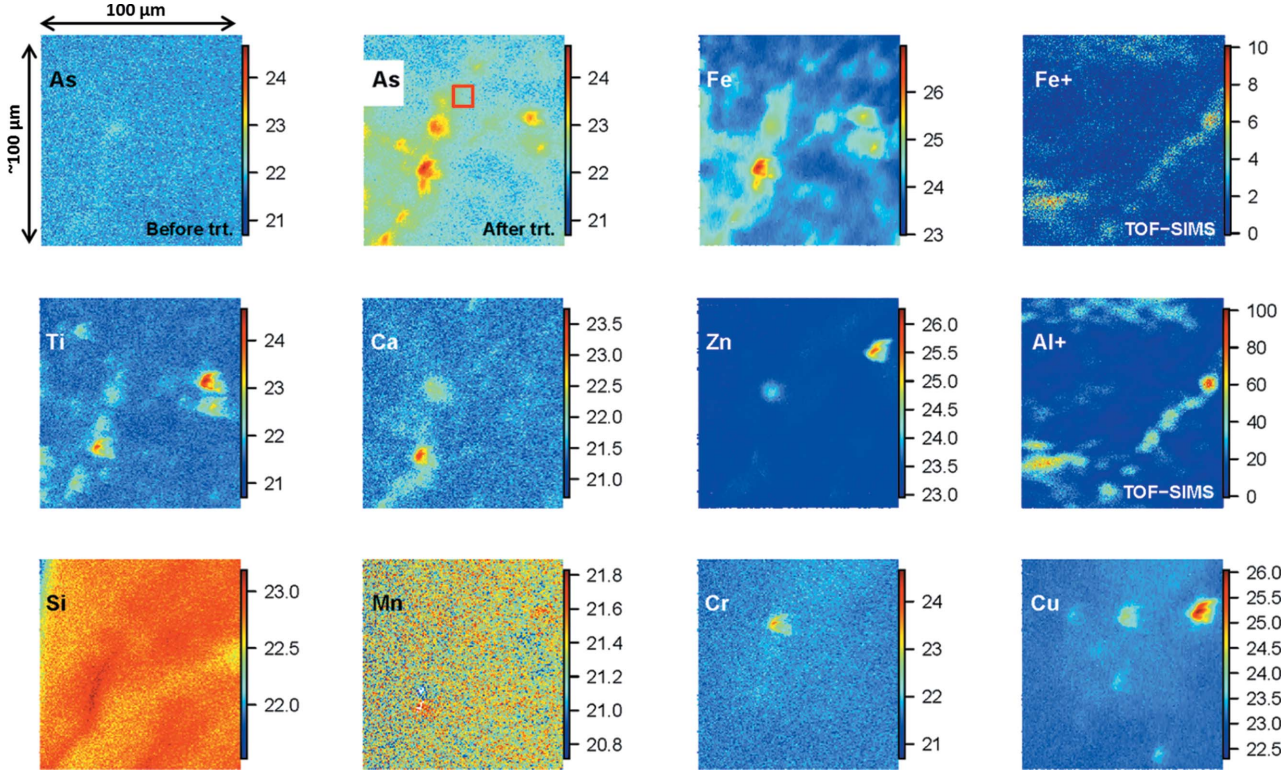


Figure 1
Micro-XRF images of As [before and after a 0.1 m*M* As(V) treatment] and soil matrix elements (before treatment) collected on the SRX beamline, along with TOF-SIMS images of Fe and Al collected on the As(V) treated sand grain. These images represent ROI-100, which had visible Fe-oxide coatings (Fig. S1). In μ-XRF images, scale bars show that brighter colors in the image represent greater natural log-transformed fluorescence signals for a given element in imaged soil volumes (microsites), and the linear scale bar in TOF-SIMS images indicates greater ion counts. The red square in the As image after treatment delineates ROI-10a, a 10 μm × 10 μm sub-region of ROI-100.

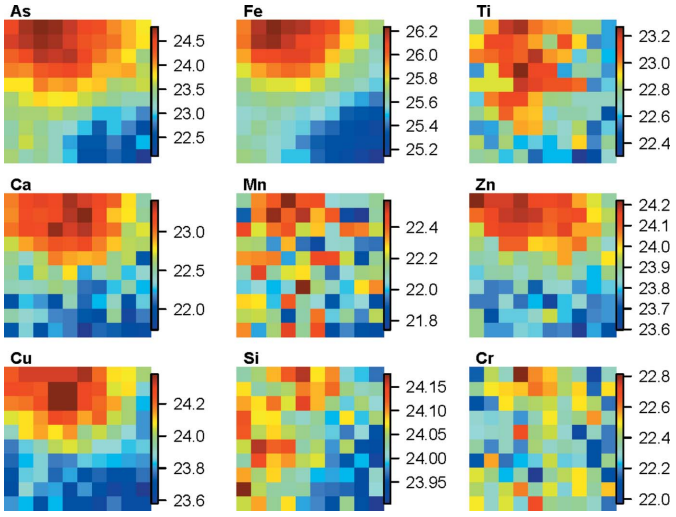


Figure 2
Spatial patterns of μ-XRF signals for As and selected soil matrix elements collected on the SRX beamline from a 10 μm × 10 μm sub-region (ROI-10a) of ROI-100 of the As(V) treated sand grain. Scale bars show that brighter colors represent greater natural log-transformed elemental fluorescence signals, generally corresponding to greater relative concentrations within imaged soil volumes (microsites).

Al^{+} collected in or near ROI-100 are included in Fig. 1 for comparison. Prior to the As(V) treatment, ROI-100 showed a uniform distribution of background As μ-XRF signals [$I_{As}/I_0 \leq \exp(22)$ counts s^{-1}, or $\ln(I_{As}/I_0) \leq 22$]. In contrast, heterogeneous spatial patterns and element-enriched 'hotspots' are seen for most of the soil-matrix elements imaged before treatment, including Fe, Ti, Ca, Zn, Cu and Cr (Fig. 1). The pattern for Si likely contains contributions from both the quartz core of the sand grain and any Si associated with minerals in the coating. After the As(V) treatment, μ-XRF signals for As increased by up to an order of magnitude or more compared with those before treatment, but still remained more than an order of magnitude lower than the dominant Fe signals. For example, the maximum As *Kα* fluorescence signal in ROI-100 after treatment was exp(24) counts s^{-1} compared with exp(27) counts s^{-1} for Fe. Overall, the spatial pattern of accumulated As was similar to those of Fe and Ti, and to a lesser extent Ca and Zn, in ROI-100 (Fig. 1). For ROI-10a, the As spatial pattern was most similar to those of Fe, Ca, Cu and Zn (Fig. 2). TOF-SIMS images showed similar spatial patterns for Fe and Al (Fig. 1), and these elements were significantly correlated ($r = 0.51$). It is noteworthy that spatial patterns of Ca, Zn, Cu and Cr showed minor visible changes after As(V) treatment, suggesting that the treatment partially mobilizes these elements (Figs. S5 and S6). Co-localization of matrix elements (*e.g.* Fe and Al, Fe and Ti, or Zn and Cu) within soil microsites illustrates the challenge of identifying which specific soil components containing these elements promoted, inhibited or had no effect on As accumulation. We used more advanced statistical analyses (partial correlation and spatial regression) of μ-XRF image data to address this challenge.

Table 1
Pearson correlation and partial correlation coefficients between natural log-transformed arsenic μ-XRF signals and those of soil matrix elements across a 100 μm × 100 μm region of a sand-grain coating (ROI-100) and 10 μm × 10 μm sub-region (ROI-10a).

ns: non-significant at $\alpha = 0.05$. *: significant at $\alpha = 0.05$.

	ROI-100 (No. of observations = 34284)		ROI-10a (No. of observations = 100)	
Matrix element	Correlation (r)	Partial correlation (r')	Correlation (r)	Partial correlation (r')
Fe	0.85*	0.77*	0.97*	0.64*
Ti	0.52*	0.05*	0.68*	-0.03^{ns}
Ca	0.33*	0.11*	0.9*	0.22*
Mn	−0.11*	-0.01^{ns}	0.38*	-0.05^{ns}
Cr	0.11*	0.08*	0.23*	-0.13^{ns}
Si	0.09*	−0.05*	0.58*	0.16^{ns}
Cu	−0.08*	−0.07*	0.92*	0.14^{ns}
Zn	0.05*	0.11*	0.85*	0.05^{ns}
Ni	−0.03*	0.07*	0.49*	-0.06^{ns}

3.2. Correlations of arsenic and soil matrix elements

Consistent with visible similarities of spatial patterns, accumulated As in ROI-100 showed the strongest simple (Pearson) correlations with Fe, Ti and Ca ($r = 0.85$, 0.52 and 0.33, respectively, see Table 1; Fig. S7). However, such pairwise correlations neglect effects of co-localization of these matrix elements themselves on As accumulation (Guinness *et al.*, 2014; Terres *et al.*, 2018). Partial correlations in Table 1 separate confounding effects of all other elements in each pairwise correlation (Guinness *et al.*, 2014). For the most highly correlated element pairs from ROI-100, partial correlation coefficients (r') decreased relative to Pearson correlation coefficients in the order As–Ti ($r = 0.52$, $r' = 0.05$) > As–Ca ($r = 0.33$, $r' = 0.11$) > As–Fe ($r = 0.85$, $r' = 0.77$), but also increased for As–Zn ($r = 0.05$, $r' = 0.11$) (Table 1). Thus, the higher Pearson correlations between As and Ca or Ti are largely due to their shared correlation with Fe (*e.g.* $r = 0.58$ for Fe–Ti and $r = 0.31$ for Fe–Ca; Fig. S7). Likewise, a decrease in the partial correlation coefficient for As–Fe relative to its simple correlation coefficient indicates that this pairwise relationship was also significantly affected by other co-localized elements, particularly Ti, Ca and Zn in ROI-100. However, an increase in the partial correlation coefficient for As–Zn relative to the simple correlation coefficient indicates that accounting for co-localized element effects amplified the As–Zn correlation. In fact, a strong linear trend between the natural-log transformed As and Zn signals for moderate Fe signals is visible in the middle cluster of the As–Zn correlation plot in Fig. S7. Partial correlations between other element pairs in ROI-100 were statistically significant ($\alpha = 0.05$) because of the large number of observations ($n = 34\,284$, Table 1).

For ROI-10a, accumulated As was most highly correlated with Fe, Ca, Zn and Cu; but partial correlations were only significant between As and Fe ($r = 0.97$, $r' = 0.64$) and As and Ca ($r = 0.90$, $r' = 0.22$; Table 1). The non-significant partial correlation between As and Ti in ROI-10a compared with ROI-100 is probably due to the order of magnitude

Table 2
Regression models developed for predicting As accumulation in ROI-100 based on μ-XRF fluorescence signals of soil matrix elements.

ns: elements that are non-significant at $\alpha = 0.05$.

Equation number	Models	Comment	AIC	R^2
(1)	As = −5.69 + 0.84 Fe + 0.19 Zn + 0.1 Ca + 0.11 Ni − 0.1 Cu + 0.08 Cr − 0.07 Si + 0.04 Ti + error	Non-spatial model (all predictors)	−23477.6	0.73
(2)	As = 13.03 + 0.35 Fe + 0.02 Ti + 0.02 $Zn^{(ns)}$ + 0.01 Ca − 0.01 $Si^{(ns)}$ − 0.003 $Cu^{(ns)}$ + 0.002 $Cr^{(ns)}$ − 0.002 $Ni^{(ns)}$ + error	Spatial model (all predictors)	−43551.3	0.47
(3)	As = 13.27 + 0.35 Fe + 0.02 Ti + 0.01 Ca + error	Spatial model (significant predictors)	−43555.3	0.47
(4)	As = −1.27 + 0.84 Fe + 0.12 Ca + 0.04 Ti + error	Non-spatial model (significant predictors)	−22453.1	0.72
(5)	As = 83.78 − 2.57 Fe − 1.84 Ti − 1.38 Ca + 0.08 Fe*Ti + 0.06 Fe*Ca + error	Spatial model (interaction terms)	−43630.8	0.48
(6)	As = 58.12 − 4.2 Ca − 1.6 Fe + 1.58 Ti + 0.18 Fe*Ca − 0.06 Fe*Ti + error	Non-spatial model (interaction terms)	−22695.7	0.73
(7)	As = 132.58 − 6.16 Fe + 0.13 Fe^2 − 2.21 Ca + 0.05Ca^2 − 1.47 Ti + 0.03 Ti^2 + error	Spatial model (polynomial)	−43761.4	0.47
(8)	As = 22.31 + 49.44 Fe − 0.08 $Fe^{2\ (ns)}$ + 4.45 Ca + 1.12 Ca^2 + 1.82 Ti + 0.10 $Ti^{2\ (ns)}$ + error	Non-spatial model (polynomial)	−22488.7	0.72

lower maximum fluorescence signals for Ti in ROI-10a (Figs. 1 and 2).

3.3. Regression models for predicting arsenic accumulation

Spatial regression models that account for autocorrelation in predicting accumulated As in ROI-100 and ROI-10a based on μ-XRF signals of soil matrix elements in Figs. 1 and 2 are compared with non-spatial models in Table 2. All eight spatial and non-spatial regression models that we evaluated for ROI-100 are parameterized in Table 2, and the predictive models compared are shown in Figs. 3 and S8. Eight different regression models were included: (1) a simple (non-spatial) multiple linear regression model (full model); (2) and (3) spatial likelihood linear models that included all predictors from equation (1) (full model) and only significant predictors; (4) a non-spatial model that included only significant predictors selected from spatial model (2); (5) and (6) spatial and non-spatial models – with and without interaction terms – that included only significant predictors from model (2); (7) and (8) spatial and non-spatial, non-linear (second-order polynomial) models. Spatial models gave minimal (more negative) AIC values but yielded a lower R^2, whereas non-spatial models gave greater AIC values relative to spatial models but necessarily yielded greater R^2 values (Table 2) and provided better in-sample predictions (Figs. 3 and S8). The improvement in R^2 occurred in the non-spatial models because the regression coefficients in these models minimize the sum of squared errors, whereas the regression coefficients in the spatial models do not minimize the sum of squared errors and thus yielded lower R^2 values.

A linear, non-spatial model [equation (1) in Table 2] showed that μ-XRF signal intensities for Fe, Ti, Ca, Cr, Si, Cu, Zn and Ni were all significant predictors of As signal intensities (R^2 = 0.73) in ROI-100. It would be difficult to draw physical inferences about possible soil components binding As from a statistical model with so many predictors. However, the number of significant predictors diminished to three (Fe, Ti and Ca fluorescence intensities) in a linear spatial model that accounted for autocorrelation [equations (2) and (3) in Table 2], *i.e.* the decrease in correlation of the error to itself as a function of distance from a given voxel (see variogram in Fig. S9). Note that including or excluding non-significant variables yielded essentially the same parameters in these spatial models. Because these models account for autocorrelation, they provide a more robust test for selecting significant predictor variables (*i.e.* μ-XRF signal intensities of matrix elements) than is obtained using a non-spatial model (Beale *et al.*, 2010). The importance of accounting for spatial autocorrelation to improve the selection of predictor variables

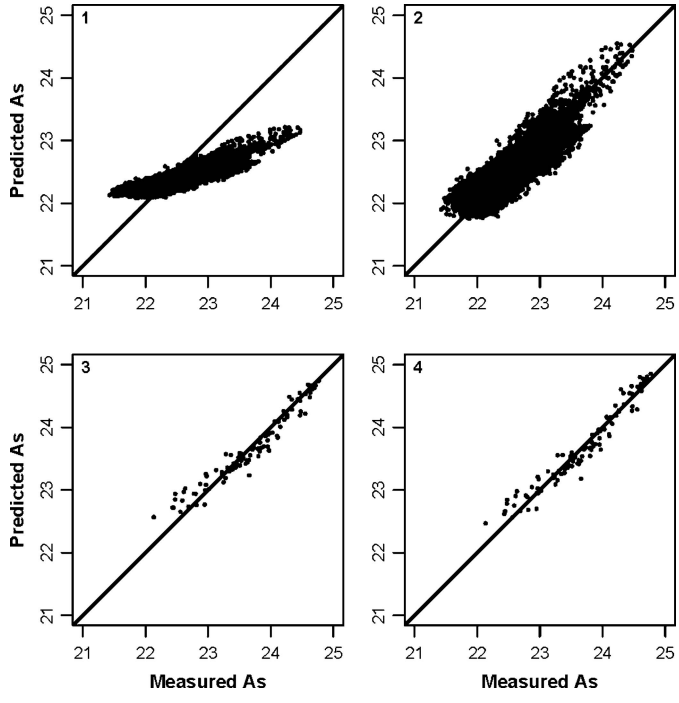


Figure 3
Scatter plots of predicted versus measured natural log-transformed arsenic μ-XRF signals for selected spatial (plots 1, 3) and non-spatial (plots 2, 4) predictive models developed for ROI-100 (plots 1, 2) and ROI-10a (plots 3, 4). Plots 1 and 2 correspond with equations (3) and (4) in Table 2 and plots 3 and 4 correspond with equations (10) and (9) in Table 3; the 1:1 lines are plotted to compare predictions.

Table 3
Regression models developed for predicting As accumulation in ROI-10a based on μ-XRF fluorescence signals of soil matrix elements.

Equation number	Models	Comment	AIC	R^2
(9)	As = −30.71 + 1.87 Fe + 0.28 Ca + error	Non-spatial model	−77.01	0.94
(10)	As = −26.49 + 1.79 Fe + 0.18 Ca + error	Spatial model	−100.58	0.93
(11)	As = −431.28 + 17.82 Ca + 17.60 Fe − 0.69 Fe*Ca + error	Spatial model	−109.29	0.95
(12)	As = −456.20 + 19.02 Ca + 18.45 Fe − 0.73 Fe*Ca + error	Non-spatial model	−104.71	0.96
(13)	As = −221.20 + 16.79 Ca − 0.37 Ca^2 + 2.13 Fe + error	Spatial model	−107.94	0.95
(14)	As = −11.53 + 2.09 Fe + 0.68 Ca − 0.85 Ca^2 + error	Non-spatial model	−105.052	0.96

in spatial models is described by Hoeting *et al.* (2006). However, the spatial model did not predict As as well as the non-spatial model [Fig. 3 (plot 1) versus Fig. S8 (plot 1)]. Therefore, we used the signals of Fe, Ti and Ca that were selected as significant predictors from equation (2) to derive the non-spatial model for ROI-100 [equation (4) in Table 2; Figs. 3 (plot 2)], which necessarily has a higher R^2. Also note that the inclusion of only three predictor elements in equation (4) explained 72% of the variation in As fluorescence, whereas the inclusion of eight predictors in equation (1) explained only 73% of the variation. This shows that the addition of more variables in a model does not necessarily improve model prediction and justifies accounting for autocorrelation via a spatial model to select significant predictor variables for inclusion in a non-spatial model.

Using the same predictor variables selected from equation (2), we also developed spatial and non-spatial models that included interaction terms for Fe*Ti and Fe*Ca, which were significant [equations (5) and (6) in Table 2]. Finally, we evaluated any nonlinear effects of Fe, Ti and Ca by developing second-order polynomial models [equations (7) and (8) in Table 2]. All three predictor variables were significant in the non-linear spatial model, but only Ca was a significant predictor of As in the non-linear, non-spatial model. Nevertheless, R^2 values in Table 2 indicate that the linear model in equation (4) gave equal or better predictions of As μ-XRF signals than linear models with interaction terms [equations (5) and (6)] or non-linear models [equations (7) and (8)], and equation (4) is the simplest model.

Similar to the approach described above, we used a linear spatial regression model to determine the significant predictors for a non-spatial model for μ-XRF data from ROI-10a [equations (9) and (10) in Table 3]. In this case, both non-spatial [equation (9)] and spatial [equation (10) in Table 3] linear models showed only Fe and Ca (not Ti or other elements) as significant (α = 0.05; n = 100) predictors of As accumulation based on μ-XRF signal intensities. Both models gave comparable predictions based on R^2 values and visibly consistent scatter plots of predicted versus measured μ-XRF signals for As (Fig. 3, plots 3 and 4; Fig. S10), which might be due to the small spatial range. Like ROI-100, the inclusion of a statistically significant Fe*Ca interaction term in the linear spatial and non-spatial models [equations (11) and (12)] and non-linear models [equations (13) and (14)] gave comparable predictions as the simple linear model [equations (9) and (10)], even though the inclusion of more predictor variables in the models necessarily gave marginally greater R^2 values (Table 3). Overall, the significant regression coefficients on μ-XRF signals of matrix elements in our predictive non-spatial models [equations (4) and (9)] inferred that Ca and Ti augmented As retention in the sand-grain coating, but these elements were only 15% and 5% as effective as Fe in promoting As accumulation, respectively.

3.4. Spatial-dependent soil arsenic speciation

Fluctuations in As *K*-edge μ-XANES spectra and LCF results from six spots across ROI-200 (Fig. 4, Table 4) and 25 voxels within ROI-10b (Figs. 5 and 6, Table S2) indicated variations in soil As speciation across imaged regions of the sand-grain coating. Most of the μ-XANES spectra had a strong white line near 11875 eV, showing a dominance of As(V). Although the spectra had only minor differences in spectral features, variations in white-line intensities or broadening of the white lines in Figs. 4(*a*) and 5 indicate differences in As speciation between the soil microsites analyzed. Consistent with the strong positive correlation of As and Fe, the level of noise in individual spectra as shown in ROI-10b increased with decreasing relative Fe (and As) content (Fig. 5). Differences in As speciation between six spots separated by up to ∼140 μm across ROI-200 (Fig. S2) are indicated by variations of up to 31% in proportions of five fitting standards (Fig. 4, Table 4). Three standards containing As(V) bonded with Fe [As(V)-ferrihydrite coprecipitate, As(V) adsorbed on goethite or As(V)–Fe(III)-peat complex] accounted for 46–77% of the μ-XANES spectral features across these spots, and two standards of As(V) bonded with Al [As(V) adsorbed on boehmite or in mansfieldite] accounted for the remaining 23–54% of spectral features. It is important to note that variations in proportions of fitting standards are best interpreted as showing variations in relative (rather than absolute) speciation of As. For example, it seems unlikely that the mineral mansfieldite ($AlAsO_4{\cdot}2H_2O$) formed during the 35 min contact time of our As(V) treatment, but perhaps this standard has spectral features that are similar to other Al-bonded As species in our sample that are not represented by our selected fitting standards. Also, it is important to point out that standards included in LCF analysis showed only minor differences in spectral features [Figs. 4(*b*) and S4], thus interpretations of fits in terms of definitive speciation should be made with caution. Nevertheless, because different fitting approaches always yielded similar LCF results, we have

Table 4
Linear combination fitting results showing best-fit combinations of standards for As μ-XANES spectra collected from six As-enriched spots across a 200 μm × 200 μm region of the sand grain (ROI-200; fits are overlaid on data in Fig. 4).

	Proportions of standards ± uncertainty					
Spot	As(V)-ferrihydrite (co-ppt)	As(V)-goethite	As(V)-Fe-peat†	As(V)-boehmite	Mansfieldite ($AlAsO_4{\cdot}2H_2O$)	*R*-factor
1L	77 ± 8			23 ± 8		0.0141
2L			61 ± 5	39 ± 5		0.0046
3L	64 ± 5			36 ± 5		0.0047
4L		46 ± 6		54 ± 4		0.0113
5L		58 ± 4		42 ± 4		0.0042
6L	59 ± 4			31 ± 4	10 ± 4	0.0027

† As(V) sorbed on peat that was pretreated with 2400 mmol Fe kg^{-1}.

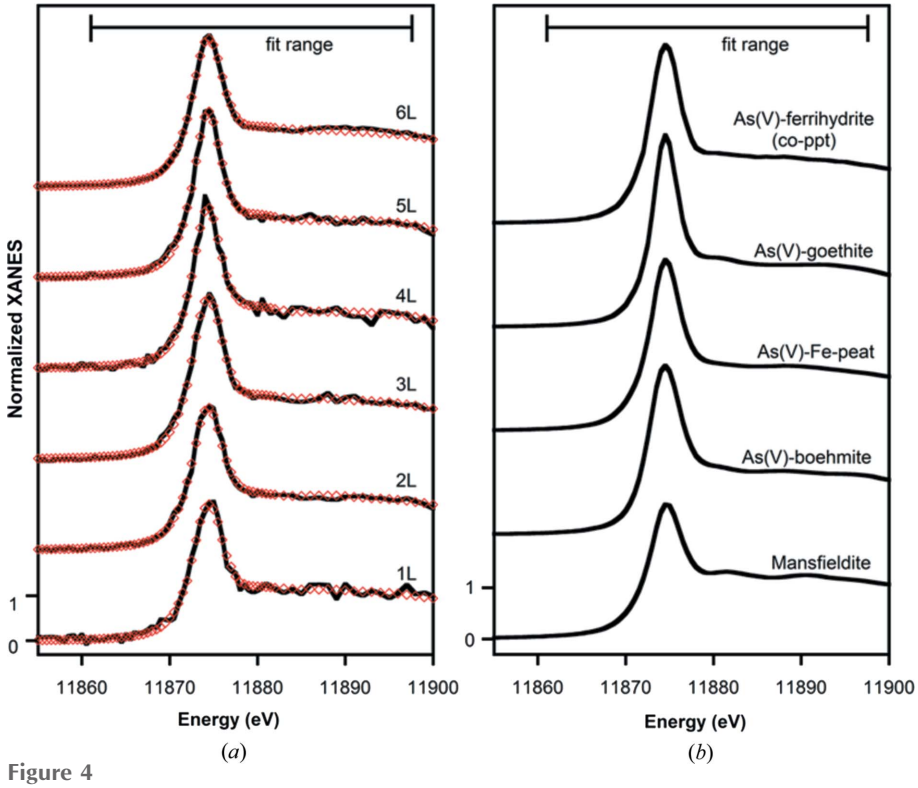


Figure 4
(*a*) Stacked As *K*-edge μ-XANES spectra collected on the XFM beamline from individual As-enriched spots across a 200 μm × 200 μm region (ROI-200, see Fig. S2) of the sand grain, along with overlaid fits (points) as given in Table 4; and (*b*) arsenic standards that were included in best fits to sample spectra from different spots.

confidence that the LCF fitting results reported are the most suitable models based on the standards used and indicate the existence of variability in As speciation across the imaged region.

Results of LCF analysis of 25 As μ-XANES spectra from within ROI-10b, which were constrained to two standards, showed means of 67 ± 14% [coefficient of variation (CV) = 21%] of an As(V)-ferrihydrite (co-ppt) standard and 33 ± 14% (CV = 42%) of a standard of As(V) adsorbed on boehmite (Table S2). Images of the proportions of these two standards in LCF results for microsites across ROI-10b (Fig. 6) showed the most similar spatial patterns between the proportions of the As(V)-ferrihydrite (co-ppt) standard and μ-XRF signals for Ca and Zn. However, we found no significant partial correlation between the fitted proportions of either of the two standards included in μ-XANES fits and μ-XRF signals from Fe or other elements imaged within ROI-10b. Also, a non-spatial regression model that included μ-XRF signals of Ca as the significant predictor yielded overall poor predictions (R^2 = 0.26) of fitting results (data not shown). These results indicate that speciation constrained to the two selected standards was not tied to any particular element in ROI-10b.

4. Discussion

4.1. Discriminating multiple-element contributions to arsenic accumulation

Our approach of using synchrotron μ-XRF to image the same area of a soil sand-grain coating before and after treatment with As(V) solution provides insights into short-term reactions that can occur in soils. Arsenic species formed in the short term can transform into other, presumably more stable, species over time (Arai *et al.*, 2006; Rahman *et al.*, 2017). Partial-correlation and spatial-regression analyses of μ-XRF image data effectively decoupled correlations between co-localized matrix elements within soil microsites, and thereby provided a more robust statistical analysis showing that Fe along with Ca, Ti and Zn to a lesser extent were most highly correlated with As accumulation, depending on the sample area analyzed [Table 1; equations (3) and (4), and (9) and (10)]. In light of model-system studies that showed high-affinity, inner-sphere surface complexation of arsenate by Fe in Fe(III) (hydr)oxides (Di Iorio *et al.*, 2018; Dzade & De Leeuw, 2018; Fendorf *et al.*, 1997; Manceau, 1995; Waychunas *et al.*, 1995), a strong As–Fe correlation for our sand-grain coating infers (but does not definitively show) that arsenate binding by Fe (hydr)oxides is a likely mechanism of As retention. This inference is corroborated by μ-XANES LCF results showing ≥50% of Fe-bonded standards contributing to fits in 5 of 6 spots analyzed across ROI-100 (Table 4) and 21 of 25 spots within ROI-10b (Table S2). However, based on a significant (r = 0.51) spatial correlation between Al and Fe from TOF-SIMS analysis (Fig. 1) and μ-XANES results showing that Al-bonded standards accounted for the remaining 31–64% of fits

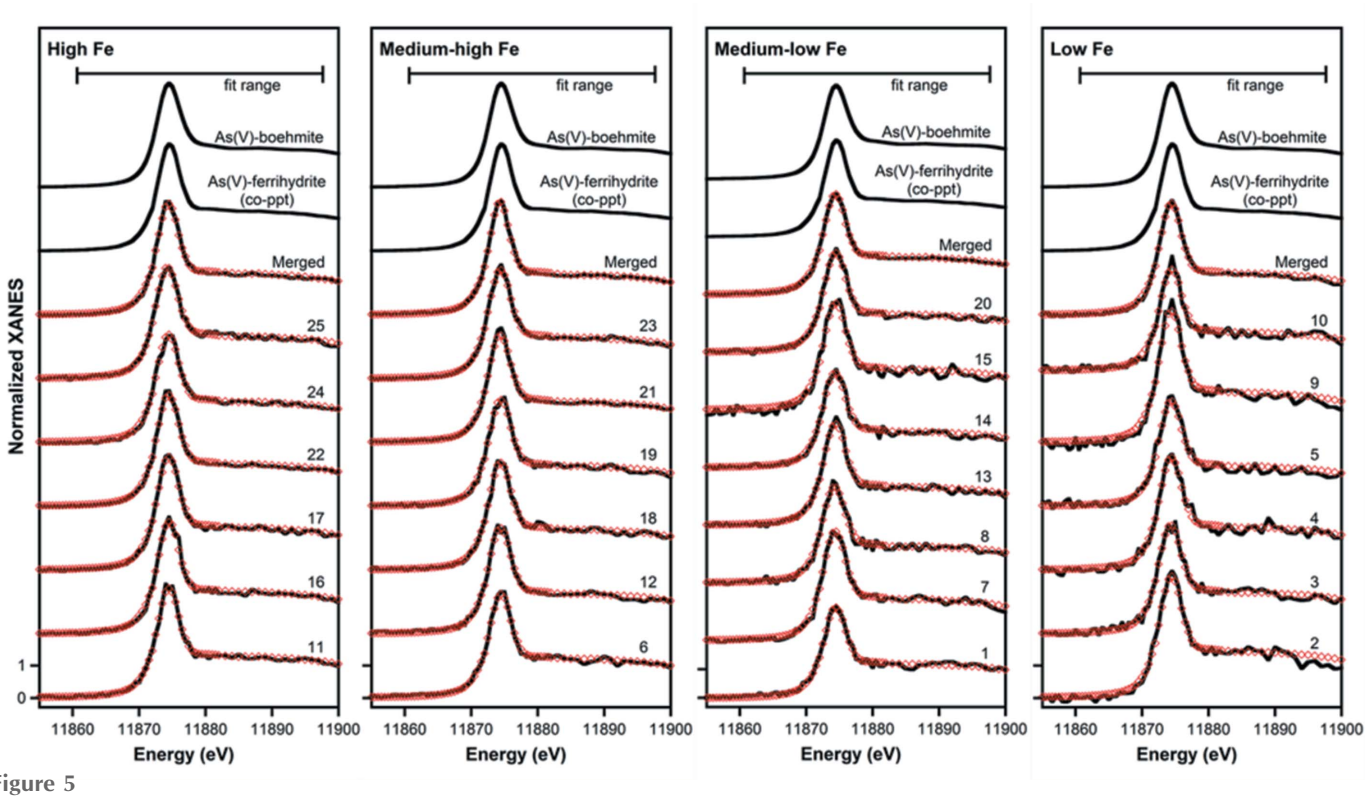


Figure 5
Twenty-five As *K*-edge μ-XANES spectra collected on the XFM beamline from ROI-10b and merged spectra from each quartile along with overlaid fits (points) as given in Table S2, and spectra of As standards included in the final linear-combination fits. Stacked spectra are plotted in four groups categorized according to the quartiles of Fe fluorescence signal.

(Tables 4 and S4), we deduced that Al (hydr)oxides also contributed substantially to As(V) accumulation in soil microsites. The significance of Al relative to Ca, Ti and Zn in augmenting As accumulation could not be determined directly from our statistical analyses because we could not detect Al signals in our μ-XRF imaging. Although Ca was also a significant predictor of accumulated As, our single Ca-arsenate standard [$Ca_3(AsO_4)_2$; Table S1] was not identified in μ-XANES fits for any of the microsites analyzed (Tables 4 and S4).

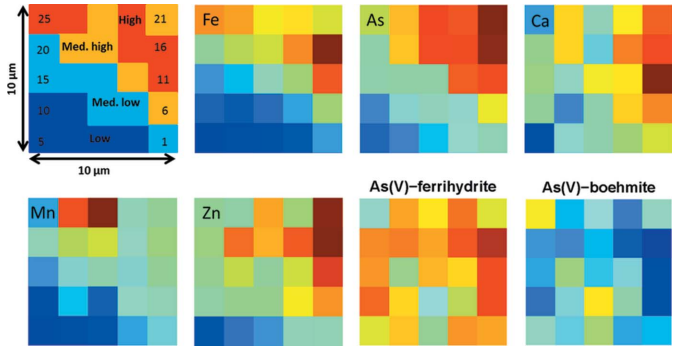


Figure 6
Spatial patterns of selected soil matrix elements and proportions of two standards – As(V) co-precipitated with ferrihydrite and As(V) adsorbed on boehmite – derived from linear combination fits to As μ-XANES spectra collected from each voxel in ROI-10b of the As(V) treated sand grain. The top left image shows the voxels belonging to high, medium-high, medium-low and low quartiles of Fe fluorescence signals. Brighter colors in the elemental images indicate greater I_0-normalized fluorescence intensities, with maximum intensities (arbitrary units) of 0.00882 (Fe), 0.00057 (As), 0.00018 (Ca), 0.00038 (Mn) and 0.00026 (Zn). Brighter colors for linear combination fits indicate greater proportions of each given standard, ranging from 5% to 95%.

Overall, our collective results suggest that up to five elements (Fe, Al, Ca, Ti and Zn) augmented As accumulation in the sand-grain coating analyzed. It is easy to infer from model-system research that Fe and Al (hydroxides) in the coatings adsorbed arsenate (Adra *et al.*, 2016; Goldberg & Johnston, 2001; Manning & Goldberg, 1996; Silva *et al.*, 2010). A significant ($\alpha = 0.05$) pairwise Pearson correlation between Fe and Al ($r = 0.51$) from TOF-SIMS analysis indicates that Fe and Al (hydr)oxides were at least partially co-localized within microsites. However, we cannot determine at the spatial scale of our μ-XRF data (~1 μm × 1 μm microsites of variable thickness) whether Fe- and Al-bearing solids were in distinctly separated phases within microsites versus associated into complex, submicrometre-sized assemblages or if Al was co-precipitated by atomic scale isomorphic substitution into Fe (hydr)oxide phases (Schwertmann & Taylor, 1989). Moreover, explaining possible mechanisms by which Ca, Ti or Zn augmented As accumulation would require analyses that are more targeted to these elements, *e.g.* μ-XAS analyses at the Ca, Ti and Zn *K*-edges. The partial correlation and multivariate spatial-regression modeling approach presented here would help to focus such studies. Finally, it is important to recognize that elements such as organic carbon that are not

detected by a given analysis might also affect As accumulation, as we deduced for Al.

4.2. Using chemical-based imaging to enhance μ-XANES analysis

Within soils, more than 95% of strongly bound trace elements typically occur in solid phases. However, minor changes in aqueous concentrations can have a disproportionate effect on element toxicity (Hesterberg, 1998; McBride, 1989). Consequently, soil regions of lower trace-element content, *i.e.* 'coldspots', could potentially be important in terms of their contribution to element mobility. For example, in a water-saturated soil of 50% porosity with an average particle density of 2.7 g cm^{-3} and containing As at the world median content of 6 mg kg^{-1} solids (Sparks, 2003), mobilization of only 0.6% of the total soil As would contribute 100 μg As l^{-1} to the soil solution. This level is one order of magnitude greater than the World Health Organization (WHO) provisional guideline value of 10 μg l^{-1} of As in drinking water (https://www.who.Int/water_sanitation_health/publications/arsenic/en/).

One challenge in μ-XANES analyses of soil trace elements is to obtain high-quality spectra from coldspots of lower-element content. Consequently, selection of sample points (microsites) for analyses are often biased toward more concentrated 'hotspots' to obtain higher-quality (lower-noise) spectra. It is difficult to know whether trace-element speciation analyses from localized hotspots are representative of the overall speciation in a geochemical matrix (Toner *et al.*, 2014), especially if contributions of species in coldspots collectively contribute a significant amount of mobilizable trace elements. Nicholas *et al.* (2017) addressed this issue by first performing LCF analysis of μ-XANES spectra from As hotspots in sediments to identify energies of unique features corresponding to different As species [As-sulfides versus As(III) versus As(V) species]. Then they performed speciation mapping at these energies across hundreds of thousands of image pixels in surrounding regions of lower As content, and applied correlation-distance hierarchical clustering to assess similarities of spectral features across these sample regions. Their results showed that As speciation was spatially correlated, but also varied with the localized geochemistry of sediment microsites, including Fe speciation and relative proportions of As, Fe and S.

In contrast to the analyses caried out by Nicholas *et al.* (2017), whereby different As species produced unique μ-XANES spectral features that were separated in energy by up to 7 eV, our data show a dominance of As(V) species with only minor differences in spectral features (Table 4; Fig. S4). To characterize As speciation in coldspots under these constraints, we proposed, based on insights from spatial statistical modeling, to first merge μ-XANES spectra from soil microsites of similar chemical composition to obtain higher-quality spectra for selecting appropriate fitting standards. Fitting of spectra from individual pixels across the imaged region, including noisy spectra from cold spots, can then be constrained to the selected standards. The underlying assumption of this approach is that arsenate reacted with geochemical microsites of similar composition will have similar speciation (Nicholas *et al.*, 2017; Toner *et al.*, 2014). Because our spatial statistical analyses showed that Fe was the strongest predictor of As accumulation (Table 1, Fig. 3), we applied this approach by grouping the 25 μ-XANES spectra from ROI-10b according to Fe μ-XRF signal intensities (Figs. 5 and 6). We hypothesized that the proportion of Fe-associated As species would change systematically with Fe content of microsites, which is illustrated by the region of high As–Zn fluorescence-signal correlation for moderate Fe signals in Fig. S7. Because all of our spectra were confined within a 10 μm × 10 μm sample area of variable chemical composition, we expected that chemical composition rather than spatial proximity would have a stronger effect on As speciation. However, a lack of significant partial correlation between the proportion of As(V)-ferrihydrite and Fe fluorescence intensity ($r' = 0.22$; $p = 0.43$), in addition to visible differences in spectra within Fe quartiles (*e.g.* high-Fe spectra in Fig. 5), suggested that microsite attributes other than Fe content affected the proportion of Fe associated As species. Microsite Al content is one potential complicating attribute, given that As(V)-boehmite was the complementary standard fit to this dataset (Fig. 6). Other possible factors include variations in Fe and Al speciation, and effects of the other elements; however, we found no significant partial correlation between speciation by μ-XANES and Ca, Ti or Zn. Although our data from 25 contiguous microsites did not support the hypothesis that As speciation varies systematically with Fe content alone, we convey this approach to illustrate that subsets of larger datasets of μ-XANES spectra can be merged based on parameters other than (or in addition to) spatial proximity (Muyskens, 2019) to improve spectral quality. We also contend that chemical-compositional grouping of geochemical microsites guided by partial correlation or multivariate spatial statistical analyses is worthy of further evaluation to improve trace-element speciation in geochemical microsites of low contents. Ideally, these analyses would be done on larger spatial datasets than we collected here, which should allow parsing based on microsite composition of multiple elements (Fe, Al, Ca, *etc.*), and also include chemical speciation analysis of matrix elements within microsites.

5. Conclusions

Synchrotron μ-XRF and μ-XANES analysis of a naturally coated sand grain from a non-contaminated soil provided a means to assess the chemical reactivity of As(V) in a soil matrix by imaging matrix composition and As accumulation from the same area before and after chemical treatment with aqueous As(V). Partial correlation and multivariate spatial regression analyses can serve as powerful data-analysis tools to separate effects of matrix-element co-localization that confound interpretation of simple Pearson correlation analyses in terms of As binding mechanisms determined on model systems. Including a spatial component in regression

modeling to remove element autocorrelations provides a statistically robust means to identify matrix elements that were significant predictors of accumulated As, which helps constrain the number of chemical elements used as predictors in non-spatial regression modeling. However, the resulting models also depend on the size of the image analyzed. For the sand-grain coating analyzed here, (spatial) statistical modeling of complementary data from μ-XRF, μ-XANES and TOF-SIMS analyses suggested that Fe, Al, Ca and Ti (and possibly Zn) all augmented accumulation of applied As(V) to varying degrees. Including these analytical and statistical approaches within a broader suite of complementary microscale and macroscale analyses should further enhance our ability to determine trace-element binding mechanisms in soils and other complex, multi-component geochemical matrices.

6. Related literature

The following reference, not cited in the main body of the paper, has been cited in the supporting information: GitHub (2018).

Acknowledgements

The authors are grateful to Dr Chuanzhen Elaine Zhou for her assistance in collecting TOF-SIMS data, to Dr Elton Alves (former NC State visiting scholar from Federal University of Viçosa, Brazil) for assistance in processing μ-XRF detector data, to Joshua Sanchez and Dr Donald Sparks (University of Delaware) for providing As(V) standards, and to Dr Keith Jones (Brookhaven National Laboratory, now deceased) for inputs on data collection.

Funding information

The following funding is acknowledged: National Science Foundation (grant No. EAR-1349374). This research used the SRX (5-ID) and XFM (4-BM) beamlines of the National Synchrotron Light Source II, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory (contract No. DE-SC0012704). This work was performed in part at the Analytical Instrumentation Facility (AIF) at NC State University, which is supported by the State of North Carolina and the NSF (award No. ECCS-1542015). The AIF is a member of the North Carolina Research Triangle Nanotechnology Network (RTNN), a site in the National Nanotechnology Coordinated Infrastructure (NNCI).

References

Adra, A., Morin, G., Ona-Nguema, G. & Brest, J. (2016). *Appl. Geochem.* **64**, 2–9.

Adra, A., Morin, G., Ona-Nguema, G., Menguy, N., Maillot, F., Casiot, C., Bruneel, O., Lebrun, S., Juillot, F. & Brest, J. (2013). *Environ. Sci. Technol.* **47**, 12784–12792.

Alcacio, T. E., Hesterberg, D., Chou, J. W., Martin, J. D., Beauchemin, S. & Sayers, D. E. (2001). *Geochim. Cosmochim. Acta*, **65**, 1355–1366.

Anderson, M. A., Ferguson, J. F. & Gavis, J. (1976). *J. Colloid Interface Sci.* **54**, 391–399.

Antelo, J., Arce, F. & Fiol, S. (2015). *Chem. Geol.* **410**, 53–62.

Arai, Y., Elzinga, E. J. & Sparks, D. L. (2001). *J. Colloid Interface Sci.* **235**, 80–88.

Arai, Y., Lanzirotti, A., Sutton, S., Newville, M., Dyer, J. & Sparks, D. (2006). *Environ. Sci. Technol.* **40**, 673–679.

Basta, N., Ryan, J. & Chaney, R. (2005). *J. Environ. Qual.* **34**, 49–63.

Beale, C. M., Lennon, J. J., Yearsley, J. M., Brewer, M. J. & Elston, D. A. (2010). *Ecol. Lett.* **13**, 246–264.

Borch, T., Kretzschmar, R., Kappler, A., Cappellen, P. V., Ginder-Vogel, M., Voegelin, A. & Campbell, K. (2010). *Environ. Sci. Technol.* **44**, 15–23.

Brown, G. E. Jr & Sturchio, N. C. (2002). *Rev. Mineral. Geochem.* **49**, 1–115.

Burton, E. D., Johnston, S. G. & Kocar, B. D. (2014). *Environ. Sci. Technol.* **48**, 13660–13667.

Chen, C. & Sparks, D. L. (2018). *ACS Earth Space Chem.* **2**, 1095–1101.

Chen-Wiegart, Y., Williams, G., Zhao, C., Jiang, H., Li, L., Demkowicz, M., Seita, M., Short, M., Ferry, S., Wada, T., Kato, H., Chou, K. W., Petrash, S., Catalano, J., Yao, Y., Murphy, A., Zumbulyadis, N., Centeno, S. A., Dybowski, C. & Thieme, J. (2016). *AIP Conf. Proc.* **1764**, 030004.

Cullen, W. & Reimer, K. (1989). *Chem. Rev.* **89**, 713–764.

Davis, J. A. (1984). *Geochim. Cosmochim. Acta*, **48**, 679–691.

Di Iorio, E., Cho, H. G., Liu, Y., Cheng, Z., Angelico, R. & Colombo, C. (2018). *Geochim. Cosmochim. Acta*, **237**, 155–170.

Dixit, S. & Hering, J. G. (2003). *Environ. Sci. Technol.* **37**, 4182–4189.

Du, H., Lin, Y., Chen, W., Cai, P., Rong, X., Shi, Z. & Huang, Q. (2017). *Eur. J. Soil Sci.* **68**, 514–523.

Dzade, N. Y. & De Leeuw, N. H. (2018). *Environ. Sci. Processes Impacts*, **20**, 977–987.

Ehlert, K., Mikutta, C., Jin, Y. & Kretzschmar, R. (2018). *Environ. Sci. Technol.* **52**, 616–627.

Fan, J.-X., Wang, Y.-J., Liu, C., Wang, L.-H., Yang, K., Zhou, D.-M., Li, W. & Sparks, D. L. (2014). *J. Hazard. Mater.* **279**, 212–219.

Fendorf, S., Eick, M. J., Grossl, P. & Sparks, D. L. (1997). *Environ. Sci. Technol.* **31**, 315–320.

Fendorf, S. & Kocar, B. D. (2009). *Adv. Agron.* **104**, 137–164.

Fendorf, S., Nico, P. S., Kocar, B. D., Masue, Y. & Tufano, K. J. (2010). *Developments in Soil Science*, Vol. 34, pp. 357–378. Amsterdam: Elsevier.

Foster, A. L., Brown, G. E. Jr & Parks, G. A. (2003). *Geochim. Cosmochim. Acta*, **67**, 1937–1953.

Foster, A. L. & Kim, C. S. (2014). *Rev. Mineral. Geochem.* **79**, 257–369.

Fritzsche, A., Rennert, T. & Totsche, K. U. (2011). *Environ. Pollut.* **159**, 1398–1405.

Gadd, G. M. (2010). *Microbiology*, **156**, 609–643.

Gamble, A. V., Givens, A. K. & Sparks, D. L. (2018). *J. Environ. Qual.* **47**, 121–128.

Gillispie, E. C., Andujar, E. & Polizzotto, M. L. (2016). *Environ. Sci. Processes Impacts*, **18**, 1090–1103.

GitHub, I. (2018). *Larch: X-ray Analysis for Synchrotron Applications Using Python*, https://github.com/xraypy/xraylarch.

Goldberg, S. & Johnston, C. T. (2001). *J. Colloid Interface Sci.* **234**, 204–216.

Grafe, M., Eick, M. & Grossl, P. (2001). *Soil Sci. Soc. Am. J.* **65**, 1680–1687.

Gräfe, M., Tappero, R. V., Marcus, M. A. & Sparks, D. L. (2008). *J. Colloid Interface Sci.* **321**, 1–20.

Guinness, J. (2018). *Technometrics*, **60**, 415–429.

Guinness, J., Fuentes, M., Hesterberg, D. & Polizzotto, M. (2014). *Spat. Stat.* **9**, 93–108.

Hesterberg, D. (1998). *Agric. Ecosyst. Environ.* **67**, 121–133.

Hesterberg, D., Duff, M. C., Dixon, J. B. & Vepraskas, M. J. (2011). *J. Environ. Qual.* **40**, 667–678.

Hoeting, J. A., Davis, R. A., Merton, A. A. & Thompson, S. E. (2006). *Ecol. Appl.* **16**, 87–98.
Hu, S., Lu, Y., Peng, L., Wang, P., Zhu, M., Dohnalkova, A. C., Chen, H., Lin, Z., Dang, Z. & Shi, Z. (2018). *Environ. Sci. Technol.* **52**, 11632–11641.
Inskeep, W. P., McDermott, T. R. & Fendorf, S. (2001). *Environmental Chemistry of Arsenic*, edited by W. T. Frankenberger, pp. 183–215. New York: Marcel Dekker.
Jiang, J., Bauer, I., Paul, A. & Kappler, A. (2009). *Environ. Sci. Technol.* **43**, 3639–3645.
Kelly, S., Hesterberg, D. & Ravel, B. (2008). *Methods of Soil Analysis*, Vol. 5, pp. 387–464. Madison: Soil Science Society of America.
Khare, N., Hesterberg, D. & Martin, J. D. (2005). *Environ. Sci. Technol.* **39**, 2152–2160.
Kim, C., Chi, C., Miller, S., Rosales, R., Sugihara, E., Akau, J., Rytuba, J. & Webb, S. (2013). *Environ. Sci. Technol.* **47**, 8164–8171.
Kizewski, F., Liu, Y.-T., Morris, A. & Hesterberg, D. (2011). *J. Environ. Qual.* **40**, 751–766.
Kopittke, P. M., Wang, P., Lombi, E. & Donner, E. (2017). *J. Environ. Qual.* **46**, 1175–1189.
Ladeira, A., Ciminelli, V., Duarte, H., Alves, M. & Ramos, A. (2001). *Geochim. Cosmochim. Acta*, **65**, 1211–1217.
Landrot, G., Tappero, R., Webb, S. M. & Sparks, D. L. (2012). *Chemosphere*, **88**, 1196–1201.
Langner, P., Mikutta, C., Suess, E., Marcus, M. A. & Kretzschmar, R. (2013). *Environ. Sci. Technol.* **47**, 9706–9714.
LeMonte, J. J., Stuckey, J. W., Sanchez, J. Z., Tappero, R., Rinklebe, J. & Sparks, D. L. (2017). *Environ. Sci. Technol.* **51**, 5913–5922.
Li, L., Yan, H., Xu, W., Yu, D., Heroux, A., Lee, W.-K., Campbell, S. I. & Chu, Y. S. (2017). *Proc. SPIE*, **10389**, 103890U.
Liu, Y.-T. & Hesterberg, D. (2011). *Environ. Sci. Technol.* **45**, 6283–6289.
Lopez, A. R., Silva, S. C., Webb, S. M., Hesterberg, D. & Buchwalter, D. B. (2018). *Environ. Toxicol. Chem.* **37**, 903–913.
McBride, M. (1989). *Adv. Soil Sci.* **10**, 1–56.
Manceau, A. (1995). *Geochim. Cosmochim. Acta*, **59**, 3647–3653.
Manceau, A., Marcus, M. A. & Grangeon, S. (2012). *Am. Mineral.* **97**, 816–827.
Manceau, A., Tamura, N., Marcus, M. A., MacDowell, A. A., Celestre, R. S., Sublett, R. E., Sposito, G. & Padmore, H. A. (2002). *Am. Mineral.* **87**, 1494–1499.
Manning, B. A., Fendorf, S. E., Bostick, B. & Suarez, D. L. (2002). *Environ. Sci. Technol.* **36**, 976–981.
Manning, B. A. & Goldberg, S. (1996). *Soil Sci. Soc. Am. J.* **60**, 121–131.
Masue, Y., Loeppert, R. H. & Kramer, T. A. (2007). *Environ. Sci. Technol.* **41**, 837–842.
Mikkonen, H. G., van de Graaff, R., Collins, R. N., Dasika, R., Wallis, C. J., Howard, D. L. & Reichman, S. M. (2019). *Sci. Total Environ.* **654**, 1072–1081.
Mikutta, C. & Kretzschmar, R. (2011). *Environ. Sci. Technol.* **45**, 9550–9557.
Minasny, B. & McBratney, A. B. (2005). *Geoderma*, **128**, 192–207.
Minasny, B. & McBratney, A. B. (2007). *Geoderma*, **140**, 324–336.
Murray, G. C. & Hesterberg, D. (2006). *Soil Sci. Soc. Am. J.* **70**, 1318–1327.
Muyskens, A. L. (2019). PhD Thesis. NC State University, USA. https://repository.lib.ncsu.edu/handle/1840.20/36314.
NCSU-AIF (2018). *Tof-Sims-General-Experiment-Description*, https://www.aif.ncsu.edu/wp-content/uploads/sites/25/2013/05/TOF-SIMS-general-experiment-description.pdf.
Nicholas, S. L., Erickson, M. L., Woodruff, L. G., Knaeble, A. R., Marcus, M. A., Lynch, J. K. & Toner, B. M. (2017). *Geochim. Cosmochim. Acta*, **211**, 228–255.
Otero-Fariña, A., Fiol, S., Arce, F. & Antelo, J. (2017). *Chem. Geol.* **459**, 119–128.
Park, J. H., Han, Y.-S. & Ahn, J. S. (2016). *Water Res.* **106**, 295–303.
Rahman, M., Clark, M. W., Yee, L. H., Comarmond, M., Payne, T. E., Kappen, P. & Mokhber-Shahin, L. (2017). *Chemosphere*, **168**, 1324–1336.
Ravel, B. & Newville, M. (2005). *J. Synchrotron Rad.* **12**, 537–541.
Raven, K. P., Jain, A. & Loeppert, R. H. (1998). *Environ. Sci. Technol.* **32**, 344–349.
Schwer, D. R. & McNear, D. H. (2011). *J. Environ. Qual.* **40**, 1172–1181.
Schwertmann, U. & Taylor, R. M. (1989). *Minerals in Soil Environments*, pp. 379–438. Madison: Soil Science Society of America.
Serrano, S., Gomez-Gonzalez, M. A., O'Day, P. A., Laborda, F., Bolea, E. & Garrido, F. (2015). *J. Hazard. Mater.* **286**, 30–40.
Silva, J., Mello, J. W., Gasparon, M., Abrahão, W. A., Ciminelli, V. S. & Jong, T. (2010). *Water Res.* **44**, 5684–5692.
Snipes, M. & Taylor, D. C. (2014). *Wine Econ. Policy*, **3**, 3–9.
Sowers, T. D., Stuckey, J. W. & Sparks, D. L. (2018). *Geochem. Trans.* **19**, 4.
Sparks, D. L. (2003). *Environmental Soil Chemistry*, 2nd ed. San Diego, London: Elsevier.
Strawn, D., Doner, H., Zavarin, M. & McHugo, S. (2002). *Geoderma*, **108**, 237–257.
Terres, M. A., Fuentes, M., Hesterberg, D. & Polizzotto, M. (2018). *Bayesian Anal.* **13**, 1–28.
Toner, B. M., Nicholas, S. L. & Wasik, J. K. C. (2014). *Environ. Chem.* **11**, 4–9.
Violante, A. & Pigna, M. (2002). *Soil Sci. Soc. Am. J.* **66**, 1788–1796.
Voegelin, A., Weber, F.-A. & Kretzschmar, R. (2007). *Geochim. Cosmochim. Acta*, **71**, 5804–5820.
Waychunas, G. A., Davis, J. A. & Fuller, C. C. (1995). *Geochim. Cosmochim. Acta*, **59**, 3655–3661.
Werner, F., Mueller, C. W., Thieme, J., Gianoncelli, A., Rivard, C., Höschen, C. & Prietzel, J. (2017). *Sci. Rep.* **7**, 3203.
Wovkulich, K., Mailloux, B. J., Bostick, B. C., Dong, H., Bishop, M. E. & Chillrud, S. N. (2012). *Geochim. Cosmochim. Acta*, **91**, 254–270.
Zeng, H., Fisher, B. & Giammar, D. E. (2008). *Environ. Sci. Technol.* **42**, 147–152.
Zhu, Y. & Elzinga, E. J. (2015). *Environ. Sci. Technol.* **49**, 13369–13377.