



Cite this: *J. Anal. At. Spectrom.*, 2020, **35**, 567

Synchrotron hard X-ray chemical imaging of trace element speciation in heterogeneous samples: development of criteria for uncertainty analysis†

Jonas Wielinski, ^{ab} Francesco Femi Marafatto, ^{ac} Alexander Gogos, ^{ad} Andreas Scheidegger, ^a Andreas Voegelin, ^{*a} Christoph R. Müller, ^e Eberhard Morgenroth ^{ab} and Ralf Kaegi ^a

Synchrotron hard X-ray spectroscopy with focussing optics allows recording X-ray fluorescence (XRF) maps at energies around element specific X-ray absorption edges. Stacking multiple XRF maps along the energy axis yields chemical images that contain spatially resolved information on the speciation of the absorber in the sample matrix at the micrometre scale. Short dwell times needed to keep measurement time and radiation-induced sample changes within acceptable limits result in spectral noise that affects the uncertainty in data analysis. In this study, we develop a model to quantify the uncertainty associated with the processing of XRF image stacks using Bayesian inference. To demonstrate the potential of our approach, the model is applied to stacks of XRF maps collected around the copper (Cu) K-edge (pixel size: $3 \times 3 \mu\text{m}^2$, map sizes: $500 \times 500 \mu\text{m}^2$). The investigated samples include digested sewage sludge spiked with either CuO nanoparticles (NP) or dissolved CuSO_4 and their corresponding ashes obtained through incineration. The chemical imaging data reveal differences in species distribution between sludge spiked with CuO NP or dissolved Cu. These differences disappear during the incineration process and the resulting ashes exhibit almost identical Cu species distribution. The uncertainty analysis approach developed in this study can be used for data interpretation, but can also be used for the planning of chemical imaging experiments at synchrotron beamlines.

Received 19th November 2019
Accepted 27th January 2020

DOI: 10.1039/c9ja00394k

rsc.li/jaas

1 Introduction

X-ray absorption spectroscopy (XAS) is widely used to investigate the speciation of major and trace elements in a wide range of sample matrices, including complex environmental samples.^{1–4} With XAS, the average speciation of selected elements in homogenised samples can be evaluated using wide spread X-ray beams with lateral and horizontal extensions of hundreds of micrometres to a few millimetres.⁴ Over the last two decades, advanced synchrotron light sources and beamlines providing higher photon fluxes and the ability to focus the beam to small sizes enabled investigations of chemical (speciation) heterogeneities at the micrometre scale.⁵ Based on elemental distribution maps (micro X-ray fluorescence (XRF) spectroscopy) that provide information

on the elemental distributions and their correlations, points of interest (POI) can be selected to investigate the speciation of individual elements by micro-focused X-ray absorption near-edge structure (XANES) or extended X-ray absorption fine structure (EXAFS) spectroscopy, *e.g.*^{6–9} The further development of dedicated beamlines capable of focussing hard ($>4.5 \text{ keV}$) X-rays¹⁰ and the advancement of detection systems¹¹ meanwhile allow to perform chemical imaging analyses. In this approach, XRF maps of the same area are recorded at multiple energies around an absorption edge to obtain spatially resolved speciation information at the micrometre scale.^{12,13} In micro-focused X-ray experiments for chemical imaging (sometimes also called ‘Chemical-state maps’, *e.g.*¹³), a large fraction of the beam is focused onto a micrometre sized spot on the sample using slits and/or Kirkpatrick-Baez (KB) mirrors.^{5,10} High photon flux densities allow users to obtain chemical images of absorbers at low concentrations in complex matrices such as environmental samples within a few hours.^{1,4,10,12} However, the complex and heterogeneous matrices and the redox sensitivity of certain absorber atoms may lead to radiation-induced speciation changes in the sample commonly referred to as beam damage, which has previously been reported for, *e.g.*, Cu.^{14,15} To limit beam-induced transformation of the sample/absorber, the exposure time to the beam can be reduced and the sample can be cooled to cryogenic

^aEawag, Swiss Federal Institute of Aquatic Science and Technology, Überlandstrasse 133, 8600 Dübendorf, Switzerland. E-mail: jonas.wielinski@eawag.ch; andreas.voegelin@eawag.ch; Tel: +41 58 765 5336; +41 58 765 5470

^bETH Zürich, Institute of Environmental Engineering, 8093 Zürich, Switzerland

^cSwiss Light Source, Paul-Scherrer-Institute (PSI), 5232 Villigen, Switzerland

^dEMPA, Swiss Federal Laboratories for Materials Science and Technology, 9014 St. Gallen, Switzerland

^eETH Zürich, Institute of Energy Technology, 8092 Zürich, Switzerland

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9ja00394k



temperatures. Further, the exposure of the sample to the X-ray beam can be reduced by collecting data only at selected energies of interest containing diagnostic spectral features instead of scanning over the whole absorption edge at high energy resolution. The most relevant energies of interest are commonly identified based on the XAS data of reference materials.¹⁶

The new possibilities offered by the chemical imaging attracted increased attention in environmental sciences^{12,13,17–21} where elements of interest often occur at low concentrations (e.g., Cu in digested sewage sludge 480–700 mg kg^{−1}) and unevenly distributed in the sample matrix.¹² Because of the potential beam damage induced by high photon flux densities, only a limited number (typically ranging from 3–10)^{13,17} of XRF maps around an X-ray absorption edge of a specific element of interest may be recorded to derive the chemical speciation of the absorber atom at the micrometre scale.²² Nevertheless, caused by the short dwell times (few milliseconds) and low element concentrations, data acquired with even the most advanced detector systems²³ contains increased amounts of noise compared to spectra recorded on bulk samples using longer dwell times (up to seconds).¹² Therefore, speciation information may be limited to the oxidation state of the absorber, or to major species classes rather than individual chemical forms. XAS and chemical imaging data processing, including the determination of the appropriate number of spectral components using principle component analysis, have considerably improved over the last decades.^{24–26} Furthermore, strategies and guidelines to reduce uncertainty by optimizing measurement conditions are available in the literature.^{27,28} Studies investigating the uncertainty associated with species information extracted from hard X-ray of chemical image stacks are still lacking, however. In this study, we therefore developed a model/algorithm to quantitatively assess the impact of uncertainty on the interpretability of chemical images derived from spatially resolved XRF maps. The usefulness of the model/algorithm was demonstrated based on chemical images recorded on digested sewage sludge that had been spiked with copper oxide nanoparticles (CuO-NP) or dissolved Cu²⁺ (CuSO₄) as well as on their corresponding ashes. These samples that have previously been examined by bulk XAS, but information on the spatial distribution of Cu and of individual Cu species in the sludge and corresponding ashes may be relevant as well for the risk and life cycle assessment of engineered nanomaterials (details in the ESI Section S1†).

2 Materials and methods

2.1 Sample preparation and characterization of experimental samples

Two digested sludge samples were spiked with either CuO-NP (SLG NP) or dissolved CuSO₄ (SLG AQ) and kept under anaerobic conditions for 24 h to allow for a sulfidation of the Cu. The sludge was dewatered, dried at 105 °C and incinerated in a pilot scale bubbling bed type fluidized bed reactor resulting in the ashes ASH NP from SLG NP and ASH AQ from SLG AQ. A detailed description of the samples and their generation can be found in Wielinski *et al.*²⁹ The samples SLG NP, SLG AQ, ASH NP

and ASH AQ correspond to D-NP, D-AQ, D-NP-ap and D-AQ-ap, in Wielinski *et al.*²⁹ The sludge and ash samples (ASH NP and ASH AQ) were dried, embedded into an epoxy resin (one part EpoFix Hardener and seven parts EpoFix Resin, both Struers) at 180 mbar (Citovac, Struers), allowed to harden for 24 h, cut into thin sections ($d = 30 \mu\text{m}$) and mounted on $1 \times 1 \text{ cm}^2$ pre-cut 250 μm -thick Si Wafers (TED PELLA, Inc.). Bulk Cu concentrations were determined using inductively coupled plasma mass spectrometry (ICP-MS, 7500cx, Agilent Technologies, Inc.) after acid digestion of 10 to 20 mg of sample (2 mL H₂O₂ and 9 mL aqua regia for the sludge samples or 9 mL HNO₃ and 200 μL HF for the ash samples) in a microwave system (ETHOS 1, MLS GmbH) for the sludge samples and in an ultraclave (MLS GmbH) for the ash samples. The visibly clear digests were diluted to 50 mL in DI water. The limit of quantification (LOQ) in the digests was 0.02 ppb, resulting in a LOQ of 0.05 mg Cu per kg sample. Chemicals for the acid digestions (37% HCl and 40% HF, both Merck, 69% HNO₃, Roth and 30% H₂O₂, Sigma-Aldrich) were used as received.

2.2 Synchrotron experiments

Synchrotron X-ray experiments were conducted at the X05LA beamline ('microXAS') at the Swiss Light Source (SLS) in Villigen, Switzerland.¹⁰ A double crystal monochromator (Si(111)) was used to select the energy of the X-ray beam produced by a minigap in-vacuum undulator. KB mirrors focused the beam to a $3 \times 3 \mu\text{m}^2$ spot on the sample. The flux at the beamline is around 2×10^{12} photons per s at a beam current of 400 mA. A 16-element (2048 channel) Silicon Drift detector (Ketek GmbH) was used to record the XRF signals. A total of seven Cu-K _{α} XRF maps recorded around the Cu K-edge were collected in 'on-the-fly' mode for each sample (area: $500 \times 500 \mu\text{m}^2$, resolution: $3 \times 3 \mu\text{m}^2$ for the samples SLG NP, SLG AQ and ASH AQ and of $350 \times 350 \mu\text{m}^2$ at the same pixel resolution for ASH NP). An additional Ti-K _{α} XRF map was recorded from the same areas for aligning the individual XRF maps. In the on-the-fly acquisition mode, the sample stage was moved in the horizontal direction perpendicular to the beam at a constant velocity and the fluorescence detector was set to accumulate the XRF signal over a time of 100 ms. The fluorescence detector and the sample were placed at 90° and 10° with respect to the incoming beam, respectively. For the normalization of the spectra, one XRF map was collected below (8950 eV) and one above (9080 eV) the Cu K-edge ($E_0 = 8979 \text{ eV}$). The remaining five maps were collected at energies representing diagnostic features observed in XANES reference spectra (8981.0, 8986.5, 8995.0, 9007.0 and 9051.5 eV) (Fig. S1†). In total, each pixel was exposed to the beam for 0.7 s.

Complementary XAS measurements of small areas ($3 \times 3 \mu\text{m}^2$) referred to as point-XANES (pXANES) were conducted on selected points of interest (POI). POI were identified based on the Cu distribution maps. pXANES spectra were recorded from 8958 eV to 9060 eV corresponding to −21 eV pre- and 81 eV post-edge. The step size was 2 eV up to 10 eV before the edge, 0.5 eV up to 60 eV past the edge and 1.5 eV up to 81 eV above the edge. The integration time was 400 ms per energy



step, leading to 88 s total exposure time per POI. Three spectra were recorded on each POI.

Reference materials (CuS (covellite), CuFeS₂ (chalcopyrite), CuO (tenorite), CuSO₄ (copper sulphate), CuFe₂O₄ (cuprospinel) and Cu₂O (cuprite)) were prepared as 7 mm pellets in a cellulose matrix and measured in transmission mode using an ion chamber (I_0) and a X-ray diode (I_t). Reference materials of Cu(+I/+II) bound to humic acid were frozen into ceramic windows (4 × 8 mm) and stored in liquid nitrogen. An aliquot of the CuO-NP that were spiked to the digested sewage sludge²⁹ were prepared as 7 mm pellet and XAS data was acquired in transmission mode. The XANES data was identical to the tenorite XANES, which was therefore used for further evaluations.

All samples (XRF maps and pXANES) and reference materials were measured at cryogenic temperatures using a liquid nitrogen cryo jet (Oxford Instruments plc.) and setting the temperature at the nozzle to 100 K.

2.3 Data treatment

XRF peak intensities (counts within regions of interest (ROI)) and ion chamber currents were stored for each position (pixel) in a text file. All XRF maps of each sample were aligned to the same spatial grid using linear interpolation. The new grid was set very close to the old grid to preserve the shape of the data as good as possible and to keep the ROI count changes through interpolation to a minimum. This alignment was necessary due to slight variations in the acceleration and velocity of the sample stage, which were not always correctly captured by the triggering mode of the detector. Additionally, changes in the incident beam energy led to sub-pixel vertical changes in the beam position. To account for these shifts, the Ti-K_α XRF maps recorded at each energy were used to calculate the shift, which was then applied to the Cu XRF maps.³⁰ A map typically contained around 28'000 pixels ((500 × 500 μm²)/(3 × 3 μm² per pixel) ≈ 27'778 pixels). The spatial distribution of Cu was obtained by plotting the difference between the XRF intensity at the post- and pre-edge energy. If not stated otherwise, XRF Cu peak intensities were pixel-wise normalized by subtracting the intensity recorded before the absorption edge (8950.0 eV) and dividing by the intensity recorded at the highest energy above the edge (9080 eV). Often a first-order polynomial is used to approximate the background, which is especially relevant for measurements conducted in transmission mode. However, our measurements were conducted in fluorescence mode and we, thus, used a constant function defined by the absorption measured close to the absorption edge for the background removal. In some cases, at low absorber concentrations and consequently small edge-jumps, the pre- and post-edge background is better characterized by a second-order polynomial compared to a linear regression. To avoid evaluating and subsequently misinterpreting such cases, only pixels with associated signal intensities satisfying certain quality criteria (e.g., if the post-edge Cu-K_α XRF intensity was at least ten times the pre-edge intensity) were considered for further evaluation of the Cu speciation.

The spectra of the reference materials and pXANES spectra were imported into Athena³¹ for merging and normalization.

For linear combination fits (LCF) to XRF maps, the reference material spectra were treated as described for the XRF maps. For LCFs to the pXANES spectra data treatment was done as for bulk XANES. LCFs were performed over the entire energy range (−21 eV ≤ E_0 ≤ 81 eV). If not indicated otherwise, data processing and evaluation was performed using Matlab R2017b.

2.4 Model for chemical imaging

The experimental XAS spectra can be decomposed into variable fractions of selected reference spectra using linear combination fitting (LCF), a technique well established in the analysis of XAS data.¹ These linear combinations are generally expressed by:

$$\vec{\mu}_{\text{sample}} = \mu_{\text{references}} \vec{x} + \vec{\varepsilon}, \quad (1)$$

where $\vec{\mu}_{\text{sample}}$ contains the normalized X-ray absorption coefficients at different energies E_1, \dots, E_7 :

$$\vec{\mu}_{\text{sample}} = [\mu_{\text{sample}}(E_1), \mu_{\text{sample}}(E_2), \dots, \mu_{\text{sample}}(E_7)]^T. \quad (2)$$

The matrix $\mu_{\text{references}}$ is a 7 × 6 matrix containing the normalized X-ray absorption coefficients at different energies E_1, \dots, E_7 of the six selected reference materials (Fig. S1†). The vector \vec{x} contains the fitting parameters (with $\forall x_i \geq 0$) in the experimental spectrum ($\vec{\mu}_{\text{sample}}$) where $\vec{\varepsilon} = [\varepsilon_1, \dots, \varepsilon_7]^T$ cumulates the experimental errors. The vector \vec{x} represent the fractions of the respective spectral components (reference spectra) best reproducing the experimental spectrum. In XAS data analyses, multiple regression tools or least square methods embedded in different software packages are routinely employed to solve such mathematical problems.^{31–33} In a previous study, we used principle component analysis and target transform to show that our samples were best described by a linear combination of six reference spectra.²⁹ We assumed that the error term $\vec{\varepsilon}$ contained all elements of $\vec{\mu}_{\text{sample}}$ that were not represented by the six references. The entries of the error term were considered as noise, although they may include contributions of additional unidentified spectral components.

2.5 Observation model for the uncertainty evaluation of chemical images

X-ray absorption data can be measured in transmission or fluorescence mode.⁴ In fluorescence mode, (energy dispersive) XRF detectors convert incoming photons into digital signals, whose integration over time (and energy) leads to a fluorescence spectrum that can be used to derive physical–chemical characteristics of the sample under investigation.²⁷ For the development of an observation model describing the noise in the data and its discussion we only consider the fluorescence mode.

Abe *et al.*²⁸ recommended categorizing the noise related to beamline properties and detection systems into (i) stochastic noise, (ii) electronic noise, (iii) X-ray instability and (iv) mechanical motion of sample and optics. These recommendations were not specifically drafted for chemical imaging but X-ray absorption fine structure (XAFS) measurements in general. For further simplification of the present computations, we combined the points (i) and (ii) and points (iii) and (iv) and refer



to them by the uncertainty resulting from the electronics/detection system and the sample matrix–beam-interactions, respectively. The former will be referred to by $\sigma^{(d)}$ (d: detector), the latter one by $\sigma_j^{(m)}$ (m: matrix). The matrix uncertainty cumulates all uncertainties related to the sample and the sample stage, e.g., uncertainty derived from the linear interpolation assumption (Section 2.3) correcting for the imprecision of the stage setting (in micrometre) compared to the heterogeneities within the sample and the dimension of the beam (here $3 \times 3 \mu\text{m}^2$). Time dependencies were neglected as all measurements were performed with constant integration times. By ‘uncertainty’ we refer to the standard deviation of a normal distribution with the mean at the determined value of the normalized X-ray absorption coefficient. This will be discussed in detail in the following paragraphs.

To derive the distribution of the uncertainty we evaluated a Cu-K $_{\alpha 1}$ and Cu-K $_{\alpha 2}$ emission spectra recorded by an energy-dispersive fluorescence detector. While the Cu-K $_{\alpha 1,2}$ emission spectrum can be described by multiple narrow Lorentzians,³⁴ the geometry of the recorded fluorescence peak is mainly determined by the energy resolution of the detector³⁵ (black curve, Fig. 1a, in more detail in Fig. S3†).

The detector noise can never be smaller than zero counts and the detection system is tuned such that most likely zero counts occur during a blank measurement. Non-zero measurements are induced through, e.g., background scatter.²⁸ Therefore, we can approximate the distribution of counts by an exponential distribution. To represent the detector induced uncertainty, a random sample drawn from an exponential distribution parameterized by λ is added at each point in E_i (or each channel of the detector, compare Fig. 1a, grey curve).

$$\text{Signal}_{\text{detector}}(E_i) \sim \lambda \exp(-\lambda \times y) \quad (3)$$

Note that we use the tilde ‘ \sim ’ to indicate that a random sample was drawn from a distribution. The expected signal can now be described as a function of E_i (Fig. 1a, red curve):

$$\text{Signal}_{\text{expected}}(E_i) = \text{Signal}_{\text{calculated}}(E_i) + \text{Signal}_{\text{detector}}(E_i) \quad (4)$$

$\text{Signal}_{\text{calculated}}$ refers to the calculated signal as a function of the energy E_i (black curve, Fig. 1a). The energies included in the data evaluation, often referred to as regions of interest (ROI) can be selected by the user (grey vertical lines, Fig. 1a). The ROI comprised a specific number of XRF detector channels $N_E \in \mathbb{N}$. To obtain the intensity of the measurement (I_k , the index k represents the k -th measurement on the same spot), we sum over the measured signal at the discrete energies $E_{1,2,\dots,N_E}$:

$$I_k = \sum_{i=1}^{N_E} \text{Signal}_{\text{expected},k}(E_i) \quad (5)$$

The intensity I_k (red shaded area, Fig. 1a) can be converted to the normalized X-ray absorption coefficient for evaluation. Further, the intensity I_k can be separated into a contribution from the absorber in the matrix $I_k^{(m)}$ and a contribution from the detection system $I_k^{(d)}$, where usually $I_k^{(m)} \gg I_k^{(d)}$. However, $I_k^{(d)}$ can still be relevant in very dilute samples.²⁸

$$I_k = I_k^{(m)} + I_k^{(d)} = \sum_{i=1}^{N_E} \text{Signal}_{\text{calculated},k}(E_i) + \sum_{i=1}^{N_E} \text{Signal}_{\text{detector},k}(E_i) \quad (6)$$

According to our model, and assuming that the measurement is repeated k times at the exact same spot on the sample and neglecting beam damage, $I_k^{(m)}$ of each repetition will have

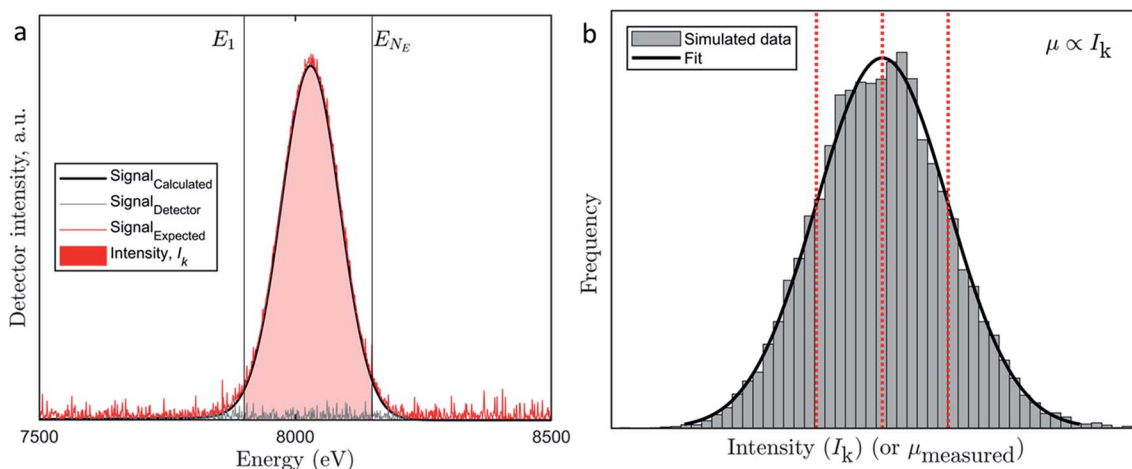


Fig. 1 (a) Graphical representation of the observation model. The black curve represents the calculated Cu K α fluorescence signal ($\text{Signal}_{\text{calculated}}$) reproduced from Sun *et al.*,³⁵ the grey curve represents the signal related to detector noise ($\text{Signal}_{\text{detector}}$) and the red curve represents the expected signal ($\text{Signal}_{\text{expected}} = \text{Signal}_{\text{calculated}} + \text{Signal}_{\text{detector}}$). The red shaded area between the two vertical grey lines E_1 and E_{N_E} represents the intensity I_k which corresponds to the integration of the expected signal ($\text{Signal}_{\text{expected}}$) over all detector channels between E_1 and E_{N_E} . (b) Histogram of the frequency of occurrence of intensities I_k . Contributions to I_k were balanced between $I_k^{(m)}$ and $I_k^{(d)}$, thus neither contribution dominated the other. The dotted lines indicate the mean value and \pm one standard deviation. Although the computation was done for I_k , the distribution also represents the distribution of μ_{measured} , because, $\mu \propto I_k$ where the mean is μ_{mean} and the standard deviation is $\sigma_j^{(m)} + \sigma^{(d)}$ (compare eqn (8)).



the exact same value (fluorescence is a stochastic process as well, however its variance is small compared to the processes discussed further down and not considered here). In contrast, $I_k^{(d)}$ (contains stochastic and electronic noise) will vary slightly as this value was sampled from an exponential distribution in each detector bin. In practice, however, the monochromator is set to a specific energy and the stage is moved with a given velocity line by line. I_k is obtained by integrating the signal over a selected time period. The 'width' of a pixel is calculated as the product of the scan speed and the integration time ($v_{\text{stage}} \Delta t_{\text{integration}} = W_{\text{pixel}}$). After completion of one XRF map at a certain energy, the energy is changed and an additional XRF map is recorded, *etc.* Ideally, the stage should be located at the exact same position at time t for every energy map. However, the locations of the stage will vary slightly, and even the smallest deviations between positions of consecutive energy maps will result in different 'absorber environments' of a specific location (pixel) at different energies. These differences, unfortunately, cannot be corrected entirely (Section 2.3). This situation is comparable to XAS at standard XAS beamlines at advanced synchrotron facilities with beam extensions of a few hundreds of micrometres, where the X-ray beam may slightly shift or change spatial flux density as a function of the energy and even tiny heterogeneities, diffraction artefacts or 'pinholes' in the sample can distort the absorption spectrum.

Furthermore, minor variations in $I_k^{(m)}$ may arise from, *e.g.*, resonant X-ray inelastic scattering (RIXS) at the absorber or elastic and inelastic scattering at any matrix element.^{28,36,37} In RIXS, the splitting of the photon emission energy at pre-edge energies of conventional XANES measurements as previously observed for CuO³⁶ may eventually increase the uncertainty in the pre-edge region. However, the distortion introduced by the RIXS contributions is small compared to the intensity of the absorption edge and further obscured by the energy resolution of the XRF detector.^{35,38} The removal of spectra distorted with contributions from elastic and inelastic scattering is discussed in Section 2.7. Differently, the change in the mean X-ray penetration depth before and after the absorption edge at spots with strong absorber concretions that extend into the Z direction, *e.g.*, over the total thin section height, might artificially enhance the contribution of the XRF intensity on the low energy side of the edge. This may be important, especially if the speciation changes in Z direction.

The resulting distribution of I_k to some extent depends on the assumption for the variation of $I_k^{(m)}$. However, the tendency towards a normal distribution is imposed by the variation of $I_k^{(d)}$. The spatial deviations of the stage position from the intended position between consecutive energy maps can be approximated by a normal distribution indicating a close to zero mean displacement (7×10^{-7} μm) and a standard deviation of 3×10^{-4} μm (Fig. S2†). Thus, for 95% of all pixels, the displacement is less than 1/2300 of the side length of a pixel. However, X-ray absorption is extremely sensitive to the absorber concentration and coordination. Therefore, larger displacements from the exact location of the pixel relative to the beam induce larger uncertainties. This consideration qualifies the

selection of a normal distribution to describe the uncertainty in $I_k^{(m)}$. Also, a normal distribution is most suitable to capture the sum of all uncertainties of the processes that were deemed relevant previously but cannot be assessed in detail.²⁸ In plain language, the size of the area underneath the black curve between the grey horizontal lines is sampled from a normal distribution (Fig. 1a). The result of the listed effects can be illustrated if $k = 1, 2, \dots, 10^4$ synthetic replicate measurements at one pixel are simulated and one arrives at another normal distribution (Fig. 1b).

By collecting sufficient repetitions of $I_k^{(m)} + I_k^{(d)}$, the mean of I_k and sum of two parameters $\sigma_j^{(m)}$ and $\sigma^{(d)}$ can be obtained (red dotted lines, Fig. 1b). These describe the most likely and mean values of the uncertainty introduced by different elements of the beamline. The uncertainty $\sigma_j^{(m)}$ will be different for each map j , but $\sigma^{(d)}$ will be the same as long as the experimental setup remains unaltered. The uncertainty in the X-ray absorption coefficient (μ) can be derived as outlined below.

$$\mu \propto \frac{I_k}{I_0} \propto I_k \quad (7)$$

The value of I_0 is obtained using an ion chamber and thus the uncertainty in the measurement of I_0 is very small and can thus be neglected here. Due to the linear relationship between μ and I_k the uncertainty in μ can also be described by a normal distribution. Thus, any measured normalized X-ray absorption coefficient (μ_{measured}) can be sampled from a normal distribution (N) where the mean is the corresponding, (presumably) true normalized X-ray absorption coefficient (μ_{mean}) and the standard deviation reflects the sum of the normalized uncertainties introduced by the matrix and the detector:

$$\mu_{\text{measured}} \sim N(\mu_{\text{mean}}, [\sigma_j^{(m)} + \sigma^{(d)}]). \quad (8)$$

Eqn (8) indicates that the total uncertainty ($\sigma_j^{(m)} + \sigma^{(d)}$) is an absolute quantity. However, in the present study, the X-ray absorption coefficients μ were normalized to values between 0 and 1, hence uncertainties obtained for different datasets are comparable.

Due to practical constraints (*e.g.* available beam time, beam damage), XRF maps at a given energy were only recorded once. This, however, hampers the determination of the measurement uncertainty through assessing the standard deviations. Consequently, we cannot determine the uncertainty directly related to the XRF intensities, only the uncertainty relative to a LCF result. Therefore, a proper choice of spectra from reference materials is critical for capturing the XAS signal originating from Cu species in the sample (compare Section 2.4 and eqn (1)). In this section, we argued that the uncertainty related to XRF signal production is comparable for each XRF measurement at every energy for all pixels. Thus, Bayesian inference can be used to evaluate the uncertainties according to eqn (8). We performed Markov Chain Monte Carlo (MC) Simulations using JAGS ("Just Another Gibbs Sampler") Version 4.3.0, based on BUGS ("Bayesian inference Using Gibbs Sampling"),³⁹ through the R library "rjags"⁴⁰ in RStudio Version 1.1.456 under R Version 3.4.4 (ref. 41) to



quantify $\sigma_j^{(m)}$ and $\sigma^{(d)}$. Details of the computation and the computer code are available in the ESI (Section S5†). Briefly, in the MC, the standard deviations $\sigma_j^{(m)}$ and $\sigma^{(d)}$ (eqn (8)) are treated as two normal distribution that are each characterized by a mean and standard deviation, which are to be determined (eqn S3 and S4†).

2.6 Synthetic data for model testing and calibration of the data interpretability

For model testing a synthetic dataset was prepared that covered a large variety of combinations of the six reference spectra. For each synthetic dataset, we created 1000 pixels (measurements or data points) where the (input) fractions were sampled from a Dirichlet distribution.

$$\sum_{i=1}^n x_i = 1 \text{ with } \forall x_i \geq 0 \quad (9)$$

A linear combination reconstruction of the normalized X-ray absorption coefficient was performed (eqn (1) without ϵ) at the seven energies at which XRF maps were recorded (Fig. S1†). Finally, the normalized X-ray absorption coefficients were modified by the introduction of variable magnitudes of noise according to eqn (8).

2.7 Quality criteria and benchmarking

Fit quality benchmarking was performed using the previously described synthetic datasets augmented with specific quantities of uncertainty (noise) (Fig. S7†). Thereafter, LCFs were performed to each pixel in each dataset, resulting in output fractions. The larger the uncertainty added to a dataset, the larger the expected discrepancy between the input and the output fractions. The generated input fractions of each data point (pixel) were sorted by their weight and compared to the computed output fractions, which were equally sorted. A score ranging from 0 to 6 was assigned to each pixel reflecting the agreement between the sorted input and the output fractions. A zero means that the first (largest) fractions is wrongly assigned, an one means that the largest fraction is correctly assigned, *etc.* Finally, a six means that the fractions of the reference spectra used for the LCF fits decreased in the same order for the input

and for the output fractions. The scores of the total dataset were then averaged. Furthermore, the percentage of pixels was calculated, for which the reference spectra with the largest spectral component (reference material spectrum) contributing to the LCF fit was correctly identified (score ≥ 1). The scores and the percentage of pixels (Correct largest Spectral Component Identified: CSCI) as described above can be used to assess the quality of the fits at different levels of uncertainty. Arguments in the favour of the introduction of the score and CSCI are given in the ESI (Section S6†).

3 Results and discussion

3.1 Validation of the uncertainty analysis technique

3.1.1 Recovery of the uncertainty from synthetic data. A synthetic dataset with four times 1000 pixels (or spectra) was compiled according to Section 2.6 with $\sigma^{(d)} = 0.1$ and $\sigma_j^{(m)} = [0, 0.01, 0.1, 0.2]$. In this way, 1000 pixels contained an uncertainty of $\sigma^{(d)} + \sigma_1^{(m)} = 0.1 + 0 = 0.1$, the following 1000 pixels of $\sigma^{(d)} + \sigma_2^{(m)} = 0.1 + 0.01 = 0.11$, *etc.* Distributions of the noise ($\sigma_j^{(m)}$ and $\sigma^{(d)}$) and the LCF fractions of reference materials (\bar{x}_i) were recovered with the MC approach by sampling 10^3 times from the model per pixel after initiating the model for 2×10^3 iterations. The long initiation phase ensured sampling under constant conditions. With this approach, the introduced uncertainties in synthetic spectra were successfully reproduced (Fig. 2), although the recovered values for the noise were slightly higher compared to the noise in the synthetic dataset ($\approx 5\%$). Also, a correlation between $\sigma^{(d)}$ and $\sigma_j^{(m)}$ was evident: lower values of $\sigma^{(d)}$ led to higher values of $\sigma_j^{(m)}$ and *vice versa* (Fig. S8†). This is caused by the linear dependence of these two parameters, which add up to the same uncertainty (eq. (8)). As a consequence, all $\sigma_j^{(m)}$ and $\sigma^{(d)}$ were subsequently combined to $\sigma^{(d)} + \sigma_j^{(m)} = \sigma_j$. Recovering σ_j from four additional datasets with $\sigma_j = [0, 0.01, 0.1, 0.2]$ during 10 repetitions yielded $\sigma_j = [0.0014 \pm 8 \times 10^{-4}, 0.0103 \pm 10^{-4}, 0.0988 \pm 2 \times 10^{-4}, 0.2032 \pm 5 \times 10^{-4}]$. The standard error varied between 1.87×10^{-3} and 1.38×10^{-2} . These results demonstrate that our approach recovers uncertainties in the synthetic dataset with high precision (standard error $\leq 1.4\%$) and accuracy (relative deviation $\leq 1.6\%$). The dataset $\sigma_j = 0$ was not included in the evaluation of

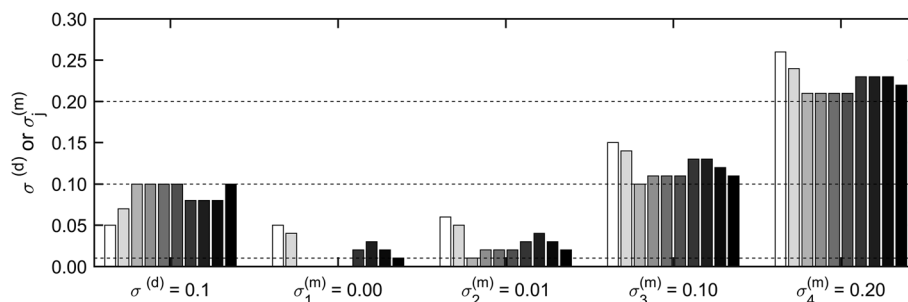


Fig. 2 The introduced uncertainty (x-axis) versus the recovered uncertainty (y-axis). Data shown are the result of ten replicates of the identical experiment. Each shade of grey is associated with one replicate. In each replicate experiment, the introduced uncertainties $\sigma^{(d)} = 0.1$ and $\sigma_j^{(m)} = [0, 0.01, 0.1, 0.2]$ were recovered for four times 1000 pixels. The horizontal, dotted lines indicate the introduced uncertainties of $\sigma_2^{(m)} = 0.01$, $\sigma_3^{(m)} = 0.10$ and $\sigma_4^{(m)} = 0.20$.



the accuracy of the uncertainty as relative deviations were difficult to calculate.

3.1.2 Recovery of the input fractions from synthetic data.

In this section the recovery of the correct spectral components of the synthetic XAS spectra with a well-defined uncertainty ranging from 0 to 0.20 (σ_j , levels according to Fig. 3 abscissa and Table S1†) were evaluated. The procedure follows the previously discussed workflow (Sections 2.6 and 2.7, outlined in Fig. S7†). In the absence of noise, and using the same reference spectra as in the synthetic data, the score reached 1.742 and the CSCI was 72.4% of the pixels (Fig. 3a and b, black circles). Performing LCF using a sequential quadratic programming (sqp) method^{42,43} (implemented through the function *fmincon* in Matlab, only constraint: $\forall x_i \geq 0$) the score for the noise-free data increased to about 3 (Fig. 3, black triangles) but the CSCI decreased to about 60%. Other commonly used algorithms, *e.g.*, simplex,⁴⁴ yielded much lower scores (data not shown). At noise levels in the synthetic spectra of 0.01 or higher, the scores and CSCI values resulting from the MC approach and the sqp method were almost identical (Fig. 3). However, both the score (a) and the CSCI (b) exponentially decreased with increasing noise. Due to the similarity of the Cu(II)-O spectra (tenorite, cuprospinel and copper sulphate) and the Cu-S spectra, (covellite, chalcopyrite and the amorphous copper sulphide⁴⁵), the respective spectra were treated interchangeably and referred to as 'Cu(II)-O' and 'Cu-S' combined. Through this procedure, the scores and the CSCIs significantly increased (Fig. 3). Although the score still showed a decreasing trend with increasing uncertainty, it remained above 1.5 at the highest uncertainty ($\sigma = 0.2$). Comparable results were observed to the CSCI, and in around 70% of all 'measurements', the most important spectral component was identified correctly. This indicates that the quality of the spectra allow to discriminate between Cu(II)-O and Cu-S species in the experimental samples, however, a further differentiation into individual Cu(II)-O and Cu-S species may not be possible. To some extent this is related to the spectral similarity of different Cu(II)-O/Cu-S

reference materials which may be compensated by recording additional XRF maps at different energies. We propose to use the described procedure to evaluate expected score/CSCI at variable energies and levels of uncertainties when planning synchrotron chemical imaging experiments.

3.2 The uncertainty in the case study datasets

The uncertainties in four experimental datasets recorded on digested sewage sludge spiked with CuO-NP or CuSO₄ and the resulting ashes were assessed using the approach described above. The levels of uncertainty were then compared to those in Section 3.1.2 and information on the quality of the extracted LCF fits were obtained. Briefly, XRF maps of the CuO spiked sludge (SLG NP), the CuSO₄ spiked sludge (SLG AQ) and ash derived from SLG AQ (ASH AQ) were $500 \times 500 \mu\text{m}^2$ in size with a $3 \mu\text{m}$ lateral resolution. The map of the ash derived from SLG NP (ASH NP) was $350 \times 350 \mu\text{m}^2$ with the same lateral resolution.

For initial and comparative analyses, all pixels with a non-normalized (NN) post edge XRF intensity of at least three times the NN pre-edge XRF intensity were included in the analysis. The correlation between $\sigma_j^{(m)}$ and $\sigma_j^{(d)}$, previously observed in the synthetic data, was also observed in the experimental data (Fig. S9†). Therefore, we combined these uncertainties into $\sigma_j(\sigma_j^{(d)} + \sigma_j^{(m)}) = \sigma_j$ and found $\sigma_j = [0.18, 0.17, 0.13, 0.33]$ for the experimental datasets SLG NP, SLG AQ, ASH NP and ASH AQ, respectively (Fig. S10†). Similar to the analysis of the synthetic datasets, the uncertainty was represented by a normal distribution around the values of σ_j . The uncertainty estimated for the two sludge samples was almost identical, likely due to the comparable sample matrix and Cu concentrations. The larger σ_j of ASH AQ may be explained by the granular texture of the ASH AQ sample resulting in an investigated area, dominated by an ash grain with rather low Cu concentrations.

Based on the results obtained with the synthetic dataset, uncertainties in the range of 0.13–0.18 as obtained for the experimental dataset translate into Cu(II)-O and Cu-S combined scores of around 1.5 and CSCI between 65 and

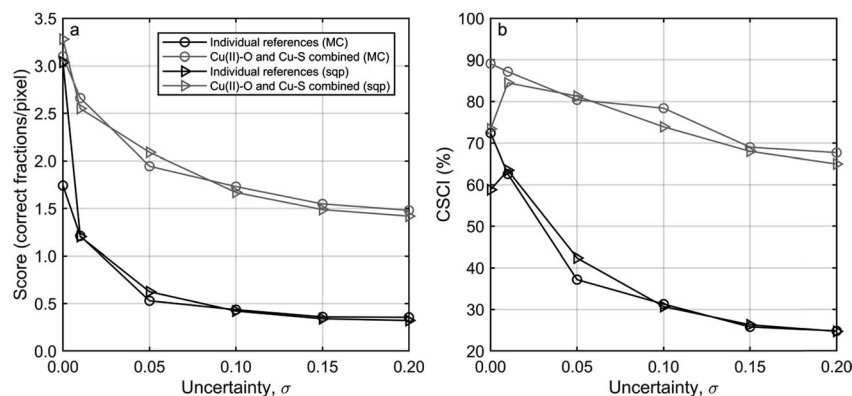


Fig. 3 (a) The score (number of correct fractions/pixel) and (b) CSCI (Correct largest Spectral Component Identified) versus the uncertainty introduced to the data. Black lines indicate that each reference was treated individually, grey lines indicate that each Cu(II)-O and Cu-S reference was treated as interchangeable as discussed in the text. Circles indicate that the score and CSCI was determined using the Markov Chain (MC), triangles represent data determined using sequential quadratic programming (sqp).



Table 1 Characteristics of the chemical images displayed in Fig. 4. Scores and CSCIs were determined by linear interpolation from values displayed in Fig. 3

Map	SLG NP	SLG AQ	ASH NP	ASH AQ
Threshold type	Relative (φ)	Relative (φ)	Absolute (χ)	Absolute (χ)
Level (φ or χ)	10	10	1000	3000
# pixels included	24 278	15 305	9198	6048
Pixels included, %	87.05	56.01	64.95	21.68
Uncertainty	0.18	0.18	0.09	0.13
Score (individual references)	0.36	0.36	0.45	0.39
CSCI (individual references) (%)	25.2	25.2	32.48	28.00
Score (combined references)	1.51	1.51	1.77	1.62
CSCI (combined references) (%)	68.22	68.22	78.80	72.76

70%. An uncertainty of 0.33, leading to a score of <1.5 and a CSCI < 60% (ASH AQ), however, is too large to make reliable statements concerning the speciation of Cu in the sample (Section 3.1.2).

To obtain more reliable datasets associated with a lower uncertainty, we introduced a second criterion for the inclusion of individual pixels. In addition to the 'relative threshold' (NN post edge XRF intensity $\geq \varphi$ times NN pre edge XRF intensity), where φ varied between 3 and 50, an 'absolute threshold' (NN post-edge XRF intensity $\geq \chi$) was introduced and varied between 500 and 1000.

The uncertainties determined at different levels of φ or χ are reported in Fig. S11 and Table S2.† Our results suggest that it is important to quantify the uncertainty as a function of different pixel inclusion criteria when analysing different samples. In each dataset, we had to make a compromise between the number of included pixels (spectra) and the resulting uncertainty by trying to include as many pixels as possible while keeping the uncertainty as low as possible. The results of this evaluation are tabulated (Table 1) and the chemical images based on the selected inclusion criteria are visualized (Fig. 4) and discussed in the following section.

Furthermore, it can be shown using the results from SLG NP that binning multiple pixels into larger pixels reduces the uncertainty in the dataset (Table S3†). Binning of $2 \times 2 = 4$ pixels reduces the uncertainty drastically ($\Delta\sigma = 0.04$). A further increase of the binning from $2 \times 2 = 4$ to $4 \times 4 = 16$ pixels results in a linear decrease of σ (Fig. S12†). Increasing the pixel binning to $5 \times 5 = 25$, however, only results in a moderate decrease in uncertainty. The large initial decrease of the uncertainty observed when 2×2 pixels were binned may be explained by the smoothing of residual shifts from the sample stage setting and from the beam location shift that are now small relative to the (binned) pixel size. The further decrease with increasing binning of pixels results from the improved counting statistics, which however, is associated with a decrease in the lateral resolution and in agreement with the dose limited resolution criteria.⁴⁶

3.3 The spatial distribution of the Cu speciation in sludge samples and corresponding ashes

Cu concentrations in the samples SLG NP, SLG AQ, ASH NP and ASH AQ were 1120 ppm, 1340 ppm, 1860 ppm and 2400 ppm, respectively. Based on the difference between the non-

normalized intensities at the post- and pre-edge energy (ΔI_i), the distribution patterns of Cu were obtained (Fig. 5). In the sludge samples, the Cu distribution was uneven and textural differences were observed between the SLG NP and the SLG AQ samples. The Cu concentration pattern of SLG NP appeared dotted with spot sizes between 10 and 20 μm (yellow/orange spots against the green/blue background, Fig. 5a). In contrast, the Cu distribution in the SLG AQ followed Schlieren-like structures with a few larger objects with diameters up to 70 μm (red circles, Fig. 5b). These observations are consistent with results from previous studies investigating the spatial Cu distribution in fresh biosolids, which also showed an uneven Cu distribution.^{12,47} The Cu distribution patterns in the sludge are in stark contrast to the Cu distribution patterns observed in both ash samples, where Cu was more evenly distributed in the ash grains (Fig. 5c and d). Therefore, the textural differences observed in the sludge samples largely disappeared during the incineration process.

LCF to bulk-EXAFS on the same samples indicated that SLG AQ can be described by 80% Cu_xS (amorphous) and 20% chalcopyrite (Table 2).²⁹ For SLG NP, identical analyses returned a similar fraction of Cu_xS (81%), but in addition to chalcopyrite (11%) a comparable fraction (8%) was associated with $\text{Cu}(\text{II})\text{-O}$ references (tenorite and copper sulphate).²⁹ To compare the integrated LCF fractions determined from the spatially resolved XAS data with the fractions derived from bulk EXAFS LCF, the spatially resolved XAS data were weighted with the intensity difference between the pre- and post-edge (ΔI_i).

Using the selected criteria for including individual XANES spectra (Table 1), chemical images that discriminate between $\text{Cu}(\text{II})\text{-O}$ (yellow pixels) and Cu-S (blue pixels) were extracted from each dataset (Fig. 4) and their integrated LCF fractions were evaluated (Table 2). In general, the integrated LCF results of SLG NP and SLG AQ were comparable, with 78 and 80% of the integrated experimental spectra assigned to the Cu-S reference spectra and 22% and 20% assigned to $\text{Cu}(\text{II})\text{-O}$ reference spectra, respectively (Table 2). If each pixel is binary associated with either Cu-S or $\text{Cu}(\text{II})\text{-O}$ spectra (Fig. 4) the fractions of Cu-S were 96% (SLG NP) and 95% (SLG AQ), respectively (Table 2).

In the chemical images of both sludge samples Cu was dominantly coordinated to sulphide, represented by the blue colour (Fig. 4a and b). However, the LCF analyses of the bulk spectra, returned a higher fraction of $\text{Cu}(\text{II})\text{-O}$ species for the



sludge samples spiked with CuO-NP and it was previously speculated that either the formation of Cu_xS coatings protected the CuO core from further sulfidation and/or that the agglomeration of CuO-NP substantially decreased the sulfidation kinetics.⁴⁸ The resolution ($3 \times 3 \mu\text{m}^2$) used in the present experiments does not allow to assess whether CuO– Cu_xS core–shell structures formed on an individual particle level, however, agglomerates of several particles in the sludge were well within the resolution limit of the experimental setup. In SLG NP, round Cu(II)–O objects were observed within the Cu–S matrix (red circles, Fig. 4a). Two of these objects were also associated with elevated Cu concentrations relative to the background Cu (red circles, Fig. 5a). We interpret these structures as agglomerates of CuO-NP that remained untransformed during the anaerobic digestion process. The other two objects were not associated with significantly elevated Cu concentrations (black circles, Fig. 5a) and may, therefore, represent CuO-NP, where Cu remained untransformed due to locally limited/restricted availability of bisulfide or Cu bound to O functional groups

from the organic matrix of the sludge. A chemical image with binned pixels ($3 \times 3 = 9$) confirmed these observations at a lower resolution but with higher score and CSCI (Fig. S13†). The three prominent spots observed on the sample SLG AQ reflecting locally elevated Cu concentrations (red circles, Fig. 5b) did not translate into local Cu(II)–O clusters (red circles, Fig. 4b) and may thus likely reflect agglomerates of CuS particles precipitated during the CuSO_4 spiking.

The locations of recorded pXANES are indicated by a series of red 'x' (Fig. 5 and 4). The spectra, reference material spectra, LCFs, resulting fractions and an in-depth analysis and discussion are given in the ESI (Section S8†). Briefly, the results derived from the evaluation of pXANES were consistent with the observations made on the chemical images (Fig. 4). In SLG NP, the pXANES p53–p55 showed coordination of Cu(II) to O, all other pXANES recorded on this sample revealed coordination of Cu to S. In SLG AQ, all recorded pXANES suggest Cu associated with S. In both ash samples the pXANES revealed strong local variations of the Cu oxidation state.

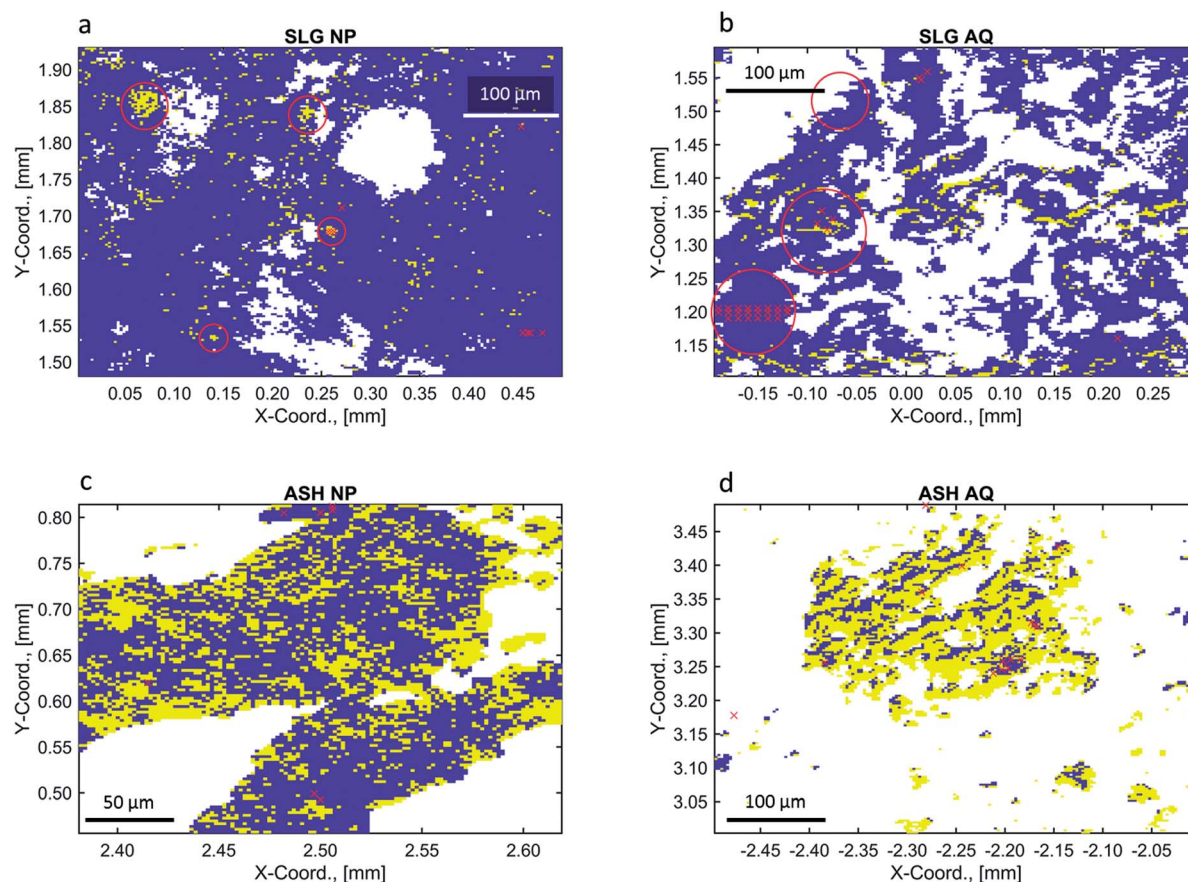


Fig. 4 Chemical images of the experimental samples ((a): SLG NP, (b): SLG AQ, (c): ASH NP and (d): ASH AQ) using combined references (yellow: fraction (Cu(II)–O) > fraction (Cu–S), blue: fraction (Cu(II)–S) > fraction (Cu–O)) and applying the exclusion criteria listed in Table 1. White pixels reflect excluded data points. The CSCIs were 68, 68, 79 and 73% for SLG NP, SLG AQ, ASH NP and ASH AQ, respectively. The coordinates (in mm) represent the coordinates relative to the beam in the centre of the sample. The red crosses and labels indicate the locations of the recorded pXANES. For their assignment to individual spectra, refer to Fig. 5. LCF was performed using a sequential quadratic programming (sqp) method.^{42,43}



Transmission electron microscopy and dynamic light scattering measurements did not show agglomeration at the micrometre scale (10–20 μm) in the spiking dispersion.²⁹ Therefore, the results of the present study support the hypothesis that the agglomeration occurred after CuO-NP were spiked to the digested sludge. Furthermore, the results showed that although the speciation of Cu in the two sludge samples was comparable, important textural differences based on the Cu concentrations (dotted *versus* Schlieren-type structures and large agglomerates) remained. Such differences were only observed in the sludge and were absent in the ash samples.

The Cu speciation in sludge samples spiked with either CuO-NP or CuSO₄ converged with the speciation of Cu related to presence of Cu in unspiked sludge within hours.⁴⁸ Nevertheless, spatial heterogeneities caused by different forms of Cu (nanoparticles *vs.* dissolved) added to the wastewater stream might outlive wastewater treatment including anaerobic digestion. A similar convergence of speciation has been observed for Zn during anaerobic sludge digestion.⁴⁹ A recent

study showed that lysimeter aged sludge obtained from a pilot wastewater treatment plant (WWTP) spiked with ZnO-NP inhibited the reproduction of earthworms to a higher degree compared to sludge produced by spiking dissolved Zn²⁺ to the same pilot WWTP.⁵⁰ The Zn speciation as well as the total Zn concentrations were almost identical in both sludges. The authors, therefore, speculated that the morphological or spatial differences between the ZnS phases formed after spiking ZnO-NP or ZnSO₄ to the pilot WWTP were responsible for the observed differences in growth inhibitions.⁵⁰ However, suitable tools to investigate spatial heterogeneities at the micrometre scale were not available and the hypothesis remained speculative. As Cu and Zn both classify as chalcophile transition metals⁵¹ the results from this study on Cu may also be transferrable to Zn. Structural differences of the sludge at the micrometre scale or even below may have influenced the Zn bioavailability and, thus, also have contributed to the increased eco-toxicity of the ZnO-NP-spiked *versus* the dissolved Zn spiked sludge reported in the study of Lahive *et al.*⁵⁰

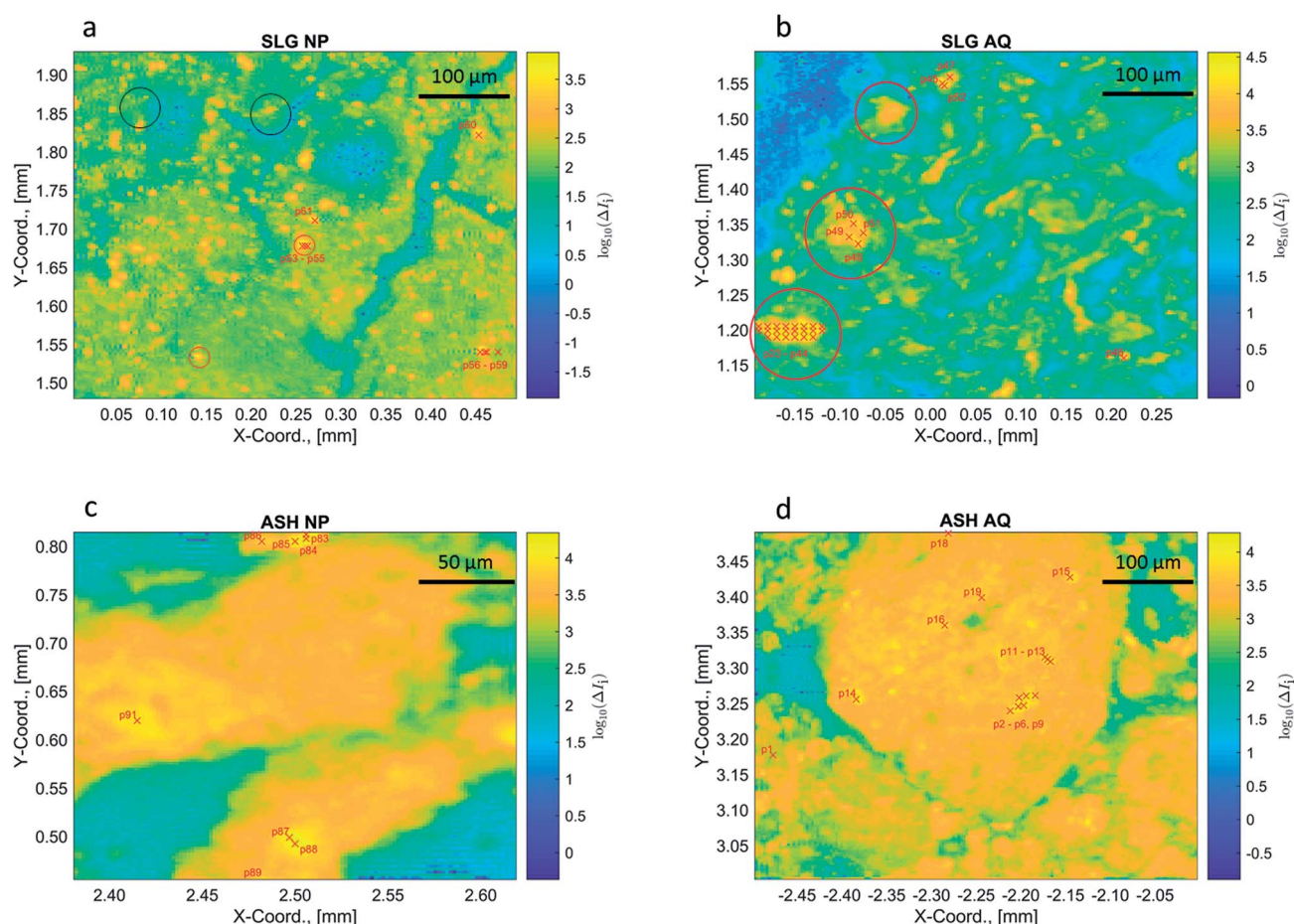


Fig. 5 An approximation of the Cu concentrations in the samples ((a): SLG NP, (b): SLG AQ, (c): ASH NP and (d): ASH AQ) reflecting the difference between the non-normalized intensity of the post- and pre-edge energy (ΔI). The intensities are displayed on a logarithmic scale. The coordinates (in mm) represent the coordinates relative to the beam in the centre of the sample. The red crosses and labels indicate the locations of the recorded pXANES. The pXANES, the reference material spectra and the results of a LCF analysis are given in the ESI (Fig. S14–S20†). Where multiple pXANES were recorded close to each other, the labels were combined, e.g., p2–p6, p9 in the sample ASH AQ. The location of the individual pXANES areas are given in the ESI (Section S8†).



Table 2 Summary of the results of the LCF analysis for SLG NP and SLG AQ from the chemical images (Fig. 4). The first column for each sample contains LCF results normalized by the intensity difference between the post-edge and pre-edge energy of the Cu K edge (ΔI_i). The second column contains the sum of the fractions either described with the spectra of Cu–S or Cu(II)–O reference materials. The third column contains the percentage of pixels either assigned to Cu–S or Cu(II)–O normalized by ΔI_i (Fig. 5)

SLG NP		SLG AQ									
Reference material group	Individual reference material	Sum of the LCF fractions over all pixels in %.		Sums of previous column, %	Fractions based on pixel association in %.	Sums of previous column, % ^a	EXAFS LCF, % ^a	Fractions based on pixel association, % normalized by ΔI_i	Sums of previous column, % ^a	EXAFS LCF, % ^a	Sums of previous column, % ^a
		Normalized by ΔI_i	Normalized by ΔI_i								
Cu-S	Cu ₃ S	39	78	96	81	92	45	80	95	85	100
	CuS	28			0		25			0	
	CuFeS ₂	11			11		10			15	
Cu(n)-O	CuO	13	22	4	7	8	10	20	5	0	0
	CuSO ₄	4			1		3			0	
	CuFe ₂ O ₄	5			0		7			0	

^a Fourth and fifth column: EXAFS LCF results of a previous study on the same samples given in ref. 29 (the samples SLG NP and SLG AQ, were referred to by D-NP and D-AQ, respectively). LCF was performed using a sequential quadratic programming (sqp) method.^{42,43}

4 Conclusion and outlook

We presented a model and related algorithms to quantify the uncertainty associated with hard X-ray derived chemical images. The model and algorithm were evaluated using a synthetic dataset. Two newly introduced data quality benchmarks (score and CSCI) to assess the reliability of the chemical images were evaluated using a synthetic dataset and applied to experimental data of a case study. Eventually, the model/algorithms can also be adapted to evaluate uncertainties associated with data interpretations in other settings where LCF methods are used, *e.g.*, electron energy loss spectroscopy.⁵²

The model/algorithm also offers a suitable tool for planning and conducting chemical imaging experiments. To the end of the latter, multiple XRF maps of a small area can be recorded and the presented algorithms can be used to quantitatively evaluate the resulting uncertainties and noise levels associated with LCF data interpretations under the given end-station settings. The model output can then be used to tune the end station settings to achieve a desired level of spectra interpretability (decision tree in Fig. S25†).

Progress in the quantitative assessment of the uncertainties may further enhance the interpretability of chemical images derived from synchrotron based hard X-rays. As demonstrated in this study, the evaluation and interpretation of chemical imaging data requires a detailed understanding of the underlying uncertainties, especially at short dwell times. The present model is purely empirical and future modifications will be directed towards a more rigorous treatment of physical processes (*e.g.*, undulator based spatial beam structure development with energy changes, beam polarisation, elastic and inelastic scattering, RIXS, self-absorption, beam damage, *etc.*), which will reveal future possibilities and also current limitations associated with the data interpretation of chemical images of heterogeneous samples with low absorber concentrations.

Code availability

The computer code extracting the uncertainty (σ) from (linearized) chemical images can be found in the Eawag Research Data Institutional Repository (ERIC) under: <https://doi.org/10.25678/0001MF>. The package includes a comprehensive README file and the data displayed in Fig. 3, 4, S12 and S13.†

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

The authors acknowledge the Paul-Scherrer-Institute (PSI) and the Swiss Light Source (SLS) for allocating beam time and Mr Dario Ferreira Sanchez (X05LA “microXAS” beamline/end station, SLS) for his support during the data acquisition. We acknowledge funding from the Swiss National Science



Foundation (SNF) through grant 5221.01038 and as well as discretionary funding from Eawag and PSI. Further, we acknowledge the library of ETH Zürich for covering the APC for open access publishing.

References

- 1 M. Gräfe, E. Donner, R. N. Collins and E. Lombi, *Anal. Chim. Acta*, 2014, **822**, 1–22.
- 2 D. E. Sayers, E. A. Stern and F. W. Lytle, *Phys. Rev. Lett.*, 1971, **27**, 1204–1207.
- 3 E. A. Stern, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1974, **10**, 3027–3037.
- 4 P. Willmott, *An introduction to synchrotron radiation: techniques and applications*, John Wiley & Sons, 2011.
- 5 M. A. Marcus, A. A. MacDowell, R. Celestre, A. Manceau, T. Miller, H. A. Padmore and R. E. Sublett, *J. Synchrotron Radiat.*, 2004, **11**, 239–247.
- 6 A. Manceau, N. Tamura, M. A. Marcus, A. A. MacDowell, R. S. Celestre, R. E. Sublett, G. Sposito and H. A. Padmore, *Am. Mineral.*, 2002, **87**, 1494–1499.
- 7 A. Manceau, N. Tamura, R. S. Celestre, A. A. MacDowell, N. Geoffroy, G. Sposito and H. A. Padmore, *Environ. Sci. Technol.*, 2003, **37**, 75–80.
- 8 S. D. Kelly, D. Hesterberg and B. Ravel, Analysis of soils and minerals using X-ray absorption spectroscopy, *Methods of soil analysis. Part 5, Mineralogical methods*, ASA-CSSA-SSSA, 5th edn, 2008.
- 9 A. Lanzirotti, L. Lee, E. Head, S. R. Sutton, M. Newville, M. McCanta, A. H. Lerner and P. J. Wallace, *Geochim. Cosmochim. Acta*, 2019, **267**, 164–178.
- 10 C. N. Borca, D. Grolimund, M. Willmann, B. Meyer, K. Jefimovs, J. Vila-Comamala and C. David, *J. Phys.: Conf. Ser.*, 2009, **186**, 012003.
- 11 R. Kirkham, P. A. Dunn, A. J. Kuczewski, D. P. Siddons, R. Dodanwala, G. F. Moorhead, C. G. Ryan, G. D. Geronimo, R. Beuttenmuller, D. Pinelli, M. Pfeffer, P. Davey, M. Jensen, D. J. Paterson, M. D. d. Jonge, D. L. Howard, M. Küsel and J. McKinlay, *AIP Conf. Proc.*, 2010, **1234**, 240–243.
- 12 B. E. Etschmann, E. Donner, J. Brugger, D. L. Howard, M. D. de Jonge, D. Paterson, R. Naidu, K. G. Scheckel, C. G. Ryan and E. Lombi, *Environ. Chem.*, 2014, **11**, 341–350.
- 13 M. A. Marcus, B. M. Toner and Y. Takahashi, *Mar. Chem.*, 2018, **202**, 58–66.
- 14 J. G. Mesu, A. M. J. van der Eerden, F. M. F. de Groot and B. M. Weckhuysen, *J. Phys. Chem. B*, 2005, **109**, 4042–4047.
- 15 J. Yang, T. Regier, J. J. Dynes, J. Wang, J. Shi, D. Peak, Y. Zhao, T. Hu, Y. Chen and J. S. Tse, *Anal. Chem.*, 2011, **83**, 7856–7862.
- 16 M. A. Marcus, *TrAC, Trends Anal. Chem.*, 2010, **29**, 508–517.
- 17 C. M. Rico, M. G. Johnson, M. A. Marcus and C. P. Andersen, *Environ. Sci.: Nano*, 2017, **4**, 700–711.
- 18 M. I. Heller, P. J. Lam, J. W. Moffett, C. P. Till, J.-M. Lee, B. M. Toner and M. A. Marcus, *Geochim. Cosmochim. Acta*, 2017, **211**, 174–193.
- 19 P. J. Lam, D. C. Ohnemus and M. A. Marcus, *Geochim. Cosmochim. Acta*, 2012, **80**, 108–124.
- 20 L. E. Mayhew, S. M. Webb and A. S. Templeton, *Environ. Sci. Technol.*, 2011, **45**, 4468–4474.
- 21 B. M. Toner, S. L. Nicholas and J. K. C. Wasik, *Environ. Chem.*, 2014, **11**, 4–9.
- 22 C. M. Rico, M. G. Johnson and M. A. Marcus, *Environ. Sci.: Nano*, 2018, **5**, 1807–1812.
- 23 C. G. Ryan, D. P. Siddons, R. Kirkham, P. A. Dunn, A. Kuczewski, G. Moorhead, G. D. Geronimo, D. J. Paterson, M. D. d. Jonge, R. M. Hough, M. J. Lintern, D. L. Howard, P. Kappen and J. Cleverley, *AIP Conf. Proc.*, 2010, **1221**, 9–17.
- 24 E. R. Malinowski, *Factor analysis in chemistry*, Wiley, 2002.
- 25 E. R. Malinowski, *Anal. Chim. Acta*, 1978, **103**, 339–354.
- 26 A. Manceau, M. Marcus and T. Lenoir, *J. Synchrotron Radiat.*, 2014, **21**, 1140–1147.
- 27 S. Heald, *J. Synchrotron Radiat.*, 2015, **22**, 436–445.
- 28 H. Abe, G. Aquilanti, R. Boada, B. Bunker, P. Glatzel, M. Nachtegaal and S. Pascarelli, *J. Synchrotron Radiat.*, 2018, **25**, 972–980.
- 29 J. Wielinski, A. Gogos, A. Voegelin, C. Müller, E. Morgenroth and R. Kaegi, *Environ. Sci. Technol.*, 2019, **53**, 11704–11713.
- 30 M. Guizar-Sicairos, S. T. Thurman and J. R. Fienup, *Opt. Lett.*, 2008, **33**, 156–158.
- 31 B. Ravel and M. Newville, *J. Synchrotron Radiat.*, 2005, **12**, 537–541.
- 32 M. Newville, *J. Phys.: Conf. Ser.*, 2013, **430**, 012007.
- 33 S. M. Webb, *Phys. Scr.*, 2005, **2005**, 1011.
- 34 M. Deutsch, G. Hölzer, J. Härtwig, J. Wolf, M. Fritsch and E. Förster, *Phys. Rev. A*, 1995, **51**, 283–296.
- 35 Y. Sun, S.-C. Gleber, C. Jacobsen, J. Kirz and S. Vogt, *Ultramicroscopy*, 2015, **152**, 44–56.
- 36 H. Hayashi, R. Takeda, Y. Udagawa, T. Nakamura, H. Miyagawa, H. Shoji, S. Nanao and N. Kawamura, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2003, **68**, 045122.
- 37 H. Hayashi, *AIP Conf. Proc.*, 2007, **882**, 833–837.
- 38 P. Glatzel, T.-C. Weng, K. Kvashnina, J. Swarbrick, M. Sikora, E. Gallo, N. Smolentsev and R. A. Mori, *J. Electron Spectrosc. Relat. Phenom.*, 2013, **188**, 17–25.
- 39 D. J. Lunn, A. Thomas, N. Best and D. Spiegelhalter, *Stat. Comput.*, 2000, **10**, 325–337.
- 40 M. Plummer, *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling*, *Proceedings of the 3rd international workshop on distributed statistical computing*, Vienna, Austria, 2003, vol. 124, No. 125, 10, pp. 1–10.
- 41 RCoreTeam, *R: A language and environment for statistical computing*, *R Foundation for Statistical Computing*, Vienna, Austria, 2017, <https://www.R-project.org/>.
- 42 M. J. Powell, in *Numerical analysis*, Springer, 1978, pp. 144–157.
- 43 M. J. Powell, in *Nonlinear programming 3*, Elsevier, 1978, pp. 27–63.
- 44 J. C. Lagarias, J. A. Reeds, M. H. Wright and P. E. Wright, *SIAM J. Optim.*, 1998, **9**, 112–147.



- 45 R. A. D. Patrick, J. F. W. Mosselmans, J. M. Charnock, K. E. R. England, G. R. Helz, C. D. Garner and D. J. Vaughan, *Geochim. Cosmochim. Acta*, 1997, **61**, 2023–2036.
- 46 A. Rose, in *Advances in Electronics and Electron Physics*, ed. L. Marton, Academic Press, 1948, vol. 1, pp. 131–166.
- 47 E. Donner, D. L. Howard, M. D. d. Jonge, D. Paterson, M. H. Cheah, R. Naidu and E. Lombi, *Environ. Sci. Technol.*, 2011, **45**, 7249–7257.
- 48 A. Gogos, A. Voegelin and R. Kägi, *Environ. Sci.: Nano*, 2018, **5**, 2560.
- 49 E. Lombi, E. Donner, E. Tavakkoli, T. W. Turney, R. Naidu, B. W. Miller and K. G. Scheckel, *Environ. Sci. Technol.*, 2012, **46**, 9089–9096.
- 50 E. Lahive, M. Matzke, M. Durenkamp, A. J. Lawlor, S. A. Thacker, M. G. Pereira, D. J. Spurgeon, J. M. Unrine, C. Svendsen and S. Lofts, *Environ. Sci.: Nano*, 2017, **4**, 78–88.
- 51 F. Albarede, in *Encyclopedia of Astrobiology*, ed. M. Gargaud, R. Amils, J. C. Quintanilla, H. J. Cleaves, W. M. Irvine, D. L. Pinti and M. Viso, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 283–283, DOI: 10.1007/978-3-642-11274-4_261.
- 52 S. Yakovlev and M. Libera, *Micron*, 2008, **39**, 734–740.

