

Robust framework and software implementation for fast speciation mapping

Serge X. Cohen,^{a,*} Samuel M. Webb,^b Pierre Gueriau,^{a,c,†} Emmanuel Curis^{d,e} and Loïc Bertrand^{a,†*}

Received 12 March 2020

Accepted 27 April 2020

Edited by I. Lindau, SLAC/Stanford University, USA

† Current affiliation: Institute of Earth Sciences, University of Lausanne, CH-1015 Lausanne, Switzerland.

Keywords: synchrotron imaging; software; chemical speciation; X-ray absorption spectroscopy; heterogeneous materials.

^aUniversité Paris-Saclay, CNRS, Ministère de la Culture, UVSQ, IPANEMA, F-91192 Saint-Aubin, France,

^bStanford Radiation Lightsource (SSRL), SLAC National Accelerator Laboratory, Menlo Park, CA, USA,

^cSynchrotron SOLEIL, F-91192 Gif-sur-Yvette, France, ^dLaboratoire de Biomathématiques, EA 7537 – BioSTM, Faculté de Pharmacie de Paris – Université Paris Descartes, 4 Avenue de l'Observatoire, F-75006 Paris, France,

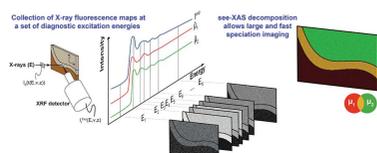
^eService de Biostatistique et d'Informatique Médicale, Hôpital Saint-Louis, Assistance Publique Hôpitaux de Paris, 1 Avenue Claude-Vellefaux, F-75010 Paris, France, and ^fUniversité Paris-Saclay, F-91190 Saint-Aubin, France.

*Correspondence e-mail: serge.cohen@synchrotron-soleil.fr, loic.bertrand@universite-paris-saclay.fr

One of the greatest benefits of synchrotron radiation is the ability to perform chemical speciation analysis through X-ray absorption spectroscopies (XAS). XAS imaging of large sample areas can be performed with either full-field or raster-scanning modalities. A common practice to reduce acquisition time while decreasing dose and/or increasing spatial resolution is to compare X-ray fluorescence images collected at a few diagnostic energies. Several authors have used different multivariate data processing strategies to establish speciation maps. In this manuscript, the theoretical aspects and assumptions that are often made in the analysis of these datasets are focused on. A robust framework is developed to perform speciation mapping in large bulk samples at high spatial resolution by comparison with known references. Two fully operational software implementations are provided: a user-friendly implementation within the *MicroAnalysis Toolkit* software, and a dedicated script developed under the *R* environment. The procedure is exemplified through the study of a cross section of a typical fossil specimen. The algorithm provides accurate speciation and concentration mapping while decreasing the data collection time by typically two or three orders of magnitude compared with the collection of whole spectra at each pixel. Whereas acquisition of spectral datacubes on large areas leads to very high irradiation times and doses, which can considerably lengthen experiments and generate significant alteration of radiation-sensitive materials, this sparse excitation energy procedure brings the total irradiation dose greatly below radiation damage thresholds identified in previous studies. This approach is particularly adapted to the chemical study of heterogeneous radiation-sensitive samples encountered in environmental, material, and life sciences.

1. Introduction

Microscale chemical speciation is essential to describe a material's local composition, structure and properties in heterogeneous samples. Synchrotron-based X-ray absorption near-edge spectroscopy (XANES) is one of the main analytical methods used for this purpose, and it has been applied in a variety of scientific fields such as environmental sciences, heritage sciences, Earth and meteorite geology, and plant sciences (Gardea-Torresdey *et al.*, 2005; Linge, 2008; Cotte *et al.*, 2010; Bertrand *et al.*, 2012; Etschmann *et al.*, 2014; Henderson *et al.*, 2014; Gueriau *et al.*, 2016). This excitation spectroscopy probes the energy levels of unoccupied electronic states from core levels upon X-ray photoexcitation. Of particular interest for the characterization of heterogeneous



materials is the use of focusing optics to acquire XANES in a spatially resolved manner. As X-ray optics have continued to improve in terms of reducing the size of focused beams, and synchrotron X-ray sources have become brighter, the ability to collect high-resolution X-ray images has become more and more routine. Two different experimental methodologies have been used to collect speciation images in transmission.

The first strategy uses a large beam and a spatially resolved detector to collect ‘instantaneous’ images of the transmitted X-ray intensity through the sample at successive given excitation energies – effectively taking a radiograph of the sample at each excitation energy. The work of Kinney *et al.* was one of the first studying chemical contrast in imaging using transmitted radiographs above and below the absorption edge of interest to create the chemical contrast (Kinney *et al.*, 1986). These concepts have further been developed with full-field XANES, where the thin section of a sample is positioned in front of an extended X-ray beam; radiographs are collected at successive excitation energy points on an area detector positioned downstream (Ade *et al.*, 1992; De Andrade *et al.*, 2011; Fayard *et al.*, 2013). This approach can be used without additional X-ray optics to collect high-resolution (typically micrometres) images using a luminescent screen and a visible microscope (Fayard *et al.*, 2013; Tack *et al.*, 2014; Rau *et al.*, 2002; Sciau & Wang, 2019), or make use of X-ray condensing optics to illuminate the sample, followed by an X-ray imaging optical element, such as a Fresnel zone plate, to image the transmission onto an area detector as in transmission X-ray microscopy (TXM) with resolutions in the tens of nanometres (Holt *et al.*, 2013; Liu *et al.*, 2012a; Nelson *et al.*, 2011; Meirer *et al.*, 2011; Li *et al.*, 2015).

A second strategy uses an X-ray focused microbeam to scan across a sample surface. In scanning transmission X-ray microscopy (STXM), the intensity of a transmitted beam is collected while varying the energy of the incoming beam and raster-scanning the sample with a fixed microscopic or nanoscopic X-ray beam (Zhang *et al.*, 1994; Flynn *et al.*, 1999; Jacobsen *et al.*, 2000).

Both strategies allow for the position-sensitive detection of transmitted intensities in the presence of and without the sample. The resulting maps allow reconstruction of a XANES spectrum at each pixel. Both are performed by measuring the transmitted X-ray intensities, which requires that samples are thinned down to allow for typically 10–20% of incident photons to cross the sample. Obviously, since the excitation energy is imposed by the absorption edge of the element whose speciation is being studied, it cannot be increased to improve transmission as in hard X-ray radiography or tomography. At low X-ray energies of a few keV or below, this imposes stringent constraints on sample thickness, equivalent to those requested in transmission electron microscopy. At greater energies of 8–22 keV, samples must still be thinned from a few dozen to a few hundred micrometres, which can be very constraining for samples. The transmission method is further constrained by the concentration of the element to be measured, as it needs to show strong enough absorption contrast across the energies to be measured.

When it is not possible or desirable to prepare thin sections, fluorescence detection of XANES signals from the surface of a thick sample is an option. This typically involves scanning the sample as in the second strategy above by detecting the fluorescent signal with an energy-dispersive X-ray fluorescence (XRF) detector generally placed at 90° with respect to the incident X-ray beam. In this geometry, a higher flux must reach the sample in order to obtain a signal-to-noise ratio identical to that of a transmission geometry. For example, the *K*-shell XRF yield for 3*d* transition metals of the fourth period (Sc–Zn) ranges from 16% to 46%. While XRF emission is isotropic, the solid angles of most fluorescence detectors are limited, and only a small fraction of the total fluorescence signal can be measured. In addition, matrix effects (re-absorption of fluoresced photons) can significantly impact the collected signal. XRF detection is usually strongly count-rate limited, further decreasing yields of detection and signal-to-noise ratios. This methodology has been applied to a wide host of samples, including geological systems to examine transition metal speciation (Sutton *et al.*, 1995, 2002) and more dilute biological systems (Pushie *et al.*, 2014; Pickering *et al.*, 2000; Oram *et al.*, 2011).

While the collection of transmission and fluorescence images as a function of incident X-ray energy is a relatively common technique in the literature (Kaulich *et al.*, 2011; Lombi *et al.*, 2011; Wang *et al.*, 2013; Collins & Ade, 2012; Liu *et al.*, 2012b; Pickering *et al.*, 2000), the treatment and analysis of the datasets is not always straightforward. Several hundred images at different energies are often collected in full-field imaging, aiming to reconstruct a ‘complete’ XANES spectrum in each pixel. However, in dilute fluorescence experiments where raster-scan XRF imaging must be performed, a full complement of energies may take a prohibitive amount of experimental time. A selection of a limited number of energies must therefore be made using the chemical knowledge of the system, in order to contrast different chemical species in the specimen. Several review articles have pioneered the idea of mapping the chemical nature of the sample by collecting XRF data at several varied excitation energies (Bertsch & Hunter, 2001; Marcus, 2010). This can be done by exploiting the contrast between different species at, for example different resonances in the S *K*-edge XAS spectra for organic and inorganic forms in fossilized leaf materials (Edwards *et al.*, 2014). Lerotic *et al.* (2004) used cluster analysis for the analysis of multiple energy datasets. Other authors have proposed statistical treatments like principal component analysis (PCA) while the data collection is on-going to help identify regions of interest in the sample in both transition metal systems (Mayhew *et al.*, 2011; Etschmann *et al.*, 2014) and complex sulfur speciation systems (Farfan *et al.*, 2018).

In this manuscript, we focus on the theoretical aspects and assumptions that are often made in the analysis of these datasets. A robust framework is developed to perform speciation mapping in large thick samples at high spatial resolution by comparison with known references. We show that, with the proper experimental setup and the measurement of images over a selected number of energy points, the linear

decomposition of the XANES signal yields the relative concentration of individual chemical species and the total concentration of the element. While using a linear decomposition, this method takes advantage of the increased number of effective data measurements at each pixel in the image, which results in improved overall counting statistics. The method minimizes beam exposure to the sample, which is particularly important in the case of samples that may be sensitive to radiation damage, such as the real paleontological sample on which the effectiveness of this approach is tested.

The low computational cost of the involved multivariate algorithm combined with the usefulness of the produced results make the proposed method suitable to inline use while collecting X-ray data and furthermore provides useful hints to the experimental design. To facilitate this aspect, the proposed approach has been integrated and deployed into a freely available software package, the *MicroAnalysis Toolkit* (Webb, 2011), and a dedicated package developed under the *R* environment. This approach is of interest for the study of a wide range of materials containing different chemical forms of an element, as frequently encountered in many fields of materials science.

2. Methods – sparse excitation energy XAS

The following approximations have been made considering that the XAS signal is detected in fluorescence, on a thick sample (an infinitely thin sample would give similar approximations), regarding a target element diluted in a matrix presenting a significant regularity.

2.1. Notations

Let I_0 be the incident beam intensity, and θ and ϕ the angles between the sample surface and the incident and detected beam, respectively (Fig. 1). For the purpose of clarity, Table 1 presents a list of notations used in the manuscript. The experimental frame is set such that the incident beam propagates along the \mathbf{x} axis, the \mathbf{z} axis is vertical pointing upward and \mathbf{y} is such that $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ forms a direct orthonormal frame. The fluorescence is measured at a right angle from the incident beam such that $\theta + \phi = \pi/2$. We consider that the measured fluoresced signal propagates along \mathbf{y} . Note that, although these constraints correspond to usual geometries and simplify the remaining computation, they have no direct consequence on the overall result.

We assume that we collect the experimental fluoresced intensity I_f^{Exp} from a rectangular parallelepiped of thickness d . The thickness in the beam direction is then $e = d/\sin\theta$. We attach to the sample a frame with the \mathbf{u} axis orthogonal to the entry face and pointing inward the sample, \mathbf{z} the unitary vertical vector pointing upward and \mathbf{v} parallel to the entry face such that $(\mathbf{u}, \mathbf{v}, \mathbf{z})$ forms a direct orthonormal basis. Both orthonormal frames have their origin on the sample surface such that voxels on the surface are characterized by $x = y = 0$ or $u = 0$ according to the set of coordinates used. We will also use \mathbf{r} to point to the position of a voxel irrespective of the

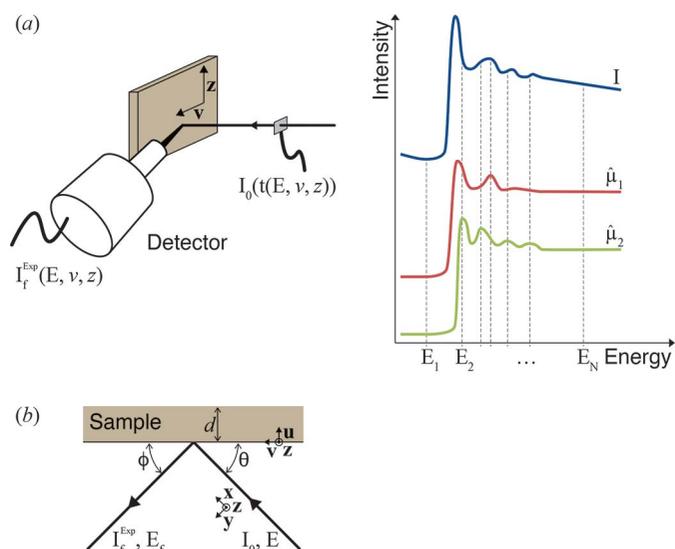


Figure 1
(a) Graphical illustration of the acquisition process. (b) Schematic of the geometry and of the geometrical notations used.

coordinate system used. The pixel coordinates on the sample surface are noted (v, z) with implied $u = 0$. The incident and fluoresced energies are denoted E and E_f , respectively. We included a possible indirect dependency of I_0 in (E, v, z) as fluctuations of I_0 occur in the scanning timescale which we express as $I_0[t(E, v, z)]$.

2.2. Algorithmic development

2.2.1. XRF from an elementary volume: general case. The linear absorption coefficient for the voxel at \mathbf{r} is noted $\mu(E, \mathbf{r})$, and can be decomposed as

$$\mu(E, \mathbf{r}) = \mu_m(E, \mathbf{r}) + \sum_{k=1}^K C_k(\mathbf{r}) \mu_k(E),$$

where $\mu_m(E, \mathbf{r})$ is the matrix absorption coefficient that varies slowly with energy E and $\mu_k(E)$ is the absorption coefficient of species k around the ‘target’ elemental edge considered – the linear combination standing if the atom probed exists as K distinct chemical forms. In the energy range which is usually considered for XAS studies, photo-absorption is by far the main photon–matter interaction mechanism leading to fluorescence. Hence we will consider all along our computation that global absorption, be it by the matrix or the target element, is photo-absorption related. Furthermore, since the incident energy E is close to the excitation energy of the target element, it is unlikely that the fluoresced signal produces any secondary sources of fluorescence from this element and hence we consider only primary fluorescence signal.

Let us first consider the general case by computing the contribution of a given voxel $\mathbf{r}_0 = (x_0, y_0, z_0)$,

Table 1

Notations used in the article.

K	Number of species present (indexed by $k \in [1, K]$)
N	Number of energies at which measurement is performed (indexed by $n \in [1, N]$)
\mathbf{r}, dV	Position and volume, respectively, of an elementary volume (<i>voxel</i>)
I_0	Incident beam intensity
I_f^{Exp}	Experimentally measured intensity of the fluorescence signal
μ_m	Mass attenuation coefficient of the matrix
l	The homogeneity length of the matrix composition, or a lower bound thereof, used to express the hypotheses of matrix homogeneity both transversely and longitudinally
μ_{cm}	Apparent <i>complete mass attenuation coefficient</i> of the matrix (accounting for the attenuation of both the incident and the fluoresced beams)
l_c	Maximum apparent attenuation length of the matrix, defined as $l_c(E_f, \nu, z) = \max_E [1/\mu_{\text{cm}}(E, E_f, \nu, z)]$
d	Sample thickness
μ_k	Mass attenuation coefficient of the species k of the probed element
$\tilde{\mu}_k$	Normalized reference spectra for the species k
C_k	Concentration of species k
$E, E^{\text{edge}}, \delta E$	Incident energy, reference energy and deviation from reference: $E = E^{\text{edge}} + \delta E$
E_f	Energy of the fluoresced signal from the excited element, energy at which the detection is performed
μ_k^{edge}	Mass attenuation contribution of excited electron in species k
μ^{edge}	Mass attenuation contribution of excited electron at energy higher than the edge
ε_f, Ω	Quantum efficiency of the fluorescence detector, and solid angle of detection
$\gamma_k, E_f, \gamma_{E_f}$	Yield of species k , and excited element (irrelevant of the species considered) to produce fluorescence at energy E_f
$\mathcal{A}_m, \mathcal{B}_m$	Coefficients in the affine model of μ_m
$\mathcal{A}^{\text{edge}}, \mathcal{B}^{\text{edge}}$	Coefficients in the affine model of <i>out of edge</i> mass attenuation of the excited element, that is contribution to the mass attenuation of non-excited electrons
$a^{\text{edge}}, b^{\text{edge}}$	Coefficients in the affine model for μ^{edge}
A, B	Affine model of the XAS baseline on <i>relative scale</i> , when using normalized reference spectra
$C_0, C_{\delta E}$	Linear regression coefficients, respectively, intercept and coefficient
C_k	Linear regression coefficients, relative to each of the K tested species
C^{sum}	Sum of the C_k ; $C^{\text{sum}} = \sum_{k \in [1, K]} C_k$
D_m	A proportionality factor depending on the experimental setup, matrix composition and used to measure the fluorescence signal

$$I_f^{\text{Exp}}(E, \mathbf{r}_0) = I_0[t(E, \nu, z)](\varepsilon_f \Omega / 4\pi) \quad (1)$$

$$\times \exp \left[- \int_{x=0}^{x_0} \mu(E, x, y_0, z_0) dx \right] \quad (2)$$

$$\times \left[\sum_{k=1}^K \gamma_{k, E_f} C_k(\mathbf{r}_0) \mu_k(E) \right] \quad (3)$$

$$\times \exp \left[- \int_{y=y_0}^{y_0+x_0 \tan \theta} \mu(E_f, x_0, y, z_0) dy \right] \quad (4)$$

$$+ \epsilon(E, \nu, z).$$

Line (1) of this equation corresponds to factors independent of (x_0, y_0, z_0) : the incident beam flux, $I_0[t(E, \nu, z)]$, and the detector characteristics, $\varepsilon_f \Omega / 4\pi$ that accounts for its quantum efficiency (ε_f) and the solid angle it covers (Ω). Note that considering that Ω is constant for all values of x in the sample assumes that the detector is sufficiently far away from the voxel of interest to consider the incoming fluorescent beam as a parallel beam coming from the infinite.

Line (2) corresponds to the decrease of the incident flux due to absorption while the incident beam penetrates in the x direction to the voxel of interest.

Line (3) accounts for the quantity of fluorescence emitted by the elementary volume irradiated by the incident microbeam. It takes into account the fluorescence yield $(\gamma_{k, E_f})_{k \in [1, K]}$ of the K species of the target element whose edges are probed and their amount in the given voxel. The latter will be called the *fluorescence source term*.

Finally, Line (4) of this equation considers that the fluorescence emission at E_f is attenuated due to reabsorption in the sample on the outgoing path expressed in its integral form arising from the Beer–Lambert law. Finally, we have the error term $\epsilon(E, \nu, z)$ that we will omit from now on to simplify notations.

2.2.2. Target element diluted within the matrix. Assuming that the K species of the target element are very diluted in the matrix, the absorption coefficient, $\mu(E, \mathbf{r})$, is largely dominated by its matrix-related contribution

$$\mu(E, \mathbf{r}) \simeq \mu_m(E, \mathbf{r}). \quad (5)$$

Under this hypothesis, the amplitude of the attenuation, both on the incident and fluoresced beam, is made independent of the concentration of the species of the target element. In this approximation, the absorption of the target element affects only the *fluorescence source term*.

2.2.3. Matrix and target element of homogeneous composition at the scale of the probed volume. We consider that both the composition of the matrix and the concentrations of the target element in its various species, $(C_k)_{k \in [1, K]}$, vary sufficiently slowly so that they can be considered as constant on the neighbourhood corresponding to the measure in each value of (ν, z) . More precisely, if we consider the $\mathbf{r}_0 = (x_0 = 0, y_0, z_0) = (u_0 = 0, v_0, z_0)$ voxel, we assume that there is a *homogeneity radius*, noted l , greater than the radius of the volume probed by the beam when the sample is fixed and such that

$$\forall \mathbf{r}, \quad \|\mathbf{r} - \mathbf{r}_0\| < l \Rightarrow [\mu_m(E, \mathbf{r}) \simeq \mu_m(E, \mathbf{r}_0)] \quad \text{and} \quad (6)$$

$$[\forall k \in [1, K], \quad C_k(\mathbf{r}) \simeq C_k(\mathbf{r}_0)].$$

Combining these hypotheses with the one developed in the previous paragraph we obtain an analytical expression of the measured fluorescence originating from a voxel centred on \mathbf{r} , showing a simple dependency on the concentrations $C_k(\mathbf{r})$,

$$I_f^{\text{Exp}}(E, \mathbf{r}) = I_0[t(E, v, z)](\varepsilon_f \Omega / 4\pi) \quad (7)$$

$$\times \exp[-\mu_m(E, \mathbf{r}) x_0] \left[\sum_{k=1}^K \gamma_{k, E_f} C_k(\mathbf{r}) \mu_k(E) \right]$$

$$\times \exp[-\mu_m(E_f, \mathbf{r}) x_0 \tan \theta].$$

2.2.4. Characteristic matrix absorption and thickness of the sample. In the previous equation, terms in the exponential can be grouped and the $-x_0$ term can be factorized and we can define $\mu_{\text{cm}}(E, E_f, v, z) = \mu_m(E, v, z) + \tan \theta \mu_m(E_f, v, z)$, such that the expression of the contribution of a voxel becomes

$$I_f^{\text{Exp}}(E, \mathbf{r}) = I_0[t(E, v, z)](\varepsilon_f \Omega / 4\pi) \exp[-\mu_{\text{cm}}(E, E_f, v, z) x_0]$$

$$\times \left[\sum_{k=1}^K \gamma_{k, E_f} C_k(v, z) \mu_k(E) \right]. \quad (8)$$

$l_c(E_f, v, z) = \max[1/\mu_{\text{cm}}(E, E_f, v, z)]$ can be interpreted as the characteristic attenuation length of the experimental setup, and we need to have $l \gg l_c(E_f, v, z)$ to satisfy the hypothesis of a homogeneous matrix and target element species composition. Furthermore, if d , the thickness of sample, satisfies $d \gg l_c(E_f, v, z)$ we can treat the sample as infinitely thick. This condition, together with the hypothesis expressed in the previous section, basically states that the matrix has a *homogeneity radius* which is both greater than the beam radius and the penetration depth of the incident beam interaction.

2.2.5. Summation along the incident beam. Equation (8) is implicitly an infinitesimal form, corresponding to

$$dI_f^{\text{Exp}} f = I_f^{\text{Exp}} f(E, \mathbf{r}) d\mathbf{r}. \quad (9)$$

This can now be summed over the path of the incident beam (in x_0), considering a semi-infinite sample, leading to the intensity of the fluorescence signal measured when the pixel (v, z) is illuminated by the probe,

$$dI_f^{\text{Exp}} f(E, v, z) = I_0[t(E, v, z)] \frac{\varepsilon_f \Omega}{4\pi} \frac{1}{\mu_{\text{cm}}(E, E_f, v, z)}$$

$$\times \left[\sum_{k=1}^K \gamma_{k, E_f} C_k(v, z) \mu_k(E) \right] dv dz, \quad (10)$$

which is a linear equation from the viewpoint of $(C_k)_{k \in [1, K]}$. We can then estimate the coefficients of these linear equations by measuring K spectra of references that are linearly independent. The simplest way would be to measure K spectra $(S_i)_{i \in [1, K]}$ such that $\forall (i, k) \in [1, K]^2$ the *a priori* known concentrations are $(C_k)_i = \delta(i, k)$; in other words, K reference spectra corresponding to a single species in a unique matrix.

2.2.6. Reference and normalized spectra. In the general case corresponding to the approximations (10) the reference

spectra allowing estimation of the parameters $[\varepsilon_f, \mu_{\text{cm}}(E, E_f, v, z), (\gamma_{k, E_f})_{k \in [1, K]}]$ through the knowledge of $\mu_k(E)_{k \in [1, K]}$ must be measured under the same conditions, the same matrix, and the same experimental setup as the measurements made during the sparse excitation energy XAS experiment. This is in the best cases unpractical and in most cases unfeasible.

However, the X-ray absorption spectra on an energy domain allowing to probe an electron edge bear specificity that will allow us to relax these constraints.

2.2.7. Approximate affine behaviour of the matrix. First, considering the matrix alone in the typical energy domain probed during an XAS experiment, we approximate the absorption coefficient using a first-order Taylor expansion leading to an affine function. This holds as long as no constituent of the matrix has a fluorescence edge in the domain,

$$\exists (\mathcal{A}_m, \mathcal{B}_m) \in \mathbb{R}^2 / \mu_m(E) \simeq \mathcal{B}_m + \mathcal{A}_m \delta E, \quad (11)$$

$$\text{for } \delta E = E - E_f,$$

$$\text{with } |\mathcal{A}_m \delta E| \ll \mathcal{B}_m. \quad (12)$$

In this setting we can approximate the value of $\mu_{\text{cm}}(E, E_f, v, z)$,

$$\mu_{\text{cm}}(E, E_f, v, z) = \mu_m(E, v, z) + \tan \theta \mu_m(E_f, v, z)$$

$$\simeq \mathcal{B}_m(v, z) + \mathcal{A}_m(v, z) \delta E + \tan \theta \mathcal{B}_m(v, z),$$

$$\text{(as } \delta E_f = 0 \text{ by definition)} \quad (13)$$

$$\simeq (1 + \tan \theta) \mathcal{B}_m(v, z) + \mathcal{A}_m(v, z) \delta E$$

$$\simeq (1 + \tan \theta) \mathcal{B}_m(v, z) \text{ since } |\mathcal{A}_m \delta E| \ll \mathcal{B}_m,$$

basically stating that $\mu_{\text{cm}}(E, E_f, v, z)$ has very small variation when E varies within a small energy domain.

2.2.8. Linear composition of near-edge absorption spectra. The contribution of the different electronic levels of an atom to the absorption signals that correspond to different absorption edges is composed by simple addition of the effects at each of the edges taken individually. Thus the approximation described above is valid and usable for the contribution of the absorption of unprobed edges, and the absorption coefficient of the K species probed can be written as the sum of an affine contribution and a contribution representing the very edge probed for each species, measured through fluorescence signal, $\gamma_{k, E_f} \mu_k^{\text{edge}}(E)$,

$$\forall E, \quad \gamma_{k, E_f} \mu_k(E) \simeq \mathcal{B}^{\text{edge}} + \mathcal{A}^{\text{edge}} \delta E + \gamma_{E_f} \mu_k^{\text{edge}}(E), \quad (14)$$

with

$$\forall k \in [1, K], \quad E \ll E^{\text{edge}} \Rightarrow \mu_k^{\text{edge}}(E) = 0 \quad (15)$$

$$E \gg E^{\text{edge}} \Rightarrow \mu_k^{\text{edge}}(E) = \mu^{\text{edge}}(E)$$

$$\simeq b^{\text{edge}} + a^{\text{edge}} \delta E.$$

In equation (15), the term $a^{\text{edge}} \delta E$ corresponds to the out-of-edge change in slope, *i.e.* the slope correction applicable to equation (14) for the part of the spectrum collected at $E > E^{\text{edge}}$. If $|a^{\text{edge}}| \ll |\mathcal{A}^{\text{edge}}|$, one can, on a small domain that does not extend too much after the edge energy E^{edge} , neglect a^{edge} and therefore write

$$\forall E, \gamma_{k,E_i} \mu_k(E) \simeq \mathcal{B}^{\text{edge}} + \mathcal{A}^{\text{edge}} \delta E + \gamma_{E_i} b^{\text{edge}} \widehat{\mu}_k(E), \quad (16)$$

with

$$\begin{aligned} \forall k \in [1, K], E \ll E^{\text{edge}} &\Rightarrow \widehat{\mu}_k(E) = 0, \\ E \gg E^{\text{edge}} &\Rightarrow \widehat{\mu}_k(E) = 1. \end{aligned}$$

Equation (16) defines $[\widehat{\mu}_k(E)]_{k \in [1, K]}$, the normalized spectra of the K reference species at the probed edge, which under the current approximations are unaffected by the matrix used to measure the reference spectra, as long they are measured in fluorescence mode at energy E_f . Indeed, under the current approximations the dependence of the measurement to the matrix is encoded within $\mathcal{A}^{\text{edge}}$ and $\mathcal{B}^{\text{edge}}$ which in turn are not affected by the species of the probed element, whilst b^{edge} depends only on the edge probed. In other words, equation (16) allows us to model the reference spectra using species-independent terms (parameters: $\mathcal{B}^{\text{edge}}$, $\mathcal{A}^{\text{edge}}$), a matrix-independent term $\{[\widehat{\mu}_k(E)]_{k \in [1, K]}\}$, and a term independent of both ($\gamma_{E_i} b^{\text{edge}}$). In this context, $[\widehat{\mu}_k(E)]_{k \in [1, K]}$ can be obtained by measuring the reference species in any matrix and at any concentration provided that the approximation of the diluted sample holds, as usually done.

Note finally that the affine absorption behaviour of the matrix, characterized in equation (11) by A_m and B_m , is accounted for with the parameters $\mathcal{A}^{\text{edge}}$ and $\mathcal{B}^{\text{edge}}$.

2.2.9. Linear combination of reference spectra. We may now rewrite equation (10) using the approximation from equation (13) for $\mu_{\text{cm}}(E, E_f, v, z)$ and equation (16) for $[\mu_k(E)]_{k \in [1, K]}$,

$$\begin{aligned} \frac{I_f^{\text{Exp}} f(E, v, z)}{I_0[t(E, v, z)]} &= \frac{\varepsilon_f \Omega}{4\pi} \frac{1}{(1 + \tan \theta) \mathcal{B}_m(v, z)} \\ &\times \left[(\mathcal{B}^{\text{edge}} + \mathcal{A}^{\text{edge}} \delta E) \left(\sum_{k=1}^K C_k(v, z) \right) \right. \\ &\left. + \gamma_{E_f} b^{\text{edge}} \left(\sum_{k=1}^K C_k(v, z) \widehat{\mu}_k(E) \right) \right], \quad (17) \end{aligned}$$

which states that for a given (v, z) position the measured signal, $I_f^{\text{Exp}} f(E, v, z)/I_0[t(E, v, z)]$ is an affine function of $[\widehat{\mu}_k(E)]_{k \in [1, K]}$ and δE with coefficients:

$C_0(v, z) = D_m(E_f, v, z) \mathcal{B}^{\text{edge}} [\sum_{k=1}^K C_k(v, z)]$ for the intercept;

$C_{\delta E}(v, z) = D_m(E_f, v, z) \mathcal{A}^{\text{edge}} [\sum_{k=1}^K C_k(v, z)]$ for the coefficient of δE ;

$C_k(v, z) = [D_m(E_f, v, z) \gamma_{E_f} b^{\text{edge}} C_k(v, z)]_{k \in [1, K]}$ for the coefficient of each of $[\widehat{\mu}_k(E)]_{k \in [1, K]}$;

where $D_m(E_f, v, z) = [(\varepsilon_f \Omega)/(4\pi)] \{1/[(1 + \tan \theta) \mathcal{B}_m(v, z)]\}$ is independent of the incident energy E .

For any given coordinate set (v, z) on the sample, one can estimate the above expressed coefficient, $\{C_0(v, z), C_{\delta E}(v, z), [C_k(v, z)]_{k \in [1, K]}\}$, by performing a linear least-squares fitting of $(N \geq K + 2)$ measurements corresponding to energies $(E_n)_{n \in [1, N]}$ such that the vector $[\widehat{\mu}_k(E_1), \dots, \widehat{\mu}_k(E_N)]_{k \in [1, K]}$ and $(\delta E_1, \dots, \delta E_N)$ are linearly

independent. The estimated coefficients $[C_k(v, z)]_{k \in [1, K]}$ correspond indeed to the relative concentration of each of the K species that could be present in the samples at coordinates (v, z) hence providing a relative estimation of $[C_k(v, z)]_{k \in [1, K]}$.

2.2.10. Further notational simplification. Since the absolute scale of equation (17) cannot be easily computed from information one usually has *a priori* on the system, there is no incentive to keep the full notations. We can simplify the notation by integrating the scaling factors within the parameters without altering the relative values of $C_k(v, z)$ taken from different pixels. For pixel (v, z) the scale is directly integrated in $C_k(v, z)$, while the slope and intersect of the affine baseline will be noted as $A(v, z)$ and $B(v, z)$ which leads to equation (20) in the main text.

2.2.11. Validation of the model and estimation of the relative accuracy. Under the model developed above, as long as the matrix of the sample is homogeneous, the estimated coefficients $C_0(v, z)$ and $C_{\delta E}(v, z)$ should be strictly correlated, within measurement noise error, as they both should be correlated to the estimated sum of relative concentrations, $C^{\text{sum}}(v, z) = \sum_{k=1}^K C_k(v, z)$. Indeed, such correlations are very unlikely to be present by chance and the quality of the correlations can be used as a marker of the validity of the proposed model. Furthermore, as the proposed method is used to perform speciation imaging, hence collecting data on a large set of (v, z) coordinates, these correlations can be verified on the basis of a significant number of data points.

Note, however, that if the hypothesis of the homogeneity of the matrix is no longer met, only the $C_0(v, z)$ and $C_{\delta E}(v, z)$ estimation are affected since only $\mathcal{A}^{\text{edge}}$ and $\mathcal{B}^{\text{edge}}$ depend on the matrix. In these circumstances $\mathcal{A}^{\text{edge}}$ and $\mathcal{B}^{\text{edge}}$ become functions of the probed coordinates, (v, z) , and the correlations that are described above no longer stand. Then $C^{\text{sum}}(v, z)$ is the only available estimate of the relative concentration of the probed element at the different measured coordinates.

Finally, within the framework of linear least squares, one can compute the condition number, κ , of the design matrix as soon as $[\widehat{\mu}_k(E_1), \dots, \widehat{\mu}_k(E_N)]_{k \in [1, K]}$ and $(\delta E_1, \dots, \delta E_N)$ are known, and before performing any measure on the actual sample. With κ known it is easy to obtain an upper bound on relative errors in relative concentrations, ϵ_{C_k} , given a level of relative errors, also termed signal-to-noise ratio, in fluorescence signal measurements,

$$\epsilon_{C_k} \leq \kappa \epsilon_{\text{fluorescence}}. \quad (18)$$

Through this simple inequality, which can be computed very early in the experiment design, one can set the appropriate exposure time to reach the fluorescence signal-to-noise ratio, $\epsilon_{\text{fluorescence}}$, so as to reach a predefined requirement in terms of accuracy of the relative species concentrations.

Under such conditions, the experimental XAS spectrum at any given pixel of coordinates (v, z) results from a linear combination of the K reference spectra $[\mu_k(E)]_{k \in [1, K]}$ and of an affine baseline,

$$\frac{I_f^{\text{Exp}}(E, v, z)}{I_0[t(E, v, z)]} \simeq \frac{\Gamma}{\mu_{\text{cm}}(v, z)} \left\{ \left[\sum_{k=1}^K C_k(v, z) dV \mu_k(E) \right] + \mathcal{A}(v, z) E + \mathcal{B}(v, z) \right\} + \epsilon(E, v, z), \quad (19)$$

where $I_f^{\text{Exp}}(E, v, z)$ is the fluorescence signal measured when illuminating pixel (v, z) at an excitation energy E , and $I_0[t(E, v, z)]$ is the incident beam intensity at the moment $t(E, v, z)$, when $I_f^{\text{Exp}}(E, v, z)$ is measured. $\mathcal{A}(v, z) E + \mathcal{B}(v, z)$ models the summed up slow varying background from XAS spectra of the sample and of the references, and $C_k(v, z) dV$ is the quantity of the species k in the elementary volume dV probed. Γ is a proportionality constant depending on the XRF yield of the probed element and on the geometry of the setup (solid angle of the detector, *etc.*), $\mu_{\text{cm}}(v, z)$ is the apparent absorption coefficient of the matrix, accounting for both incident beam and fluoresced signal attenuation, at pixel (v, z) , and $\epsilon(E, v, z)$ is a zero mean stochastic error term which is independent from one measurement to the next.

The parameters $[C_k(v, z)]_{k \in [1, K]}$, $\mathcal{A}(v, z)$ and $\mathcal{B}(v, z)$ can be estimated using the linear least-squares fit of the equation system to the measurements. With the previous assumption of a zero-mean $\epsilon(E, v, z)$, these estimates are unbiased; with the additional assumption of independence, constant variance, and Gaussian (normal) distribution, they correspond to maximum-likelihood estimation. From now, we will simplify the notations by keeping the error term out of all expressions. Note that by normalizing the reference spectra, $[\widehat{\mu}_k(E)]_{k \in [1, K]}$, we bring C_k on an arbitrary unit and as such there is not much interest in obtaining C_k rather than $C_k \propto [\Gamma/\mu_{\text{cm}}(v, z)] C_k dV$. Finally the very limited dependence that $\mu_{\text{cm}}(v, z)$ has in terms of E can be accounted for in the affine baseline [equation (13)], thus we can simplify our equation to

$$\frac{I_f^{\text{Exp}}(E, v, z)}{I_0[t(E, v, z)]} \simeq \left[\sum_{k=1}^K C_k(v, z) \widehat{\mu}_k(E) \right] + A(v, z) E + B(v, z). \quad (20)$$

For a given (v, z) pixel position and using N probing energies $(E_n)_{n \in [1, N]}$, equation (20) produces a linear system of N equations and $K + 2$ unknowns $[C_k(v, z), A(v, z)$ and $B(v, z)]$.

The linear least-squares solution requires that the number of equations is at least equal to the number of unknowns, $N \geq K + 2$.

Interestingly, when all needed references are identified, $\sum_{k=1}^K C_k(v, z) = C^{\text{sum}}(v, z)$ is proportional to the absorbing element concentration. This provides an elemental *relative concentration* map as a side result of the full procedure.

3. Open source software implementation

The XAS procedure described here has been integrated into the most recent versions of the *MicroAnalysis Toolkit* package (Webb, 2011) under the ‘XANES Fitting’ analysis option. The latter package utilizes a combination of self-written routines in

Python (VanRossum & Drake, 2010) with the *SciPy* (Jones *et al.*, 2001) package used for minimization. Compiled versions and source code for the software updated with these routines is made available as of version 1.5 under an open source licence in parallel to this publication (Webb, 2010). File formats from many synchrotron sources worldwide are able to load, and the list of compatible data formats is expanding upon request. The program is freely available and can be run on multiple operating system platforms.

In addition, a specific script was developed under the *R* statistical environment (R Development Core Team, 2011) and is made freely available with this publication. The script is annotated so as to provide a step-by-step description of the data processing, in such a way that it is easy for a user to adapt it to its own data. The script starts from raw XRF data and covers all steps up to decomposition and statistical analysis. In particular, several diagnosis steps are present in this script, as a contribution to data validation, while not all described in this manuscript.

The script and all datasets needed to reproduce the presented results can be found at <https://doi.org/10.5281/zenodo.3748752>.

4. Application to a fossil sample

The proposed procedure was tested in a study of the distribution of the oxidation states of cerium in a paleontological sample at micrometric scales, as a typical heterogeneous sample containing an element with two distinct valencies.

An exception among rare-earth elements, the element cerium exists in both the +III and +IV oxidation states. The resulting sensitivity to redox conditions makes the oxidation state of cerium a frequently used proxy of the chemistry of past environments (German & Elderfield, 1990). Synchrotron XAS mapping allows studying directly the speciation of cerium within ancient samples (such as sediment or bone; Gueriau *et al.*, 2015). Yet, a significant difficulty is that a high radiation dose can result in the photo-oxidation of Ce(III) to Ce(IV) during collection of entire XANES spectra, thereby biasing the measured data (Gueriau *et al.*, 2015).

We studied a loose fragment (not a cross section) of well preserved soft tissues (mineralized muscles) around a bone of a yet undescribed ?gonorynchiform fish (where ‘?’ indicates that the supposed identification has not yet been corroborated by a complete anatomical description) coming from the Late Cretaceous levels of the Djebel Oum Tkout locality in Morocco [*ca.* 95 million years old, Figs. 2(a) and 2(b)]. We produced rapid maps of cerium speciation of the sample at the Lucia beamline of the SOLEIL synchrotron source (Vantelon *et al.*, 2016) over a 6466 pixel area (120 $\mu\text{m} \times 210 \mu\text{m}$ with a scan step of 2 $\mu\text{m} \times 2 \mu\text{m}$; beam size 3.0 $\mu\text{m} \times 2.5 \mu\text{m}$ (H \times V); flux: 2.8×10^{10} photons s^{-1} in the focused spot) using the sparse excitation energy procedure presented above. Spectra of cerium III and IV reference compounds were collected using the procedure given by Gueriau *et al.* (2015).

Gueriau *et al.* (2015) studied a series of cross sections from well preserved fossil teleost fishes and crustaceans from the

same locality [see Fig. 4 of Gueriau *et al.* (2015)] using a limited number of excitation energy points to limit photo-oxidation changes they observed within a few tens seconds during XAS acquisition. They selected six excitation energies: below the L_3 absorption edge of cerium (5680 eV), in the post-edge region (5765 and 5900 eV) and at energies corresponding to three main spectral features that allow distinguishing Ce(III) from Ce(IV) in cerium-containing compounds – a white line at 5726 eV resulting from the $2p \rightarrow (4f^1)5d^*$ electron transition and two peaks that partly overlap with the former attributed to the $2p_{3/2} \rightarrow (4f^1L)5d^*$ and the $2p_{3/2} \rightarrow (4f^0)5d^*$ transitions, where L denotes a hole in ligand orbitals (at 5733 and 5736 eV, respectively) (Takahashi *et al.*, 2000).

In the loose soft tissue fragment of the ?gonorynchiform fish, the procedure developed here (using the same excitation energies) reveals the presence of a fraction of cerium in the redox +IV state $\{[Ce(IV)]/[Ce(tot)] = 25.6 \pm 8.3\%$ over the whole mapped area, suggesting slightly oxidative local conditions of burial [Figs. 2(c)–2(f)]. This value agrees well with the indirect data obtained from micro- and macroscale elemental concentration profiles measured at the site (so-called ‘REE patterns’), which also display a slight negative cerium anomaly attributed to oxidative burial conditions (Gueriau *et al.*, 2015). The total cerium concentration, $[Ce(tot)]$, was mapped in the sample using two different approaches. The cerium distribution shown in Fig. 2(d) is the result of the usual processing of an XRF dataset collected

after the cerium edge (5900 eV), while Fig. 2(e) shows the $[Ce(tot)]$ map derived from the proposed XAS procedure. Since the latter takes advantage of the total collection time of the several XRF maps at different energies, it led to a far higher signal-to-noise ratio than the single-energy XRF map, without the need to switch to an additional incident energy significantly above the absorption edge associated with the element. While the total concentration of cerium differs significantly between bone and surrounding mineralized muscles, the distribution of the Ce(IV) fraction relative to total cerium appears very stable and homogeneous between bone ($27.2 \pm 2.9\%$) and the surrounding mineralized muscles [$27.7 \pm 5.2\%$, Fig. 2(g)], which corroborates the validity of the redox proxy at different length scales in the presence of varying cerium ion concentrations.

A dwell time of 1 s per pixel was used. The dead-times associated with the movements of the scanning motors led to a total collection time of about 165 min per excitation energy, *i.e.* about 16.5 h for the full dataset (collection at six excitation energies). The typical acquisition time for a single full XAS spectrum was 243 s, dead-times included: the collection of the same dataset with a complete XAS spectrum per pixel would therefore have taken 436 h, *i.e.* about 26 times longer. While the collection of a full XAS spectrum under these conditions would have resulted in a *surface* irradiation dose (Howells *et al.*, 2009; Bertrand *et al.*, 2015) of 5.3 GGy, this value was reduced by a similar factor of 26 to

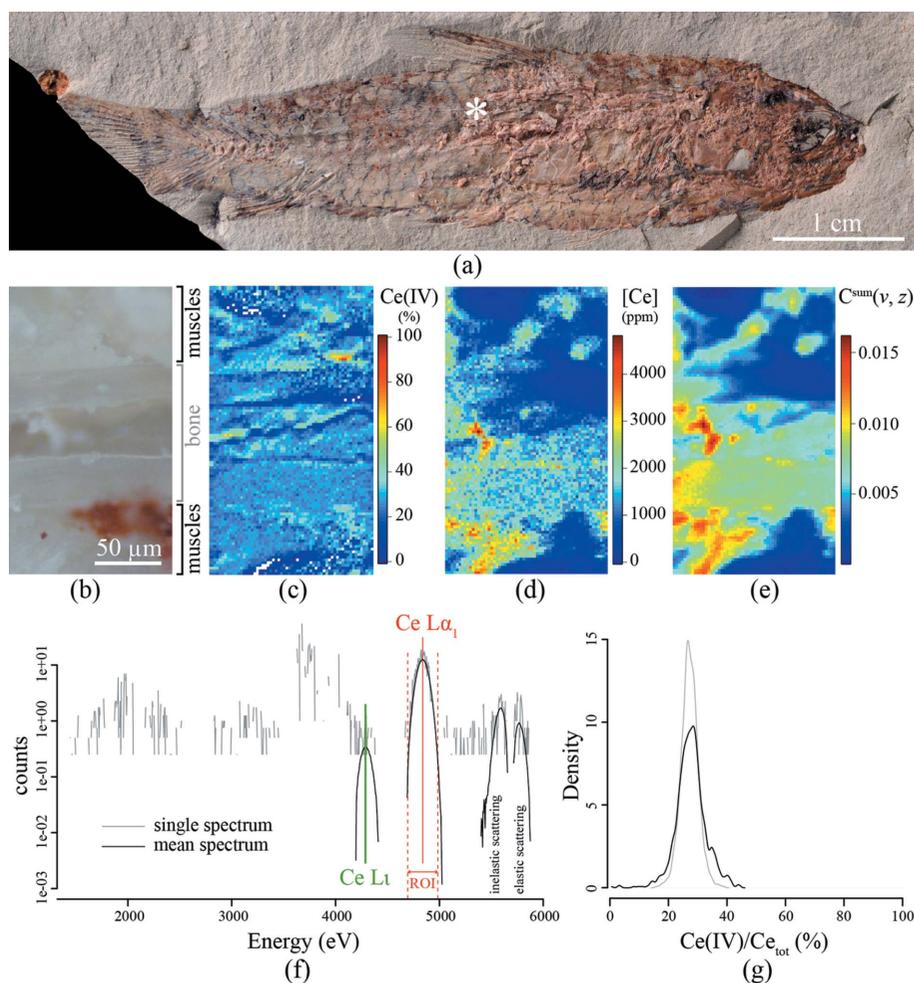


Figure 2 Mapping of the two speciation states of cerium in a well preserved Cretaceous teleost ?gonorynchiform fossil fish from Morocco (MHNM-KK-OT 10, ~100 million years old). (a) Photograph of the fossil. (b) Light microscopy image of the spot denoted by a star in (a). (c) Corresponding $[Ce(IV)]/[Ce(tot)]$ ratio (scan step: $2 \mu\text{m} \times 2 \mu\text{m}$, 6466 pixels). (d) Total cerium content $[Ce(tot)]$ determined from post-edge spectral decomposition (5900 eV). (e) $C^{sum}(v, z)$ map. Note the similarity and improved signal-to-noise with respect to the cerium map in (d). (f) μXRF spectrum from a single spot and mean spectrum from the full map, calculated from subtraction of the spectra collected at 5726 eV from that collected before the edge at 5680 eV. The difference spectrum reveals the specific contribution of the excited fluorescence lines from cerium. (g) Kernel density estimates of the $[Ce(IV)]/[Ce(tot)]$ distribution from areas corresponding, respectively, to fossilized muscles (black curve; 915 pixels, bandwidth: 0.009) and to bone (grey curve; 915 pixels, bandwidth: 0.006). Data collected at the Lucia beamline of the SOLEIL synchrotron.

about 200 MGy per pixel, with a counting time of 9.2 s per pixel, significantly less than the dose leading to a change in speciation, as shown by the photo-oxidation kinetics of cerium in comparable specimens (Gueriau *et al.*, 2015, Fig. 3C).

The sparse excitation energy XAS approach produced maps of cerium speciation and total concentration allowing comparison of different physiological zones and showing homogeneous speciation, even if the concentration of the trace element varies considerably on the map.

5. Conclusion

X-ray absorption imaging can be used to establish chemical speciation at sub-micrometric to centimetric spatial scales over large sample areas. However, as observed in many samples, including the example discussed here of cerium speciation in fossil paleontological materials, the traditional way of collecting XAS spectra, which requires the accumulation of data for sometimes several minutes per point, can create changes in the redox state of the sample at the points of analysis. With the increased brightness of synchrotron X-ray sources, this approach leads to the deposition of large radiation doses in small sample volumes. Such alterations can occur in many samples and present a significant risk to the quality of the measurement. The procedure presented here with the implementation of the provided software allows to produce speciation and total elemental composition maps by collecting several individual XRF maps across the absorption edge of the element of interest at relatively low total irradiation doses. Although similar procedures have been reported many times in the past, we show with a rigorous framework that the determination of proportions of individual species can be solved accurately, in particular by making explicit the typical assumptions and conditions used when collecting XRF data on thick samples. A particular advantage of this low-dose XAS imaging approach, that we propose to denominate as see-XAS (sparse excitation energy XAS), is that not only can speciation of the sample be obtained at thousands or millions of micro-scale locations to reconstruct high-resolution images of the speciation but also that the actual dose in a given pixel is much lower (up to several orders of magnitude) owing to the selective nature of the number of energies collected. This constitutes a true risk mitigation procedure that can be invaluable for radiation-sensitive samples in the fields of cultural heritage, environment, life and materials science.

Acknowledgements

PG thanks D. B. Dutheil, S. Charbonnier, G. Clément, N.-E. Jalil, H. Bourget, B. Khalloufi (MNHN), A. Tourani, F. Khaldoune (Cadi Ayyad University, Marrakesh) and P. M. Brito (Rio de Janeiro State University) for collecting the fossils. E. Janots (Université Grenoble Alpes, Grenoble) is warmly acknowledged for providing us Ce(III) and Ce(IV) standard XANES data. We thank D. Vantelon and N. Trcera (SOLEIL) for their support at the Lucia beamline.

Funding information

SW acknowledges support from the Jean d'Alembert fellowship program at Université Paris-Saclay. Study of heritage materials at IPANEMA and the SOLEIL synchrotron was supported by the Research Infrastructures activity IPERION CH of the Horizon2020 program of the European Commission (Grant Agreement No. 654028). Beam time was funded by the SOLEIL synchrotron under project no. 20141143. Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under contract no. DE-AC02-76SF00515.

References

- Ade, H., Zhang, X., Cameron, S., Costello, C., Kirz, J. & Williams, S. (1992). *Science*, **258**, 972–975.
- Bertrand, L., Cotte, M., Stampanoni, M., Thoury, M., Marone, F. & Schöder, S. (2012). *Phys. Rep.* **519**, 51–96.
- Bertrand, L., Schöeder, S., Anglos, D., Breese, M. B. H., Janssens, K., Moïni, M. & Simon, A. (2015). *TrAC Trends Anal. Chem.* **66**, 128–145.
- Bertsch, P. M. & Hunter, D. B. (2001). *Chem. Rev.* **101**, 1809–1842.
- Collins, B. A. & Ade, H. (2012). *J. Electron Spectrosc. Relat. Phenom.* **185**, 119–128.
- Cotte, M., Susini, J., Dik, J. & Janssens, K. (2010). *Acc. Chem. Res.* **43**, 705–714.
- De Andrade, V., Susini, J., Salomé, M., Beraldin, O., Rigault, C., Heymes, T., Lewin, E. & Vidal, O. (2011). *Anal. Chem.* **83**, 4220–4227.
- Edwards, N., Manning, P., Bergmann, U., Larson, P., van Dongen, B., Sellers, W., Webb, S., Sokaras, D., Alonso-Mori, R., Ignatyev, K., Barden, H. E., van Veelen, A., Anné, J., Egerton, V. M. & Wogelius, R. A. (2014). *Metallomics*, **6**, 774–782.
- Etschmann, B. E., Donner, E., Brugger, J., Howard, D. L., de Jonge, M. D., Paterson, D., Naidu, R., Scheckel, K. G., Ryan, C. G. & Lombi, E. (2014). *Environ. Chem.* **11**, 341–350.
- Farfan, G., Apprill, A., Webb, S. M. & Hansel, C. M. (2018). *Anal. Chem.* **90**, 12559–12566.
- Fayard, B., Pouyet, E., Berruyer, G., Bugnazet, D., Cornu, C., Cotte, M., Andrade, V. D., Chiaro, F. D., Hignette, O., Kieffer, J., Martin, T., Papillon, E., Salomé, M. & Sole, V. A. (2013). *J. Phys. Conf. Ser.* **425**, 192001.
- Flynn, G., Keller, L., Jacobsen, C. & Wirrick, S. (1999). *Meteorit. Planet. Sci. Suppl.* **34**, A36.
- Gardeatorresdey, J., Peraltavidea, J., Delarosa, G. & Parsons, J. (2005). *Coord. Chem. Rev.* **249**, 1797–1810.
- German, C. R. & Elderfield, H. (1990). *Paleoceanography*, **5**, 823–833.
- Gueriau, P., Bernard, S. & Bertrand, L. (2016). *Elements*, **12**, 45–50.
- Gueriau, P., Mocuta, C. & Bertrand, L. (2015). *Anal. Chem.* **87**, 8827–8836.
- Henderson, G. S., de Groot, F. M. & Moulton, B. J. (2014). *Rev. Mineral. Geochem.* **78**, 75–138.
- Holt, M., Harder, R., Winarski, R. & Rose, V. (2013). *Annu. Rev. Mater. Res.* **43**, 183–211.
- Howells, M., Beetz, T., Chapman, H., Cui, C., Holton, J., Jacobsen, C., Kirz, J., Lima, E., Marchesini, S., Miao, H., Sayre, D., Shapiro, D., Spence, J. & Starodub, D. (2009). *J. Electron Spectrosc. Relat. Phenom.* **170**, 4–12.
- Jacobsen, C., Wirrick, S., Flynn, G. & Zimba, C. (2000). *J. Microsc.* **197**, 173–184.
- Jones, E., Oliphant, T., Peterson, P. *et al.* (2001). *SciPy: Open source scientific tools for Python*, <http://www.scipy.org/>.
- Kaulich, B., Thibault, P., Gianoncelli, A. & Kiskinova, M. (2011). *J. Phys. Condens. Matter*, **23**, 083002.

- Kinney, J., Johnson, Q., Nichols, M. C., Bonse, U. & Nusshardt, R. (1986). *Appl. Opt.* **25**, 4583–4585.
- Lerotic, M., Jacobsen, C., Schäfer, T. & Vogt, S. (2004). *Ultramicroscopy*, **100**, 35–57.
- Li, L., Chen-Wiegart, Y., Wang, J., Gao, P., Ding, Q., Yu, Y. S., Wang, F., Cabana, J., Wang, J. & Jin, S. (2015). *Nat. Commun.* **6**, 6883.
- Linge, K. L. (2008). *Crit. Rev. Environ. Sci. Technol.* **38**, 165–196.
- Liu, Y., Meirer, F., Wang, J., Requena, G., Williams, P., Nelson, J., Mehta, A., Andrews, J. C. & Pianetta, P. (2012a). *Anal. Bioanal. Chem.* **404**, 1297–1301.
- Liu, Y., Meirer, F., Williams, P. A., Wang, J., Andrews, J. C. & Pianetta, P. (2012b). *J. Synchrotron Rad.* **19**, 281–287.
- Lombi, E., de Jonge, M. D., Donner, E., Ryan, C. G. & Paterson, D. (2011). *Anal. Bioanal. Chem.* **400**, 1637–1644.
- Marcus, M. A. (2010). *TrAC Trends Anal. Chem.* **29**, 508–517.
- Mayhew, L., Webb, S. & Templeton, A. (2011). *Environ. Sci. Technol.* **45**, 4468–4474.
- Meirer, F., Cabana, J., Liu, Y., Mehta, A., Andrews, J. C. & Pianetta, P. (2011). *J. Synchrotron Rad.* **18**, 773–781.
- Nelson, G. J., Harris, W. M., Izzo, J. R. Jr, Grew, K. N., Chiu, W. K., Chu, Y. S., Yi, J., Andrews, J. C., Liu, Y. & Pianetta, P. (2011). *Appl. Phys. Lett.* **98**, 173109.
- Oram, L. L., Strawn, D. G. & Möller, G. (2011). *Environ. Sci. Technol.* **45**, 870–875.
- Pickering, I. J., Prince, R. C., Salt, D. E. & George, G. N. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 10717–10722.
- Pushie, M. J., Pickering, I. J., Korbas, M., Hackett, M. J. & George, G. N. (2014). *Chem. Rev.* **114**, 8499–8541.
- Rau, C., Somogyi, A., Bytchkov, A. & Simionovici, A. S. (2002). *Proc. SPIE*, **4503**, 249–255.
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Sciau, P. & Wang, T. (2019). *Synchrotron Radiation*, ch. 3. IntechOpen.
- Sutton, S., Bajt, S., Delaney, J., Schulze, D. & Tokunaga, T. (1995). *Rev. Sci. Instrum.* **66**, 1464–1467.
- Sutton, S. R., Bertsch, P. M., Newville, M., Rivers, M., Lanzirotti, A. & Eng, P. (2002). *Rev. Mineral. Geochem.* **49**, 429–483.
- Tack, P., Garrevoet, J., Bauters, S., Vekemans, B., Laforce, B., Van Ranst, E., Banerjee, D., Longo, A., Bras, W. & Vincze, L. (2014). *Anal. Chem.* **86**, 8791–8797.
- Takahashi, Y., Shimizu, H., Kagi, H., Yoshida, H., Usui, A. & Nomura, M. (2000). *Earth Planet. Sci. Lett.* **182**, 201–207.
- VanRossum, G. & Drake, F. L. (2010). *The Python language reference*. Python software foundation Amsterdam, The Netherlands.
- Vantelon, D., Trcera, N., Roy, D., Moreno, T., Mailly, D., Guilet, S., Metchalkov, E., Delmotte, F., Lassalle, B., Lagarde, P. & Flank, A.-M. (2016). *J. Synchrotron Rad.* **23**, 635–640.
- Wang, J., Chen-Wiegart, Y. & Wang, J. (2013). *Chem. Commun.* **49**, 6480–6482.
- Webb, S. (2010). *Sam's X-rays – Applications for Synchrotrons*, <https://www.sams-xrays.com/>.
- Webb, S. (2011). *AIP Conf. Proc.* **1365**, 196–199.
- Zhang, X., Ade, H., Jacobsen, C., Kirz, J., Lindaas, S., Williams, S. & Wirick, S. (1994). *Nucl. Instrum. Methods Phys. Res. A*, **347**, 431–435.