Available online at www.sciencedirect.com

SciVerse ScienceDirect

Geochimica et Cosmochimica Acta 99 (2012) 206–223

Geochimica et Cosmochimica Acta

www.elsevier.com/locate/gca

Quantitative analysis of sulfur functional groups in natural organic matter by XANES spectroscopy

Alain Manceau [a,*], Kathryn L. Nagy [b]

[a] ISTerre, CNRS and Université de Grenoble 1, F-38041 Grenoble Cedex 9, France

[b] Department of Earth and Environmental Sciences, MC-186, 845 West Taylor Street, University of Illinois at Chicago, Chicago, IL 60607, United States

Received 6 April 2012; accepted in revised form 19 September 2012; Available online 28 September 2012

Abstract

Two new approaches to quantify sulfur functionalities in natural organic matter from S K-edge XANES spectroscopy are presented. In the first, the K-edge spectrum is decomposed into Gaussian and two arctangent functions, as in the usual Gaussian curve fitting (GCF) method, but the applicability of the model is improved by a rigorous simulation procedure that constrains the model-fit to converge toward chemically and physically realistic values. Fractions of each type of functionality are obtained after spectral decomposition by correcting Gaussian areas for the change in X-ray absorption cross-section with increasing oxidation state. This correction is made using published calibration curves and a new curve obtained in this study. Calibration-induced errors, inherent to the choice of a particular curve, are typically lower than 5% of total sulfur for oxidized species (e.g., sulfate), may reach 10% for organic reduced sulfur, and may be as high as 30–40% for inorganic reduced sulfur. A generic curve, which reduces the calibration-induced uncertainty by a factor of two on data collected to avoid X-ray over-absorption, is derived.

In the second analytical scheme, the K-edge spectrum is partitioned into a weighted sum of component species, as in the usual linear combination fitting (LCF) method, but is fit to an extended database of reference spectra under the constraint of non-negativity in the loadings (Combo fit). The fraction of each sulfur functionality is taken as the sum of all positive fractions of references with similar oxidation state of sulfur.

The two proposed methods are applied to eight humic and fulvic acids from the International Humic Substances Society (IHSS). The nature and fractions of sulfur functionalities obtained by the two analytical approaches are consistent with each other. The accuracy of the derived values, expressed as the difference in values of a fraction obtained on the same material by the two independent methods, is on average $4.5 \pm 3.0\%$ of total sulfur for exocyclic reduced sulfur, $4.1 \pm 2.1\%$ for heterocyclic reduced sulfur, and $1.6 \pm 1.4\%$ for sulfate. Total reduced sulfur has a better accuracy of estimation ($2.4 \pm 1.6\%$) than either exocyclic and heterocyclic sulfur, because the errors on the two reduced pools have opposite sign. Experimental difficulties and uncertainties of the results associated with the analysis of concentrated and heterogeneous samples are discussed. The spectra of the IHSS materials and the reference compounds are made available as an open source for interlaboratory testing.
© 2012 Elsevier Ltd. All rights reserved.

1. INTRODUCTION

Sulfur is present in natural organic matter (NOM) in a range of oxidation states from −2 to +6 and numerous chemical forms that are difficult to characterize. Among them, reduced organic sulfur species with electronic oxidation states (EOS) (Frank et al., 1987; Waldo et al., 1991; Vairavamurthy, 1998) between 0 and 1 form a major class of compounds that is important in controlling metal speciation and mobility in aqueous and soil environments. These compounds include thiol moeities (R-SH) which display an affinity for soft Lewis acids such as Hg(II), Cd(II), Pb(II),

* Corresponding author.
E-mail address: alain.manceau@ujf-grenoble.fr (A. Manceau).

0016-7037/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.gca.2012.09.033

Ag(I), and Cu(I) (Smith et al., 2002). Improvements in quantifying as precisely as possible the nature and abundance of each oxidation state and the main chemical forms of sulfur are essential to advance knowledge of the reactivity of NOM.

Sulfur speciation in natural organic matter is commonly determined from the variation in energy of the s → p electronic transition for different types of functional groups with distinct oxidation states as measured by X-ray absorption near-edge structure (XANES) spectroscopy (George and Gorbaty, 1989; George et al., 1991; Huffman et al., 1991; Waldo et al., 1991; Taghiei et al., 1992; Vairavamurthy et al., 1993, 1994; Kasrai et al., 1994, 1996; Morra et al., 1997; Pickering et al., 1998; Xia et al., 1998; Frank et al., 1999; Hundal et al., 2000; Skyllberg et al., 2000; Hutchison et al., 2001; Szulczewski et al., 2001; Beauchemin et al., 2002, 2004; Myneni, 2002; Olivella et al., 2002a,b; Qian et al., 2002; Sarret et al., 2002; Solomon et al., 2003, 2005; Bostick et al., 2005; Schäfer et al., 2005; Wiltfong et al., 2005; Zhao et al., 2006; Einsiedl et al., 2007; Kelemen et al., 2007, 2012; Pattanaik et al., 2007; Prietzel et al., 2007, 2009; Schroth et al., 2007; Cai et al., 2009; Orthous-Daunay et al., 2010; Nagy et al., 2011). The desirable attributes of this technique are (1) a mass sensitivity of about one part per ten thousand (100 ppm or 100 mg S/kg NOM), (2) good chemical sensitivity based on an edge shift of up to 12 eV with a change in the oxidation state of sulfur from −2 to +6, (3) the capability to quantify the sulfur forms in a homogeneous matrix without any special preparation or destruction of the sample, and (4) the possibility of analyzing heterogeneous materials in their natural state using micro-scale or nanoscale X-ray probes (Manceau et al., 2002).

However, as with any experimental technique, care must be taken to minimize errors during data acquisition and analysis to obtain meaningful values from S-XANES measurements (Jalilehvand, 2006). The main experimental problem of "overabsorption" (Manceau et al., 2002), also called "self-absorption" (Tröger et al., 1992), occurs with data recorded in fluorescence-yield detection mode, either when the sulfur is distributed homogeneously in the analyzed volume and exceeds about 20,000 ppm (depending on matrix composition), or when the sample is dilute on the macro scale and the sulfur is heterogeneously distributed in small particles. For example, calculation shows that the fluorescence signal of a sphere of α-S_8 0.2 μm in radius is attenuated by 17% and, therefore, no longer proportional to the absorption coefficient (or absorbance) (Pickering et al., 2001). Also, peak amplitudes are distorted non-linearly. Peaks with normalized X-ray absorbance less than one are enhanced more the smaller their amplitude, whereas those with normalized absorbance greater than one are attenuated to a greater extent the larger their amplitude (Manceau and Gates, 1997; Manceau et al., 2002). Use of total electron-yield (TEY) detection can eliminate the non-linear distortion, with the caveat that chemical sensitivity is not as good as in fluorescence mode. Therefore, homogeneous samples or reference compounds with sulfur concentrations between 20,000 and 100,000 ppm are paradoxically more difficult to measure than diluted materials, and a sample that is dilute at the macro scale may not be appropriate for X-ray absorption measurements if the sulfur distribution is inhomogeneous.

Typically, six types of sulfur functionality in addition to mineral sulfides and elemental sulfur are determinable from S-XANES spectra of natural organic matter: exocyclic sulfur, heterocyclic sulfur, sulfoxide, sulfone, sulfonate, and sulfate (Vairavamurthy, 1998) (Fig. 1a). Quantifying the proportions of each type of functionality relative to the total amount of S (S_i%) is the most difficult part of analyzing the spectra. A single spectrum represents the weighted sum of individual contributions from all sulfur species, which ideally should be quantifiable by performing a linear combination fit (LCF) with spectra of model compounds (Vairavamurthy et al., 1997). However, acquiring reference spectra that represent all of the dominant sulfur species likely present is time-consuming. The acquisition is also technically difficult because these spectra should be recorded with an electron yield detector on pure compounds, or by fluorescence detection on compounds homogeneously diluted in an inert matrix (Jalilehvand, 2006; Prietzel et al., 2011). In most cases the cost of collecting the data is prohibitive and, therefore, the data generally are reconstructed with a limited number of appropriate standard sulfur compounds (e.g., Waldo et al., 1991; Sarret et al., 1999, 2002; Hutchison et al., 2001; Beauchemin et al., 2004; Bostick et al., 2005; Schroth et al., 2007; Cai et al., 2009; Fors et al., 2012; Wijaya and Zhang, 2012).

Alternatively, the main types of sulfur functionalities can be obtained by decomposing a XANES spectrum with a series of Gaussian functions, assuming that each grouping of sulfur oxidation state produces a single peak. In this case, each type of functionality should, in principle, be modeled with one Gaussian centered at the energy of the corresponding s → p photoelectron transition, one or several broad Gaussian peaks representing the resonance scattering that reflects the average binding environment of sulfur in this functionality, and an arctangent step function representing the transition of the photoelectron to the continuum (Fig. EA-1). This analytical scheme has too many unknown parameters, so one Gaussian per functionality and two arctangent functions, one at low energy for all reduced sulfur forms (S_{Red}), and another at higher energy for all oxidized forms (S_{Ox}) are usually applied (Huffman et al., 1991; Vairavamurthy, 1998) (Fig. 1b). The residuals of such Gaussian curve fits (GCF) are in most cases extremely small (Xia et al., 1998; Martinez et al., 2002; Olivella et al., 2002a), suggesting that the fit quality is insensitive to the intrinsic complexity and variability of sulfur species in NOM. A potential issue is that the positions of the arctangents are somewhat arbitrary. The positions either vary inconsistently by several electron-volts among samples in the same study (Martinez et al., 2002), or are placed to the low-energy side of the main Gaussian peak (Solomon et al., 2005; Cai et al., 2009). Both observations suggest that model-fits are not well-constrained (Electronic annex). In fact, the amount of variability in position of the arctangent function in published studies exceeds the difference in excitation energy between any two types of functionality successive in oxidation state (Table EA-1), which increases the uncertainty on the derived S_i fractions.

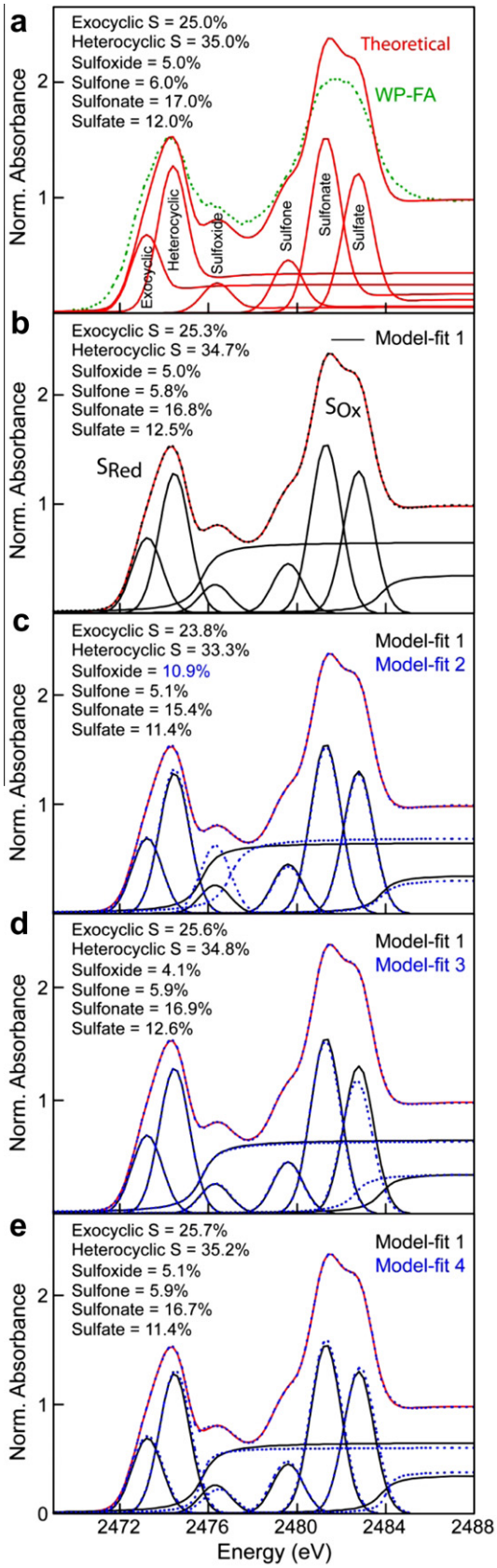


Fig. 1. Validation of the decomposition of a XANES spectrum into several Gaussians and two arctangent curves. (a) A theoretical spectrum made of the six most common sulfur functionalities, each modeled by one Gaussian and one arctangent (6(G + A)), superimposed by the experimental spectrum of the fulvic acid fraction of Waskish peat (WP-FA, green dotted line). (b–e) Model-fits of the theoretical spectrum (red line) to a sum of six Gaussians and two arctangents (6G + 2A). The theoretical spectrum can be fit with a sum of 6G + 2A, but the regression is ill-posed mathematically and the solution is not unique. (b) Model-fit 1 (solid line) = 1st arctangent to the left of sulfoxide (normalized sum-of-squares $NSS = 7.3 \times 10^{-6}$). (c) Model-fit 1 compared to model-fit 2 = 1st arctangent to the right of sulfoxide ($NSS = 7.4 \times 10^{-6}$) (blue dotted line). Misplacing the first arctangent increases the sulfoxide fraction from 5.0% to 10.9%. (d) Model-fit 1 compared to model-fit 3 = 2nd arctangent to the left of sulfate ($NSS = 1.1 \times 10^{-5}$) (blue dotted line). (e) Model-fit 1 compared to model-fit 4 = height of the two arctangents fixed at half (0.62 eV) their optimal width (1.25 eV, model-fit 1) ($NSS = 3.7 \times 10^{-5}$) (blue dotted line). All G positions were optimized. Parameter values are given in Table EA-2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The fractions of different sulfur functionalities are not directly proportional to the relative areas of each Gaussian because the probability of the s → p transition, or absorption cross-section, increases approximately linearly with the density of low-lying unoccupied states, and thus with oxidation state (Waldo et al., 1991). Scaling factors accounting for the change in absorption cross-section are required to convert the areas of the Gaussian curves to fractional amounts of sulfur functionalities. Such scaling factors can be derived from calibration curves measured on model compounds (Huffman et al., 1991; Waldo et al., 1991; Sarret et al., 1999; Orthous-Daunay et al., 2010; Prietzel et al., 2011) (Fig. 2). In these investigations of sulfur-containing natural organic materials, the area of each Gaussian function fit to the spectrum of a natural sample was divided by a derived scaling factor to obtain a quantity proportional to the fraction of S_i.

Using scaling factors determined for individual S reference compounds introduces uncertainty to calculations of S_i functionalities in NOM. The main uncertainty is associated with the measurements of the S compounds, notably if performed in fluorescence-yield detection and the data affected by overabsorption. Slopes of scaling factors vs. absorption edge energy (or equivalently oxidation state of the sulfur functionality) derived from published calibration curves differ by nearly fivefold (Fig. 2). Because the slope is positive, and the fractional amount of S_i ($=S_i/S_T$ where S_T is total S) is proportional to the peak area of S_i divided by the scaling factor, a lower slope results in comparatively overestimated S_{Ox} forms and underestimated S_{Red} forms. The uncertainty caused by an error in slope increases with decreasing oxidation state of the sulfur species. The deviation in S_i caused by a factor of, for example, 2.5 in slope is generally less than 5% of S_T for sulfate (EOS = + 6), may reach 10% for exocyclic sulfur (EOS ~ +0.2 to +0.5), and may be as high as 30–40% for iron monosulfide (EOS = −2) (Electronic annex).

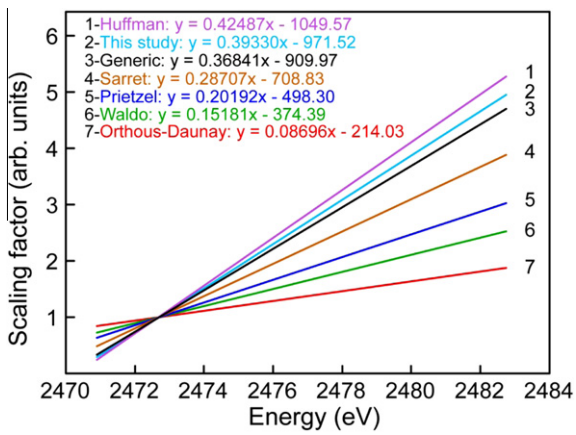


Fig. 2. Scaling factor curves used to calculate the sulfur fractions from the Gaussian areas in the GCF method. The curve from this study was calculated using the approach of Sarret et al. (1999). The Generic curve is the average of the curves of Huffman, Sarret and this study. Scaling factors are normalized to 1.0 at the energy of elemental sulfur (2472.70 eV). The Orthous-Daunay curve has a lower slope because it was measured in fluorescence-yield without correction for overabsorption.

The goal of this article is to improve the reliability of the Gaussian curve fitting (GCF) and linear combination fitting (LCF) methods for S speciation through rigorous analysis of spectra for humic and fulvic acids from the International Humic Substances Society (IHSS). The data for the humic substances and reference compounds are provided in an Electronic annex as a resource and to standardize practices among laboratories.

We test the validity of the Gaussian plus two arctangent model-fit procedure and improve its applicability by showing how to constrain the model to converge towards the correct solution using both theoretical and experimental spectra. The uncertainty associated with use of the scaling factor is estimated by normalizing the Gaussian areas with the two most extreme calibration curves published in the literature (Fig. 2). A generic curve of scaling factor vs. absorption edge energy is generated from the average of the three most realistic curves. Curves with lower slope result from incomplete or lack of correction for overabsorption. This effect is discussed in some detail, and a physically-realistic empirical function is given to correct data from overabsorbing materials.

The new LCF approach, referred to as "Combo method" (Manceau et al., 2012), is applied to the spectra of the IHSS humic substances using an extended database of twenty-nine model compounds. The Combo method consists of performing an unconstrained linear fit to all the references, progressively eliminating references with negative proportions, and then summing the fractions of all remaining references having S in a similar oxidation state and coordination structure to obtain values, S_i, for the types of functionalities. This approach is more robust than regression analysis with a limited set of model compounds, because irrelevant references, which give negative loadings, are removed from the fit. Lastly, the merits and limitations of the parameter-constrained GCF and Combo LCF approaches are discussed.

2. EXPERIMENTAL SECTION

2.1. Materials

Twenty-nine organic and inorganic pure reference compounds were considered (Table EA-1). Eight humic and fulvic acids (Table 1) were obtained from the IHSS and measured by XANES spectroscopy as provided. The S content of the IHSS materials ranges from 2,900 to 8,900 ppm, except for Pony Lake fulvic acid, which has about 30,000 ppm sulfur.

2.2. Data acquisition and treatment

Data were collected on beamline 10.3.2 at the Advanced Light Source (Marcus et al., 2004). Each spectrum is the average of three to five scans for references and at least eight for samples. The references were run in total electron yield mode to reduce overabsorption (Manceau et al., 2002), and the samples in fluorescence yield. Energy calibration was carried out using elemental sulfur, with an absorption maximum taken as 2472.70 eV (inflection point at 2472.0 eV). A monochromator glitch was used as an internal calibration point (Marcus et al., 2008). The uncertainty in energy is 0.10 eV.

Spectra were recorded from well below the edge to at least 2520 eV to provide sufficient pre-edge and post-edge data for precise determination of the background. A smooth pre-edge background was removed from all spectra by fitting a polynomial to the pre-edge region and subtracting this polynomial from the entire spectrum. Normalization of the data was accomplished by fitting a first or second order polynomial to the post-edge region and normalizing the edge jump to 1.0 beyond 2510 eV where the variation of the absorption cross-section is small. Normalized spectra were analyzed over the 2466–2489 eV energy range (GCF and LCF) and the 2466–2520 eV energy range (LCF). Note that the absorption coefficient at $E_{max} = 2489$ eV in the reduced energy interval is higher than one, because of the contribution from post-edge resonance scattering. In the GCF method, a scaling different from unity at E_{max} does not change the fractions of S species because the peak areas for the different oxidation states are divided by the total area for all species. In the LCF method, the sum of components was not constrained to one to evaluate the reliability of the data analysis (Electronic annex). The sum of component fractions for all samples over the two fit intervals had a mean of 1.03 and a standard deviation of 0.03. The largest discrepancy from unity was 1.09 (for Elliott Soil humic acid (ES-HA) analyzed in the reduced energy interval). The GCF fits were carried out using an in-house program, and the LCF fits using the Combo method implemented in a LabView-based program (Manceau et al., 2012). The fit quality of the GCF and LCF reconstructions were evaluated with the normalized sum-of-squares ($NSS = \sum(y_{exp} - y_{fit})^2/\sum y_{exp}^2$) parameter.

The steps of the LCF method have been described previously (Manceau et al., 2012). Briefly, the data were fit to a sum of all references. References with negative loadings

Table 1
Samples names, sulfur content by weight, and atomic fractions (%) of sulfur species as determined by sulfur K-edge XANES spectroscopy.

Code		Total S	Exocyclic S		Heterocyclic S		Sulfoxide		Sulfone		Sulfonate		Sulfate		Sum
Name	IHSS #	% g/g	%	Δ/2	%	Δ/2	%	Δ/2	%	Δ/2	%	Δ/2	%	Δ/2	
Suwannee River HA II	SR-HA	2S101H	0.54	23.6	0.3	40.4	0.7	3.5	0.1	6.4	1.0	19.1	0.1	7.0	0.3	100.0
Suwannee River FA II	SR-FA	2S101F	0.46	25.8	2.7	27.6	3.6	2.8	0.0	9.1	1.8	25.0	0.4	9.7	0.4	100.0
Waskish Peat HA	WP-HA	1R107H	0.36	23.5	2.7	34.0	1.5	4.1	0.0	3.6	0.2	14.6	1.0	20.1	2.0	100.0
Waskish Peat FA	WP-FA	1R107F	0.29	30.5	3.4	39.7	2.4	4.1	0.5	5.0	0.2	13.4	0.3	7.3	0.4	100.0
Nordic Aquatic HA	NA-HA	1R105H	0.58	32.2	4.2	39.6	3.2	3.8	0.2	5.4	0.6	13.6	0.9	5.4	0.4	100.0
Pony Lake FA	PL-FA	1R109F	3.00	46.9	0.0	21.3	2.7	3.9	1.4	5.5	1.4	17.8	0.5	4.5	0.7	100.0
Elliott Soil HA I	ES-HA	1S102H	0.44	18.4	3.1	28.2	1.2	5.8	0.6	4.4	0.2	25.1	0.8	18.1	0.3	100.0
Elliott Soil FA II	ES-FA	2S102F	0.89	15.3	1.5	19.5	1.2	3.2	0.4	6.3	0.4	29.6	1.3	26.1	1.8	100.0

The atomic fraction was taken as the arithmetic mean of the GCF and LCF average values, and the error is the half difference (Table EA-5).

were successively eliminated in ascending order of weight (i.e., starting from the most negative loading). After each negative component was eliminated, the loadings of all others were automatically re-optimized, and the elimination iterated until all references with negative loadings were removed. At this point, each reference that had been deleted was added randomly one-by-one, while maintaining the constraint of non-negativity, to ensure that the fit had reached a global rather than local minimum. Uncertainties in the fractions of reference species were discussed previously (Manceau et al., 2012). Unlike in the usual type of linear combination fitting, we do not assert that the unknowns are mixtures of the actual reference compounds. Therefore, there is no meaning to assigning error bars to the loadings of individual components. Rather, what we have verified with the GCF method is that the empirical Combo method yields the correct partitioning of sulfur functionalities after grouping individual fitted components with similar electronic oxidation state.

3. RESULTS AND DISCUSSION

We first evaluate application of the GCF method, which is the most commonly used in the literature, to determine amounts of different types of sulfur functionalities characterized by sulfur oxidation state. The critical issue in previous usage is the placement of the arctangent functions, which can lead to erroneous results. The method is tested against a theoretical spectrum and then applied to IHSS humic and fulvic acids. We next report results on the IHSS humic and fulvic acids obtained using the LCF method. Similarities and differences between the two approaches lead to a recommendation for the preferred use of the LCF method. We show in the last section with examples and a calculation that overabsorption can have a significant effect on quantifying sulfur functionalities, and explain how data corrupted by overabsorption may, to some extent, be empirically corrected.

3.1. The Gaussian curve fitting (GCF) method

3.1.1. Validation and limitations

A theoretical spectrum made of six Gaussian and six arctangent functions ($\sum_{i=1}^{6}(\mathrm{G}+\mathrm{A})_i$) with parameters chosen to reproduce the main features of a typical XANES spectrum for a humic substance was calculated, then fit using least-squares to six Gaussians and two arctangents ($2A + \sum_{i=1}^{6}\mathrm{G}_i$) (Fig. 1a, Table EA-2). The theoretical mixture is composed of 25% exocyclic reduced sulfur (S_{Exo}), 35% heterocyclic reduced sulfur (S_{Hetero}), 5% sulfoxide (S_{Sulfx}), 6% sulfone (S_{SO_2}), 17% sulfonate (S_{SO_3}), and 12% sulfate (S_{SO_4}). These specific percentages were chosen to be similar to the average sulfur functionalities of the eight IHSS humic and fulvic acids (Table 1). The positions in units of energy (eV) and heights in units of normalized absorbance of the Gaussian and arctangent functions used to fit the theoretical spectrum were optimized independently. The widths of all the G functions (G_w) were covaried, and the widths of the A functions (A_w) covaried and also fixed. The influence of the initial positions and width

of the two arctangents on the fit values were tested with parametric settings chosen to reproduce the main diversity of the fit strategies reported in the literature:

Model-fit 1. The initial energy of the first arctangent function (A_E^1) was set below the energy of the Gaussian of the sulfoxide functionality (G_{Sulfx}), and the initial energy of the second arctangent function (A_E^2) was set above the energy of the sulfate Gaussian (G_{SO_4}), thus $A_E^1 < G_E^{Sulfx}$ and $A_E^2 > G_E^{SO4}$;

Model-fit 2. $A_E^1 > G_E^{Sulfx}$ and $A_E^2 > G_E^{SO4}$;

Model-fit 3. $A_E^1 < G_E^{Sulfx}$ and $A_E^2 < G_E^{SO4}$;

Model-fit 4. $A_E^1 < G_E^{Sulfx}$ and $A_E^2 > G_E^{SO4}$, and the width of the first A function was held constant at half the optimal value obtained in the previous trials: $A_{w=0.62}^1$ instead of $A_{w=1.25}^1$.

The theoretical spectrum was reproduced almost equally well by all four model-fits. In the first three model-fits $NSS \leqslant 10^{-5}$, and in the fourth $NSS = 3.7 \times 10^{-5}$ (Fig. 1b–e). The small difference observed in the last fit would be undetectable on a real spectrum, because experimental data always contain noise. The best-fit fractions of S_i deviate at most by 0.5% from their nominal values in model-fit 1, and in the 1% range in the other models, except for S_{Sulfx}, which deviates by as much as 5% (100% error) when the A_1 function is to the right of the sulfoxide component (Model-fit 2) (Table EA-2). In general, S_{Red} species are more sensitive to the position of A_1 than S_{Ox} species are to the position of A_2 because the reduced species have smaller absorption cross-sections, and therefore, relatively higher A_1/G amplitude ratios. Overall, these simulations show that the $2A + \sum_{i=1}^{6} G_i$ model-fit recovers the correct sulfur fractions from the theoretical XANES spectrum as long as the position of the first arctangent is constrained to be close to the Gaussians of the reduced sulfur functionalities (i.e., $A_E^1 < G_E^{Sulfx}$).

In the above four model-fits we assigned the same widths to all Gaussian functions for the sake of simplicity. In reality, the widths vary because the number, type and electronegativity of the atoms coordinated to sulfur vary among and within each type of functionality. For example, sulfur is bonded to four oxygens in sulfate and to three oxygens and one carbon in sulfonate (RSO_2O^-). A caveat in assigning the same Gaussian width for SO_4^{2-} and SO_3^- species is the fact that the electronegativities for oxygen and carbon are extremely different (3.44 vs. 2.55), which is reflected as an asymmetrical broadening of the s → p peak to higher energy for sulfonate (Fig. EA-1). Two Gaussians are needed to reproduce the shape of the main peak of sulfonate, with the second located close to the energy of sulfate (2482.95 eV vs. 2482.75 eV). In sulfone, the S atom is attached to two oxygens and two carbons, usually in two separate hydrocarbon substituents (RSO_2R', Table EA-1). Thus, the main peak of sulfone can be decomposed into either two to three Gaussians or one single Gaussian with a larger width of 2.9 eV compared to 1.9 eV in sulfate (Fig. EA-1). Finally, each type of functionality is in reality a group of sulfur compounds with a range of 'oxidation index' (Vairavamurthy, 1998). The distribution of compounds within the functionality determines the energy of the observed peak maximum. Therefore, a spectrum from NOM is more complex than our theoretical spectrum. We show below, with the example of the humic acid (HA) fraction for Elliott Soil, that the natural heterogeneity of sulfur speciation introduces more correlations in the fit parameters.

Six statistically equivalent fits to the S-XANES spectrum of Elliott Soil HA (ES-HA) were obtained by varying the initial parametric settings (Fig. 3; Table EA-3). These fits represent the diversity of spectral decomposition reported in the literature for natural organic materials. Two different widths for the Gaussian functions were necessary, one for the low-oxidation-state species (exocyclic and heterocyclic sulfur, sulfoxide) and one for the high-oxidation-state species (sulfone, sulfonate, sulfate, and sulfite (the latter was included only in Model-fit 6)).

Model-fit 1. 6G, $A_E^1 < G_E^{Sulfx}$ and $A_E^2 < G_E^{SO4}$, five G positions fixed to nominal values (±0.1 eV) for specific types of functionality (Table EA-2). This scheme is the most reliable; the best-fit values are close to the values derived from the Combo LCF method, discussed later.

Model-fit 2. 6G, $A_E^1 > G_E^{Sulfx}$ and $A_E^2 < G_E^{SO4}$, all G positions unconstrained. Shifting A_E^1 to the right of G_E^{Sulfx} increases the fraction of sulfoxide from 5.1% (Model-fit 1) to 13.0%.

Model-fit 3. 6G, $A_E^1 < G_E^{Sulfx}$ and $A_E^2 > G_E^{SO4}$, four G positions fixed to nominal values (±0.1 eV). Shifting A_E^2 to the right of G_E^{SO4} increases the fraction of sulfate from 17.8% (Model-fit 1) to 22.1%.

Model-fit 4. 6G, $A_{w=1.0}$, $A_E^1 < G_E^{Sulfx}$ and $A_E^2 < G_E^{SO4}$, all G positions unconstrained. When the widths of the two arctangent functions are fixed to 1.0, instead of 0.42 (Model-fit 1 value), a good fit is still obtained by adjusting the positions of the Gaussians. However, the sulfone peak is shifted to the left from 2479.60 to 2479.10 eV, a value which is unrealistic for this type of functionality (Vairavamurthy, 1998) because it approaches the nominal position of the sulfite peak at 2478.55 eV.

Model-fit 5. 6G, $A_{h=0.5}^1$, $A_E^1 < G_E^{Sulfx}$ and $A_E^2 < G_E^{SO4}$, all G positions unconstrained. Lowering arbitrarily the height of A_1 from 0.69 (Model-fit 1 value) to 0.5 is compensated mathematically by a shift of the sulfone peak even further to the left to 2478.90 eV.

Model-fit 6. 7G, $A_E^1 < G_E^{Sulfx}$ and $A_E^2 < G_E^{SO4}$, all G positions fixed to nominal values (±0.1 eV). The number of degrees of freedom exceeds the number of independent data when a seventh Gaussian is added in the refinement; the data are overfit. The fit presented here was obtained by adding sulfite and fixing the seven Gaussians to their nominal values. This solution is chemically plausible, but the presence of sulfite is unsupported by the Combo fit as shown below (Section 3.2).

In summary, although each S_i functionality should, in principle, have its own Gaussian and arctangent step function, meaningful fits can be obtained by using only two arctangents, one representing the transitions to the continuum for the S_{Red} species and the other this transition for S_{Ox} species. However, even the $2A + \sum_{i=1}^{6} G_i$ model is overparameterized and several initial conditions and constraints on parameter values must be introduced in order to obtain meaningful fits. The main bias is in the position of the first

arctangent which should be between the energies of the G_{Hetero} and G_{Sulfx} peaks to obtain realistic reduced sulfur and sulfoxide values. In certain previous studies the first arctangent has been placed either below or above this range of energy (Huffman et al., 1991, 1995; Taghiei et al., 1992; Skyllberg et al., 2000; Szulczewski et al., 2001; Martinez et al., 2002; Solomon et al., 2003; Wiltfong et al., 2005; Pattanaik et al., 2007; Huggins et al., 2009) (Electronic annex).

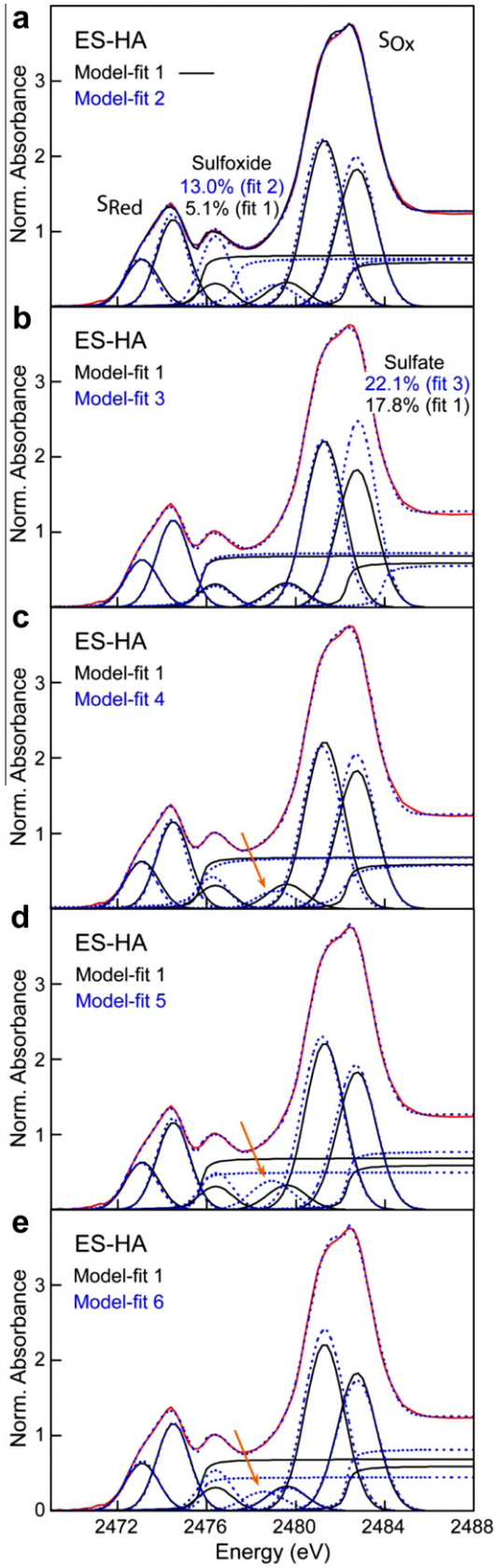


Fig. 3. Influence of initial parameter settings and constraints on best-fit values in the Gaussian decomposition of the XANES spectrum for Elliott Soil humic acid (ES-HA, red line). (a) Model-fit 1 (black solid line) = convergence-optimized method ($NSS = 2.8 \times 10^{-4}$) (black line); Model-fit 2 = 1st arctangent to the right of sulfoxide and all G positions optimized ($NSS = 1.9 \times 10^{-4}$) (blue dotted line). (b) Model-fit 1 compared to model-fit 3 = 2nd arctangent to the right of sulfate, several G positions fixed ($NSS = 2.4 \times 10^{-4}$) (blue dotted line). (c) Model-fit 1 compared to model-fit 4 = FWHM of the two arctangents fixed at twice (1.0 eV) their optimal width (0.42 eV, model-fit 1), and heights fixed to their optimal values (0.69 and 0.60), all G positions optimized ($NSS = 2.4 \times 10^{-4}$) (blue dotted line). (d) Model-fit 1 compared to model-fit 5 = height of the 1st arctangent fixed to 0.50 (instead of 0.69) and width fixed to 0.42, all G positions optimized ($NSS = 1.6 \times 10^{-4}$) (blue dotted line). (e) Model-fit 1 compared to model-fit 6 = addition of a seventh component (sulfite), all G positions fixed ($NSS = 2.1 \times 10^{-4}$) (blue dotted line). Arrows in (c–e) point to regions of significant fit error in the mid-energy range. Parameter values are given in Table EA-3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.1.2. Description of optimized GCF procedure

Numerous fitting schemes, in addition to those discussed above, were tested until a general minimization procedure that converged to a unique numerical solution with results in close correspondence to those from the LCF method (*v.i.*) was devised. Consistent with the simulations of the theoretical and ES-HA spectra, fitting one or two widths to all the Gaussian functions, and a separate width to both arctangent functions, decreased most, but not all, parameter correlations. In addition to these constraints on widths of the functions, we determined that the positions of G_{Sulfx}, G_{SO_2}, G_{SO_3}, and G_{SO_4} could be fixed safely at their nominal values without unduly affecting the goodness-of-fit parameter (*NSS*). When varied, G_{Exo} shifted to the left by at most −0.2 eV, and G_{Hetero} shifted to the right by at most +0.3 eV relative to their nominal positions. Lastly, by imposing $A_E^1 < G_E^{Sulfx}$ and $A_E^2 > G_E^{SO4}$ when [sulfate] > [sulfonate] or $G_E^{SO3} < A_E^2 < G_E^{SO4}$ when [sulfate] < [sulfonate] as initial conditions for arctangent positions, the fit converged to the same mathematical solution regardless of the starting parametric values, and the regressed S_i values were all close to those derived from the LCF method.

3.1.3. Application to IHSS humic substances and determination of generic scaling factors

Exocyclic and heterocyclic reduced sulfur, sulfoxide, sulfone, sulfonate, and sulfate, but not sulfite ($E = 2478.55$ eV) or sulfonium ($E = 2475.60$ eV), were identified in all materials, in agreement with previous studies (Xia et al., 1998; Yoon et al., 2005) (Tables 1, EA-4 to

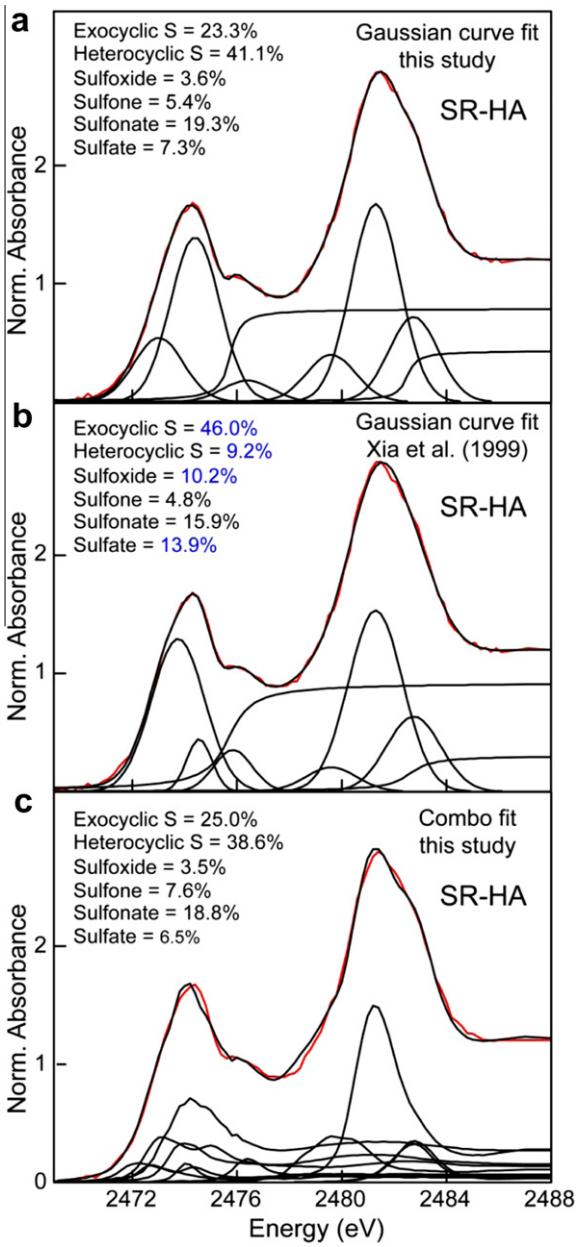


Fig. 4. Three simulations of the XANES spectrum of Suwannee River humic acid (SR-HA). (a) Gaussian decomposition using the convergence-optimized method. (b) Gaussian decomposition of Xia et al. (1998). (c) Combo fit. Significant deviations in the fractional amounts of sulfur functionalities in blue arise from use of unrealistic Gaussian fit and the Waldo scaling factor curve.

EA-6, and Figs. 4 and EA-2). The S_i values were calculated from the Gaussian areas using five peak area vs. peak energy calibration curves (Fig. 2; Table EA-4). In order of increasing slope the curves are denoted as the Waldo curve (#6 in Figs. 2, Waldo et al., 1991), which was reproduced by Xia et al. (1998); the Prietzel curve (#5, Prietzel et al., 2011); the Sarret curve (#4, Sarret et al., 1999); the curve from this study (#2); and, the Huffman curve (#1, Huffman et al., 1991) (Figs. 2 and EA-3). The X-ray energies in the previous studies were shifted to match our calibration. The Orthous-Daunay curve (#7 in Fig. 2, Orthous-Daunay et al., 2010) has the lowest slope and were omitted because the data were recorded in fluorescence-yield detection mode without correcting for overabsorption.

For calibration curves with lower slopes, the scaling factor for the S_{Red} species (i.e., $S_{Exo} + S_{Hetero}$) is relatively larger compared to those from calibration curves with higher slopes. Thus, lower slopes yield smaller proportions of S_{Red} species, and correlatively, greater proportions of S_{Ox} species. This effect is seen by comparing all the S_i values for the eight IHSS samples (Table EA-4). The S_{Exo} fraction is on average $-7.3 \pm 1.3\%$ lower when calculated with the Waldo curve than with the Huffman curve, and the S_{SO_4} fraction is on average $3.8 \pm 1.6\%$ higher. One might consider that these differences, calculated with the two most extreme curves, are within the precision of 10% of total S generally stated for the method, and thus not significant (Huffman et al., 1991; Prietzel et al., 2011). However, other sources of errors (e.g., data measurement and normalization, parameterization of the model-fit), when added to the imprecision arising from this correction factor, can stretch the error bar well beyond 10% and in certain cases to the point of rendering the results meaningless. Functionalities with intermediate oxidation states are less sensitive to variability in the slope. For example, the value of S_{Sulfx} is essentially independent of the calibration curve used to normalize the areas under the Gaussian peaks (Table EA-4).

Definition of a generic calibration curve would decrease uncertainty in determining the fractional amounts of sulfur functionalities, for data that are not compromised by overabsorption. The Waldo and Prietzel curves both have lower slopes than the Sarret curve and the curve from this study, likely in part because they were measured in fluorescence-yield detection mode, instead of TEY as for the two others. Therefore, the uncertainty of calculated values of S_i for any natural organic material should be minimized by calibrating the data with the mean of the three steepest curves. Its equation is $y = 0.36841x - 909.97$, where x is the energy of the absorption maximum (i.e., corresponding to the oxidation state of the sulfur functionality) and y the scaling factor normalized to $y = 1$ at the energy of elemental sulfur ($E = 2472.70$ eV). The confidence interval of this calibration procedure is estimated to be $\Delta S_i = |S_i(\text{Huffman}) - S_i(\text{Sarret})|/2)$. Results (Table EA-4) show that S_i calculated with this generic curve has a mean precision of $\Delta S_i = 1.6 \pm 0.2$ atom% for the eight S_{Exo} values and 0.7 ± 0.3 atom% for S_{SO_4}. The accuracy of S_i is estimated below from its independent quantification by the LCF method.

In the previous discussion we showed that model bias is the main source of error in the GCF method: if the selected model is incorrect or the number of degrees of freedom in the regression are too high, then the error on S_i can exceed 20% of S_T, and some species can be overlooked or misidentified. An example from the literature for an IHSS humic substance, Suwannee River humic acid (SR-HA) (Xia et al., 1998), is shown in Figs. 4b and EA-5. The amount of S_{Exo} was overestimated by 23% and that of S_{Hetero} underestimated by 32% (Tables EA-4 and EA-6), at least partly

because the widths of the G_{Exo} and G_{Hetero} components were varied independently, and possibly because G_{Hetero} was nested within G_{Exo}. In addition, the sum of the two reduced sulfur functionalities ($S_{Exo} + S_{Hetero}$) was underestimated by 9%, and S_{SO_4} correlatively overestimated by 7%, because the Gaussian areas were normalized with the Waldo curve. Similar, although slightly smaller, deviations were obtained for Suwannee River fulvic acid (SR-FA) (Xia et al., 1998) (Fig. EA-6). In contrast, differences in the fractions of sulfur functionalities for Nordic Aquatic humic acid (NA-HA) between published GCF results (Yoon et al., 2005) and those obtained by the GCF method in this study are at most 3.9%, and the main cause is use of the Waldo curve (Fig. EA-7).

3.2. The linear combination fitting (LCF) method

In the GCF method, sulfur functionalities in humic substances are determined implicitly with a curve-fitting model that is physically realistic. A more direct approach is to model the data explicitly with a database of model compounds. Results show that the LCF method provides a similarly good description of the data with *NSS* values from 0.47×10^{-3} to 1.25×10^{-3} (Table EA-7, Figs. 4c and EA-4). Decreasing the fit interval from 2466–2520 eV to 2466–2489 eV had little effect on the results. The precision, expressed as $\Delta S_i/2$ between the two fits, is always better than 0.5% of S_T for oxidized sulfur species and 2% for reduced sulfur species (Tables EA-4 and EA-8). Expressed as the variation of S_i ($(\Delta S_i/2)/S_i$), the relative precision is always better than 7%, and in most cases as good as 1% for minor species, such as S_{Sulfx}. The good precision on minor species indicates that the method is robust because calculated errors on small numbers are often much greater than the numbers themselves. The accuracy of the LCF method is, like for the GCF method, difficult to evaluate because it also depends on several factors, such as systematic errors, data quality (signal-to-noise ratio, linearity of the detector, overabsorption, the range of fitting), and how well the standards represent the unknown sample (Jalilehvand, 2006; Prietzel et al., 2011).

3.3. Comparison of GCF and LCF results for IHSS humic substances

The convergence-optimized GCF and the Combo LCF methods were used to make duplicate estimates of S_i fractions and their uncertainties (omitting experimental errors) of the eight IHSS standard and reference materials. The six sulfur functionalities quantified for each of the eight samples yield a set of 48 values. The S_i values recovered by the two independent methods differ at most by 4% of S_T in 38 cases, and by 4–8% in the 10 other cases, all of which are for reduced sulfur functionalities (Table EA-5). Therefore, each oxidized sulfur functionality is quantified in the two analytical methods with an accuracy of about 4%. Although each S_{Red} species has an accuracy of about 8%, the sum of the two reduced sulfur functionalities ($S_{Hetero} + S_{Exo}$) is accurate to within <4%, except for Pony Lake fulvic acid (PL-FA), which has a slightly higher error of 5.4%. The two reduced sulfur functionalities have a lower accuracy in part because major thiol/sulfide species with electronic oxidation state of 0 to 0.5, such as cystine, often have a split white line (Chauvistré et al., 1997; Prange et al., 2002), which is assigned to a mixture of S_{Exo} and S_{Hetero} in the GCF method and to S_{Exo} only in the LCF method (Szulczewski et al., 2001; Prietzel et al., 2011). Accordingly, $S_{Exo}(GCF) < S_{Exo}(LCF)$ and $S_{Hetero}(GCF) > S_{Hetero}(LCF)$ in all samples rich in reduced sulfur (Table EA-4). This deficiency of the GCF method is the reason for the lower calculated accuracy of the two reduced species.

3.4. Sulfur fractions in IHSS humic substances

The final S_i values (Tables 1 and EA-5) for the eight IHSS samples are the average of (1) the GCF values corrected with the generic calibration curve and (2) the mean LCF values of the fits over the extended (2466–2520 eV) and reduced (2466–2489 eV) energy intervals. The precision, estimated as the magnitude of one-half of the difference between the GCF and LCF values ($\Delta S_i = |[S_i(GCF) - S_i(LCF)]/2|$) ranges from a mean of 2.2 ± 1.5 atom% for the eight S_{Exo} values to 0.8 ± 0.7 atom% for S_{SO_4}. The final S_i values with their uncertainties represent the best description of the sulfur XANES data for the IHSS samples with the sulfur compound database recorded in this study. Results show that dissolved organic matter from aquatic or organic soil environments, such as peats, contain more reduced sulfur, whereas dissolved organic matter from the Elliott Soil, a mineral-dominated environment, contains more oxidized sulfur, in agreement with previous studies (Xia et al., 1998; Qian et al., 2002; Solomon et al., 2003, 2005; Zhao et al., 2006; Prietzel et al., 2007, 2009).

Pony Lake fulvic acid (PL-FA), from a coastal pond near McMurdo, Antarctica (Brown et al., 2004), has the highest amount of strongly reduced sulfur ($S_{Exo} = 46.9\%$), which was fit in the LCF simulation with a combination of dibenzyl sulfide, cystine, pyrite, and pyrrhotite. The fraction of cystine, equal to 12%, is compatible with the dominantly microbial origin of the dissolved organic matter in Pony Lake (Brown et al., 2004). The fractions of Fe-sulfide minerals, also equal to approximately 11–12%, is supported by the anaerobic microbial assemblage and odor of hydrogen sulfide gas observed in bottom ice extracted from Pony Lake in November 2004 when the lake was frozen solid (Foreman et al., 2011). We note that the main spectral peaks fit with pyrite and pyrrhotite occur at energies similar to those reported for ferredoxins (Anxolabéhère-Mallart et al., 2001). Given the microbial and algal origin of PL-FA (Brown et al., 2004; Mao et al., 2007), the components fit as the Fe-sulfide minerals may also be attributable to ferredoxins, proteins with Fe_xS_y clusters bonded to cysteinyl sulfur. Because cysteine (and its oxidized form cystine) is more stable than Fe sulfides in oxidizing environment, the ferredoxin hypothesis was tested by measuring the same sample stored at ambient conditions twice over a three-year interval, in 2009 and 2012 (Fig. 5a), The broad left tail of the S_{Exo} component clearly apparent on the 2009 spectrum, and fit as pyrite/pyrrhotite in the LCF simulation, is no longer observed on the 2012 spectrum, and the S_{Ox} peak

is more intense. The LCF amount of S_{Exo} decreased from 47.0 to 39.6%, and the amount of S_{Ox} increased correspondingly from 22.2% to 28.9%. The decrease of S_{Exo} is reflected in the LCF simulation by a diminution from 11–12% to 3–4% of the pyrrhotite + pyrite component, thereby supporting the inorganic nature of the strongly reduced sulfur form in PL-FA.

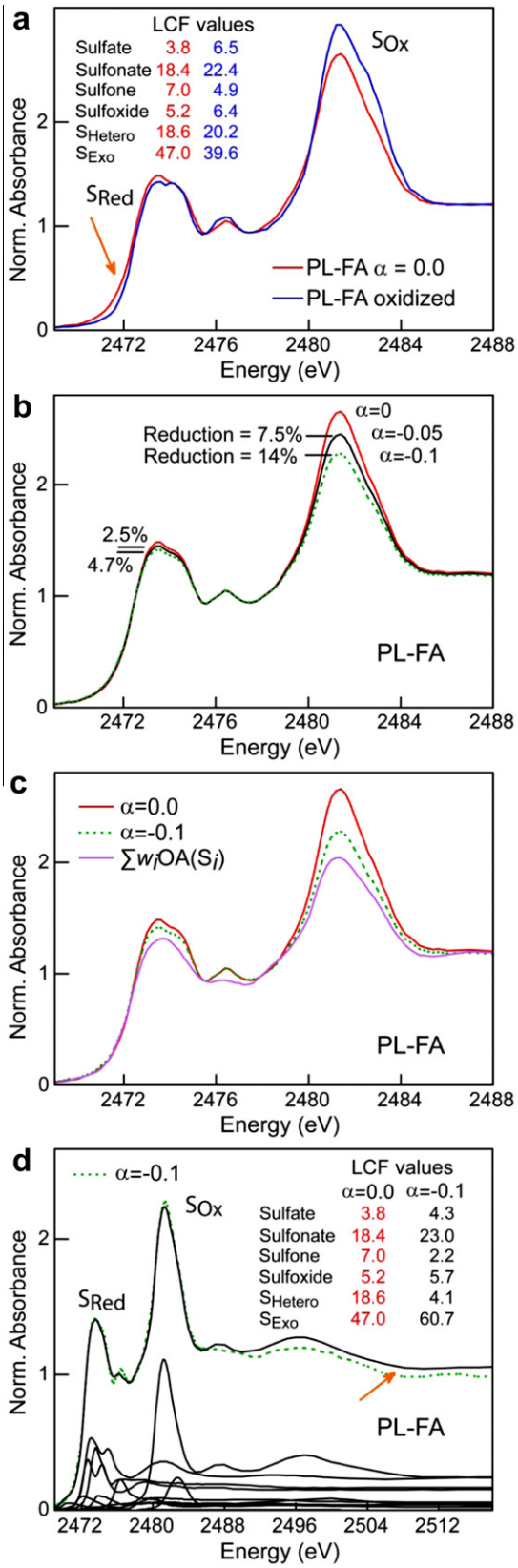


Fig. 5. (a) Sulfur K-edge XANES spectrum for Pony Lake fulvic acid (PL-FA) recorded in 2009 and again in 2012 after storage at ambient conditions in a capped tube. A large fraction of the strongly reduced pool (arrow) has been oxidized with time. The S_i values were calculated by LCF simulation. (b) Simulation of the overabsorption effect on PL-FA. The absorbance for the S_{Ox} and S_{Red} fractions are reduced non-linearly, which modifies the S_{Ox}/S_{Red} ratio in the data analysis if this effect is overlooked. (c) Reconstruction of the $\alpha = 0.0$ spectrum obtained in 2009 with the best-fit S_i components from Table EA-7 (PL-FA long), each of which was modified intentionally by the overabsorption factor $\alpha = -0.1$, then weighted by w_i. This reconstruction reproduces what the simulation would look like if the reference spectra were affected by overabsorption (i.e., were measured in fluorescence-yield mode). (d) Combo fit of the $\alpha = -0.1$ PL-FA spectrum to the reference spectra from the database, each of which was modified intentionally before the fit by the overabsorption factor $\alpha = -0.1$. The fit has a higher absorbance at high energy (arrow), because the amplitude of the reference spectra is more attenuated by overabsorption than the amplitude of the measured spectrum. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Seasonal melting and formation of the ice cover on Pony Lake and wind patterns during the summer induce changes in ecological dynamics, and also oxygenation state of the pond water (McKnight et al., 1994; Foreman et al., 2011). The overall S speciation of PL-FA samples collected at any one time, therefore, is likely a function of the fluctuating biogeochemical system in this relatively small, shallow body of water. The initial total amount of S_{Red} (65.6% from LCF, Fig. 5a), when renormalized to exclude sulfate, is 68.2%, compared to the value of 58.9% (uncertainty not reported) determined by XPS (Fimmen et al., 2007). The difference may be true variation, again related to different sampling times (January 2005, Fimmen et al., 2007 vs. December 2005–February 2006, IHSS PL-FA reference), and may also come partly from the analysis by XPS of a more oxidized PL-FA powder.

3.5. The special case of a dominant reduced sulfur species

Natural organic matter from aquatic environments and soils typically contains a large number of individual, but no single dominant, sulfur species. In this case, the weighted summation of all post-edge resonances from constitutive S_{Red} species is smeared out and does not overlap to a great extent with the white-line of more oxidized species, a condition necessary for successful application of the GCF method. However, in cases where one or two particular reduced sulfur species dominate, only the LCF method will produce the correct proportions of sulfur functionalities.

Two examples are provided: soil organic matter enriched in cysteinyl sulfur species and insoluble organic matter from a carbonaceous chondrite (Fig. 6, Electronic annex), both of which have a broad upward slope in their spectra between approximately 2476–2480 eV. The LCF fits identify

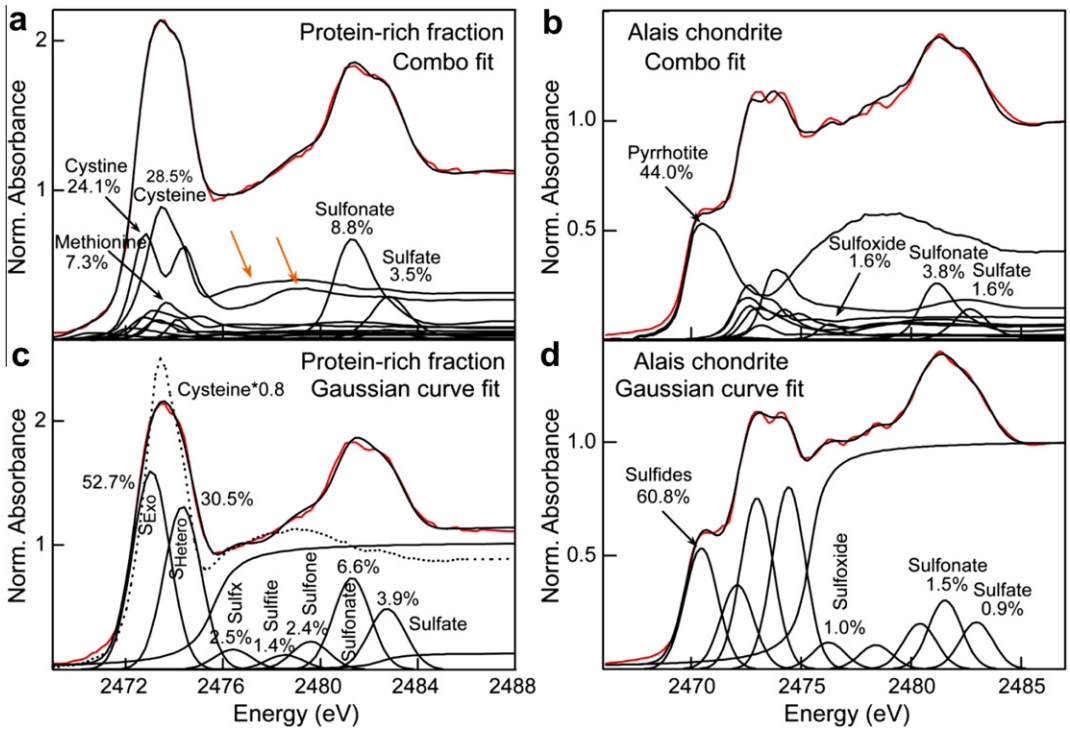


Fig. 6. (a) Combo fit of the sulfur K-edge XANES spectrum for the protein-rich fraction of a humic substance. (b) Combo fit of the sulfur K-edge XANES spectrum of the Alais meteorite (spectrum digitized from Orthous-Daunay et al., 2010). The high amounts of cysteine and cystine (colored arrows in a) and pyrrhotite (b) are reflected phenomenologically in the positive slope of the XANES spectra at 2476–2479 eV. (c) Fitting the humic material with the GCF method leads one to conclude incorrectly that the sample contains sulfoxide, sulfite and sulfone. Eighty percent of the cysteine reference spectrum is superimposed for comparison. (d) Fitting the meteorite spectrum with the GCF method results in a high percentage of multiple reduced sulfur species when in reality only one is dominant. Only one arctangent was fit as in the model of Orthous-Daunay et al. (2010). Details are given in the Electronic annex. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

predominantly cysteine and cystine in the soil organic matter, and pyrrhotite in the meteorite. All three reference compounds have a broad absorption hump at ~2475–2482 eV, in a spectral region intermediate between the reduced and oxidized sulfur pools where the absorption usually goes through a minimum in a more heterogeneous material. These broad and intense spectral features overlap with the first arctangent function in the GCF method and suppress the minimum in absorbance between the Gaussians of sulfoxide and sulfone components. In this region, the LCF model for the protein-rich organic matter, which incorporates the detailed post-edge features of the S_{Red} species, shows that the three species, sulfoxide, sulfite, and sulfone, fit by the GCF model are spurious. The same region in the meteorite spectrum poses even greater challenges when applying a GCF model. This region is described with one arctangent, the tail of one Gaussian function with relatively large amplitude, and three small-amplitude Gaussians using our optimized GCF application of the model of Orthous-Daunay et al. (2010) (Fig. 6d, Electronic annex). In addition, the asymmetrical shape of the main pyrrhotite component is described with two Gaussian peaks, clearly showing the limit of the GCF approach. Other examples of upward slopes between 2476 and 2480 eV caused by the prevalence of one reduced sulfur species include thiols and organic monosulfides in chick corneas (Koudouna et al., 2012), and elemental sulfur α-S_8 in marine archeological wood from shipwrecks and pyrite in reduced sediments (Fors et al., 2012).

For these types of organic materials, only the LCF method will give the correct S_i partitioning, assuming that the predominant S_{Red} species is not fine-grained which would corrupt the amplitude through overabsorption (*v.i.*). For example, the GCF method indicates that the protein-rich humic sample contains approximately 52.7% S_{Exo} and 30.5% S_{Hetero}, and the LCF method 83.0% S_{Exo} and 4.2 S_{Hetero}. However, despite the large uncertainty on S_i, in this case, the sum of the reduced sulfur obtained with the GCF method is only 5% less than that obtained using the LCF method because (1) the total area of the reduced pool is measured correctly with GCF, and (2) the generic scaling factors used to correct for absorption cross-sections are consistent with the spectral database used to derive the GCF values (Fig. 2). For the same reasons, both methods also yield near-identical total reduced sulfur in the meteorite sample, being equal to 93.0% with LCF and 94.8% with GCF.

3.6. Overabsorption

The somewhat lower slope of the calibration curve of Prietzel et al. (2011) shown in Fig. 2 suggests that it may be affected by overabsorption. This calibration curve was

obtained from standards prepared as mill-ground mixtures of pure sulfur compounds (FeS, FeS_2, S°, cysteine, sulfoxide, Na_2SO_4,...) and finely-ground quartz in proportions yielding total S concentrations of 200, 2,000 and 20,000 ppm. The authors reported that the edge-normalized spectra of the 200 and 2,000 ppm series had white-lines with similar magnitudes and shapes, as did most of the 20,000 ppm spectra. They concluded from this observation that overabsorption was negligible in their mixtures, in support of the view (Waldo et al., 1991) that overabsorption is significant only above 50,000 ppm (5 wt.% S). We show below, with a calculation performed using Hephaestus (Ravel and Newville, 2005), that the 50,000 ppm limit in concentration is misleading for two reasons. First, the actual upper limit in S concentration depends on the matrix composition, and is much lower for an organic than a mineral matrix. Second, this limit is for samples which are homogeneous, not for samples in which sulfur is nanoparticulate, which occurs in many, if not most, environmental materials.

3.6.1. Overabsorption in matrices of quartz and humic substances with homogeneously distributed sulfur

The reduction of the X-ray absorbance measured in fluorescence-yield resulting from overabsorption can be calculated from the edge jump vs. μ(res)/μ(tot) contour map of Manceau et al. (2002), where μ(res) is the resonant absorption at the energy of the absorption edge and μ(tot) is the total absorption. This map was determined for a planar film of variable thickness (i.e., edge jump) that approximates a thin layer coating or a slice of a larger particle. It shows that the fluorescence intensity measured on a thick sample (relative to the penetration depth of the incident beam) is attenuated by less than 10% when $\mu(res)/\mu(tot) < \sim 0.09$ and by 20% when $\mu(res)/\mu(tot) = 0.19$. Mass absorption coefficients (μ) reported in Table 2 show that 20,000 ppm S in SiO_2 has a resonant absorption equal to 0.017 of the total absorption, most of which comes from Si. Absorption by S is negligible, therefore the fluorescence intensity is always proportional to absorbance. In a humic sample devoid of Si, we find that the contribution of sulfur to the total absorption is 3–4 times higher (0.058) than in a quartz matrix for the same sulfur concentration in mg/kg. Overabsorption should be noticeable if the white-line is intense, for example in a humic substance rich in S_{Ox} (Fig. 5b). At a concentration of 50,000 ppm, the reduction in amplitude approaches 20% in an organic matrix ($\mu(res)/\mu(tot) = 0.156$), which is huge, but is still not an issue in a sandy matrix ($\mu(res)/\mu(tot) = 0.042$), especially if it contains reduced sulfur. Overall, sulfur concentration in most natural matrices does not cause significant overabsorption, as long as the distribution of sulfur is uniform. In general this condition is realized in humic substances and their dissolved organic matter fractions, HA and FA, but rarely in soils and sediments, as discussed below.

3.6.2. Overabsorption in matrices with nanoparticulate sulfur

Commonly in reduced soil and sediment sulfur is speciated as pyrite (FeS_2) with a bimodal grain size distribution (Isaure et al., 2002). In the example shown in Fig. 7a, the bigger grains, which also contain Zn, have a diameter of about 30 μm and the smaller grains, more uniformly distributed in the fine clay matrix, have a dimension of about 3–5 μm. Manceau et al. (2002) calculated from the edge jump vs. μ(res)/μ(tot) contour map that a 1.5 μm thick grain of ZnS causes an absorbance reduction of 10% at the Zn K-edge (9659 eV) by overabsorption. When applied to the S K-edge of FeS_2 ($\mu(res)/\mu(tot) = 0.51$, Table 2), the calculation gives a maximum thickness of 0.1 μm for the same signal attenuation. This result is consistent with the 17% absorbance reduction calculated by Pickering et al. (2001) for 0.2 μm spheres of α-S_8, which approximate tiny round-shaped particles, and the 35% reduction measured by Burton et al. (2009) on 0.1–0.2 μm α-S_8 particles. Because overabsorption increases with sample thickness in concentrated samples, the deviation from the proportionality of the fluorescence signal measured on thicker particles to the resonant absorption (i.e., true absorbance) becomes

Table 2
Resonant absorption fraction for 20,000 ppm sulfur in quartz and in a humic substance, and for pyrite.

Atomic formula	Formula weight	μ(S−)[a]	μ(S+)[a]	μ(F)[a]	bkd[b]	μ(res)[c]	μ(res)/μ(tot)[d]
Quartz							
Si	28	1688		2017	103740		
O_2	32	380		461	26912		
S0.0381	1.22	217	2070	260	582	2261	0.017
Humic substance							
C4	48	162		197	17232		
O3	48	380		461	40368		
N0.07	0.980	258		314	561		
S0.062	1.98	217	2070	260	944	3669	0.058
Pyrite							
Fe	55.8	930		1112	113944		
S_2	64.1	217	2070	260	30576	118777	0.510

[a] Mass absorption coefficients (μ in cm^2/g) calculated below (2471 eV) and above (2472 eV) the S K-edge, and at the S(Kα) energy (2310 eV).
[b] bkd is (S−+F) * FW.
[c] Resonant part (defined in Manceau et al., 2002) of the total absorption above the edge. For the quartz matrix = (2070–217) × 1.22.
[d] Resonant absorption fraction to the total absorption. For the quartz matrix = 2261/(103740 + 26912 + 582 + 2261).

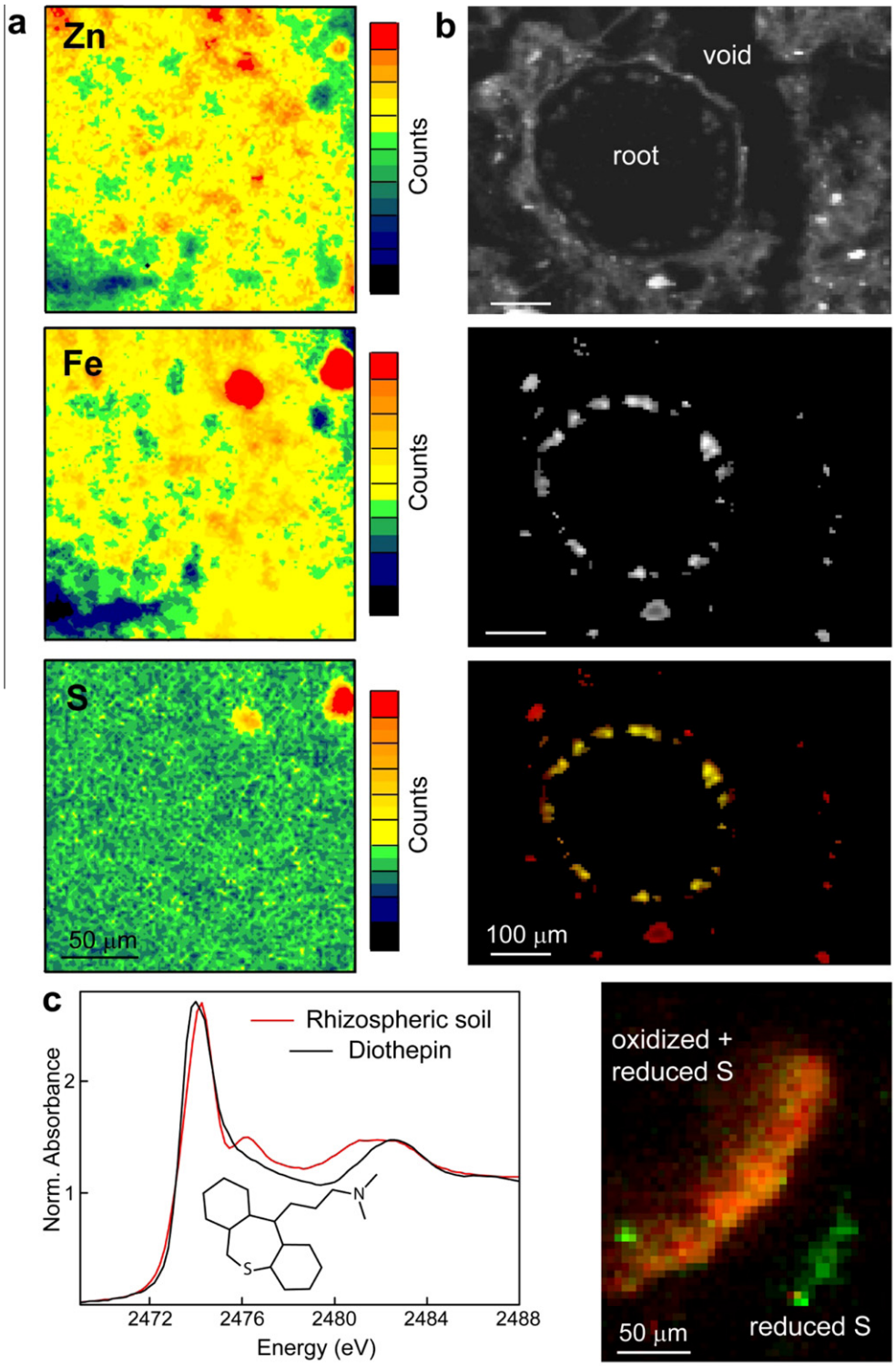


Fig. 7. (a) Micro X-ray fluorescence maps of the <2 μm fraction from a canal sediment. The clay suspension was deposited as an oriented preparation on a glass slide. $2 \times 2\ \mu m^2$ pixel size. Data from Isaure et al. (2002). (b) Distribution and speciation of S in a rhizospheric soil prepared as a 30 μm thick petrographic thin section. From top to bottom: Root and surrounding soil imaged with the elastic beam; distribution of S (white); bicolor-coded distribution of S (red) and Ca (green); bicolor-coded chemical map of oxidized (red) and reduced (green) sulfur. $5 \times 5\ \mu m^2$ pixel size. (c) Micro-XANES spectrum from the reduced sulfur form speciated as heterocyclic sulfur. $5 \times 5\ \mu m^2$ pixel size. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

rapidly large, and the measured spectrum increasingly distorted until the saturation limit is reached (Fig. 5b). Therefore, we propose that the spectral similarity of the 200, 2,000, and 20,000 ppm standard mixtures reported by Prietzel et al. (2011) to a large extent is a result of overabsorption by infinitely-thick samples. The pure sulfur compounds would have to be nanoparticulate in size for the XANES spectra to be essentially free of distortion. Uniform nanoparticles are impossible to obtain by conventional grinding techniques and some samples were likely infinitely thick with respect to the penetration depth of the incident beam (Pickering et al., 2001). This example for FeS_2 highlights the need to assess sample homogeneity by X-ray micro/nano-fluorescence or scanning electron microscopy before measurement.

Bulk XANES measurements on oxidized soils and sediments are also prone to the overabsorption distortion described here because sulfate is a major species most often present in concentrated grains, such as gypsum. Rhizospheric soils are even more heterogeneous spatially with, in addition, steep redox gradients between the near-root environment and the surrounding soil as a result of microbial activity and root respiration. This case of heterogeneity is exemplified in Fig. 7b and c with the distribution and variation of sulfur oxidation state in a rhizospheric soil from the floodplain of East Fork Poplar Creek downstream from Oak Ridge, TN. Sulfur is heterogeneously distributed with a distinct pattern of S and Ca co-localization inside the root, and patches of S-rich aggregates in the soil matrix. The oxidation state of sulfur near the root was imaged with a chemical X-ray micro-fluorescence map (Pickering et al., 2000; Freeman et al., 2006; Marcus, 2010; Banuelos et al., 2011) recorded at incident energies of 2470.40, 2474.40, and 2482.40 eV to obtain the background, and the S_{Red} and S_{Ox} forms, respectively. The chemical map displays two areas, one with a high concentration of S_{Red}, and another where the S_{Red} and S_{Ox} forms are co-associated spatially, with a predominance of the oxidized form. In addition, a hot-spot of ~10 μm of reduced sulfur is observed on the edge of the mixed-valence area. Micro-XANES was used to positively identify heterocyclic sulfur as the reduced sulfur functionality in this region of the rhizosphere. The immunity of the measurement to overabsorption was ascertained by measuring the micro-XANES spectra of several tiny particles and comparing the normalized amplitudes of their peak maxima. Comparison of the transmitted and fluorescence signals is not a good test for overabsorption, as explained below.

In summary, the previous examples of sulfur distribution and speciation, which are representative of the diversity of natural systems, show that the difficulty of modeling sulfur XANES spectra correctly is compounded by the measurement task itself as well as preparation and pre-characterization of the sample. Acquisition and interpretation of a meaningful S-XANES spectrum is not straightforward.

3.6.3. Correction for overabsorption

Accurate mathematical correction of overabsorption relies on precise knowledge of the nature and amount of all atoms in the path of the X-rays, which is difficult to measure on a real sample. Alternatively, an empirical correction function can be derived from Eq. (28) of Manceau et al. (2002) for a thick, flat sample. A physically-realistic form for thickness effects in fluorescence-yield measurement is given by:

$$y_{OA} = \frac{y}{1 - \beta + \beta y}$$

where y_{OA} is the pre-edge subtracted, post-edge normalized measured absorbance in the presence of overabsorption, y is the spectrum which would be measured in the absence of overabsorption, and β is a dimensionless parameter which measures the strength of the effect and depends on the absorption coefficients below the edge, above the edge, and at the fluorescence energy (Table 2), as well as the incidence and takeoff angles from the sample surface (generally 45°). After inversion, the equation is written as:

$$y = \frac{y_{OA}}{1 - \alpha + \alpha y_{OA}}$$

where $\alpha = -\beta/(1 - \beta)$, and has a positive value. The undesired amount of overabsorption can be suppressed from the measured spectrum by adjusting α. Overabsorption also may be simulated upon performing the correction to the spectrum using a negative α value.

3.6.4. Analysis of data affected by overabsorption

Fig. 5b shows the Pony Lake spectrum calculated with the correction factors $\alpha = -0.05$ and -0.1 to model spectral distortion caused by overabsorption. The calculation was carried out with the overabsorption algorithm implemented in the data analysis programs available on the 10.3.2 beamline where the data were collected (Marcus et al., 2004) (http://xraysweb.lbl.gov/uxas/index.htm). In agreement with theory, overabsorption produces a non-linear reduction of the signal amplitude. The reduction is higher the greater the normalized absorbance is than one. Because the S_{Ox} peak has a higher amplitude than the S_{Red} peak, overabsorption leads to underestimation of the S_{Ox}/S_{Red} ratio. Qualitatively, overabsorption produces the opposite effect on calculated values of S_i from the normalization of the Gaussian areas with calibration curves of lower slopes. One may conclude incorrectly that data collected in fluorescence-yield detection on concentrated samples can be safely analyzed with either the Orthous-Daunay, Waldo or Prietzel curves (Fig. 2). The two effects generally do not compensate each other, as shown with the $\alpha = -0.1$ PL-FA spectrum analyzed with the GCF method (Table EA-4). The sum of the two S_{Ox} functionalities (sulfate plus sulfonate) is underestimated by -16% and the sum of the two S_{Red} functionalities ($S_{Exo} + S_{Hetero}$) overestimated by +5% when the oversaturated fit values are normalized with the generic calibration curve. The −16% vs. +5% asymmetry of the two S_i% values reflects the non-linearity effect mentioned previously. The new S_i values match best the original S_i values calculated for the undistorted PL-FA spectrum when the Gaussian areas are normalized to the Sarret calibration curve. The Waldo curve yields S_{Ox} and S_{Red} values which differ from the real values (i.e., the undistorted

spectrum) by +16% and −8%. The correction is now too high and produces relative S_{Ox} and S_{Red} fractions which are, in this example, similarly biased as those obtained with the generic curve, but in the opposite direction.

In general, data affected by overabsorption can neither be analyzed with the LCF method using reference spectra that are also affected by overabsorption. Letting the spectral distortion caused by overabsorption at some concentration be represented by the operator OA(spectrum), one can assess the equality:

$$\mathrm{OA}(w_1\mathrm{S}_1 + w_2\mathrm{S}_2 + \ldots) = w_1\mathrm{OA}(\mathrm{S}_1) + w_2\mathrm{OA}(\mathrm{S}_2) + \ldots$$

where w_i and S_i are the fractions and spectra for species i. Unless one of the $w_i = 1$ and all others $= 0$, $OA(\sum w_i S_i) \neq \sum w_i OA(S_i)$, as shown in Fig. 5c with the $\alpha = -0.1$ PL-FA spectrum. The $\sum w_i \mathrm{OA}(\mathrm{S}_i)$ spectrum has a lower amplitude because each of the unweighted, undistorted S_i reference spectra has a higher amplitude, and hence, is more distorted by overabsorption than the humic substance spectrum. More problematic, the shape of the S_{Red} component is also modified because it is decomposed into several reduced S_i species, each of which often has a complex spectral shape not simply described by one single Gaussian function. Furthermore, if a fit is performed:

$$\mathrm{OA}(w_1\mathrm{S}_1 + w_2\mathrm{S}_2 + \ldots) = x_1\mathrm{OA}(\mathrm{S}_1) + x_2\mathrm{OA}(\mathrm{S}_2) + \ldots$$

with x_i as adjustable parameters, the OA(S_{Red}) and OA(S_{Ox}) components can be reproduced reasonably well, but with x_i values different from the w_i values. The S_{Exo} and S_{Hetero} fractions (60.7% and 4.1%, respectively) deviate significantly from the real values derived from the undistorted spectrum (47.0% and 18.6%, respectively). However, the total amount of reduced sulfur (S_{Red}) recovered from the fit of overabsorbed data to the combination of overabsorbed reference spectra is correct. In summary, it is advisable to always use reference spectra measured in TEY mode in the LCF analysis of a sample spectrum, whether or not it is affected by overabsorption. If some overabsorption occurs in the unknown, then its amplitude can be adjusted with the α correction factor described previously.

The same samples which tend to show overabsorption caused by spatial heterogeneity will often show a hole effect in transmission measurement, for the particles will never be uniform in thickness across the beam (Stern and Kim, 1981; Manceau et al., 2002). Hole effect distorts signal in the same way as overabsorption. Thus, in principle, comparison between transmission and fluorescence spectra (Prietzel et al., 2007) does not guarantee that the data are free of overabsorption. The empirical function defined above may be used to correct for hole effect. However, because it is difficult to correct overabsorbed data and hole effect accurately, it is preferable to avoid or at least minimize these problems at the sample preparation and measurement stages.

4. CONCLUSION

The GCF method of quantifying groups of sulfur species in humic substances is valid if the following three conditions are met: (1) the initial mathematical model is physically and chemically realistic and not overparameterized, (2) each oxidation state produces a single peak and, (3) each oxidation state has a weak post-edge oscillation which does not overlap greatly with the white-line of more oxidized species.

The first condition is realized by fixing the Gaussians to their nominal positions (Table EA-2) and constraining their widths. The second condition is never fully satisfied, at least in humic substances, because the spectra of several constitutive S_{Exo} moieties exhibit doublets. This causes an artificial increase in the S_{Hetero}/S_{Exo} ratio, especially when S_{Hetero} is minor. The third condition is always satisfied in IHSS samples, but not necessarily in samples in which one particular S_{Red} species prevails. In the latter case, only the LCF method will give correct results.

Although the LCF method has the superior advantage to be model-independent, its accuracy is limited by how well the reference spectra represent the unknown sample. The extended database compiled in this study covers a large range of sulfur moieties found in humic substances, earth and extraterrestrial materials, coal, bitumen and by-products, and biological samples (Sarret et al., 1999; Pickering et al., 2001; Fleet, 2005; Struis et al., 2008; Cai et al., 2009), and therefore can be used to analyze many types of materials without preliminary knowledge of the sample. Its main limitation is the transferability of the reference spectra. Caution should be taken to not use the reference spectra in the database to simulate data collected at lower or higher spectral resolution (Electronic annex). It is recommended to verify the transferability of the reference spectra on one or several reference compounds before fitting an unknown spectrum collected on another X-ray absorption spectrometer to the database.

ACKNOWLEDGMENTS

Matthew Marcus is thanked for his advice during data collection and analysis and for useful suggestions on general evaluation of overabsorption effects. Two anonymous reviewers provided thoughtful comments that were used in revising the manuscript. This work was supported by the Geobiology and Low-Temperature Geochemistry Program of the U.S. National Science Foundation under Grant EAR-0952311 and the Subsurface Biogeochemical Research Program of the Office of Biological and Environmental Research, Office of Science, U.S. Department of Energy under Grant DE-SC0001730. The ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of Energy under Contract DE-AC02-05CH11231 at the Lawrence Berkeley National Laboratory.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.gca.2012.09.033.

REFERENCES

Anxolabéhère-Mallart E., Glasser T., Frank P., Aliverti A., Zanetti G., Hedman B., Hodgson K. O. and Solomon E. I. (2001) Sulfur K-edge X-ray absorption spectroscopy of 2Fe-2S ferredoxin: covalency of the oxidized and reduced 2Fe forms and comparison to model complexes. *J. Am. Chem. Soc.* **123**, 5444–5452.

Banuelos G. S., Fakra S. C., Walse S. S., Marcus M. A., Yang S. I., Pickering I. J., Pilon-Smits E. A. H. and Freeman J. L. (2011) Selenium accumulation, distribution, and speciation in spineless prickly pear cactus: a drought- and salt-tolerant, selenium-enriched nutraceutical fruit crop for biofortified foods. *Plant Physiol.* **155**, 315–327.

Beauchemin S., Hesterberg D. A. and Beauchemin M. (2002) Principal component analysis approach for modeling sulfur K-XANES spectra of humic acids. *Soil Sci. Soc. Am. J.* **66**, 83–91.

Beauchemin S., Hesterberg D., Nadeau J. and McGeer J. C. (2004) Speciation of hepatic Zn in trout exposed to elevated waterborne Zn using X-ray absorption spectroscopy. *Environ. Sci. Technol.* **38**, 1288–1295.

Bostick B. C., Theissen K. M., Dunbar R. B. and Vairavamurthy M. A. (2005) Record of redox status in laminated sediments from Lake Titicaca: a sulfur K-edge X-ray absorption near edge structure (XANES) study. *Chem. Geol.* **219**, 163–174.

Brown A., McKnight D. M., Chin Y. P., Roberts E. C. and Uhle M. (2004) Chemical characterization of dissolved organic material in Pony Lake, a saline coastal pond in Antarctica. *Mar. Chem.* **89**, 327–337.

Burton E. D., Bush R. T., Sullivan L. A., Hocking R. K., Mitchell D. r. G., Johnston S. G., Fitzpatrick R. W., Raven M., McClure S. and Jang L. Y. (2009) Iron-monosulfide oxidation in natural sediments: Resolving microbially mediated S transformations using XANES, electron microscopy, and selective extractions. *Environ. Sci. Technol.* **43**, 3128–3134.

Cai J. H., Morris E. and Jia C. Q. (2009) Sulfur speciation in fluid coke and its activation products using K-edge X-ray absorption near edge structure spectroscopy. *J. Sulfur Chem.* **30**, 555–569.

Chauvistré R., Hormes J., Hartmann E., Etzenbach N., Hosch R. and Hahn J. (1997) Sulfur K-shell photoabsorption spectroscopy of the sulfanes R–Sn–R, $n = 2$–4. *Chem. Phys.* **223**, 293–302.

Einsiedl F., Schäfer T. and Northrup P. (2007) Combined sulfur K-edge XANES spectroscopy and stable isotope analyses of fulvic acids and groundwater sulfate identify sulfur cycling in a karstic catchment area. *Chem. Geol.* **238**, 268–276.

Fimmen R. L., Cory R. M., Chin Y. P., Trouts T. D. and McKnight D. M. (2007) Probing the oxidation–reduction properties of terrestrially and microbially derived dissolved organic matter. *Geochim. Cosmochim. Acta* **71**, 3003–3015.

Fleet M. E. (2005) XANES spectroscopy of sulfur in earth materials. *Can. Mineral.* **43**, 1811–1838.

Foreman C. M., Dieser M., Greenwood M., Cory R. M., Laybourn-Parry J., Lisle J. T., Jaros C., Miller P. L., Chin Y. P. and McKnight D. M. (2011) When a habitat freezes solid: microorganisms over-winter within the ice column of a coastal Antarctic lake. *FEMS Microbiol. Ecol.* **76**, 401–412.

Fors Y., Jalilehvand F., Risberg E. D., Björdal C., Phillips E. and Sandström M. (2012) Sulfur and iron analyses of marine archaeological wood in shipwrecks from the Baltic Sea and Scandinavian waters. *J. Archeol. Sci.* **39**, 2521–2532.

Frank P., Hedman B., Carlson R. M. K., Tyson T. A., Roe L. and Hodgson K. O. (1987) A large reservoir of sulfate and sulfonate resides within plasma cells from Ascidia ceratodes, revealed by X-ray absorption near-edge structure spectroscopy. *Biochemistry* **26**, 4975–4979.

Frank P., Hedman B. and Hodgson K. O. (1999) Sulfur allocation and vanadium-sulfate interactions in whole blood cells from the tunicate Ascidia ceratodes, investigated using X-ray absorption spectroscopy. *Inorg. Chem.* **38**, 260–270.

Freeman J. L., Zhang L. H., Marcus M. A., Fakra S., McGrath S. P. and Pilon-Smits E. A. (2006) Spatial imaging, speciation, and quantification of selenium in the hyperaccumulator plants *Astragalus bisulcatus* and *Stanleya pinnata*. *Plant Physiol.* **142**, 124–134.

George G. N. and Gorbaty M. L. (1989) Sulfur K-edge x-ray absorption spectroscopy of petroleum asphaltenes and model compounds. *J. Am. Chem. Soc.* **111**, 3182–3186.

George G. N., Gorbaty M. L., Kelemen S. R. and Sansone M. (1991) Direct determination and quantification of sulfur forms in coals from the Argonne Premium Sample Program. *Energy Fuels* **5**, 93–97.

Huffman G. P., Mitra S., Huggins F. E., Shah N., Vaidya S. and Lu F. (1991) Quantitative analysis of all major forms of sulfur in coal by X-ray absorption fine structure spectroscopy. *Energy Fuels* **5**, 574–581.

Huffman G. P., Shah N., Huggins F. E., Stock L. M., Chatterjee K., Kilbane J. J., Chou M. M. and Buchanan D. H. (1995) Sulfur speciation of desulfurized coals by XANES spectroscopy. *Fuel* **74**, 549–555.

Huggins F. E., Seidu L. B. A., Shah N., Huffman G. P., Honaker R. Q., Kyger J. R., Higgins B. L., Robertson J. D., Pal S. and Seehra M. S. (2009) Elemental modes of occurrence in an Illinois #6 coal and fractions prepared by physical separation techniques at a coal preparation plant. *Int. J. Coal Geol.* **78**, 65–76.

Hundal L. S., Carmo A. M., Bleam W. L. and Thompson M. L. (2000) Sulfur in biosolids-derived fulvic acid: characterization by XANES spectroscopy and selective dissolution approaches. *Environ. Sci. Technol.* **34**, 5184–5188.

Hutchison K. J., Hesterberg D. and Chou J. W. (2001) Stability of reduced organic sulfur in humic acid as affected by aeration and pH. *Soil Sci. Soc. Am. J.* **65**, 704–709.

Isaure M. P., Laboudigue A., Manceau A., Sarret G., Tiffreau C., Trocellier P., Lamble G., Hazemann J. L. and Chateigner D. (2002) Quantitative Zn speciation in a contaminated dredged sediment by μPIXE, μSXRF, EXAFS spectroscopy and principal component analysis. *Geochim. Cosmochim. Acta* **66**, 1549–1567.

Jalilehvand F. (2006) Sulfur: not a "silent" element any more. *Chem. Soc. Rev.* **35**, 1256–1268.

Kasrai M., Bancroft G. M., Brunner R. W., Jonasson R. G., Brown J. R., Tan K. H. and Feng X. H. (1994) Sulphur speciation in bitumens and asphaltenes by X-ray absorption fine structure spectroscopy. *Geochim. Cosmochim. Acta* **58**, 2865–2872.

Kasrai M., Brown J. R., Bancroft G. M., Yin Z. and Tan K. H. (1996) Sulphur characterization in coal from X-ray absorption near edge spectroscopy. *Int. J. Coal Geol.* **32**, 107–135.

Kelemen S. R., Afeworki M., Gorbaty M. L., Sansone M., Kwiatek P. J., Walters C. C., Freud H., Siskin M., Bence A. E., Curry D. J., Solum M., Pugmire R. J., Vandenbroucke M., Leblond M. and Behar F. (2007) Direct characterization of kerogen by X-ray and solid-state ^{13}C nuclear magnetic resonance methods. *Energy Fuels* **21**, 1548–1561.

Kelemen S. R., Sansone M., Walters C. C., Kwiatek P. J. and Bolin T. (2012) Thermal transformations of organic and inorganic sulfur in Type II kerogen quantified by S-XANES. *Geochim. Cosmochim. Acta* **83**, 61–78.

Koudouna E., Veronesi G., Patel I. I., Cotte M., Knupp C., Martin F. L. and Quantock A. (2012) Chemical composition and sulfur speciation in bulk tissue by X-ray spectroscopy and X-ray

microscopy: corneal development during embryogenesis. *Biophys. J.* **103**, 357–364.

Manceau A. and Gates W. (1997) Surface structural model for ferrihydrite. *Clay Clay Mineral.* **43**, 448–460.

Manceau A., Marcus M. A. and Grangeon S. (2012) Determination of Mn valence states in mixed-valent manganates by XANES spectroscopy. *Am. Mineral.* **97**, 816–827.

Manceau A., Marcus M. A. and Tamura N. (2002) Quantitative speciation of heavy metals in soils and sediments by synchrotron X-ray techniques. In *Applications of Synchrotron Radiation in Low-Temperature Geochemistry and Environmental Science*, vol. 49 (eds. P. A. Fenter, M. L. Rivers, N. C. Sturchio and S. R. Sutton). Mineralogical Society of America, Washington, DC.

Mao J., Cory R. M., McKnight D. M. and Schmidt-Rohr K. (2007) Characterization of a nitrogen-rich fulvic acid and its precursor algae from solid state NMR. *Org. Geochem.* **38**, 1277–1292.

Marcus M. A., MacDowell A. A., Celestre R., Manceau A., Miller T., Padmore H. A. and Sublett R. E. (2004) Beamline 10.3.2 at ALS: a hard X-ray microprobe for environmental and materials sciences. *J. Synch. Rad.* **11**, 239–247.

Marcus M. A., Westphal A. J. and Fakra S. C. (2008) Classification of Fe-bearing species from K-edge XANES data using two-parameter correlation plots. *J. Synch. Rad.* **15**, 463–468.

Marcus M. A. (2010) X-ray photon-in/photon-out methods for chemical imaging. *Trends Analyt. Chem.* **29**, 508–517.

Martinez C. E., McBride M. B., Kandianis M. T., Duxbury J. M., Yoon S. J. and Bleam W. F. (2002) Zinc-sulfur and cadmium-sulfur association in metalliferous peats: Evidence from spectroscopy, distribution coefficients, and phytoavailability. *Environ. Sci. Technol.* **36**, 3683–3689.

McKnight D. M., Andrews E. D., Spulding S. A. and Aiken G. R. (1994) Aquatic fulvic acids in algal-rich Antarctic ponds. *Limnol. Oceanogr.* **39**, 1972–1979.

Morra M. J., Fendorf S. E. and Brown P. D. (1997) Speciation of sulfur in humic and fulvic acids using X-ray absorption near-edge structure (XANES) spectroscopy. *Geochim. Cosmochim. Acta* **61**, 683–688.

Myneni S. C. B. (2002) Soft X-ray spectroscopy and spectromicroscopy studies of organic molecules in the environment. In *Applications of Synchrotron Radiation in Low-Temperature Geochemistry and Environmental Science*, vol. 49 (eds. P. A. Fenter, M. L. Rivers, N. C. Sturchio and S. R. Sutton). Mineralogical Society of America, Washington, DC.

Nagy K. L., Manceau A., Gasper J. D., Ryan J. N. and Aiken G. R. (2011) Metallothionein-like multinuclear clusters of mercury(II) and sulfur in peat. *Environ. Sci. Technol.* **45**, 7298–7306.

Olivella M. A., del Rio J. C., Palacios J. M., Vairavamurthy A. and de las Heras F. X. C. (2002a) Characterization of humic acid from leonardite coal: an integrated study of PY-GC-MS, XPS and XANES techniques. *J. Anal. Appl. Pyrol.* **63**, 59–68.

Olivella M. A., Palacios J. M., Vairavamurthy A., del Rio J. C. and de las Heras F. X. C. (2002b) A study of sulfur functionalities in fossil fuels using destructive- (ASTM and Py-GC-MS) and non-destructive- (SEM-EDX, XANES and XPS) techniques. *Fuel* **81**, 405–411.

Orthous-Daunay F. R., Quirico E., Lemelle L., Beck P., deAndrade V., Simionovici A. and Derenne S. (2010) Speciation of sulfur in the insoluble organic matter from carbonaceous chondrites by XANES spectroscopy. *Earth Planet. Sci. Lett.* **300**, 321–328.

Pattanaik S., Huggins F. E., Huffman G. P., Linak W. P. and Miller C. A. (2007) XAFS studies of nickel and sulfur speciation in residual oil fly-ash particulate matters (ROFA PM). *Environ. Sci. Technol.* **41**, 1104–1110.

Pickering I. J., Prince R. C., Salt D. E. and George G. N. (2000) Quantitative, chemically specific imaging of selenium transformation in plants. *Proc. Natl. Acad. Sci. USA* **97**, 10717–10722.

Pickering I. J., George G. N., Yu E. Y., Brune D. C., Tuschak C., Overmann J., Beatty J. T. and Prince R. C. (2001) Analysis of sulfur biochemistry of sulfur bacteria using X-ray absorption spectroscopy. *Biochemistry* **40**, 8138–8145.

Pickering I. J., Prince R. C., Divers T. and George G. N. (1998) Sulfur K-edge X-ray absorption spectroscopy for determining the chemical speciation of sulfur in biological systems. *FEBS Lett.* **441**, 11–14.

Prange A., Dahl C., Trüper H. G., Behnke M., Hahn J., Modrow H. and Hormes J. (2002) Investigation of S–H bonds in biologically important compounds by sulfur K-edge X-ray absorption spectroscopy. *Eur. Phys. J. D* **20**, 589–596.

Prietzel J., Botzaki A., Tyufekchieva N., Brettholle M., Thieme J. and Klysubun W. (2011) Sulfur speciation in soil by S K-edge XANES spectroscopy: comparison of spectral deconvolution and linear combination fitting. *Environ. Sci. Technol.* **45**, 2878–2886.

Prietzel J., Thieme J., Salome M. and Knicker H. (2007) Sulfur K-edge XANES spectroscopy reveals differences in sulfur speciation of bulk soils, humic acid, fulvic acid, and particle size separates. *Soil Biol. Biochem.* **39**, 877–890.

Prietzel J., Thieme J., Tyufekchieva N., Paterson D., McNulty I. and Kögel-Knabner I. (2009) Sulfur speciation in well-aerated and wetland soils in a forested catchment assessed by sulfur K-edge X-ray absorption near-edge spectroscopy (XANES). *J. Plant Nutr. Soil Sci.* **172**, 393–409.

Qian J., Skyllberg U., Frech W., Bleam W. F., Bloom P. R. and Petit P. E. (2002) Bonding of methyl mercury to reduced sulfur groups in soil and stream organic matter as determined by X-ray absorption spectroscopy and binding affinity studies. *Geochim. Cosmochim. Acta* **66**, 3873–3885.

Ravel B. and Newville M. (2005) ATHENA, ARTEMIS, HEPHAESTUS: data analysis for X-ray absorption spectroscopy using IFEFFIT. *J. Synch. Rad.* **12**, 537–541.

Sarret G., Conna J., Kasrai M., Bancroft G. M., Charrié-Duhaut A., Lemoine S., Adam P., Albrecht P. and Eybert-Bérard L. (1999) Chemical forms of sulfur in geological and archeological asphaltenes from Middle East, France, and Spain determined by sulfur K- and L-edge X-ray absorption near-edge structure spectroscopy. *Geochim. Cosmochim. Acta* **63**, 3767–3779.

Sarret G., Mongenot T., Connan J., Derenne S., Kasrai M., Bancroft G. M. and Largeau C. (2002) Sulfur speciation in kerogens of the Orbagnoux deposit (Upper Kimmeridgian, Jura) by XANES spectroscopy and pyrolysis. *Org. Geochem.* **33**, 877–895.

Schäfer T., Buckau G., Artinger R., Kim J. I., Gyer S., Wolf M., Bleam W. F., Wirick S. and Jacobsen C. (2005) Origin and mobility of fulvic acids in the Gorleben aquifer system: implications from isotopic data and carbon/sulfur XANES. *Org. Geochem.* **36**, 567–582.

Schroth A. W., Bostick B. C., Graham M., Kaste J. M., Mitchell M. J. and Friedland A. J. (2007) Sulfur species behavior in soil organic matter during decomposition. *J. Geophys. Res.* **112**, 1–10.

Skyllberg U., Xia K., Bloom P. R., Nater E. A. and Bleam W. F. (2000) Binding of mercury(II) to reduced sulfur in soil organic matter along upland-peat soil transects. *J. Environ. Qual.* **29**, 855–865.

Smith D. S., Bella R. A. and Kramerb J. R. (2002) Metal speciation in natural waters with emphasis on reduced sulfur groups as strong metal binding sites. *Comp. Biochem. Physiol. C* **133**, 65–74.

Solomon D., Lehmann J., Lobe I., Martinez C. E., Tveitnes S., Du Preez C. C. and Amelung W. (2005) Sulphur speciation and biogeochemical cycling in long-term arable cropping of subtropical soils: evidence from wet-chemical reduction and S K-edge XANES spectroscopy. *Eur. J. Soil Sci.* **56**, 621–634.

Solomon D., Lehmann J. and Martinez C. E. (2003) Sulfur K-edge XANES spectroscopy as a tool for understanding sulfur dynamics in soil organic matter. *Soil Sci. Soc. Am. J.* **67**, 1721–1731.

Stern E. A. and Kim K. (1981) Thickness effect on the extended X-ray absorption fine structure amplitude. *Phys. Rev. B* **23**, 3781–3787.

Struis R. P. W. J., Ludwig C., Barrelet T., Krähenbühl U. and Rennenberg H. (2008) Studying sulfur functional groups in Norway spruce year rings using S L-edge total electron yield spectroscopy. *Sci. Total Environ.* **403**, 196–206.

Szulczewski M. D., Helmke P. A. and Bleam W. F. (2001) XANES spectroscopy studies of Cr(VI) reduction by thiols in organo-sulfur compounds and humic substances. *Environ. Sci. Technol.* **35**, 1134–1141.

Taghiei M. M., Huggins F. E., Shah N. and Huffman G. P. (1992) In situ X-ray absorption fine structure spectroscopy investigation of sulfur functional groups in coal during pyrolysis and oxidation. *Energy Fuels* **6**, 293–300.

Tröger L., Arvanitis D., Baberschke K., Michaeis H., Grimm U. and Zschech E. (1992) Full correction of the self-absorption in soft-fluorescence extended x-ray absorption fine structure. *Phys. Rev. B* **46**, 3283–3289.

Vairavamurthy A. (1998) Using X-ray absorption to probe sulfur oxidation states in complex molecules. *Spectrochem. Acta A* **54**, 2009–2017.

Vairavamurthy A., Manowitz B., Luther G. W. and Jeon Y. (1993) Oxidation state of sulfur in thiosulfate and implications for anaerobic energy metabolism. *Geochim. Cosmochim. Acta* **57**, 1619–1623.

Vairavamurthy A., Zhou W. Q., Eglinton T. and Manowitz B. (1994) Sulfonates: a novel class of organic sulfur compounds in marine sediments. *Geochim. Cosmochim. Acta* **58**, 4681–4687.

Vairavamurthy M. A., Maletic D., Wang S., Manowitz B., Eglinton T. and Lyons T. (1997) Characterization of sulfur-containing functional groups in sedimentary humic substances by X-ray absorption near-edge structure spectroscopy. *Energy Fuels* **11**, 546–553.

Waldo G. S., Carlson R. M. K., Moldowan J. M., Peters K. E. and Penner-Hahn J. E. (1991) Sulfur speciation in heavy petroleums: Information from X-ray absorption near-edge structure. *Geochim. Cosmochim. Acta* **55**, 801–814.

Wijaya N. and Zhang L. (2012) Generation of ultra-clean fuel from Victorian brown coal – synchrotron XANES study on the evolution of sulphur in Victorian brown coal upon hydrothermal upgrading treatment and thermal pyrolysis. *Fuel* **99**, 217–225.

Wiltfong R., Mitra-Kirtley S., Mullins O. C., Andrews B., Fujisawa G. and Larsen J. W. (2005) Sulfur speciation in different kerogens by XANES spectroscopy. *Energy Fuels* **19**, 1971–1976.

Xia K., Weesner F., Bleam W. F., Bloom P. R., Skyllberg U. L. and Helmke P. A. (1998) XANES studies of oxidation states of sulfur in aquatic and soil humic substances. *Soil Sci. Soc. Am. J.* **62**, 1240–1246.

Yoon S. J., Diener L. M., Bloom P. R., Nater E. A. and Bleam W. F. (2005) X-ray absorption studies of CH_3Hg^+-binding sites in humic substances. *Geochim. Cosmochim. Acta* **69**, 1111–1121.

Zhao F. J., Lehmann J., Solomon D., Fox M. A. and McGrath S. P. (2006) Sulphur speciation and turnover in soils: evidence from sulphur K-edge XANES spectroscopy and isotope dilution studies. *Soil Biol. Biochem.* **38**, 1000–1007.

Associate editor: Stephan Kraemer

Electronic Annex

Quantitative analysis of sulfur functional groups in natural organic matter by XANES spectroscopy

Alain Manceau [a,*] and Kathryn L. Nagy [b]

[a] *ISTerre, CNRS and Université de Grenoble 1, F-38041 Grenoble Cedex 9, France.*

[b] *Department of Earth and Environmental Sciences, MC-186, 845 West Taylor Street, University of Illinois at Chicago, Chicago, IL 60607, United States.*

Table of Contents

1. Tables ……..…….....….. 2

2. Gaussian decomposition of some references ……. 14
Figure EA-1

3. Gaussian decomposition of the IHSS references with the convergence-optimized method ……. 14
Figure EA-2

4. Calibration curves ……...… 15
Figure EA-3

5. Combo fit of the IHSS references …….. 16
Figure EA-4

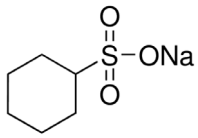
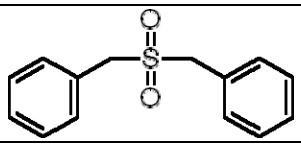
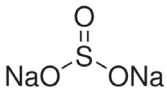
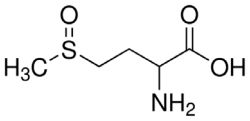
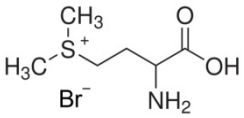
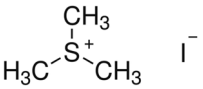
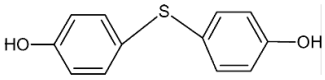
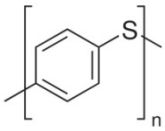
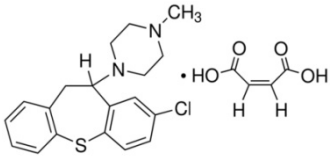
6. Evaluation and discussion of previous S-XANES data analyses …… 18
6.1. IHSS humic substances ……...… 18
Figure EA-5
Figure EA-6
Figure EA-7
6.2. Alais meteorite (carbonaceous chondrite) …….… 21
Figure EA-8
6.3. Sulfur speciation in soil ……... 22
Figure EA-9
6.4. Incorrect positioning of the arctangents ……...… 23
Figure EA-10
Figure EA-11

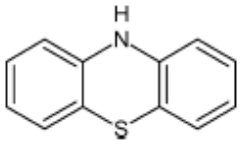
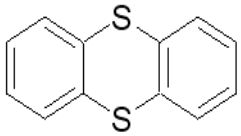
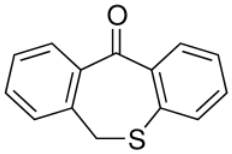
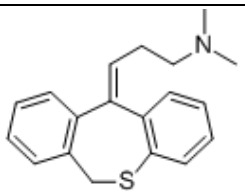
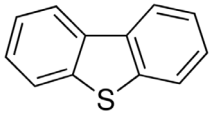
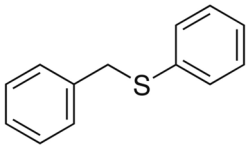
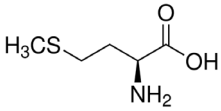
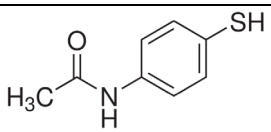
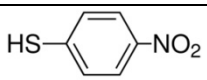
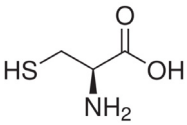
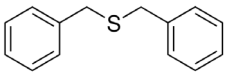
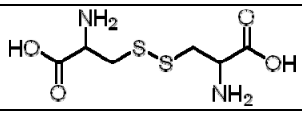
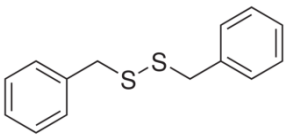
7. Effect of spectrometer resolution on the shape of S-XANES spectra …… 24
Figure EA-12

8. References …… 25

1. Tables

Table EA-1. List of sulfur model compounds contained in the database. Filenames are in parentheses.

Compound	CAS number /Origin	(Molecular) formula	Structural formula	Peak max. (eV)
Sodium sulfate (Sulfate-Na)[1]	7757-82-6	Na_2SO_4		2482.75
Magnesium sulfate (Sulfate-Mg)[1]	7487-88-9	$MgSO_4$		2482.75
Gypsum (Gypsum)[1]	Naica, Chihuahua, Mexico	$CaSO_4{\cdot}2H_2O$		
Sodium cyclohexanesulfonate (Sulfonate)[1]	13893-74-8	$C_6H_{11}NaO_3S$		2481.30
Dibenzyl sulfone (Sulfone)[1]	620-32-6	$C_{14}H_{14}O_2S$		2479.60
Sodium sulfite (Sulfite)	7757-83-7	Na_2SO_3		
DL Methionine sulfoxide (Sulfoxide)[1]	62697-73-8	$C_5H_{11}NO_3S$		2476.40
DL-Methionine methylsulfonium bromide (Sulfonium-Met)	2766-51-0	$C_6H_{14}NO_2S$		
Trimethylsulfonium iodide (Sulfonium-TriMet)[1]	2181-42-2	C_3H_9IS		2475.60
4,4'-thiodiphenol (PhSPh)	2664-63-3	$C_{12}H_{10}O_2S$		
Poly(1,4-phenylene sulphide) (PhSn)	25212-74-2	$(C_6H_4S)_n$		
Octoclothepin maleate (Clothepin)	4789-68-8	$C_{19}H_{21}ClN_2S{\cdot}C_4H_4O_4$		

Phenothiazine (Thiazine)	92-84-2	$C_{12}H_9NS$		
Thianthrene (Thianthrene)	92-85-3	$C_{12}H_8S_2$		
6,11-Dihydrodibenzothiepin (Thiepin)	1531-77-7	$C_{14}H_{10}OS$		
Dothiepin hydrochloride (Dothiepin)[1]	897-15-4	$C_{19}H_{21}NS{\cdot}HCl$		2474.00
Dibenzothiophene (Thiophene)[1]	132-65-0	$C_{12}H_8S$		2474.10
Benzyl phenyl sulfide (PhCH2SPh)[1]	831-91-4	$C_{13}H_{12}S$		2473.75
L-methionine (Methionine)[1]	63-68-3	$C_5H_{11}NO_2S$		2473.65
4-Acetamidothiophenol (Ph-SH-1)	1126-81-4	C_8H_9NOS		
4-Nitrobenzenethiol (Ph-SH-2)	1849-36-1	$C_6H_5NO_2S$		
L-cysteine (Cysteine)[1]	52-90-4	$C_3H_7NO_2S$		2473.45
Dibenzyl sulfide (Sulfide)[1]	538-74-9	$C_{14}H_{14}S$		2473.30
DL-cystine (Cystine)	923-32-0	$C_6H_{12}N_2O_4S_2$		
Dibenzyl disulfide (Disulfide)[1]	150-60-7	$C_{14}H_{14}S_2$		2473.10

Elemental S (S0)	Baja, California	S		
Pyrite (Pyrite)	Huanuco, Peru	FeS_2		
Pentlandite (Pentlandite)	Sudbury District, Canada	$(Fe,Ni,Co)_{9\pm x}S_8$		
Pyrrhotite (Pyrrhotite)	Chihuahua, Mexico	$Fe_{1-x}S$		

1- References with a single maximum used to calculate the new calibration curve shown in Fig. EA-3.

Table EA-2. Sulfur species and atomic fractions (%) used to calculate the theoretical spectrum shown in Figure 1a and parameter values from the simulations shown in Figures 1b to 1e.

	Theoretical spectrum			Model-fit 1			Model-fit 2			Model-fit 3			Model-fit 4		
	Energy[a]	FWHM	%	Energy	FWHM	%	Energy	FWHM	%	Energy	FWHM	%	Energy	FWHM	%
Arctg1				2475.70	**1.25**		2477.00	**1.25**		2475.70	**1.25**		2475.65	**0.62**	
Arctg2				2483.80	**1.25**		2483.80	**1.25**		2482.60	**1.25**		2484.05	**0.62**	
Exocyclic	2473.20	1.60	25.0	2473.25	**1.60**	25.3	2473.25	**1.60**	23.8	2473.25	**1.60**	25.6	2473.25	**1.66**	25.7
Heterocyclic	2474.40	1.60	35.0	2474.45	**1.60**	34.7	2474.45	**1.60**	33.3	2474.45	**1.60**	34.8	2474.50	**1.66**	35.2
Sulfoxide	2476.40	1.60	5.0	2476.30	**1.60**	5.0	2476.30	**1.60**	10.9	2476.35	**1.60**	4.1	2476.50	**1.66**	5.1
Sulfone	2479.60	1.60	6.0	2479.60	**1.60**	5.8	2479.60	**1.60**	5.1	2479.60	**1.60**	5.9	2479.55	**1.66**	5.9
Sulfonate	2481.30	1.60	17.0	2481.30	**1.60**	16.8	2481.30	**1.60**	15.4	2481.30	**1.60**	16.9	2481.30	**1.66**	16.7
Sulfate	2482.75	1.60	12.0	2482.80	**1.60**	12.5	2482.80	**1.60**	11.4	2482.70	**1.60**	12.6	2482.85	**1.66**	11.4

[a] Nominal energy position determined from the analysis by micro-XANES of soil humic substances enriched in one type of functionality. Sulfite is at 2478.55 eV and sulfonium at 2475.60 eV . Energy and full width at half-maximum (FWHM) are in eV. Each species was represented by one Gaussian (FWHM=1.6 eV) centered at its nominal position and one arctangent function (FWHM=0.9 eV) positioned 1.3 eV above the energy of the corresponding Gaussian. In all Tables, parameters in bold were co-varied and those in bold and underlined were fixed. The fit values highlighted in green are identical or close to real values, and those highlighted in magenta are much different.

Table EA-3. Six, statistically equivalent, decompositions of the sulfur K-edge XANES spectrum for the Elliott Soil humic acid (ES-HA) shown in Figure 3. See Table EA-2 footnote for additional coding.

	Model-fit 1				Model-fit 2				Model-fit 3				Model-fit 4				Model-fit 5				Model-fit 6			
	Energy	FWHM	Amp	%	Energy	FWHM	Amp	%	Energy	FWHM	Amp	%	Energy	FWHM	Amp	%	Energy	FWHM	Amp	%	Energy	FWHM	Amp	%
Arctg1	2475.74	**0.42**	0.69		2477.20	**0.63**	0.65		2475.75	**0.51**	0.73		2475.75	**1.00**	**0.69**		2475.70	**0.42**	**0.50**		2475.75	**0.41**	0.45	
Arctg2	2482.30	**0.42**	0.60		2482.25	**0.63**	0.64		2483.90	**0.51**	0.57		2482.20	**1.00**	**0.60**		2482.30	**0.42**	0.78		2482.30	**0.41**	0.82	
Exocyclic	2473.10	**1.75**	0.63	21.4	2473.00	**1.62**	0.64	18.7	**2473.10**	**1.73**	0.63	20.9	2473.00	**1.60**	0.63	20.2	2473.05	**1.70**	0.62	19.4	**2473.10**	**1.71**	0.66	19.7
Heterocyclic	**2474.50**	**1.75**	1.16	27.1	2474.45	**1.62**	1.23	24.7	**2474.50**	**1.73**	1.16	26.4	2474.45	**1.60**	1.19	26.0	2474.45	**1.70**	1.21	25.6	**2474.50**	**1.71**	1.17	24.1
Sulfoxide	**2476.40**	**1.75**	0.31	5.1	2476.40	**1.62**	0.94	13.0	**2476.40**	**1.73**	0.28	4.5	2476.25	**1.60**	0.42	6.6	2476.45	**1.70**	0.49	7.2	**2476.40**	**1.71**	0.54	7.8
Sulfite																					**2478.55**	**2.16**	0.25	3.5
Sulfone	**2479.60**	**2.05**	0.33	4.2	2479.10	**2.04**	0.29	3.6	**2479.60**	**1.92**	0.30	3.6	2479.10	**2.03**	0.24	3.2	2478.90	**2.14**	0.39	5.2	**2479.60**	**2.16**	0.31	3.8
Sulfonate	**2481.30**	**2.05**	2.22	24.3	2481.20	**2.04**	2.23	22.3	2481.20	**1.92**	2.22	22.6	2481.15	**2.03**	2.16	24.0	2481.20	**2.14**	2.30	24.5	**2481.30**	**2.16**	2.41	25.1
Sulfate	**2482.75**	**2.05**	1.83	17.8	2482.70	**2.04**	1.99	17.6	2482.80	**1.92**	2.48	22.1	2482.70	**2.03**	2.05	20.0	2482.70	**2.14**	1.92	18.0	**2482.75**	**2.16**	1.74	16.0

Table EA-4. Optimal parameter values for the GCF model fits to the IHSS data, and atomic fractions (%) of sulfur functionalities after correction of the absorption cross-sections with five different calibration curves. See Table EA-2 footnote for additional coding.

SR-HA	Energy	FWHM	Amp	Area	Generic (%)	Δ/2 (%)[a]	Precision (%)[b]	This study (%)	Huffman (%)	Sarret (%)	Prietzel (%)	Waldo (%)
A	2475.75	**0.66**	0.79									
A	2482.50	**0.66**	0.43									
G	2473.00	**2.21**	0.54	1.27	23.3	1.5	6.5	23.8	24.4	21.4	19.1	17.5
G	**2474.40**	**2.21**	1.39	3.28	41.1	0.5	1.3	41.2	41.4	40.3	38.9	37.6
G	**2476.40**	**2.21**	0.18	0.42	3.6	0.1	2.4	3.6	3.6	3.8	3.9	3.9
G	**2479.60**	**2.21**	0.40	0.94	5.4	0.3	5.2	5.3	5.2	5.7	6.2	6.6
G	**2481.30**	**2.21**	1.67	3.94	19.3	1.2	6.1	18.9	18.4	20.8	23.0	24.7
G	**2482.75**	**2.21**	0.72	1.69	7.3	0.5	6.7	7.2	7.0	8.0	8.9	9.7
SR-FA												
A	2475.55	**0.83**	0.68									
A	2482.45	**0.83**	0.55									
G	2473.15	**2.12**	0.55	1.23	23.1	1.5	6.7	23.7	24.3	21.2	18.8	17.1
G	**2474.40**	**2.12**	1.03	2.32	31.3	0.7	2.2	31.5	31.7	30.3	28.8	27.5
G	**2476.40**	**2.12**	0.14	0.31	2.8	0.1	1.5	2.8	2.8	2.9	2.9	2.9
G	**2479.60**	**2.12**	0.52	1.18	7.3	0.3	4.3	7.2	7.1	7.7	8.2	8.6
G	**2481.30**	**2.12**	2.14	4.83	25.4	1.3	5.1	25.0	24.5	27.1	29.4	31.2
G	**2482.75**	**2.12**	0.96	2.16	10.1	0.6	5.7	9.9	9.7	10.8	11.9	12.7
WP-HA												
A	2475.77	**0.76**	0.73									
A	2484.40	**0.76**	0.45									
G	**2473.20**	**2.32**	0.47	1.17	20.8	1.4	6.5	21.3	21.8	19.1	17.0	15.6
G	**2474.40**	**2.32**	1.11	2.74	35.5	0.8	2.3	35.8	36.1	34.5	32.7	31.2
G	**2476.40**	**2.32**	0.19	0.46	4.1	0.1	1.4	4.1	4.1	4.2	4.2	4.2
G	**2479.60**	**2.32**	0.26	0.64	3.8	0.2	4.2	3.7	3.7	4.0	4.3	4.5
G	**2481.30**	**2.32**	1.08	2.69	13.6	0.7	5.1	13.4	13.1	14.5	15.8	16.7
G	2482.75	2.32	1.99	4.92	22.1	1.2	5.6	21.7	21.2	23.7	26.0	27.9
WP-FA												
A	2475.80	**0.93**	0.83									
A	2482.37	**0.93**	0.37									
G	2473.05	**2.30**	0.63	1.55	27.1	1.6	5.7	27.7	28.3	25.1	22.7	21.1
G	**2474.40**	**2.30**	1.42	3.47	42.1	0.3	0.8	42.2	42.3	41.6	40.6	39.5
G	**2476.40**	**2.30**	0.23	0.55	4.6	0.1	2.9	4.5	4.5	4.8	4.9	5.0
G	**2479.60**	**2.30**	0.35	0.85	4.7	0.3	5.8	4.6	4.5	5.1	5.5	5.9
G	**2481.30**	**2.30**	1.19	2.89	13.7	0.9	6.6	13.4	13.1	14.9	16.6	18.0
G	**2482.75**	**2.30**	0.75	1.85	7.8	0.6	7.2	7.6	7.4	8.5	9.6	10.5
NA-HA												
A	2475.80	**0.75**	0.77									
A	2482.50	**0.75**	0.38									
G	2473.10	**2.39**	0.62	1.58	28.1	1.5	5.4	28.6	29.2	26.2	23.8	22.2

G	**2474.40**	**2.39**	1.34	3.41	42.8	0.3	0.7	42.9	42.9	42.4	41.4	40.4
G	**2476.40**	**2.39**	0.18	0.46	4.0	0.1	3.0	4.0	3.9	4.1	4.3	4.4
G	**2479.60**	**2.39**	0.33	0.83	4.8	0.3	5.9	4.7	4.6	5.2	5.7	6.0
G	**2481.30**	**2.39**	1.16	2.96	14.5	1.0	6.7	14.2	13.8	15.8	17.6	19.1
G	**2482.75**	**2.39**	0.53	1.34	5.8	0.4	7.3	5.7	5.5	6.4	7.2	7.9
PL-FA												
A	2475.85	**0.85**	0.93									
A	2482.67	**0.85**	0.32									
G	2473.14	**2.12**	1.15	2.60	46.9	2.2	4.6	47.7	48.5	44.2	40.6	38.0
G	2474.68	**2.12**	0.88	1.98	24.0	0.2	0.6	24.0	23.9	24.2	24.0	23.7
G	**2476.40**	**2.12**	0.13	0.28	2.5	0.1	3.6	2.5	2.4	2.6	2.7	2.8
G	**2479.60**	**2.12**	0.31	0.69	4.1	0.3	6.5	4.0	3.9	4.4	4.9	5.2
G	**2481.30**	**2.12**	1.52	3.43	17.3	1.3	7.3	16.9	16.4	18.9	21.3	23.2
G	**2482.75**	**2.12**	0.51	1.15	5.1	0.4	7.9	5.0	4.9	5.7	6.4	7.1
PL-FA_with_overabsorption_α = -0.1												
A	2475.84	**0.85**	0.93									
A	2482.66	**0.85**	0.29									
G	2473.06	**2.12**	1.09	2.46	48.1	2.2	4.5	48.9	49.7	45.3	41.7	39.1
G	2474.66	**2.12**	0.90	2.04	26.2	0.3	1.1	26.1	26.0	26.5	26.6	26.4
G	**2476.40**	**2.12**	0.13	0.28	2.7	0.1	4.1	2.6	2.6	2.8	3.0	3.1
G	**2479.60**	**2.12**	0.30	0.68	4.2	0.3	7.0	4.4	4.0	4.6	5.2	5.6
G	**2481.30**	**2.12**	1.18	2.66	14.1	1.1	7.8	13.8	13.4	15.6	17.7	19.4
G	**2482.75**	**2.12**	0.43	0.98	4.6	0.4	8.4	4.5	4.3	5.1	5.9	6.5
ES-HA												
A	2475.74	**0.42**	0.69									
A	2482.30	**0.42**	0.60									
G	2473.10	**1.75**	0.63	1.18	21.4	1.6	7.4	22.0	22.6	19.5	17.0	15.4
G	**2474.50**	**1.75**	1.16	2.16	27.1	0.7	2.4	27.3	27.5	26.2	24.7	23.5
G	**2476.40**	**1.75**	0.31	0.58	5.1	0.1	1.0	5.1	5.1	5.2	5.2	5.2
G	**2479.60**	**2.05**	0.33	0.72	4.2	0.2	3.7	4.2	4.1	4.4	4.7	4.8
G	**2481.30**	**2.05**	2.22	4.85	24.3	1.1	4.6	23.9	23.5	25.7	27.7	29.1
G	**2482.75**	**2.05**	1.83	4.01	17.8	0.9	5.2	17.5	17.2	19.0	20.7	22.0
ES-FA												
A	2475.51	**1.01**	0.56									
A	2483.51	**1.01**	0.73									
G	2473.15	**1.55**	0.53	0.87	16.9	1.4	8.4	17.4	18.0	15.1	13.0	11.6
G	**2474.50**	**1.55**	0.81	1.34	18.2	0.7	3.7	18.5	18.7	17.3	16.0	14.9
G	**2476.40**	**1.55**	0.18	0.30	2.9	0.0	0.4	2.9	2.9	2.9	2.8	2.7
G	**2479.60**	**1.94**	0.45	0.92	5.9	0.1	2.4	5.8	5.8	6.1	6.3	6.4
G	**2481.30**	**1.94**	2.52	5.20	28.3	0.9	3.2	28.0	27.6	29.4	30.9	32.0
G	**2482.75**	**1.94**	2.80	5.77	27.8	1.1	3.8	27.4	27.1	29.2	31.0	32.4

[a] $\Delta/2$ = |%(Huffman-Sarret)/2| = ΔS_i. [b] Precision = [($\Delta/2$)/(% Generic)]*100

Table EA-5. Atomic fractions (%) of sulfur functionalities as determined by the GCF and LCF methods and average values. The difference of % S_i between the two methods (Δ) is in most cases ≤ 4%, and in a few cases (all reduced sulfur, S_{red}, species) ranges between 4 and 8% (values highlighted in yellow). Values from the literature (denoted in blue text) which differ with the new values by more than 4% and 8% are highlighted in magenta.

	Exocyclic	Heterocyclic	Sulfoxide	Sulfone	Sulfonate	Sulfate	Sum	Sred[a]
SR-HA								
GCF - % Generic	23.3	41.1	3.6	5.4	19.3	7.3	100.0	64.4
GCF - Δ/2	1.5	0.5	0.1	0.3	1.2	0.5		5.0
LCF – average[b]	23.9	39.6	3.4	7.4	19.0	6.7	100.0	63.5
LCF - Δ/2	1.1	1.0	0.0	0.2	0.2	0.2		0.1
Average [GCF+LCF]	23.6	40.4	3.5	6.4	19.1	7.0	100.0	63.9
Δ/2	0.3	0.7	0.1	1.0	0.1	0.3		0.4
Precision%	1.3	1.8	2.8	15.6	0.7	4.8		0.6
Xia et al. (1998)	46.0	9.2	10.2	4.8	15.9	13.9	100.0	55.2
SR-FA								
GCF - % Generic	23.1	31.3	2.8	7.3	25.4	10.1	100.0	54.4
GCF - Δ/2	1.5	0.7	0.0	0.3	1.3	0.6		5.3
LCF - average	28.6	24.0	2.8	10.8	24.6	9.2	100.0	52.6
LCF - Δ/2	0.3	0.3	0.0	0.1	0.1	0.1		0.0
Average [GCF+LCF]	25.8	27.6	2.8	9.1	25.0	9.7	100.0	53.5
Δ/2	2.7	3.6	0.0	1.8	0.4	0.4		0.9
Precision%	10.6	13.2	1.0	19.5	1.7	4.4		1.7
Xia et al. (1998)	34.7	12.6	7.3	6.5	24.0	14.9	100.0	47.3
WP-HA								
GCF - % Generic	20.8	35.5	4.1	3.8	13.6	22.1	100.0	56.4
GCF - Δ/2	1.3	0.8	0.1	0.2	0.7	1.2		5.2
LCF - average	26.2	32.5	4.2	3.5	15.5	18.1	100.0	58.7
LCF - Δ/2	1.7	1.8	0.0	0.1	0.0	0.1		0.1
Average [GCF+LCF]	23.5	34.0	4.1	3.6	14.6	20.1	100.0	57.5
Δ/2	2.7	1.5	0.0	0.2	1.0	2.0		1.2
Precision%	11.5	4.5	0.8	4.2	6.5	10.0		2.0
WP-FA								
GCF - % Generic	27.1	42.1	4.6	4.7	13.7	7.8	100.0	69.2
GCF - Δ/2	1.6	0.3	0.1	0.3	0.9	0.6		4.6
LCF - average	33.9	37.4	3.6	5.2	13.1	6.9	100.0	71.2
LCF - Δ/2	1.4	1.3	0.0	0.2	0.2	0.2		0.1
Average [GCF+LCF]	30.5	39.7	4.1	5.0	13.4	7.3	100.0	70.2
Δ/2	3.4	2.4	0.5	0.2	0.3	0.4		1.0
Precision%	11.1	6.0	12.8	4.9	2.3	5.9		1.4
NA-HA								
GCF - % Generic	28.1	42.8	4.0	4.8	14.5	5.8	100.0	70.9
GCF - Δ/2	1.5	0.3	0.1	0.3	1.0	0.4		4.5

LCF - average	36.4	36.4	3.5	6.0	12.8	4.9	100.0	72.8
LCF - Δ/2	0.4	0.1	0.1	0.4	0.1	0.1		0.3
Average [GCF+LCF]	32.2	39.6	3.8	5.4	13.6	5.4	100.0	71.8
Δ/2	4.2	3.2	0.2	0.6	0.9	0.4		1.0
Precision%	12.9	8.1	6.6	11.2	6.4	8.3		1.3
Yoon et al. (2005)	34.7	30.2	5.6	3.2	14.2	12.2	100.1	64.9
PL-FA								
GCF - % Generic	46.9	24.0	2.5	4.1	17.3	5.1	100.0	71.0
GCF - Δ/2	2.2	0.1	0.1	0.3	1.3	0.4		5.0
LCF - average	47.0	18.6	5.2	7.0	18.4	3.8	100.0	65.6
LCF - Δ/2	0.3	0.3	0.0	0.1	0.1	0.0		0.0
Average [GCF+LCF]	46.9	21.3	3.9	5.5	17.8	4.5	100.0	68.3
Δ/2	0.0	2.7	1.4	1.4	0.5	0.7		2.7
Precision%	0.0	12.7	35.2	26.1	3.0	14.7		3.9
ES-HA								
GCF - % Generic	21.4	27.1	5.1	4.2	24.3	17.8	100.0	48.5
GCF - Δ/2	1.6	0.7	0.0	0.2	1.1	0.9		5.2
LCF - average	15.3	29.4	6.4	4.7	25.9	18.3	100.0	44.7
LCF - Δ/2	0.7	0.6	0.0	0.2	0.2	0.2		0.1
Average [GCF+LCF]	18.4	28.2	5.8	4.4	25.1	18.1	100.0	46.6
Δ/2	3.1	1.2	0.6	0.2	0.8	0.3		1.9
Precision%	16.8	4.1	10.7	5.2	3.2	1.5		4.1
ES-FA								
GCF - % Generic	16.9	18.2	2.9	5.9	28.3	27.8	100.0	35.1
GCF - Δ/2	1.4	0.7	0.0	0.1	0.9	1.1		4.7
LCF - average	13.8	20.7	3.6	6.8	30.9	24.3	100.0	34.4
LCF - Δ/2	0.4	0.3	0.0	0.2	0.2	0.1		0.1
Average [GCF+LCF]	15.3	19.5	3.2	6.3	29.6	26.1	100.0	34.8
Δ/2	1.5	1.2	0.4	0.4	1.3	1.8		0.3
Precision%	10.1	6.2	10.9	7.0	4.4	6.8		1.0

[a] $S_{red} = S_{Hetero} + S_{Exo}$

[b] Average of fractional amounts derived from the spectral fits in the 2466-2520 eV (long) and 2466-2489 eV (short) intervals (Table EA-8).

Table EA-6. Parameter values for the GCF model fitted to NA-HA with the Yoon model (Yoon et al., 2005), to SR-HA and SR-FA fitted with the Xia model (Xia et al., 1998), to the protein-rich fraction of a humic substance fitted with the Generic model, and to the Alais meteorite (chondrite) fitted with the Orthous-Daunay model (Orthous-Daunay et al., 2010). Fits are shown in sections 6.1 and 6.2.

SR-HA_Xia Model (Model-fit 1)

	Energy	Amp	Width	Area	Waldo (%)	Xia (%)	Δ (%)
A	2475.70	**0.92**	**1.53**				
A	2482.61	0.30	**1.53**				
G	2473.75	1.29	2.38	3.28	44.9	46.0	-1.1
G	2474.56	0.44	1.12	0.53	6.5	9.2	-2.7
G	2475.85	0.35	1.74	0.65	7.0	10.2	-3.2
G	**2479.60**	0.20	2.31	0.50	3.9	4.8	-0.9
G	**2481.30**	1.53	2.44	3.99	27.4	15.9	11.5
G	**2482.75**	0.64	2.43	1.65	10.3	13.9	-3.6

SR-HA_Modified Xia Model (Model-fit 2)

	Energy	Amp	Width	Area	Waldo (%)	% Xia	Δ (%)
A	2475.70	**0.92**	**0.73**				
A	2482.56	0.29	**0.73**				
G	2473.98	1.42	2.76	4.17	54.5	46.0	8.5
G	2474.62	0.28	1.30	0.38	4.6	9.2	-4.6
G	2475.85	0.05	1.68	0.08	0.8	10.2	-9.4
G	**2479.60**	0.19	2.01	0.41	3.1	4.8	-1.7
G	**2481.30**	1.59	2.40	4.06	27.4	15.9	11.5
G	**2482.75**	0.62	2.36	1.55	9.6	13.9	-4.3

SR-FA_Xia Model

	Energy	Amp	Width	Area	Waldo (%)	% Xia	Δ (%)
A	2475.70	**0.80**	**1.11**				
A	2482.51	0.42	**1.11**				
G	2473.51	1.00	2.08	2.24	31.4	34.7	-3.3
G	2474.68	0.67	1.42	1.01	12.2	12.6	-0.4
G	2475.85	0.27	1.37	0.40	4.3	7.3	-3.0
G	**2479.60**	0.38	2.04	0.83	6.4	6.5	-0.1
G	**2481.30**	2.04	2.18	4.74	32.3	24.0	8.3
G	**2482.75**	0.92	2.19	2.15	13.4	14.9	-1.5

NA-HA_Yoon Model

	Energy	Amp	Width	Area	Waldo (%)[a]	Yoon (%)	Δ (%)
A	2475.67	**0.84**	**0.99**				
A	2481.54	0.20	**0.99**				
G	2473.27	0.98	2.21	2.21	37.9	34.7	3.2
G	2474.56	1.01	1.67	1.81	26.3	30.2	-3.9
G	2475.88	0.24	1.67	0.43	5.4	5.6	-0.2

G	2479.47	0.22	1.53	0.35	3.2	3.2	0.0
G	2480.90	0.83	1.76	1.55	12.8	14.2	-1.4
G	2482.48	0.77	2.36	1.92	14.4	12.2	2.2

Protein-rich fraction of a humic substance

	Energy	Amp	Width	Area	Generic (%)
A	2475.63	1.03	**1.66**		
G	2482.50	0.13	**1.66**		
G	**2473.05**	1.60	**1.78**	3.03	52.7
G	**2474.34**	1.31	**1.78**	2.49	30.5
G	**2476.40**	0.16	**1.78**	0.31	2.5
G	**2478.55**	0.12	**1.78**	0.22	1.4
G	**2479.60**	0.23	**1.78**	0.43	2.4
G	**2481.30**	0.74	**1.78**	1.40	6.6
G	**2482.75**	0.49	**1.78**	0.93	3.9

Alais Chondrite_Orthous-Daunay Model (Model-fit 1)

	Energy	Amp	Width	Area	Orthous-Daunay (%)[b]	Generic (%)	Δ (%)
A	2476.96	1.01	**1.20**				
G	2470.47	0.54	**1.77**	1.01	18.4	57.0	38.7
G	2472.10	0.37	**1.77**	0.69	10.7	9.1	-1.6
G	2473.00	0.77	**1.77**	1.45	20.7	13.4	-7.3
G	2474.44	0.84	**1.77**	1.58	20.1	9.9	-10.2
G	2476.20	0.77	**1.77**	1.44	16.1	6.4	-9.7
G	2478.36	0.15	**1.77**	0.28	2.8	0.9	-1.8
G	2480.37	0.20	**1.77**	0.39	3.4	1.0	-2.3
G	2481.53	0.31	**1.77**	0.58	4.8	1.4	-3.4
G	2482.97	0.21	**1.77**	0.39	3.0	0.8	-2.2

Alais Chondrite_Modified Orthous-Daunay Model (Model-fit 2)

	Energy	Amp	Width	Area	Orthous-Daunay (%)	Generic (%)	Δ (%)
A	2475.39	1.01	**1.20**				
G	2470.47	0.53	**1.75**	0.99	21.7	60.8	39.0
G	2472.10	0.37	**1.75**	0.69	12.8	9.8	-3.0
G	2473.00	0.75	**1.75**	1.41	24.2	14.1	-10.1
G	2474.44	0.80	**1.75**	1.50	22.9	10.1	-12.8
G	2476.26	0.12	**1.75**	0.22	2.9	1.0	-1.9
G	2478.41	0.11	**1.75**	0.20	2.3	0.7	-1.6
G	2480.41	0.20	**1.75**	0.37	3.9	1.1	-2.8
G	2481.55	0.30	**1.75**	0.57	5.6	1.5	-4.1
G	2482.97	0.21	**1.75**	0.39	3.6	0.9	-2.7

[a] Correction of the absorption cross-section using the Waldo calibration curve (Fig. 2).
[b] Correction of the absorption cross-section using the Orthous-Daunay calibration curve (Fig. 2)

Table EA-7. Atomic fractions (%) of sulfur references (Table EA-1) derived from the Combo fit of K-edge XANES spectra in the 2466-2520 eV (long) and 2466-2489 eV (short) interval.

	SR-HA long	SR-HA short	SR-FA long	SR-FA short	WP-HA long	WP-HA short	WP-FA long	WP-FA short	NA-HA long	NA-HA short	PL-FA long	PL-FA short	ES-HA long	ES-HA short	ES-FA long	ES-FA short
Sulfate Na																
Sulfate Mg	0.043	0.033			0.049	0.01	0.05	0.027	0.039	0.03	0.024	0.019	0.138	0.104	0.117	0.081
Gypsum	0.027	0.034	0.095	0.094	0.133	0.176	0.022	0.043	0.01	0.018	0.015	0.02	0.06	0.094	0.137	0.174
Sulfonate	0.196	0.195	0.251	0.252	0.157	0.159	0.135	0.134	0.126	0.125	0.188	0.187	0.28	0.28	0.324	0.324
Sulfone	0.073	0.079	0.109	0.113	0.036	0.035	0.051	0.056	0.055	0.063	0.07	0.072	0.048	0.053	0.069	0.073
Sulfite																
Sulfoxide	0.035	0.036	0.028	0.029	0.042	0.043	0.036	0.037	0.033	0.036	0.053	0.054	0.068	0.07	0.037	0.038
Sulfonium-Met, Sulfonium-TriMet, PhSPh, PhSn, Clothepin[a]																
Thiazine					0.056	0.045	0.021									
Thianthrene																
Thiepin	0.269	0.241	0.139	0.137	0.065	0.096	0.222	0.211	0.171	0.174			0.194	0.172	0.103	0.103
Dothiepin	0.131	0.122	0.05	0.038	0.226	0.145	0.138	0.121	0.137	0.131	0.188	0.177	0.123	0.119	0.108	0.09
Thiophene	0.016	0.037	0.059	0.068		0.027	0.013	0.041	0.047	0.056	0.004	0.011	0.004	0.023	0.007	0.022
PhCH2SPh																
Methionine																
Ph-SH-1						0.045										
Ph-SH-2	0.123	0.151	0.253	0.255	0.139	0.141	0.191	0.22	0.212	0.192	0.15	0.158	0.029	0.047	0.085	0.095
Cysteine																
Sulfide											0.093	0.09				
Cystine	0.065	0.05			0.071	0.053	0.095	0.077	0.081	0.043	0.121	0.114	0.127	0.12	0.054	0.044
Disulfide																
Elemental S										0.076						
Pyrite	0.045	0.058	0.035	0.042	0.026	0.033	0.045	0.068	0.067	0.045	0.094	0.105		0.007	0.001	0.01
Pentlandite																
Pyrrhotite					0.012	0.014					0.017	0.016				
Sum	1.02	1.04	1.02	1.03	1.01	1.02	1.02	1.04	0.98	0.99	1.02	1.02	1.07	1.09	1.04	1.05
***NSS* x 10^3**	1.01	0.80	1.01	0.85	1.05	0.89	0.90	0.47	0.80	0.76	0.50	0.49	1.25	0.85	1.15	0.95
ΔE (eV)	-0.01	0.02	0.04	0.06	0.01	0.04	0.02	0.07	0.05	0.10	0.04	0.06	-0.02	0.00	-0.02	0.00

[a] None of these components were fit; they all had negative loadings.

Table EA-8. Atomic fractions (%) of sulfur functionalities derived from the Combo fit of K-edge XANES spectra in the 2466-2520 eV (long) and 2466-2489 eV (short) interval. The fraction of each functionality was taken as the sum of all positive fractions of references with similar oxidation states. The groupings of reference compounds are listed in the footnote.

	Exocyclic	Heterocyclic	Sulfoxide	Sulfone	Sulfonate	Sulfate	Sum (%)
SR-HA long	22.8	40.7	3.4	7.1	19.2	6.8	100.00
SR-HA short	25.0	38.6	3.5	7.6	18.8	6.5	100.00
SR-FA long	28.3	24.3	2.7	10.7	24.6	9.3	100.00
SR-FA short	28.9	23.6	2.8	11.0	24.5	9.1	100.00
WP-HA long	24.5	34.3	4.1	3.6	15.5	18.0	100.00
WP-HA short	28.0	30.6	4.2	3.4	15.6	18.2	100.00
WP-FA long	32.5	38.7	3.5	5.0	13.2	7.1	100.00
WP-FA short	35.3	36.0	3.6	5.4	12.9	6.8	100.00
NA-HA long	36.8	36.3	3.4	5.6	12.9	5.0	100.00
NA-HA short	36.0	36.5	3.6	6.4	12.6	4.8	100.00
PL-FA long	46.7	18.9	5.2	6.9	18.5	3.8	100.00
PL-FA short	47.2	18.4	5.3	7.0	18.3	3.8	100.00
ES-HA long	14.6	30.0	6.3	4.5	26.1	18.5	100.00
ES-HA short	16.0	28.8	6.4	4.9	25.7	18.2	100.00
ES-FA long	13.4	20.9	3.5	6.6	31.1	24.4	100.00
ES-FA short	14.1	20.4	3.6	6.9	30.7	24.2	100.00

Sulfate = sulfate-Na + sulfate-Mg + gypsum
Heterocyclic S = thiazine + thianthrene + thiepin + dothiepin + thiophene
Exocyclic = methionine + Ph-SH-1 + Ph-SH-2 + cysteine + sulfide + cystine + disulfide + S0 + pyrite + pyrrhotite

2. Gaussian decomposition of some references

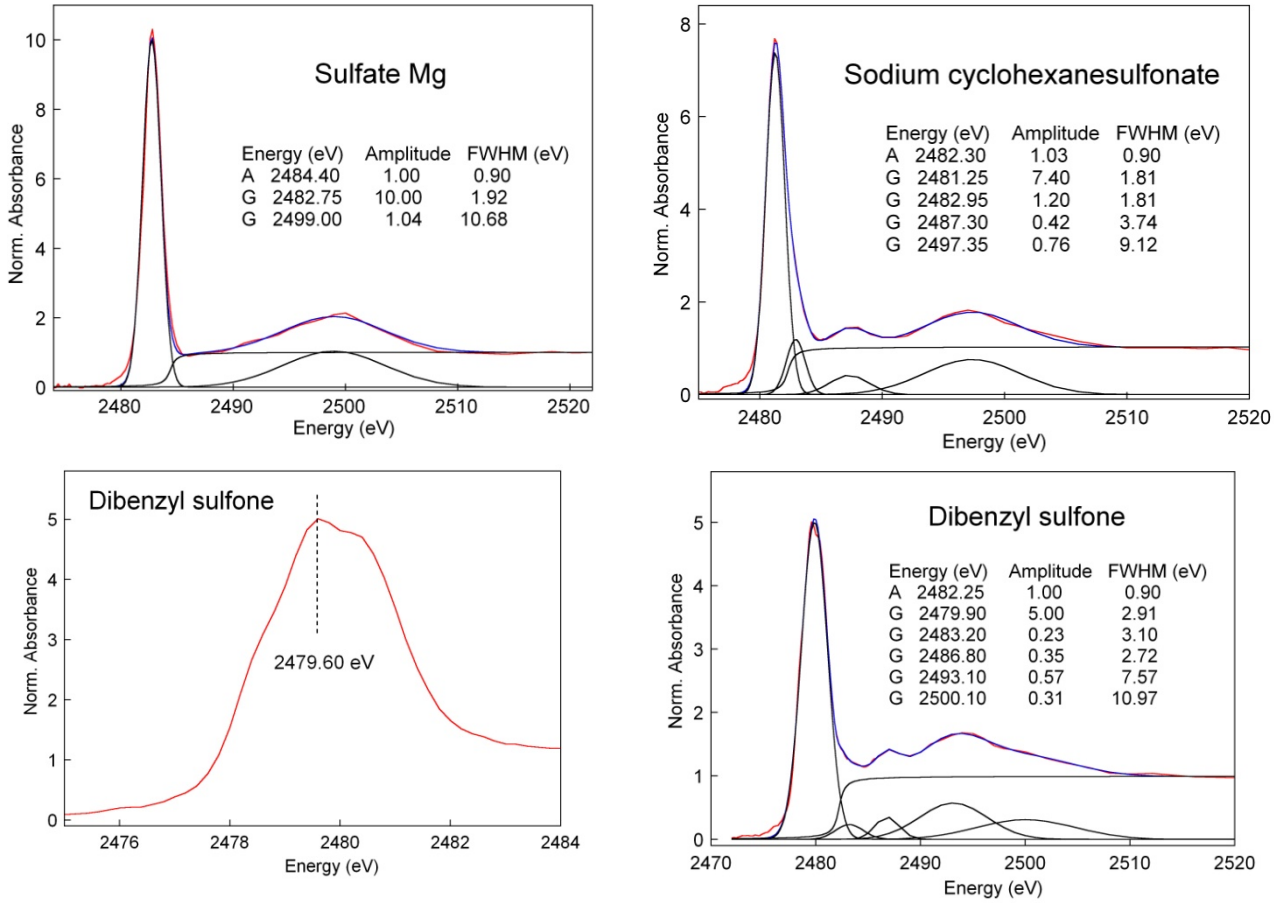
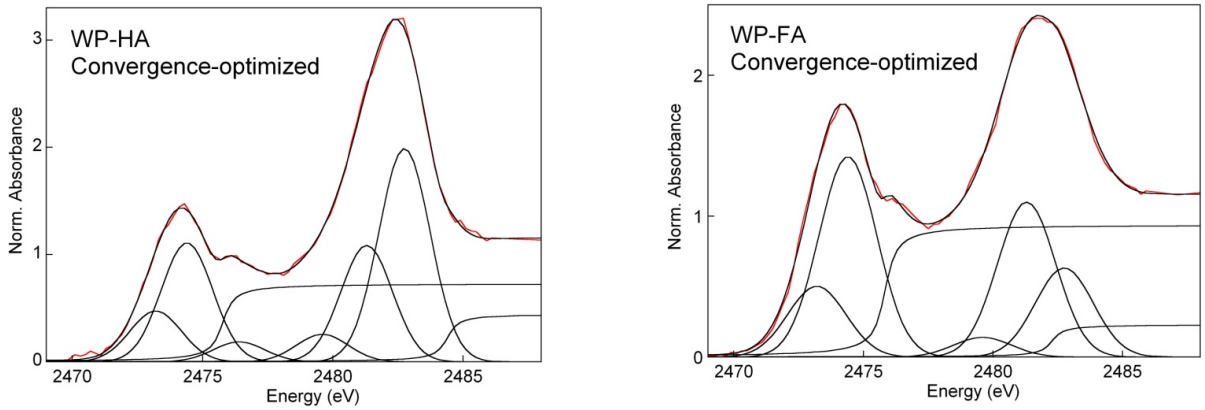


Fig. EA-1. Oxidized sulfur spectra fit to a sum of Gaussian and arctangent functions. The main peak of sulfonate and sulfone is broadened as a result of the asymmetry of the S 3p molecular orbitals (σ* type) hybridized with O and C orbitals. The peak of sulfone can be decomposed into either two to three Gaussians or one single Gaussian with a larger width of 2.9 eV compared to 1.9 eV in sulfate.

3. Gaussian decomposition of the IHSS references with the convergence-optimized method



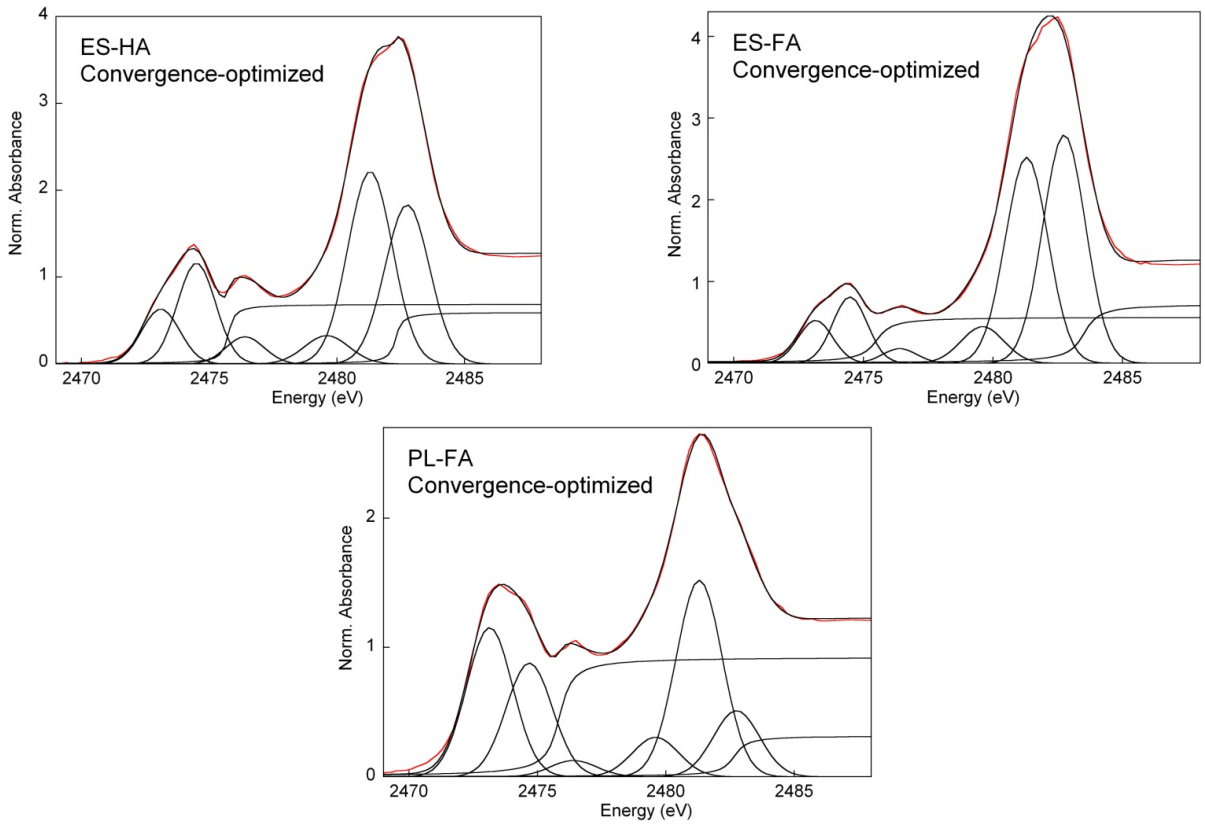


Fig. EA-2. Spectral decomposition of IHSS samples with Gaussians and two arctangents (GCF method). Parameter values and S_i fractions are given in Table EA-4. The fits of NA-HA, SR-HA and SR-FA are shown in section 6.1.

4. Calibration curves

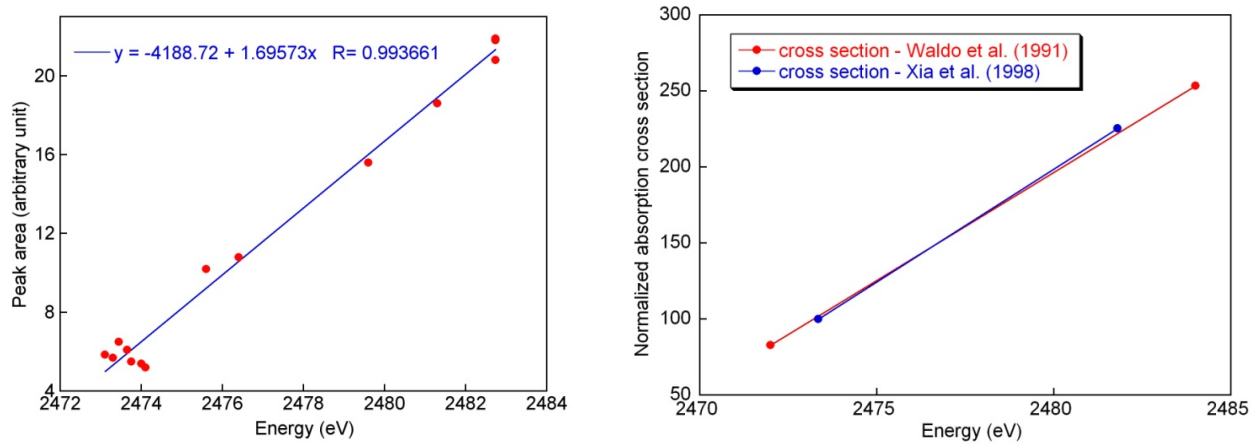
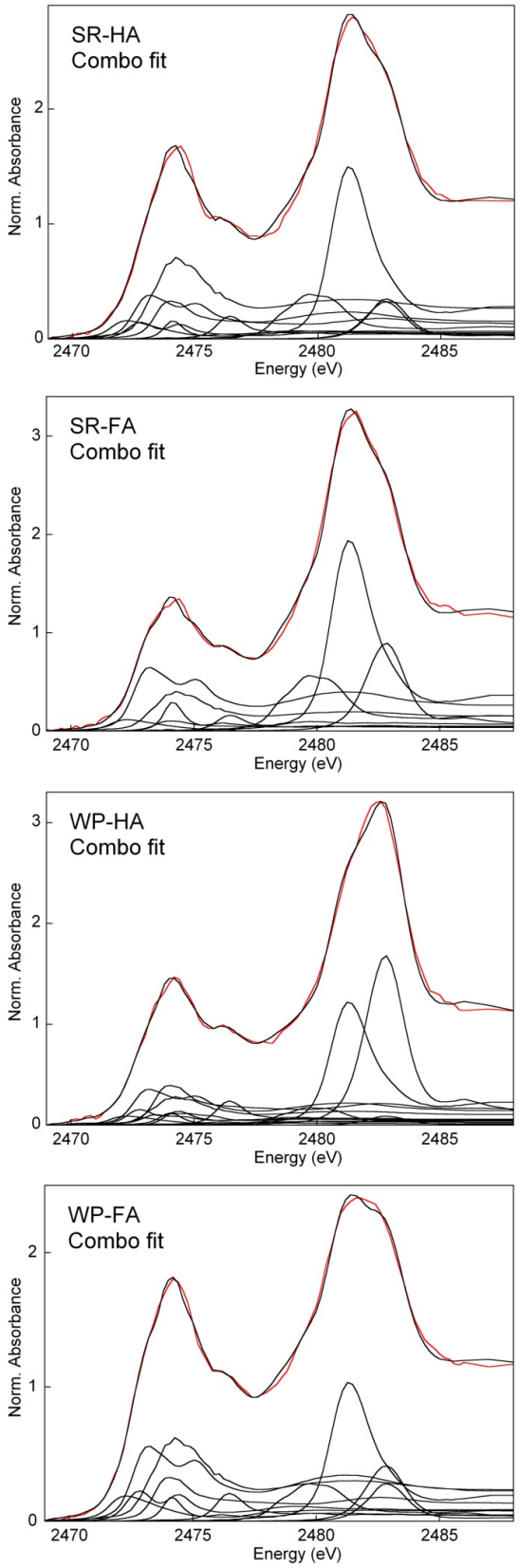


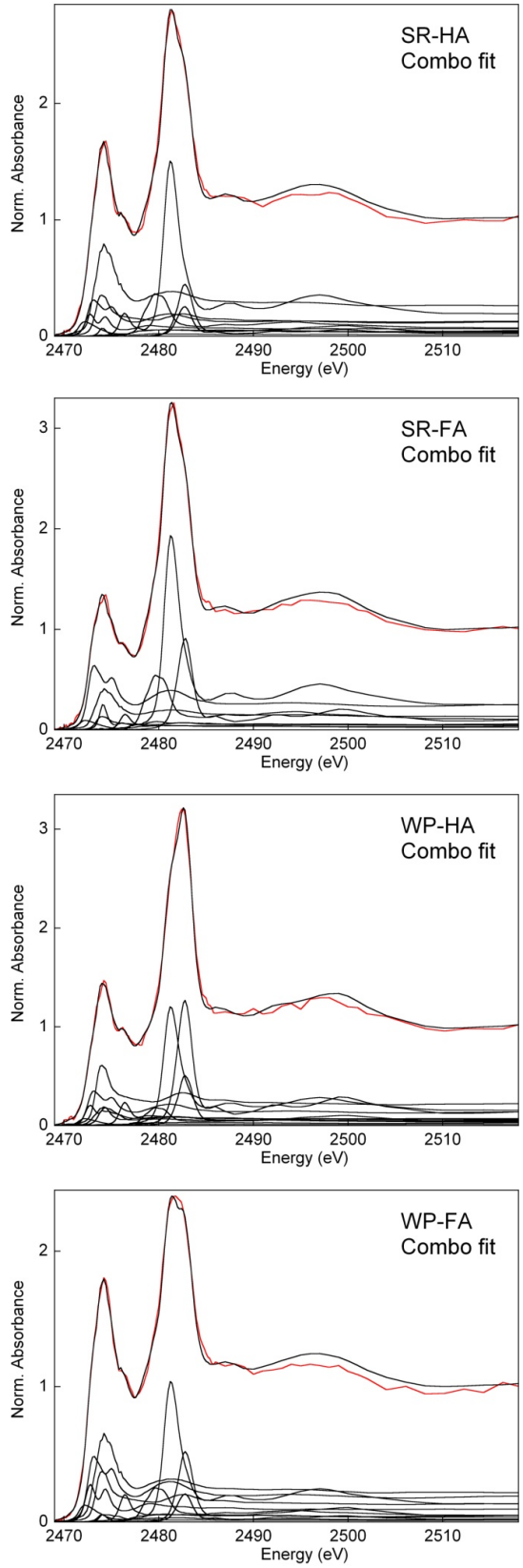
Fig. EA-3. Left: Calibration curve determined in this study from the white line peak area of model compounds using the same method as Sarret et al. (1999). The references used are: sulfate-Na, sulfate-Mg, gypsum, sulfonate, sulfone, sulfoxide, sulfonium-TriMet, dothiepin, thiophene, PhCH2SPh, methionine, cysteine, sulfide, disulfide. Right: Digitization of calibration curves from Figure 4 of Waldo et al. (1991) and Figure 2 of Xia et al. (1998) verifies that the two curves are the same (within the precision of the data reproduction).

5. Combo fit of the IHSS references

Fit in the 2466-2489 eV interval



Fit in the 2466-2520 eV interval



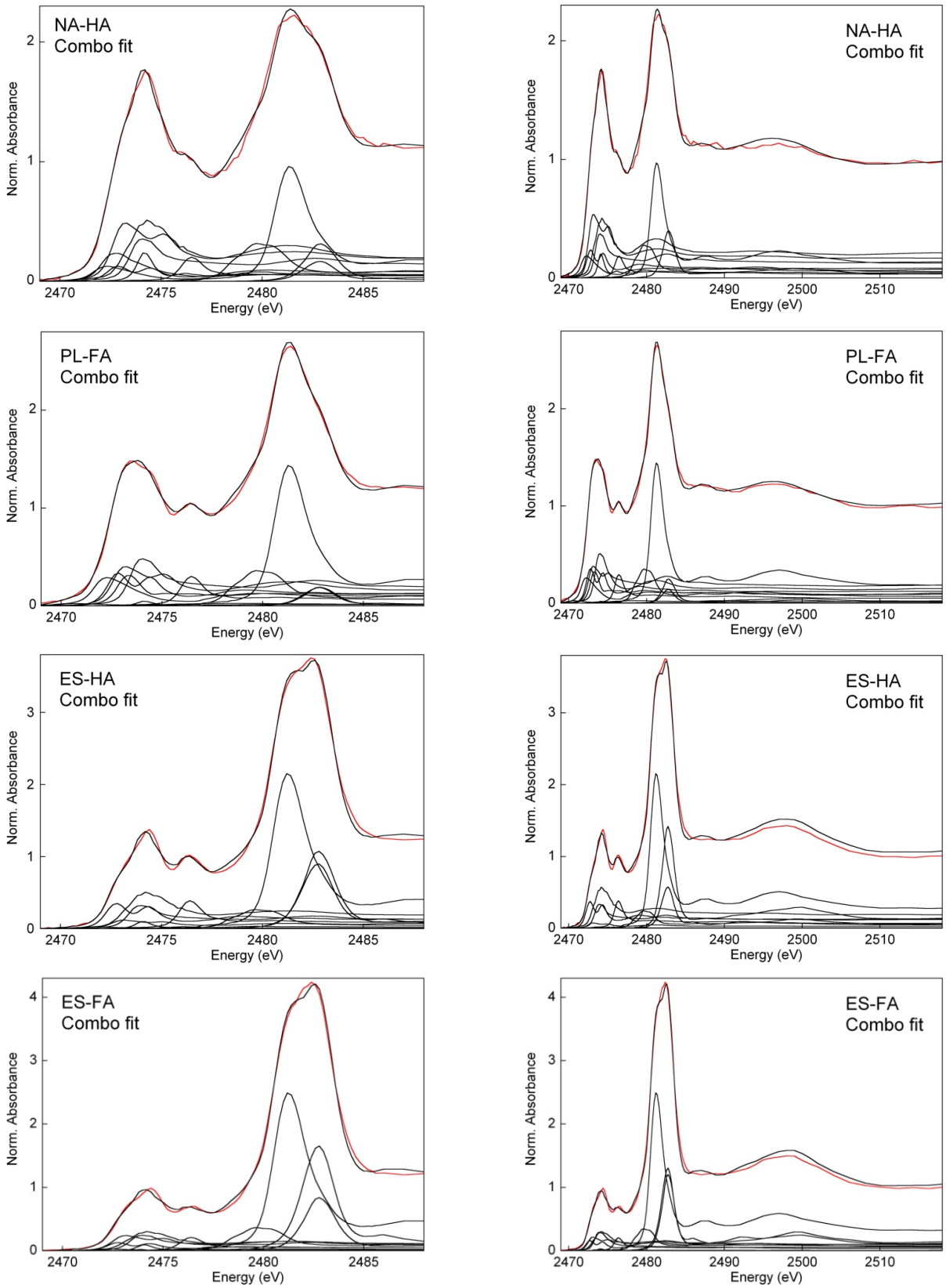


Fig. EA-4. Spectral decomposition of IHSS samples with a linear combination of reference spectra (LCF method). Parameter values and S_i fractions are given in Table EA-6 and synthesized in Tables EA-5 and EA-8.

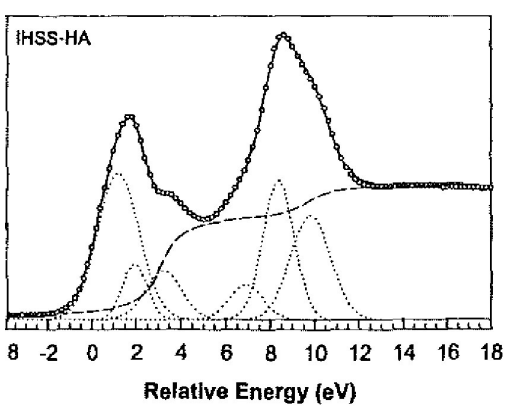
6. Evaluation and discussion of previous S-XANES data analyses

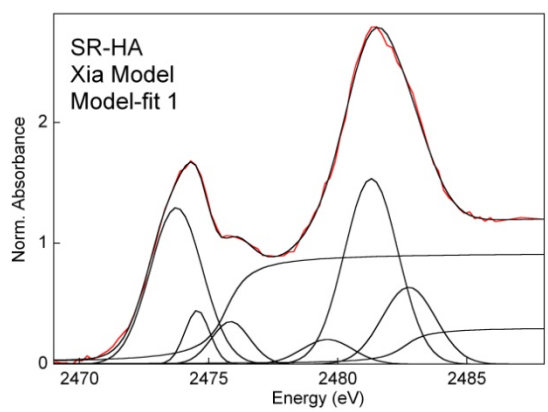
6.1. IHSS humic substances

Three IHSS samples were studied previously by S-XANES: SR-HA and SR-FA (Xia et al., 1998), and NA-HA (Yoon et al., 2005). The spectra were modeled with a $2A+\sum_{i=1}^{6} G_i^{Ei,wi}$ model, in which both the energies and widths of the Gaussians were varied, and their areas normalized with scaling factors from the Waldo calibration curve. To evaluate model fit uncertainty and bias on sulfur fractions (S_i) in these three materials, our measured spectrum for each sample was analyzed in three ways: (1) with the same strategy as in the previous studies, introducing when needed some constraints to reproduce as closely as possible the original fits; (2) with the convergence-optimized GCF procedure, and, (3), with the LCF "Combo fit" method. Results are displayed in Figures EA-5 to EA-7, and the corresponding fit parameters in Table EA-6.

SR-HA. The Xia model applied to our data gives results in good accordance with their values (ΔS_i < 3.6%), except for S_{SO3} which is 11.5% higher in our case. Fig. EA-5 shows two mathematical solutions of the Gaussian curve fit, among many others possible, in which the width of the first arctangent was fixed successively to 1.53 eV and 1.14 eV. The S_{Exo}, S_{Hetero} and S_{sulfx} fractions are particularly sensitive to the parametric values of the step-function, with a variation in S_{red} of about 10% between the two fits (~50 vs. 60%). The convergence-optimized Gaussian curve fit provides, in contrast, a robust structural solution with all S_i values in close agreement with the Combo fit values ($\Delta S_i \leq 2\%$, Table EA-4). The width of the G_{SExo} component using the Xia model is unrealistically large (~2.4 eV), and the width of the $G_{SHetero}$ unrealistically narrow (~1.1 eV).

With respect to the results, the amount of S_{Hetero} is 40% - 9% = 31% of S_T higher in our study, and the amount of S_{Exo} is 46% - 24% = 22% lower. These differences are attributed to the model-fit. In contrast, the S_{red} fraction is only 9% lower in their study, a difference attributed to the calibration curve (Waldo vs. Generic curve), and not to the model-fit because the total area of the S_{red} species is independent of the widths and relative heights of the G_{Hetero} and G_{Exo} functions.





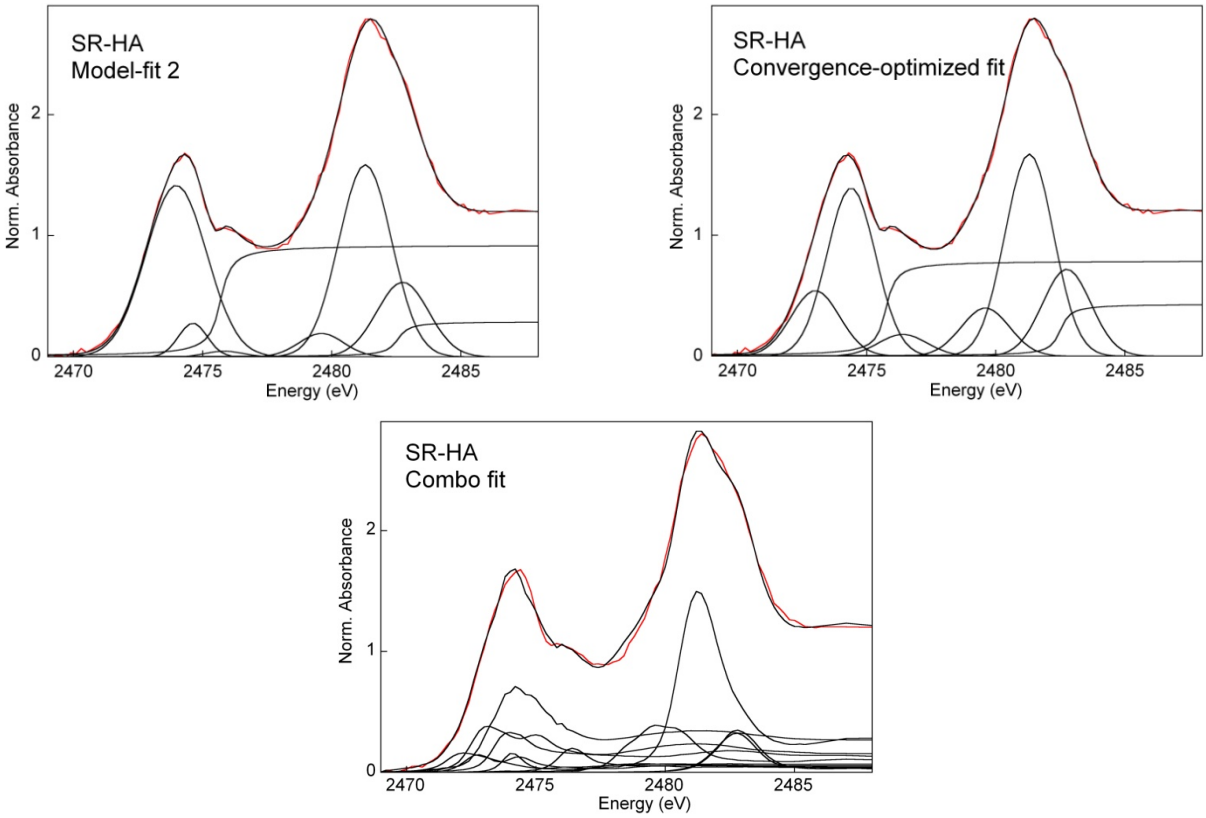
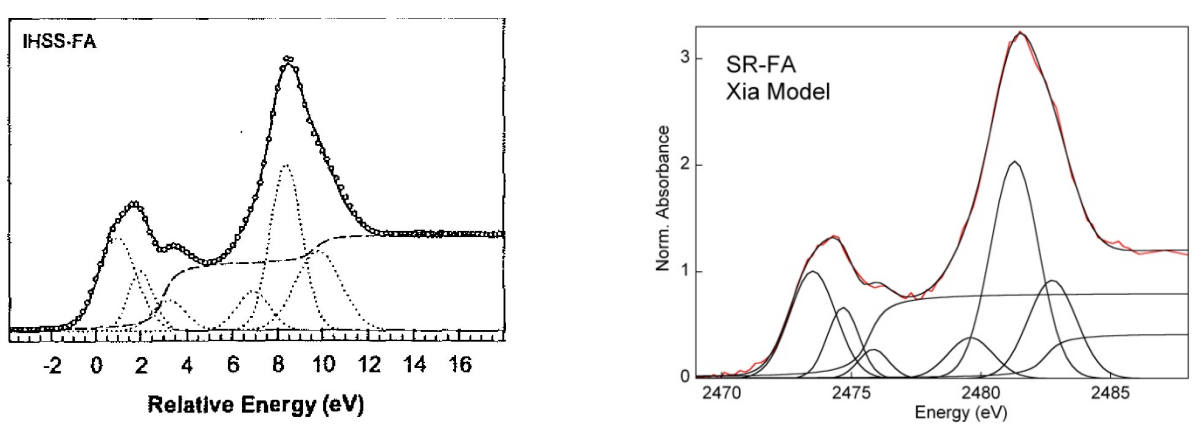


Fig. EA-5. Sulfur K-edge XANES spectra of Suwannee River humic acid (SR-HA, IHSS 2S101H). Upper left: Data and model-fit of Xia et al. (1998). It is unknown if Xia et al. (1998) measured 1S101H or 2S101H. Upper right: Data of this study and same model-fit as in Xia et al. (1998). Middle left: Another solution of the Xia model-fit, in which the second Gaussian is nested within the first one, with almost no sulfoxide. Varying the position and width of the Gaussians introduces many parameter correlations, so that some can be increased and others decreased with no significant change in quality of a model-fit to data. Middle right: Convergence-optimized fit to data of this study. Lower: Combo fit to data of this study.

SR-FA. The Xia model applied to our data gives results in good accordance with their values (ΔS_i < 3.3.%), except for S_{SO3} (ΔS = 8.3%). The amount of S_{Hetero} is 28% - 13% = 15% higher, and S_{Exo} 35% - 26% = 9% lower in our study. Similarly to SR-HA, S_{red} was underestimated by about 53% - 47% = 6% in the study of Xia et al. (1998), as a result of using the Waldo calibration curve, which has a smaller slope.



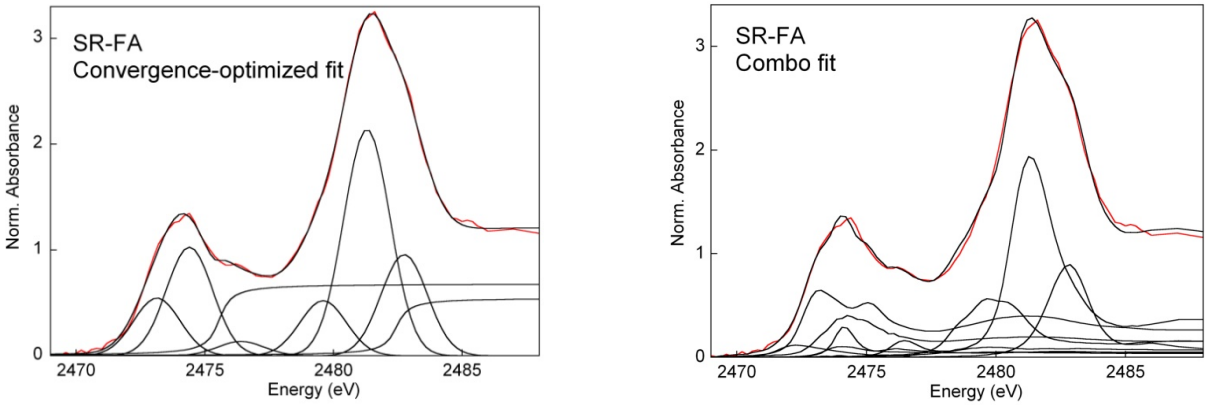


Fig. EA-6. Sulfur K-edge XANES spectra of Suwannee River fulvic acid (SR-FA, IHSS 2S101F). Upper left: Data and model-fit of Xia et al. (1998). It is unknown if Xia et al. (1998) measured 1S101F or 2S101F. Upper right: Data of this study and same model-fit as in Xia et al. (1998). Lower left: Convergence-optimized Gaussian curve fit to data of this study. Lower right: Combo fit to data of this study.

NA-HA. The S_i fractions obtained on our own data with the Yoon model differ at most by 3.9% (i.e., for S_{Hetero}) with the values reported by Yoon et al. (2005), which engenders confidence in the new calculations. The differences discussed below are analytical and not experimental. The convergence-optimized and Combo fits yield S_{SO4} values ~7% lower, $S_{red} = S_{Hetero+SExo}$ ~6% higher than the Yoon model. Overall, the Yoon mathematical solution seems reasonable; the main difference between their and our results comes from application of the Waldo calibration curve in their study, which overestimates S_{SO4} and underestimates S_{red}.

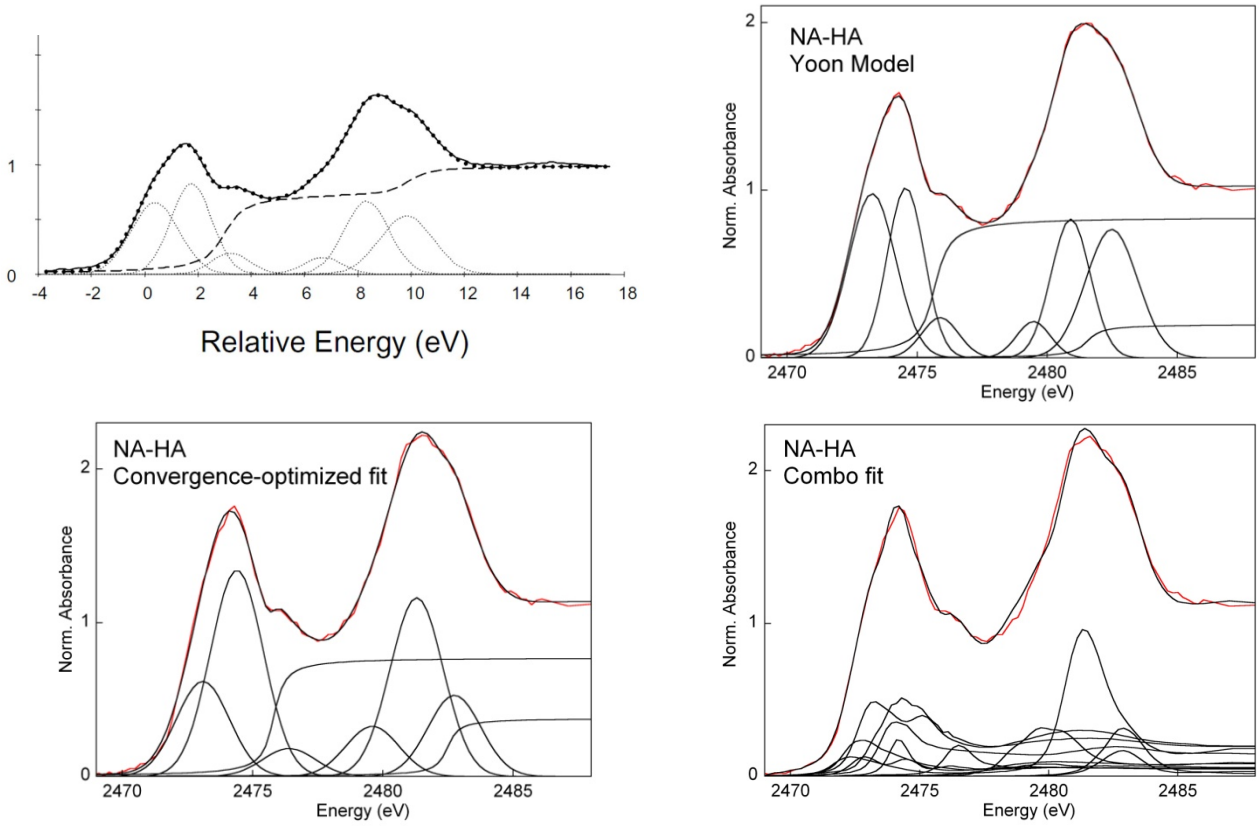


Fig. EA-7. Sulfur K-edge XANES spectra of Nordic aquatic humic acid (NA-HA, IHSS 1R105H). Upper left: Data and model-fit of Yoon et al. (2005). Upper right: Data of this study and same model-fit as in Yoon et al. (2005). To facilitate comparison, the spectrum was normalized to mu=1 at 2489 eV. Lower left: Convergence-optimized Gaussian curve fit to data of this study. Lower right: Combo fit to data of this study. Data normalized to mu=1 at 2510-2520 eV. This

parameter does not change the final S_i values, because their sum is re-normalized to 1 after correction for the absorption cross-sections.

6.2. Alais meteorite (carbonaceous chondrite)

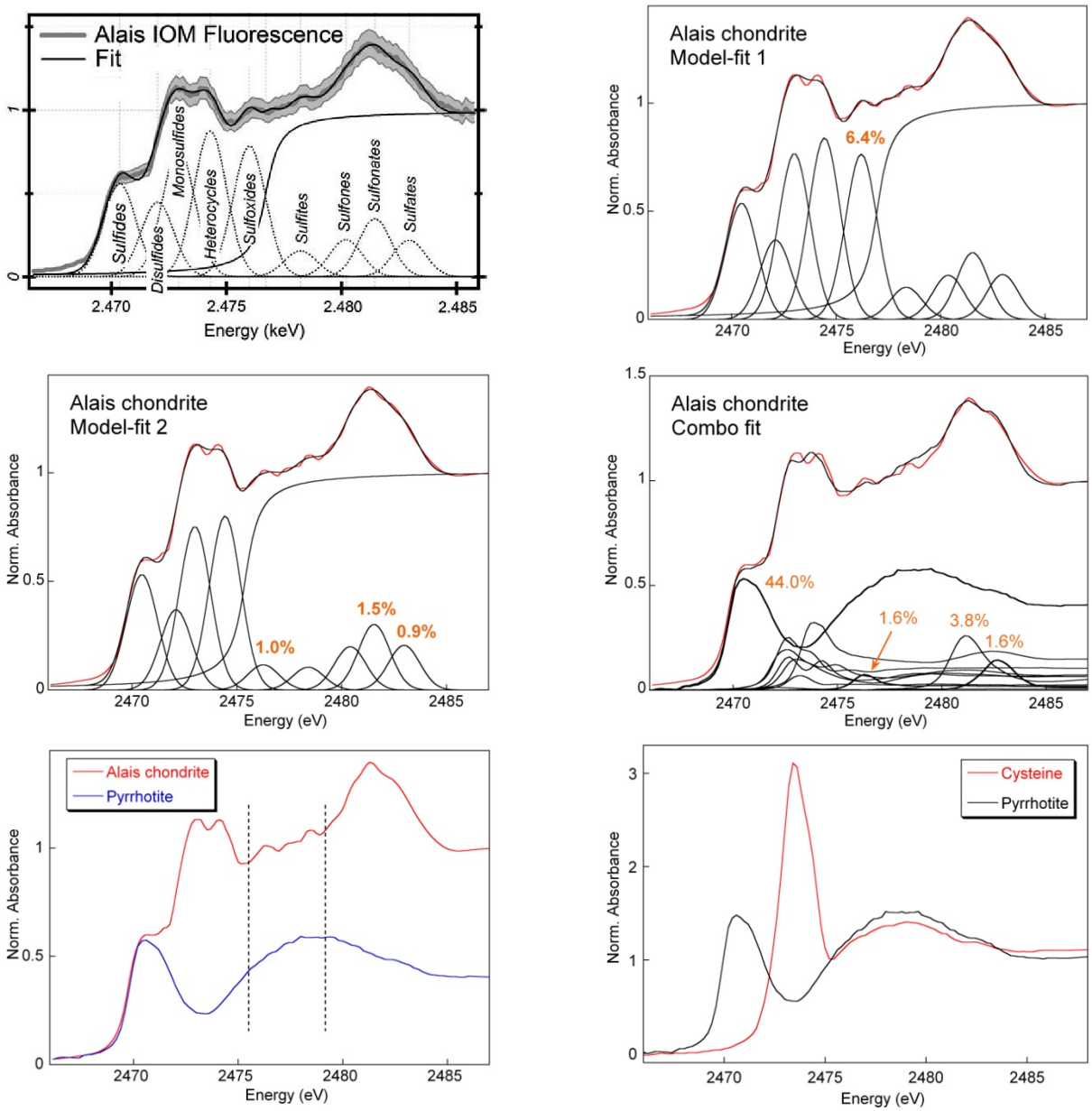


Fig. EA-8. Sulfur K-edge XANES spectrum of the Alais chondrite measured by Orthous-Daunay et al. (2010). Upper left: Model-fit of Orthous-Daunay et al. (2010). Upper right: Our version of their model-fit to a digitized version of their spectrum. Middle left: Another solution of the Othous-Daunay model-fit with the arctangent function shifted to the left of the sulfoxide peak. This second solution gives a sulfoxide fraction in line with the result from the Combo fit. Middle right: Combo fit, showing that 44% of sulfur is in the form of pyrrhotite (pentlandite has negative loading). Lower left: The 2475-2482 eV region is shaped by the broad oscillation of the dominant pyrrhotite component. In the GCF approach, this region is modelled by adding Gaussian functions for sulfite and sulfone, which in reality are absent. Lower right: The same phenomenon is observed in protein-rich fractions of natural organic matter (Fig. 6), because cysteine, and to a lesser extent cystine, also have a relatively intense oscillation after the main peak.

Two equivalent mathematical solutions of the Orthous-Daunay data (2010) are presented in Figure EA-8, one (Model-fit 1) in which the first arctangent is placed to the right of the sulfoxide peak (original fit by the authors) and another in which the arctangent is placed to the left (Model-fit 2). The difference in S_{sulfx} after correction of the absorption cross-sections with the generic scaling factors is 6.4% – 1.1% = 5.3% (Table EA-6). The second model-fit gives a S_{Sulfx} value in good agreement with the Combo fit (1.6%). The rise in absorbance between 2476 eV and 2479 eV was reproduced by introducing sulfite and sulfone in the fits. The Combo fit simulation shows that these two species are in reality absent, and that the particular spectral shape in the region between reduced (S_{red}) and oxidized (S_{ox}) sulfur species is caused by the broad resonance oscillation of pyrrhotite, which is the most abundant species in this chondrite, making up 44% of total S. The uncommon spectral shape in the intermediate S-XANES region of natural organic matter is also observed in fractions enriched in cysteine, such as proteinaceous fractions (Fig. 6). Lastly, the fractions of inorganic sulfides (pyrrhotite) and sulfates in the previous study (Orthous-Daunay et al., 2010) were 29% and 4%, respectively, compared to 44% and 1.6% from the Combo fit. This difference is mainly because of the fact that the calibration curve in the previous study (Fig. 2) was measured in fluorescence-yield detection mode and the data for the reference compounds not corrected for overabsorption (Manceau et al., 2002).

6.3. Sulfur speciation in soil

The accuracy and precision of LCF and GCF methods were evaluated recently on binary, ternary and quaternary mixtures of FeS, L-cysteine, and Na_2SO_4 (set 1), and FeS, FeS_2, S°, and L-cysteine (set 2), containing variable proportions of each constituent. Then the two methods were applied to speciate sulfur in soils (Prietzel et al., 2011). In the GCF method, all spectra were analyzed by Prietzel et al. (2011) with the $\sum_{i=1}^{j}(G_{w=1}+A_{w=0.3})_i$ model-fit shown in Figure 1a and discussed in their article. Gaussian areas were normalized with scaling factors derived from calibration curve (Fig. 2) calculated from the measurement in fluorescence-yield of ten pure compounds diluted in a finely-ground quartz matrix (2 mg S g^{-1}) to prevent overabsorption. Spectra from the first series have the characteristic shape in the 2475-2482 eV region of samples rich in pyrrhotite and cysteine described previously (Fig. EA-9).

The Figure on the left below shows that the $\sum_{i=1}^{3}(G_{w=1}+A_{w=0.3})_i$ model-fit failed to reproduce data from the FeS-cysteine-sulfate mixture, because the slope in the 2475-2482 eV region, arising from the post-edge of the two reduced sulfur species, cannot be modelled with an asymptotic function. In addition, the arctangents were moved by several eV away from their respective Gaussian functions, causing the total $(G+A)_i$ absorbance to drop to zero after the white-line. This mathematical model is clearly incorrect because the absorbance never returns to its pre-edge value after the main edge, even when the peak-to-background amplitude is high (Fig. EA-1). Applied to soil samples, this flawed model-fit dramatically overestimates the fraction of sulfoxide, and does not reproduce well the data, as seen for example in the misfit of the S_{ox} region.

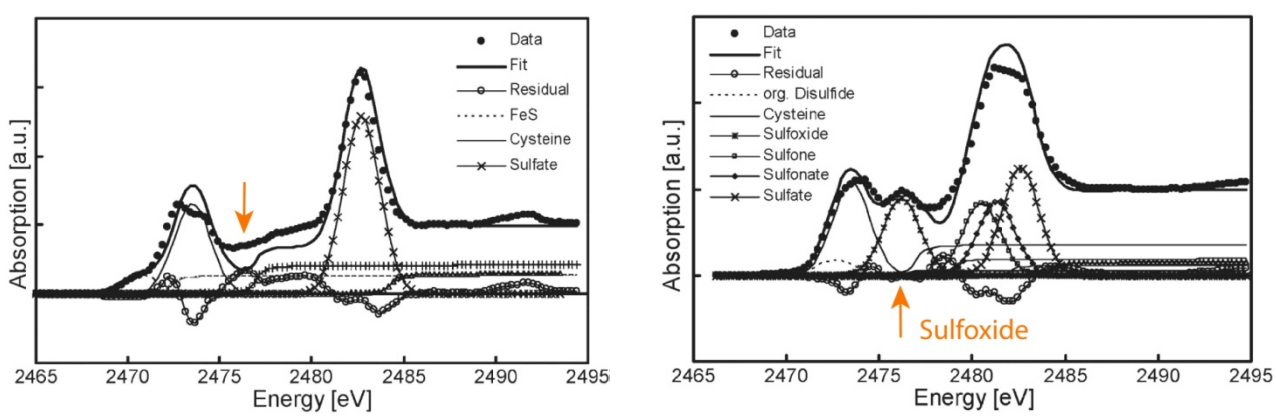


Fig. EA-9. Left: GCF decomposition of a mixture containing equal proportions of S in FeS, L-cysteine, and Na sulfate. Right: GCF decomposition of a soil spectrum. Reproduction from Prietzel et al. (2011). The orange arrows point out the dips in absorbance caused by the incorrect positioning of the arctangent function.

6.4. Incorrect positioning of the arctangents

One of the most common types of errors encountered in the literature is the incorrect positioning of the arctangent(s) in the final fit. Some additional examples are given below to caution modelers against this problem.

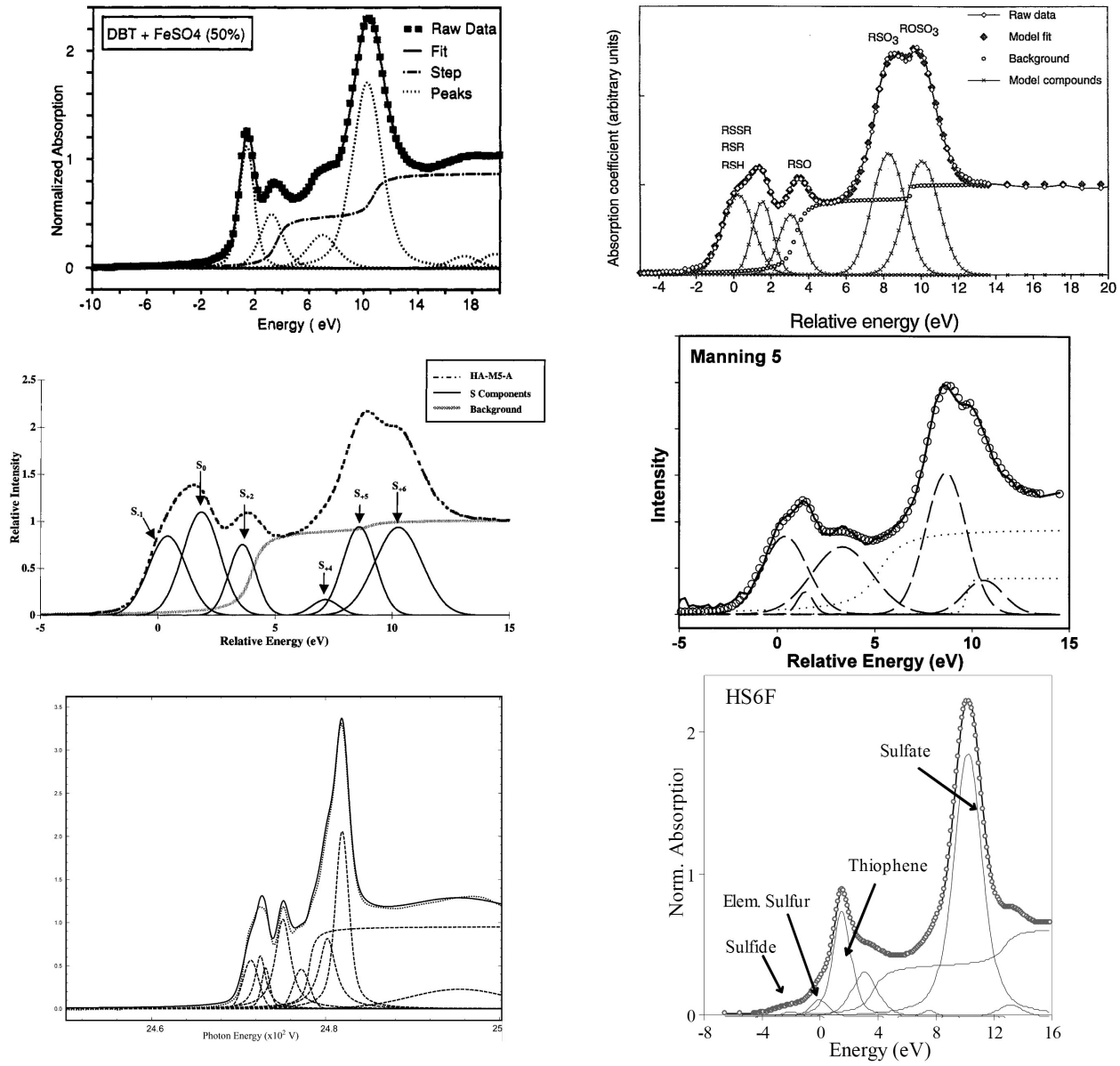


Fig. EA-10. Cases in which the first arctangent is at too high energy. Upper left: From Huffman et al. (1991). Upper right: From Skyllberg et al. (2000). Middle left: From Szulczewski et al. (2001). Middle right: From Martinez et al. (2002). Lower left: From Wiltfong et al. (2005). Lower right: From Pattanaik et al. (2007).

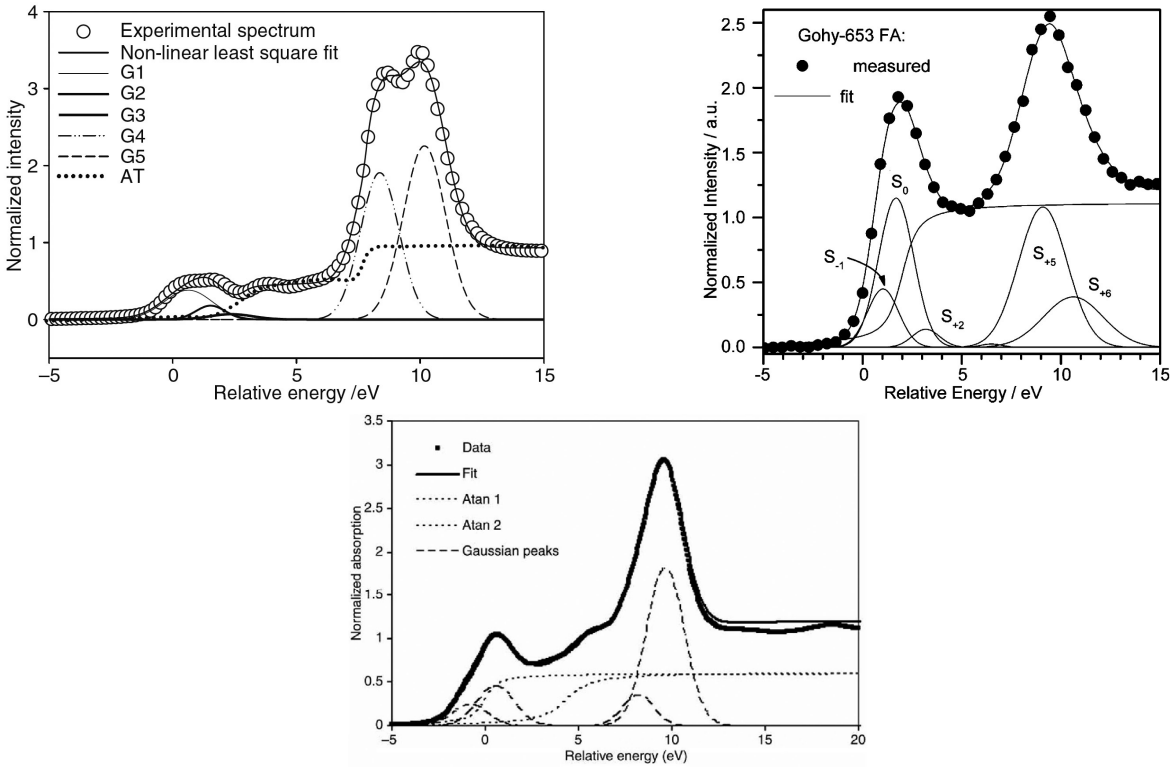
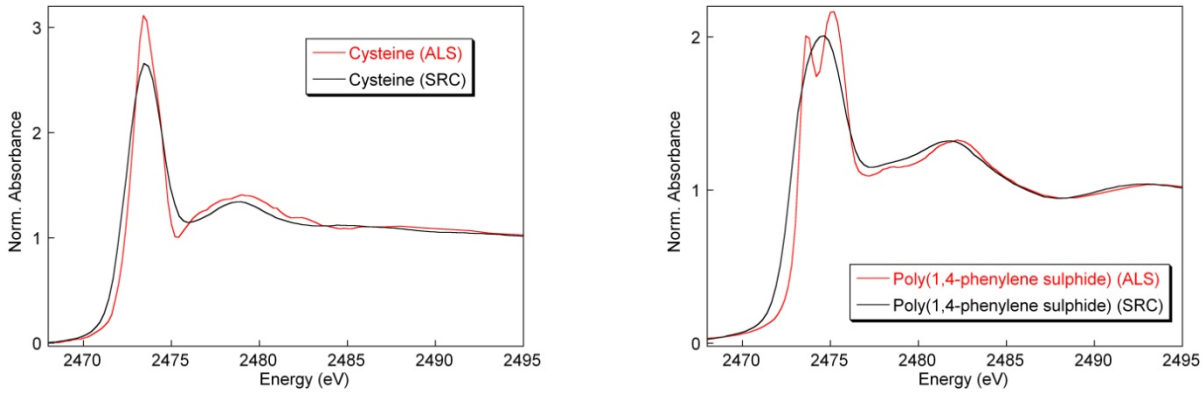


Fig. EA-11. Cases in which at least one arctangent is at too low energy. Upper left: From Solomon et al. (2005). Upper right: From Schäfer et al. (2005). Lower: From Cai et al. (2009). The Gaussian describes the transition of the photoelectron to bound states, and the arctangent to the continuum. Therefore the second function cannot be placed to the left of the first.

7. Effect of spectrometer resolution on the shape of S-XANES spectra



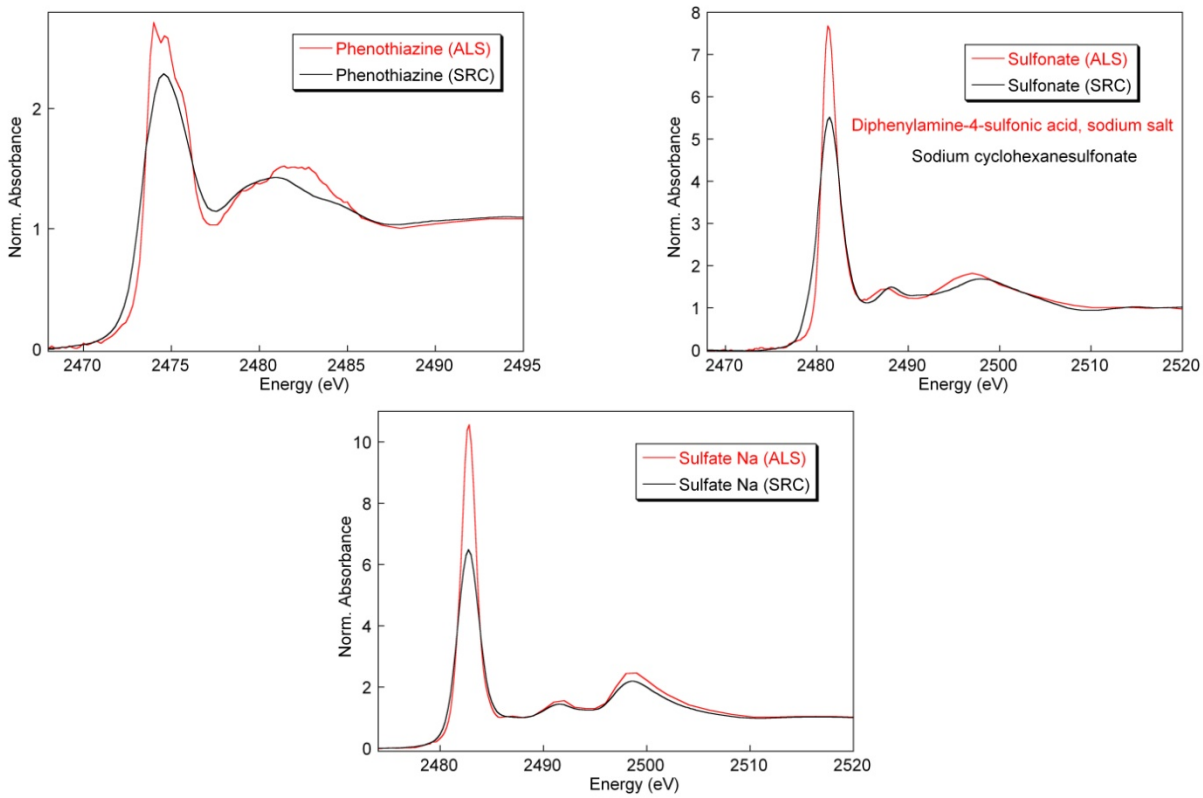


Fig. EA-12. Effect of spectral resolution on the shape of sulfur K-edge XANES spectra for the same compounds recorded at the Advanced Light Source (ALS) and the Canadian Synchrotron Radiation Facility (CSRF) located at the Synchrotron Radiation Center (SRC) of the University of Wisconsin, Madison. The CSRF data are from Cai et al. (2009), courtesy of Eric Morris.

8. References

Cai J. H., Morris E. and Jia C. Q. (2009) Sulfur speciation in fluid coke and its activation products using K-edge X-ray absorption near edge structure spectroscopy. *J. Sulfur Chem.* **30**, 555-569.

Huffman G. P., Mitra S., Huggins F. E., Shah N., Vaidya S. and Lu F. (1991) Quantitative analysis of all major forms of sulfur in coal by X-ray absorption fine structure spectroscopy. *Energy Fuels* **5**, 574-581.

Manceau A., Marcus M. A. and Tamura N. (2002) Quantitative speciation of heavy metals in soils and sediments by synchrotron X-ray techniques. In *Applications of Synchrotron Radiation in Low-Temperature Geochemistry and Environmental Science*, vol. 49 (eds. P. A. Fenter, M. L. Rivers, N. C. Sturchio and S. R. Sutton). Mineralogical Society of America, Washington, DC.

Martinez C. E., McBride M. B., Kandianis M. T., Duxbury J. M., Yoon S. J. and Bleam W. F. (2002) Zinc-sulfur and cadmium-sulfur association in metalliferous peats: Evidence from spectroscopy, distribution coefficients, and phytoavailability. *Environ. Sci. Technol.* **36**, 3683-3689.

Orthous-Daunay F. R., Quirico E., Lemelle L., Beck P., deAndrade V., Simionovici A. and Derenne S. (2010) Speciation of sulfur in the insoluble organic matter from carbonaceous chondrites by XANES spectroscopy. *Earth Planet. Sci. Let*. **300**, 321-328.

Pattanaik S., Huggins F. E., Huffman G. P., Linak W. P. and Miller C. A. (2007) XAFS studies of nickel and sulfur speciation in residual oil fly-ash particulate matters (ROFA PM). *Environ. Sci. Technol.* **41**, 1104-1110.

Prietzel J., Botzaki A., Tyufekchieva N., Brettholle M., Thieme J. and Klysubun W. (2011) Sulfur speciation in soil by S K-edge XANES spectroscopy: Comparison of spectral deconvolution and linear combination fitting. *Environ. Sci. Technol*. **45**, 2878-2886.

Sarret G., Conna J., Kasrai M., Bancroft G. M., Charrié-Duhaut A., Lemoine S., Adam P., Albrecht P. and Eybert-Bérard L. (1999) Chemical forms of sulfur in geological and archeological asphaltenes from

Middle East, France, and Spain determined by sulfur K- and L-edge X-ray absorption near-edge structure spectroscopy. *Geochim. Cosmochim. Acta*. **63**, 3767-3779.
Schäfer T., Buckau G., Artinger R., Kim J. I., Gyer S., Wolf M., Bleam W. F., Wirick S. and Jacobsen C. (2005) Origin and mobility of fulvic acids in the Gorleben aquifer system: implications from isotopic data and carbon/sulfur XANES. *Org. Geochem.* **36**, 567-582.
Skyllberg U., Xia K., Bloom P. R., Nater E. A. and Bleam W. F. (2000) Binding of mercury(II) to reduced sulfur in soil organic matter along upland-peat soil transects. *J. Environ. Qual*. **29**, 855-865.
Solomon D., Lehmann J., Lobe I., Martinez C. E., Tveitnes S., Du Preez C. C. and Amelung W. (2005) Sulphur speciation and biogeochemical cycling in long-term arable cropping of subtropical soils: evidence from wet-chemical reduction and S K-edge XANES spectroscopy. *Eur. J. Soil Sci*. **56**, 621-634.
Szulczewski M. D., Helmke P. A. and Bleam W. F. (2001) XANES spectroscopy studies of Cr(VI) reduction by thiols in organosulfur compounds and humic substances. *Environ. Sci. Technol*. **35**, 1134-1141.
Wiltfong R., Mitra-Kirtley S., Mullins O. C., Andrews B., Fujisawa G. and Larsen J. W. (2005) Sulfur speciation in different kerogens by XANES spectroscopy. *Energy Fuels* **19**, 1971-1976.
Xia K., Weesner F., Bleam W. F., P.R. B., Skyllberg U. and Helmke P. A. (1998) XANES studies of oxidation states of sulfur in aquatic and soil humic substances. *Soil Sci. Soc. Am. J.* **62**, 1240-1246.
Yoon S. J., Diener L. M., Bloom P. R., Nater E. A. and Bleam W. F. (2005) X-ray absorption studies of CH_3Hg^+-binding sites in humic substances. *Geochim. Cosmochim. Acta*. **69**, 1111-1121.