JOURNAL OF SYNCHROTRON RADIATION

ISSN 1600-5775

Received 22 August 2018
Accepted 14 January 2019

Edited by R. W. Strange, University of Essex, UK

**Keywords:** X-ray fluorescence; per-pixel baselines; X-ray microprobes; elemental maps; X-ray nanoprobes.

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

# *M-BLANK*: a program for the fitting of X-ray fluorescence spectra

**Andrew M. Crawford,[a,b] Aniruddha Deb[a] and James E. Penner-Hahn[a,c]***

[a]Department of Chemistry, University of Michigan, 930 N. University Avenue, Ann Arbor, MI 48109-1055, USA, [b]Department of Geology, University of Saskatchewan, Saskatoon, SK S7N 5E2, Canada, and [c]Department of Biophysics, University of Michigan, 930 N. University Avenue, Ann Arbor, MI 48109-1055, USA.
*Correspondence e-mail: jeph@umich.edu

The X-ray fluorescence data from X-ray microprobe and nanoprobe measurements must be fitted to obtain reliable elemental maps. The most common approach in many fitting programs is to initially remove a per-pixel baseline. Using X-ray fluorescence data of yeast and glial cells, it is shown that per-pixel baselines can result in significant, systematic errors in quantitation and that significantly improved data can be obtained by calculating an average blank spectrum and subtracting this from each pixel.

## 1. Background

Metals perform a multitude of functions inside cells. With the development of third-generation synchrotron sources, it has become possible to use X-ray fluorescence (XRF) to image the elemental composition of individual cells with a resolution of tens of nanometres (Pushie *et al.*, 2014; Penner-Hahn, 2014; Paunesku *et al.*, 2006; Punshon *et al.*, 2013; Ortega *et al.*, 2009; Vogt & Ralle, 2013; Zhao *et al.*, 2014; West *et al.*, 2014). XRF, which can also be used with flow cytometry (Crawford *et al.*, 2016), is the only approach that can determine total metal content independent of the chemical form of the metal and its surrounding environment, making XRF an ideal tool for interrogating the cellular metallome.

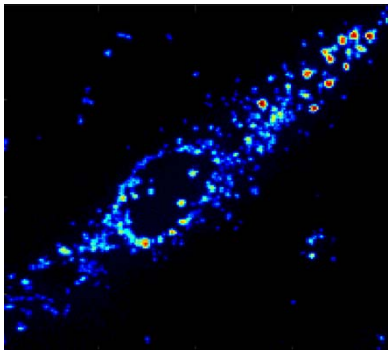
A variety of programs are in wide use at synchrotron laboratories to perform the fitting that is necessary to obtain quantitative elemental compositions from XRF data sets, including: *Axil* (Vekemans *et al.*, 1994), *GeoPIXE/Dynamic Analysis* (Ryan *et al.*, 2005), *MAPS* (Vogt, 2003), *PyXRF* (Li *et al.*, 2017) and *PyMCA* (Solé *et al.*, 2007). While these programs differ in detail, one common feature is that they all use some form of background subtraction (see below for more details). In the following, we show that, under certain conditions, background subtraction can result in significant errors for overlapping peaks. We describe a different approach to the quantitative analysis of XRF data based on blank rather than background removal and present a *MATLAB* program, *M-BLANK*, to implement this approach.

### 1.1. XRF data analysis

A raw XRF spectrum records X-ray counts as a function of energy, as described by equation (1),

$$\mathrm{XRF}_i = S_i + B_i, \tag{1}$$

where $S_i$ is the signal counts, or characteristic X-rays, from the sample and $B_i$ is the background counts arising from every-



thing else at each energy $i$. For an ideal detector, the signal $S_i$ would be Gaussian with a peak width defined by the detector resolution. In practice, the peak shape is distorted because of incomplete charge collection (see the supporting information for more detail on peak shapes). The background $B$ may include counts from elastic and inelastic scattering, sample holder fluorescence, secondary fluorescence from the hutch, and any detector dark current. For non-synchrotron excitation sources, Bremsstrahlung radiation would also contribute to the background.

The simplest analytical approach for XRF data is to sum all the counts in a region of interest (ROI, $i = a$ to $b$) corresponding to a particular elemental emission,

$$\mathrm{XRF} = \sum_{a}^{b} \mathrm{XRF}_i = \sum_{a}^{b} S_i + \sum_{a}^{b} B_i. \qquad (2)$$

In principle, conversion from fluorescence counts to elemental mass could be as simple as comparing this sum with the corresponding sum for a standard after correcting each for the background counts. In practice, the ROI approach will not work for most samples because of peak overlap (Crawford, Sylvain *et al.*, 2018). For example, most energy-resolving fluorescence detectors cannot distinguish the manganese $K\beta$ peak from the iron $K\alpha$ peak. For samples containing both, the iron ROI will include some fraction of manganese counts and vice versa; accurate quantitative analysis requires that the full peak shape be fit.

For this reason, XRF spectra are typically fitted as the sum of a background function together with one or more fluorescence peaks, often modeled as modified Gaussians, as illustrated in Fig. S1 in the supporting information (van Espen & Lemberge, 2000; Campbell & Maxwell, 1997; van Grieken & Markowicz, 2002). This approach depends on having an accurate background function. Although scattering and dark-current give a relatively smooth background, the background also often contains discrete impurity peaks (Bos *et al.*, 1984; Gordon, 1982), making modeling of the XRF background challenging. Most analytical packages, *e.g. MAPS* (Vogt, 2003), *PyMCA* (Solé *et al.*, 2007), *PyXRF* (Li *et al.*, 2017) and *Axil* (Vekemans *et al.*, 1994; Janssens *et al.*, 1996), start with calculating a smooth baseline (Vogt, 2003; Solé *et al.*, 2007; Janssens *et al.*, 1996; Brunetti & Steger, 2000; Ren *et al.*, 2014; Yi *et al.*, 2015). Conceptually, this amounts to replacing equation (1) with equation (3),

$$\mathrm{XRF}_i = S_i + F_i + L_i, \qquad (3)$$

where the background, $B$, has been partitioned into a smooth baseline, $L$, and one or more background fluorescence peaks, $F$ (*e.g.* from the sample holder or secondary fluorescence from the hutch). Once fitted, $L$ can be subtracted from the data, leaving the fluorescence peaks ($S$ and $F$) to be fitted. After fitting, the background fluorescence, $F$, can be determined from the signal in certain (*e.g.* cell free) regions of the sample. One of the most common baseline calculations uses a peak-stripping function (van Grieken & Markowicz, 2002). Each channel, $n$, of the spectrum is compared with its two neighboring channels, $n \pm 1$; if the intensity at channel $n$ is greater than the mean intensity of channels $n \pm 1$ then the intensity at channel $n$ is set to the mean intensity of its two neighbors. Performed iteratively, this process results in peak-stripping and creates a baseline passing through local minima. For the remainder of this article, we use 'baseline' to refer to 'per pixel baseline' (*i.e.* a baseline that is calculated independently for each pixel without any *a priori* knowledge of whether the pixel represents sample or blank).

There are a number of alternatives for background removal. For example, the statistics-sensitive non-linear iterative peak-clipping SNIP) algorithm (Ryan *et al.*, 1988) is widely used. This algorithm has a smoothing window with a width that is indirectly proportional to the total counting statistics within a region. This acts to smooth low-count regions (*e.g.* the transition elements for the data in this article) while leaving high-count regions [*e.g.* calcium (4 keV) and below] alone. After the spectrum is smoothed the data are then processed using an iterative peak-stripping (or clipping) algorithm. Another common approach is to use some sort of filtering, such as the top-hat filter (Schamber, 1977) which is widely used in energy-dispersive X-ray spectrometer (EDS) software (Schamber, 2009). This is a linear filter which removes background at the expense of spectral distortion. However, by applying the same filter to both the data and the standards, it is possible to recover accurate quantitation. As such filtering is not widely used in synchrotron XRF fitting programs, we have restricted our comparison to baseline-fitting algorithms.

### 1.2. Blank subtraction

Rather than baseline subtraction, we have taken the approach of blank subtraction. This is possible because our XRF images inevitably contain some cell-free regions. The cell-free regions include contributions from the background fluorescence [$F$ in equation (3)] but not from the sample. We calculated the mean X-ray fluorescence spectrum from these regions in order to define a blank which can be subtracted, instead of a baseline, directly from equation (2), leaving only the signal.

In order to be valid, this approach depends on the blank being uniform. As shown in Fig. S2, we find excellent uniformity aside from small variations in the silicon signal, which arises from the $Si_3N_4$ substrate upon which the cells are mounted. The variation in the silicon signal is somewhat larger than would be expected from counting statistics; it most likely represents small variations in the thickness of the $Si_3N_4$ since cellular debris from damaged cells would have to form a relatively thick layer (>1 µm) to significantly attenuate the Si $K\alpha$ fluorescence. Based on these observations, it would be possible for these data to use a single blank, averaged over all of the non-cell regions of all of the images [and such observations are in agreement with Fig. 1(*a*)]. However, we have instead calculated a blank for each image, using the cell-free pixels in that image; we believe that this is best practice when possible.

We describe a program, *M-BLANK*, that implements the blank-subtraction approach and demonstrate that it provides

better accuracy, avoiding systematic errors associated with baseline subtraction. In some cases, *M-BLANK* also provides better precision than baseline-fitting programs. In order to compare blank and baseline-analysis procedures, we present here two specific cases where the XRF data for yeast and glial cells were analyzed by both methods. In addition to being important model systems, both yeast (Wang *et al.*, 2014; Crawford & Penner-Hahn, 2018; Crawford *et al.*, 2016; Crawford, 2015; Zhao *et al.*, 2015; Dima *et al.*, 2017; Ballo *et al.*, 2017) and glial cells (Al-Ebraheem *et al.*, 2016; Busse *et al.*, 2017; Ide-Ektessabi *et al.*, 2002; Stamenković *et al.*, 2017) have been studied extensively using X-ray fluorescence. As such, the artifacts demonstrated here may be more broadly represented across the biological XRF literature.

## 2. Consequences of per-pixel baseline subtraction

Ideally, baselines should be insensitive to signal details [*i.e.* $L$ in equation (3) should not depend on $S$ or $F$]. In contrast, we find baselines can be quite sensitive, as illustrated in Fig. 1(*a*). The calculated baselines for the XRF from cell (black) and blank (red) pixels are quite different. Even more striking is the pixel-to-pixel variation (gray and pink lines). This variability reflects the sensitivity of the fitted baseline to even small changes in elemental composition.

It is noteworthy that the calculated baselines for blank pixels are systematically lower than those for cell pixels. This happens because cell pixels, but not blank pixels, have significant counts from phosphorus, sulfur and potassium (peaks at ∼2.0, 2.3 and 3.3 keV), resulting in the former having significantly higher baselines than the latter. This difference is large enough that the baseline amplitudes alone give a good reproduction of the budding yeast cell in this sample [Fig. 1(*b*): budding refers to the asexual reproduction in some yeast, where the new cell is formed as an outgrowth or 'bud' from the parent cell]. An elevated baseline such as that in Fig. 1(*b*) causes a decrease in the apparent intensity from the 'cell' pixels since a larger baseline is subtracted from cell pixels than from the non-cell pixels [this is the black region in Fig. 1(*b*)]. Consequently, the apparent silicon and chlorine signals (that is, the measured signals minus the fitted per-pixel baselines) are smaller for the cell-containing pixels than they are for the non-cell regions. This results in an apparent decrease in chlorine concentration in the cell-containing regions [see Fig. 1(*c*)]. This is not only biologically suspect [cells typically have m$M$ chloride concentrations (Fraústo da Silva & Williams, 1993), although these can vary depending on growth conditions] but is also at odds with the clear observation [Fig. 1(*a*)] that the cell pixel has slightly more chlorine counts than the non-cell pixel. The latter point can be seen by comparing the cell pixel (black line, ∼19 counts $s^{-1}$ at the maximum in the chlorine $K\alpha$ peak) with the non-cell pixel (red line, ∼9 counts $s^{-1}$ at the maximum in the chlorine $K\alpha$ peak).

This behavior can be understood from equation (3): since $L_{\rm blank} < L_{\rm cell}$, the correction $\mathrm{XRF}_i - L_i$ over-corrects the data by $L_{\rm cell} - L_{\rm blank}$. If $L_{\rm cell} - L_{\rm blank}$ is larger than the actual chlorine fluorescence, the cell pixels appear to have less chlorine than the blank pixels, giving an apparent negative chlorine concentration [Fig. 1(*c*)]. In contrast, *M-BLANK* gives physiologically reasonable chlorine concentrations [Fig. 1(*d*), the integrated areal chlorine signal of ∼0.10 μg $cm^2$ gives an estimated cellular concentration of 4.5 m$M$ assuming a cellular diameter of 6 μm].

Sample-dependent variation in baseline appears to affect only the chlorine quantitation for the cell in Fig. 1 (silicon is also affected, but this is not biologically relevant). However, the effect is likely to occur any time that there are two adjacent peaks with significantly different amplitudes. For example, this artifact may account for the apparent decrease in calcium content that was seen for cell pixels having high potassium concentrations in a study of the effect of chemical fixation on cellular composition (Jin *et al.*, 2017).

This artifact also occurs for glial-cell images. Fig. 2(*a*) shows sulfur and chlorine images obtained using blank correction (*M-BLANK*) and per-pixel

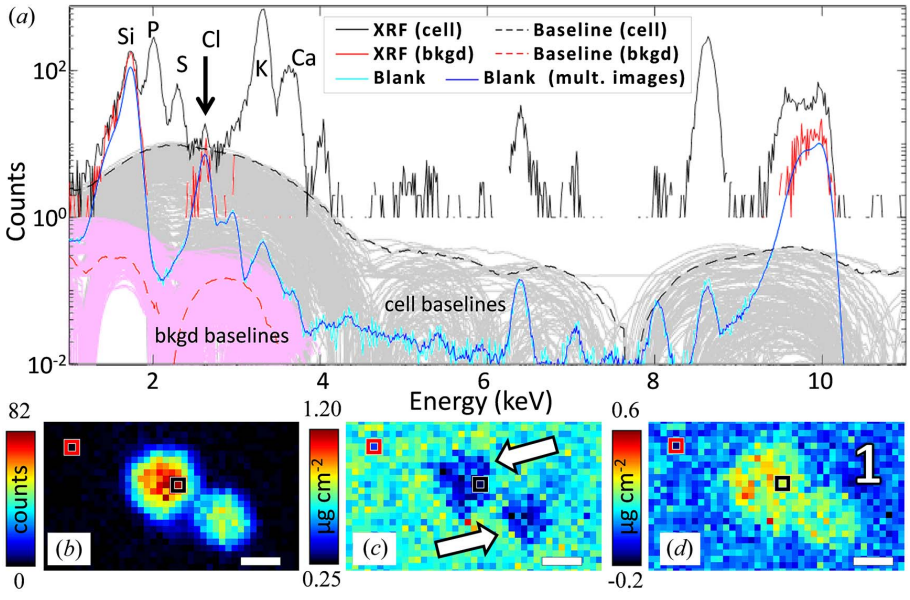



**Figure 1**
Effect of baseline variation on chlorine quantitation. (*a*) Representative XRF signals from individual cell (black) and blank (red) pixels with the calculated blank from this image (cyan) and multiple images (blue). Individual per-pixel baselines for all of the pixels are shown in gray and pink; baselines for the two pixels where the XRF is shown are the black and red dashed lines. The black arrow indicates the position of the chlorine peak. (*b*) Integrated baseline amplitude at the chlorine $K\alpha$ peak ± 50 eV. (*c*) Quantitation for chlorine using per-pixel baseline correction, showing an apparent decrease in chlorine signal in the cell relative to the background (after removing the background, this would give an apparent negative chlorine concentration). (*d*) Quantitation for chlorine using *M-BLANK*. Black and red squares in (*b*)–(*d*) identify the pixels used in (*a*). The scale bar represents 2 μm.

baseline correction (*MAPS*). These samples were contaminated with small sodium chloride crystals, as a result of the sample-preparation method. The crystals give sharp spikes in the chlorine distribution which perturb the per-pixel baseline for that pixel, causing that pixel's baseline to be higher than it would be otherwise. When this happens, the per-pixel baseline correction over-corrects the sulfur signal, with the result that there are wells in the sulfur distribution (*i.e.* pixels with negative apparent sulfur concentration). This can be seen, for example, by comparing the chlorine and sulfur values marked by the white arrows; blank-corrected images do not contain this artifact.

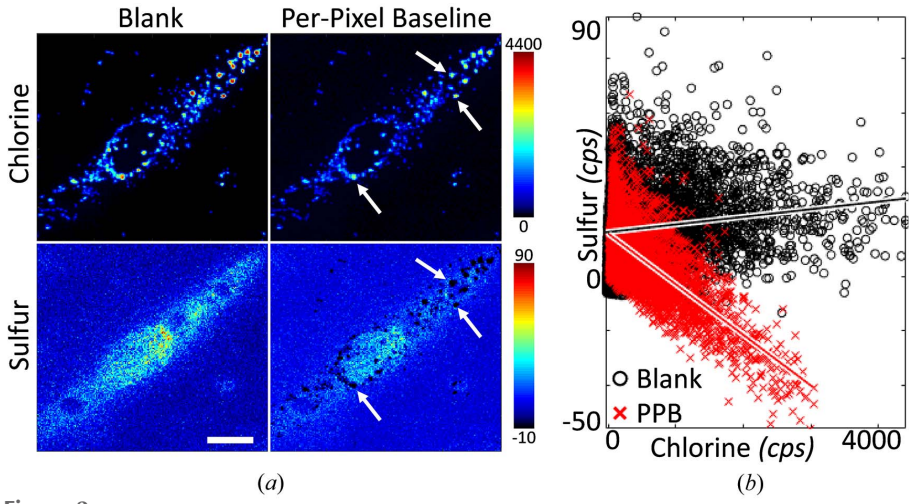



**Figure 2**
Chlorine and sulfur in glial cells. (*a*) Apparent chlorine (top) and sulfur (bottom) quantitation for glial cells as a result of blank correction (left) and per-pixel baseline correction (right). (*b*) Correlation plots (in units of counts $s^{-1}$) for sulfur plotted as a function of chlorine for blank (○) and per-pixel baseline (×) correction. The intensity bars are in counts $s^{-1}$ and the scale bar represents 10 μm.

The negative correlation between apparent sulfur and chlorine concentration for baseline-corrected data can also be seen in the scatter plot of sulfur versus chlorine counts [Fig. 2(*b*)]. Since the chlorine in these samples is largely an artifact of sample preparation, one would not necessarily expect there to be any correlation between sulfur and chlorine. That is, there is no reason to expect the sodium chloride crystals to form at positions of low or high sulfur. This is what is observed for the blank-corrected data. In contrast, the baseline-corrected data show a strong negative correlation between the apparent sulfur and chlorine concentrations, consistent with our interpretation that high chlorine concentration leads to over-correction, and thus lower than expected apparent sulfur concentration.

Although baseline distortion is particularly severe for adjacent peaks with very different amplitudes (*e.g.* potassium and chlorine), it is unlikely to be limited to these. This phenomenon could occur any time that amplitude variation in one peak causes a significant change in the fitted baseline, for example in studies of copper/zinc or arsenic/selenium interactions, or in studies of silver and cadmium, where the *L* emission lines overlap with the chlorine and potassium *Kα* lines.

The examples in Figs. 1 and 2 both result from the distortions that occur when two adjacent elements have very different concentrations. However, the phenomenon of baseline distortion is a more general problem. In Fig. 3, the per-pixel baseline and blank-corrected concentrations are compared for yeast cells. At high concentrations, both methods give similar results. However, at low concentrations, the lower-*Z* elements [Fig. 3(*d*)] show systematically decreased concentrations for per-pixel baseline correction relative to blank correction. We attribute this to baseline distortions similar to those shown in Fig. 1. This can be modeled by rewriting equation (3) to compare the baseline-corrected signal [left-hand side of equation (4)] with the blank-corrected signal (right-hand side),

$$(\mathrm{XRF}_i - L_i)_{\mathrm{cell}} - \overline{(\mathrm{XRF} - L)_{\mathrm{blank}}} = \mathrm{XRF}_{i,\mathrm{cell}} - \mathrm{XRF}_{\mathrm{blank}}. \tag{4}$$

If the apparent mass calibrations between the two techniques are identical and baselines in cell pixels [$L_i$ in equation (4)] and in blank regions [$L$ in equation (4)] are identical, equation (4) will be true. Otherwise, we find instead

$$\mathrm{XRF}_{i_{\mathrm{baseline\ corrected}}} = m\mathrm{XRF}_{i_{\mathrm{blank\ corrected}}} - \Delta L, \tag{5}$$

where $\Delta L = L_{i,\mathrm{cell}} - \overline{L_{\mathrm{blank}}}$ and $m$ is the ratio $(\mathrm{mass}_{\mathrm{baseline\ corrected}})/(\mathrm{mass}_{\mathrm{blank\ corrected}})$. Equation (5) is the functional form representing the fitted masses from baseline correction as a function of blank correction. Dividing both sides of equation (5) by $\mathrm{XRF}_{i_{\mathrm{blank\ corrected}}}$, this can be rearranged to equal

$$y_i = m - \frac{\lambda}{x_i}, \tag{6}$$

where

$$x_i = \frac{\mathrm{XRF}_{i_{\mathrm{blank\ corrected}}}}{\max(\mathrm{XRF}_{\mathrm{blank\ corrected}})}, \qquad y_i = \frac{\mathrm{XRF}_{i_{\mathrm{baseline\ corrected}}}}{\mathrm{XRF}_{i_{\mathrm{blank\ corrected}}}},$$

and

$$\lambda = \frac{\Delta L}{\max(\mathrm{XRF}_{\mathrm{blank\ corrected}})}$$

with $y_i$ having units identical to $m$, and $x_i$ and λ being fractional masses relative to the maximum fitted value for the given element [*e.g.* for Fig. 3(*a*) this is sulfur]. In equation (6), the mass and baseline are normalized to the maximum in order to facilitate comparison between elements that are present at very different levels (see below).

If λ is zero then a plot of equation (6) will yield a constant-value curve at $m$. If the per-pixel baseline over compensates ($\lambda > 0$), the ratio, $y_i$, will asymptotically approach $m$ (the ratio of calibration terms) from underneath, as seen for the low-*Z*

elements [Fig. 3(*d*)]. Alternatively, if the per-pixel baseline under compensates ($\lambda < 0$), the ratio, $y_i$, will asymptotically approach $m$ (the ratio of calibration terms) from above, as seen for the transition elements [Fig. 3(*e*)].

Applying equation (6), we see that the behavior for sulfur is especially complex. Not only do we see a nonzero value for $\lambda$, we see eight different sub-populations, or groupings in the plot, with each one of these groupings indicating a different distinct distortion, or $\lambda$ [Fig. 3(*a*)].

Since the data comprising Fig. 3 came from multiple yeast images, it is possible that the different populations arose from differential treatment of the per-pixel baselines across samples caused by variable background continuums. Such a scenario would allow for the comparison within a single image but make it difficult to compare across images. To test for this possibility, we color-coded the eight sub-populations and used these colors to redraw the original images [Fig. 3(*a*), inset images]. From these data, it is obvious that the different sub-populations correspond to different regions within each cell. This is inconsistent with image-to-image variability; rather, this suggests that the sub-populations arise from sample-dependent baseline distortions, $\lambda$, which are systematically reproduced across all of the samples.

The variability shown in Fig. 3 makes it difficult to compare sulfur content when per-pixel baseline correction is used. An example of a difficulty arising from this artifact can be seen by comparing points a, b and c [Fig. 3(*a*)], corresponding to cell border, cell periphery just beyond the border and bud, respectively. These points have identical sulfur concentrations when fitted using blank correction. In contrast, the sulfur concentration that is determined using per-pixel baseline subtraction varies significantly from point to point; the baseline-corrected values at points a, b and c are 82, 65 and 49%, respectively, of those determined with blank subtraction. Not only does this artifact interfere with quantitation, it can also make identification of cell boundaries, especially for budding cells, more difficult.

Each of the eight sub-populations from Fig. 3(*a*) was fit using equation (6) to solve for each population's $\lambda$ value while constraining $m$ (the ratio of calibration terms should not change) to be constant across all eight populations. For nonzero $\lambda$, the errors within a sub-population (*i.e.* deviation of $y_i$ from 1) will always get worse as the concentration decreases. Equation (6) can be rewritten to solve for the limiting $x_i$ value, $x_i(\mathrm{lim})$, above which the error remains below some threshold. Equation (7) gives these limiting values for a 5% threshold [the absolute value sign in equation (7) is necessary to account for the possibility that $\lambda$ could be negative; see the supporting information for the derivation],

$$x_i(\mathrm{lim}) \geq \frac{20\,|\lambda|}{m}, \qquad \lambda > 0. \tag{7}$$

For each sub-population, equation (7) was used to calculate the limits of $x_i$ for which the baseline-subtracted estimate will be within 5% of $m$, and these values have been tabulated in the inset table of Fig. 3(*a*). For sub-population 1 we see that the baseline over-compensates by 1.2% of the maximum sulfur (*i.e.* $\lambda = 0.012$) and that the two approaches will not agree for sulfur concentrations that are <24% of the maximum sulfur value. Unfortunately, sub-population 1 does not persist much beyond $x_i = 0.1$, and consequently the baseline-subtracted value is never within 5% of the blank-corrected data. Although sub-populations 3–7 have average errors that are somewhat smaller than those for sub-populations 1 and 2 (*i.e.* average $y_i$ closer to 1), even the former never get within 5% of the blank-corrected data.

The different distinct regions, or sub-populations, arise from variations in the fitted per-pixel baselines caused by overlapping emission lines in high-statistics regions of the X-ray fluorescence spectrum. Since it is caused by overlapping emissions in high-statistics regions, the SNIP (Ryan *et al.*, 1988) algorithm will not avoid this. This can be seen by

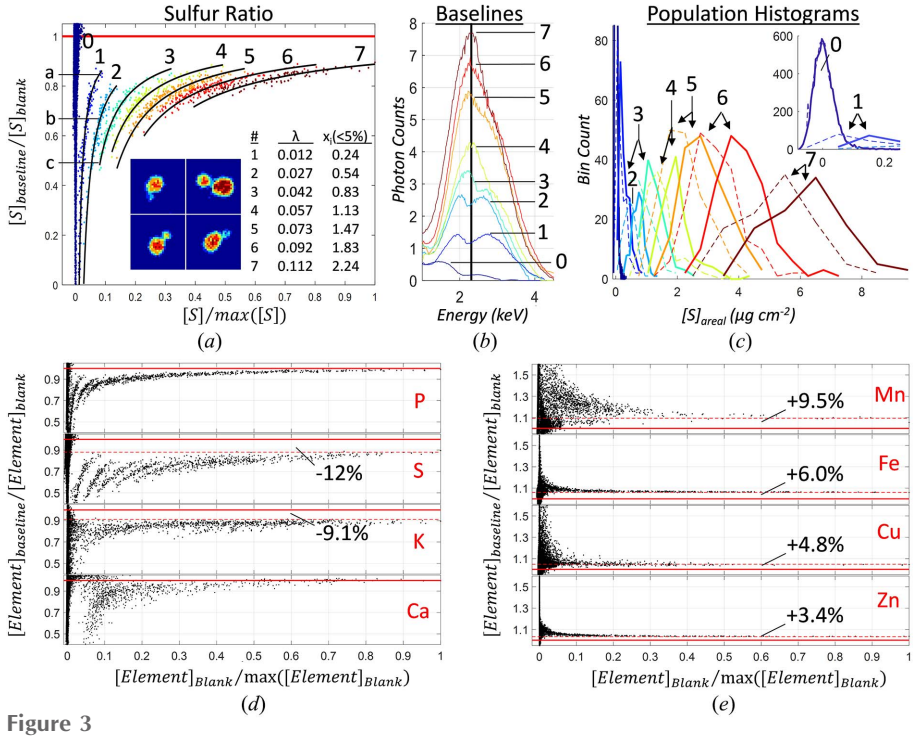



**Figure 3**
Comparison of per-pixel baseline and blank removal. (*a*) Ratio of per-pixel to blank-corrected fits for sulfur. Seven sub-populations (see the supporting information) are color coded on both the scatter plot and images. Black lines: fits to equation (6). (*b*) Representative per-pixel baselines for each sub-population (1–7) and blank pixels (0); the black line marks the sulfur *Kα* emission centroid. (*c*) Histograms of fitted sulfur concentrations for each sub-population. Solid = blank corrected; dash = per-pixel baseline corrected. (*d*) Comparison of per-pixel and blank removal for low-*Z* elements. (*e*) Comparison for transition metals. Solid red lines in (*d*) and (*e*) are a ratio of 1; dashed red lines mark the apparent asymptote.

comparing representative baselines for each sub-population [Fig. 3(*b*)]. Instead of a smooth, monotonic baseline, there are peaks at ~2 and ~2.6 keV (corresponding to $K\alpha$ emissions for phosphorus and chlorine, respectively, still present at varying degrees for the different baselines). As a consequence of these peaks, baseline subtraction consistently underestimates the amount of sulfur, but this underestimation is systematically different for each of the different sub-populations. This negative correlation between apparent sulfur concentration and phosphorus and chlorine concentration is equivalent to, although smaller than, that shown in Fig. 2. Even for the data in Fig. 3, the percent error in sulfur concentration can become significant as the amount of sulfur decreases. Only for blank pixels (sub-population 0), where there is no phosphorus, chlorine or potassium background, do baseline and blank subtraction agree [Fig. 3(*c*)]. A similar negative correlation is seen between the apparent chlorine concentration and the potassium concentration when using baseline subtractions (see Fig. S3).

Although they are not as numerous and well resolved, sub-populations are also seen for phosphorous, potassium and calcium [see Fig. 3(*d*)]. In addition to these errors in accuracy, the presence of sub-populations will decrease the precision of baseline-corrected data, since different sub-populations have different apparent baselines [as discussed above for points a, b and c in Fig. 3(*a*)].

In contrast to the lower-*Z* elements, ratio plots of the transition metals [Fig. 3(*e*)] do not show any systematic distortion at low concentration. The reason is twofold. First, there is very minimal overlap of the elemental emissions for these samples. This means there is very minimal baseline distortions of a given element by neighboring elements (*e.g.* Mn and Fe). Second, the energy region of the X-ray fluorescence spectrum corresponding to the transition elements is a very low statistics region. This results in many baselines that are essentially equal to 0 for most of the elements. There is, however, a small systematic error, with $(\mathrm{XRF}_{\mathrm{baseline\ corrected}})/(\mathrm{XRF}_{\mathrm{blank\ corrected}})$ asymptotically approaching, from above, values that are 3–10% greater than 1. We attribute this to a similar, but smaller, distortion in the baseline for the elemental standards.

## 3. Non-negativity

In addition to the improved quantitation discussed above, an advantage of *M-BLANK* is that it avoids artifacts that can result from non-negativity constraints. While non-negativity constraints are not an essential part of a per-pixel baseline correction, they are widely used in order to avoid parameter correlation (*e.g.* non-physical minima with alternating large positive and large negative amplitudes for adjacent peaks). They should not, however, be used with blank-corrected data as presented here with *M-BLANK* (see the supporting information for further discussion). The inability to use non-negativity constraints has the unanticipated benefit of allowing blank-corrected fitting in *M-BLANK* to avoid certain false-positive signals. Even if an element is not present in the sample, attempts to fit the peaks for this element will inevitably give occasional positive signals as a result of random noise, particularly for small signals in the presence of large signals. One such example is the detection of μ*M* cadmium in the presence of m*M* potassium (Fig. S4). As shown in Fig. S5, the use of non-negativity constraints can give false positives under such conditions, an artifact which does not affect *M-BLANK*. In addition, of course, the linear least-squares fits used in *M-BLANK* are much faster than non-linear least squares that are required for non-negativity.

## 4. Experimental details

Experimental details and calculation procedures are available in the supporting information.

## 5. Conclusions

We have demonstrated that per-pixel baseline subtraction can introduce artifacts in the apparent concentration of dilute elements, particularly when there are strong fluorescence peaks nearby, and have shown that this is important under biologically relevant conditions. We showed examples comparing *M-BLANK* with *MAPS*, but have found comparable results with other programs, including *PyMCA* and *PyXRF* (data not shown). *M-BLANK*, a *MATLAB* program which uses the background-subtraction methodology described here, is available on request from the authors. In addition to being both more accurate and more precise, blank subtraction has the additional advantage of permitting linear-least-squares fits, as compared with the non-linear fitting that is required by the non-negativity constraints typically used with baseline removal. This allows for much faster fitting and is more easily adapted to real-time analysis.

## 6. Related literature

The following citations refer to references used in the supporting information: Carter *et al.* (2010).

## Acknowledgements

The authors thank Andrea Stoddard, Nathan Zahler, Lubos Dostal and Carol Fierke for sample data. The yeast data used for testing *M-BLANK* were measured with the assistance of Nathan Zahler and Andrea Stoddard, Matthew Kidd and Marcy Hernick; the glial cell data were measured by Kelly Summers. Anna Turyanskaya assisted in the comparison of *M-BLANK* with *PyMCA*.

## Funding information

This research was funded in part by the National Science Foundation under NSF-IDBR-0852802 to JEPH, and used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract No. DE-AC02-06CH11357. This project was

supported by grant 9P41 GM103622 from the National Institute of General Medical Sciences of the National Institutes of Health.

## References


Al-Ebraheem, A., Miller, L., Geraki, K., Desmond, K. L., Stanisz, G. J., Bock, N. A. & Farquharson, M. J. (2016). *ICXOM23: International Conference on X-ray Optics and Microanalysis*, edited by J. Thieme & D. P. Siddons, September 14–18, 2015, Brookhaven National Laboratory, New York, USA.

Ballo, M. K. S., Rtimi, S., Kiwi, J., Pulgarin, C., Entenza, J. M. & Bizzini, A. (2017). *J. Photochem. Photobiol. B*, **174**, 229–234.

Bos, A. J. J., Vis, R. D., Verheul, H., Prins, M., Davies, S. T., Bowen, D. K., Makjanić, J. & Valkovicf, V. (1984). *Nucl. Instrum. Methods Phys. Res. B*, **3**, 232–240.

Brunetti, A. & Steger, T. J. (2000). *Nucl. Instrum. Methods Phys. Res. A*, **441**, 504–509.

Busse, M., Windsor, M. S. A., Tefay, A. J., Kardashinsky, M., Fenton, J. M., Morrison, D. E., Harris, H. H. & Rendina, L. M. (2017). *J. Inorg. Biochem.* **177**, 313–321.

Campbell, J. L. & Maxwell, J. A. (1997). *Nucl. Instrum. Methods Phys. Res. B*, **129**, 297–299.

Carter, E. A., Rayner, B. S., McLeod, A. I., Wu, L. E., Marshall, C. P., Levina, A., Aitken, J. B., Witting, P. K., Lai, B., Cai, Z. H., Vogt, S., Lee, Y. C., Chen, C. I., Tobin, M. J., Harris, H. H., Lay, P. A. (2010). *Mol. Biosystems*, **6**, 1316–1322.

Crawford, A. M. (2015). PhD thesis, University of Michigan, USA.

Crawford, A. M., Kurecka, P., Yim, T. K., Kozemchak, C., Deb, A., Dostál, L., Sun, C.-J., Brewe, D. L., Barrea, R. & Penner-Hahn, J. E. (2016). *J. Synchrotron Rad.* **23**, 901–908.

Crawford, A. M. & Penner-Hahn, J. E. (2018). *Cellular Heterogeneity*, *Methods in Molecular Biology*, Vol. 1745, 1st ed., edited by N. S. Barteneva & I. A. Vorobjev, pp. 97–112. New York: Humana Press.

Crawford, A. M., Sylvain, N. J., Hou, H., Hackett, M. J., Pushie, M. J., Pickering, I. J., George, G. N. & Kelly, M. E. (2018). *J. Synchrotron Rad.* **25**, 1780–1789.

Dima, S. O., Panaitescu, D. M., Orban, C., Ghiurea, M., Doncea, S. M., Fierascu, R. C., Nistor, C. L., Alexandrescu, E., Nicolae, C. A., Tricā, B., Moraru, A. & Oancea, F. (2017). *Polymers*, **9** 374.

Espen, P. van & Lemberge, P. (2000). *Advances in X-ray Analysis*, Vol. 43, pp. 560–569. Denver: JCPDS International Centre for Diffraction Data.

Fraústo da Silva, J. J. R. & Williams, R. J. P. (1993). *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*, pp. 1–561. Oxford: Clarendon Press.

Gordon, B. M. (1982). *Nucl. Instrum. Methods*, **204**, 223.

Grieken, R. E. van & Markowicz, A. A. (2002). *Handbook of X-ray Spectrometry: Methods and Techniques*, *Practical Spectroscopy Series*, Vol. 29, 2nd ed. New York: Marcel Dekker.

Ide-Ektessabi, A., Fujisawa, S. & Yoshida, S. (2002). *J. Appl. Phys.* **91**, 1613–1617.

Janssens, K., Vekemans, B., Adams, F., van Espen, P. & Mutsaers, P. (1996). *Nucl. Instrum. Methods Phys. Res. B*, **109–110**, 179–185.

Jin, Q. L., Paunesku, T., Lai, B., Gleber, S. C., Chen, S., Finney, L., Vine, D., Vogt, S., Woloschak, G. & Jacobsen, C. (2017). *J. Microsc.* **265**, 81–93.

Li, L., Yan, H., Xu, W., Yu, D., Heroux, A., Lee, W., -k, Campbell, S. I. & Chu, Y. S. (2017). *Proc. SPIE*, **10389**, 103890U.

Ortega, R., Devès, G. & Carmona, A. (2009). *J. R. Soc. Interface*, **6**, s649.

Paunesku, T., Vogt, S., Maser, J., Lai, B. & Woloschak, G. (2006). *J. Cell. Biochem.* **99**, 1489–1502.

Penner-Hahn, J. E. (2014). *Metallomics and the Cell*, *Metal Ions in Life Sciences*, edited by L. Banci, pp. 15–40. Dordrecht: Springer.

Punshon, T., Ricachenevsky, F. K., Hindt, M. N., Socha, A. L. & Zuber, H. (2013). *Metallomics*, **5**, 1133–1145.

Pushie, M. J., Pickering, I. J., Korbas, M., Hackett, M. J. & George, G. N. (2014). *Chem. Rev.* **114**, 8499–8541.

Ren, L. Q., Wu, D., Li, Y. H., Chen, W. R. & Liu, H. (2014). *Proc. SPIE*, **8944**, 89440H.

Ryan, C. G., Clayton, E., Griffin, W. L., Sie, S. H. & Cousens, D. R. (1988). *Nucl. Instrum. Methods Phys. Res. B*, **34**, 396–402.

Ryan, C. G., Etschmann, B. E., Vogt, S., Maser, J., Harland, C. L., van Achterbergh, E. & Legnini, D. (2005). *Nucl. Instrum. Methods Phys. Res. B*, **231**, 183–188.

Schamber, F. H. (1977). *X-ray Analysis of Environmental Samples*, edited by T. G. Dzubay, pp. 241–257. Ann Arbor: Ann Arbor Science Publishers.

Schamber, F. H. (2009). *Microsc. Microanal.* **15**, 491–504.

Solé, V. A., Papillon, E., Cotte, M., Walter, P. & Susini, J. (2007). *At. Spectrosc.* **62**, 63–68.

Stamenković, S., Dučić, T., Stamenković, V., Kranz, A. & Andjus, P. R. (2017). *Neuroscience*, **357**, 37–55.

Vekemans, B., Janssens, K., Vincze, L., Adams, F. & van Espen, P. (1994). *X-ray Spectrom.* **23**, 278–285.

Vogt, S. (2003). *J. Phys. IV Fr.* **104**, 635–638.

Vogt, S. & Ralle, M. (2013). *Anal. Bioanal. Chem.* **405**, 1809–1820.

Wang, S., Ward, J., Leyffer, S., Wild, S. M., Jacobsen, C. & Vogt, S. (2014). *J. Synchrotron Rad.* **21**, 568–579.

West, M., Ellis, A. T., Potts, P. J., Streli, C., Vanhoof, C. & Wobrauschek, P. (2014). *J. Anal. At. Spectrom.* **29**, 1516–1563.

Yi, L. T., Liu, Z. G., Wang, K., Chen, M., Peng, S. Q., Zhao, W. G., He, J. L. & Zhao, G. C. (2015). *Nucl. Instrum. Methods Phys. Res. A*, **775**, 12–14.

Zhao, F. J., Moore, K. L., Lombi, E. & Zhu, Y. G. (2014). *Trends Plant Sci.* **19**, 183–192.

Zhao, J. T., Pu, Y. X., Gao, Y. X., Peng, X. M., Li, Y. Y., Xu, X. H., Li, B., Zhu, N. L., Dong, J. C., Wu, G. & Li, Y. F. (2015). *J. Anal. At. Spectrom.* **30**, 1408–1413.