

# An evaluation of methodologies for calibrating Itrax X-ray fluorescence counts with ICP-MS concentration data for discrete sediment samples

Braden R.B. Gregory<sup>a,\*</sup>, R. Timothy Patterson<sup>a</sup>, Eduard G. Reinhardt<sup>b</sup>, Jennifer M. Galloway<sup>a,c,d</sup>, Helen M. Roe<sup>e</sup>

<sup>a</sup> Ottawa-Carleton Geoscience Center, Department of Earth Sciences, Carleton University, 1125 Colonel By Dr., Ottawa, Ontario K1S 5B6, Canada

<sup>b</sup> School of Geography and Earth Sciences, McMaster University, 1280 Main St. W., Hamilton, Ontario K1S 5B6, Canada

<sup>c</sup> Geological Survey of Canada, Calgary 3303 33rd St., Calgary, Alberta T2L 2A7, Canada

<sup>d</sup> Aarhus Institute of Advanced Studies, Aarhus University, Aarhus 8000, Denmark

<sup>e</sup> School of Natural and Built Environment, Queen's University Belfast, Belfast BT7 1NN, United Kingdom

## ARTICLE INFO

Editor: G. Jerome

Keywords:

Itrax

μXRF

X-ray fluorescence

Multivariate Log-ratio

Elemental normalization

Discrete sediment samples

## ABSTRACT

Core-scanning X-ray Fluorescence (XRF-CS) is a well-established technique for rapid (< 30 s/interval) analysis of sediment core geochemistry at sub-mm resolution with substantially less analytical cost compared to methods that rely on physical sub-sampling. Due to issues inherent in analyzing wet sediment of heterogeneous particle size and composition with irregular surface topography using XRF, XRF-CS results are generally considered semi-quantitative. The result of early efforts to calibrate XRF-CS data with conventional geochemical results (e.g. WD- or ED-XRF, ICP-AES, ICP-MS) showed weak correlations for less abundant or poorly detectable elements, however, more recent methods have been proposed to improve accuracy. These methods include: 1) converting XRF-CS results to dry mass concentration; 2) normalizing XRF-CS data to conservative elements (Si, Ca), total counts/second, or X-ray scatter (CIR); and 3) calibration of data using multivariate analysis of elemental log-ratios (MLC). These approaches are not yet widely employed, and require additional testing on a variety of sediment compositions. Recently developed equipment enables analysis of discrete sediment samples, providing > 30 replicate analyses for up to 180 samples in a single XRF-CS run. These replicate measurements allow for rigorous testing of precision and accuracy of XRF-CS data. To determine the ideal method of data transformation to improve XRF-CS calibration to quantitative geochemical concentration, 100 lake sediment-surface samples collected from Harvey Lake, New Brunswick, Canada, were analyzed using Itrax-XRF-CS, and then with ICP-MS analysis after multi-acid digestion. Normalization using the CIR and correction for water content showed strong correlation coefficients (Kendall's  $\tau$ ) for elements with atomic number > 18 and high concentrations in the sediment. Results for lighter elements and those with lower concentrations did not perform well using these calibration methods. The MLC provided the most accurate reproduction of observed ICP-MS trends and strong correlations (R) between predicted and actual geochemical concentrations. Based on these results, CIR-normalized or wet-corrected calibrations are ideal for studies where absolute geochemical values are of lesser importance, and the MLC method is appropriate for studies with large numbers of sediment samples ( $n > 100$ ), or those where absolute concentrations of elements are of greater importance.

## 1. Introduction

The use of high-resolution, non-destructive, X-ray fluorescence core scanners (XRF-CS) has become commonplace in the study of paleoenvironments and paleoclimate (Turner et al., 2010; Giralt et al., 2011; Gregory et al., 2015; Profe et al., 2016; Peros et al., 2017), sediment provenance (Kujau et al., 2010; Hunt et al., 2015; Zielhofer et al., 2017) and soil and sediment contamination (Guyard et al., 2007;

Miller et al., 2015; Rodríguez-Germade et al., 2015; Lintern et al., 2016; Schillereff et al., 2016). X-ray fluorescence core scanners are capable of analyzing sediment cores at up to 0.1 mm resolution at < 30 s per interval, offering a major advantage in speed, cost, and resolution over conventional geochemical analysis (ED- or WD-XRF, ICP-MS, INAA). A consequence of this rapid, high-resolution analysis is loss of geochemical accuracy and precision (Tjallingii et al., 2007; Hennekam and de Lange, 2012; MacLachlan et al., 2015). Traditional XRF analysis uses

\* Corresponding author.

E-mail address: [braden.gregory@carleton.ca](mailto:braden.gregory@carleton.ca) (B.R.B. Gregory).

<https://doi.org/10.1016/j.chemgeo.2019.05.008>

Received 9 October 2018; Received in revised form 21 February 2019; Accepted 4 May 2019

Available online 11 May 2019

0009-2541/© 2019 Elsevier B.V. All rights reserved.

sediment which is ground to homogenous particle size and either fused to glass beads or compressed into a puck with a smooth surface to minimize matrix and specimen effects (International Atomic Energy Association, 1997; Weltje and Tjallingii, 2008). XRF-CS measurements can be influenced by heterogeneities in the sample surface, such as an uneven topography or variation in particle size (specimen effects), or by reabsorption, re-emission and scattering of primary X-radiation by some component of the sediment being analyzed, such as interstitial water content, sediment porosity, or the sediment itself (matrix effects; Croudace et al., 2006; Tjallingii et al., 2007; MacLachlan et al., 2015). Additional issues arise due to matrix effects influencing different elements to different degrees, with lighter elements commonly experiencing more pronounced deviations from actual geochemical values compared to heavier elements (Tjallingii et al., 2007). XRF-CS data is thus considered semi-quantitative, particularly for light elements.

Analysis of sediment using ICP-MS is relatively expensive and often requires large sample volumes for standard commercial analysis, which can restrict the resolution of temporal or spatial studies, and it is time consuming in terms of sampling. Many researchers therefore use XRF-CS data to examine relative changes in elemental concentrations at high resolution in core samples while using conventional geochemical analysis of select samples to accurately determine absolute geochemical values. In environmental studies, XRF-CS is used to identify the timing of impacts, or the spatial and temporal extent of contaminated horizons, which are then verified using conventional geochemical results (Guyard et al., 2007; Miller et al., 2015; Rodríguez-Germade et al., 2015; Lintern et al., 2016; Schillereff et al., 2016). If it were possible to obtain more reliable geochemical values using XRF-CS, it would represent an enormous advancement in the ability of researchers to study environmental systems, contamination, and high-resolution climate events in terms of rapidity and cost-effectiveness.

Since the introduction of the first XRF-CS in 1998, researchers have attempted to calibrate XRF-CS data to total elemental concentrations derived from conventional geochemical analytical methods (Jansen et al., 1998). Early attempts that used linear regression to convert XRF-CS data to absolute concentrations showed substantial scatter about regression lines. These variations in data distribution were likely due to matrix and specimen effects that impacted elemental detection by XRF in a non-linear fashion, making it difficult to correct for their influence (Kido et al., 2006). Several methods for reducing the influence of matrix and specimen effects before calibration or interpretation have been developed. These methods can be broadly categorized into three groups: 1) correction for a variable that cannot be measured directly by XRF-CS, typically water or organic content (Tjallingii et al., 2007; Boyle et al., 2015; Chen et al., 2016); 2) normalization of data to a variable measured by the XRF-CS device, such as a conservative element, X-ray scatter, or total counts per second (Wei et al., 2004; Guyard et al., 2007; Turner et al., 2010; Bouchard et al., 2011; Cuven et al., 2011; Kylander et al., 2011; Löwemark et al., 2011; Berntsson et al., 2014; Chawchai et al., 2016); and 3) data transformation aimed at normalizing or scaling data, such as log-normalized (Weltje and Tjallingii, 2008; Hennekam and de Lange, 2012; Weltje et al., 2015) or median-normalized calibration methods (Lyle and Backman, 2013; Lyle et al., 2014; Shackford et al., 2014). Application of these methods has generally resulted in stronger correlations and reduced scatter in linear regressions, yet they have not yet been widely adopted, and still need to be tested on a variety of substrate types.

Although XRF-CS is designed for analysis of split cores, recent studies have begun to adapt core scanners for other purposes including the analysis of discrete samples. A study by Huang et al., (2016) analyzed discrete sediment standards to determine optimal count times for XRF-CS. Profe et al. (2017) analyzed discrete sub-samples of a 20-m loess-paleosol sequence from Hungary to characterize climatic changes in the region. Gregory et al. (2017) developed the design specifications and recommendations for use of a discrete sample holding device that could accommodate up to 180 sediment samples for analysis using the Itrax-

XRF-CS (Sequential Sample Reservoirs - SSR). Analyzing discrete sediment samples using XRF-CS has the potential to enable rapid characterization of sediment geochemistry. The resultant reduced analytical costs can increase both the scale and resolutions of research projects. Discrete sediment samples are ideal for testing calibration methods as they do not exhibit geochemical variations inherent to sediment cores caused by temporal changes in sedimentation and allow for multiple measurements of each sample. This permits for a more direct comparison to the larger, homogenous samples analyzed using conventional geochemical techniques, and enables evaluation of analytical precision that is difficult to quantify when attempting XRF-CS calibration using core samples.

To compare the accuracy of various existing calibration methodologies on XRF-CS data using discrete sediment samples, we obtained 100 sediment-water interface samples from Harvey Lake, New Brunswick, Canada. Representatives from the Harvey Lake Association, the New Brunswick Alliance of Lake Associations, and the adjacent Village of Harvey Station are interested in assessing the current and past health of Harvey Lake to elucidate how any contamination risks may hinder the lakes' recreation use by locals and tourists visiting the region due to the occurrence of a major cyanobacterial bloom in Harvey Lake during the summer of 2016. The ability to quantify the spatial variability in metal concentrations in lacustrine environments would be of considerable utility in identifying contaminants that pose a health risk, and facilitates detection of contaminations sources. Harvey Lake is a typical lake found in developed regions: a relatively small water body, irregular in shape, with moderate depth (~ 11 m), and varied land use around the periphery (natural, recreational, residential and formerly industrial). It is thus an ideal site for testing calibration methods.

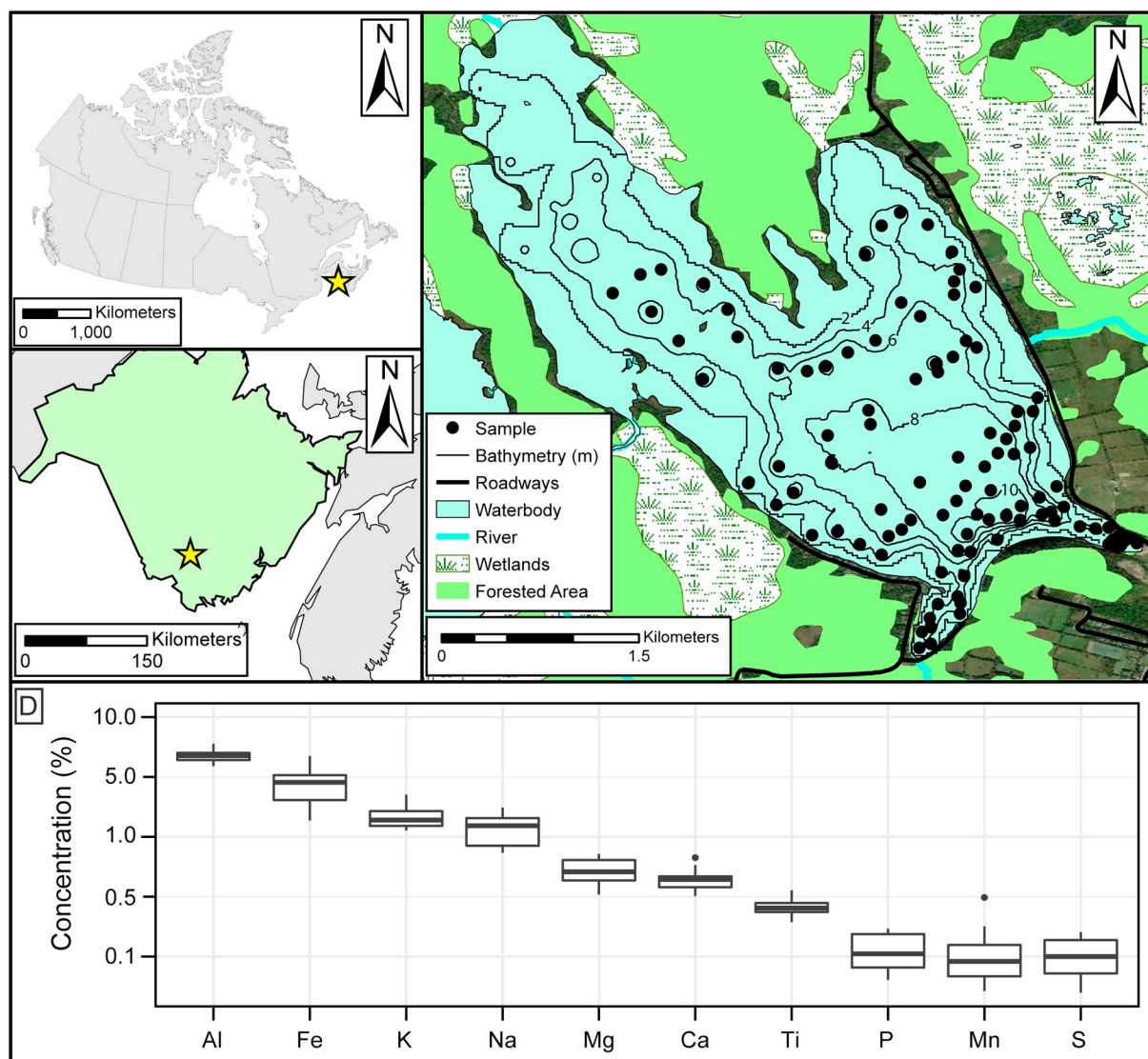
## 2. Methods

### 2.1. Study area and sample collection

Harvey Lake (45°43'45" N, 67°00'25" W) is irregularly shaped, has a surface area of 6.9 km<sup>2</sup>, was non-thermally stratified during late summer sample collection, and has a maximum depth of 11 m (Fig. 1). Water flows into the lake from a stream emptying into Herbert Cove in the south-western part of the lake, from a second stream in the north-western cove, and from a third stream located midway along the eastern shore (Fig. 1). There is a small outflow stream located on the northwestern margin of the lake, which empties through marshland to the adjacent Mud Lake (Fig. 1). Harvey lake is surrounding by wetlands and forest on the western, northern and eastern shores, with a strip of cottages running the length of the eastern shore and in the area surrounding Herbert Cove. The Village of Harvey Station is just south of the lake adjacent to Cherry Mountain, an ~100 m topographic high at the southern end of the lake. The northern portion of Harvey Lake is underlain by Silurian- to Devonian-aged sedimentary rocks including red sandstones, conglomerates and shales (Bottomley, 1984; Payette and Martin, 1986). The southern end of the lake is underlain by Mississippian volcanic rocks including rhyolitic lava flows, and ash-fall/ash-flow tuffs that outcrop on Cherry Mountain (Bottomley, 1984; Payette and Martin, 1986). One hundred samples (HV-01 to HV-100) were collected from the sediment-water interface using an Ekman grab sampler in September 2016. Sample locations were recorded using a Garmin GPSMAP 76CSx GPS. The upper 0.5 cm of sediment collected was sampled into plastic bags using a plastic spoon. Sediment samples were stored in coolers for transport to Carleton University for subsequent analysis.

### 2.2. Laboratory methods

The Harvey Lake samples were divided into sub-samples for laboratory analysis. For ICP-MS analysis, sediment was homogenized with a spatula, then ~10 mL of sediment were subsampled into 50 mL



**Fig. 1.** The location of Harvey Lake within Canada (A), and within New Brunswick (B). (C) Shows the location of surface sediment samples within Harvey Lake as well as vegetation type and inflow/outflow for Harvey Lake. Black lines within the Lake represent bathymetry contours at 2-m depth intervals that were interpolated using depths of sediment samples. (D) Boxplot of elemental concentration of 10 most abundant elements in sediment determined through ICP-MS analysis.

falcon centrifuge tubes and sent to Bureau Veritas, Vancouver, BC, where the sediment was dried, ground to a homogenous powder and digested using a multi-acid wash before analysis (DYAIR, PULCB, and MA250 packages). Bureau Veritas measured 10% of samples in triplicate as well as two quality control standards OREAS45E and OREAS25A-4A, to ensure accurate analysis. All elements studied were within two standard deviations of certified elemental concentrations. The use of multi-acid wash ensured that the results of this analysis best represent whole rock geochemistry measured by XRF-CS (Parsons et al., 2012). A total of 48 samples were analyzed using ICP-MS, of which 28 were unique sediment samples, with 10 being analyzed as blind triplicates for statistical control purposes. Sub-samples selected for ICP-MS evenly covered the area sampled, and included all sediment lithologies observed in samples. Sediment was typically dark-brown, organic-rich gyttja, with only a few samples collected nearer the periphery and the south-east corner of the lake showing slight increases in silt and sand content. The full results of the ICP-MS analysis are available in Gregory et al., 2018a.

Preparation of subsamples for XRF-CS followed Gregory et al. (2017). Sediment was stirred using a spatula to homogenize material, then approximately 20 mL of wet sediment (more for watery samples)

was subsampled into 50 mL falcon tubes, centrifuged at 4000 rpm for 4 min and the supernatant decanted. If time allows, samples can be dried at room temperature to the required consistency to ensure no material is lost during decanting of samples. Each sample was covered to minimize the possibility of contamination, and then left at room temperature to dry until sediment samples reached a paste-like consistency that enables easy loading and reduces the risk of cross contamination (Gregory et al., 2017). The dried samples were transported to the McMaster University Core Scanning Facility, homogenized by stirring with a spatula, loaded into an acrylic SSR ( $1 \times 1 \times 1.5 \text{ cm} = 1.5 \text{ cm}^3$ ), and analyzed using a Cox Analytics Itrax core scanning X-ray fluorescence (Itrax-XRF-CS) device. Some sediment samples did not completely fill sediment reservoirs in the SSR. In these cases, the excess lateral volume was filled with phenolic plastic foam (Oasis Floral Foam) to ensure that the sediment surface did not subside during scanning as this can cause inaccuracies in the XRF-CS analysis. Samples were analyzed using a Mo-anode X-ray tube at 30 kV and 19 mA at 0.2 mm resolution for 15 s per interval. These parameters were selected as they represent the typical settings for Itrax-XRF-CS analysis of lacustrine material. Although count time, scanning resolution, or XRF voltage/amperage may be modified if experimental design



has a specific goal (e.g. detection of rare earth elements [Hennekam et al., 2018]), the results presented here are meant to inform typical XRF-CS researchers as to the best methods for calibration under normal conditions. Subsequent to analysis, the Itrax-XRF-CS results were batch analyzed using RediCore software created by Cox Analytical Solutions. The ICP-MS analysis results in conjunction with known XRF-CS elemental detection limits were used to inform the selection of elements for batch analysis using RediCore. The processed XRF-CS results are available in Gregory et al., 2018b.

Once the non-destructive XRF-CS analysis was complete, sample water content was measured. Material was extracted from the SSR, placed into pre-weighed crucibles, and the wet weight of sediment was measured using a Sartorius BP 121 S scale. Material was placed in a 110 °C oven for 24 h then weighed again to determine the dry weight. The difference between wet and dry weight was considered to be the weight % water content of the sample (Heiri et al., 2001). The organic matter content in each sample was derived using loss on ignition (LOI; after Heiri et al., 2001). Approximately 5 cm<sup>3</sup> of sediment was placed into a weighed crucible, dried in an oven at 110 °C, and then combusted at 550 °C for four hours, with the sample being weighed between each step. The difference between dry weight and weight after combustion was considered to be the weight % of organic matter in the sample. Data regarding water and organic content in sediment is available in Gregory et al., 2018c.

### 2.3. Analytical methods

XRF-CS sediment data was isolated from the extraneous measurements of the SSR acrylic reservoir dividers and floral foam before calibration attempts. Gregory et al. (2017) were able to easily discriminate acrylic dividers from sediment on the basis of the abrupt shifts in the ratio of coherent to incoherent X-ray scatter (CIR), known width of the acrylic compartment walls, total counts per second (kcps), and abundance of the elements (Fe and Ca) in sediment. Subsequent to removal of the non-sediment XRF-CS data points, at least 15 and up to 67 discrete sample measurements remained for each sample depending on whether or not floral foam had been required to fill excess space in sample reservoirs. The median XRF elemental values obtained for each sample were taken as a measure of central tendency because several elements, notably those nearer instrument detection limits, showed non-normal distributions. The relative standard deviation (RSD), expressed as percent relative to them mean value of an elements, was calculated after the formula:

$$\%RSD_{jn} = \frac{\sigma_j}{\mu_j} \times 100 \quad (1)$$

where  $\sigma$  is the standard deviation of element  $j$  for sample  $n$ , and  $\mu$  is the mean of element  $j$  for sample  $n$ . To express the relative standard deviation of elements in the XRF-CS dataset, the mean of the relative standard deviation across all samples was used.

The standard deviation for triplicate measurements of elemental concentration using ICP-MS was calculated using the following equation (after Mccurdy and Garrett, 2016):

$$\sigma_j = \sqrt{\left( \frac{\sum_{i=1}^3 (x_{ijn} - \mu_{jn})^2}{3N} \right)} \quad (2)$$

where  $x$  represents replicate measurement  $i$  for element  $j$  in triplicate set  $n$ ,  $\mu$  represents the mean concentration of element  $j$  in a triplicate set  $n$ , and  $N$  equals the total number of triplicate sets measured. The RSD of elements for ICP-MS data was calculated following Eq. (1);  $\mu_j$  was considered the mean of all triplicate measurements for element  $j$ .

#### 2.3.1. Simple linear calibration

All 28 unique ICP-MS sample measurements were compared to Itrax-XRF values and a linear equation was derived on an element-by-

element basis after the formula:

$$W_{ij} = a_j + I_{ij}B_j \quad (3)$$

where  $W_{ij}$  represent the weight proportion of element  $j$  at interval  $i$ ,  $I_{ij}$  represents the XRF-CS-measured concentration of element  $i$  at interval  $j$ ,  $a_j$  represents the y intercept of element  $j$ , and  $B_j$  the slope of the regression line for element  $j$ . Correlation between ICP-MS and XRF-CS data was examined using Kendall's Tau ( $\tau$ ) as data were non-normally distributed.

#### 2.3.2. Cox Analytics Q-spec calibration

Cox Analytics Qspec software was used for the batch evaluation of XRF-CS results and to calculate absolute concentrations (ppm or weight %) of analyzed XRF-CS data based on results of previous XRF-CS analysis of geological standards. To test the accuracy of this automated calibration in comparison to other recommended methods, the XRF-CS data from the Harvey Lake samples were calibrated using Qspec software. Whereas typical calibration methods (e.g. Jansen et al., 1998; Boyle et al., 2015; Weltje et al., 2015; Chen et al., 2016) require analysis of sediment using conventional geochemical techniques as well as XRF-CS, calibration using Qspec software requires no additional analysis for conversion of XRF-CS results to absolute concentrations as it bases calibration equations on elemental standards previously measured using XRF-CS. The XRF-CS results were converted to absolute concentrations using the Green River Shale standard (USGS SGR-1). The median value of each element in each sample was taken as a measure of central tendency. The results were then compared to the ICP-MS values using linear regression and Kendall's  $\tau$ ; no calibration equation was derived. The Qspec-calibrated XRF-CS results are available in Gregory et al., 2018.

#### 2.3.3. Normalized calibrations

It is common practice to normalize XRF core-scanner data to a conservative element (Wei et al., 2004; Guyard et al., 2007; Turner et al., 2010; Löwemark et al., 2011; Chawchai et al., 2016), or other variables measured by Itrax-XRF-CS, such as total counts per second (kcps; Turner et al., 2010; Bouchard et al., 2011; Cuven et al., 2011; Haenssler et al., 2013), or X-ray scatter (Kylander et al., 2011; Marshall et al., 2011; Berntsson et al., 2014). In theory, such normalization procedures should minimize the influence of matrix and specimen effects that are known to affect XRF-CS measurements, as both the element of interest and the normalization variable should be subject to similar matrix/specimen effects at every measured point. Thus, normalized XRF-CS data should better approximate actual sediment geochemical concentrations measured by conventional geochemical analysis than raw XRF-CS data. Itrax-XRF-CS data were normalized to variables commonly used in XRF-CS research (Si, Ca, kcps and CIR). Silicon and Ca were selected for calibration over other elemental denominators as they are relatively insensitive to redox changes, are abundant in most lacustrine settings, and have been shown in other studies to be ideal denominators for calibration (Weltje and Tjallingii, 2008). Normalized XRF-CS results were compared to ICP-MS data to derive linear equations using a formula similar to Eq. (3) above:

$$W_{ij} = a_j + (I_{ij}/I_{ki})B_j \quad (4)$$

where  $I_{kj}$  represents the value of normalization factor  $k$  at interval  $i$ . Correlation coefficients between normalized XRF-CS data and ICP-MS were calculated using Kendall's  $\tau$ .

#### 2.3.4. Water content correction

Several authors have recognized the problem of water content in a sample resulting in the attenuation of the XRF signal, both due to interstitial water, and a thin film of water that occurs between the sediment surface and the thin mylar film placed on the core before scanning (Kido et al., 2006; Tjallingii et al., 2007; Boyle et al., 2015). Studies have shown that accounting for water content improves the

relationship between the XRF-CS and conventional geochemical data (Kido et al., 2006; Tjallingii et al., 2007; Hennekam and de Lange, 2012; Boyle et al., 2015; Chen et al., 2016). Following Boyle et al. (2015), we derived an equation for determining the water content observed at each point by comparing the water content of the surface samples post XRF-CS-analysis to median values of CIR for all samples. Although Boyle et al. (2015) converted geochemical values to wet percent and then compared these wet values to XRF-CS data to derive calibration equations, we alternatively converted the XRF-CS values to dry mass concentrations for comparison to dry mass ICP-MS data. The procedure was carried out using the equation provided by Boyle et al. (2015).

$$D_{ji} = I_{ji} / (100 / 100 - \omega_i) \quad (5)$$

where  $D_{ji}$  represents the dry mass concentration of a given element  $j$  at interval  $i$  as measured by XRF-CS, and  $\omega_i$  represents the percent dry mass at interval  $i$ . At each interval measured by XRF-CS,  $\omega$  was calculated based on linear regression between median values of CIR for each surface sample and the weight % water content as measured in the lab following the formula:

$$\omega = a_{ij} + (CIR_i)B_{ij} \quad (6)$$

where CIR is the ratio of coherent to incoherent X-ray scatter measured by the Itrax-XRF-CS at interval  $i$ . After raw XRF-CS data was corrected for water content using Formula (5), median values of sediment were taken from the wet-corrected XRF-CS data and compared to ICP-MS data to create elemental calibration equations using Formula (3) above. Kendall's  $\tau$  was used to examine correlation between the two datasets.

### 2.3.5. Multivariate Log-ratio Calibration

Weltje et al. (2015) developed a multivariate method of calibrating XRF-CS data based on the use of log ratios. Alongside recommendations on how to best approach this new Multivariate Log-ratio Calibration (MLC) method, Weltje et al. (2015) created the program ItraXelerate to automatically perform the calibration. The ItraXelerate software uses data from replicate measurements of sediment analyzed using XRF-CS and conventional geochemical methods to calculate absolute geochemical concentrations of select elements. As the XRF-CS data from the Harvey Lake samples represents multiple measurements of the same sample, and as Weltje et al. (2015) recommend at least triplicate scans of several intervals throughout the core, three new XRF-CS datasets were created based on the original XRF-CS data. Three measurements were randomly sampled (with replacement) from each sediment sample output and placed into a new XRF-CS dataset. This process was repeated three times to create triplicate datasets for each discrete sample. As ItraXelerate was designed for use with XRF-CS core data and not discrete samples, each data point within the triplicate datasets was by necessity assigned an arbitrary depth starting at 0 mm for the first point in the dataset and increasing by an interval of 10 mm for each subsequent data point. For example, the triplicate data from HV-01 were assigned the depths 0, 10, and 20 mm, and data from HV-02 were assigned depths of 30, 40 and 50 mm. These three XRF-CS datasets were compared to the ICP-MS dataset containing triplicate analysis of roughly 30% of samples. Each ICP-MS sample was assigned a depth corresponding to the centroid of the depths for the same sediment sample in the XRF-CS datasets (i.e. the ICP-MS depth of HV-02 would be 40 mm). An additional alteration to ICP-MS data was necessitated prior to loading the Harvey Lake ICP-MS data into the ItraXelerate software. The ItraXelerate software will not load geochemical data that is missing elements always observed in conventional XRF analysis as it was originally designed for comparison of ED or WD-XRF to XRF-CS. As ICP-MS does not detect silicon, and because this element is so ubiquitous as to almost always be found in XRF analysis of sediment, ItraXelerate would not load the ICP-MS dataset from Harvey Lake with zero values for SiO<sub>2</sub>. Therefore, an arbitrary value of 0.00001 ppm SiO<sub>2</sub> was substituted into our ICP-MS dataset for all sediment samples analyzed.

ItraXelerate measures the accuracy of calibrations using a goodness-of-fit measurement (R) that uses the mean squared Euclidean distance between predicted and measured compositions (i.e. distance of a given point from the regression line; Weltje et al., 2015). Based on the recommendations of Weltje et al. (2015) the number of elements calibrated should equal roughly one third the number of ICP-MS samples analyzed, rounded down. Nine elements can therefore be calibrated from our 28 ICP-MS samples. As this study is meant to act as a foundation for future paleolimnological work, common elements used as paleolimnological proxies were selected for calibration (Al, Ca, Ti, Mn, Fe, Zn, K) alongside an element of interest for Harvey Lake (As). Copper was also added as, in spite of the fact that it occurred in concentrations much higher than the detection limit of Itrax-XRF-CS, it performed poorly in initial experimental calibrations using the MLC method. ItraXelerate automatically calculated elemental concentrations in proportions that summed to unity. The actual proportion of sediment that the nine selected elements comprised was near to 9.8% on average. As such, XRF-CS concentrations output by ItraXelerate were multiplied by the mean proportion of sediment samples comprised by the nine selected elements (0.098) to facilitate comparison with ICP-MS data.

## 3. Results

Analysis of sediment using ICP-MS detected 44 elements representing a median value of 11.6 weight percentage of sediment (Gregory et al., 2018a). Of this ~11%, Al, Fe, K, Na, Mg, Ca, Ti, P, Mn, and S comprised a median proportion of 99.1% of the measured distribution of elements. Samples were consistently co-dominated by Al (median = 4.75%) and Fe (median = 2.86%), followed by K (median = 1.38%), Na (median = 1.23%), Mg (median = 0.51%) Ca (median = 0.44%), Ti (median = 0.25%), P (median = 0.11%), Mn (median = 0.09%) and S (median = 0.06%; Fig. 1D). Elemental composition of samples remained relatively stable, with similar elemental dominance structure observed in all samples across the basin. Relative standard deviation of elemental concentrations showed a median of 21.9%, although S and Al were notable exceptions with a relative standard deviation of 91.3% and 4.92%, respectively.

Based on comparison of Itrax-XRF-CS detection limits and ICP-MS results, as well as working to maximize the fit between observed and predicted XRF spectra, 28 elements were detected in Harvey Lake sediment using XRF-CS, including Al, Si, P, Cl, Ar, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, Ga, As, Rb, Sr, Y, Zr, Ce, Sm, Tm, Yb, Ta, W, and Pb (Gregory et al., 2018b). The parameters set in preliminary analysis of the sum spectra were applied to each interval analyzed and re-evaluated using QSpec software, which automatically determined ideal XRF parameters for each interval through 10 iterations of a pre-programmed wiggle-fitting procedure. The MSE value for the resultant dataset after plastic sections were removed ranged from 0.92 to 2.42 (median = 1.12,  $n = 4270$ ). Of the 28 elements detected in the sum spectra, 12 were selected for discussion to represent elements with low detectability and high abundance (Al), moderate- to high-detectability and high abundance (K, Ca, Ti, Fe, Mn, Zn), and moderate- to high-detectability and low abundance (Cu, Ni, Zr, As, Ce; Table 1).

### 3.1. Raw XRF-CS

Comparison of raw-XRF-CS data to ICP-MS values ( $n = 28$ ) showed weak to moderate positive Kendall's  $\tau$  ( $0.3 < \tau < 0.6$ ) for most elements (Table 2). Copper and Ce were the only two elements that showed negative correlations. Arsenic showed the strongest correlation observed ( $\tau = 0.69$ ), followed by K, Ca, and Fe ( $\tau_K = 0.59$ ,  $\tau_{Ca} = 0.55$ ,  $\tau_{Fe} = 0.52$ ;  $n = 28$ ). Aluminum had a weak positive correlation ( $\tau = 0.26$ ). Despite moderate correlations, several elements showed considerable deviation from inferred linear relationships (Figs. 2, A1).

**Table 1**

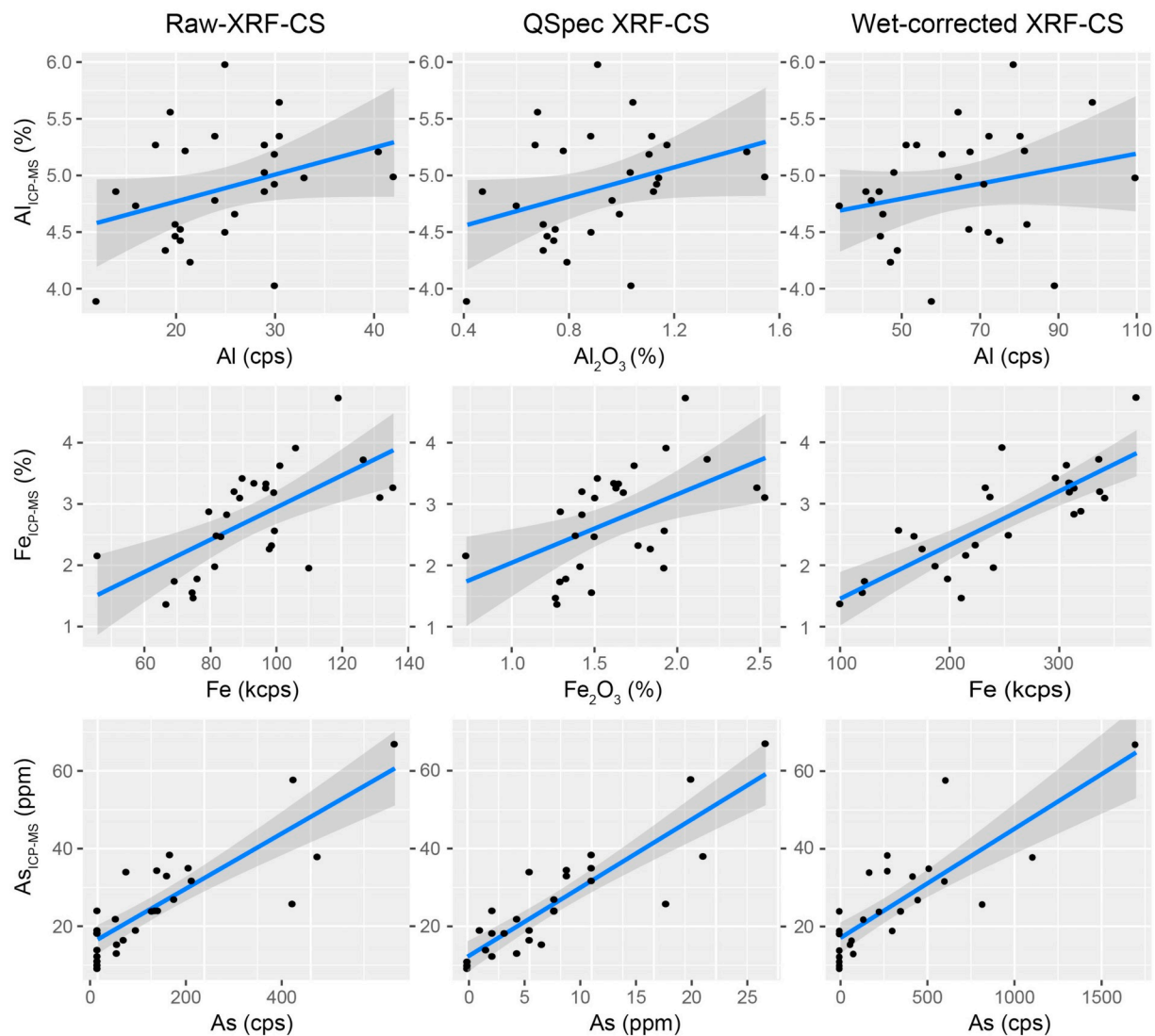
Depiction of the ability of Itrax-XRF-CS to detect elements versus their concentration in sediment based on median ICP-MS concentrations for all elements included in XRF-CS analysis. Elements included for the comparison of different calibration methods are shown in bold. Argon, Si and Cl are excluded from the table as they were not included in the suite of elements analyzed using ICP-MS.

Concentration in sediment	Itrax-XRF-CS Detection Limit		
	High (< 10 ppm)	Moderate (< 100 ppm)	Low (> 100 ppm)
	High (> 1%)	K	Al,
	Moderate (100 ppm – 1%)	Mn	Ca, Ti
	Low (< 100 ppm)	Ni, Cu, Zn, Ga, As, Rb, Sr,	P
		V, Cr, Y, Zr, Ce, Sm, Tm, Yb, Ta, W, Pb	

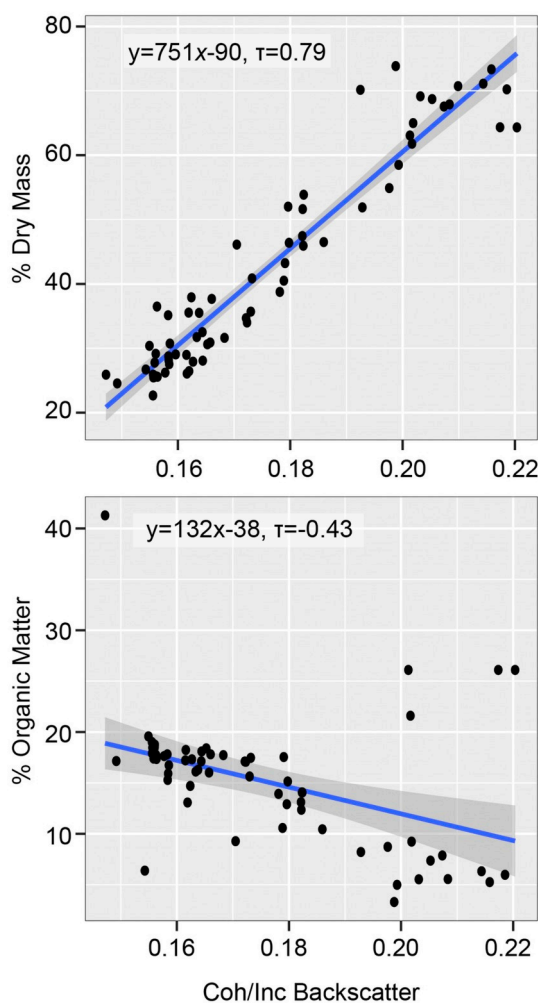
**Table 2**

Kendall's  $\tau$  correlations between raw, Qspec calibrated, and wet-corrected XRF-CS data and ICP-MS results. Mean elemental abundances were calculated using ICP-MS data. Italicized, values indicate an increase in Kendall's  $\tau$  value from raw-XRF-CS data, bold values indicate a decrease in Kendall's  $\tau$  value. Detection limits shown are based on analysis of USGS standard SGR-1.

Element	Detection Limit (ppm)	Mean Elemental Abundance	Atomic Number	Raw XRF-CS	Qspec XRF-CS	Wet-Corrected XRF-CS
Al	7704	4.8 %	13	0.26	<b>0.24</b>	<b>0.25</b>
K	43	1.5 %	19	0.59	<i>0.64</i>	<b>-0.15</b>
Ca	27	0.43 %	20	0.55	<b>0.51</b>	<b>0.39</b>
Ti	13	0.26 %	22	0.06	<b>-0.03</b>	<i>0.44</i>
Mn	7	940 ppm	25	0.35	<b>0.29</b>	<i>0.82</i>
Fe	6	2.7 %	26	0.52	<b>0.44</b>	<i>0.80</i>
Ni	5	26 ppm	28	0.49	<i>0.53</i>	<i>0.62</i>
Cu	5	14 ppm	29	-0.04	<b>-0.50</b>	<i>0.78</i>
Zn	5	110 ppm	30	0.27	<b>0.19</b>	<i>0.86</i>
As	5	25 ppm	33	0.69	<i>0.72</i>	<i>0.81</i>
Zr	27	80 ppm	40	0.31	<b>0.27</b>	<i>0.54</i>
Ce	28	51 ppm	58	-0.40	<b>-0.41</b>	<i>0.04</i>



**Fig. 2.** Linear regressions comparing raw XRF-CS data (left column), data calibrated using Qspec software (middle column) and data corrected for water content after Boyle et al., (2015; right column) to ICP-MS-derived elemental concentrations. Regression lines are represented by blue lines surrounded by darker grey shaded area indicating 95% confidence interval of regression. Only select elements (Al, Fe, As) are compared here, see Supplementary Figs. A1–A3 for regression of all elements included in XRF-CS analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** Comparison of wt% dry mass and wt% organic material measured using LOI to coherent/incoherent X-ray scatter. The median value of the CIR was used for comparison of x-ray scatter in a sample to water and organic content. Regression lines are shown in blue bordered by 95% confidence interval indicated by darker grey area. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### 3.2. Qspec calibration

In comparison to the raw-XRF-CS data, Qspec-corrected data show slightly weaker correlations to ICP-MS results ( $n = 28$ ), although the observed decrease in correlation was often minor (Table 2, Figs. 2, A2; Gregory et al., 2018). For example, Ca decreased from  $\tau = 0.55$  to 0.51 and Al from  $\tau = 0.26$  to 0.24 from the raw-XRF-CS data to Qspec-Calibrated data, respectively. Kendall's  $\tau$  for a select few of the elements (As, Ni, and K) did increase, but this change was minor. Almost one third of all the elements analyzed, including Cu, Ti, and Ce, showed negative correlations.

### 3.3. Water content correction

Comparison of the coherent/incoherent X-ray scatter ratio (CIR) to percent dry mass measured immediately after XRF-CS analysis showed a strong, positive correlation ( $\tau = 0.79$ ,  $n = 66$ ), with little scatter about the regression line (Fig. 3). Although Boyle et al. (2015) used X-ray scatter as a proxy for water content, several authors have suggested that X-ray scatter can be used as a proxy for organic content (Brown et al., 2007; Sáez et al., 2009; Burnett et al., 2011; Giralte et al., 2011). Comparison of LOI-derived organic content to CIR (the inverse of more commonly used Incoherent/Coherent ratio) showed a moderate

strength negative correlation ( $\tau = -0.43$ ,  $n = 61$ ), suggesting a weaker influence of organic matter on X-ray scatter in samples (Fig. 3; Gregory et al., 2018). Organic matter comprises a much smaller proportion of sediment than water content.

Based on the  $a_j$  and  $B_j$  values derived from comparison of percent dry mass to CIR, the raw Itrax-XRF-CS data were adjusted at each measured interval to account for variations in water content. After correction for changes in water content within samples, Kendall's  $\tau$  correlations for all elements became much stronger, with half of the samples showing correlations above  $\tau = 0.60$  ( $n = 28$ ; Table 2, Figs. 2, A3). Only Al, K, and Ce had correlation coefficients below  $\tau = 0.3$ . Potassium exhibited the only negative correlation of all elements analyzed.

### 3.4. Normalization of Itrax-XRF data

Normalization of XRF-CS data to Ca proved moderately successful (Table 3, Figs. 4, A4). On average, Kendall's  $\tau$  for comparison of Ca-normalized XRF-CS data to ICP-MS data ( $n = 28$ ) increased by 0.06 points compared to raw XRF-CS data, with Cu and Zn showing an increase of at least 0.25 points and only two elements (As, and K) showing decreased correlations. Only Ce showed a negative correlation. Normalization of data to Ca also reduced the spread about the regression line for many elements (e.g. Fe; Fig. 4).

Normalization of XRF-CS data to Si showed similar improvements as were observed for normalization to Ca (Table 3, Figs. 4, A5). Only two elements had negative correlations in the Si-normalized data, K and Ca. Many elements showed large improvements in Kendall's  $\tau$ , including improvements of over 0.5 points for Cu, and Ce. Kendall's  $\tau$  for Si-normalized data was, on average, 0.04 points greater than the raw XRF-CS correlations. Six of the elements in the Si-normalized data did not show improved correlations in comparison to raw-XRF-CS data.

In comparison to the raw-XRF-CS data, normalization of data to the total counts per second (kcps) showed an overall decrease in correlation to the observed ICP-MS data ( $n = 28$ ; Table 3, Figs. 4, A6). Cerium, and K showed minor increases in Kendall's  $\tau$ , while all other elements showed a decrease in  $\tau$  by an average of 0.05. Titanium, Cu, and Ce had negative correlations. Several elements showed spread far from regression lines with little discernable trend (Fig. 4, Supp. Fig. A6).

Normalization of elements to the CIR showed improvement in Kendall's  $\tau$  values in comparison to raw XRF-CS (Table 3, Figs. 4, and A7). Some elements showed as much as 0.47 point increase in Kendall's  $\tau$  (mean 0.219 point increase,  $n = 28$ ). All elements showed positive correlations except for Ce ( $\tau = -0.42$ ,  $n = 28$ ). Seven of 13 elements considered had correlation coefficients greater than  $\tau = 0.5$ , with As ( $\tau = 0.83$ ) and Fe ( $\tau = 0.82$ ) showing the highest correlations. There was a reduced spread of data points about the regression line for abundant, well detectable elements and moderate improvement in poorly detectable, yet abundant Al (Fig. 4).

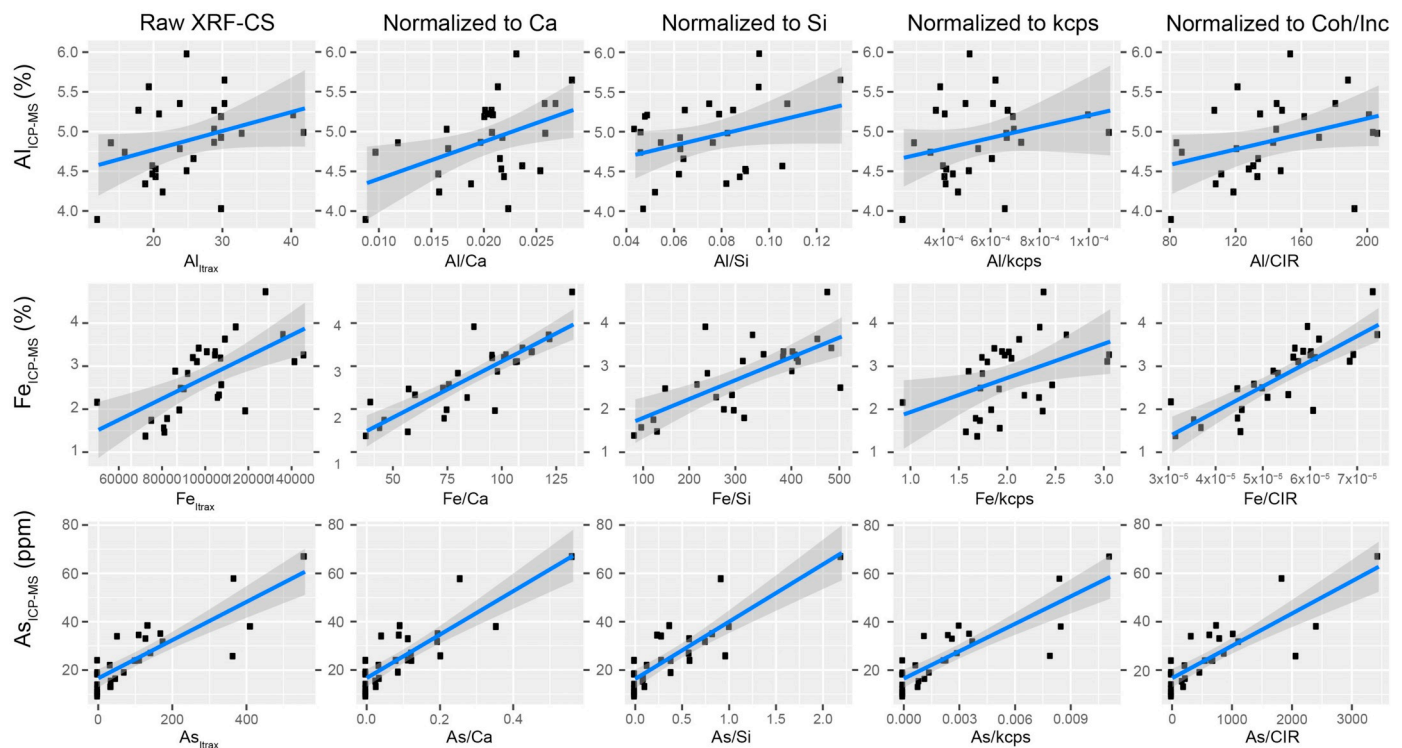
### 3.5. Multivariate Log-ratio Calibration

Correlations for eight out of nine elements using MLC showed moderate to strong positive R values ( $R > 0.6$ ,  $n = 28$ ; Fig. 5). Aluminum, As, Mn, Ti and Zr had moderate R values  $> 0.64$ , and Fe, Ca, and K showed strong correlations with  $R > 0.8$ . Copper showed a weaker correlation with an  $R = 0.47$ . Most values plotted near the regression line for all elements, and all regression lines passed through the origin. There were no negative correlations.

### 3.6. Comparison of calibrated XRF-CS to ICP-MS data

To assess the accuracy of each method, the ICP-MS-derived elemental concentrations were compared to the elemental concentrations for all samples predicted by the three XRF-CS calibration methods with that showed strong, positive correlations. The selected methods were:





**Fig. 4.** From left to right, graphs represent comparison of the ICP-MS data to raw-XRF-CS data, XRF-CS data normalized to Ca, XRF-CS data normalized to Si, XRF-CS data normalized to the total counts per second (kcps) and XRF-CS data normalized to Coherent/Incoherent X-ray scatter (CIR). Regression lines are shown in blue. The dark-grey shaded areas represent the 95% confidence interval of the regression. Only select elements (Al, Fe, As) are compared here, see Figs. A1 and A4–A7 for regression of all elements included in XRF-CS analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

conversion of XRF-CS data to dry mass concentrations after Boyle et al., (2015); CIR-Normalization of XRF-CS data; and the MLC after Weltje et al., (2015). The raw XRF-CS data were included as a baseline for comparison to other methods of calibration. The precision values of the three best Itrax-XRF-CS methods were compared to values observed in a multi-laboratory study by the United States Environmental Protection Agency (USEPA) of the accuracy and precision of ICP-MS analysis of sediment (USEPA, 1998), as well as a more recent analysis of the precision and accuracy of handheld-XRF geochemical analysis of sediment (USEPA, 2007; Table 4).

The raw XRF-CS data showed a moderate degree of overlap with the ICP-MS concentrations of sediment samples (Figs. 6, A8). The concentration of Ni, As, Al and Zr were the only elements of the 12 examined that showed ICP-MS concentrations within 1 SD of the median XRF-CS data for 75% of samples compared. Potassium, Ca, Mn and Ce

showed overlap between ICP-MS and XRF-CS in 50% of sediment samples compared, whereas the remaining 4 elements examined (Cu, Zn, Ti and Fe) showed agreement in < 50% of samples. For some elements, notably Al, the variations in ICP-MS results barely exceeded the observed XRF-CS standard deviation. For Cu and Ti, the XRF-CS results greatly underestimated the variability of ICP-MS results.

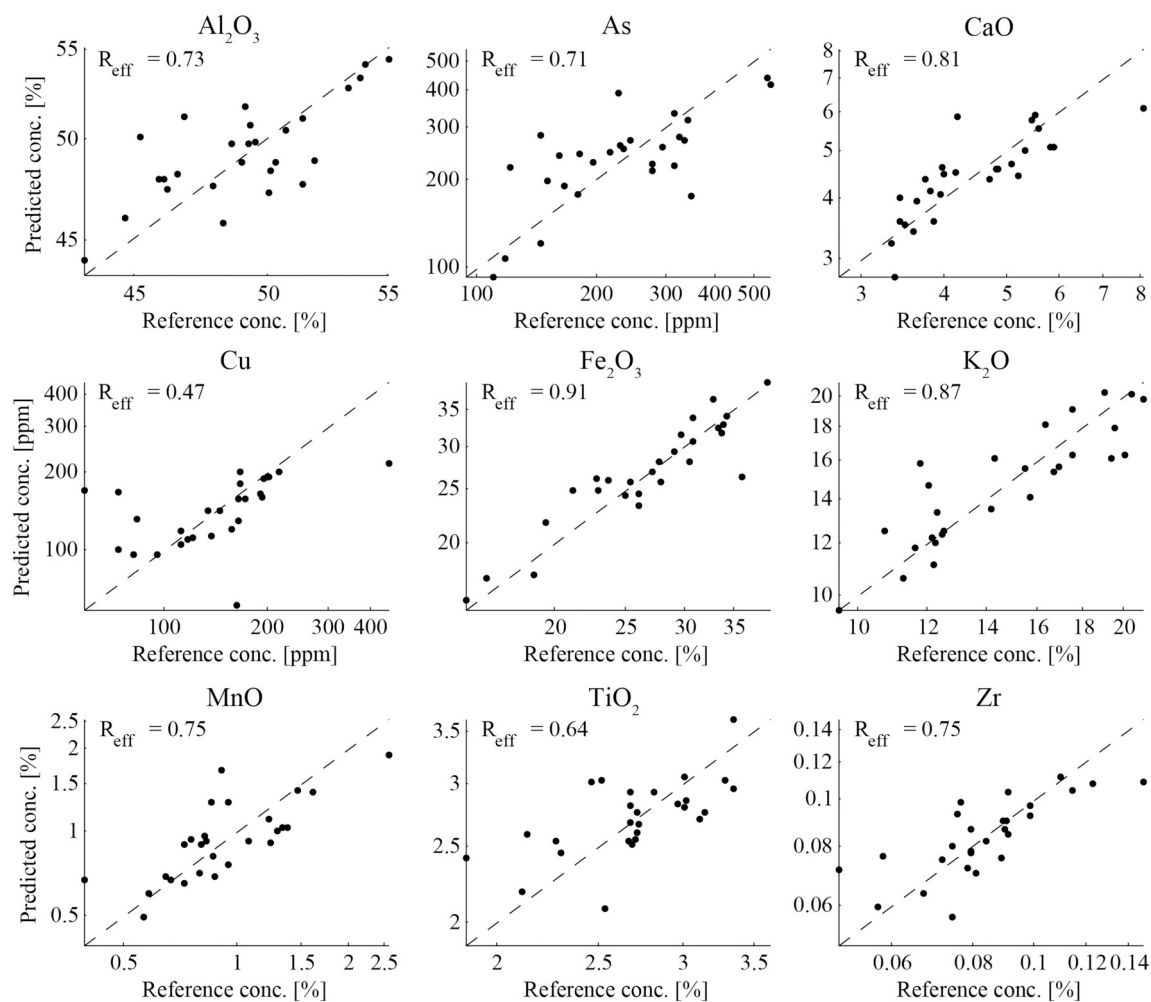
Normalization of the raw XRF-CS data to CIR showed good agreement between the transformed XRF-CS and ICP-MS data. Aluminum, Fe, Cu, Zr and As concentrations showed overlap between ICP-MS and CIR-normalized XRF-CS in 75% of samples compared. Only Zn showed agreement between the two geochemical methods in < 50% of samples; K, Ca, Ti, Mn and Ce all showed agreement between ICP-MS and CIR-normalized XRF-CS within 1 SD in over 50% of samples (Figs. 6, A9). The variability of ICP-MS-measured Al and Cu was lower than standard deviation of CIR-normalized XRF-CS results (e.g. CIR-normalized Cu

**Table 3**

Kendall's  $\tau$  correlations for comparison of ICP-MS data to Ca-, Si-, kcps- and CIR-normalized XRF-CS data. Mean elemental abundances are based on ICP-MS results. Italicized values indicate an increase in Kendall's  $\tau$  values from raw-XRF-CS data, bold values indicate a decrease in Kendall's  $\tau$  values. Detection limits shown are based on analysis of USGS standard SGR-1.

Element	Detection Limit (ppm)	Mean Elemental Abundance	Atomic Number	Raw XRF-CS	Ca-normalized	Si-normalized	kcps-normalized	CIR-normalized
Al	7704	4.8 %	13	0.26	<i>0.27</i>	<b>0.21</b>	<b>0.18</b>	<i>0.35</i>
K	43	1.5 %	19	0.59	<b>0.32</b>	<b>-0.18</b>	0.62	0.65
Ca	27	0.43 %	20	0.55	NA	<b>-0.11</b>	<b>0.54</b>	0.65
Ti	13	0.26 %	22	0.06	<i>0.16</i>	<i>0.30</i>	<b>-0.04</b>	0.25
Mn	7	940 ppm	25	0.35	<i>0.51</i>	<i>0.54</i>	<b>0.25</b>	0.70
Fe	6	2.7 %	26	0.52	<i>0.70</i>	<b>0.48</b>	<b>0.37</b>	0.82
Ni	5	26 ppm	28	0.49	<i>0.52</i>	<i>0.51</i>	<b>0.36</b>	0.70
Cu	5	14 ppm	29	-0.04	<i>0.49</i>	<i>0.61</i>	<b>-0.21</b>	0.43
Zn	5	110 ppm	30	0.27	<i>0.57</i>	<i>0.63</i>	<b>0.14</b>	0.69
As	5	25 ppm	33	0.69	<b>0.65</b>	<b>0.64</b>	0.69	0.83
Zr	27	80 ppm	40	0.31	<i>0.32</i>	<b>0.30</b>	<b>0.29</b>	0.53
Ce	28	51 ppm	58	-0.40	<i>-0.19</i>	<i>0.23</i>	<b>-0.38</b>	<b>-0.42</b>





**Fig. 5.** Comparison of MLC-XRF-CS (predicted) and ICP-MS derived (actual) geochemical concentrations of nine elements selected for the multivariate log-ratio calibration (MLC) using logarithmic graphs. The dashed lines represent linear regression of the data. This plot was generated using ItraXelerate software.

**Table 4**

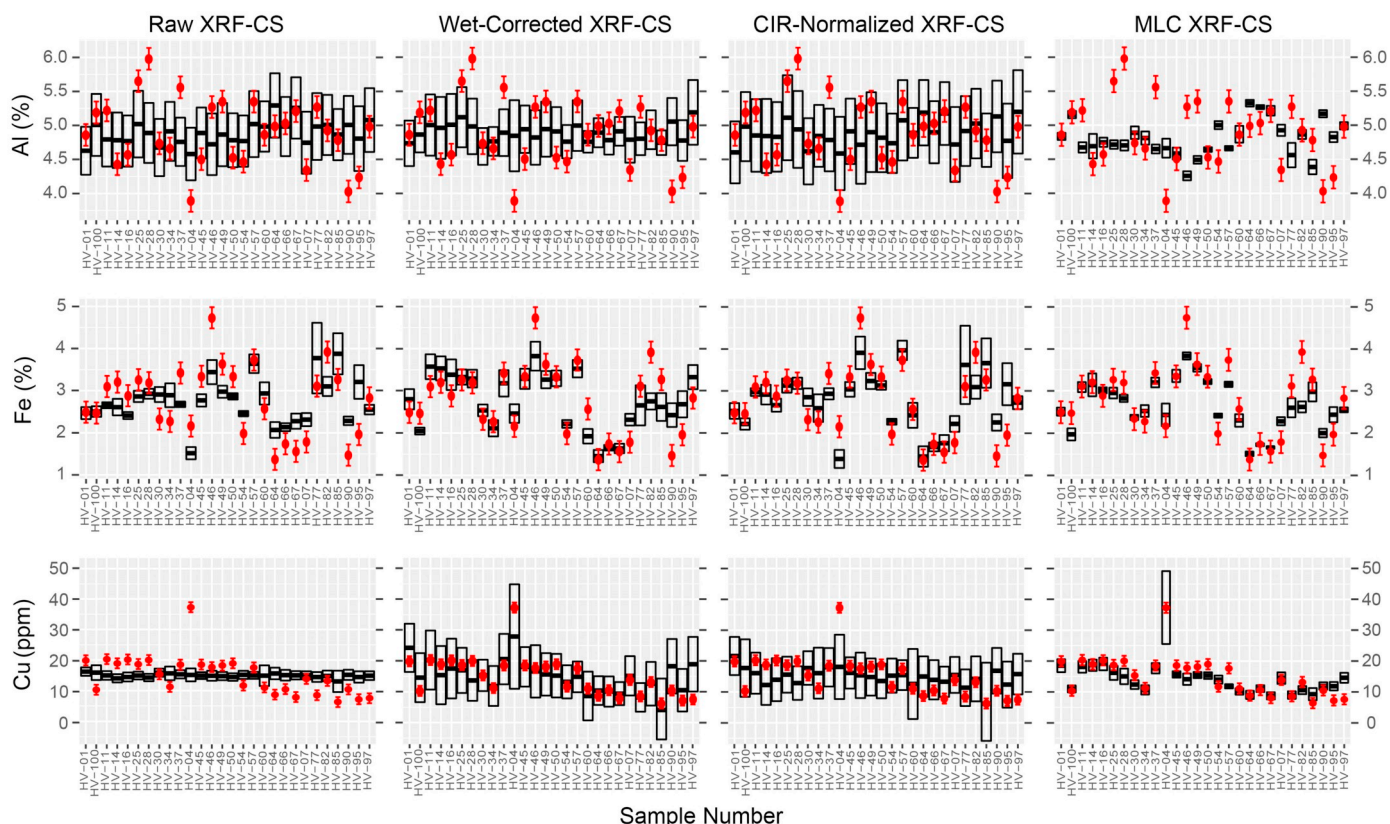
Comparison of relative standard deviation for the best calibration methods and an inter-laboratory study on ICP-MS conducted by the USEPA (USEPA, 1998), and results from a study of handheld-XRF results (USEPA, 2007). The range of data for handheld-XRF results represent the range RSD calculated for analysis of dried, ground sediments (lower range) and unprocessed samples (higher range). Detection limits shown are based on analysis of USGS standard SGR-1.

Element	Detection Limit (ppm)	Itrax Relative Standard Deviation			Harvey Lake ICP-MS (%)	USEPA ICP-MS (%)	USEPA handheld-XRF (%)
		Wet corrected XRF-CS (%)	CIR-normalized XRF-CS (%)	MLC XRF-CS (%)			
Al	7704	6.7	10	1.6	3.6	11–39	–
K	43	2.4	5.8	5.6	5.4	11–62	2.6–18.6
Ca	27	3.4	6.3	5.4	4.2	4.1–27	1.2–17.5
Ti	13	3.1	2.1	5.4	7.5	–	3.74–13.3
Mn	7	12	8.3	12	17.0	11–40	12.3–40.5
Fe	6	8.1	8.8	4.5	9.4	6.7–21	1.6–10.3
Ni	5	7.9	16	–	3.2	9.2–29	18.2–29.8
Cu	5	56	51	12	12.5	9.0–25	7.9–26.4
Zn	5	11	10	–	5.0	14–42	11.1–26.6
Zr	27	8.6	8.5	9.0	15.3	–	5.18–20.2
As	5	36	36	9.8	19.3	12–23	3.8–22.5
Ce	28	1.7	14	–	11.9	–	–

concentrations in Fig. 6). In the case of several elements where extreme highs in elemental concentrations occurred in ICP-MS results, notably Mn, Fe and Ca concentrations in sample HV-46, the high values were underestimated by CIR-normalized XRF-CS data. However, there did not appear to be a systematic over- or underestimation of values in this

dataset. The RSD of CIR-normalized XRF-CS results was generally slightly higher than the relative standard deviation of Harvey Lake ICP-MS results, and comparable to ICP-MS and handheld-XRF measures of precision as determined by the USEPA (Table 4)

Correction of the XRF-CS data for water content (hereafter termed



**Fig. 6.** Comparison of median elemental concentrations predicted by calibrated XRF-CS data  $\pm$  the standard deviation of data (black boxes) and ICP-MS concentrations (red dots). Red crossbars show two relative standard deviations for elements calculated based on triplicate ICP-MS measurements. From left to right, graphs show comparison of ICP-MS data to raw-XRF-CS data, wet-corrected XRF-CS data after Boyle et al., (2015), CIR-normalized XRF-CS data, and data corrected using the MLC method after Weltje et al., (2015). Only select elements (Al, Fe, Cu) are shown here, see Figs. A8–A11 for all the elements included in the calibration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

wet-corrected data) also showed good agreement between ICP-MS and XRF-CS data (Figs. 6, A10). Aluminum, Mn, Cu and As showed agreement between CIR-normalized XRF-CS and ICP-MS data in over 75% of samples compared; Fe, Ti and Zn showed similar concentrations in both datasets for > 50% of samples compared, whereas K, Ni, Ca, and Ce showed agreement below this threshold (Figs. 6, A9). Copper concentrations for wet-corrected XRF-CS data show low precision, with almost the entire range of ICP-MS-measured Cu concentrations falling within 1 SD of XRF-CS results. For some elements, notably Ce, K and to a lesser degree Al, the CIR-normalized XRF-CS data underestimated the variability of ICP-MS results. However, the wet-corrected data successfully approximated most extreme highs in elemental concentrations. The relative standard deviation of wet-corrected XRF-CS data generally remained below 10% with the exception of Cu (56%) and As (36%). The precision of the wet-corrected data within the range of observed precision values for handheld-XRF and ICP-MS, and were similar to the precision of ICP-MS replicate samples from Harvey Lake with the exception of As and Cu (Table 4).

The MLC-XRF-CS data showed good agreement with ICP-MS results, with Mn, Fe, Cu, As and Zr showing overlapping concentrations in 75% of samples compared between the two methods. K, Ca, Ti showed good agreement between MLC-XRF-CS data in > 50% of samples compared, with Al being the only element that did not show a high degree of overlap between ICP-MS and MLC-XRF-CS data (Fig. 6). The MLC-XRF-CS data has a low relative standard deviation, with values lower than the relative standard deviation measured using ICP-MS for all elements except for potassium and copper; precision tended to be greater than the ranges provided by the USEPA for handheld XRF and ICP-MS (Table 4).

## 4. Discussion

The 2nd generation Itrax is capable of analyzing a range of elements with moderate atomic weight ranging from Al to Rb, although the detection limit for Al is generally only reliable at concentrations of > 22,000 ppm (2.2%; Croudace et al., 2006). Attempts to calibrate lighter elements have been met with success when they are present in higher concentrations (Chéron et al., 2016; Huang et al., 2016; Theden-Ringl and Gadd, 2017). It is expected that a successful calibration should have strong positive correlations for most elements with moderate to high atomic weight that are present in higher concentrations, and that those with lower concentration, lower atomic weight, or both, may show weaker positive correlations due to “noise” induced by various matrix and specimen effects interfering with elemental detection. In addition to strong correlations across a range of values, an ideal linear regression of XRF-CS data and conventional geochemical data should have a regression line that passes through the origin, or else through the minimum detection limit of both techniques. Below, the merits of each calibration method are discussed further in light of the aforementioned criteria regarding optimal characteristics of calibrated data.

### 4.1. XRF-CS calibration

#### 4.1.1. Raw XRF-CS calibration

The raw XRF-CS analyses showed a moderate degree of success. There were moderate-strength positive correlations ( $\tau > 0.35$ ) for most elements considered to be well detectable with Itrax XRF-CS using a Mo-tube. Lighter elements, such as Al, had weak positive correlations as expected. Several heavier elements showed weak negative or weak positive correlations in this calibration method in spite of their atomic

weight and the moderate abundance, notably Zn, Cu and Ce. ICP-MS showed that Ce had a mean concentration of 51 ppm, Cu a mean concentration of 14 ppm and Zn a mean concentration of 100 ppm. The mean concentrations of Cu and Ce are close to the lower detection limits of Itrax-XRF-CS for Cu and Ce, which are 6 ppm and 36 ppm, respectively (Table 2). It is possible that the poor correlations are a result of poor elemental detection by Itrax-XRF-CS as the elemental detection limits for Itrax XRF-CS were tested on ideal samples for exposures times much greater than the average XRF-CS runs (100 s exposure time). The lower count time and composition of sediment may have resulted in higher than expected minimum detection limit for several elements. However, Cu and Ce were still fairly well detected by Itrax XRF-CS, with average count rates of 184 cps and 72 cps ( $n = 4270$ ), respectively, and  $< 5\%$  non-detect values (4.1% for Cu and 0.81% for Ce). These values are comparable to As, which had a mean count rate of 129 cps ( $n = 4270$ ), a higher proportion of non-detect values (36%), a low mean concentration (25 ppm) that approached the detection limit of Itrax-XRF-CS (5 ppm), and the highest correlation observed in this dataset ( $\tau = 0.68$ ). Furthermore, Zn had a weak correlation, but a mean concentration much higher than the stated detection limit. It is therefore unlikely that the weak correlations are the result of either imprecise or inaccurate detection of elements due to elemental concentrations near detection limits as similar effects would be observed in As, but not in Zn. The weak correlations could be the result of sampling issues. ICP-MS was measured on larger sample volumes, nearer to  $3\text{ cm}^3$ , whereas even the multiple measurements of XRF-CS likely constituted  $< 1\text{ cm}^3$  ( $0.2\text{ mm} \times 0.2\text{ mm} \times$  critical depth of a given element). However, we would expect that the same sampling bias effect would similarly be present in As, Cu and Ce due to similarly low concentrations in the ICP-MS dataset, which was not observed. Another possible explanation for the lower  $\tau$  correlations in heavier elements is that overlapping X-ray escape lines may have interfered with accurate detection of select elemental concentrations. We reject this hypothesis as both As and Pb are known to have overlapping X-ray escape lines (Bearden, 1967) and the inclusion of lead in our data did not reduce the correlation for As. It is also possible that absorption of the XRF signal by water could produce spurious correlations, but water should affect lighter elements to a greater degree than heavier elements (Kido et al., 2006; Tjallingii et al., 2007), which is not observed in the Harvey Lake data XRF-CS data. In all likelihood, an unpredictable combination of matrix and specimen effects, dilution effects, variations in actual elemental concentration, and limitations of Itrax-XRF-CS itself combined to create unexpected results that deviated from actual geochemical concentrations in difficult to predict ways. A better correlation co-efficient for these select elements may be possible with experimental variation of Itrax-XRF settings during analysis, or if a greater emphasis is placed on these energies during signal processing.

#### 4.1.2. Qspec XRF-CS calibration

Weaker and more often negative correlations were observed in the Qspec-calibrated XRF-CS data in comparison to raw-XRF-CS, although these decreases were often minor. The calibrated standards the Qspec software uses for calibration are based on analysis of solid materials, which would reduce errors associated with matrix/specimen effects. Analysis of wet sediment from Harvey Lake inherently showed deviations from actual geochemistry because of these effects, possibly resulting in negative or weak correlations. The Green River Shale standard selected for calibration may also not have had a similar enough sedimentary composition to that of the Harvey Lake sediment samples to allow for adequate comparison. Improvements may have been possible if additional standards were added based on analysis of powdered sediment, or if sediment standards were analyzed alongside sediment samples and these results used for calibration. It is also possible that variations in exposure time between analysis of standards (100 s) and the Harvey Lake samples (15 s) may also have contributed to the generation of less precise data, negatively influencing correlations.

However, Huang et al. (2016) found that there was no appreciable improvement in correlations between actual and XRF-CS-derived elemental concentrations for most elements of moderate to high atomic number when exposure time was increased from 15 s to 100 s, and decreases were observed in the dataset even for well detectable, abundant elements, such as Fe, Mn, and Ca. Ultimately, the similarity in correlations observed in both the Qspec-XRF-CS and raw-XRF-CS results is likely due to lack of any transformation of the data. All matrix, or specimen present in the raw-XRF-CS dataset were still present in the Qspec-XRF-CS as the data was not transformed before calibration attempted. This suggests that data transformation or normalization, either before or after calibration, is required for XRF-CS results to accurately reproduce geochemical trends in sediment.

#### 4.1.3. Wet-corrected XRF-CS calibration

Boyle et al. (2015) showed that water content is well approximated by the CIR if water content is the most significant influence on mean atomic number in the sediment. Considering that the dominant elements in sediment that comprised  $\sim 91\%$  of the element composition remained fairly stable, varying by only  $\sim 20\%$  of their median values, it is likely that variations in the proportion of water or organic content would exert greater control on the CIR. Comparison of dry mass percentage to CIR showed strong positive correlation, with a  $\tau = 0.79$  ( $n = 66$ ), indicating that the CIR approximated variations in water content within the Harvey Lake sediment samples analyzed (Fig. 3). Elsewhere, researchers have used X-ray scatter as a proxy for organic content since organic material comprises predominantly lighter elements (C, H, O and N) not detected by XRF-CS methods that would decrease the mean atomic number and thus alter the X-ray scatter (Brown et al., 2007; Sáez et al., 2009; Burnett et al., 2011; Giralt et al., 2011). Comparison of wt% organic matter to Coh/Inc. scatter from the Harvey Lake samples did not show a strong correlation ( $\tau = -0.43$ ,  $n = 61$ ; Fig. 3). Other researchers have shown that there is negligible impact on the XRF-CS signal if water content or large sediment grains constitute  $< 25\text{ wt\%}$  of a sample (MacLachlan et al., 2015). It is possible that a similar relationship is true for organic content, although this was not quantified by MacLachlan et al. (2015). However, recent research has shown that a ratio of Inc./Coh X-ray scatter poorly approximates variations in organic content in sediment with higher organic content, but provides a good approximation organic content when present in lower proportions (Chawchai et al., 2016). The substantially larger proportion (mean = 52 wt%,  $n = 66$ ) and the greater variability (range 22.7–73.8%) of water than organic content (mean = 15 wt%,  $n = 61$ ) in samples analyzed in this study likely resulted in water having a greater impact on X-ray scatter than organic matter. More organic-rich sediment needs to be tested to fully understand how well X-ray scatter approximates water or organic content for a variety of sedimentary compositions, or how various types of organic matter may alter sediment porosity and thus both interstitial water and organic content.

Using the CIR to convert raw XRF-CS data to dry weight concentrations (wet-corrected dataset) improved the observed correlations between ICP-MS and XRF-CS results. The wet-corrected data performed poorly for some lighter elements, including a significant decrease in the correlations in the values for K and Ca relative to raw XRF-CS data. The decrease in K and Ca correlations was unexpected as both these elements would be expected to suffer to a greater degree than heavier elements from bias induced by high water content, and thus correction of water content should better improve the accuracy of lighter elements. After correction for water content by Tjallingii et al. (2007) and Boyle et al. (2015) both observed minor increases in Pearson's R value between XRF-CS and ED-XRF of Ca, whereas in MacLachlan et al. (2015) and this study, Ca correlation coefficients decreased slightly after water content correction. Carbonate content in cores examined by Tjallingii et al. (2007) was high at 30–65 wt%, whereas cores examined by Boyle et al. (2015) and MacLachlan et al. (2015) generally ranged from about 5–15 wt% and carbonate content in Harvey Lake was  $< 1$



wt%. Standards analyzed by Bureau Veritas showed that ICP-MS-measured Ca concentrations are accurate to actual concentrations within the 90% confidence interval, indicating ICP-MS accurately reproduced Ca trends in sediment. There are apparent issues with accurately recreating Ca values when correcting XRF-CS for water content when Ca is present in lower proportions. It is unclear why this relationship exist; some research has shown a negative correlation between Ca and water content ( $R^2 = 0.85$ ; Maclachlan et al., 2015). It is possible that water content influences Ca detection by XRF-CS to such a degree that, below a certain concentration of Ca, a simple shift in absolute concentrations to account for water content does not properly account for the issues induced by changes in sediment water content.

The low K correlation may have been the result of variations in Mg, as Mg can absorb XRF emitted from K (Kido et al., 2006), thus inducing noise in the XRF-CS signal not reproduced by ICP-MS. However, it is unclear whether this would have had a substantial effect as ICP-MS results show Mg is not particularly abundant in sediment (mean = 0.52%,  $n = 28$ ), even in comparison to K (mean = 1.47%,  $n = 28$ ). Potassium may also be carried to the surface of the material undergoing analysis by the movement of interstitial water towards the sediment surface where the water forms a thin film between the sediment and the mylar covering used to protect the sediment from dessication during analysis (Tjallingii et al., 2007). This may result in the XRF-CS instrument providing an over-estimation of K concentrations, as well as altering actual variations in K in an unpredictable fashion, resulting in spurious correlations.

Tjallingii et al. (2007) observed that only elements lighter than Si are significantly influenced by dilution of water content, yet in this study and elsewhere (Hennekam and de Lange, 2012; Chen et al., 2016), correction for water content was shown to improve the relationship between the heavier elements analyzed with XRF-CS. In this study, elements heavier than Ca showed substantial improvements in  $\tau$  in comparison to raw XRF-CS data. It is apparent that, although water content may not substantially impact the detection of these heavier elements, correction for water content is still able to improve the correspondence between XRF-CS and results from conventional geochemical analyses.

#### 4.1.4. Normalized XRF-CS Calibration

Normalization of data to Si or Ca showed an overall increase in  $\tau$  correlations by 0.04 and 0.06 for Si and Ca, respectively. Decreases in Kendall's  $\tau$  in comparison to raw XRF-CS data were more frequent Si-normalized dataset than were seen in the Ca-normalized XRF-CS data. Although elemental normalization did increase observed correlations, in most cases the increases were minor, with only Cu and Zn showing an increase in  $\tau > 0.2$ . Both normalization techniques reduced the spread of data points about the regression lines of multiple elements, although this influence was clearly more prevalent for heavier elements, like Fe, than lighter elements, like Al (Fig. 4). Weltje and Tjallingii (2008) found that lighter elements, or those with low abundance, were not suitable as elemental denominators during calibration. In their study, elements lighter than Si were suggested to be unsuitable for elemental denominators in log-ratio calibrations, with Ca being considered the ideal denominator. For the Harvey Lake sediments, Ca barely performed better than Si. This result emphasizes the need for consideration of multiple possible elemental denominators before any calibration attempts are made.

In XRF-CS, normalization to an element can help to reduce the influence of matrix/specimen effects under the assumption that both elements will be affected by said effects to the same degree. Although it is expected that correction for these effects would substantially increase correlations for all elements, the minor improvements observed for the normalization of XRF-CS data to Si and Ca may have been caused by subtle variations in how elements respond to matrix and specimen effects. Absorption of emitted X-radiation occurs predominantly at lower energies, and energy emitted by atoms increases with atomic number

(Potts, 1987), thus resulting in a greater influence of absorption effects on lighter elements. Furthermore, lighter elements have a smaller critical depth and should be more strongly influenced by a small water film forming between mylar covering samples and the sediment surfaces being analyzed (Kido et al., 2006; Tjallingii et al., 2007). Therefore, using a lighter element as the denominator in elemental ratios may induce larger errors into heavy element concentrations. The samples from Harvey Lake show a slight decrease in correlations for As, and Zr. Similar results were observed in Hennekam and de Lange (2012), where only log ratios of elements with similar atomic number (Ca/K and Ca/Ti ratio) showed a strong correlation between XRF-CS data and weak correlation to water content proxies. We observed a smaller increase in the correlation coefficients for XRF-CS data normalized to Si than to Ca. This may be caused by Si suffering from greater water-induced error. Normalization issues may also arise because elemental denominators respond not just to variations in matrix and specimen effects, but to environmental parameters that change over time and across a study area. Normalization to conservative elements, especially lighter elements susceptible to water content variations, is not ideal for calibration of heavier elements. Furthermore, use of elemental ratios with disparate atomic numbers may induce water-related errors, influencing paleoenvironmental interpretations using elemental ratios even in datasets where calibration is not attempted. Finally, application of elemental ratios to calibration attempts may suffer from comparison of normalized elements to un-normalized elements. Elemental ratios may remove some of the variability associated with matrix/specimen effects, but inherently introduces variability associated with environmental parameters influencing the denominator element, creating unpredictable and possibly incorrect calibrations. The success of the calibrations in this paper indicates calibration using elemental ratios may be possible when sediment geochemistry remains stable. However, in samples where major changes in sediment geochemistry occur (e.g. a shift from carbonate to silicate dominated system), calibration attempts using elemental ratios may not be as successful.

Several researchers have normalized XRF-CS data to the total counts per second (kcps) measured by Itrax-XRF-CS (Turner et al., 2010; Bouchard et al., 2011; Cuven et al., 2011) under the assumption that the kcps approximates the overall density of sediment. This effect is regularly observed, and was used here and previously (Gregory et al., 2017) in the analysis of discrete sediment samples to distinguish lighter acrylic-reservoir walls from denser sediment. By normalizing XRF-CS data to variations in density, density-related matrix effects should be minimized and XRF-CS data should better approximate dry-mass concentrations recorded in the ICP-MS data. This is especially true if components of the sample that cannot be measured by XRF-CS, like water and organic content, have a significant impact on overall sediment density. Correlations between kcps-normalized XRF-CS data and ICP-MS for the Harvey Lake samples decreased for almost all elements, and more negative correlations occurred. This may suggest that kcps poorly approximates major changes in sediment density in sediment samples from Harvey Lake. The low variability of kcps (RSD of only 8%) may have resulted from relatively minor variations in the elemental composition of sediment across samples. Iron comprised ~85% of the signal recorded by XRF-CS, and did not substantially change between samples (SD = 2%). The method of densely packing the sediment into reservoirs prior to analysis may have contributed to relatively uniform kcps readings observed. If sediment density and packing remained relatively uniform, kcps may have been approximating variations in other variables, such as minute variations in the distance between the XRF detector and the sediment surface (Rothwell and Croudace, 2015), or sample heterogeneity. If sediment density is not approximated by kcps, then normalization to this variable likely induced additional noise into the dataset independent of geochemical variability, resulting in the decrease in correlations observed for this method.

Boyle et al. (2015) used a ratio of Compton (incoherent) to Rayleigh (coherent) X-ray scatter to correct for water content in sediment as this



ratio is proportional to the mean atomic mass of material analyzed (Duvauchelle et al., 1999). Assuming that use of elemental ratios eliminates matrix effects for a given sample, and that matrix effects are predominantly caused by variations in water content, sediment density, elemental composition of sediment, or other variables controlled by the mean atomic mass of sediment, normalizing elemental ratios to CIR values should improve the observed relation between XRF-CS data and dry-mass ICP-MS concentrations. Normalization to X-ray scatter proved to be one of the best methods to approximate quantitative geochemical results. All elements showed positive correlations, commonly with correlation coefficients > 0.5. However, Ce had moderate-strength negative correlation ( $\tau = -0.42$ ). Cerium was an element that commonly had weak correlations between XRF-CS and ICP-MS data across all calibration methods, suggesting possible issues associated with detection of the element rather than issues with the calibration methods. ICP-MS showed a mean Ce concentration of 51 ppm, higher than the stated detection limit of Itrax-XRF-CS of 36 ppm. Although similar issues are not uniformly observed in elements with concentrations that approach the detection limits of XRF-CS, it seems that lower concentration elements are more susceptible to inconsistencies at lower concentrations. It may be possible to better replicate actual Ce concentrations in sediment by refining XRF-CS parameters during scanning or post processing (e.g. longer exposure times, a different combination of elements or further refining XRF-CS parameters during batch evaluation) or possibly performing additional sample preparation, such as freeze-drying samples or embedding them in epoxy. As this study is meant to test ability of XRF-CS to be calibrated under typical conditions, additional preparations or optimization of XRF peak fitting for lower concentration elements were not undertaken.

Both the CIR-normalized data and the wet-corrected data showed similar increases in the correlations between corrected XRF-CS data and ICP-MS values, however, the CIR-normalized data shows higher correlations for elements from Fe to Zn (Atomic numbers 26–40) as well as Ca and K. The improved  $\tau$  values of the CIR-normalized data indicates that, for heavier elements where water content may have less of an effect, or where water content may induce detection errors (e.g. K), CIR-normalization better approximates actual geochemical values. This may be caused by shifts in sediment composition of a given interval exerting stronger influence on matrix effects than interstitial water content alone, or else from reduced influence of the water film that forms between the sediment surface and the protective mylar film covering the sediment as the critical depth of X-ray increases with increasing atomic number (Tjallingii et al., 2007).

#### 4.1.5. Multivariate Log-ratio Calibration (MLC)

The MLC-XRF-CS data was generated from ItraXelerate software. The output from ItraXelerate assumes that the elements used in calibration sum to unity, however, the nine elements used in our calibration constitute an average of 9.8 wt% of sediment ( $n = 28$ ). To account for this discrepancy, the Harvey Lake geochemical data were adjusted both by the actual proportion of sediment accounted for by the nine elements calibrated, as well as by the mean proportion of sediment accounted for by the nine elements calibrated (Fig. 7). When adjusting MLC-output by the actual proportion of sediment represented by the nine elements selected for calibration, the XRF-CS results fell within 1 SD of median ICP-MS concentrations most of the time for all elements calibrated. However, as the purpose of these methods is to calibrate larger datasets, the mean proportion of sediment represented by the nine elements used for calibration must be used for application to the entire dataset.

Direct comparison between MLC and other methods of calibration presented in this paper were difficult as only nine elements were calibrated using the MLC. The correlation coefficients observed in MLC-calibrated data were comparable with those in most successful calibration techniques tested, notably wet-corrected data or CIR-normalized data. However, the ItraXelerate using a goodness-of-fit R values as the log-ratios better approximate a normal distribution, while other

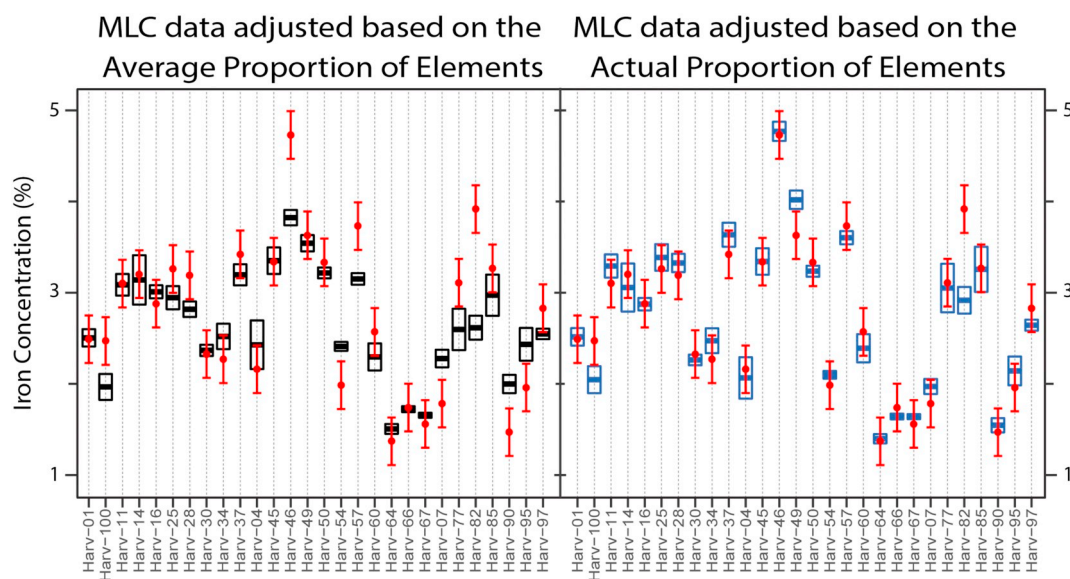
methods discussed in this paper use Kendall's  $\tau$  as raw XRF-CS data had non-normal distributions. The lower correlation coefficients in the MLC dataset than observed for wet-corrected and CIR-normalized data for select elements may also have been due to water content remaining an influence on MLC-XRF-CS data. Previous work has shown that correction for water content before MLC can improve correlations in XRF-CS analysis of sediment cores (Chen et al., 2016). In spite of some lower correlation coefficients, graphs of linear regression showed little spread about the regression line. Copper, for example, had an R value of 0.47, yet showed little spread about the regression line.

Furthermore, the regression line for all elements calibrated using the MLC passed through origin not observed for any element the other calibration methods. This suggests that the MLC method is the soundest method for calibrating XRF-CS data.

#### 4.2. Accuracy and precision of calibration methods

For a calibrated XRF-CS data point to be considered “accurate” the median value of the calibrated dataset had to fall within 1 SD of measured ICP-MS data for a given sediment sample. Comparison of CIR-normalization, wet-corrected data and the MLC elemental concentrations to ICP-MS data showed improved calibration accuracy across all three methods compared to raw XRF-CS data (Fig. 6). Wet-corrected data was accurate for more elements than the CIR-normalized data, however, the wet-corrected data had a greater tendency to underestimate actual variability of elemental concentrations in sediment. The elements that showed less variability in wet-corrected XRF-CS data than in the ICP-MS data were elements that had moderate detectability and abundance with low correlations (K, Ca), or had high detectability, but had low abundance (Ce). For the CIR-normalized data, elements that underestimated variability were those that were poorly detected or had low abundance and weak correlations across all methods (e.g. Ti). Elements with high abundance and poor detectability remained problematic. For example, calibrated Itrax-XRF results for Al exhibited a higher error than the variability than the ICP-MS data (Fig. 6). This suggests that in spite of concentration well above the detection limit of Itrax-XRF, Al data may be too noisy to reliably recreate actual geochemical trends. Both the CIR-normalized and the wet-corrected XRF-CS data underestimated some extreme ICP-MS values, although the wet-corrected data was generally closer to extreme values than the CIR-normalized data. For those elements where regressions were observed to be driven by a few anomalously high values, notably As and Mn, ICP-MS trends were still well replicated for both CIR-normalized and wet-corrected data. These results suggested that these were not spurious correlations in spite of being driven by relatively few data points, and that these calibration methods are applicable to contaminated sites that have skewed distributions that are inherently have extremely high values alongside lower background concentrations. Both CIR-normalization and wet-correction before calibration provided reliable results for elements with moderate- to good-detection by the Itrax-XRF-CS using a Mo-tube, although issues may arise when attempting to calibrate elements with lower detection and abundance.

The MLC substantially out-performed both the CIR-normalized and the wet-corrected data when the MLC results were shifted by the average proportion of nine elements used in the calibration. It is clear that the MLC better approximated elemental trends, even for elements with poorer detection, such as Al. In spite of the low R value for Cu (0.47), MLC was able to accurately predict Cu concentrations. For the MLC-XRF-CS, the deviation about the median was substantially smaller, yet MLC-XRF-CS data still showed greater overlap with ICP-MS than other methods. The higher precision of this method may be due to the artificial creation of triplicate datasets wherein there are only nine replicate measurements compared to > 15 measurements for the other calibration methods. In select cases, the MLC-XRF-CS underestimated extreme values. This was likely caused by shifting MLC results by the average proportion of elements in sediment; the samples that strongly



**Fig. 7.** Comparison of MLC-derived median elemental concentrations  $\pm$  one standard deviation (black and blue boxes) to ICP-MS data (red dots) based on average and actual correction of MLC data. Black boxes represent XRF-CS data transformed to represent the actual proportion of sediment that the nine elements selected for MLC constitute. Blue boxes represent XRF-CS data adjusted to represent the average proportion of the nine elements represented in sediment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

deviate from this mean proportion of elements had reduced accuracy. For example, XRF-CS data overestimated the elemental concentrations for most elements for sample HV-54 (Figs. 6, A8–A11). The proportion of the nine selected elements in this sample was 8.45%, lower than the average proportion of 9.8%. This issue impacted all elements, and correction for the actual proportion of 9-element concentrations in sediment instead of the average brings these anomalous results into line with ICP-MS concentrations (Fig. 7). For most elements analyzed and most samples, the influence of this effect is minimal, and ICP-MS results fell within 1 SD of MLC-XRF-CS results.

Comparison of ICP-MS to MLC-XRF-CS data showed that ICP-MS values fell within 1 SD of MLC-XRF-CS data most of the time for all elements, indicating that it was possible to perform MLC on XRF-CS scans of discrete samples by randomly sampling each sample to produce replicate scans attempted for the first time in this paper. Furthermore, it appears that substitution of Si values into our dataset did not substantially alter results, even though Si might have been more abundant than Al, based on XRF-CS results.

The precision of all three XRF-CS calibration methods was comparable to USEPA ICP-MS inter-laboratory comparison (USEPA, 1998). The recorded range of the RSD in USEPA-ICP-MS methods is quite outdated, however, and modern ICP-MS provides much higher degrees of precision. This partially explains why the RSD of triplicate ICP-MS measurements of Harvey Lake sediment surpassed the highest precision in the USEPA multi-laboratory study for all elements except Cu, As, Pb and Ca. The more recent review of handheld-XRF by the USEPA (USEPA, 2007) showed smaller RSD more comparable to Harvey Lake ICP-MS. Reported RSD in Table 4 for USEPA handheld XRF represent the range observed in samples that have undergone no preparation as well as those that had been dried and ground to homogeneous texture. The RSD of Itrax-XRF-CS calibrated data is more often near the upper limit of handheld-XRF RSD commonly associated with analysis of unprepared samples, as is expected, whereas RSD of Harvey Lake ICP-MS results were commonly nearer the lower limit of handheld-XRF RSD measured on prepared sediment samples. The mean RSD of Harvey Lake ICP-MS results was 9.5% for the 12 elements examined considered in this study, whereas the mean RSD for the three ideal calibration methods were 13%, 15% and 7.2% for the wet-corrected, CIR-normalized and MLC XRF-CS data, respectively. In most cases, XRF-CS RSD was only slightly less precise than the Harvey Lake ICP-MS results.

Major issues were present in the As and Cu data, with RSD > 35% in the CIR-normalized and wet-corrected datasets. Although both ICP-MS and calibrated XRF-CS data show increase in RSD for these elements, possibly driven by a broader range of concentrations across the lake basin than other elements, the magnitude of the RSD increase is much larger for the calibrated XRF-CS data than the ICP-MS data. The low precision of As and Cu may have been caused by errors in the XRF quantification of these elements caused by overlapping escape lines. It is unclear as to the cause of high RSD in Cu values. Regardless, the low precision of these low-concentration elements could likely have been improved through refining the XRF spectral fit, or else through longer exposure times. Although increasing exposure has shown little influence on the ability of XRF-CS to accurately recreate the concentration of elements that are readily detected and present in high concentration, minor improvements have been observed for elements present in lower concentrations (Huang et al., 2016). If the goal of a study is to assess rarer elements, in contaminations studies for example, than longer exposure times may help increase the accuracy and precision of data.

#### 4.3. Recommended method for calibration of surface samples

Analysis of sediments using XRF core scanners offers increases in attainable sampling resolution and speed of analysis, as well as a decrease in analytical costs in exchange for providing semi-quantitative data when uncalibrated. An ideal XRF-CS calibration methodology not only preserves the high-resolution, rapid, cost-effective nature of XRF-CS data, but results in accurate and precise measurement of geochemical concentrations from XRF-CS data. In this study, we observed improved relationships between XRF-CS and ICP-MS data using three methods of post-analysis calibration: normalization of XRF-CS data to CIR; transforming data to dry-weight proportions using methods of water content correction modified from Boyle et al. (2015); and using the Multivariate Log-ratio Calibration method with the ItraXelate software outlined in Weltje et al. (2015). Although all three methods provided precision comparable to USEPA ICP-MS and handheld-XRF, the accuracy of the three calibration methods varied. For our samples, CIR-normalization and correction for water content provided good correlations for most well detectable elements in each analysis that had higher concentrations in sediment. Light elements with higher concentration in sediment (Al) performed poorly for these methods.

However, some studies have been able to reliably estimate trends of lighter elements in sediment using XRF-CS (Chéron et al., 2016; Theden-Ringl and Gadd, 2017), suggesting it may be an issue with abundance or possibly XRF-CS parameters. Heavier elements that were less abundant provided mixed results depending on the method used. For example, Ce was heavy, and well above the detection limit for both methods, but was poorly predicted by the wet-corrected method and well approximated by the CIR-normalized data. If either of these methods were used in, for example, analysis of heavy metal pollutants in sediment, relationships between observed and actual values should be examined carefully before undertaking an interpretation of the results. It is likely that additional processing of the data, or modification of analytical procedures (e.g. longer count times) may improve the precision or accuracy of said elements.

The MLC method is the method that performed the best of all those tested (Weltje et al., 2015). This method showed the most accurate results after adjusting the ItraxExelate output to account for the average proportion of the calibrated elements in sediment as measured using ICP-MS. However, this accuracy comes with an increased temporal and monetary cost, somewhat diminishing the advantages provided by XRF-CS. This diminished advantage is further emphasized by the inherent issues associated with analysis of discrete samples using XRF-CS that remove the advantage of improved resolution, unless the purpose of a study is to better understand the homogeneity of sediment samples themselves. For analysis of relatively few discrete samples using XRF-CS ( $n < 30$ ), the MLC would thus offer no substantial improvement over conventional geochemical methods as most samples would need to be analyzed to calibrate a reasonable number of elements. However, for projects involving a larger numbers of surface samples ( $n > 100$ ), the MLC offers substantial advantages over conventional geochemical analysis. Furthermore, if a study requires more reliable results for analysis of pollutants in a system, the MLC method offers clear advantages over the other XRF-CS calibration methods considered here. If a research project is investigatory in nature, or is interested in sedimentary processes or trends vs measuring pollutants, XRF-CS with CIR-normalization or wet-correction methods allow sufficient data correction to profit from the time/resolution/cost advantages of the instrumentation.

## 5. Conclusions

In this study, 100 surface sediment samples were analyzed with XRF-CS and ICP-MS, and multiple methods of calibrating XRF-CS data were tested to determine which would provide the highest accuracy, highest precision and most cost-effective results. Linear regression of raw-XRF-CS results showed moderate strength correlations for heavier elements with good detection. Calibration using Qspec software designed by Cox Analytics resulted in decreased correlations, possibly due to disparities between geochemical composition of standards and sediment analyzed, or the material analyzed (pressed pellet vs wet sediment). Normalization of XRF-CS data to Si and Ca showed slight improvement in correlation coefficients for most elements, but was outperformed by normalization to the CIR. Of the normalization methods, only normalization to kcps did not show substantial improvement in correlation coefficients. Correction of XRF-CS data for water content based on the ratio of coherent to incoherent X-ray scatter showed substantial improvement on a similar scale as CIR-normalized data, likely due to the correction being based on the same parameter. Of the simple linear calibration methods, CIR-normalization and water-content correction performed best, recreating trends in elements with good detection. These two methods, however, had minor issues when used to recreate observed geochemical trends in elements that were well detected by XRF-CS but had low concentration in sediment, or those that were poorly detected by XRF-CS. The multivariate log-ratio calibration method performed the best of all tested methods after results were adjusted to account for the proportion of elements in a sample. The

MLC reliably recreated trends for elements with lower detection or low concentration, such as Al and Cu. Calibration of XRF-CS data after normalization to CIR or conversion to dry mass concentrations took full advantage of the speed and cost advantages XRF-CS offers, and are ideal for preliminary investigation of study areas, or studies that do not require rigorous quantification of absolute values. Although the MLC method reduced the advantages of XRF-CS by increasing cost and time required for analysis, it greatly outperformed the other methods of calibration in terms of accurately recreating observed ICP-MS trends. Itrax-XRF-CS of discrete samples enables rapid characterization and comparison within or between environmental systems. Although the calibration of data in this study showed precision comparable to other commonly used methods of geochemical determination used by the USEPA, calibration does not yet show the accuracy that may be required in some contamination studies that have to withstand legal or regulatory scrutiny. The results of the calibration experiments presented here indicate that properly calibrated XRF-CS analyses can be used even in cases where regulatory or planning policy is under consideration to provide high resolution and inexpensive reconnaissance or pilot results to inform subsequent sampling protocols, or to provide approximate elemental concentrations that can later be verified with quantitative analysis.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.chemgeo.2019.05.008>.

## Acknowledgements

We gratefully acknowledge funding provided for this project by a Polar Knowledge Canada grant to RTP and JMG (Grant #1516-149), the Geological Survey of Canada Environmental Geoscience Program, Metal Mining Project, Northern Baselines Activity (JMG), and Canadian Foundation of Innovation Infrastructure operating funds from EGR (Grant #2003785). We also acknowledge the contributions of Dr. Sheryl Bartlett, the Harvey Lake Association, and the New Brunswick Alliance of Lake Associations for facilitating the research on Harvey Lake. We thank Roy T. Patterson for contributing his time and a boat during sample collection and to Zacchareus Compton of the University of New Brunswick for providing assistance in the field during collection of samples. Thank you as well to Omid Haeri Ardakani (GSC) and Dr. Elizabeth Patterson for helping to refine the paper for submission. This publication represents NRCan contribution number 20170346.

## References

- Bearden, J.A., 1967. X-ray wavelengths. *Rev. Mod. Phys.* 39, 78–124.
- Berntsson, A., Rosqvist, G.C., Velle, G., 2014. Late-Holocene temperature and precipitation changes in Vindelfjällen, mid-western Swedish Lapland, inferred from chironomid and geochemical data. *The Holocene* 24, 78–92.
- Bottomley, D.J., 1984. Origins of some arseniferous groundwaters in Nova Scotia and New Brunswick, Canada. *J. Hydrol.* 69, 223–257.
- Bouchard, F., Francus, P., Pienitz, R., Laurion, I., 2011. Sedimentology and geochemistry of thermokarst ponds in discontinuous permafrost, subarctic Quebec, Canada. *J. Geophys. Res. Biogeosci.* 116, 1–14.
- Boyle, J.F., Chiverrell, R.C., Schillereff, D., 2015. Approaches to water content correction and calibration for uXRF core scanning: comparing X-ray scattering with simple regression of elemental concentrations. In: Croudace, I.W., Rothwell, R.G. (Eds.), *Micro-XRF Studies of Sediment Cores: Applications of a Non-destructive Tool for Environmental Sciences*. Springer Dordrecht, NL, pp. 373–390.
- Brown, E., Johnson, T., Scholz, C., Cohen, A., King, J., 2007. Abrupt change in tropical African climate linked to the bipolar seesaw over the past 55,000 years. *Geophys. Res. Lett.* 34, 1–5.
- Burnett, A.P., Soreghan, M.J., Scholz, C.A., Brown, E.T., 2011. Tropical East African climate change and its relation to global climate: a record from Lake Tanganyika, Tropical East Africa, over the past 90 + kyr. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 303, 155–167.
- Chawchai, S., Kylander, M.E., Chabangborn, A., Löwemark, L., Wohlfarth, B., 2016. Testing commonly used X-ray fluorescence core scanning-based proxies for organic-rich lake sediments and peat. *Boreas* 45, 180–189.
- Chen, Q., Kissel, C., Govin, A., Liu, Z., Xie, X., 2016. Correction of interstitial water changes in calibration methods applied to XRF core-scanning major elements in long sediment cores: case study from the South China Sea. *Geochemistry, Geophys. Geosystems* 17, 1–10.



- Chéron, S., Etoubeau, J., Bayon, G., Garziglia, S., Boissier, A., 2016. Focus on sulfur count rates along marine sediment cores acquired by XRF core scanner. *X-Ray Spectrom.* 45, 288–298.
- Croudace, I.W., Rindby, A., Rothwell, R.G., 2006. ITRAX: description and evaluation of a new multi-function X-ray core scanner. *Geol. Soc. London* 267, 51–63.
- Cuven, S., Francus, P., Lamoureux, S., 2011. Mid to Late Holocene hydroclimatic and geochemical records from the varved sediments of East Lake, Cape Bounty, Canadian High Arctic. *Quat. Sci. Rev.* 30, 2651–2665.
- Duvauchelle, P., Peix, G., Babot, D., 1999. Effective atomic number in the Rayleigh to Compton scattering ratio. *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms* 155, 221–228.
- Giralt, S., Rico-Herrero, M.T., Vega, J.C., Valero-Garcés, B.L., 2011. Quantitative climate reconstruction linking meteorological, limnological and XRF core scanner datasets: the Lake Sanabria case study, NW Spain. *J. Paleolimnol.* 46, 487–502.
- Gregory, B.R.B., Peros, M., Reinhardt, E.G., Donnelly, J.P., 2015. Middle-late Holocene Caribbean aridity inferred from foraminifera and elemental data in sediment cores from two Cuban lagoons. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 426, 229–241.
- Gregory, B.R.B., Reinhardt, E.G., Macumber, A.L., Nasser, N.A., Patterson, R.T., Kovacs, S.E., Galloway, J.M., 2017. Sequential sample reservoirs for Itrax-XRF analysis of discrete samples. *J. Paleolimnol.* 57, 287–293.
- Gregory, B.R.B., Patterson, R.T., Reinhardt, E.G., Galloway, J.M., Roe, H.M., 2018. Harvey Lake QSpec Calibration. *Mendeley Data*. Version 1.0 1, 0. 10.17632/pf7f7tzt4.1.
- Gregory, B.R.B., Patterson, R.T., Reinhardt, E.G., Galloway, J.M., Roe, H.M., 2018a. Harvey Lake ICP-MS. *Mendeley Data*, Version 1.0. 10.17632/2yj6kpcrg.1.
- Gregory, B.R.B., Patterson, R.T., Reinhardt, E.G., Galloway, J.M., Roe, H.M., 2018b. Harvey Lake Itrax XRF-CS. *Mendeley Data*, Version 1.0. 10.17632/3fpdwctg.1.
- Gregory, B.R.B., Patterson, R.T., Reinhardt, E.G., Galloway, J.M., Roe, H.M., 2018c. Harvey Lake Loss on Ignition. *Mendeley Data*, Version 1.0. 10.17632/bshvrgxfz.1.
- Guyard, H., Chapron, E., St-Onge, G., Anselmetti, F.S., Arnaud, F., Magand, O., Francus, P., Mélières, M.A., 2007. High-altitude varve records of abrupt environmental changes and mining activity over the last 4000 years in the Western French Alps (Lake Bramant, Grandes Rousses Massif). *Quat. Sci. Rev.* 26, 2644–2660.
- Haenssler, E., Nadeau, M.J., Vött, A., Unkel, I., 2013. Natural and human induced environmental changes preserved in a Holocene sediment sequence from the Etoliko Lagoon, Greece: New evidence from geochemical proxies. *Quat. Int.* 308, 89–104.
- Hennekam, R., de Lange, G., 2012. X-ray fluorescence core scanning of wet marine sediments: methods to improve quality and reproducibility of high-resolution paleoenvironmental records. *Limnol. Oceanogr. Methods* 10, 991–1003.
- Hennekam, R., Sweere, T., Tjallingii, R., de Lange, G.J. and Reichert, G.-J. (2018). Trace metal analysis of sediment cores using novel X-ray fluorescence core scanning method. *Quat. Int.*, (In Press).
- Heiri, O., Lotter, A.F., Lemcke, G., 2001. Loss on ignition as a method for estimating organic carbonate content in sediments: reproducibility and comparability of results. *J. Paleolimnol.* 25, 101–110.
- Huang, J., Lowemark, L., Chang, Q., Lin, T.-Y., Chen, H.-F., Song, S., Wei, K., 2016. Choosing optimal exposure times for XRF core-scanning: Suggestions based on the analysis of geological reference materials. *Geochemistry Geophys. Geosystems* 17.
- Hunt, J.E., Croudace, I.W., MacLachlan, S.E., 2015. Use of calibrated Itrax-XRF data in determining turbidite geochemistry and provenance in Agadir Basin, Northwest African Passive Margin. In: Croudace, I.W., Rothwell, R.G. (Eds.), *Micro-XRF Studies of Sediment Cores: Applications of a Non-Destructive Tool for Environmental Sciences*. Springer Dordrecht, NL, pp. 127–146.
- International Atomic Energy Association, 1997. Sampling, Storage and Sample Preparation Procedures for X Ray Fluorescence Analysis of Environmental Materials.
- Jansen, J.H.F., Van Der Gaast, S.J., Koster, B., Vaars, A.J., 1998. CORTEX, a shipboard XRF-scanner for element analyses in split sediment cores. *Mar. Geol.* 151, 143–153.
- Kido, Y., Koshikawa, T., Tada, R., 2006. Rapid and Quantitative Major Element Analysis Method for Wet Fine-Grained Sediments Using an XRF Microscanner. vol. 229. pp. 209–225.
- Kujau, A., Nürnberg, D., Zielhofer, C., Bahr, A., Rohl, U., 2010. Mississippi River discharge over the last ~560,000 years—indications from X-ray fluorescence core-scanning. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 298, 311–318.
- Kylander, M.E., Ampel, L., Wohlfarth, B., Veres, D., 2011. High-resolution X-ray fluorescence core scanning analysis of Les Echets (France) sedimentary sequence: new insights from chemical proxies. *J. Quat. Sci.* 26, 109–117.
- Lintern, A., Leahy, P.J., Hejnis, H., Zawadzki, A., Gadd, P., Jacobsen, G., Deletic, A., McCarthy, D.T., 2016. Identifying heavy metal levels in historical flood water deposits using sediment cores. *Water Res.* 105, 34–46.
- Löwemark, L., Chen, H.-F., Yang, T.-N., 2011. Normalizing XRF-scanner data: a cautionary note on the interpretation of high-resolution records from organic-rich lakes. *J. Asian Earth Sci.* 40, 1250–1256.
- Lyle, M., Backman, J., 2013. Data report: calibration of XRF-estimated CaCO<sub>3</sub> along the Site U1338 splice. *Proc. Integr. Ocean Drill. Progr.* 320, 1–16.
- Lyle, M., Lyle, A.O., Gorgas, T., Holbourn, A., Westerhold, T., Hathorn, E., Kimoto, K., Yamamoto, S., 2014. Data report: raw and normalized elemental data along the Site U1335, U1336, and U1337 splices from X-ray fluorescence scanning. *Proc. Integr. Ocean Drill. Progr.* 320.
- MacLachlan, S.E., Hunt, J.E., Croudace, I.W., 2015. An empirical assessment of variable water content and grain-size on X-ray fluorescence core-scanning measurements of deep sea sediments. In: Croudace, I.W., Rothwell, R.G. (Eds.), *Micro-XRF Studies of Sediment Cores: Applications of a Non-Destructive Tool for Environmental Sciences*. Springer Dordrecht, NL, pp. 173–185.
- Marshall, M.H., Lamb, H.F., Huws, D., Davies, S.J., Bates, R., Bloemendal, J., Boyle, J., Leng, M.J., Umer, M., Bryant, C., 2011. Late Pleistocene and Holocene drought events at Lake Tana, the source of the Blue Nile. *Glob. Planet. Change* 78, 147–161.
- McCurdy, M.W., Garrett, R.G., 2016. Geochemical data quality control for soil, till and lake and stream sediment samples. In: Geological Survey of Canada Open File Report 7944.
- Miller, H., Croudace, I.W., Bull, J.M., Cotterill, C.J., Dix, J.K., Taylor, R.N., 2015. Modern pollution signals in sediments from Windermere, NW England, determined by micro-XRF and lead isotope analysis. In: Croudace, I.W., Rothwell, R.G. (Eds.), *Micro-XRF Studies of Sediment Cores: Applications of a Non-Destructive Tool for Environmental Sciences*. Springer, Dordrecht, NL, pp. 1–18.
- Parsons, M.B., LeBlanc, K.W.G., Hall, G.E.M., Sangster, A.L., Vaive, J.E., Pelchat, P., 2012. Environmental Geochemistry of Tailings, Sediments and Surface Waters Collected from 14 Historical Gold Mining Districts in Nova Scotia. *Geological Survey of Canada Open File* 7150. pp. 1–312.
- Payette, C., Martin, R.F., 1986. The Harvey volcanic suite, New Brunswick. II. Postmagmatic adjustments in the mineralogy and bulk composition of a high-fluorine Rhyolite. *Can. Mineral.* 24, 571–584.
- Peros, M., Collins, S., G'Meiner, A.A., Reinhardt, E.G., Pupo, F.M., 2017. Multistage 8.2 kyr event revealed through high-resolution XRF core scanning of Cuban sinkhole sediments. *Geophys. Res. Lett.* 7374–7381.
- Potts, P.J., 1987. A Handbook of Silicate Rock Analysis. Blackie Glasgow, pp. 622.
- Profe, J., Zolitschka, B., Schirmer, W., Frechen, M., Ohlendorf, C., 2016. Geochemistry unravels MIS 3/2 paleoenvironmental dynamics at the loess–paleosol sequence Schwalbenberg II, Germany. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 459, 537–551.
- Profe, J., Neumann, L., Zolitschka, B., Frechen, M., Rolf, C., Barta, G., 2017. Geochemical record of the loess–paleosol sequence Sütő (Hungary) derived from X-ray fluorescence scanning of discrete samples. *Geophysical Research Abstracts* 19, 9085.
- Rodríguez-Germade, I., Rubio, B., Rey, D., Borrego, J., 2015. Detection and monitoring of REEs and related trace elements with an itrax core scanner in the Ria de huelva (SW Spain). *Water Air Soil Pollut.* 226.
- Rothwell, R.G., Croudace, I.W., 2015. Micro-XRF studies of sediment cores: A perspective on capability and application in the environmental sciences. In: *Micro-XRF studies of sediment cores: Applications of a non-destructive tool for environmental sciences*. Springer Dordrecht, NL, pp. 1–21.
- Sáez, A., Valero-Garcés, B., Giralt, S., 2009. Glacial to Holocene climate changes in the SE Pacific. The Raraku lake sedimentary record (Easter Island, 27 S). *Quat. Sci. Rev.* 28, 2743–2759.
- Schillereff, D.N., Chiverrell, R.C., Macdonald, N., Hooke, J.M., Welsh, K.E., 2016. Quantifying system disturbance and recovery from historical mining-derived metal contamination at Brotherswater, northwest England. *J. Paleolimnol.* 56, 205–221.
- Shackford, J.K., Lyle, M., Wilkens, R., Tian, J., 2014. Data Report: Raw and Normalized Elemental Data along the Site U1335, U1336, and U1337 Splices from X-Ray Fluorescence Scanning.
- Theden-Ringl, F., Gadd, P., 2017. The application of X-ray fluorescence core scanning in multi-element analyses of a stratified archaeological cave deposit at Wee Jasper, Australia. *J. Archaeol. Sci. Reports* 14, 241–251.
- Tjallingii, R., Röhl, U., Kölling, M., Bickert, T., 2007. Influence of the water content on X-ray fluorescence core-scanning measurements in soft marine sediments. *Geochemistry, Geophys. Geosystems* 8 (2), 1–12.
- Turner, J., Davis, S., Langdon, C., Scaife, R., Holmes, N., Leng, M., Mulrooney, G., Cummins, T., 2010. A multiproxy (pollen, stable isotope, chironomid and  $\mu$ XRF) record for the Late Glacial to Holocene transition from Thomastown Bog, Ireland. *J. Quat. Sci.* 6, 514–528.
- US Environmental Protection Agency, 1998. Method 6020A: Inductively Coupled Plasma-Mass Spectrometry. *Revision 1 - US Environmental Protection Agency*. pp. 1–23.
- US Environmental Protection Agency, 2007. Method 6200: Field Portable X-Ray Fluorescence Spectrometry for the Determination of Elemental Concentrations in Soil and Sediment. *Revision 0 - US Environmental Protection Agency*. pp. 1–32.
- Wei, G., Liu, Y., Li, X., Shao, L., Fang, D., 2004. Major and trace element variations of the sediments at ODP Site 1144, South China Sea, during the last 230 ka and their paleoclimate implications. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 212, 331–342.
- Weltje, G.J., Tjallingii, R., 2008. Calibration of XRF core scanners for quantitative geochemical logging of sediment cores: theory and application. *Earth Planet. Sci. Lett.* 423–438.
- Weltje, G.J., Bloemsa, M.R., Tjallingii, R., Heslop, D., Rohl, U., Croudace, I.W., 2015. Prediction of geochemical composition from XRF core scanner data: a new multivariate approach including automatic selection of calibration samples and quantification of uncertainties. In: Croudace, I.W., Rothwell, R.G. (Eds.), *Micro-XRF Studies of Sediment Cores: Applications of a Non-Destructive Tool for Environmental Sciences*. Springer, Dordrecht, NL, pp. 507–534.
- Zielhofer, C., von Schudodoletz, H., Fletcher, W.J., Schneider, B., Dietze, E., Schlegel, M., Schepanski, K., Weninger, B., Mischke, S., Mikdad, A., 2017. Millennial-scale fluctuations in Saharan dust supply across the decline of the African Humid Period. *Quat. Sci. Rev.* 171, 119–135.