Trace element analyses of carbonates using portable and micro-X-ray fluorescence: Performance and optimization of measurement parameters and strategies.

De Winter, Niels; Sinnesael, Matthias; Makarona, Christina; Vansteenberge, Stef; Claeys, Philippe

Published in:
Journal of Analytical Atomic Spectrometry

DOI:
10.1039/C6JA00361C

Publication date:
2017

License:
Unspecified

Document Version:
Accepted author manuscript

JAAS

TECHNICAL NOTE

Cite this: *J. Anal. At. Spectrom.*, 2017, **32**, 1211

Trace element analyses of carbonates using portable and micro-X-ray fluorescence: performance and optimization of measurement parameters and strategies†

Niels J. de Winter, ‡* Matthias Sinnesael,‡ Christina Makarona, Stef Vansteenberge and Philippe Claeys

Variations in elemental abundances in carbonate archives offer a wealth of information that can be used as a proxy for the palaeoenvironment and diagenetic history. The state-of-the-art portable handheld X-ray Fluorescence (pXRF) and laboratory micro X-ray Fluorescence (μXRF) instruments provide a relatively inexpensive, fast and non-destructive way of acquiring these trace element composition data. However, there are well-known issues and limitations regarding the method of spectrum acquisition and the conversion of XRF spectra into quantitative elemental mass fractions. This study offers a guideline for the appropriate use of these XRF techniques for the study of carbonates. Using certified calcium carbonate and dolomite standards, accuracy and reproducibility of a pXRF (Bruker AXS Tracer IV) and a μXRF (Bruker M4 Tornado) device are tested under various measurement conditions. The experimental set-up allowed for the variation of several parameters, including the measurement area, integration time, quantification method and measurement strategy. The effects on the accuracy and reproducibility of the quantified elemental abundance results are examined to assess the optimal performance conditions for both devices for the determination of trace element abundances in natural carbonates. The limits of detection and quantification are evaluated for both instruments for a range of trace elements commonly used as palaeoenvironmental proxies (*e.g.* Sr, Mn and Fe). The quality of the XRF spectra is evaluated using spectral processing software. As a result, two new methods for the determination of optimized parameter combinations are proposed for a range of commonly used elements. The Time of Stable Reproducibility (TSR) is based on optimizing the measurement reproducibility by examining the change of the relative standard deviation per time unit and proposing an integration time threshold for reproducible measurements. The Time of Stable Accuracy (TSA) is based on optimizing the measurement accuracy by studying changes in accuracy as a function of increasing integration time and defining an integration time threshold for accurate measurements. An overview table including minimum integration times by which a reliable measurement is achieved is provided for all analyzed elements and experimental set-ups for this study. However, the methodological approach that is developed here is applicable to other (carbonate) materials as well. A comparison between the two X-ray fluorescence instruments allows the evaluation of their respective advantages and disadvantages. Finally, we recommend optimal measurement strategies and techniques for specific research questions.

Received 29th September 2016
Accepted 21st March 2017

DOI: 10.1039/c6ja00361c

rsc.li/jaas

1. Introduction

Modern X-ray Fluorescence (XRF) based techniques provide an adequate, low-cost and rapid answer to many analytical problems in a wide variety of research disciplines.[1] In the case of energy-dispersive XRF, multiple elements can be measured simultaneously with high sensitivity. The non-destructive aspect of the technique is a major advantage when working with precious and/or rare museum or collection samples (*e.g.* fossils, art pieces and precious artifacts). Laboratory-scale XRF instruments are now widely available and offer a high sample output

Earth System Science, Analytical-, Environmental-, and Geo-Chemistry, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium. E-mail: Niels.de.Winter@vub.be

† Electronic supplementary information (ESI) available: Supplementary data 1: Table showing the certified values of CRM393, CRM512 and ECRM782 standards in μg g^{-1} as well as the weighing factors used for the FP quantification method based on the CRM393 standard. Supplementary data 2: table showing the results of all measurements with XRF and ICP-MS methods used in this study in %. Means, standard deviations and RSDs are given. Supplementary data 3: table showing all mass fractions measured for the three standards using ICP-MS measurements. Average values (μ) and standard deviations (σ) of the measurements are given. See DOI: 10.1039/c6ja00361c

‡ These authors contributed equally to this study.

alternative, which does not require the extensive sample preparation of other trace element analysis techniques (*e.g.* Inductively Coupled Plasma Mass Spectrometry (ICP-MS), Inductively Coupled Plasma Optical Emission Spectrometry (ICP-OES) and Synchrotron Radiation induced X-ray Fluorescence (SR-XRF)). XRF techniques are also competitive in terms of cost per measurement.

A broad range of approaches and instruments is currently available for laboratory-scale XRF. Over the last few decades, micro-X-ray Fluorescence (μXRF) was made possible by the improvement of X-ray optics allowing for a micrometer scale spatial resolution.[2] Small spot sizes allow for more precise sampling strategies in the case of inhomogeneous materials. Portable X-ray Fluorescence (pXRF) devices make *in situ* measurements possible, paving the path for direct field analyses.[3,4] Moreover, an additional advantage of pXRF is that it can be used under controlled lab conditions as a less expensive alternative to exclusively laboratory-based devices.

Variations in elemental abundances in calcium carbonate and dolomite materials are frequently studied, offering a wealth of paleoenvironmental proxy information. They are an interesting substrate for laboratory scale XRF analysis in the field of paleoenvironmental reconstruction. Their applications in earth and environmental sciences are numerous: *e.g.* (1) XRF core scanning yields near-continuous records of elemental intensities and ratios,[5–8] (2) elemental mapping or imaging has the potential to reveal the spatial distribution of elements in a sample,[9] (3) studying the composition of (carbonate or bioapatite) skeletons can reveal seasonal records of paleoclimate proxies,[10–13] (4) determination of variations in trace elements in speleothems to trace back hydrogeochemical processes,[14–16] (5) ancient artefacts can be characterized for archeological purposes[17–20] and many more.

Despite the substantial convenience of using XRF methods on a broad variety of samples, one drawback is the plethora of potential uses and experimental set-ups that can hamper a correct interpretation and comparison of results from different studies. For this reason, profound characterization of the effects influencing the reliability of the final results is essential.[21,22] On the one hand, there are variations in results caused by differences in sample properties such as water contents,[23,24] sample matrix effects,[1] sample (in)homogeneity and different sample geometries[25] as well as surface effects.[26] On the other hand, there are variations in measurement conditions, including different beam spot sizes,[18] bulk *versus* single-particle analysis,[27] different detection limits and mass fractions of desired elements[18] and different instrumental set-ups.[9] These changing parameters are equally important for geological XRF.[6,8,28–32] Specifically, for carbonates, Wheeler[28] carried out an exploratory study on limestones and dolomites while Quye-Sawyer *et al.*[32] evaluated the use of handheld XRF on carbonate field samples. Due to the variety in measurement parameters, it is crucial to define criteria for a reliable measurement that are independent of these parameters.

To determine thresholds for the reliable detection and quantification of XRF results in geological studies, it is common practice to calculate lower limits of detection and quantification based on the error of deconvolution. This error of deconvolution is the error of calculating elemental composition based on element peaks in the spectrum and on solutions of the Sherman equation (on which the fundamental parameter quantification is based[33–36]). Such thresholds include the Lower Limit of Detection (LLD[36]) which is approximated by three standard deviations of deconvolution and the Lower Limit of Quantification (LLQ[38]) in turn approximated by 10 standard deviations of deconvolution. However, the aforementioned range of different applications and samples calls for a method to determine such thresholds in a way that takes into account not only errors related to the deconvolution of the XRF spectra, but also errors that are invoked by differences in the measurement conditions, such as sample geometry, sample preparation and matrix effects. The matrix effect is an important issue in X-ray based analyses. The matrix effect is the combined effect of the sample matrix on the absorption, penetration and fluorescence of X-rays that occurs when travelling through the sample.[1] Because of the importance of the matrix effect on XRF measurements, it is difficult to determine the right measurement conditions for any sample. This is why, in any study, quality tests for different measurement conditions should be conducted specifically for the material studied.

This study explores a new method of determining minimum requirements for a reliable XRF measurement based on total accuracy (*i.e.* the agreement of measured value with certified value) and reproducibility (*i.e.* the closeness of agreement between independent results obtained with the same method, IUPAC[39]) errors rather than solely on errors of deconvolution. Therefore, the effect of actual measurement time and experimental set-up using both μXRF and pXRF on calcium carbonate and dolomite materials is investigated. Certified powdered calcium carbonate and dolomite standards were repeatedly measured under laboratory conditions with a state-of-the-art μXRF device using different set-ups and different measurement times. The same standards were repeatedly measured under laboratory conditions with the pXRF for different measurement times. The reproducibility and accuracy for both the μXRF and pXRF were compared with each other and with High Resolution Inductively Coupled Plasma Mass Spectrometry (HR-ICP-MS) measurements. In doing so, we propose two new criteria (the Time of Stable Reproducibility, TSR, and the Time of Stable Accuracy, TSA) to determine which measurement strategy is optimal in terms of achieving reproducibility and accuracy for both commercially available instruments and for a range of elements. These two proposed ways of evaluation can be applied to other XRF systems, other materials and even other measurement techniques where the measurement time also plays an important role. Finally, as an example, a recommendation of a minimal measurement time for each element of interest in the samples of this study analyzed under laboratory conditions using various μXRF and pXRF based sampling strategies is provided.

2. Materials and methods

2.1 Standards and instrumentation

All measurements were performed on ISO-certified CRM393 (ECRM 752-1), CRM512 and ECRM 782-1 powder graded ($D < 75$ μm, 200 mesh) standards (Bureau of Analyzed Samples Ltd.,

UK). The CRM393 is a standard for limestone, while CRM512 and ECRM782 are dolomite standards. Mass fractions of ten different elements (K, Ca, Ti, Cr, Mn, Fe, Ni, Cu, Zn and Sr) were determined using a Bruker Tracer IV Hand Held portable XRF device (HHpXRF, hereafter: pXRF). The pXRF is equipped with a 2 W Rh anode X-ray tube and a 10 mm^2 Silicon Drift Detector (SDD) with a resolution of 145 eV (Mn-Kα). The X-ray beam was focused on a 6 mm by 8 mm integrated area using a Pd collimator. X-ray spectra from the pXRF were deconvoluted and quantified using the standard factory "Soil Fundamental Parameters" method. The fundamental parameters (FP) method makes use of the theoretical relationship between X-ray fluorescence and material composition as determined by Sherman.[33] The factory-calibrated quantification method of the pXRF uses this fundamental principle with a correction based on a soil standard.

Mass fractions of 21 elements (Cl, Br, Na, Mg, Al, Si, P, S, K, Ca, Ti, Cr, Mn, Fe, Ni, Cu, Zn, Rb, Sr, Ba and Pb) were determined using the Bruker M4 Tornado micro XRF (hereafter: μXRF) under near-vacuum conditions (20 mbar). The μXRF used a 30 W Rh anode metal-ceramic X-ray tube and a 30 mm^2 SSD with a resolution of 145 eV (Mn-Kα). The X-ray beam was focused by a poly-capillary lens on a spot with a diameter of 25 μm (Mo-Kα). X-ray spectra of the μXRF were deconvoluted and quantified with Bruker Esprit software using FP quantification calibrated with one-standard calibration based on the CRM393 limestone standard. Calibration of the quantification procedure was performed by determining the offset of measured values using a long (1800 s) integration time and certified values of the standard and adjusting the weighing factors in the Esprit software accordingly (see ESI 1 and 2†).

Mass fractions of 19 elements (Na, Mg, Al, Si, P, S, K, Ca, Ti, Cr, Mn, Fe, Ni, Cu, Zn, Rb, Sr, Ba and Pb) in all three standards were independently determined by High Resolution Inductively Coupled Plasma Mass Spectrometry (HR-ICP-MS ELEMENT2, Thermo Finnigan, Bremen, Germany). Duplicate samples of ~10 mg were digested in 2 mL 16 M ultrapure nitric acid (HNO_3) on an 80 °C hotplate until completely dissolved, after which the solution was left to evaporate. The residue was redissolved in 100 mL 2% HNO_3, resulting in a total dilution factor of 10 000, which was necessary for mass fractions of the elements of interest to be detectable (ng g^{-1} range) by the HR-ICP-MS. Each sample was measured twice and HR-ICP-MS results were corrected using a linear regression with 5 in-house standard solutions. Drift of the machine was corrected by spiking each sample with 100 μg g^{-1} indium-solution. For all elements, except Na and Cu, the difference in result between duplicate measurements (reproducibility) was found to be on average 3.5% relative to the measured value.

2.2 Measurement strategies

To test the effect of integration time on the reproducibility of the XRF results, measurements were carried out on a flattened surface of all standard powders using a range of integration times. A fixed amount of powder was placed in cylindrical plastic sample holders with a diameter of 13.7 mm and a depth of 7 mm (~1 cm^3). The powder was manually pressed to ensure homogeneous packing. Examination of the spatial variation in the size of the Rh Compton peaks was used to ensure that packing was homogeneous.[2] The means and standard deviations of reproducibility were calculated based on 30 repeated measurements for each integration time. To facilitate direct comparison between different measurement strategies and elements, the Relative Standard Deviation (RSD, or coefficient of variation; $\mathrm{RSD} = \frac{\sigma}{\mu}$ in which σ represents the SD and μ the average value of the measurement) was calculated. The RSD is a unit-less measure for the variation within repeated measurements and can therefore be used to compare the precision of measurements whose results are not in the same order of magnitude. It is used in this study to directly compare results from pXRF and μXRF and between different elements. The ARTAX spectral analysis software (Bruker, Germany) was used to extract signal-to-noise ratios and RSDs of the spectral deconvolution method (RSD of deconvolution) from each spectrum measured by pXRF and μXRF. The effect of different strategies of measuring powdered samples in the μXRF was tested by comparing the result of these strategies amongst themselves and with pXRF measurements. Samples were not prepared as pressed powder pellets (as is common for XRF measurements) to allow sample retrieval after measurement without contamination by pellet binders or change of matrix, retaining the non-destructive aspect of the technique. pXRF measurements were repeated on the same area with integration times of 3 s, 10 s, 30 s, 60 s, 90 s and 120 s. The following four measurement strategies for μXRF measurements were compared:

(1) Mapping strategy: an 8.5 mm × 8.5 mm (72.25 mm^2) area of the flattened surface of CRM393 standard powder was mapped at 25 μm (spot size) spatial resolution (340 × 340 points), and the XRF sum spectra of maps were quantified. This procedure was repeated for total integration times of 300 s, 600 s and 1800 s. Shorter integration times were not possible for the given surface due to limitations of the movement speed of the μXRF's XYZ stage.

(2) Single spot strategy: a single 25 μm spot on the CRM 393 powder surface was measured with varying integration times of 3 s, 10 s, 30 s, 60 s, 120 s, 300 s, 600 s, and 1800 s.

(3) Multiple spot strategy: to measure the effect of sample surface heterogeneity, measurements were executed on 30 different 25 μm spots on the powder surface of all three standards. This procedure was repeated for a range of integration times (3 s, 10 s, 30 s, 60 s, 120 s, 300 s, 600 s, and 1800 s). The locations of the multiple spots were kept constant between different integration times.

(4) Average spot strategy: to simulate the strategy of measuring 10 different spots on a powdered surface and averaging the results, the measurements obtained from 10 spots, which were sub-sampled randomly from the multiple spot strategy described above, were averaged for each integration time.

For the comparison of the results of the different set-ups the elements Ca, Fe, Sr and Mn were chosen, because they are elements that are often measured, cover a wide range of mass

fractions in the standards, and are therefore fit to illustrate the limitations of the different measurement strategies.

3. Results

3.1 Elemental mass fractions of the standards

Fig. 1 illustrates the changes in the quantified mass fractions of Sr, Mn, Fe and Ca with increasing integration times, measured by μXRF using the single spot strategy on the CRM393 standard. Results for all elements and all measurement strategies are given in ESI 2.† The mass fractions determined by HR-ICP-MS are shown as well and fall within one SD of the certified values of the CRM393 standard, showing that HR-ICP-MS measurements successfully reproduce certified values of the standards (see also ESI data 3†). Mass fractions measured by μXRF describe an asymptotic evolution towards a stable value (close to the standard certified mass fraction) with increasing measurement time. Variation within 30 repeated measurements decreases with longer integration times (>60 s) for all four elements. The spread reached in the μXRF measurements after 60 seconds is lower than the error on the certified values and the error on the HR-ICP-MS measurements.

The average mass fractions measured using short integration times vary significantly from those obtained using long integration times. For example, Ca mass fractions of all 30 measurements with a 3 second integration time are significantly higher than all 30 samples measured with a 300 s integration time or longer. The μXRF measurements taken with short integration times (<60 s) also often result in values outside the statistical range of the certified value. For Sr and Fe, the average value calculated from 30 measurements with less than 60 s integration-time is more than one SD off the certified value. Mass fractions of elements measured using a long integration time (>60 s) are closer to certified mass fractions of the standard. The integration time for which a statistically acceptable value (within the error of the certified values) is reached differs between different elements.

3.2 Deconvolution and reproducibility error

An overview of the evolution of RSDs of deconvolution and RSDs of reproducibility through increasing integration time for measurements using the multiple spot strategy on the CRM393 standard is shown for all elements in Fig. 2. The multiple spot strategy was applied here because the precisions include errors caused by heterogeneity of the pressed powder surface. Therefore, the multiple spot strategy simulates a more realistic approximation of the precisions reached by measuring a single,

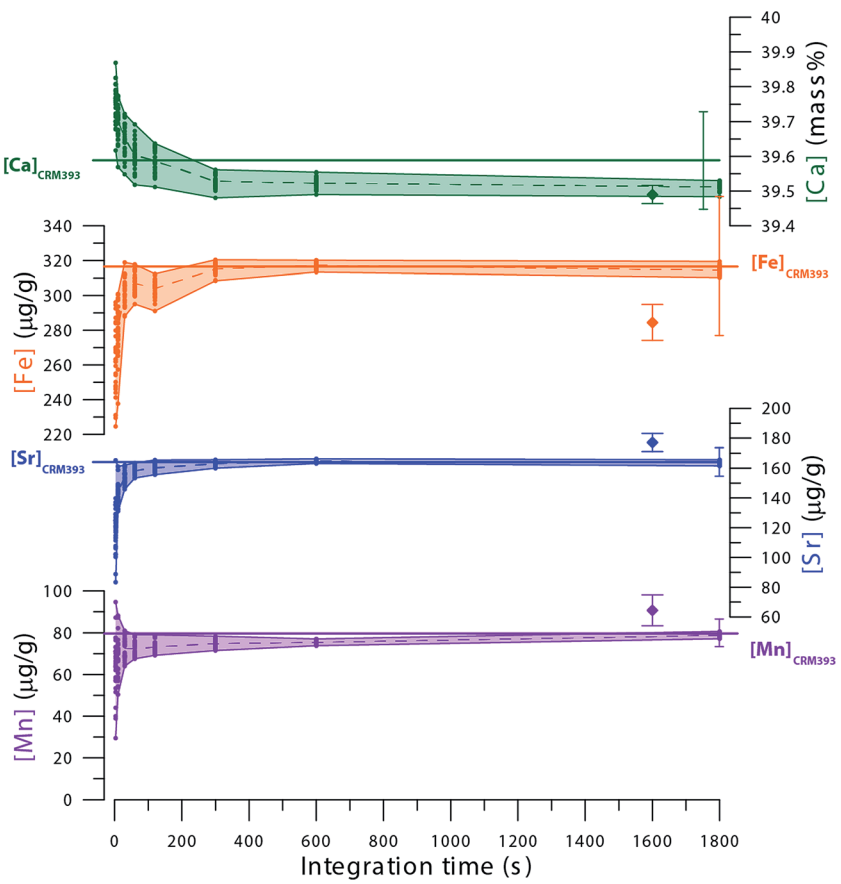

Fig. 1 Change in the measured mass fraction of Ca, Fe, Sr and Mn with increasing integration times using the same spot method on μXRF. Horizontal lines are certified values and narrow error bars are certified errors (1σ) on the standard. Shaded regions indicate the spread of the μXRF measurements. Diamonds and error bars on the right side of the graph show the results of HR-ICP-MS measurements.

random spot on the sample surface using the μXRF. It is evident that RSD values of reproducibility vary largely between different elements, with a difference of 4 orders of magnitude between the most reproducible element (Ca, RSD of 0.03) and the least reproducible element (Cl, RSD > 600). Fig. 2 also shows that for most elements, the RSDs for both deconvolution and reproducibility decrease asymptotically with increasing integration time. Elements not following this asymptotic trend include Cl, K, Ca, Ti, Cr, Ni and Rb. Except for Ca, all these elements have very high reproducibility RSD values (>2) and their deconvolution RSDs are generally higher than those of reproducibility. For Cl, Ti, Cr and Rb, RSDs of deconvolution increase with longer integration times, showing an opposite trend compared to other elements. Elements with asymptotically decreasing RSD values have reproducibility RSDs that are higher than the deconvolution RSDs. For Na, Si, S, Ni and Ba reproducibility RSDs are always above 1. Other elements, such as Mg, Cu and Pb, have RSDs greater than 1 for short integration times but RSDs decrease to values smaller than 1 over longer integration times. Br, Al, P, Mn, Fe, Zn and Sr on the other hand have RSDs that never exceed 1 even for short integration times. For some elements (*e.g.* Mg, P and Sr), the difference between deconvolution RSD and reproducibility RSD is very small, while other elements (*e.g.* Si, Ca and Fe) show relatively large differences between their RSD of deconvolution and their actual reproducibility errors.

3.3 Comparison of measurement strategies

Fig. 3 shows the evolution of RSDs of reproducibility for Ca, Fe, Sr and Mn in the CRM393 standard for each of the four different measurement strategies on the μXRF and for the measurements on the pXRF. The highest RSD, and therefore the lowest precision, is found for the multiple spot and pXRF measurements. Even for long integration times (>60 s), the pXRF retains a low reproducibility, while the RSDs of the μXRF average spot, single spot and map measurements decrease to lower values with longer integration times. μXRF measurements taken with the multiple spot strategy also have high RSD values even for long integration times (see also Fig. 2). As mentioned before, the single spot strategy has lower RSDs of reproducibility than the multiple spot strategy but is less a realistic approximation of measuring one random spot on a sample surface because it does not take into account changes in surface properties across the sample. The average spot strategy performs better (lower RSDs) for Ca, Fe and Mn and yields one of the lowest RSD values

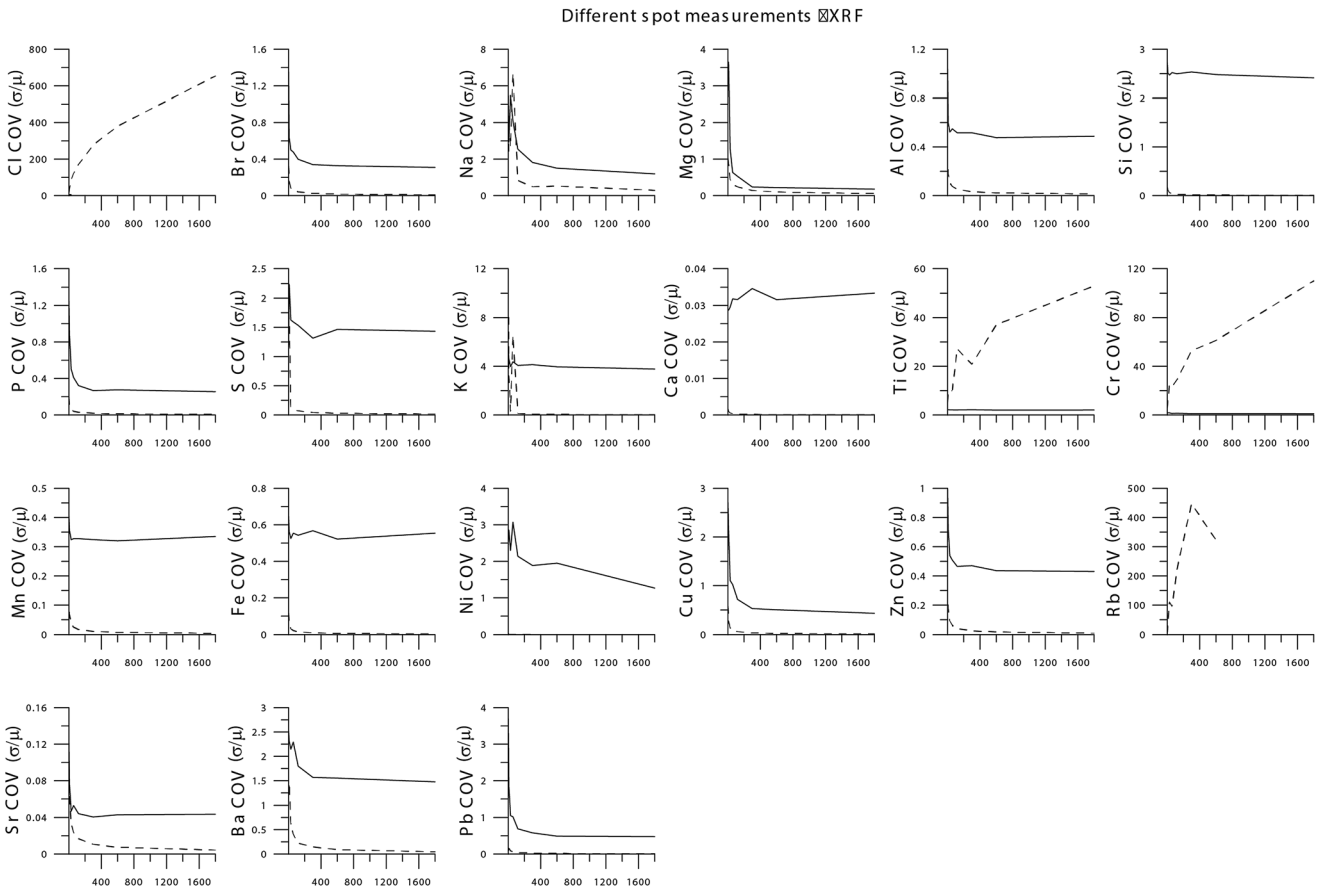

Fig. 2 Overview of the changes in RSD with increasing integration time for all elements using the different spot measurement strategy on μXRF. Solid lines are RSDs of reproducibility of the mass fractions and dashed lines are RSDs of deconvolution errors. Errors of deconvolution are generally lower than errors of reproducibility and decrease with longer integration times. Note that the errors of deconvolution are higher than the reproducibility errors and increase with the measurement time for elements with high RDSs, such as Cl, Ti and Cr.

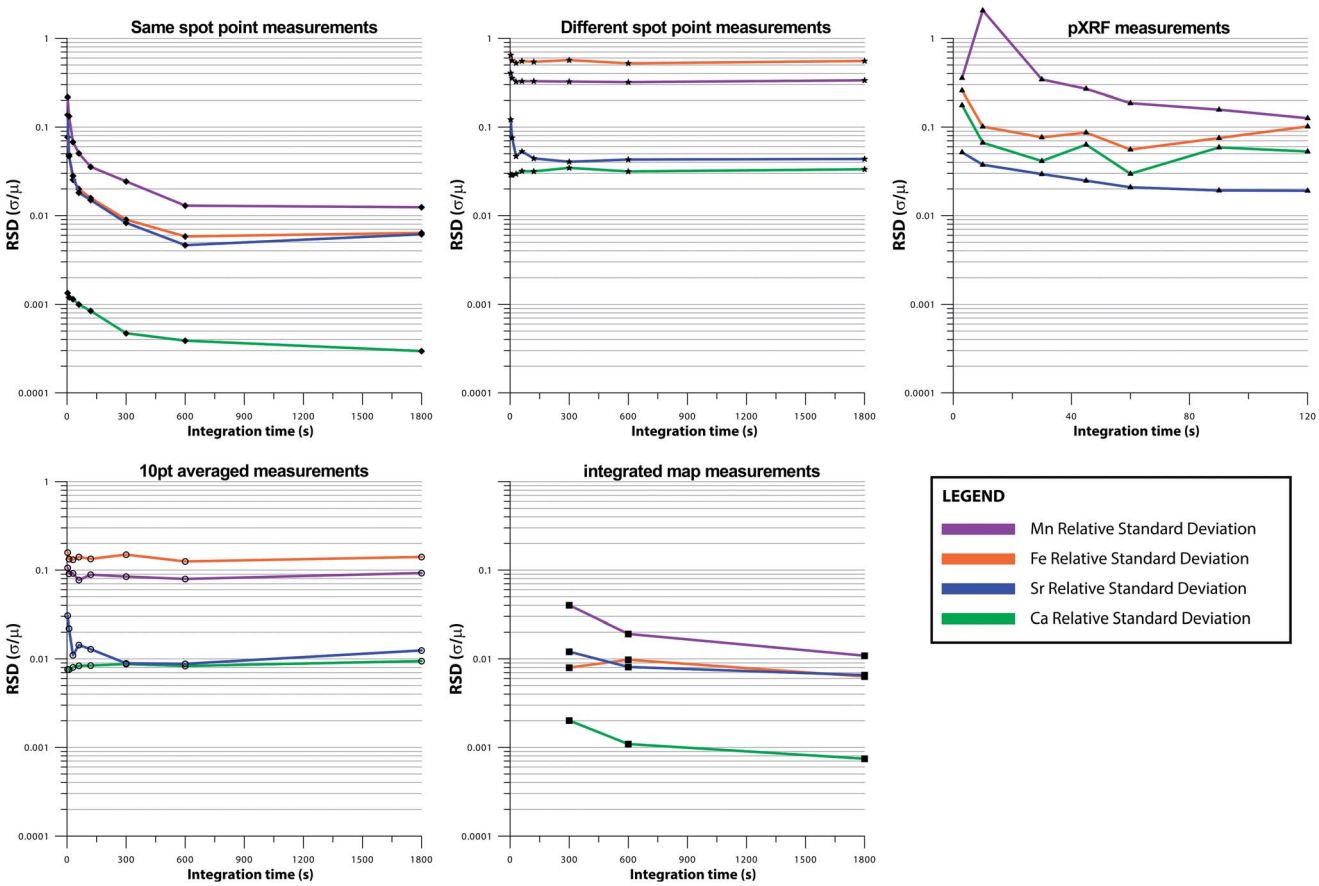

Fig. 3 Overview of the evolution of RSDs with increasing integration time for all four μXRF measurement strategies and the pXRF measurements. The elements Mn (purple), Fe (orange), Sr (blue) and Ca (green) are shown in different colors and the measurement strategies are represented by different symbols. Note that the integration time axis is shorter for pXRF measurements, which were not performed for times longer than 2 minutes. The vertical scale is the same for each graph.

for Sr. For all elements, the highest precisions are reached with the single spot strategy and map strategies. Fig. 3 further demonstrates that, while all strategies yield lower RSDs with increasing integration times, the shape of the logarithmic curve describing this evolution is different for different strategies. The pXRF measurements are described by flatter curves, starting out with high RSD values and reaching stable precisions at shorter integration times than the μXRF measurements. Repeated measurements on the same spot of the sample powder show the best increase in precision after long integration times. In most of the other strategies, the increase in precision of the measurements occurs rather quickly and increases at longer integration times (>60 s) are small or absent. Fig. 3 shows clearly that results with the single spot strategy are much more repeatable than those using the multiple spot strategy as a result of the effect of sample surface heterogeneity on reproducibility of multiple spot measurements. The average spot strategy also yields more repeatable results than the multiple spot strategy. For some strategies (*e.g.* pXRF method and multiple spot strategy) the difference in precision between elements is less than for other strategies (*e.g.* single spot and map strategies), although most of the variation seems to be caused by the very low RSDs for Ca in these single spot and map strategies.

Relative errors are variable for different elements. Of the selected elements in Fig. 2, Mn and Fe have the highest RSD values in most strategies, while Ca and Sr are measured with more precision. There is a difference between the examined elements in terms of the minimum integration time that is needed for the RSD value to stabilize. Some elements, like Fe and Ca, show RSDs that decrease steeply and reach their asymptotic value at low integration times, while Sr and Mn RSD values only stabilize with higher integration times.

3.4 Accuracy of XRF methods

To illustrate the accuracy of μXRF measurements, the values obtained for five elements (Ca, Fe, Sr, Mn and Mg) with different measurement times on all three standards using the multiple spot strategy are plotted against the certified values for these elements (Fig. 4). Mg is added to show how the factory one-point calibration FP quantification method of the Bruker M4 Tornado performs over large mass fraction differences (*i.e.* between the dolomite and limestone standards). Fig. 4 shows that the CRM393 standard used for the one-point calibration falls on or

near the 1 : 1 line, showing that FP quantification gives accurate results for samples with mass fractions close to the calibration standard. The measured mass fractions are generally less accurate when the elemental mass fractions in the measured samples (CRM512 and ECRM782) deviate further from the mass fractions of the same elements in the calibration standard (CRM393). Fig. 4 also illustrates that the mass fraction curves generally approach the 1 : 1 line with increasing integration time, resulting in better accuracies for measurements with longer integration times. Variations between lines of different integration times become progressively smaller, illustrating the asymptotic character displayed by most XRF results with increasing integration time (see Fig. 2). Elements with low mass fractions (*e.g.* Sr in ECRM782) are measured by using the μXRF plot relatively far away from the 1 : 1 line, while higher mass fractions (*e.g.* Ca and Mg) are generally measured more accurately, except when mass fractions deviate far from the mass fraction in the calibration standard (CRM393).

4. Discussion

4.1 Measurement strategies

Fig. 2–4 illustrate that the reproducibility and accuracy of XRF measurements varies with the applied measurement strategy, the integration time, the element that is measured and its mass fraction in the sample. While some elements (*e.g.* Cl, Ti and Cr) never reach satisfactory (RSD < 0.33) values, the reproducibility of other elements can be greatly improved by choosing the appropriate measurement strategy and integration time.

Fig. 2 demonstrates that the error of deconvolution always underestimates the (real) error of reproducibility for repeatable element mass fractions. The reason for this difference is that the error of deconvolution measures the error in the calculation of mass fractions from XRF spectra,[33,34] without taking into account the errors associated with the measurement conditions and sample properties. Errors of deconvolution will not include variation caused by differences in sample preparation and measurement strategy, and are solely dependent on the total amount of XRF counts in the entire spectrum, and therefore on the integration time. This difference is clearly illustrated in Fig. 3, where the difference in RSD between single spot measurements and multiple spot measurements is shown. Changes in the properties of the surface of the material that is being measured (in this case CRM393 limestone powder) cause this offset. The multiple spot strategy is a better approximation of the reproducibility error that has to be taken into account when measuring a random spot on the sample surface. Surface effects are of greater relative importance for some elements (*e.g.* Ca) than for others (*e.g.* Sr), and therefore need to be taken into

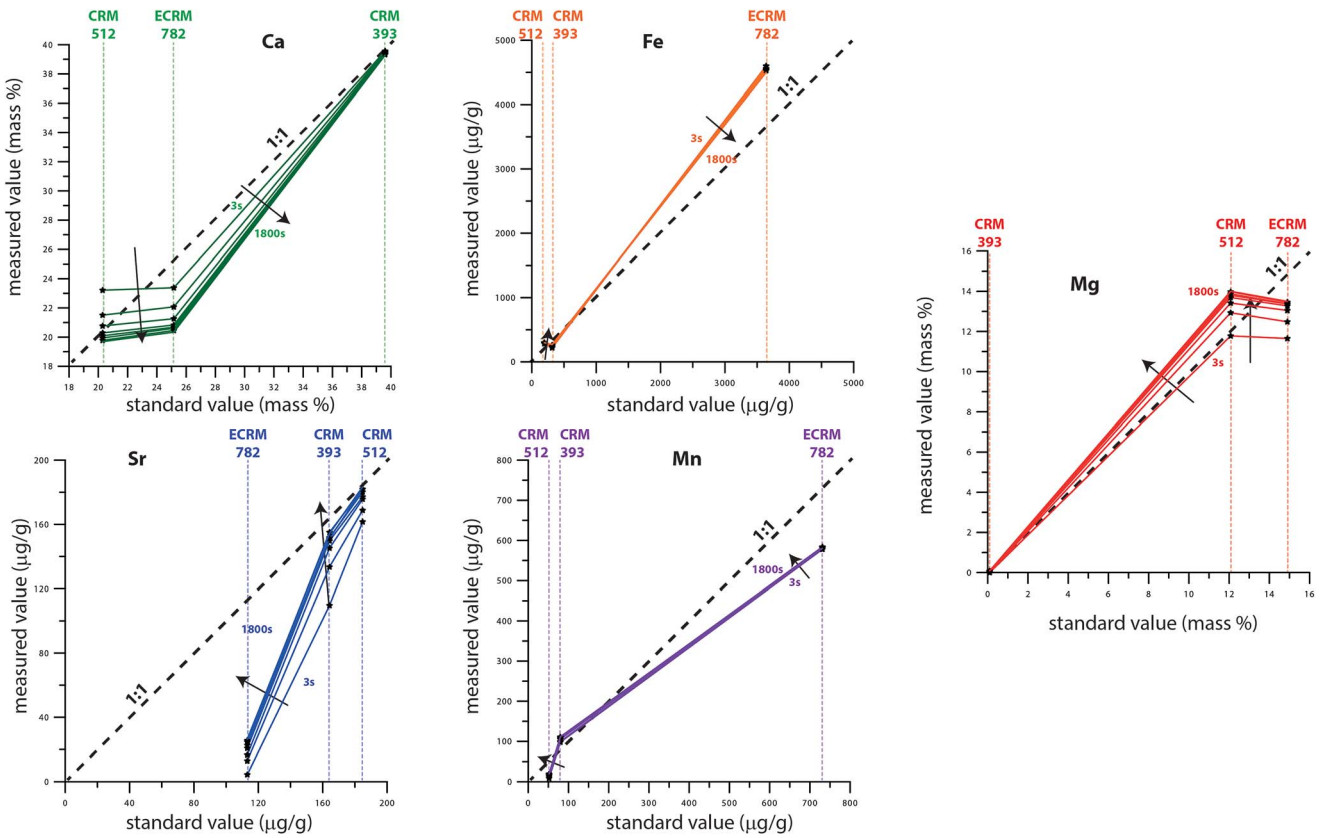

Fig. 4 Overview of the change in accuracy (represented by the distance to the 1 : 1 line) of different spot measurements on the μXRF with changing integration times and for different elements. Note that CRM393 values are always closest to the 1 : 1 line because the fundamental parameters method was calibrated using this standard. Note also that measurements generally approach the 1 : 1 line (become more accurate) with increasing integration time, and that elements with a low mass fraction (*e.g.* Sr in ECRM782) are less accurate.

account in the measurement of individual trace element records as well as in the calculation of trace element ratios.

The pXRF clearly performs worse than all μXRF sampling strategies, except for the multiple point strategy, in terms of reproducibility (Fig. 3). The shallow RSD curve of the pXRF shows that not much improvement can be made by extending the measurement time beyond ~60 seconds. A similar conclusion was reached by Ross *et al.*[8] while measuring a fine-grained marine sediment standard with a pXRF with similar varying integration times as used in this study. However, reproducibility values for Fe, Ca and Sr do reach satisfactory values (<0.1) already after 10 seconds of measurements, showing that the pXRF is a fast and robust means of measuring the mass fractions of common trace and major elements in carbonate powders. The high RSD of Mn shows that the reproducibility of the pXRF quickly diminishes for elements with lower mass fractions (<30 μg g^{-1}).

Because the single spot strategy is not a realistic approximation of choosing one measurement spot on a sample surface, three sampling strategies for the μXRF remain to be discussed. Based on the data presented in Fig. 3, the strategy of choosing one spot on the sample surface (represented by multiple spot measurements) yields repeatable values (RSD < 0.33) though with relatively high errors of reproducibility for Mn and Fe and higher RSDs than other strategies for Ca and Sr. Much better results are reached with the strategy of averaging 10 point measurements. Nevertheless, mapping a part of the surface and averaging out over this area yields the best result in terms of reproducibility (Fig. 3). The success of these strategies lies in their ability to reduce the influence of heterogeneities in the sample surface and matrix, which is crucial for a small-spot, surface-based method such as μXRF.[2] For the same reason, the larger spot size of the pXRF (6–8 mm) allows it to compete with the multiple spot strategy in terms of reproducibility for the aforementioned selection of elements.

4.2 Accuracy

The data presented in Fig. 4 illustrates another issue that needs to be addressed carefully in XRF measurements: the sample matrix. The effects of the different matrices between the calcium carbonate and dolomite standards result in an offset of measured values from the certified values. Although an increase in integration time seems to reduce this offset, using longer integration times does not entirely cancel matrix-related differences (Fig. 4). This illustrates that XRF measurements must always be calibrated using standards with the same matrix as the samples (matrix-matched). High accuracies can be reached using the μXRF method, as shown by the results of the CRM393 standard, which was calibrated using different measurements on sub-samples of the same standard. Trace and major elemental mass fractions of the other standards, which are close to those of CRM393, have higher accuracies (Fig. 4). Therefore, the calibration standard for XRF measurements should be as close as possible to the measured sample in matrix and composition.[1,2] This limitation poses problems for XRF analysis of samples with a heterogeneous matrix, such as sedimentary records with alternations of different lithologies using the fundamental parameter quantification supported by only one standard.

4.3 Integration time

As shown in Fig. 1, accuracy and reproducibility of XRF measurements generally tend to increase with integration time. This behavior is also observed in Fig. 2–4, showing a decrease of the RSD and a shift of the results towards certified values with increasing integration time (see also ref. 8, 31 and 32). Elements not following this evolution through integration time in μXRF measurements (*e.g.* in this case Cl, K, Ti, Cr, Ni and Rb) generally do not produce repeatable results (*i.e.* RSD > 1, Fig. 2). The error of deconvolution for these elements increases with the integration time, showing that their peaks in the XRF spectrum diminish with higher integration times. These peaks are not likely to represent fluorescence of the element in the sample, and may instead be noise in the spectrum or secondary peaks induced by other elements or the Rh source (*e.g.* the peak of Cl-Kα overlaps with that of Rh-L series, Table 1). Other phenomena that can cause these peaks to appear in regions of the spectra associated with the aforementioned elements are the occurrence of sum-peaks and escape-peaks.[1,2] These artefacts can originate due to the interference of two X-ray photons (sum peaks) or of X-ray photons with the detector material (escape peaks). While the Esprit software corrects for the presence of such artefacts, complex interference and artefacts of the correction itself can still produce XRF peaks that do not represent the fluorescence of an element. The example in this study shows that peak identification remains a user-guided process and that care must always be taken to only identify peaks of elements that are really present to avoid false positive results. One straightforward way to check whether an XRF peak represents an element is to identify a second fluorescence peak of this element in the spectrum.[1,2] The actual mass fractions of elements that are affected by these phenomena in the standards used in this study is too low for a reliable measurement. The evolution of the RSD of Ca is an exception to this asymptotic pattern. Both the values of RSD for Ca and the changes in RSD for Ca are very low on account of the high mass fraction of Ca in the standards used in this study. Furthermore, the fundamental parameter quantification procedure of the Esprit software uses Ca to calculate the mass fractions of the elements C and O by stoichiometric proportions as well as to complete the sum of mass to 100%, assuming a carbonate matrix. This so-called quantify-per-difference methodology yields more accurate mass fractions of trace elements in carbonate samples, but limits the variability of Ca mass fraction resulting in the low variation in RSD of Ca. After calculating an initial mass fraction for Ca based on the Ca-peak in the XRF spectrum, the method adjusts elemental mass fractions to sum the total mass to 100%. Because Ca is identified as the main measurable matrix element, its mass fraction is used to complete the sum of all mass fractions to 100%. This also explains why, contrary to Sr, Mn and Fe (see Fig. 1), Ca mass fractions do not converge to a value as close to the certified value. While Ca mass fractions of long integration time measurements are still accurate within

Table 1 Table showing the calculated minimum integration times by which a reliable measurement is reached and after which no significant improvement occurs with increasing measurement time for all measurement strategies and methods applied in this study. Red text and empty positions indicate measurements for which the threshold could not be calculated because it is not reached within the measured time frame. Zero values indicate that the threshold for reliable measurement was reached at the shortest integration time (3 s)

Element	M (u)	E peak (keV)	CRM393																			
			Same-spot				Different spot				pXRF				Maps				Average spot			
			LLD	LLQ	TSR	TSA	LLD	LLQ	TSR	TSA	LLD	LLQ	TSR	TSA	LLD	LLQ	TSR	TSA	LLD	LLQ	TSR	TSA
C	12.00	-	0	0	4	45	0	0	8	0	0	0		0	0	0	158	1384	0	0	2	0
O	16.00	-	0	0	1	162	0	0	1	0	0	0		0	0	0	214	0	0	0	0	0
Na (Kα)	22.99	1.04	INF	INF	3623	0	1480	INF	3186	0	0	14		0	INF	INF		0	1539	INF		70
Mg (Kα)	24.31	1.25	57	538	675	298	56	580		151	14	104		0	0	1481	1183	442	56	584	201	545
Al (Kα)	26.98	1.49	0	26	189	266	3	37	456	0	62	INF		0	0	0	229	1052	0	37		10
Si (Kα)	28.09	1.74	0	10	147	120	0	14	302	0	60	INF		0	0	0	53	1524	0	14	978	0
P (Kα)	30.97	2.01	0	12	319	1090	0	12	667	15	20	INF		0	0	0	396	522	0	12		110
S (Kα)	32.07	2.31	57	250	1407	5	27	55	342	0	INF	INF		0	0	0	1987	1424	26	55	135	0
Cl (Kβ)	35.45	2.81	INF	INF		0	INF	INF		120	INF	INF		0	300	300		0	INF	INF	27	120
K (Kβ)	39.10	3.59	26	138		0	30	220	1605	0	INF	INF	161	0	0	349		0	30	221	210	0
Ca (Kα)	40.08	3.69	0	0	4	278	0	0	8	0	0	0	123	0	0	0	32	0	0	0	5	0
Ti (Kα)	47.87	4.51	INF	INF	754	71	INF	INF	58	0	INF	INF	375	0	0	377	243	1416	INF	INF	208	0
Cr (Kα)	52.00	5.41	4	30	211	724	INF	INF	1126	0	119	INF	623	0	INF	INF	1936	0	3	3	2	0
Mn (Kβ)	54.94	6.49	0	9	155	1103	0	6	91	0	15	116	421	0	0	0	275	1434	0	6	55	0
Fe (Kα)	55.85	6.40	0	0	78	225	0	5	101	0	0	0	169	0	0	0	21	839	0	5	365	0
Ni (Kβ)	58.69	8.26	0	0	1171	641	0	0	1620	0	0	46	26	0	0	0		0	0	0	23	20
Cu (Kβ)	63.55	8.91	0	26	689	418	9	52	1242	7	4	81	304	0	0	0	1256	844	9	53	16	213
Zn (Kβ)	65.39	9.57	11	92	417	1260	0	28	539	0	60	INF	163	0	0	0	472	1582	0	28	401	27
Br (Kα)	79.90	11.92	4	37	445	435	0	26	654	9	INF	INF		0	0	0	613	1231	0	26	460	200
Rb (Kα)	85.47	13.39	1395	1677		0	38	40		0	INF	INF		0	0	0		0	INF	INF	152	0
Sr (Kα)	87.62	14.17	0	4	89	204	0	4	82	22	0	0	63	6	0	0	62	1525	0	5		59
Rh (Lβ)	102.91	2.83	3	3		0	0	0		0	INF	INF		0	0	0		0	0	0	27	0
Ba (Lα)	137.33	4.46	INF	INF	754	588	86	543	1381	0	INF	INF		0	0	455	623	1542	86	546	167	0
Pb (Lα)	207.20	12.14	6	32	376	1300	0	4	1272	4	INF	INF		0	0	0	718	481	0	26	447	224

Element	M (u)	E peak (keV)	CRM512								ECRM782							
			Different spot				Average spot				Different spot				Average spot			
			LLD	LLQ	TSR	TSA	LLD	LLQ	TSR	TSA	LLD	LLQ	TSR	TSA	LLD	LLQ	TSR	TSA
C	12.00	-	INF	INF	5	0	INF	INF	1	0	INF	INF	2	0	INF	INF	0	44
O	16.00	-	INF	INF	7	8	INF	INF	2	49	INF	INF	1	0	INF	INF	0	21
Na (Kα)	22.99	1.04	INF	INF		0	INF	INF		276	1799	INF		0	1798	INF	36	245
Mg (Kα)	24.31	1.25	0	0	80	9	0	0	189	55	0	0	39	0	0	0	101	25
Al (Kα)	26.98	1.49	24	150	555	97	24	154		566	17	120	482	121	17	120		554
Si (Kα)	28.09	1.74	0	60	613	0	0	60		8	12	106	611	10	13	105		160
P (Kα)	30.97	2.01	0	29	467	114	0	29	26	547	0	37	416	58	0	38	10	565
S (Kα)	32.07	2.31	537	INF	1019	96	542	INF	215	526	30	287	822	65	31	289	195	539
Cl (Kβ)	35.45	2.81	INF	INF		120	INF	INF	199	120	0	18	158	0	0	18	203	16
K (Kβ)	39.10	3.59	101	453		0	105	462	178	0	10	47	623	9	10	46	159	107
Ca (Kα)	40.08	3.69	0	0	49	10	0	0	367	85	0	0	3	0	0	0	297	36
Ti (Kα)	47.87	4.51	INF	INF	12	0	INF	INF		4	INF	INF	353	0	INF	INF	246	0
Cr (Kα)	52.00	5.41	INF	INF	1113	0	INF	INF	17	0	INF	INF	1250	0	INF	INF	0	0
Mn (Kβ)	54.94	6.49	0	45	552	0	0	45	3	32	0	0	31	0	0	0	137	0
Fe (Kα)	55.85	6.40	0	0	73	0	0	0	225	0	0	0	23	0	0	0	279	0
Ni (Kβ)	58.69	8.26	0	0	447	0	0	0	160	9	0	0	508	0	0	0	1	78
Cu (Kβ)	63.55	8.91	296	403	2699	0	296	423	29	20	214	631	2157	0	214	647	4	4
Zn (Kβ)	65.39	9.57	118	427	1392	0	31	519	119	10	0	8	112	0	0	8	98	25
Br (Kα)	79.90	11.92	0	16	638	9	0	16	841	207	0	10	328	3	0	10	672	22
Rb (Kα)	85.47	13.39	594	594		0	590	577	432	0	INF	INF		0	INF	INF	39	134
Sr (Kα)	87.62	14.17	0	0	192	0	0	0		0	6	54	526	33	6	54		208
Rh (Lβ)	102.91	2.83	29	INF		0	1467	INF	55	0	279	1445		0	243	1449	205	0
Ba (Lα)	137.33	4.46	346	1592	2187	0	327	1594	571	0	53	394	1871	0	52	397	605	7
Pb (Lα)	207.20	12.14	78	952	906	0	79	936	264	83	403	1683	547	6	375	1753	184	107

one standard deviation, the quantify-per-difference method adds or removes Ca to complete the sum to 100% for each measurement, creating the slight offset observed in Fig. 1. The (soil) fundamental parameter quantification of the pXRF uses the SiO_2 mass fraction to sum the total weight to 100%. As a result, Si mass fractions from pXRF measurements will show a similar lack of improvement in RSD and accuracy with increasing measurement time.

4.4 Thresholds for reliable measurements

Several authors have proposed thresholds (for mass fractions and measurement times) for the reliable detection and quantification of XRF spectra based on the error of deconvolution, such as LLD and LLQ.[29,37,38,40] Fig. 5 shows an example of where these thresholds are found relative to the asymptotic shape of the RSD graph and how minimum measurement times for the detection and quantification of elements by XRF can be derived based on these thresholds. An overview of the integration times by which LLD and LLQ are achieved using different measurement strategies is found in Table 1. However, as discussed above, the error of deconvolution is not a good predictor of the total error of measurement and thresholds of measurement based on errors of real accuracy and reproducibility should be preferred.

One common way to calculate measurement error based on repeated analyses is to use one or more SD as the confidence level.[29,41–43] Fig. 5 shows the point in the RSD graph where a 3 SD threshold is reached, and illustrates that the associated integration time is not an acceptable threshold for a minimum measurement time, as much better precisions are reached by increasing the integration time. Instead, the shape of the asymptotic RSD curve can be used to obtain a recommended minimum integration time. For elements yielding repeatable

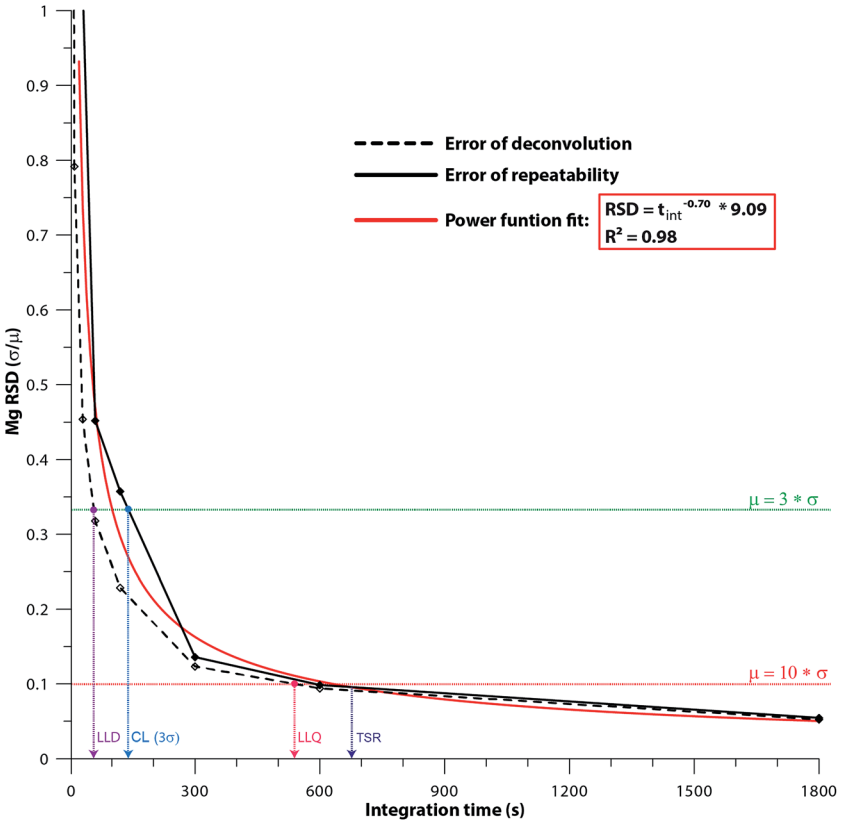

Fig. 5 Example of the change in RSD with integration time for Mg using the different spot method on μXRF. The solid black line shows RSDs of reproducibility and the dashed line shows RSDs of deconvolution. In red is the power curve fitted through the data (equation and coefficient of determination are reported in the legend) to calculate the Time of Stable Reproducibility (TSR). Horizontal red and green line show positions of 3 and 10 standard deviation thresholds used to calculate the time of lower limit of detection, lower limit of quantification and 3 standard deviations confidence level. TSR is defined as the time by which the slope of the power curve reaches 10−4 RSD per s.

measurements, this curve describes a negative power function of the form $y = x^a \times b$, in which y is the RSD, x is the integration time and a and b are constants. In this function, a is negative for elements for which RSD decreases with increasing integration time (Fig. 5). To use this model, a conservative threshold of 10^{-4} RSD per s is proposed to represent a negligible rate of improvement of RSDs with increasing measurement time. When this threshold is reached, an unrealistic measurement period of 10 000 s (about 3 hours) would be needed to improve the RSD by one unit. The time by which this happens is proposed as an adequate threshold for reliable measurements and will be referred to as the Time of Stable Reproducibility (TSR). An overview of TSR values for all elements measured by various strategies on the three standards is given in Table 1. Note that the decision of this threshold is arbitrary, as the power function dictates that precision will keep increasing even for very long integration times and will never become completely stable. Depending on the research question and the available time on the machine, less conservative thresholds could be deemed sufficient as long as the reason for choosing these thresholds is motivated.

Alternatively, a threshold may be proposed based on the accuracy of measurement rather than the reproducibility. Such a threshold is reached for a particular measurement strategy and element when the measured mass fraction falls within a confidence level of two SDs of reproducibility of the most accurate measurement (the measurement using 1800 s integration time in the case of μXRF measurements, Fig. 1). After this threshold, an increase in measurement time does not imply a significant increase in the accuracy of the measurement, and is therefore not necessary. The integration times by which this threshold is reached is referred to as the Time of Stable Accuracy (TSA). As with the TSR, this confidence level can be altered (*e.g.* to 3 SDs) in function of the needs of the study. Values for TSA for all measurement strategies and elements were calculated by interpolation of the intersection of mass fraction graphs (Fig. 1) with 2 SD confidence levels of the 1800 s measurements, and are shown in Table 1.

4.5 Minimum integration time

The integration time thresholds calculated using the four different strategies are shown in Table 1. From this table, it is clear that thresholds based on error of deconvolution (LLD and LLQ) generally underestimate the time needed to reach a stable reproducibility. As discussed above, a threshold for minimum measurement time based on error of deconvolution cannot be

used to recommend a measurement strategy and integration time. Therefore, it is recommended that another threshold based on total errors of measurement, defined by reproducibility (TSR) or accuracy (TSA), is used instead of LLD or LLQ to determine the appropriate measurement strategy and integration time in carbonate XRF analysis. While detection and quantification limits based on errors of deconvolution seem to provide a method-independent indication of measurement quality, they fail to report the total error of the measurement: they should therefore be avoided as indicators of measurement precision.

The comparison between TSR and TSA is not straightforward, because the total error on the 1800 s measurement is large for some measurement strategies, resulting in high TSR times but low TSA times. Also, for some elements like Sr in CRM393 (see Fig. 1), the mean mass fraction of the 30 measurements does not show much change with longer integration times, while the reproducibility of the measurement increases significantly. Since for most measurements reproducibility shows more improvement with increasing integration time than accuracy, the most conservative integration time threshold for reliable measurement is the TSR.

Based on TSR, the minimum integration time recommended for most μXRF measurement strategies discussed in this study is in the order of minutes (100–300 seconds for most elements). Table 1 shows that for carbonate powders, the recommended strategy with the highest accuracy and reproducibility on common trace elements is to either map the surface of the powder for 300 to 1800 s (mapping strategy) or to combine 10 random spots on the surface using an integration time of 60 to 300 s per spot (average point strategy). Measurements of one random spot (illustrated by the multiple spot strategy) either do not reach significant reproducibility (RSD < 0.33) or require very long integration times to achieve accurate results. This difference is a result of heterogeneity in the sample surface and in the properties of the powder and should also be taken into account for other surface based methods (*i.e.* X-ray diffraction, color analysis and Fourier Transform Infrared Spectroscopy).

Results in Table 1 also demonstrate that the pXRF can be used for the detection and even for the quantitative measurement of some elements (*e.g.* Ca, Fe and Sr) in a calcium carbonate material using a ~120 s integration time, but the precisions reached by the handheld instrument are not sufficient to quantify most other elements with confidence within a realistic timeframe. However, the handheld technique can still be used in a semi-quantitative way for elements that can be detected but not quantified (above the LLD but below the LLQ; Table 1). This makes the pXRF a useful tool to detect and discuss relative variations of elements, which can be very useful for initial survey studies in the field. However, attention should always be given to the surface and weathering conditions of the measured samples in the field.[3,31,32]

4.6 Applications and further research

This study focuses on XRF measurements on calcium carbonate and dolomite materials, commonly investigated in paleoclimate studies. It must be emphasized that the intention of this work is not to provide minimal measurement times for all elements for all types of carbonate materials. Calcium carbonates and dolomites have specific matrix effects and ranges of mass fractions of elements. Because these matrix effects have a large impact on XRF measurements, they need to be isolated specifically for the studied material. Investigations of TSR and TSA using appropriate, matrix-matched, reference materials should be set-up for different types of materials and should be the basis of every thorough XRF study. If such standards are not available, ICP-MS measurements on homogenous samples can be used to calibrate in-house XRF standards to serve this purpose. In order to improve the accuracy of XRF measurements of a specific material, a calibration curve for this material can be constructed. A more extended study using a broader range of (certified) carbonate materials with this goal could certainly be a good addition to the discussion of the reliability of XRF measurements on carbonate and other materials. In addition, since different materials will have different matrix effects, single-point, multiple-point or surface integration measurement strategies could be compared in other studies to further isolate the effect of sample matrix on XRF measurement strategies.

Other applications of the methodological approach presented in this study could not just focus on optimizing reproducibility or accuracy but rather on other research questions where the integration time is a critical parameter. Examples include the calculation of minimal or maximal acquisition times for methods where long exposure can damage a sample or where long measurement times are expensive (*e.g.* SR-XRF). Alternatively, the TSR and TSA approach can be applied to any type of transient signal acquisition where acquisition time per point influences the reliability of the result (*e.g.* XRF line scanning, depth profiling, laser ablation, *etc.*). In the example of LA-ICP-MS or XRF line scanning, the TSR and TSA can be applied to calculate scanning speeds and spatial resolution of reproducible measurements.

5. Conclusions

In this study we present a novel approach to assign appropriate measurement strategies for XRF measurements on calcium carbonate and dolomite materials. We also demonstrate that conventional limits of detection and quantification, which are solely based on the XRF spectrum properties, underestimate appropriate measurement times as they do not take into account, for example, matrix and surface effects. Our methodological approach is easily applicable to other types of measurements and materials as well.

Systematic testing of the effect of changing measurement strategy and integration time on XRF measurement results shows the advantages and limitations of using handheld portable XRF and laboratory scale μXRF methods to determine mass fractions of trace elements in calcium carbonate and dolomite powders. These results show that accuracy and reproducibility can be improved significantly, though not indefinitely, by increasing the integration time of the XRF

measurement. However, this improvement remains limited and thresholds based on reproducibility and accuracy of the measurements can be established to show after which length of time an increase in measurement time cannot improve the quality of the measurement significantly. For the determination of such thresholds, two new techniques are proposed, namely the Time of Stable Reproducibility (TSR) and the Time of Stable Accuracy (TSA). These thresholds can be determined for a given measurement strategy, instrument setup, sample preparation and sample matrix and used as a guideline for measuring samples under these conditions. As surface properties and matrix effects have a significant effect on the quality of XRF results, the abovementioned thresholds for the measurement time provide a better estimation of the time needed to achieve reliable XRF results than the conventional limits of detection and quantification, which are solely based on the XRF spectrum properties. According to these thresholds for reliable measurements, it can be concluded that the most precise results in terms of quantitative trace element analysis are achieved by averaging the results of multiple μXRF point measurements or integrating a μXRF-mapped surface of the sample. Most common trace elements (*e.g.* Sr, Mn and Fe) can be quantified using the Bruker M4 Tornado μXRF with an integration time of 60 s to 300 s. Satisfactory results are obtained using the Bruker Tracer pXRF for Ca, Fe and Sr, though less precise than the μXRF, after 120 s of measurement, and it is suitable for semi-quantitative analysis of other detectable trace elements.

Acknowledgements

Research by N. J. de Winter is funded by an awarded IWT PhD Flanders fellowship (IWT700-SB-141047). Matthias Sinnesael thanks the Research Foundation of Flanders (FWO) for the awarded PhD fellowship (FWOTM782). Funding for Christina Makarona and Stef Vansteenberge comes from VUB Strategic Research. Technical support of Dr Roald Tagle (Bruker) and Luc Deriemaker (VUB) is acknowledged. Ph. Claeys thanks the FWO-Hercules foundation for financing the XRF analytical platform at the VUB.

References

1 *Handbook of Practical X-ray Fluorescence Analysis*, ed. B. Beckhoff, Springer, Berlin, New York, 2006, p. 863.
2 M. Haschke, *Laboratory Micro-x-Ray Fluorescence Spectroscopy*, Cham: Springer International Publishing, 2014, Springer Series in Surface Sciences, vol. 55.
3 P. J. Potts and M. West, *Portable X-ray Fluorescence Spectrometry: Capabilities for in situ Analysis*, Royal Society of Chemistry, 2008.
4 E. Frahm and R. C. P. Doonan, The technological *versus* methodological revolution of portable XRF in archaeology, *J. Archaeol. Sci.*, 2013, **40**(2), 1425–1434.
5 S. E. Calvert and T. F. Pedersen, Chapter fourteen elemental proxies for palaeoclimatic and palaeoceanographic variability in marine sediments: interpretation and application, *Developments in Marine Geology*, 2007, vol. 1, pp. 567–644.
6 G. J. Weltje and R. Tjallingii, Calibration of XRF core scanners for quantitative geochemical logging of sediment cores: theory and application, *Earth Planet. Sci. Lett.*, 2008, **274**(3), 423–438.
7 H. Rowe, N. Hughes and K. Robinson, The quantification and application of handheld energy-dispersive x-ray fluorescence (ED-XRF) in mudrock chemostratigraphy and geochemistry, *Chem. Geol.*, 2012, **324**, 122–131.
8 P.-S. Ross, A. Bourke and B. Fresia, Improving lithological discrimination in exploration drill-cores using portable X-ray fluorescence measurements: (1) testing three Olympus Innov-X analysers on unprepared cores, *Geochem.: Explor., Environ., Anal.*, 2014, **14**(2), 171–185.
9 D. S. Gholap, A. Izmer, B. De Samber, J. T. van Elteren, V. S. Šelih, R. Evens, K. De Schamphelaere, C. Janssen, L. Balcaen, I. Lindemann, L. Vincze and F. Vanhaecke, Comparison of laser ablation-inductively coupled plasma-mass spectrometry and micro-X-ray fluorescence spectrometry for elemental imaging in Daphnia magna, *Anal. Chim. Acta*, 2010, **664**(1), 19–26.
10 S. Kurunczi, S. Török and P. Chevallier, A micro-XRF study of the element distribution on the growth front of mussel shell (species of Unio Crassus Retzius), *Microchim. Acta*, 2001, **137**(1–2), 41–48.
11 M. S. Hull, P. Chaurand, J. Rose, M. Auffan, J.-Y. Bottero, J. C. Jones, I. R. Schultz and P. J. Vikesland, Filter-feeding bivalves store and biodeposit colloidally stable gold nanoparticles, *Environ. Sci. Technol.*, 2011, **45**(15), 6592–6599.
12 N. J. Winter and P. Claeys, Micro X-ray fluorescence (μXRF) line scanning on Cretaceous rudist bivalves: A new method for reproducible trace element profiles in bivalve calcite, *Sedimentology*, 2017, **64**(1), 231–251.
13 N. J. de Winter and C. Snoeck, Claeys, Ph., Seasonal cyclicity in trace elements and stable isotopes of modern horse enamel, *PLoS One*, 2016, **11**(11), 1–29.
14 A. Borsato, S. Frisia, I. J. Fairchild, A. Somogyi and J. Susini, Trace element distribution in annual stalagmite laminae mapped by micrometer-resolution X-ray fluorescence: Implications for incorporation of environmentally significant species, *Geochim. Cosmochim. Acta*, 2007, **71**, 1494–1512.
15 M. Finné, M. Kylander, M. Boyd, H. S. Sundqvist and L. Löwemark, Can XRF scanning of speleothems be used as a non-destructive method to identify paleoflood events in caves?, *Int. J. Speleol.*, 2014, **44**(1), 17–23.
16 J. Buckles and H. D. Rowe, Development and optimization of microbeam X-ray fluorescence analysis of Sr in speleothems, *Chem. Geol.*, 2016, **426**, 28–32.
17 D. N. Papadopoulou, G. A. Zachariadis, A. N. Anthemidis, N. C. Tsirliganis and J. A. Stratis, Comparison of a portable micro-X-ray fluorescence spectrometry with inductively coupled plasma atomic emission spectrometry for the ancient ceramics analysis, *Spectrochim. Acta, Part B*, 2004, **59**(12), 1877–1884.
18 D. Papadopoulou, G. Zachariadis, A. Anthemidis, N. Tsirliganis and J. Stratis, Development and optimisation

of a portable micro-XRF method for *in situ* multi-element analysis of ancient ceramics, *Talanta*, 2006, **68**(5), 1692–1699.

19 A. J. Nazaroff, K. M. Prufer and B. L. Drake, Assessing the applicability of portable X-ray fluorescence spectrometry for obsidian provenance research in the Maya lowlands, *J. Archaeol. Sci.*, 2010, **37**(4), 885–895.

20 A. M. W. Hunt and R. J. Speakman, Portable XRF analysis of archaeological sediments and ceramics, *J. Archaeol. Sci.*, 2015, **53**, 626–638.

21 N. Goodale, D. G. Bailey, G. T. Jones, C. Prescott, E. Scholz and N. Stagliano, Lewis, C. pXRF: a study of inter-instrument performance, *J. Archaeol. Sci.*, 2012, (4), 875–883.

22 N. W. Brand and C. J. Brand, Performance comparison of portable XRF instruments, *Geochem.: Explor., Environ., Anal.*, 2014, 2012–2172.

23 L. Ge, W. Lai and Y. Lin, Influence of and correction for moisture in rocks, soils and sediments on *in situ* XRF analysis, *X-Ray Spectrom.*, 2005, **34**(1), 28–34.

24 R. Tjallingii, U. Röhl, M. Kölling and T. Bickert, Influence of the water content on X-ray fluorescence core-scanning measurements in soft marine sediments, *Geochem., Geophys., Geosyst.*, 2007, **8**(2), Q02004.

25 N. Forster, P. Grave, N. Vickery and L. Kealhofer, Non-destructive analysis using PXRF: methodology and application to archaeological ceramics, *X-Ray Spectrom.*, 2011 Sep, **40**(5), 389–398.

26 P. J. Potts, F. Bernardini, M. C. Jones, O. Williams-Thorpe and P. C. Webb, Effects of weathering on *in situ* portable X-ray fluorescence analyses of geological outcrops: dolerite and rhyolite outcrops from the Preseli Mountains, South Wales, *X-Ray Spectrom.*, 2006, **35**(1), 8–18.

27 L. Vincze, A. Somogyi, J. Osán, B. Vekemans, S. Török, K. Janssens and F. Adams, Quantitative Trace Element Analysis of Individual Fly Ash Particles by Means of X-ray Microfluorescence, *Anal. Chem.*, 2002, **74**(5), 1128–1135.

28 B. D. Wheeler, Analysis of limestones and dolomites by X-ray fluorescence, *Rigaku J.*, 1999, **16**(1), 16–25.

29 H. Rowe, R. Loucks, S. Ruppel and S. Rimmer, Mississippian Barnett Formation, Fort Worth Basin, Texas: Bulk geochemical inferences and Mo–TOC constraints on the severity of hydrographic restriction, *Chem. Geol.*, 2008, **257**(1–2), 16–25.

30 T. W. Dahl, M. Ruhl, E. U. Hammarlund, D. E. Canfield, M. T. Rosing and C. J. Bjerrum, Tracing euxinia by molybdenum concentrations in sediments using handheld X-ray fluorescence spectroscopy (HHXRF), *Chem. Geol.*, 2013, **360–361**, 241–251.

31 Z. Yuan, Q. Cheng, Q. Xia, L. Yao, Z. Chen, R. Zuo and D. Xu, Spatial patterns of geochemical elements measured on rock surfaces by portable X-ray fluorescence: application to hand specimens and rock outcrops, *Geochem.: Explor., Environ., Anal.*, 2014, **14**(3), 265–276.

32 J. Quye-Sawyer, V. Vandeginste and K. J. Johnston, Application of handheld energy-dispersive X-ray fluorescence spectrometry to carbonate studies: opportunities and challenges, *J. Anal. At. Spectrom.*, 2015, **30**(7), 1490–1499.

33 J. Sherman, The theoretical derivation of fluorescent X-ray intensities from mixtures, *Spectrochim. Acta*, 1956, **7**, 283–306.

34 R. M. Rousseau, Fundamental algorithm between concentration and intensity in XRF analysis 2—practical application, *X-Ray Spectrom.*, 1984, **13**(3), 121–125.

35 N. Broll, Quantitative x-ray fluorescence analysis. Theory and practice of the fundamental coefficient method, *X-Ray Spectrom.*, 1986, **15**(4), 271–285.

36 R. Sitko and B. Zawisza, *Quantification in X-Ray Fluorescence Spectrometry [Internet]*, INTECH Open Access Publisher, 2012.

37 R. M. Rousseau, Detection limit and estimate of uncertainty of analytical XRF results, *Rigaku J.*, 2001, **18**(2), 33–47.

38 L. H. Keith, W. Crummett, J. Deegan Jr, R. A. Libby, J. K. Taylor and G. Wentler, Principles of environmental analysis, *Anal. Chem.*, 1983, **55**(14), 2210–2218.

39 IUPAC, *Compendium of Chemical Terminology (The "Gold Book")*, ed. A. D. McNaught and A. Wilkinson, Blackwell Scientific Publications, Oxford, 2nd edn, 1997.

40 L. Bray, M. A. Pancucci-Papadopoulou and J. M. Hall-Spencer, Sea urchin response to rising pCO_2 shows ocean acidification may fundamentally alter the chemistry of marine skeletons, *Mediterr. Mar. Sci.*, 2014, **15**(3), 510–519.

41 D. J. Kalnicky and R. Singhvi, Field portable XRF analysis of environmental samples, *J. Hazard. Mater.*, 2001, **83**(1), 93–122.

42 M. E. Kylander, L. Ampel, B. Wohlfarth and D. Veres, High-resolution X-ray fluorescence core scanning analysis of Les Echets (France) sedimentary sequence: new insights from chemical proxies, *J. Quaternary Sci.*, 2011, **26**(1), 109–117.

43 R. Hennekam and G. de Lange, others. X-ray fluorescence core scanning of wet marine sediments: methods to improve quality and reproducibility of high-resolution paleoenvironmental records, *Limnol. Oceanogr.: Methods*, 2012, **10**, 991–1003.