

Some Updated Quality Concerns on Non-Destructive Geochemical Analysis with XRF Spectrometry

Detlef Wilke

Dr. Wilke Management and Consulting GmbH, Wennigsen, Germany

ABSTRACT

Non-destructive geochemical analysis with portable X-ray fluorescence (XRF) spectrometers is probably one of the most controversially discussed methodologies, when valid quantitative, inter-instrument and inter-laboratory comparable results are requested by the nature of the study subject. This is particularly evident in compositional analysis of ancient pottery for provenance allocation. Data acquisitions with portable XRF analysers is broadly discredited as point & shoot approach, and there is indeed justified concern about a simplistic reading and trusting of concentration figures generated by the preinstalled calibration software, which instrument manufacturers provide for the convenience of the user. Professional calibration and validation concepts are discussed which overcome this critique and allow scholars to exploit the full potential of non-destructive analysis, namely the reliable large scale, high-throughput generation of compositional data for archaeological settlement analysis or geochemical field studies. Since it is unrealistic that from one day to the other, and despite the substantial critique, a majority of interested scholars will develop into professionally trained analysts, a fit for purpose approach to data quality is suggested, which may bring such results into accordance with the accepted rules of the applied sciences.

Keywords: Calibration, Validation, Archaeometry, Ceramics, Provenancing

INTRODUCTION

Over the past 20 years portable or handheld XRF (pXRF, hhXRF) analysers have received increasing attention by archaeometrists and archaeologists looking for a fast and cost-efficient method for compositional analysis of obsidian tools and chip debris, stone tools, ancient pottery sherds, technical ceramics and construction materials [1,2]. Similarly pXRF instruments are frequently used for on-site as well as laboratory geochemical analysis for geological purposes [3]. The design of the instrument nose and window allows for positioning of even large and bulky specimens like sizeable vessel sherds and intact jars, as well as on-site analysis of architectural elements, soils, sediments, rocks and minerals. In geological exploration and sedimentology semi-quantitative data are often fully sufficient for orientation surveys and YES/NO categorizations are also appropriate for many archaeological research tasks, like for determining threshold values of Zn, As, Sn and Pb in corroded copper alloy objects [4]. In compositional provenancing of artefacts made from geological raw materials accurate values of minor and trace element concentrations, however, are mandatory (for the diverging provenance term in geology and archaeology cf [5]). Numerous studies worldwide with NAA, ICP-MS, powder XRF and non-destructive XRF have successfully provenanced obsidian tools to a limited number of glassy volcanic outcrops, whether from the narrow surroundings of the find spot or as a result of long distance transport and exchange [6]. Provided that a complete inventory of all regional flows of amorphous, silicon oxide rich volcanic outcrops has been established, obsidian sourcing is methodologically comparatively easy, since the number of potential geological sources is modest, and the compositional homogeneity within an individual obsidian flow is high [7]. By contrast the number of potential flint and basalt sources is much larger, and the local reservoirs exhibit considerable micro-heterogeneity. Finally, the number of small local clay pits useful for preindustrial potting is almost infinite, and as heterogenous microcrystalline

geological sediments, clay deposits exhibit an even larger small scale compositional variability. Thereby the scatter of minor and trace element concentrations within each local clay source is accompanied by a low overall compositional differentiation of clay sediments, which makes local and regional source discrimination often very difficult. Thus the more discriminative elements available and the more accurate its quantitative determination, the more likely it gets that the raw material provenance of stone and ceramic artifacts can be successfully elucidated.

The general perception of non-destructive compositional analysis

Micro-invasive samplings with laser ablation techniques are often considered as non-destructive too. Like PIXE, however, they are less appropriate for compositional analysis of heterogenous substrates, which natural and processed geological specimens typically are, and whose microcrystalline heterogeneity is not fully reflected by the upper 50-100 μm surface layer of the study material. Hence true non-destructive compositional analysis is synonymous with pXRF spectrometry, which for the mid-Z elements of Ga to Nb allows for reliable quantification within the upper 3 mm of sample thickness. For the “digging archaeologist” non-destructive analysis is typically not an issue, since for pottery sherds as the bulk of archaeological finds, surface abrasion, producing fresh breaking edges or cracking part of a larger sherd for grinding, is conceptually accepted. As a result of new cultural heritage regulations, however, this is under change now from country to country. By contrast archaeological material already deposited in a museum or an official collection is typically no longer accessible to any kind of destructive treatment, since curators refuse to take responsibility for any artifact damage. Apart from conservational aspects the main benefit of non-destructive XRF analysis, which avoids laborious sample preparation steps, is the high sample throughput. This is particularly advantageous in compositional settlement analysis of thousands of small pottery sherds, but also in any kind of geological exploration.

A recent comprehensive compilation by Holmqvist [8] of perceptions, experiences and achievements of scholars using pXRF analysers for non-destructive pottery analysis, accounts for far more cons than pros: unresolved surface, particle size and mineralogical effects and unresolved calibration issues. Killick [9] in his general critique of the “awkward adolescence of archaeological science” took pXRF as an explicit example stating, that for inhomogenous polycrystalline samples, non-destructive XRF analysis is “inevitably inaccurate”. With respect to instrument calibration and the quantification software provided by the manufacturers he concluded, that each pXRF study of earthenware pottery and the like is an “island unto itself”. By contrast Wilke [10] and Wilke et al. [11] addressing these issues in a dedicated ceramic study came to the conclusion that true non-destructive compositional pottery analysis with pXRF is suitable. These authors have particularly focussed on instrument calibration, precision and accuracy of concentration values of selected discriminative minor and trace elements, and on the operating procedure necessary for a reliable analysis of intact specimens.

The innocent user

XRF is a sophisticated surface analytical method which as a precondition requires homogenous sample material of infinite thickness, larger than the penetration depth of the emitted photons of the heaviest element to be quantified in the sample matrix. Furthermore XRF, like many other analytical procedures, suffers from matrix and interference phenomena which aggravate data processing, calibration and quantification. In this respect a non-destructive approach adds few complications to general XRF physics, which, with the exception of a few classes of substrate materials, can all be managed. What makes the difference between XRF and non-destructive XRF is the easy access of non-trained scholars to the portable type of instruments. Shackley [12] opened the respective debate by complaining about the lack of calibration and validation of the portable spectrometers by inexperienced users. Speakman and Shackley [13] argued, that the easy handling of portable XRF spectrometers allows users, who do not “care to learn how and why the instrument functions” to “play scientist”. Unfortunately there seems to be little progress in the scientific community – at least in archaeology and archaeometry, but the situation in geochemistry is probably quite similar – since even today many oral and poster presenters on international conferences provide “numbers” rather than data and results, and if asked by senior attendants for the detailed methodological background of their presented pXRF studies, they are often unable to give a reply. An anonymous reviewer recently told me “More and more unqualified researchers need publications, and XRF is an easy way to generate scientific-ish data.” Even if we ignore scientific rules, it is a tremendous waste of time to students when providing them with a handy instrument but without proper guidance, training and supervision.

Calibration and validation

pXRF instruments have been developed for quality assurance and fast sorting of raw materials in the metal processing industry, as well as for fast semi-quantitative environmental monitoring. The instruments are provided by the vendors

with preinstalled quantification software and pass/fail evaluations of selected metal alloys and inorganic toxic thresholds, respectively, for the convenience of the user, who typically has no specific experience in instrumental analytics at all. Geochemical analysis can thus be considered as a nearby, but far more sophisticated spin-off of the bulk applications. Indeed many instrument vendors admit that the portable spectrometers have never been developed for quantitative analysis in a scientific or academic context. Opponents coming from laboratory/stationary analytics have therefore criticised pXRF analytics as point and shoot approach. However, there is nothing bad with pointing and shooting, if the agent knows where to point and when to shoot. The real problem with pXRF analysis by inexperienced analysts is point, shoot and believe.

The preinstalled quantification software provided by the vendors converts the primary signals received by the detector into channel specific net counts – areas under element specific spectral peaks. Net counts are arbitrary units, which i.a. depend on the instrument, the lifetime of the measurement, current and voltage, and the beam filter used. Net counts are not even “internally consistent”, unless the individual instrument is run under identical operating conditions, and due to matrix effects on X-ray photon attenuation, only within a set of samples with very narrow matrix differences. In order to make the raw data comparable with other instruments, other conditions and other samples differing in its matrix composition, the vendor software typically applies fundamental parameter algorithms to convert net counts per energy channel range into mass concentrations of the respective elements in % w/w or ppm. The calibration, i.e. the respective multiplication factor for converting net counts into % w/w or ppm, is determined by mathematical approximations of the characteristic photon emissions of the pure elements, and fine-tuned with a set of standard reference materials, which in case of geochemical calibrations are well characterized soils, sediments and rock powders. Since XRF is photon physics, the correlation of signal size and element concentration is constant, but only for a given instrument condition and sample matrix. In order to develop a universal quantification program, fundamental parameter software normalizes the net counts of each specific element (area under peak) to the net counts of the Compton scatter (area under the Compton peak). The X-ray beam when hitting the sample not only excites element specific photons which can be quantified, but also two characteristic “internal” signals, the Compton and the Rayleigh scatter of the tube element, typically Rh or Ag. Both source signals respond to matrix differences and to the distance of the sample from the detector in a fairly similar way like the respective photons of the mid-Z elements [11,14,15]. Vendor installed calibration software is proprietary and thus a black box for the user. Relying on the performance of these manufacturer solutions without proof of accuracy must thus be considered as a lack of diligence. Indeed recent studies have demonstrated for all major instrument brands that it is difficult, and for certain elements like Ga, Y, Zr and Nb perhaps even impossible to validate the factory calibrations [16-20], which is most probably due to unresolved matrix and spectral interference issues, but also to uncertainties in the given values of the certified reference materials which users applied for validation.

Alternatively analysts have developed “matrix matched calibrations” by selecting a set of reference materials, which are similar to the matrix of their study material, either selected certified reference materials or well-characterized “secondary standards” like obsidian samples from different sources [21], mud rock samples [22] and soils [20] and used the given element concentrations determined by NAA, ICP-MS or laboratory XRF for developing calibration factors for their portable instruments. Thereby the more well-characterized certified or secondary standards are available, the more reliable is the resulting calibration factor. Wilke et al. [11] have developed a calibration based on fired clay samples spiked with each four concentrations of Ti, Fe, Ga, Rb, Sr, Y, Zr and Nb oxides/salts, which subsequently have been validated with certified reference materials for confirmation of proper weighing and of the analytical purity statement given by the chemical provider. The authors argued that their clay matrix is representative for mass absorption of mid-Z elements in a broad range of clays (and other aluminosilicates) with a total matrix composition of elemental O, Al and Si greater than 90 %. The respective Fe concentration series allowed for element specific correction of Ga, Rb, Sr, Y, Zr and Nb photon absorption by Fe, and for correction of Ti photon enhancement. Furthermore, empirical influence coefficients for spectral interference correction of Rb $K\beta/Y K\alpha$ and Y $K\beta/Nb K\alpha$ [11], and of the Pb L peaks on Ga, Rb and Y [23] have been developed. There is no alternative empirical development and validation of proper interference correction algorithms available than with spiked samples, since natural geochemical reference materials with just one interfering element of variable concentration, but with the effected elements having a fixed concentration, do not exist. When using the fired spiked clay samples for validating the preinstalled calibration of selected pXRF instrument brands, it was confirmed that calibration factors of individual elements deviated by up to +/- 25%, the Ti concentrations were not corrected for Fe enhancement, and the Fe calibration curve in the concentration range relevant for natural potting clays (0.1-5% elemental Fe) was polynomial rather than linear (unpublished results by the author). Whereas simple slope correction can be made on the level of concentration values, complex interference adjustments have to be performed with Lachance-Traill type influence coefficients [24], which need access to the raw

data - spectra and deconvoluted net counts, respectively. At least for artifact provenancing coherent inter-instrument and inter-laboratory data are indispensable, since nobody can seriously claim to be in a position to establish hundreds of local reference groups to work fully self-contained in someone's individual data room.

Fit for purpose or How to make the best out of it?

In his above cited critique Killick argued "Anyone who has reviewed for JAS for as long as I have, has a list of papers that should never have been accepted for publication by the journal – either because they were fundamentally flawed or because they were trivial (analyses made for no good archaeological purpose)". This statement is for sure not specific to the Journal of Archaeological Science, published in association with the Society for Archaeological Sciences, but editors of any other renown scientific journal probably have similar black lists, even without admitting it. Indeed many papers, whether related to pXRF or any other analytical methodology, seem to have escaped the critical radar of the reviewers, perhaps since the reviewers were simply too overwhelmed by the number of manuscripts they got, or because they were not enough familiar in detail with the specific subject matter.

It looks like many authors using this and similarly complex analytics have difficulties to comply with the routines of established analytical laboratories. If scholars are unable to generate accurate data, or to properly validate their methods, this is per se no reason to reject their work, as long as the data are clearly marked as "informational", "semi-quantitative", "non-validated", "approximate" or the like. It is then up to the reviewers and editors to evaluate, whether the quality of the presented data fits for purpose, whether the "numbers" support the conclusions drawn from the study, and whether the paper presents any relevant results at all.

CONCLUSION

Non-destructive compositional analysis with XRF spectrometry provides great potential for study tasks which realistically can only be tackled, when a large number of samples is analysed - a few hundred to several thousand specimens. In archaeology and museum studies this versatility argument is as relevant as the physical integrity of the sample material for conservational reasons, and as a precondition for future analyses by new and improved methods, respectively. Compliance with established scientific standards is thereby not just an academic exaggeration but a measure to protect scholars from an investment of labour and efforts into the generation of data which are predictably insufficient. Thus when deviating from these standards, data fit for purpose must be the criterion for the acceptance of works and manuscripts.

REFERENCES

1. Tykot RH. Nondestructive portable X-ray fluorescence spectrometers on stones, ceramics, metals, and other materials in museums: advantages and limitations. *Applied Spectroscopy*, **2016**, 70: 42-56.
2. William-Thorpe O. The Application of Portable X-Ray Fluorescence Analysis to Archaeological Lithic Provenancing. In: Portable X-ray Fluorescence Spectrometry – Capabilities for In Situ Analysis. Potts PJ, West M (edtrs), Cambridge: The Royal Society of Chemistry, **2008**, 174-205.
3. Simandl GJ, Stone RS, Paradis S, Fajber R, Reid HM, et al. An assessment of a handheld X-ray fluorescence instrument for use in exploration and development with an emphasis on REEs and related specialty metals. *Mineralium Deposita*, **2014**, 49: 999-1012.
4. Colomban P, Tournié A, Maucuer M, Meynard P. On-site Raman and XRF analysis of Japanese/Chinese bronze/brass patina – the search for specific Raman signatures. *Journal of Raman Spectroscopy*, **2012**, 43: 799-808.
5. Hunt AMW. On the origin of ceramics: moving toward a common understanding of 'provenance'. *Archaeological Review from Cambridge*, **2012**, 27: 85-97.
6. Galipaud JC, Reepmayer C, Torrence R, Kelloway S, White P. Long-distance connections in Vanuata: New obsidian characterisations for the Makué site, Aore Island. *Archaeology in Oceania*, **2014**, 49: 110-116.
7. Brown FH, Nash BP. Correlation: Volcanic Ash, Obsidian. In: Archaeology and Anthropology. Cerling TE (edtr), Amsterdam: Elsevier, **2014**, Pp: 63-80.

8. Holmqvist E. Handheld portable energy-dispersive X-ray fluorescence spectrometry (pXRF). In: The Oxford Handbook of Archaeological Ceramic Analysis, Hunt A (edtr), Oxford: University Press, **2017**, 363-381.
9. Killick D. The awkward adolescence of archaeological science. *Journal of Archaeological Science*, **2015**, 56: 242-247.
10. Wilke D. Using spiked, fired clay samples for developing robust quantification algorithms for pXRF of pottery. Oral presentation on the 80th Annual Meeting of the Society for American Archaeology, San Francisco, **2015**.
11. Wilke D, Rauch D, Rauch P. Is non-destructive provenancing of pottery possible with just a few discriminative trace elements? *Sci Technol Archaeol Res*, **2016**, 2:141-158.
12. Shackley MS. Is there reliability and validity in portable X-ray fluorescence spectrometry (PXRF)? SAA Archaeological Record, **2010**, 10: 17-20.
13. Speakman RJ, Shackley MS. Silo science and portable XRF in archaeology: A response to Frahm. *Journal of Archaeological Science*, **2013**, 40: 1435-1443.
14. Reynolds RC. Matrix corrections in trace element analysis by X-ray fluorescence: Estimation of the mass absorption coefficient by Compton scattering. *The American Mineralogist*, **1963**, 48: 1133-1143.
15. Potts PJ, Webb PC, Williams-Thorpe O. Investigation of a correction procedure for surface irregularity effects based on scatter peak intensities in the field analysis of geological and archaeological rock samples by portable X-ray fluorescence spectrometry. *Journal of Analytical Atomic Spectrometry*, **1997**, 12: 769-776.
16. Brand NW, Brand CJ. Performance comparison of portable XRF instruments. *Geochemistry: Exploration, Environment, Analysis*, **2014**, 14: 125-138.
17. Piercey SJ, Devine MC. Analysis of powdered reference materials and known samples with a benchtop, field portable X-ray fluorescence (pXRF) spectrometer: Evaluation of performance and potential applications for exploration litho geochemistry. *Geochemistry: Exploration, Environment, Analysis*, **2014**, 14: 139-148.
18. Hall GEM, Bonham-Carter GF, Buchar A. Evaluation of portable X-ray fluorescence (pXRF) in exploitation and mining: Phase 1, control reference materials. *Geochemistry: Exploration, Environment, Analysis*, **2014**, 14: 99-123.
19. Conrey RM, Goodman-Elgar M, Bettencourt N, Seyfarth A, Van Hoose A, et al. Calibration of a portable X-ray fluorescence spectrometer in the analysis of archaeological samples using influence coefficients. *Geochemistry: Exploration, Environment, Analysis*, **2014**, 14: 291-301.
20. Hunt AMW, Speakman RJ. Portable XRF analysis of archaeological sediments and ceramics. *J Archaeol Sci*, **2015**, 53: 626-638.
21. Ferguson JR. X-ray fluorescence of obsidian: approaches to calibration and the analysis of small samples. In: Studies in Archaeological Sciences – Handheld XRF for Art and Archaeology, Shugar AN, Mass JL (edtrs), Leuven: Leuven University Press, **2012**, Pp: 401-422.
22. Rowe H, Hughes N, Robinson K. The quantification and application of handheld energy-dispersive x-ray fluorescence (ED-XRF) in mudrock chemostratigraphy and geochemistry. *Chem Geol*, **2012**, 324: 122-131.
23. Wilke D. Pb correction algorithms for non-destructive provenancing of lead and tin glazed slip wares. *Universal Journal of Material Science*, **2016**, 4: 125-132.
24. Lachance GR. Correction procedures using influence coefficients in X-ray fluorescence spectrometry. *Spectrochimica Acta*, **1993**, 48B: 343-357.