Measurement uncertainty in total reflection X-ray fluorescence

G.H.Floor[a,1], I.Queralt[b], M.Hidalgo[c], E.Marguí[c]

[a] GFZ German Research Centre for Geosciences Section 3.4. Earth Surface Geochemistry, Telegrafenberg, 14473 Postdam, Germany

[b] Institute of Earth Sciences Jaume Almera ICTJA-CSIC. Solé Sabaris s/n, 08028 Barcelona, Spain

[c] Department of Chemistry. University of Girona. Campus Montilivi s/n, 17071-Girona, Spain

Abstract

Total reflection X-ray fluorescence (TXRF) spectrometry is a multi-elemental technique using micro-volumes of sample. This work assessed the components contributing to the combined uncertainty budget associated with TXRF measurements using Cu and Fe concentrations in different spiked and natural water samples as an example. The results showed that an uncertainty estimation based solely on the count statistics of the analyte is not a realistic estimation of the overall uncertainty, since the depositional repeatability and the relative sensitivity between the analyte and the internal standard are important contributions to the uncertainty budget. The uncertainty on the instrumental repeatability and sensitivity factor could be estimated and as such, potentially relatively straightforward implemented in the TXRF instrument software. However, the depositional repeatability varied significantly from sample to sample and between elemental ratios and the controlling factors are not well understood. By a lack of theoretical prediction of the depositional repeatability, the uncertainty budget can be based on repeat measurements using different reflectors. A simple approach to estimate the uncertainty was presented. The measurement procedure implemented and the uncertainty estimation processes developed were validated from the agreement with results obtained by inductively coupled plasma – optical emission spectrometry (ICP-OES) and/or reference/calculated values.

Keywords: TXRF, measurement uncertainty, data quality, method validation.

1. Introduction

Total reflection X-ray fluorescence spectrometry (TXRF) is a multi-elemental and micro-analytical technique. The TXRF technique is a variation of energy-dispersive

[1] Corresponding author: geerke.floor@gfz-potsdam.de

XRF where the primary beam strikes the sample at a very small incident angle (≈ 0.1°) leading to lower scattering and an improvement of detection limits. To perform analysis under total-reflection conditions, samples must be provided as thin films [1]. For liquid samples, this is done by depositing as little as 5–50 µL of sample on a reflective carrier with a subsequent drying by applying heat or vacuum. Preparation of samples as thin layer exclude absorption and secondary excitation and thus, the quantification in TXRF analysis can be done simply by the addition of an internal standard to the sample prior to deposition [2-3].

Most of the published TXRF analyses in the last decades, were performed using large-scaled instruments with high-power X-ray tubes, demanding water-cooling systems and liquid-nitrogen cooled detectors. However, in recent years, the development and commercialization of benchtop TXRF instrumentation, which offer extreme simplicity of operation, in a low-cost compact design, have promoted its application in industry as well as in research activities for trace element analysis [4-6]. However, despite the efficiency and simplicity of the benchtop TXRF instrumentation, this technique is very little employed for environmental and geochemical analyses compared to other techniques, such as inductively coupled plasma - mass spectrometry or optical emission spectrometry (ICP-MS or ICP-OES) and Atomic Absorption Spectrometry (AAS).

To ensure the validation of the data interpretation in environmental analyses, the performance and 'fit-for-purpose' of the method should be demonstrated and sufficient information should be provided about the followed methodology [7-10]. The international standard ISO/IEC 17025 states that the performance of a measurement procedure should be evaluated based on one or a combination of the following approaches: a) the use of reference materials, b) the comparison of results achieved with other methods, c) inter-laboratory comparison, d) systematic assessments of the factors influencing the result and e) the assessment of the uncertainty of the results [11]. Therefore, in this paper, we contribute to the validation of TXRF analyses by assessing the factors contributing to the measurement uncertainty. In previous works theoretical models have been combined with the empirical uncertainty based on repetitive measurements of standards [12, 13]. However, the theoretical uncertainties and empirical model did not always give the same result. In this study, we want to present a mathematical model explaining the measurement procedure which can be applied in a practical way to samples. For that, an example of copper (Cu) and iron (Fe) in natural waters is used. Copper enters drinking water primarily through plumbing materials. Copper is highly toxic as it is

carcinogens and mutagens in nature [14-15]. Hence, the European Water Framework Directive (WFD) [16] and the guidelines of the World Health Organization (WHO) handle a maximum Cu concentration of 2 $mg \cdot L^{-1}$. Iron is an unaesthetic parameter in drinking water and therefore included in the WFD [16] as an indicator parameter at 0.2 $mg \cdot L^{-1}$. Iron plays an important role in element cycling and therefore often studied within geochemical studies [17-18]. In this manuscript, the focus is on which factors, including the sample characteristics and the measurement protocol, influence the results and the uncertainty budget of Cu and Fe determinations by TXRF to provide an example how a realistic uncertainty can be estimated in TXRF measurements.

2. Method and materials

2.1 Reagent, materials and samples

Stock solutions of 1000 ± 0.5 $mg \cdot kg^{-1}$ (Spectroscan, TECKNOLAB A/S, Norway) of Cu, Fe, Rh and Y were used to prepare standard solutions and spiked samples. High purity water used for dilution of stock solutions was obtained from a Milli-Q purifier system operated at 18 $M\Omega \cdot cm$ (Millipore Corp., Bedrod, MA).

Several water samples with different matrices have been used for analyses: the certified reference material SPS-WW2 ("Reference Material for Measurement of Elements in Wastewaters", Spectrapure Standard, Manglerud, Oslo, Norway), spiked tap/river water samples at the level of around 2 $mg \cdot kg^{-1}$ of Cu and 4 $mg \cdot kg^{-1}$ of Fe, a municipal waste water sample and a mine water sample. All samples (except the certified reference material) were filtered through a 0.45 µm cellulose acetate filter (Millipore) before TXRF analysis.

In TXRF analysis, the sample carrier plays an important role with regard to the achievement of optimal analytical results. In this work, taking into account the higher resistance and the lower background, quartz glass discs (Bruker AXS Microanalysis GmH, Berlin, Germany) with a diameter of 30 mm and a thickness of 3±0.1 mm were used as sample holders for introducing the sample into the TXRF equipment.

2.2 Instrumentation

TXRF analysis of standards and samples was performed with the benchtop S2 PICOFOX™ spectrometer (Bruker AXS Microanalysis GmbH, Berlin, Germany). The spectrometer specifications and operating conditions used are summarized in Table 1. One of the advantages of this spectrometer compared to other existing systems is

that is equipped with an air-cooled low-power X-ray tube and a Peltier cooled silicon drift detector and thus, no cooling media and gas consumption are required. As it is shown in Table 1, the anode of the X-ray tube in the TXRF instrument is made of W. This fact allows performing TXRF analysis using K-lines of high atomic number elements such as Sn and Cd (in conventional Mo-based X-ray tubes less intense L-lines have to be used for this purpose) and thus the limits of detection for heavy elements are improved. However, limits of detection for the 4th period elements (Z=19-36) are higher than those associated with Mo anode X-ray tubes (between 1 and 10 $\mu g \cdot kg^{-1}$). Limits of detection calculated for the determination of Cu and Fe using the operating conditions displayed in Table 1 were around 30 and 40 $\mu g \cdot kg^{-1}$, respectively.

For comparison, the samples were also analyzed by using an ICP-OES spectrometer Varian Liberty (Springvale, Australia) with a V-groove nebulizer of which the instrumental parameters and measurement conditions are also shown in Table 1.

2.3 TXRF measurements

Samples and standard solutions were prepared as following for TXRF analysis: an appropriate amount of a 1000 ± 0.5 $mg \cdot kg^{-1}$ Y or Rh solution (to reach a final concentration of 15 $mg \cdot kg^{-1}$) was added to 2mL of the target sample or standard for internal standardization. Afterwards, the resulting solution was thoroughly homogenized using a Vortex device and an aliquot of 10μL was transferred onto a quartz glass sample carrier and dried using an infrared lamp (T ≈ 80-100ºC) placed under a laminar flow hood. Subsequently, the sample was analyzed using a measurement time of 1000s if not otherwise indicated within the manuscript.

2.4 Data treatment

The evaluation of TXRF spectra and calculation of the analyte net peak area were performed using the provided software (Spectra Plus 5.3, Bruker AXS Microanalysis GmbH, Berlin, Germany). For the peak integration, the software applies a deconvolution routine which uses measured mono-element profiles for the evaluation of peak areas. The Fe and Cu concentrations were calculated using the equations in Table 2. The combined uncertainties indicated are expanded uncertainties $U = ku_c$ where u_c is the combined standard uncertainty and k is a coverage factor equal to 2. All intermediate steps are standard uncertainties (k=1). Combined uncertainties were obtained by propagating together individual uncertainty components according to the Guide for Uncertainty in Measurements [19] using the GUM workbench software [20].

3. Results and discussion

3.1 Mathematical description of the measurement procedure

The basic purpose of an uncertainty statement is to propose a range of possible 'true' values. There are various ways of estimating uncertainties. In the uncertainty estimation proposed in the GUM, the measurement procedure is described by a mathematical model and the values and associated standard uncertainties of the different components in the model must be established [19]. Therefore, a set of mathematical equations to describe the TXRF measurement processes is given in Table 2. In the TXRF software, the analyte concentration is calculated using the intensity ratio between the analyte (A) and the internal standard (IS), the concentration of the IS (C_{IS}) and the sensitivity factor (S) between A and IS. In equation 1 (Table 2) we also included the dilution factor (DF) arising from the addition of the IS to the samples, which has a value just above 1. The ratio between A and IS was calculated using equation 2a. The counts on A and IS are introduced as constants, not carrying an uncertainty. Instead, two unity multiplicative factors were introduced to carry standard uncertainties associated with the instrumental repeatability and depositional repeatability, which are discussed in more detail in Section 3.2. The software has build-in sensitivity factors. However, to be able to obtain their uncertainty, in this work the sensitivity factors were determined using a linear regression (equation 3a, Table 2) using triplicate measurements at 8 different levels of analytes (Cu and Fe, 25-200 $mg \cdot kg^{-1}$) with constant level of the internal standard elements (Y and Rh, ≈100 $mg \cdot kg^{-1}$). Lastly, the concentration of the IS was calculated using the concentration of the stock and the masses of respectively the IS solution and the sample (equation 4b). Additionally, this equation includes a correction for the initial IS concentration in the sample. Although the internal standard concentrations in the original sample prior to spiking were below detection limit, the initial concentrations do carry an uncertainty which was taken into account in the uncertainty budget. The masses of the IS stock solution and the sample were also used to determine the DF (equation 4a).

3.2 Assessing standard uncertainties

3.2.1 Intensity ratios

3.2.1.1 Instrumental repeatability

In Section 3.1, a unity multiplicative factor carrying the standard uncertainty associated to the instrumental repeatability was introduced. In order to test the instrumental repeatability, 10 µL of a standard solution containing ≈ 2 $mg \cdot kg^{-1}$ of Cu, ≈10 $mg \cdot kg^{-1}$ of Fe (analytes of interest), ≈10 $mg \cdot kg^{-1}$ of Rh and ≈20 $mg \cdot kg^{-1}$ of Y (elements for internal standardization) was analysed for 200, 500, 1000, 1500 and

2000 seconds. The same sample on the same reflector was used throughout this experiment, which was removed and re-introduced into the total-reflection chamber between each measurement.

The Poisson statistic on intensity level was calculated using:

$$\sigma = \sqrt{N_X + 2N_{BCK}} \quad (1)$$

in which N is the number of counts of respectively element X and its background (BCK). The standard deviation of the 20 replicates against the average counts for Fe and Cu is compared with the theoretical value obtained using Poisson statistics (Figure 1). It can be seen that although the measured standard deviation is much more scattered, the general trends are similar to those expected based on Poisson statistics.

Since the TXRF procedure uses intensity ratios, a regression between the experimental standard deviations and the calculated Poisson statistics for the different analyte/internal standard ratios was performed. A slope of 0.97±0.10 and an intercept of 0.0005±0.0007 (k=1) was obtained. Therefore, we can conclude that there is no evidence of another contribution for the range of intensities investigated and that the Poisson distribution can be used to describe the uncertainty on the unity multiplicative instrumental repeatability.

$$u_{inst}(\%) = \sqrt{\frac{\sigma^2_A}{N_A} + \frac{\sigma^2_{IS}}{N_{IS}}} \quad (2)$$

The Poisson distribution on intensity level for each element is given in the TXRF software (sigma values). However, both the analyte and internal standard contribute to the counting statistics of the ratio (equation 2). This has consequences for the optimum of internal standard added. Commonly, the internal standard is added at similar levels as the analyte of interest. In this case, the internal standard contributes for 50% to the standard uncertainty of the ratio. On the other hand, an internal standard at much higher levels than the analyte of interest can cause interferences and loss of sensitivity. Using the Poisson distribution and uncertainty propagation, it was calculated that having an internal standard intensity 3 times higher than the intensity of the analyte of interest, the counting statistics of the internal standard contributes <10 % to the standard uncertainty of the ratio. Therefore, for the analyses of samples Y and Rh were added at 3 times higher concentrations than the Fe concentration, the most abundant analyte, to obtain a concentration of ≈ 15 $mg \cdot kg^{-1}$.

3.2.1.2 Depositional repeatability

In addition to the instrumental repeatability, small geometric and homogeneity variations of the deposition on the reflector surface influence the repeatability of the analyses. In order to test the depositional repeatability and to test if there was an effect of the sample matrix, different solutions with a concentration of about 2 $mg \cdot kg^{-1}$ of Cu, 5 $mg \cdot kg^{-1}$ Fe and 15 $mg \cdot kg^{-1}$ Rh and Y were deposited on 10 different reflectors and analysed using 1000 seconds measurement time. The experimental standard deviations obtained are the combination of the instrumental and the depositional repeatability. The results in Table 3 show that the signal is strongly influenced by the deposition. Experimental standard deviations are 4 to 9 times higher than the theoretical counting statistics. It is often assumed that internal standardization corrects for such depositional effects [2]. Nevertheless, also using the ratio (element/internal standard) the experimental standard deviation (between 2.8 and 9.4%) was higher than the Poisson distribution (between 2.4 and 4.0 %). The depositional repeatability was calculated by subtracting the effect of the instrumental repeatability from the empirical repeatability based on the standard deviation:

$$u_{dep}(\%) = \sqrt{u^2_{emp}(\%) - u^2_{instr}(\%)} \quad (3)$$

It can be observed that the depositional repeatability is high for the mine water and much lower for the CRM (waste water).

In a previous work [10] the variations for multiple measurements were explained by differences in internal standard concentrations. However, we used different aliquots of the same spiked solution for the different reflectors. Therefore, the internal standard concentration is the same for each repeat measurements. Additionally, the average counts are within uncertainty the same for each solution analyses and can therefore not explain the differences in the experimental standard deviation. It was previously observed that a more complex matrix increases the background in the spectra [2, 21]. Moreover, different reflectors having an altered roughness can increment the variability of the background associated with the Compton peak [13]. In the samples, the relative background (background/net signal) ranges from 3.6% for Fe in the tap and river water up to 15% for Cu in the waste water. Nevertheless, there was no correlation between the spectral background and the experimental standard deviation. Additionally, no link between the experimental standard deviation and the presence of certain elements (e.g. Ca, Na) was present. Lastly, the depositional effects depend on the elemental ratio considered (Table 3). This all

indicates that the sample matrix alone cannot explain the depositional repeatability. Since the controlling factors of the depositional repeatability are not well understood, it is difficult to come up with one general standard uncertainty value of the unity multiplicative factor associated to the depositional repeatability.

3.2.1.3 Peak deconvolution

The peak deconvolution is performed in an automatic way by the software. In this work, it is assumed that the uncertainty associated within this process is low and taken into account by the empirical repeatability. However, in this light it is important to check for peak overlap. For this reason, Ga was discarded as potential internal standard from this study since at the concentration levels of interest it overlaps with the Cu peak (Ga-K_α: 9.251 keV, Cu-K_β:8.907 keV). Usually, the fit quality is tested using a statistical parameter named Chi, χ. This value is calculated as the standardized square sum of the differences between the measured and the calculated, deconvoluted intensities for all channels [2]. The value for the fit quality should preferably be smaller than 10. Higher values are an indication of a poor fit quality [22]. For our samples, chi values were <4, indicating that no major issues with peak overlap or peak fitting were present.

3.2.2 Internal standardization

It is assumed that when working under conditions of total reflection, the elemental intensity ratios measured are directly proportional to the mass fractions of elements present. Therefore, quantification, can be done by using a single set of relative sensitivities and a known concentration of an internal standard (see equation 1 in Table 2). Additionally, as discussed above, internal standardization helps to eliminate some of the variations caused by effects from deposition on the reflector.

In this work the sensitivity factors were determined using a linear regression (equation 3a, Table 2) for different analyte/internal ratios. The standard uncertainties of the sensitivity factors were obtained using the statistical uncertainty of the regression between N_E and $N_{RF}*C_E*C_{RF}$ using 24 data points (8 levels x 3 replicates), as for example described in [23, 24]. The resulting standard uncertainties (Table 3) were between 1.3 and 1.8%.

Different factors contribute to the uncertainty on the concentration of the IS (equation 4b, Table 2). Firstly, the addition of the internal standard can be performed by volume or by weight. When using an analytical balance the contribution of the weighing error to the total combined uncertainty using our dataset was <0.1%. When

volumetric rather than gravimetric dilution data were used, uncertainties of volumes contribute several percent to the total combined uncertainty. Secondly, any IS present in the original sample will contribute to the final concentration in the spiked samples. Therefore, all samples were tested without any addition of IS. As an example, in Figure 2 TXRF spectra for the certified reference material SPS-WW2 (waste water) with and without addition of Y and Rh as internal standards is displayed. It was found that for our sample the signal was <0.2% than the signal in the spiked samples. Therefore, the correction for the signal deriving from the original sample causes <0.05% uncertainty on the measured ratio (k=1) and is therefore negligible compared to the effect of other factors. Thirdly, the stated uncertainty (5 $mg \cdot kg^{-1}$) on the concentration (1000 $mg \cdot kg^{-1}$) of the stock solution has to be taken into account.

In this study both Y and Rh were used for internal standardization. The results and uncertainties (see Section 3.2 and Table 4) do not show a systematic difference between the IS. Both elements were present below TXRF detection limits in our samples prior to spiking (Figure 2). However, since Y concentration is general more abundant in natural samples [25], it might be better to use Rh. Yet, the proper choice of the internal standard should be carried out for each analysed sample.

3.3 Measurement uncertainty

The uncertainty budget was calculated using the equations and standard uncertainties as described above and the GUM workbench software for uncertainty propagation. The total combined uncertainty (coverage factor k=2, corresponding to a confidence interval of 95%) was between 6 and 19% (Table 4). The uncertainty budget for Cu in the municipal waste and mine water samples is illustrated in Figure 3. For all samples and both analytes, the ratio repeatability (assessed by measuring the same solution deposited at different reflectors) is the main contributor to the uncertainty budgets. However, the relative contribution of the instrumental and depositional varies. For example, for the SPS-WW2 the main contributor is the instrumental repeatability, while for the mine water the depositional repeatability is the main contribution (Figure 3). The sensitivity factor contributes up to 20% to the combined expanded uncertainty. The concentration of the IS present in the sample (i.e. weighing, concentration of the stock and the original IS concentration in the samples) seems only a minor contribution to the total combined uncertainty for our measurements.

The provided sigma-concentration by the TXRF software, which is based on the counting statistics of the analyte, is much lower compared to the combined expanded uncertainty, since the counting statistics of the analyte just contribute between 23-70% to the total combined uncertainty. Moreover, in the case the internal standard has similar counts to the analyte, the software will further underestimate the uncertainty by at least a factor of 2. This could be one of the reasons why intercomparison studies lead to statistically different results [9]. It would be beneficial if TXRF software could take the effect of the counting statistics of the internal standard and the uncertainty of the sensitivity factor into account. However, at present it will be a major challenge to include the depositional repeatability in the software, since it varies from sample to sample and it is unclear which factors influence this uncertainty.

Therefore, an alternative approach was tested for the waste and mine water sample in which the empirical standard uncertainty of the repeatability (which includes both the instrumental and depositional repeatability) was calculated based on 3 measurements on 3 different reflectors. The uncertainty budgets using Rh as an internal standard are given in Figure 4. It can be observed that the concentrations results are the same within uncertainty as the previous uncertainty budget. However, uncertainties are lower since the repeatability was divided by the square root of the number of independent measurements. Additionally, the contribution of the sensitivity factor does increase following this approach. This approach allows the estimation of the uncertainty without understanding the depositional repeatability. Nevertheless, as this depositional repeatability is 9-68% of the total combined uncertainty, future research on the factors influencing the depositional repeatability and ways how to lower this uncertainty would be a major step forward in improving the uncertainty for TXRF analyses.

3.4 Accuracy of TXRF

The obtained values using TXRF could be compared with ICP-OES analyses and reference or calculated values, the latter based on the amount of spike added (Table 4). In general there is no evidence for bias since TXRF and ICP-OES or reference/calculated values and are the same at a 95% confidence interval. However, for the spiked tap water, the Cu content is higher than the theoretical value. Given the good results for the other samples, we believe this is related to presence of Cu in the non-spiked matrix. In summary, there is no evidence that one of the main factors influencing the uncertainty was not taken into account in this study.

3.5 Practical approach to obtain the measurement uncertainty

A detailed evaluation of the uncertainty budget, showed that 2 or 3 parameters form >95% of the uncertainty budget. Therefore, the relative combined expanded uncertainty ($U_{Ca}(\%)$), can be approximated by combining the uncertainty of the instrumental repeatability (u_{inst}), the deposition ($u_{dep)}$ and the sensitivity factor (u_S), using a simple set of equations:

$$U_{C_A}(\%) \approx 2 \cdot \sqrt{u^2_{inst} + u^2_{dep} + u^2_S} = 2 \cdot \sqrt{u^2_{emp} + u^2_S} \quad (4)$$

As discussed above the standard uncertainty on the instrumental repeatability can be described using the Poisson distribution (equation 2). This are all parameters which can be easily deducted from the software. The standard uncertainty on the depositional repeatability can be determined using the difference between the empirical repeatability (standard deviation of replicate measurements) and the instrumental repeatability (equation 3). Alternatively, the instrumental and depositional repeatability can be taken together if the multiple measurements were performed for one sample on different reflectors. In this case, the standard empirical uncertainty of the repeatability (u_{emp}) can be determined by the standard deviation divided by the square root of the number of replicates (n).

For the uncertainty on the sensitivity factor (u_S), a regression between the analyte and the internal standard should be performed. In principle, this only needs to be performed once for each TXRF system. General statistics can then be used to determine the slope and its relative uncertainty.

It is important to realize that this approximation of the uncertainty budget is only valid if the concentration of the internal standard does not significantly contribute to the uncertainty budget. This means that the internal standard additions should be performed by weight by a calibrated balance, that the presence of the IS in the non-spiked sample should be checked and that the concentration of the IS stock solution should have a low uncertainty. Although this is not a complete uncertainty budget, this practical approach provides a more realistic uncertainty estimate compared to the sigma value of the analyte provided by most softwares.

4. Conclusion

The parameters contributing to the measurement of uncertainty are summarized in Figure 5. The instrument variability seems to be mainly controlled by counting statistics for studied range (up to 5000 counts). In order to avoid that the counting

statistics of the internal standard significantly contribute to the total combined uncertainty, it is recommended that the concentration of the internal standard is about 3 times higher than the analytes of interest. Nevertheless, the empirical repeatability of the ratio is higher than the Poisson distribution. Although it is generally believed that depositional effects can be eliminated by normalisation to an internal standard, in this study it is shown that the ratio repeatability using different reflectors is higher than expected based on Poisson statistics. The factors influencing this depositional repeatability could not be determined. In order to get a realistic uncertainty budget with one replicate, future work should try to determine the controlling factors of the ratio repeatability. Moreover, since the depositional repeatability is a main contribution to the uncertainty budget, a mechanistic understanding of the depositional effects could potentially be used to decrease the uncertainty associated to TXRF measurements. By a lack of theoretical prediction of the depositional repeatability, the uncertainty budget can be based on replicate measurements (using different reflectors). Another important contribution to the uncertainty is the sensitivity factor, which contributes up to 50%. Currently, most software provides the sigma-concentration, which reflects the influence of the counting statistics of the analyte on uncertainty. However, this is not a good representation of the total combined uncertainty. Therefore, we provided a simple approach to obtain a more realistic uncertainty budget.

Acknowledgments

This work has been funded by the Spanish National Research Programme (CGL-2010-22168-C03-01).

References

[1] C.Streli, Recent advances in TXRF, Appl. Spectrosc. Rev. 41 (2006) 473-489.

[2] R. Klochenkämper, Total reflection X-ray fluorescence analysis, vol. 140 in Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications, John Wiley & Sons, New York, 1997.

[3] E.Marguí, R.Van Grieken, X-ray fluorescence spectrometry and related techniques: An introduction, Momentum Press, New York, 2013.

[4] E.Marguí, J.C.Tapias, A.Casas, M.Hidalgo, I.Queralt, Analysis of inlet and outlet industrial wastewater effluents by means of benchtop total reflection X-ray fluorescence spectrometry, Chemosphere 80 (2010) 263-270.

[5] E.Marguí, I.Queralt, M.Hidalgo, Determination of platinum group metal catalyst residues in active pharmaceutical ingredients by means of total reflection X-ray spectrometry, Spectrochim. Acta B 86 (2013) 50-54.

[6] S.Kunimura, D.Watanabe, J.Kawai. Optimization of a glancing angle for simultaneous trace elemental analysis by using a portable total reflection X-ray fluorescence spectrometer. Spectrochimica Acta Part B 64 (2009) 288-290.

[7] P. Deines, S.L. Goldstein, E.H. Oelkers, R.L. Rudnick, L.M. Walter. Standards for publication of isotope ratio and chemical data in Chemical Geology, Chemical Geology 202 (2003) 1-4.

[8] H. Egli, M. Dassenakis, H. Garelick, R. van Grieken, W. J. G. M. Peijnenburg, L. Klasinc, W. Kördel, N. Priest, T. Tavares. Minimum requirements for reporting analytical data for environmental samples (IUPAC Technical Report).Pure and Applied Chemistry 75 (2009): 999-1155

[9] P.J. Potts, Proposal for the Publication of Geochemical Data in the Scientific Literature. Geostandards and Geoanalytical Research 36 (2012) 225–230

[10] L. Borgese, F. Bilo, K. Tsuji, R. Fernández-Ruiz, E. Margui, C. Streli, G. Pepponi, H. Stosnach, T. Yamada, P. Vandenabeele, D.M. Maina, M. Gatari, K.D. Shepherd, E.K. Towett, L. Bennun, G. Custo, C. Vasquez, L.E. Depero. First Total Reflection X-Ray Fluorescence round-robin test of water samples: Preliminary results, Spectrochimica Acta Part B: Atomic Spectroscopy 101 (2014): 6-14

[11] ISO/IEC, General requirements for the competence of testing and calibration laboratories, ISO/IEC 17025:2005, International Organization for Standardization/International Electrotechnical Commission, Geneva, 2005.

[12] O.K. Owoade, F.S. Olise, H.B. Olaniyi, D. Wegrzynek. Model estimated uncertainties in the calibration of a total reflection x-ray fluorescence spectrometer using single-element standards. X-Ray Spectrometry 35 (2006) 249-252.

[13] R. Fernández-Ruiz. Uncertainty in the multielemental quantification by total-reflection X-ray fluorescence: theoretical and empirical approximation. Analytical chemistry 80.22 (2008): 8372-8381.

[14] B.P. Zietz, H.H. Dieter, M. Lakomek, H. Schneider, B. Keßler-Gaedtke, B., H. Dunkelberg. Epidemiological investigation on chronic copper toxicity to children exposed via the public drinking water supply. Science of the total environment 302 (2003), 127-144.

[15] T. Theophanides, J. Anastassopoulou, Copper and carcinogenesis, Critical Reviews in Oncology/Hematology 42 (2002): 57-64,

[16] Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, Official Journal of the European Communities L 327/1, 22/12/2000.

469 [17] T. D. Jickells, Z. S. An, K. K. Andersen, A. R. Baker, G. Bergametti, N. Brooks, J. J. Cao, P.
470 W. Boyd, R. A. Duce, K. A. Hunter, H. Kawahata, N. Kubilay, J. laRoche, P. S. Liss, N.
471 Mahowald, J. M. Prospero,A. J. Ridgwell, I. Tegen, R. Torres. Global Iron Connections between
472 desert dust, ocean biogeochemistry and climate. Science 308 (2005): 67-71.

473 [18] A.B. Cundy, L. Hopkinson, R.L.D. Whitby. Use of iron-based technologies in contaminated
474 land and groundwater remediation: A review. Science of The Total Environment 400 (2008)42-
475 51
476

477 [19] JCGM 100, Evaluation of measurement data — Guide to the expression of uncertainty in
478 measurement; Joint Committee for Guides in Metrology (BIPM, IEC, IFCC, ISO, IUPAC, IUPAP,
479 OIML, ILAC), BIPM, www.bipm.org, Paris, 2008.

480 [20] Metrodata GmbH, 2003. Gum Workbench. The software tool for the expression of
481 uncertainty in measurement. Metrodata GmbH, D-79639 Grenzach-Wyhlen, Germany.

482 [21] E. Marguí, G. H. Floor, M. Hidalgo, P. Kregsamer, G. Román-Ross, C. Streli and I. Queralt.
483 Applicability of direct total reflection X-ray fluorescence analysis for selenium determination in
484 solutions related to environmental and geochemical studies. Spectrochimica Acta B 25 (2010):
485 1002-1007

486 [22] S2 PICOFOX™ User Manual. Bruker advanced X-ray solutions. Bruker AXS Microanalysis
487 GmH, Berlin (Germany), 2007.

488 [23] Miller, J.N., 1991. Basic statistical methods for analytical chemistry. Part 2. Calibration
489 and regression methods. A review. Analyst 116: 3-14.

490 [24] Press, W.H., 2012. Numerical Recipes in C: The Art of Scientific Computing.

491 [25] K. Hans Wedepohl. The composition of the continental crust. Geochimica et cosmochimica
492 Acta 59 (1995) 1217-1232.

493 [26] http://www.lgcstandards.com/CA/en/Waste-water-Trace-metals/p/SPS-WW2. Accessed
494 16-06-2015.

495

496

497

498 **Table 1.**

499 Instrumental parameters and measurement conditions

S2 PICOFOX TXRF benchtop spectrometer	
X-Ray tube	W
Rating	50 kV, 1mA (maximum power 50 W)
Optics	Multilayer monochromator (33.0 keV)
Detector	Si drift detector, 10mm^2, <160eV resolution Mn-K_α
Working environment	Air
Sample station	Cassette changer for 25 samples
Size, weight	600x300x450mm, 37kg
Measurement time	1000 s
Analytical lines	Cu, Fe, Y, Rh: K-lines
Varian Liberty RL ICP-OES spectrometer	
Element wavelength	Fe:238.204nm, Cu:324.754nm
RF power	1000 W
Plasma gas flow rate	12 L min^{-1}
Auxiliary gas flow rate	1.5 L min^{-1}
Nebulizer	V-groove

500

501

Table 2: mathematical model of the measurement procedure

Key equation
Concentration of the analyte using internal standardization 1: $C_A = R \cdot \frac{C_{IS} \cdot DF}{S}$
Intensity ratio
Ratio between analyte and internal standard (1 measurement): 2a: $R = \frac{N_A}{N_{IS}} \cdot \delta_{inst} \cdot \delta_{dep}$
Unity multiplicative factor arising with instrumental repeatability: 2b: $\delta_{inst} = 1$
Unity multiplicative factor arising with depositional repeatability: 2c: $\delta_{dep} = 1$
Sensitivity factor
Least Square regression of N_E versus $N_{RF}*C_E*C_{RF}$: 3a: $S = \frac{n \cdot \sum \left[N_{RF\_i} \cdot \frac{C_{E\_i}}{C_{RF\_i}} \cdot N_{E\_i} \right] - \sum \left[N_{RF\_i} \cdot \frac{C_{E\_i}}{C_{RF\_i}} \right] \cdot \sum N_{E\_i}}{n \cdot \sum \left[N_{RF\_i} \cdot \frac{C_{E\_i}}{C_{RF\_i}} \cdot \right]^2 - \left(\sum N_{RF\_i} \cdot \frac{C_{E\_i}}{C_{RF\_i}} \cdot \right)^2}$
Concentration internal standard
Dilution factor of the solution 4a: $DF = \left(\frac{m_{smp} + m_{IS}}{m_{smp}} \right)$
Concentration of the IS 4b: $C_{IS} = C_{Stock} \frac{m_{IS}}{m_{smp}} + C_0$

Parameter		Index	
C	Concentration (mg·kg^{-1})	A	Analyte of interest (in sample)
N	Intensity (counts)	IS	Internal standard (in sample)
R	Intensity ratio	E	Element of interest (for determination of sensitivity, same element as A)
S	Sensitivity factor	RF	Reference element (for determination of sensitivity, same element as IS)
m	Mass (in kg)	smp	sample
DF	Dilution factor	stock	Stock solution
δ	Multiplicative unity factors	0	Sample without spike
		x,i	Generic indices

Table 3: input parameters for the set of equations of Table 2. An uncertainty of 0.2 mg was taken for the weighing, the uncertainty of the stock solution of the internal standards was 5 mg·kg^{-1} (k=1).

Sample	Parameter		Signal		Ratio				Sample prep; weight in g		
			Fe	Cu	Fe/Y	Fe/Rh	Cu/Y	Cu/Rh	m_{sample}	m_Y	m_{Rh}
	Sensitivity factor	Value			0.267	0.318	0.446	0.531			
		u(k=1)			0.004	0.005	0.006	0.010			
CRM SPS-WW2	Intensity		1513	1022	0.081	0.096	0.055	0.065	2.964	0.046	0.046
	Empirical rep. (%)		25.6	26	2.8	4.5	2.7	3.5			
	Instrumental rep. (%)		2.8	3.5	2.7	2.7	3.3	3.3			
	Depositional rep. (%)		25.4	25.8	0.7	3.6	<0	1.2			
Spiked tap water	Intensity		1618	1377	0.078	0.093	0.065	0.079	3.112	0.046	0.044
	Empirical rep. (%)		20.5	18.6	7.4	6.9	4.4	3.8			
	Instrumental rep. (%)		2.8	3.1	2.4	2.4	2.6	2.6			
	Depositional rep. (%)		20.3	18.3	7.0	6.5	3.5	2.8			
Spiked river Water	Intensity		1589	1197	0.080	0.092	0.060	0.069	3.059	0.045	0.046
	Empirical rep. (%)		12.5	12.4	4.7	4.5	5.2	5.2			
	Instrumental rep. (%)		2.8	3.3	2.6	2.6	2.9	3.0			
	Depositional rep. (%)		12.2	12.0	3.9	3.7	4.3	4.2			
Waste water	Intensity		1555	1201	0.070	0.082	0.054	0.063	15.65	0.235	0.236
	Empirical rep. (%)		13	12.7	4.9	4.5	4.2	4.4			
	Instrumental rep. (%)		2.4	2.9	3.2	3.3	3.9	4.0			
	Depositional rep. (%)		12.8	12.4	3.7	3.1	1.6	1.8			
Mine water	Intensity		1600	1271	0.076	0.087	0.061	0.069	15.12	0.233	0.235
	Empirical rep. (%)		10.3	11.7	9.4	9.1	7.6	6.4			
	Instrumental rep. (%)		2.7	3.1	2.6	2.9	2.9	3.2			
	Depositional rep. (%)		9.9	11.3	9.0	8.6	7.0	5.5			

Table 4. TXRF results for the analysis of different types of water samples. Concentrations ($mg \cdot kg^{-1}$) obtained by ICP-OES and theoretical values (added amount of analyte to the sample matrix or consensus values)[26] are given for comparison.

	Analyte	TXRF						ICP-OES	References/ Calculated
		IS=Rh			IS=Y				
		Value	U (k=2)		Value	U (k=2)			
CRM SPS-WW2	Cu	1.96	0.16	8.2%	1.96	0.14	7.1%	1.9±0.1	2.0±0.01
	Fe	4.85	0.47	9.7%	4.70	0.30	6.4%	4.8±0.3	5.0±0.025
Spiked tap water	Cu	2.18	0.19	8.7%	2.23	0.21	9.4%	-	1.87
	Fe	4.31	0.61	14.2%	4.26	0.64	15.0%	-	4.53
Spiked river water	Cu	2.03	0.22	10.8%	2.06	0.22	10.7%	-	1.90
	Fe	4.49	0.43	9.6%	4.42	0.44	10.0%	-	4.58
Waste water	Cu	1.91	0.18	9.4%	1.91	0.17	8.9%	1.8±0.1	-
	Fe	4.13	0.40	9.7%	4.01	0.41	10.2%	4.3±0.2	-
Mine water	Cu	2.16	0.29	13.4%	2.23	0.34	15.2%	2.0±0.1	-
	Fe	4.55	0.84	18.5%	4.54	0.86	18.9%	4.7±0.2	-

FIGURE CAPTIONS

Figure 1. Comparison of measured standard deviation based on 20 replicate measurements of the same reflector (Cu ≈ 2 mg·kg^{-1} and Fe ≈ 10 mg·kg^{-1}) and Poisson statistics.

Figure 2. TXRF spectra for the certified reference material SPS-WW2 (Waste water, 2 and 5 mg kg^{-1} of Cu and Fe) with and without addition of Y and Rh as internal standards (15 mg kg^{-1}).

Figure 3. Relative contributions to the Cu uncertainty budget using the instrumental and depositional repeatability and Rh as internal standard the waste and mine water sample.

Figure 4. Relative contributions to the Cu and Fe uncertainty budget using the empirical repeatability and Rh as internal standard for the waste water and mine water sample.

Figure 5. Ishikawa diagram for the TXRF measurement procedure.

Figure 1

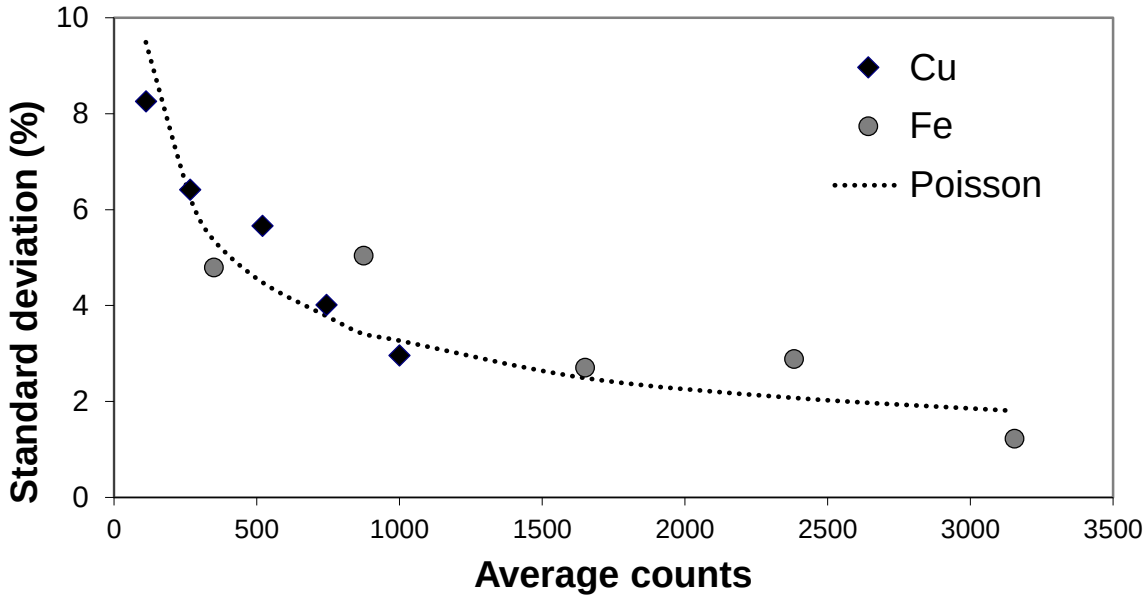


Figure 2

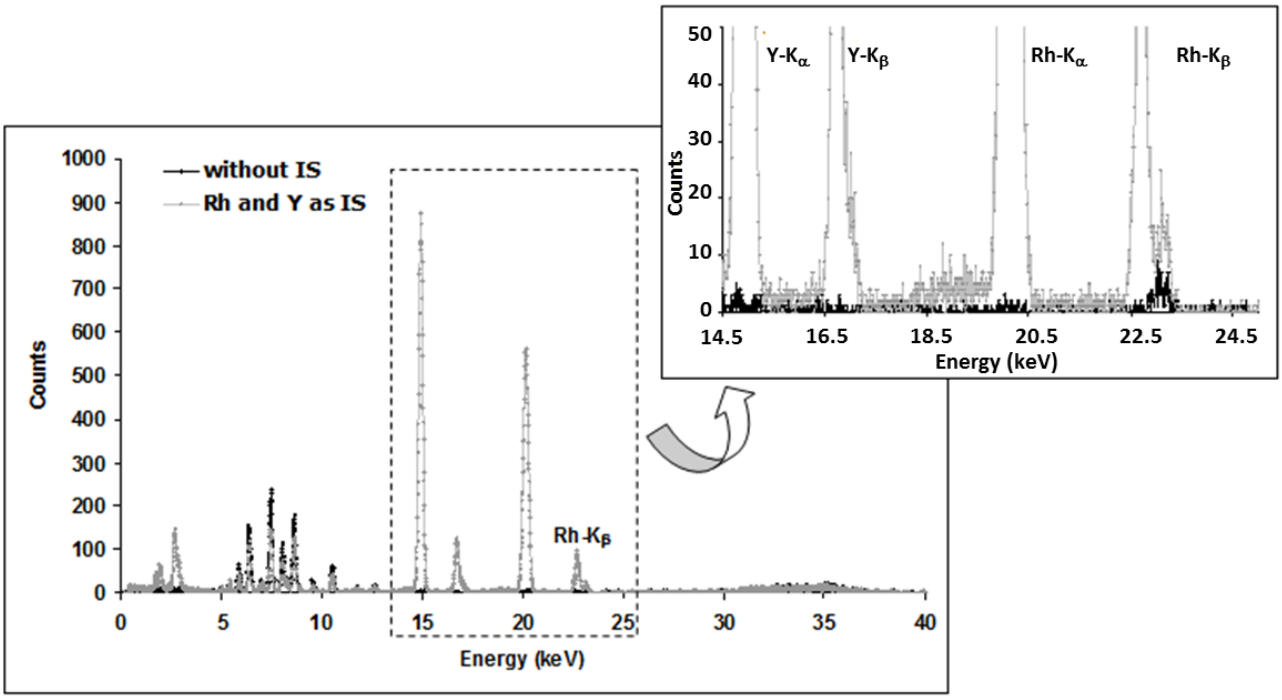

without IS
Rh and Y as IS
Counts
Energy (keV)
Rh-Kβ
Y-Kα
Y-Kβ
Rh-Kα
Rh-Kβ

Figure 3

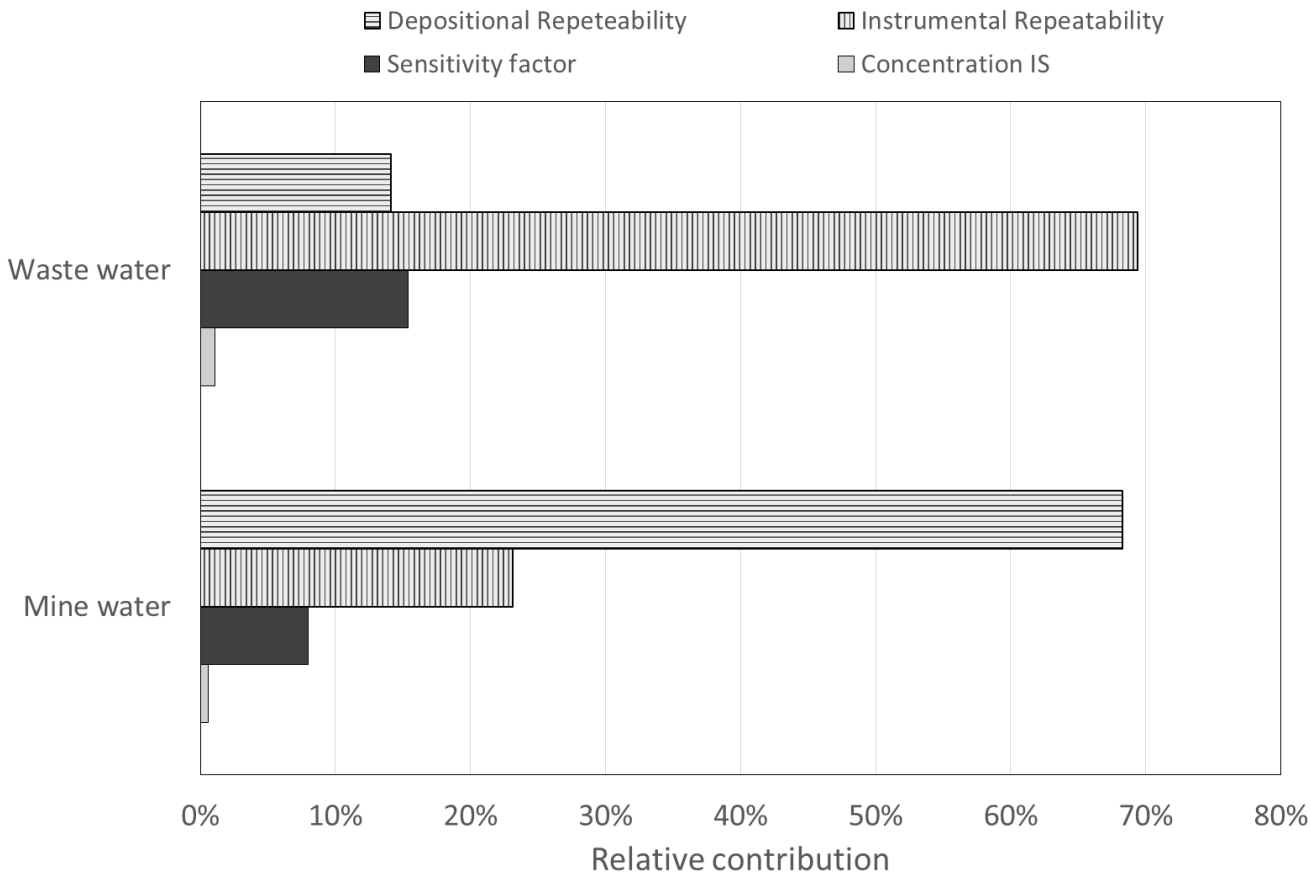

Depositional Repeteability
Instrumental Repeatability
Sensitivity factor
Concentration IS
Waste water
Mine water
0%
10%
20%
30%
40%
50%
60%
70%
80%
Relative contribution

Figure 4

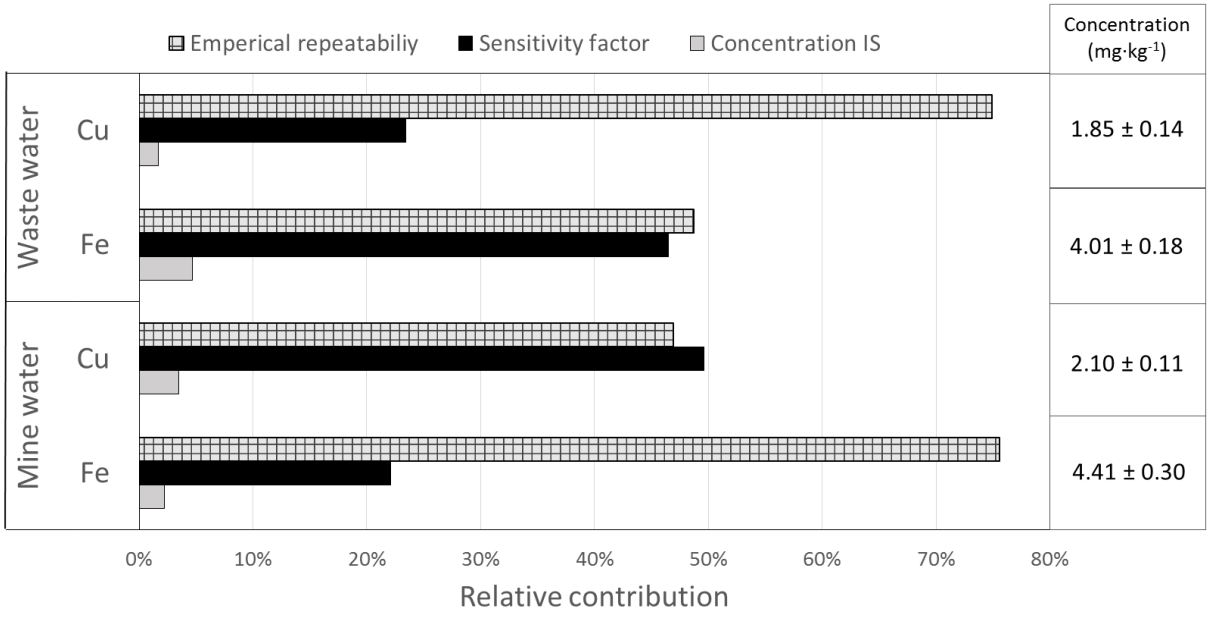
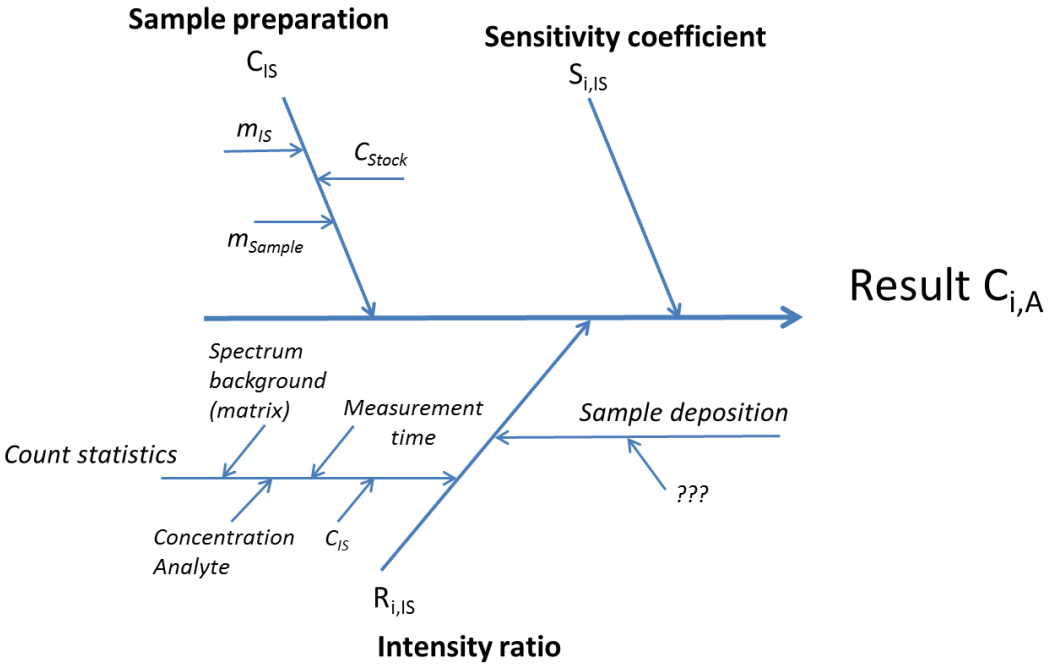

Emperical repeatabiliy
Sensitivity factor
Concentration IS
Concentration (mg·kg-1)
Waste water
Cu
Fe
Mine water
Cu
Fe
1.85 ± 0.14
4.01 ± 0.18
2.10 ± 0.11
4.41 ± 0.30
0%
10%
20%
30%
40%
50%
60%
70%
80%
Relative contribution

Figure 5


Sample preparation
Sensitivity coefficient
C_{IS}
$S_{i,IS}$
m_{IS}
C_{Stock}
m_{Sample}
Result $C_{i,A}$
Spectrum background (matrix)
Measurement time
Sample deposition
Count statistics
???
Concentration Analyte
C_{IS}
$R_{i,IS}$
Intensity ratio