Available online at www.sciencedirect.com
SCIENCE DIRECT®

Spectrochimica Acta Part B 59 (2004) 185–197

SPECTROCHIMICA
ACTA
PART B

www.elsevier.com/locate/sab

Validation of an X-ray methodology with environmental concern

P. Morgenstern, L. Brüggemann, R. Wennrich*

UFZ-Centre for Environmental Research, Leipzig-Halle, Department of Analytical Chemistry, Permoserstrasse 15, D-04318 Leipzig, Germany

Received 24 March 2003; accepted 11 November 2003

Abstract

The capabilities of test procedures toward the validation of X-ray fluorescence analysis with environmental concern were investigated. The applied analytical scheme was based on the combination of fusion and pelletizing technique for sample preparation, while matrix correction was performed by means of the fundamental parameter and the scattered radiation method. Validation was carried out in terms of precision, trueness, measurement uncertainty, limit of detection and test for homogeneity of the sample material. To confirm trueness for the examined calibration strategy, the bias of the recovery functions for each analyte under study was evaluated by means of orthogonal regression analysis. The investigations were focused on single laboratory validation. Consequently, the presented validation procedures were based nearly exclusively on certified reference material.

Keywords: X-ray fluorescence analysis; Environmental samples; Method validation; Precision studies; Estimation of trueness

1. Introduction

Introducing aspects of quality assurance into the fields of research and development is a frequently discussed matter. These discussions apparently suffer from the misapplication of well-established quality systems, developed for routine processes, to non-routine processes [1].

X-ray fluorescence (XRF) analysis is meanwhile a well-established method in the field of environmental research. It is frequently used, e.g.

- for screening of contaminated sediments and soils to investigate anthropogenic impacts;
- to get an insight in both the sources of environmental pollution and the distribution of analytes in contaminated areas;
- to examine and optimize remediation processes.

The validation of the analytical method is nevertheless a challenge for the analyst. The main problems are related to the large ranges of analyte concentrations as well as unexpected changes of the sample morphology and matrix composition, which can occur within the investigated material.

When XRF is applied to the analysis of solids, the response of the spectrometer is strongly influenced by the matrix composition of the sample. Consequently, the functional relationship between the measured fluorescence intensity I_i^{mea} and the analyte concentration C_i leads to a multivariate problem:

$$I_i^{\mathrm{mea}} \approx f(C_1,\ C_2,\ldots C_i,\ldots C_n) \qquad (1)$$

To start a successful calibration routine it is generally recommended to apply at first a suitable matrix correction algorithm to the raw data to calculate matrix corrected intensities I_i^{corr}. Under the best circumstances such transformation procedures restore (although commonly only for a restricted range of analyte concentrations and matrix compositions) linear relationships between the corrected fluorescence intensity and analyte concentration, so that finally for each individual analyte univariate relations of the type

$$I_i^{\mathrm{corr}} \approx a_i + b_i C_i \qquad (2)$$

are available, where b_i and a_i represent the element-specific sensitivity and the intercept of the individual calibration lines, respectively.

---

*Corresponding author. Tel.: +49-341-235-2915; fax: +49-341-235-2625.

E-mail addresses: rainer.wennrich@ufz.de (R. Wennrich), morgen@ana.ufz.de (P. Morgenstern).

0584-8547/04/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.sab.2003.11.007

In this context the present study deals with the use of tools of regression analysis as a reliable base to quantify the performance capabilities of the analytical method with regard to the principles of fitness for purpose [2].

2. Validation concept related to the used calibration strategy

2.1. Used validation parameters

In the field of XRF analysis a calibration is based in general on a linear regression between certified concentration values and the matrix corrected fluorescence intensities I_i^{corr} (Eq. (2)). The validation parameters for the calibration defined in the German Standards DIN32645 [3] or DIN51418-2 [4] are consequently expressed in terms of measured and calculated fluorescence intensities.

Commercial XRF spectrometers commonly do not explicitly display the intensity of the matrix corrected signal I_i^{corr} but rather the concentration values C_i^{cal}, calculated on the base of the evaluation function

$$C_i^{\mathrm{cal}}=\frac{I_i^{\mathrm{corr}}-a_i}{b_i}. \tag{3}$$

With respect to this, we re-arranged the corresponding expressions, given in the standards DIN32645 or DIN51418-2, in consideration of Eq. (3). The estimation of the validation parameters finally can be attributed as desired to a linear regression based on certified and calculated concentration values (Eqs. (4)–(6)).

2.1.1. Residual standard deviation

$$s_{e,i}=b_i\sqrt{\frac{\sum_{j=1}^{n}(C_{i,j}^{\mathrm{cal}}-C_{i,j}^{\mathrm{cert}})^2}{n-p}} \tag{4}$$

where n is the total number of calibration samples, p is the number of coefficients according to the type of regression (linear: $p=2$; squarely: $p=3$), $C_{i,j}^{\mathrm{cal}}$ is the calculated concentration for analyte i in the calibration sample j (based on the evaluation function Eq. (3)) and $C_{i,j}^{\mathrm{cert}}$ is the certified concentration value for analyte i in the calibration sample j.

The expression for the analyte-specific residual standard deviation $s_{e,i}$ Eq. (4) represents the integral influence of the individual uncertainty components of the laboratory, e.g. precision variation, sample preparation and also matrix correction. Its estimation based on a regression analysis of the certified concentrations of the calibration samples ($C_{i,j}^{\mathrm{cert}}$) vs. the corresponding concentrations ($C_{i,j}^{\mathrm{cal}}$) was calculated by means of the evaluation function (Eq. (3)).

2.1.2. Combined measurement uncertainty

$$u_i(C_x)=\frac{s_{e,i}}{b_i}\sqrt{1+\frac{1}{n}+\frac{(C_{i,x}-C_{i,m})^2}{\sum_{j=1}^{n}(C_{i,j}^{\mathrm{cert}}-C_{i,m})^2}} \tag{5}$$

where $C_{i,m}$ is the mean of the analyte concentration used in the calibration $(1/n(\sum_{j=1}^{n}C_{i,j}))$ and $C_{i,x}$ is the mean of the calculated concentration value for analyte i in the unknown sample.

When calculating the measurement uncertainty u_i (expressed as standard deviation) by means of Eq. (5), the term $s_{e,i}/b_i$ proves to be as the most significant influence factor. Because it reflects the integral influence of all uncertainty components on the analytical result, Eq. (5) represents an approach for the combined measurement uncertainty (Eurachem Guide E3.5 [5]), taking into account all influence factors associated with the actual concentration range and matrix composition, the used matrix correction algorithm and the applied technique for sample preparation.

2.1.3. Limit of detection

In the present paper the performance parameter limit of detection (LOD) is focused on the definition, reported in the German DIN 32645 [3]. Consequently, the amount of LOD ($C_{\mathrm{LOD},i}$) for analyte i was calculated to

$$C_{\mathrm{LOD},i}=t_{f,\alpha}\frac{s_{e,i}}{b_i}\sqrt{1+\frac{1}{n}+\frac{C_{i,m}^2}{\sum_{j=1}^{n}(C_{i,j}^{\mathrm{cert}}-C_{i,m})^2}} \tag{6}$$

where $t_{f,\alpha}$ is the critical value of the one-sided t-distribution at $f=n-p$ degrees of freedom and significance level α.

The use of LOD according to Eq. (6) in principle is consistent with the recommendations of the IUPAC [6]. Only the applied procedures for the estimation of the critical values are different. The concept of IUPAC is based on the evaluation of the measurements of blanks. In contrast to this, the used option of DIN32645 refers to the parameters of the calibration function and the confidence band belonging to it. Consequently, the LOD values recommended by DIN32645 reflect to a higher degree the experimental conditions of the measurement as the corresponding values that base on the measurements of blanks only. However, to justify the use of the approach (Eq. (6)), one has to ensure both homoscedasticity of the calibration data and linearity of the calibration function. In addition to this, it is imperative that the calibration samples were prepared individually, i.e. independent of each other.

2.2. Estimation of trueness

Trueness is the closeness of agreement between a test result and the accepted value of a certified material and is quantified in terms of bias [7,8].

If an analytical problem refers to the determination of analytes under well-defined matrix conditions and appropriate reference material is available, then trueness is simply given in terms of the difference between the mean of the experimental result and the certified value.

Trueness will then be evident if the relation

$$-a_2-2\sigma_R < C_m-\mu < a_1+2\sigma_R, \tag{7}$$

where σ_R is precision of the measurement under reproducibility conditions, C_m is average of replicate measurements, μ is certified concentration, a_1, a_2 are adjustment values chosen by the user with regard to economic or technical limitations, is fulfilled.

In contrast to this, when validating an analytical method in consideration of the whole range of calibration, it is imperative to demonstrate the absence of systematic errors not only for a single reference material but also with regard to an extended range of concentrations including also variations of the sample matrix, i.e. one has to ensure that the investigated calibration strategy is not biased over the whole range. For this purpose the parameters b_0 and b_1 of the individual recovery functions

$$C_i^{\text{cal}}=b_{0,i}+b_{1,i}C_i^{\text{cert}} \tag{8}$$

were estimated by means of regression analysis of the certified concentrations of the calibration samples (C_i^{cert}) vs. the corresponding concentrations (C_i^{cal}) calculated by means of the evaluation function.

With respect to this the absence of systematic errors is indicated, when no statistical significant deviations of the slope from 1 (proportional systematic error) and no intercept different from 0 (constant systematic error) will be detected in the recovery function Eq. (8). In accordance with that the following conditions must be fulfilled [9]:

$$b_0-\text{conf}(b_0)<0<b_0+\text{conf}(b_0) \quad b_1-\text{conf}(b_1)<1<b_1+\text{conf}(b_1). \tag{9}$$

Because the estimates of slope b_1 and intercept b_0 are not independent (correlation between b_1 and b_0), it is more convenient to take into account a joint confidence region, which takes the form of an ellipse. As a consequence, all sets of (b_0, b_1), accepted by the joint test, fall within an ellipse with the center at (b_1, b_0) [10].

The common practice that the concentration values of the certified reference material will be taken into account without error is no longer true in the investigated methodology. This is on the one hand because the analytes under study are presented under significantly changed matrix conditions in the available reference materials and on the other hand it must be emphasized that quite different analytical methods and procedures of sample preparations are involved within the collaborative validation processes.

We focused on both an orthogonal distance regression routine that takes into account the error in both axes and a joint confidence region for slope b_1 and intercept b_0. As a reliable approach we assumed equal errors in both variables—in the calculated as well as in the certified concentration values.

The assumption that slope and intercept are simultaneously 1 and 0 is then accepted when the calculated value of the F-test is smaller than the tabulated one-sided F-value with 2 and $(n-2)$ degrees of freedom at the significance level α, i.e. if

$$F^{\text{cal}}<F^{\text{tab}}_{2,(n-2),\alpha}. \tag{10}$$

F^{cal} was computed by means of the expression given by Hartmann et al. [11]

$$F^{\text{cal}}=\frac{(C_m^{\text{cal}}-B_1C_m^{\text{cert}})^2+2(C_m^{\text{cert}}+B_1C_m^{\text{cal}})(-(C_m^{\text{cal}}-B_1C_m^{\text{cert}}))\left(\frac{1-B_1}{1+B_1}\right)}{2\cdot S_E^2/n}+\frac{\left(\frac{1}{n}\sum_i(C_i^{\text{cert}}+B_1C_i^{\text{cal}})^2\right)\cdot\left(\frac{1-B_1}{1+B_1}\right)^2}{2\cdot S_E^2/n} \tag{11}$$

where

$$B_1=\frac{Q_y-Q_x+\sqrt{(Q_x-Q_x)^2+4Q_{xy}^2}}{2Q_{xy}}$$

and

$$S_E=\sqrt{\frac{1}{n-2}\cdot(B_1^2Q_x-2B_1Q_{xy}+Q_y)}$$

represent estimates of the slope b_1 and the standard deviation σ_E respectively, with

$$Q_x=\sum_i(C_i^{\text{cert}}-C_m^{\text{cert}})^2;\quad Q_y=\sum_i(C_i^{\text{cal}}-C_m^{\text{cal}})^2;$$

$$Q_{xy}=\sum_i(C_i^{\text{cert}}-C_m^{\text{cert}})\cdot(C_i^{\text{cal}}-C_m^{\text{cal}})$$

$$C_m^{\mathrm{cert}}=\frac{1}{n}\cdot\sum_i C_i^{\mathrm{cert}};\quad C_m^{\mathrm{cal}}=\frac{1}{n}\cdot\sum_i C_i^{\mathrm{cal}}$$

If Eq. (10) is fulfilled, it can be concluded that the analytical method is not biased and trueness can be stated also in consideration of all influences caused by the laboratory.

2.3. Investigation of precision

2.3.1. Estimates of the precision of the analytical method

Accordingly with the different techniques used for sample preparation (for the determination of main components and trace analytes), the estimation of the precision was carried out also for the different preparation procedures.

For this purpose two sets of sediment samples were prepared, each consisting of 10 specimens, which were individually prepared by means of fusion technique and pelletizing, respectively (Section 3.2.2). While the set of glass discs was based on the standard reference material GBW07309 (river sediment), the set of pressed pellets was prepared using an in-house standard river sediment.

One sample of each set was selected to be measured under short-term conditions (within a time scale of 1 day) and under long-term conditions (within a time scale of 2 months), including also recalibration procedures of the spectrometer. In any case 12 analytical runs were performed, yielding the repeatability precision (RSD_1) and an approximation for the laboratory precision (RSD_3), attributed to the respective preparation techniques.

The investigations of the precision were completed by each of the 12 analytical runs applied to the whole sets (for the glass discs and for the pellets) under short-term conditions (within a time scale of 1 day), to examine the influence of the preparation technique on the precision (RSD_2).

2.3.2. Dependence of the precision on the concentration

It is well known that the precision of an analytical method depends on the concentration level of the analyte (Thompson [12]). To give a realistic estimation for uncertainty of the analytical result, its precision must be known not only as a single value. It has to be estimated over the entire range of concentrations of interest. Especially, in the field of environmental research it is very important to be able to predict the precision at any given concentration level to compare the analytical performance of the method with the analytical requirements. In the literature the dependence of the precision σ_{c} on the concentration is often fitted by a linear model function of the form

$$\sigma_{\mathrm{c}}=\sigma_0+k_1C \qquad (12)$$

where σ_0 and k_1 are the standard deviation at zero concentration and a proportionally constant, respectively.

Because Eq. (12) represents the sum of analytical errors of two independent terms, it should be theoretically more satisfactory to combine for this purpose the variances rather than the standard deviations [13], i.e.

$$\sigma_{\mathrm{c}}^2=\sigma_0^2+k_1^2C^2. \qquad (13)$$

Following the procedure proposed by Thompson and Howarth [13], which expressed the precision by the twofold relative standard deviation

$$p_{\mathrm{c}}=\frac{2\sigma_{\mathrm{c}}}{C} \qquad (14)$$

and using the definition that the detection limit C_{d} is set to concentration at which p_{c} is unity, one finally finds the equation

$$p_{\mathrm{c}}=\sqrt{\left(\frac{(1-k^2)C_{\mathrm{d}}^2}{C^2}+k^2\right)} \qquad (15)$$

with $k^2=4k_1^2$.

3. Experimental

3.1. Instrumental

The measurements were carried out on a Siemens SRS 3000 X-ray fluorescence spectrometer equipped with an Rh X-ray tube (125-μm Be window), 60-kV generator, eight-position crystal changer and 58-position sample changer. The spectrometer was controlled by a PCD-3M computer with the software package SPECTRA 3000. The spectrometer operating conditions were vacuum, 34(23) mm collimator mask and the following analyzing crystals: OVO 55, Ge, LIF 100 and LIF 110.

The Rh Kα Compton scattered tube emission line, used for analytical purposes, was recorded with the 0.15° collimator in combination with the LIF 110 crystal and the NaI scintillation counter.

3.2. Sample preparation and matrix correction

To prepare unknown samples for the analysis, the original material was dried at 105 °C and following ground using an agate ball mill (Retsch). In contrast to this, the certified reference materials were used without any changes to manufacture the calibration samples.

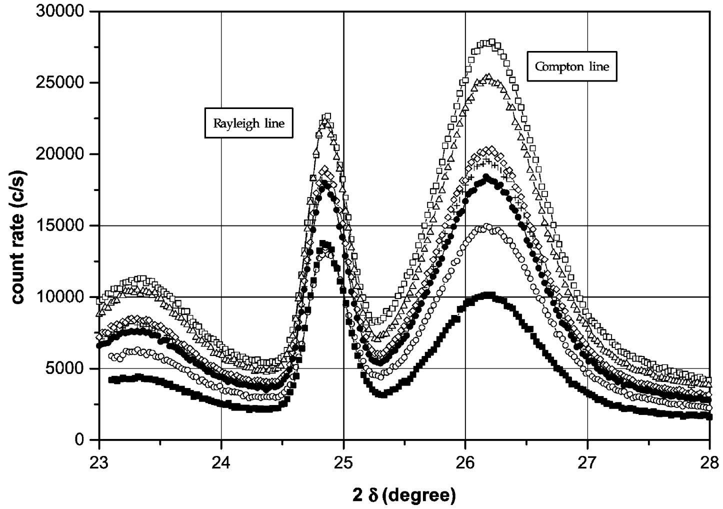

Fig. 1. Scans of the Rh Kα scattered tube line in dependence of the sample matrix (□□□ SiO2, △△△ STSD1, ●●● STSD2, + + + GBW07309, ○○○ GBW07310, ◇◇◇ GBW07311, ■■■ ML107).

3.2.1. Major components

The reference materials NIST-SRM2689 and NIST-SRM2691 (coal fly ashes), CANMET-LKSD1-LKSD4 (lake sediments), CANMET-STSD1-STSD4 and GBW07310-11 (stream sediments) were diluted with $Li_2B_4O_7$ (1 g sample + 7 g $Li_2B_4O_7$) to prepare glass discs by fusion at 1200 °C in a 95% platinum–5% gold crucible for 20 min. The application of this preparation technique to the selected reference materials yielded a set of glass discs, which were used to calibrate the compounds Na_2O, MgO, Al_2O_3, SiO_2, P_2O_5, K_2O, CaO, TiO_2, MnO and Fe_2O_3. For matrix correction the fundamental parameter option of the spectrometer software package SPECTRA 3000 was applied to the measured fluorescence intensities.

3.2.2. Trace constituents

Aliquots (4 g) of the reference material were mixed with 20% wax (Hoechst wax for XRF analysis) as a binder, poured into moulds and compacted in a hydraulic press at a pressure of 100 MPa to prepare pellets.

The calibration investigated in this study was intented for the determination of traces of heavy metals in a wide range of matrix compositions. The following reference materials (including silicate-, carbon- and iron-hydroxide matrix) were used for calibration: GBW07309-11, CANMET-STSD1-3, NIST-SRM2710 and SRM8407-08 (stream sediments), GBW07402/07 (soils), SY-2 (Syenite), BCR-CRM 40 (coal) and a set of in-house reference samples ML107 (sediment of acid mine lakes [14]) and mercury standards prepared by dilution of the reference material NIST-SRM8408 with powdered quartz of analytical grade quality.

For matrix correction the Compton Kα scattered tube line emission (Fig. 1) was used as an internal standard [14]. In Fig. 2 the performance of this correction algorithm is exemplarily demonstrated for the recovery function of Zn. The regression analysis, based on the matrix corrected response and the data of the certified reference material, yielded in a regression coefficient of $r = 0.998$. The results for the reference materials BAF BE-1 and BAF BE-1 (soft coal ashes), also displayed in Fig. 2, were not considered in the evaluation routine, because their inclusion led to a significant deterioration of the validation parameters. The necessity to exclude a reference sample from the list of the calibration samples commonly indicates that the rules of the correction algorithm apparently were disregarded. The outlying behavior of the samples BAF BE-1 and BAF BS-1, for example, may occur from incomplete matrix correction caused by the unusual high contents of Ba and Sr in these materials.

3.3. Test of homogeneity

When dealing with the analysis of solids, in general, increased requirements to the process of sample preparation must be taken into account to meet the conditions of homogenized and analytical grade material. The problem appears, for example, if a bulk of sample material is divided into sub-samples, designated for further investigations. Variations in composition (σ_{sam}) will arise due to the splitting procedure. In the ISO/IUPAC/AOAC Harmonised Protocol for Proficiency Testing [15], it is recommended that for sufficient homogeneity of distributed material the following condition must be fulfilled:

$$\frac{\sigma_{sam}}{\sigma_{tar}} < 0.3 \qquad (16)$$

where σ_{tar} represents a given target standard deviation.

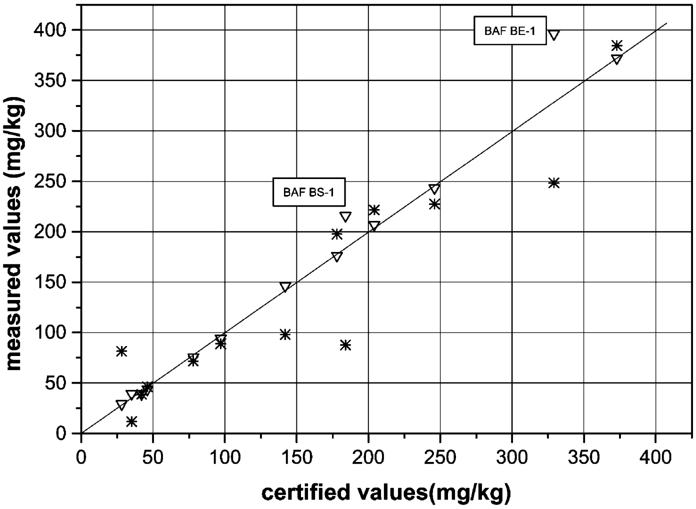

Fig. 2. Effect of matrix correction (based on the scattered radiation method) for the determination of the Zn content in the calibration samples (*** uncorrected data, ▽▽▽ including matrix correction).

As result of an analysis of variance (one way ANOVA, accomplished by an add-in program of Microsoft Excel), one obtains the mean sum of squares between the *p* groups (samples variation) and within the groups (analytical variation due to *m* measurements) MS_{sam} (mean of the between variance) and MS_{an} (mean of the within variance), respectively. The test value

$$F^{cal} = MS_{sam}/MS_{an} \tag{17}$$

is calculated and compared with the value of *F*-distribution at $f_1 = p-1$ and $f_2 = p(m-1)$ degrees of freedom and the significance level α. If $F^{cal} > F^{tab}_{f1,f2,\alpha}$, the hypothesis H_0 (no significant differences between the samples) is rejected.

The sample standard deviation σ_{sam} is estimated by

$$s_{sam} = \sqrt{(MS_{sam} - MS_{an})/2} \tag{18}$$

and for σ_{tar} we have used the within laboratory standard deviation.

As an example the homogeneity test was applied to geological material (originating from a tin mine tailings), which is in general strongly affected by mineralogical effects.

For this purpose two individual charges (*X*, *Y*) of the test material were prepared, which differed only with regard to the grinding time ($t_X = 5$ min, $t_Y = 15$ min) applied to the raw material. Each charge was at first divided into $p = 12$ sub-samples and subsequently analyzed in duplicate (i.e. $m = 2$), yielding two samples (*X*1, *X*2) and (*Y*1, *Y*2) for each group.

Table 1
Fe and Sn responses for two charges (*X*, *Y*) of geological material analyzed in duplicate

Groups (*p*)	Fe Kα				Sn Kα			
	Charge *X* ($t_X = 5$ min)		Charge *Y* ($t_Y = 15$ min)		Charge *X* ($t_X = 5$ min)		Charge *Y* ($t_Y = 15$ min)	
	*X*1 (counts)	*X*2 (counts)	*Y*1 (counts)	*Y*2 (counts)	*X*1 (counts)	*X*2 (counts)	*Y*1 (counts)	*Y*2 (counts)
1	44 562	44 058	44 000	43 462	30 801	30 101	30 031	30 543
2	45 073	45 358	43 624	43 596	31 895	31 671	30 558	30 291
3	44 833	45 003	43 961	44 459	29 523	30 266	29 888	30 549
4	43 241	43 493	44 377	44 852	27 842	27 398	30 502	30 855
5	44 412	44 757	42 856	44 005	30 837	30 074	30 385	30 396
6	43 934	44 836	44 286	44 019	30 015	30 288	30 604	29 996
7	44 229	44 674	43 538	43 764	27 744	28 018	30 459	30 423
8	44 079	44 069	43 718	44 159	29 374	29 522	29 742	30 040
9	44 068	43 827	44 005	44 432	26 428	26 546	30 325	30 389
10	44 111	44 294	43 090	43 856	29 772	29 724	29 971	30 025
11	45 000	45 360	44 534	43 987	29 876	30 126	30 341	29 730
12	43 900	44 800	43 476	44 321	28 798	29 111	30 617	30 432

Table 2
ANOVA results for the Fe and Sn responses with reference to the test of sufficient homogeneity, applied to the analytical data of the charges *X* and *Y*

	Fe		Sn	
	Charge *X* (t_X=5 min)	Charge *Y* (t_Y=15 min)	Charge *X* (t_X=5 min)	Charge *Y* (t_Y=15 min)
MS_{sam}	546 098.5	266 926.7	4 245 562.4	104 635.7
MS_{an}	107 898.7	174 301.0	92 489.0	73 725.2
F_{test}^{cal}	**5.06**	**1.53**	**45.90**	**1.42**
$F_{11;12;0,05}^{tab}$	2.717			
σ_{tar}	700	700	450	450
σ_{sam}	468	215	1441	124
$\sigma_{sam}/\sigma_{tar}$	**0.67**	**0.31**	**3.20**	**0.28**

The experimental data of the duplicate measurements, displayed in Table 1, were accumulated under repeatability conditions and focused on the most intense X-ray lines in the measured spectrum of Fe and Sn. The corresponding ANOVA results are summarized in Table 2.

Consequently, it must be stated that for the charge *X* (5 min grinding time) of the investigated test material, the requirements for sufficient homogeneity are rejected for the analytes under study by both the F-test ($F_{test}^{cal} > F_{f1,f2,\alpha}^{tab}$) and by Eq. (16).

The detected inhomogeneity in charge *X* apparently is caused by incomplete elimination of particle size or mineralogical effects within the original sample powder. It confirms that the preparation of homogenized and analytical grade sample material proves to be just as the matrix correction procedure, a further but often suppressed crucial problem, when dealing with methods of solid state analysis. This behavior must be considered to a high degree, when analytical data have to be compared within a collaborative experiment (involving even different analytical methods) with one another. To overcome this shortcoming in practice, it may be required to use more refined and individual adapted techniques of sample preparation in dependence of the quality of the sample material. With reference to the investigated geological material, for example, the requirements for sufficient homogeneity of distributed material could be confirmed, if an increased grinding time (charge *Y*) was applied to the raw material (Table 2).

4. Application of the presented validation concept

4.1. Precision

4.1.1. Precision of the method

The experimental set-up used for evaluation of the precision variation, based on the sample sets, have been described in Section 2.3.1. The results of the statistical treatment of the corresponding data (summarized in Tables 3 and 4) indicated that, in general, for all analytes the precision is substantially caused by sample preparation.

To quantify the influence of the preparation technique on the precision of the analytical results, the ratios RSD_1/RSD_2 were evaluated in dependence of the analytes of interest (Fig. 3).

Table 3
Experimental estimates for precision based on glass discs (fusion technique)

GBW07309	Na_2O	MgO	Al_2O_3	SiO_2	P_2O_5	K_2O	CaO	TiO_2	MnO	Fe_2O_3
C_{cert} (wt.%)	1.44	2.39	10.58	64.89	0.15	1.99	5.35	0.92	0.08	4.86
C_m (wt.%)	1.37	2.31	10.50	64.10	0.16	1.96	5.37	0.95	0.08	4.99
RDS_1 (%)	0.59	0.30	0.10	0.13	0.98	0.14	0.12	0.22	0.49	0.11
RDS_2 (%)	0.72	0.45	0.23	0.29	0.97	0.23	0.38	0.35	0.87	0.54
RDS_3 (%)	0.46	0.53	0.17	0.39	0.98	0.20	0.15	0.34	0.81	0.20

RDS_1—repeatability precision for an individual specimen (i.e. without the influence of sample preparation); RDS_2—repeatability precision for the whole set of samples (i.e. including sample preparation); RDS_3—intermediate precision with reference to an individual specimen (i.e. without influence of sample preparation but including long-term conditions as recalibration of the spectrometer).

Table 4
Experimental estimates for precision based on pellets (pelletizing technique)

In-house standard	Fe_2O_3	Ni	Cu	Zn	As	Rb	Zr	Pb	Sb
C_m (mg/kg)	223 000	67	36	70	83	47	129	7	15
RDS_1 (%)	0.29	3.42	4.47	1.65	0.85	1.02	0.47	12.56	3.16
RDS_2 (%)	0.87	4.60	5.49	5.13	1.27	1.27	2.19	19.96	3.87
RDS_3 (%)	0.86	3.14	5.00	2.10	1.17	1.56	0.96	16.07	3.17

RDS_1—repeatability precision for an individual specimen (i.e. without the influence of sample preparation); RDS_2—repeatability precision for the whole set of samples (i.e. including sample preparation); RDS_3—intermediate precision for an individual specimen (i.e. without influence of sample preparation but including long-term conditions as recalibration of the spectrometer).

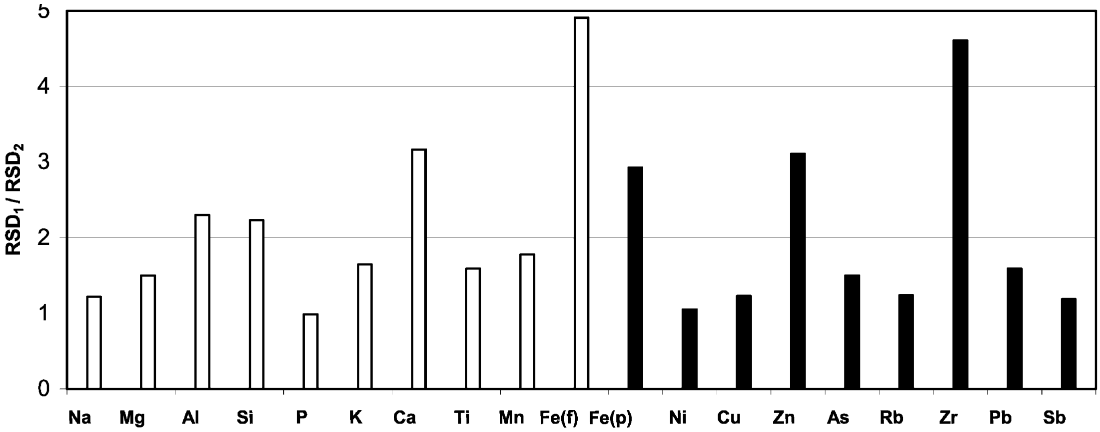

Fig. 3. Ratios of the repeatability precision RSD_1/RSD_2 (without and including the effect of sample preparation: for Na···Fe(f) glass discs and for Fe(p)···Sb pressed pellets).

Table 5
Parameters for the model function (Eq. (15)) (variance function) estimated by means of non-linear regression analysis applied to the experimental precision data $s_{\mathrm{exp}}(C)$ of Cu, Zn, As and Pb at different concentration levels

Analyte	C_d^2	k^2
Cu	37.9	0.00088
Zn	57.3	0.00032
As	6.38	0.00105
Pb	11.8	0.00017

Consequently, the precision for the individual analytes is differently affected by the sample preparation. It was found that above all the analytes Fe and Zr are concerned. While the effect for Fe (measured as in the glass discs [Fe(f)], as in the pellets [Fe(p)]) is apparently due to specific problems caused by the fusion procedure, the different behavior of Zr must be attributed to incomplete elimination of particle size or mineralogical effects in the sample powder.

Summarizing, it can be stated that despite the increased precision values for the analytes Fe and Zr, the investigated methodology should justify the 'customer's trust' to provide reliable data to the analytical part of such projects as mentioned in the introduction, at least with regard to the precision of the analytical data.

4.1.2. Dependence of the precision on the concentration

To investigate the general behavior of the approach (Eq. (15)), we estimated the parameters C_d^2 and k^2 in the model function for the analytes Cu, Zn, As and Pb, acting as analytes with elevated environmental concern. For this purpose a set of samples was prepared on the basis of reference materials, which covered a wide range of the corresponding analyte concentrations, starting at the LOD up to approximately 2000 mg/kg. From each selected reference material, four specimens were prepared as pressed pellets and measured under conditions that are closely related to the term intermediate precision. Analyzing the front side as well as the back side of the pellets, the experiment yielded eight analytical results for each concentration level, which were used to estimate the experimental precision $s_{\mathrm{exp}}(C)$.

The determination of the parameters of the model function (Eq. (15)) was finally performed by a non-linear regression routine applied to the experimental data s_{exp}. The results are summarized in Table 5.

In Table 6 the experimental and computed data for

Table 6
Comparison of precision, expressed as RSD (%), as a function of analyte concentration (As in silicate matrix)

C_{As} (mg/kg)	RSD_r (experim.) $2s_{\mathrm{exp}}/C$ (%)	RSD_r (Eq. (15)) p_c (%)	RSD_r (Horwitz Eq. (19)) (%)
16.4	20.11	15.73	10.50
22.5	5.13	11.68	10.01
91.4	6.85	4.26	8.11
639	2.83	3.26	6.05
691	3.37	3.26	5.98
1150	3.31	3.25	5.54
1735	2.44	3.24	5.21

Table 7
Validation parameters concerning the calibration of the main components (residual standard deviation in concentration units $s_{e,i}/b_i$, relative measurement uncertainty $u_i(C_{i,m})/C_{i,m}$ and mean concentrations $C_{i,m}$)

Compound	Range of concentration (wt.%)	$C_{i,m}$ (wt.%)	$s_{e,i}/b_i$ (wt.%)	$u_i(C_{i,m})/C_{i,m}$
Na_2O	0.30–2.00	1.33	0.04	0.028
MgO	0.10–5.00	1.92	0.04	0.026
Al_2O_3	3.00–25.00	11.84	0.12	0.011
SiO_2	36.00–100	54.31	0.40	0.015
P_2O_5	0.05–1.20	0.33	0.02	0.055
K_2O	0.40–3.50	1.61	0.03	0.02
CaO	0.50–26.00	5.45	0.13	0.03
TiO_2	0.20–1.50	0.72	0.05	0.07
MnO	0.03–0.50	0.19	0.026	0.09
Fe_2O_3	4.00–15.00	6.20	0.10	0.015

precision of arsenic, as an example, are displayed and compared with the precision data estimated by means of the well-known Horwitz relation [16]

$$RSD_r(\%) \approx 16C_{(mg/kg)}^{-0.1505}, \qquad (19)$$

which can be regarded as a rugged and universal estimate of precision for high analyte concentrations.

Comparing the precision values of both approaches, it must be emphasized that the Horwitz relation is related to the conditions of reproducibility, while the RSD values presented in this study were produced under intermediate conditions. From this point of view, the higher precision values, provided by the Horwitz relation, are explainable.

In contrast to this, a remarkable agreement could be stated with regard to the asymptotic behavior of the function p_c (C). For both approaches the beginning of the asymptotic region of the function p_c (C) was found to be approximately at analyte concentrations near 100 mg/kg. Therefore, it can be concluded that with regard to the investigated sample material for concentrations above 100 mg/kg the precision varies only slowly in dependence of the concentration level and can be in most cases taken into account as a constant.

4.2. Estimates of the validation parameters

With respect to the estimation of the validation parameters, a pool of analytical data was collected including the certified and the calculated concentration values (Eq. (3)) of the calibration samples, listed in Sections 3.2.1 and 3.2.2. The application of Eqs. (4)–(6) to this bulk of data finally yielded estimates for the validation parameters $(s_{e,i}/b_i)$, $u_i(C_{i,m})/C_{i,m}$, $C_{i,m}$ and LOD, associated with investigated calibration.

In Tables 7 and 8 the corresponding results of such a evaluation procedure are summarized for both the major and the trace constituents, respectively.

To obtain reliable estimates of the validation parameters also under conditions of heteroscedasticity of the calibration data, it is normally recommended to use weighted least-squares regression (WLS) instead of the more simple to handle ordinary least-squares regression (OLS). But because the use of WLS requires additional information on the variances occurring at different concentration levels (variance function), this type of regression is commonly less used as OLS [10].

In the present study heteroscedasticity (examined by means of the F-test, applied to the analytical data, accumulated at the lower- and upper-limit of the working region of the calibration line) was only indicated for the calibration data of copper, zinc and arsenic (apparently due to the more extended concentration range for these analytes). To estimate the influence of the detected heteroscedasticity on the calculated validation data, we applied both OLS and WLS, respectively, to the calibration data. The required information on the variance function was gained from Eq. (15). The numerical results of the corresponding regression procedures, also presented in Table 8, indicated that (with the exception of the As-calibration) there is no evidence for the necessity to prefer WLS instead of OLS.

Table 8
Validation parameters concerning the calibration of the trace components (residual standard deviation in units of concentrations $(s_{e,i}/b_i)$, relative measurement uncertainty $u_i(C_{i,m})/C_{i,m}$, limits of detection (LOD) and mean concentrations $C_{i,m}$)

Analyte	Range of concentration (mg/kg)	$C_{i,m}$ (mg/kg)	$s_{e,i}/b_I$ (mg/kg)	$u_i(C_{i,m})/C_{i,m}$	LOD (mg/kg)
Ni	13– 280	45	2.0	0.040	5
Cu	16–2950	339	3.9* (3.07)	0.015	7
Zn	42–3800	605	3.1* (2.88)	0.005	7
As	5– 630	90	4.5* (1.98)	0.019	5
Pb	14– 640	136	2.4* (2.39)	0.017	6
Hg	5– 110	47	1.9	0.040	5
U	4– 290	44	1.4	0.048	4

The marked results (*) refer to weighted regression.

Table 9
Recovery data for the major components and the corresponding ratios of the calculated and tabulated F-values ($F^{cal}/F^{tab}_{2,n-2.5\%}$), estimated by means of regression analysis (joint test)

Reference	Na_2O		MgO		Al_2O_3		SiO_2		P_2O_5	
Material	$C^{cert}_{Na_2O,j}$ (wt.%)	$C^{cal}_{Na_2O,j}$ (wt.%)	$C^{cert}_{MgO,j}$ (wt.%)	$C^{cal}_{MgO,j}$ (wt.%)	$C^{cert}_{Al_2O_3,j}$ (wt.%)	$C^{cal}_{Al_2O_3,j}$ (wt.%)	$C^{cert}_{SiO_2,j}$ (wt.%)	$C^{cal}_{SiO_2,j}$ (wt.%)	$C^{cert}_{P_2O_5,j}$ (wt.%)	$C^{cal}_{P_2O_5,j}$ (wt.%)
NIST-SRM2704	0.739	0.730	2.000	1.968	11.52	11.82	62.80	61.87	0.229	0.230
NIST-SRM2709	1.560	1.500	2.510	2.472	14.15	14.26	63.50	61.95	0.142	0.148
NIST-SRM2689	0.330	0.323	1.010	0.961	24.41	24.58	51.96	51.21	0.229	0.230
NIST-SRM2691	1.470	1.464	5.200	5.175	18.51	18.41	36.04	36.26	1.170	1.183
GBW07402	1.620	1.484	1.040	0.980	10.31	10.51	73.35	73.14	0.102	0.102
GBW07406	0.190	0.124	0.340	0.272	21.23	21.22	56.93	56.17	0.069	0.065
GBW07407	0.074	0.024	0.260	0.173	29.26	29.48	32.69	32.50	0.263	0.266
GBW07409	3.310	3.289	0.490	0.453	12.91	13.36	73.28	73.17	0.073	0.077
GBW07309	1.440	1.358	2.390	2.352	10.58	10.84	64.89	64.57	0.154	0.159
GBW07310	0.040	0.000	0.120	0.086	2.84	3.01	88.89	88.47	0.062	0.069
GBW07311	0.460	0.417	0.620	0.568	10.37	10.56	76.25	76.13	0.059	0.059
BAFBS-1	0.190	0.131	4.910	5.017	5.15	4.95	15.05	15.74	0.070	0.098
CTA-FFA-1	2.950	2.923	2.570	2.839	28.10	27.76	48.55	49.71	0.166	0.182
$F^{cal}/F^{tab}_{2,n-2.5\%}$	$9.96/3.98 \approx 2.5$		$1.72/3.98 < 1$		$1.99/3.98 < 1$		$2.10/3.98 < 1$		$3.52/3.98 < 1$	

	K_2O		CaO		TiO_2		MnO		Fe_2O_3	
Reference material	$C^{cert}_{K_2O,j}$ (wt.%)	$C^{cal}_{K_2O,j}$ (wt.%)	$C^{cert}_{CaO,j}$ (wt.%)	$C^{cal}_{CaO,j}$ (wt.%)	$C^{cert}TiO_2,j$ (wt.%)	$C^{cal}TiO_2,j$ (wt.%)	$C^{cert}_{MnO,j}$ (wt.%)	$C^{cal}_{MnO,j}$ (wt.%)	$C^{cert}_{Fe_2O_3,j}$ (wt.%)	$C^{cal}_{Fe_2O_3,j}$ (wt.%)
NIST-SRM2704	2.400	2.396	3.660	3.664	0.760	0.791	0.072	0.074	5.870	6.027
NIST-SRM2709	2.450	2.378	2.660	2.654	0.570	0.600	0.070	0.070	5.000	5.090
NIST-SRM2689	2.650	2.643	3.070	2.934	1.250	1.284	0.030	0.037	13.310	13.287
NIST-SRM2691	0.400	0.406	25.980	26.094	1.500	1.514	0.025	0.021	6.310	6.540
GBW07402	2.540	2.510	2.360	2.302	0.450	0.437	0.066	0.065	3.520	3.539
GBW07406	1.700	1.698	0.220	0.109	0.733	0.743	0.187	0.190	8.090	8.109
GBW07407	0.200	0.192	0.160	0.047	3.370	3.325	0.229	0.238	18.760	18.777
GBW07409	3.370	3.413	1.350	1.295	0.416	0.444	0.034	0.036	2.080	2.023
GBW07309	1.990	1.995	5.350	5.456	0.916	0.921	0.081	0.079	4.860	5.042
GBW07310	0.125	0.120	0.700	0.626	0.211	0.259	0.131	0.131	3.860	3.910
GBW07311	3.280	3.233	0.470	0.395	0.350	0.376	0.323	0.321	4.390	4.464
BAFBS-1	0.195	0.208	26.930	26.547	0.253	0.297	0.361	0.365	26.930	27.082
CTA-FFA-1	2.700	2.835	3.200	3.094	0.970	1.157	0.138	0.143	6.980	7.617
$F^{cal}/F^{tab}_{2,n-2.5\%}$	$0.04/3.98 < 1$		$2.12/3.98 < 1$		$2.76/3.98 < 1$		$1.71/3.98 < 1$		$2.69/3.98 < 1$	

To overcome the detected heteroscedasticity in the As calibration, a restriction of the working range is most recommended. In accordance with that, homoscedasticity could be restored in the more limited concentration range from 10 up to 200 mg/kg, excluding the reference sample with the highest As concentration ($C^{cert}_{As} = 628$ mg/kg) from the list of the calibration samples.

The estimation of the LOD values, calculated by means of Eq. (6), was performed in generally for a restricted working range, to ensure the requirement of homoscedasticity.

4.3. Examination of trueness

Examining the analytical method under study with regard to trueness, we focused on the regression analysis of the recovery functions (Section 2.2) based on complex sample material covering a wide range of matrix compositions and analyte concentrations. While the estimation of the validation parameters (Section 4.2) based exclusively on the use of the calibration samples, the examination of trueness is focused on additional reference material, not included in the calibration routine.

In Tables 9 and 10 the results of the joint test applied to the recovery function for both the major and the trace components, respectively, are summarized.

With regard to the major components, the reference materials NIST-SRM 2704, GBW07309 (river sediment), NIST-SRM2709, GBW07402, -07406, -07407, -07409 (contaminated soils), BAF BS-1 (lignite filter ash) and CTA-FFA-1 (fly ash) were additionally included into the evaluation routine (Table 9).

The recovery functions for the trace elements (Table 10) were performed including also the reference materials CANMET-LKSD1-4 (lake sediment), PACS-1

Table 10
Recovery data for the trace components and the corresponding ratios for the calculated and tabulated F-values ($F^{\mathrm{cal}}/F^{\mathrm{tab}}_{2,n-2.5\%}$), estimated by means of regression analysis (joint test)

Ni		Cu		Zn		As		Hg		Pb	
$C^{\mathrm{cert}}_{\mathrm{Ni,j}}$ (mg/kg)	$C^{\mathrm{cal}}_{\mathrm{Ni,j}}$ (mg/kg)	$C^{\mathrm{cert}}_{\mathrm{Cu,j}}$ (mg/kg)	$C^{\mathrm{cal}}_{\mathrm{Cu,j}}$ (mg/kg)	$C^{\mathrm{cert}}_{\mathrm{Zn,j}}$ (mg/kg)	$C^{\mathrm{cal}}_{\mathrm{Zn,j}}$ (mg/kg)	$C^{\mathrm{cert}}_{\mathrm{As,j}}$ (mg/kg)	$C^{\mathrm{cal}}_{\mathrm{As,j}}$ (mg/kg)	$C^{\mathrm{cert}}_{\mathrm{Hg,j}}$ (mg/kg)	$C^{\mathrm{cal}}_{\mathrm{Hg,j}}$ (mg/kg)	$C^{\mathrm{cert}}_{\mathrm{Pb,j}}$ (mg/kg)	$C^{\mathrm{cal}}_{\mathrm{Pb,j}}$ (mg/kg)
13	10	5	5	42	42	5	7	5	4	14	16
14	14	16	13	46	47	7	9	6	5	16	14
16	16	23	24	78	79	8	11	30	33	20	19
19	20	32	34	230	232	8	10	33	34	23	21
24	26	47	45	248	248	17	18	40	41	27	27
24	27	65	66	350	346	23	21	50	48	80	81
26	29	75	71	373	369	25	25	58	56	85	87
30	33	79	74	438	437	42	40	90	92	146	142
30	30	83	88	498	498	47	45	107	106	161	160
30	29	99	98	547	553	54	53			285	289
31	36	114	112	3800	3800	105	102			636	635
32	33	102	104			205	205				
53	56	158	161			628	628				
53	54	1230	1234								
276	277	2950	2948								
$F^{\mathrm{cal}}/F^{\mathrm{tab}}_{2,n-2.5\%}=$ 2.87/3.80		$F^{\mathrm{cal}}/F^{\mathrm{tab}}_{2,n-2.5\%}=$ 0.01/3.80		$F^{\mathrm{cal}}/F^{\mathrm{tab}}_{2,n-2.5\%}=$ 0.01/4.26		$F^{\mathrm{cal}}/F^{\mathrm{tab}}_{2,n-2.5\%}=$ 0.09/3.98		$F^{\mathrm{cal}}/F^{\mathrm{tab}}_{2,n-2.5\%}=$ 0.01/4.74		$F^{\mathrm{cal}}/F^{\mathrm{tab}}_{2,n-2.5\%}=$ 0.05/4.26	

(marine sediment), GBW 07409-11, SOIL5, SOIL7 (soils), SL1, NIST-SRM 2704 (river sediments), BCR-CRM320 (river sediment), BCR-CRM277 (estuarine sediment).

In Figs. 4 and 5 intercept b_0, slope b_1 and the corresponding confidence intervals conf(b_1) and conf(b_0) are displayed, estimated by means of orthogonal regression analysis. The graphs give a useful qualitative survey as to what extent the individual recovery functions are affected by which kind of systematic error. Alternatively to this, an orthogonal regression analysis applied to the recovery data (Tables 9 and 10) provides a quantitative statement for bias based on Eq. (10).

In accordance with that, the absence of systematic errors is accepted if the ratio $F^{\mathrm{cal}}/F^{\mathrm{tab}}_{2,n-2.5\%} < 1$.

Applying this test procedure to the analytical method under study, it could be demonstrated that with the exception of sodium for all other analytes no bias for the recovery functions was detected. This could be immediately concluded as a result of the joint test that yielded only for the recovery function of sodium a test value greater than 1 ($F^{\mathrm{cal}}/F^{\mathrm{tab}}_{2,n-2.5\%} \approx 2.5$). The outstanding behavior of sodium probably could be attributed to an incorrect or incomplete correction for the interference between the Na Kα and the Zn Kβ lines.

A qualitative evaluation of Fig. 4 indicated in addition to this that the bias for sodium results from a constant systematic error because the significance interval of the intercept does not include 0, while the slope of the recovery function includes 1. For the recovery function of TiO_2 the significance interval of the intercept also does not include 0 but nevertheless the joint test indi-

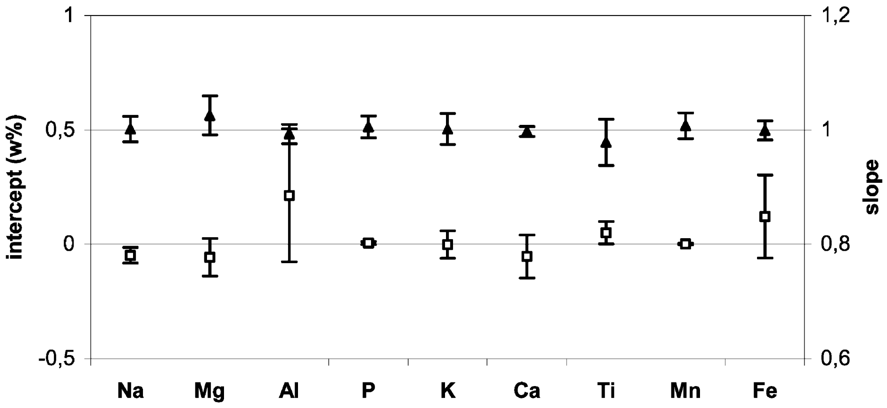

Fig. 4. Intercept $b_{0,i}$ (□□□), slope $b_{1,i}$ (▲▲▲) and the individual confidence intervals conf($b_{0,i}$) and conf($b_{1,i}$) for the recovery functions of the major components.

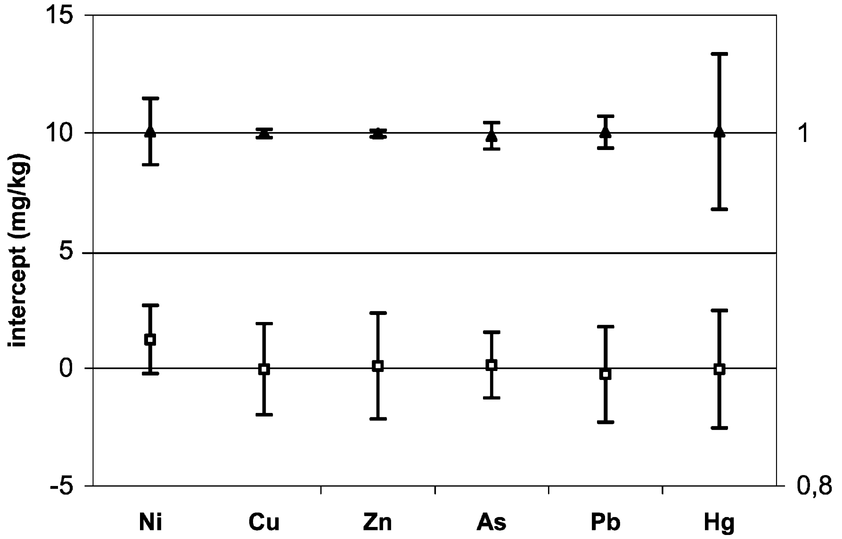

Fig. 5. Intercept $b_{0,i}$ (□□□), slope $b_{1,i}$ (▲▲▲) and the individual confidence intervals conf($b_{0,i}$) and conf($b_{1,i}$) for the recovery functions of the trace components.

cated still acceptance for the absence of a bias

$$\frac{F^{\mathrm{cal}}}{F^{\mathrm{tab}}_{2,n-2.5\%}} \approx \frac{2.76}{3.98}.$$

5. Conclusions

Steps toward the validation of XRF analysis can be performed with relatively reasonable effort applying tools of regression analysis in consideration of the recommended guide lines.

The application of validation procedures to the experimental data yielded reliable information about precision, measurement uncertainty, limit of detection and trueness with reference to the whole concentration range and within the scope of matrices defined by the used reference materials. Both the estimation of the validation parameters and the examination of trueness were performed by regression analysis applied to scatter diagrams (certified vs. measured concentration values), which are on the one hand based on the calibration samples and on the other hand on a set of additional reference samples not included in the calibration routine.

The results of the validation procedures meet the requirements to quantify the performance of the used XRF method (including the influence of sample preparation) and to examine the investigated calibration strategy with regard to fitness for purpose.

Concerning the determination of traces of heavy metals, the test procedures yielded according to the expectations somewhat increased values for the limits of detection (~5 mg/kg) and combined measurement uncertainties (~3 mg/kg), as commonly reported with respect to specified types of matrices.

Orthogonal regression analysis applied to the validation data indicated trueness for nearly all analytes of interest.

References

[1] M. Krapp, Quality assurance in research and development: an insoluble dilemma?, Fresenius J. Anal. Chem. 371 (2001) 704–713.

[2] Eurachem guide, The fitness for purpose of analytical methods, Edition 1.0, 1998.

[3] DIN32645 Nachweis-, Erfassungs- und Bestimmungsgrenze, Beuth Verlag GmbH Mai 1994.

[4] DIN51418-2 Röntgenemissions- und Röntgenfluoreszenz-Analyse (RFA), Beuth Verlag GmbH September 1996.

[5] EURACHEM/CITAC Guide, Quantifying uncertainty in analytical measurement, second edition, Final Draft, April 2000. EURACHEM: Available from: http://www.measurementuncertainty.org.

[6] IUPAC, L.A.Currie, IUPAC Commission on Analytical Nomenclature Recommendations in Evaluation of Analytical Methods including Detection and Quantification Capabilities, Pure Appl. Chem. 67 (1995) 1699–1723.

[7] ISO5725-1, Accuracy (trueness and precision) of measurements methods and results (1994).

[8] M. Thompson, S.L.R. Ellison, R. Wood, Harmonized guidelines for single laboratory validation of methods of analysis, Pure Appl. Chem. 74 (5) (2002) 835–855.

[9] K. Danzer, L.A. Currie, Guidelines for calibration in analytical chemistry, Pure App. Chem. 70 (1998) 993–1014.

[10] D.L. Massart, B.G.M. Vandeginste, S. DeJong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics—Part A, Chapter 8, Straight line regression and calibration, 1997, pp. 171–230.

[11] C. Hartmann, J. Smeyers-Verbeke, W. Penninckx, D.L. Massart, Detection of bias in method comparison by regression analysis, Anal. Chim. Acta 338 (1997) 19–40.

[12] M. Thompson, Variation of precision with concentration in an analytical system, Analyst 113 (1988) 1579–1587.

[13] M. Thompson, R.J. Howarth, Duplicate analysis in geochemical practice, Analyst 101 (1976) 690–698.
[14] P. Morgenstern, K. Friese, K. Wendt-Potthoff, R. Wennrich, Bulk chemistry analysis of sediments from acid mine lakes by means of X-ray fluorescence, Mine Water Environ. 20 (2000) 105–113.
[15] M. Thompson, R. Wood, IUPAC Protocol for the Proficiency Testing of (Chemical) Analytical Laboratories, 1993.
[16] W. Horwitz, R. Albert, Precision in analytical measurements: expected values and consequences in geochemical analysis, Fresenius J. Anal. Chem. 351 (1995) 507–515.