

# Performance and robustness of a multi-user, multi-spectrometer system for INAA

Peter Bode · Menno Blaauw

Received: 12 May 2011 / Published online: 1 June 2011  
© The Author(s) 2011. This article is published with open access at Springerlink.com

**Abstract** The laboratory for instrumental neutron activation analysis at the Reactor Institute Delft, Delft University of Technology uses a network of 3 gamma-ray spectrometers with well-type detectors and 2 gamma-ray spectrometers with coaxial detectors, all equipped with modern sample changers, as well as 2 spectrometers with coaxial detectors at the two fast rabbit systems. A wide variety of samples is processed through the system, all at specific optimized (and thus different) analytical protocols, and using different combination of the spectrometer systems. The gamma-ray spectra are analyzed by several qualified operators. The laboratory therefore needs to anticipate on the occurrence of random and systematic inconsistencies in the results (such as bias, non-linearity or wrong assignments due to spectral interferences) resulting from differences in operator performance, selection of analytical protocol and experimental conditions. This has been accomplished by taking advantage of the systematic processing of internal quality control samples such as certified reference materials and blanks in each test run. The data from these internal quality control analyses have been stored in a databank since 1991, and are now used to assess the various method performance indicators as indicators for the method's robustness.

**Keywords** Neutron activation analysis · Validation · Reference materials · Multi-user system · Accreditation

## Introduction

Instrumental neutron activation analysis (INAA) distinguishes itself from other methods for multi-element determination by, amongst others, the fact that the chemical matrix of the test portion has no, or barely noticeable influence on the trueness of the results. Calibration, based on the comparative method or on the single comparator method [1, 2], is therefore 'once and for all', making the technique suitable for analyzing a large variety of matrices. The chemical matrix has often no effect to the degree of trueness of the results, which also implies that there is no need for having quality control ("trueness control") samples of a matrix, closely matching the unknown sample matrix. The quality control sample is selected on basis of (i) known amount of the element(s) of interest, (ii) suitability for irradiation and measurement of the material under the same conditions as the real samples and (iii) detectability of the radionuclide(s) of the element(s) of interest under these conditions.

The laboratory for instrumental neutron activation analysis at the Reactor Institute Delft, Delft University of Technology uses a network of 3 gamma-ray spectrometers with well-type detectors and 2 gamma-ray spectrometers with coaxial detectors, all equipped with modern sample changers, as well as 2 spectrometers with coaxial detectors at the two fast rabbit systems. Details on the system have been published in the past [3]. Activated samples are often measured using more than one detector, and all spectra recorded are analysed and interpreted simultaneously [4]. The spectrometers with the sample changers are automated and can all run 24 h per day and 7 days per week. The quality management system, implemented at the end of the 1980 s [5] and accredited since 1993, currently meets the requirements of the NEN-EN-ISO/IEC 17025:2005.

---

P. Bode (✉) · M. Blaauw  
Reactor Institute Delft, Faculty of Applied Sciences,  
Delft University of Technology, Mekelweg 15, 2629JB Delft,  
The Netherlands  
e-mail: p.bode@tudelft.nl

The in-house developed gamma-ray spectrum analysis and interpretation software [4] has an module for graphical inspection of the outcome of quality control sample analyses [6]. The control charts could display (i) the mass fraction of an element as a function of analysis date, (ii) the mass fraction of an element for all types of quality control materials analyzed and (iii) the (normalized) mass fraction of all elements as a function of their value. The underlying databases with measurement results of any quality control sample have been kept up-to-date since the introduction in the early 1990s, and the number of options has been expanded for selecting parameters to be displayed in a control chart. Also the number of quality control samples has increased considerably, with (certified) reference materials forming the largest category.

A wide variety of samples is processed through the system, all at specific optimized (and thus different) analytical protocols and using different combination of the spectrometer systems. The gamma-ray spectra are analyzed by several qualified operators; between 1991 and 2010 approximately 15 persons, some of them for a period of 1–2 years; others during a longer period. The operators select the analytical protocol (selection of irradiation, decay, counting time and detector geometry(ies) to be used. Though the spectrum analysis and peak fitting is highly automated, sometimes manual intervention is necessary, e.g. for complex multiplets. The laboratory has to anticipate continuously on a situation that the human factor affects the quality of the results. Demonstrating the robustness of this system is therefore a continuous challenge.

The software and charts allow for trend analysis and for assessment of analytical performance indicators such as trueness, precision, reproducibility and linearity. In addition, the bias of the method, the linearity and repeatability

are assessed quantitatively. These indicators are averaged values derived from analyses of many different types of control samples, and for many elements also at a range of mass fractions and/or in the presence of interferences. As such, they implicitly reflect to some extent the robustness of the INAA method itself.

## Quality assessment

### Data sorting

An overview of the selectable parameters and the sorting functions for the control charts is given in Table 1, together with an indication of the information that may be derived. Examples of control charts have been published before [6].

### Trueness

The bias of the method is assessed though the analysis over time of a large set of certified reference materials. For each element in each material, the weighted mean of the measured mass fractions and its external and internal standard errors are calculated. The mean mass fraction is then divided by the certified mass fraction to obtain a bias factor with its standard error—the larger of the internal and external standard errors is used to this end. The weighted mean of the bias factors for a given element in all reference materials is then calculated, with its internal standard error. Also, the overall  $\chi_r^2$  is calculated from the zeta-scores of the group averages and the certified mass fractions.

The data taken into consideration are the grand total, the data from the last 10 years, the last 5 years, the last 2 years and the data from the last year. This is necessary because

**Table 1** Parameters in the control charts that can be generated by the software, in use at Reactor Institute Delft

Y axis	X axis	Selectable parameter	Sort function for selectable parameter	Indication on
Mass fraction	Measurement chronology	Control sample and element	a. Mass fraction b. Date c. Sample code	a. Trueness and linearity b. Time trends c. Human influence, robustness
Mass fraction	Type of control sample	1. Element 2. Spectral interference	a. Mass fraction b. Ratio of mass fraction of element and mass fraction of interference c. Zeta score	a. Linearity b. Selectivity and Robustness c. Estimation of uncertainty
Normalized mass fraction	All elements	1. Control sample 2. Measurement	a. Atomic number b. Relative combined standard uncertainty c. Zeta score	a. Analytical protocol b. Performance at extremes c. Estimation of uncertainty

the method and its calibration factors are continuously evolving.

The test is performed in the course of the annual evaluation of the first line internal quality control. The Laboratory for INAA has defined an acceptance criterion for the bias: an element passes the test when the observed bias factor with its standard error implies a 67% probability of the true value lying between 0.95 and 1.05 (so that an observed bias factor of  $1.00 \pm 0.05$  just passes the test). An element fails the trueness test when the bias factor deviates from unity by more than 5%, and at the same time deviates from unity by more than 3 standard deviations. In other cases, the test is inconclusive.

In any case, if the data from the last year prove the method to be unbiased for some element, failure to meet a criterion in earlier years is acceptable. The data obtained in the last year might be too few to draw a conclusion for some elements, in which case data from previous years will do.

### Range and linearity

The INAA method is, from first principles, linear in that the induced radioactivity is linear proportional to the amount of element present. This linearity applies also to the measured peak area, and the algorithm that determines the peak area [7, 8]. The Laboratory for INAA's requirement for the linearity therefore is more stringent than for the trueness: the contribution of non-linearity to the bias at any concentration level must be smaller than 0.25%.

Even at very low levels, accurate measurements of elemental mass fractions can be obtained by determining peak areas at the energies in the spectrum where the radionuclide of interest is expected to show up. The precision of such measurements below the detection limits will be very poor, but some end-users of the INAA data prefer this to only having a detection limit at their disposal. This leads to the Laboratory's requirement for the working range: it should be from 0 to 1 kg/kg.

Still there are a few caveats. If the element of interest is a major constituent of the material to be analyzed, the neutron self-shielding behavior and/or the gamma-ray self-absorption properties of the sample may become dependent on the mass of the element present, and the induced activity and/or the measured peak area would no longer be linearly proportional to its mass. Also, if the element to be determined is the major source of radioactivity in the material, the count rates in the measurement will be determined by the element concentration and linearity might be at risk due to dead time.

Neutron self-shielding, gamma-ray self-absorption and dead time and the requirements for the trueness of the corresponding correction methods are discussed separately below.

The verification of the linearity is based on a similar test as for the trueness: the bias factors found in that test are correlated to the certified concentrations by linear regression.

An element passes the test when the slope of the regression line with its combined standard uncertainty implies a 67% probability of the true value lying between  $-0.005$  and  $0.005$  (so that the slope  $0.000 \pm 0.005$  just passes the test). An element fails the linearity test when the slope deviates from zero by more than 0.005, and at the same time deviates from zero by more than 3 standard deviations. In other cases, the test is inconclusive.

### Repeatability and reproducibility

The INAA software reports the combined standard uncertainty for each element. The combined standard uncertainty, or "precision" is calculated from the known sources of the type A evaluations of uncertainty in the analysis procedure (counting statistics being the main one). If a sample is analyzed repeatedly by the same analysis protocol, this "precision" can be interpreted as the repeatability of the method.

The requirement for the precision is that the standard error of the mean value,  $SEM_{ext}$  (i.e. observed variation) should not significantly exceed the  $SEM_{int}$  (variation expected from calculations) for analysis runs of a homogeneous material, performed the same way.

If  $SEM_{ext} > SEM_{int}$ , then the implied unexplained variation should be less than 0.5%.

If  $SEM_{int} > SEM_{ext}$ , then, for the test to be conclusive, it should be demonstrated that an additional source of variation of 0.5% would have increased  $SEM_{ext}$  significantly in the  $\chi_r^2$ -sense, using

$$\chi_r^2 = \left( \frac{\sqrt{V^2/N + SEM_{ext}^2}}{SEM_{ext}} \right)^2,$$

where  $N$  is the number of observations, and  $V$  is the extra variability.

Reproducibility is tested the same way and at the same time, by selecting a homogeneous material and measurements thereof by different procedures, detectors and technicians. The requirement is that the unexplained variation arrived at from the results should be less than 1%.

### Robustness

The robustness of the method depends mostly on the performance of the operators, as outlined in the Introduction. The related induced variance is already incorporated in the variance of the trueness, repeatability and reproducibility. The only external factor potentially affecting the INAA

results is the climate in the counting room. Temperature has an effect on the energy calibration of the detectors, that might lead to misidentification of radionuclides. Since this is the most serious error that could be made, the requirement is that it should never happen. This requirement is tested along with the trueness of the method as described above—if robustness should be lacking, the method would fail that test.

### Regular review

All tests described here are performed annually as integral part of the annual evaluation of the results of the first line internal quality control. This is necessary because the method and its calibration factors are continuously evolving.

The data taken into consideration are the grand total, the data from the last 10 years, the last 5 years, the last 2 years and the data from the last year. If a test for a more recent dataset is inconclusive, the dataset from the next larger time span can be used to decide on validity. Failures to pass a test are studied in depth and corrective actions may be defined.

### Results

#### Trueness

An example of the raw data from the program is shown in the Table 2. The data shown (here for an evaluated

**Table 2** Excerpt of the performance test report, both for the period 1991–2010 (top) and for the year 2010 only (bottom)

Elem.	Bias	Onz	Chisq		Nonlin	Onz		<i>N</i>	min	Max
1991–2010										
Ag	1.058	0.02	15.1	*	−0.005	0.006		15	1.70E − 02	3.53E + 01
Al	1.009	0.003	4.3	VLD	0	0	VLD	34	3.00E + 00	9.99E + 05
As	0.997	0.005	1.8	VLD	0	0	VLD	39	3.80E − 02	6.25E + 02
Au	1.009	0.008	5.5	VLD	0	0	VLD	18	3.30E − 04	2.67E + 06
Ba	0.976	0.006	2.1	VLD	0	0	VLD	32	2.16E + 00	1.50E + 03
Br	0.991	0.009	2.4	VLD	0	0	VLD	30	3.60E − 01	1.00E + 06
Ca	1.011	0.003	2.4	VLD	0	0	VLD	37	1.16E + 02	3.27E + 05
Cd	0.965	0.005	5.9	VLD	0	0	VLD	30	1.30E − 02	4.74E + 02
Ce	1.022	0.009	0.7	VLD	0	0	VLD	28	1.20E − 01	1.56E + 07
Cl	1.031	0.008	1.4	VLD	0	0	VLD	23	3.50E + 01	3.10E + 04
Co	0.987	0.007	1.1	VLD	0.002	0.001	VLD	36	5.12E − 02	5.00E + 01
Cr	1.018	0.005	10.4	VLD	0	0	VLD	36	3.00E − 01	4.74E + 02
Cs	1.003	0.012	0.9	VLD	0.001	0.002	VLD	25	5.30E − 02	1.07E + 02
Cu	1.081	0.008		***	0	0	VLD	38	2.34E + 00	1.00E + 06
2010										
Ag	0.993	0.029	0.1	VLD	−0.001	0.003	VLD	3	1.34E − 01	3.53E + 01
Al	0.975	0.007	2.1	VLD	0	0	VLD	8	3.00E + 00	1.51E + 05
As	0.978	0.009	0.4	VLD	0	0	VLD	10	5.00E − 02	6.26E + 02
Au	0.987	0.024	1	VLD	0.052	0.041		5	2.10E − 03	3.93E + 00
Ba	0.961	0.01	1.1	VLD	0	0	VLD	8	4.11E + 01	1.50E + 03
Br	0.986	0.014	0.5	VLD	0	0	VLD	8	5.90E − 01	1.00E + 06
Ca	1.006	0.009	1.2	VLD	0	0	VLD	9	1.16E + 02	2.88E + 04
Cd	1.012	0.013	1.3	VLD	0	0	VLD	3	7.20E − 02	4.74E + 02
Ce	0.945	0.018	0.4	***	0	0	VLD	7	1.19E + 01	1.56E + 04
Cl	0.975	0.011	3.8	VLD	0	0	VLD	4	5.10E + 01	2.78E + 03
Co	1.002	0.016	0.5	VLD	0	0.003	VLD	9	1.53E − 01	5.00E + 01
Cr	1.029	0.008	2.8	VLD	0	0	VLD	8	8.74E + 00	4.74E + 02
Cs	1.007	0.025	0	VLD	0	0.002	VLD	6	5.94E − 01	1.07E + 02
Cu	1.018	0.017	2.3	VLD	0	0	VLD	6	3.70E + 00	2.95E + 03

The columns list for each element the bias, the degree of non-linearity, the number (*N*) of reference materials in the database used in the evaluation and the minimum and maximum mass fraction of the element in these reference materials

VLD means that the result is in agreement with the respective criterion, \*\*\* that it does not; \* that the result is borderline, and − that the test is inconclusive

period of one year only, 2010) are, for each element, the bias factor and its standard error, an indication of the element passing the trueness test or not, the slope of the regression line when correlating bias factors to certified concentrations and its uncertainty, an indication of the element passing the linearity test or not, the number of reference materials involved, and the lowest and highest certified concentrations for the element present in the dataset.

Out of the 67 in the scope, 51 elements pass the trueness test for the entire 1991–2010 period. For some, the test is inconclusive (the bias factor found is too imprecise; 2 cases, i.e. Er and Re), or there are no reference data at all (9 cases, i.e. Ge, Ir, Nb, Os, Rh, Ru, Te, Tm and Y). Other elements (5 cases, i.e. Cu, Gd, S, Si and U) failed the test. Over the year 2010, 38 elements passed the trueness test; the test was inconclusive for 13 elements and no reference data were available for 16 other elements.

### Linearity

The same raw data as discussed above demonstrate the linearity of the method. Not a single element fails the test in any of the time spans.

### Repeatability, reproducibility

In order to exclude sample-to-sample variations from this test, a “multi-standard” in-house material was used for this test, which is based on a solution with 20 elements present, routinely used as an internal stability check in a wide range of analysis procedures.

Of the 20 elements, over e.g., the period 2007–2008, Cr exhibits the smallest  $SEM_{int} = 0.096\%$ , with a  $SEM_{ext} = 0.085\%$ , at  $N = 169$  observations, including all analysis procedures, detectors and analysts. A  $\chi_r^2 = 0.78$  was calculated from

$$\chi_r^2 = \left( \frac{SEM_{ext}}{SEM_{int}} \right)^2$$

Since  $SEM_{int}$  is (slightly) larger than  $SEM_{ext}$  in this case, the question is what amount of extra variation  $V$  would raise  $SEM_{ext}$  significantly in the  $\chi_r^2$  sense. The answer is  $V = 0.475\%$ , which would give a  $\chi_r^2 = 1.18$ , calculated from

$$\chi_r^2 = \left( \frac{\sqrt{V^2/N + SEM_{ext}^2}}{SEM_{ext}} \right)^2$$

which is the threshold value for significance at  $\alpha = 0.05$  at  $N = 169$ .

The requirement for the magnitude of unexplained variation is that it should be less than 5%, so that the requirement for reproducibility turns out to be fulfilled.

For the repeatability test, the Cr results in multi-standard were used again, but this time only the results obtained in one specific end-user oriented project were utilized (all analyses by same person, same detector and only with slight variations in decay time).

$SEM_{ext} = 0.09\%$ , and  $SEM_{int} = 0.22\%$ , at  $N = 36$  observations. By the same procedure as described above, the amount of extra variation that would significantly increase  $SEM_{ext}$  is  $V = 0.35\%$  (which would give a  $\chi_r^2 = 1.42$ , being the threshold value for  $\alpha = 0.05$  at  $N = 36$ ), so that the repeatability criterion is met.

It may be noted that the reproducibility of these measurements is  $0.085\% \times \sqrt{169} = 1.1\%$ , and that the repeatability is  $0.09\% \times \sqrt{36} = 0.5\%$ . In both cases, the RID software slightly overestimates these values with  $0.096\% \times \sqrt{169} = 1.2\%$  and  $0.22\% \times \sqrt{36} = 1.2\%$ : It stays on the safe side when reporting “precision”.

### Cross-sensitivity

Some typical examples of potential errors due to presence of interferences are shown in Table 3. It is well-known that in NAA the determination of La by  $^{140}\text{La}$  is interfered by the presence of U because of the resulting  $^{140}\text{Ba}$  fission product formed, decaying via  $^{140}\text{La}$ . The determination of Zn via  $^{65}\text{Zn}$  is interfered by the presence of Sc, resulting in  $^{46}\text{Sc}$  and causing difficulties due to the limited energy resolution of the detectors (gamma-ray energies of  $^{65}\text{Zn}$  and  $^{46}\text{Sc}$  1115 and 1120 keV, respectively). The determination of Co via  $^{60}\text{Co}$  is interfered by the presence of Br and the formation of  $^{82}\text{Br}$ , leading to a spectral interference at the 1332.5 keV line of  $^{60}\text{Co}$  by a coincidence summing peak at 1330.8 keV (554.3 + 776.5 keV). Deviations from linearity by 0.0005 and less are considered negligible in views of the validity criterion of 0.005, so effort was made to show the values beyond four significant digits.

### Discussion

Out of the 67 in the scope, 51 elements as determined within the Laboratory for INAA passed the trueness test. The 16 elements that did not pass the trueness test receive currently special treatment until they pass the test at some point in the future, e.g. as a result of improvements to the system:

- When one of the seven elements (i.e. Er, Re, Cu, Gd, S, Si and U) with bias information available is reported to an end-user, it must be ascertained that the reported

**Table 3** Examples of assessments of potential cross sensitivity for different mass fraction ratios of interfering elements

Period	Nonlinearity	N	Mass fraction ratio		Conclusion
			From	To	
La in the presence of U, 1991–2010 dataset					
1991–2008	0.000 ± 0.000	28	0.05	3300	VLD
1999–2008	0.000 ± 0.000	26	0.05	3300	VLD
2004–2008	0.000 ± 0.000	21	0.54	3300	VLD
2007–2008	0.000 ± 0.000	10	7.7	600	VLD
2008–2010	0.000 ± 0.000	10	7.7	600	VLD
Verdict					VLD
Zn in the presence of Sc, 1991–2010 dataset					
1991–2008	0.000 ± 0.000	44	2.2	$6.6 \times 10^4$	VLD
1999–2008	0.000 ± 0.000	41	2.2	$6.6 \times 10^4$	VLD
2004–2008	0.000 ± 0.000	33	2.2	$6.6 \times 10^4$	VLD
2007–2008	0.000 ± 0.000	19	4.2	$1.2 \times 10^4$	VLD
2008–2010	0.000 ± 0.000	16	4.2	$1.2 \times 10^4$	VLD
Verdict					VLD
Co in the presence of Br, 1991–2010 dataset					
1991–2008	0.001 ± 0.001	37	$4.4 \times 10^{-4}$	63	VLD
1999–2008	−0.008 ± 0.007	34	$4.4 \times 10^{-4}$	17	
2004–2008	−0.007 ± 0.009	29	$6.1 \times 10^{-3}$	17	
2007–2008	−0.019 ± 0.018	14	$6.4 \times 10^{-3}$	17	
2008–2010	−0.012 ± 0.016	13	$6.4 \times 10^{-3}$	17	
Verdict					VLD

N number of different reference materials used. VLD means the test was passed

trueness does not imply a smaller possible bias than the one found in this test.

- For the nine elements (i.e. Ge, Ir, Nb, Os, Rh, Ru, Te, Tm and Y) where there are no data in this test: aliquots from single-element calibrators with metrologically traceable property values are included in the analysis runs if mass fractions are requested. These calibrators have been added to the database of reference materials, so that information that can be used in this test will be collected over time.
- One way or the other, efforts are being made to validate the method for these 16 elements.

The performance of INAA passes the linearity test without exception. For some elements, not enough data is available to draw a conclusion. However, not a single element fails the test. The mass fraction ranges used for the test span 6 orders of magnitude for some elements (like Al and Br). This result was expected from the scientific understanding of the technique. The INAA method also passes the cross-sensitivity, reproducibility and repeatability tests.

The approach described in this paper was accepted by the Dutch Council for Accreditation for compliance with the requirements in Clause 5.4 “Test and Calibration Methods and Method Validation” of the ISO/IEC17025:2005.

However, it should be noted that in principle the tests does not provide a conclusive answer on the validity of the routine test results since this validity depends on the target uncertainty as defined by the customer.

The methods described in this paper are also implemented in the  $k_0$ -IAEA software [9].

## Conclusion

Having a database of analysis results obtained from CRMs available makes it possible to assess a variety of quality parameters such as robustness, trueness, repeatability, and linearity. Such databases allow for inspection of trends but also for the effectiveness of corrective actions, e.g. by renewed calibration following for instance an unacceptable bias. Together, these assessments constitute the “validation” for the method. This paper demonstrates how that was done to the satisfaction of the Dutch Council for Accreditation.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

**References**

1. De Bruin M, Korthoven PJM (1972) *Anal Chem* 44:2382–2385
2. Obrušník I, Blaauw M, Bode P (1991) *J Radioanal Nucl Chem* 152:507–518
3. Bode P (2000) *J Radioanal Nucl Chem* 245:127–132
4. Blaauw M (1994) *Nucl Instr Meth* 353:269–271
5. Bode P, Van Dalen JP (1994) *J Radioanal Nucl Chem* 179:141–148
6. Bode P (1997) *J Radioanal Nucl Chem* 215:87–94
7. Blaauw M, Gelsema SJ (1999) *Nucl Instr Meth A* 422:417–422
8. Blaauw M (1999) *Nucl Instr Meth A* 432:74–76
9. Rossbach M, Blaauw M, Bacchi M, Lin X (2007) *J Radioanal Nucl Chem* 274:657–662