Supporting information for 'Round robin test of secondary raw materials: a systematic review of performance parameters'

0. Document information

This document relates to the following article:

Authors	Nathalie Korf, Paul Martin Mählitz, Vera Susanne Rotter
Affiliation	Technische Universität Berlin, Chair of Circular Economy and Recycling Technology, Straße des 17. Juni 135, 10623 Berlin, Germany
Title	Round robin test of secondary raw materials: a systematic review of performance parameters
Publication date	[insert when published]
Journal	Reviews in Analytical Chemistry
Volume	[insert when published]
Issue	[insert when published]
DOI	[insert when published]

It provides supporting, additional information on

1. round robin tests in general,
2. the methodology applied for the systematic literature review,
3. description and categorization of performance parameters for round robin tests,
4. detailed results of the systematic literature review, and
5. the bibliography with all references used in the systematic literature review.

Content

1. Introduction: round robin test

1.1. Definition and use

Round robin test (RRT), also interlaboratory comparison (ILC) or proficiency testing (PT), is a testing approach with a certain, pre-defined purpose, focusing on one or more samples and one or more measurement or testing methods, with the involvement of multiple labs and one organizing and coordinating institution. The purposes for which RRT are conducted can generally be categorized into purposes addressing a) lab performance and b) method performance. The coordinating institution organizes an RRT by i) inviting to an RRT, stating the purpose of the RRT and the benefit for participating labs, ii) sending out samples and rules/regulations for participation, including data requirements and deadline, iii) processing and assessing the submitted RRT data, and iv) informing the participating labs about the overall and the lab-specific result. (DIN EN ISO/IEC 17043, 2010; Prichard and Barwick, 2007)

1.2. Overview norms

This section describes international norms, effective in the EU, regulating good laboratory practice in general and RRT in particular.

DIN EN ISO/IEC 17025 'General requirements for the competence of testing and calibration laboratories' (17025, 2018) sets requirements for labs in general. Compliance with this norm can be understood as equal to good laboratory practice.

DIN EN ISO/IEC 17000 'Conformity assessment - Vocabulary and general principles' (17000, 2020) sets terms and definitions for all steps of an RRT.

DIN EN ISO/IEC 17011 'Conformity assessment - Requirements for accreditation bodies accrediting conformity assessment bodies' (2018) defines the requirements for institutions, which accredit laboratories. Both norms are mainly important on a regulatory and organisational level.

The two most important norms for the design, planning, execution, and evaluation of RRT are DIN EN ISO/IEC 17043 and DIN EN ISO/IEC 13528.

DIN EN ISO/IEC 17043 'Conformity assessment - General requirements for proficiency testing' (DIN EN ISO/IEC 17043, 2010) is the international norm describing the use and application of RRT. It gives an overview over the types of RRT and proficiency testing, the design, necessary terms, definitions, and statistical parameters as well as technical and management requirements to conduct a well-designed and sucessful RRT.

DIN EN ISO/IEC 13528 'Statistical methods for use in proficiency testing by interlaboratory comparison' (DIN ISO 13528, 2020) complements DIN EN ISO/IEC 17043 with a detailed description of statistical methods for RRT, guidelines for the design of RRT, determination of the reference/assigned value, determination of criteria for performance evaluation as well as calculation and graphical methods of performance parameters.

The norms DIN ISO 5725-1 to 5725-6 define principles of trueness and precision of measurement methods and hence set the standard how to assess the accuracy of an analysis. (5725-3, 2003; 5725-4, 2003; 5725-2, 2002; 5725-5, 2002; 5725-6, 2002; 5725-1, 1997)

2. Methodology systematic literature review

Figure 1 shows the flow chart for the systematic literature review applied in (Korf et al., 2022), based on (Liberati et al., 2009). It shows the four steps of the review, 1) identification of publications, 2) screening of the publications for eligibility, 3) extraction of pre-defined relevant information from the publications, and 4) assessment of the extracted information.

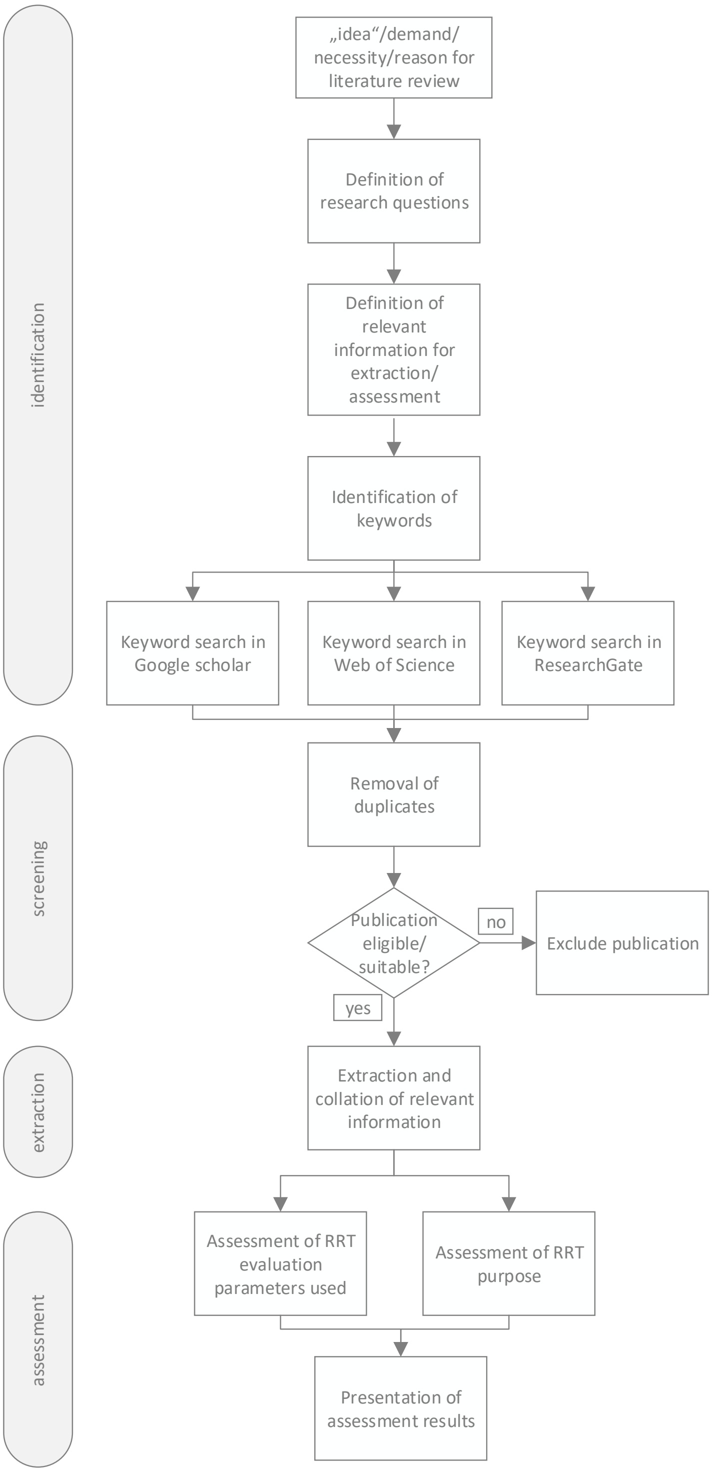

Figure 1: Flow chart methodology for systematic literature review, based on the PRISMA guidelines (Liberati et al., 2009)

3. Description performance parameters

3.1. Levels and cells: categorization of lab data

In the context of an RRT there is a multitude of data on different levels. For a better overview and a correct addressing of the data, we use the terms ‘level’ and ‘cell’ (Mandel, 1991). ‘Level’ describes the results for one element in one sample over all labs, e.g. results of four labs (mean or single values) for copper (Cu) in printed circuit boards (PCB). ‘Cell’ describes the results for the data of one lab in a level, e.g. results of lab 3 for Cu in PCB. Figure 2 shows the hierarchical relation between level and cell. The total number of elements is given with n, the related index is the chemical symbol (e.g. Cu). The total number of labs is indicated with p, the related index is j. For replicates, the number of the replicate is used as index (1, 2, 3…); the total number of replicates is m. Hence, the mass fraction of the second replicate for Cu in PCB of lab 3 is $x_{Cu,3,2}$. For mean values or standard deviations, no replicate index is necessary. This nomenclature is used for all equations throughout this document.

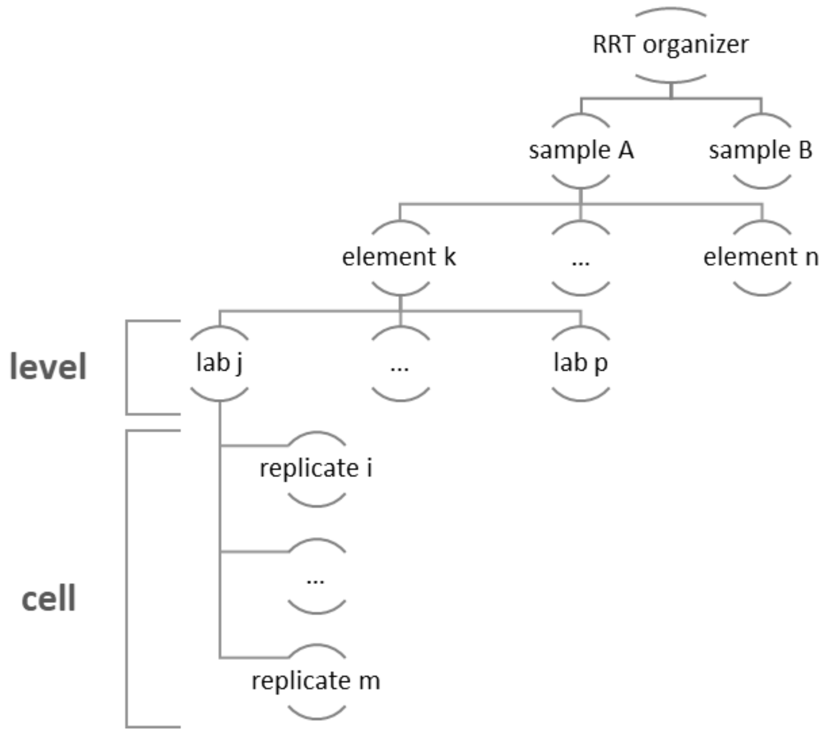

Figure 2: Schematic description of level and cell in the context of RRT data. Level describes the results for one element in one sample over all labs. Cell describes the results for the data of one lab in a level. The element index is k, the total number of elements n; the lab index is j, the total number of labs is p; the replicate index is i, the total number of replicates in one cell is m.

3.2. Accuracy: trueness and precision

DIN ISO 5725 describes accuracy of an analysis result using two parameters, ‘trueness’ and ‘precision’. The trueness of a measurement/analysis result refers to the extent of the deviation from the ‘true value’ or accepted reference value (i.e. “the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value”). The precision describes the variation within one data set (i.e. “the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value”). (5725-1, 1997)

The trueness is affected by e.g. selection of an unsuitable analysis method or inaccurate execution of an analysis. Analytical effects leading to a bias, i.e. deviation from the ‘true value’, can be categorized in systematic effects and random effects (JCGM 100:2008, 2008).

The precision is affected e.g. by the reliability of an analysis method and/or the homogeneity/ heterogeneity of the sample material (JCGM 100:2008, 2008).

3.3. Performance parameters for quality control of lab data

In general, performance parameters to determine the accuracy of chemical analysis results can be categorised with regard to two perspectives, i) the scope of the quality control measures (internal or external) and ii) the type of performance parameter (trueness or precision).

All trueness performance parameters (TPP) identified in (Korf et al., 2022) describe the (absolute or relative) deviation of the lab result to a pre-defined value (e.g. reference, assigned, or consensus value). This deviation is accompanied by different denominators, such as a mean value or a standard deviation. Precision performance parameters (PPP) describe the variability of a data set, which is oftentimes expressed as a (version of the) standard deviation.

Table 2 in (Korf et al., 2022) gives an overview over the most common performance parameters within these categories.

The following boxes 1 to 4 give an overview over general descriptive statistics (box 1) and all PP identified in (Korf et al., 2022), categorized in TPP (PP assessing the deviation of a value in box 2; significance tests in box 3) and PPP (parameters of variation and uncertainty in box 4).

Box 1: Descriptive statistics

Equations

All descriptive statistics are extracted from [Wilcox, 2009 and Wilcox, 2010].

Location parameters:

- arithmetic mean

  $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (1)

- median

  if data set has odd number n of values:

  $\tilde{x} = x_{(n+1)/2}$ (2)

  if data set has even number n of values:

  $\tilde{x} = \frac{x_{n/2} + x_{(n+1)/2}}{2}$ (3)

Parameters of variation:

- range

  $range = max - min$ (4)

- relative range

  $rel.range = \frac{range}{\bar{x}} \cdot 100\%$ (5)

- standard deviation

  $s_i = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$ (6)

- relative standard deviation

  $RSD = \frac{s_i}{\bar{x}}$ (7)

- MAD

  $MAD = median(x_i - \tilde{x})$ (8)

Parameter descriptions

Parameter	Description
n	number of values in a data set
x_i	single value in a data set
max	largest value in a data set
min	smallest value in a data set

Box 2: Parameters assessing the deviation of a value

Equations

- absolute deviation
  (DIN ISO 13528, 2020)

$$D = \left(\bar{x}_{j,k} - X_k\right) \quad (9)$$

- relative deviation
  (DIN ISO 13528, 2020)

$$RD = \left(\frac{\bar{x}_{j,k} - X_k}{X_k}\right) \cdot 100\% \quad (10)$$

- standardized bias
  (Johnston and Daniel, 1982)

$$e_{j,k} = \frac{\bar{x}_{j,k} - X_k}{s_{i,j}\sqrt{\frac{1}{m}\frac{p-1}{p}}} \quad (11)$$

- z score
  (DIN ISO 13528, 2020)

$$z = \frac{\bar{x}_{j,k} - X_k}{s_{RRT}} \quad (12)$$

- robust z score
  (Da Silva Dias et al., 2015)

$$z_{robust} = \frac{\sum_{i=1}^{m} x_i - median(\sum_{i=1}^{m} x_i)}{IQN(\sum_{i=1}^{m} x_i)} \quad (13)$$

- z' score
  (DIN ISO 13528, 2020)

$$z' = \frac{\bar{x}_{j,k} - X_k}{\sqrt{s_{RRT}^2 + u^2(\bar{x}_{j,k})}} \quad (14)$$

- z_U score
  (DIN 38402-45, 2014)

$$z_U = \begin{cases} \frac{g}{k_1} \cdot z\,, for\ z < 0 \\ \frac{g}{k_2} \cdot z\,, for\ z \geq 0 \end{cases} \quad (15)$$

- zeta score
  (DIN ISO 13528, 2020)

$$\zeta = \frac{\bar{x}_{j,k} - X_k}{\sqrt{u_{test}^2 + u_{RRT}^2}} \quad (16)$$

- E_n score
  (DIN ISO 13528, 2020)

$$E_n = \frac{\bar{x}_{j,k} - X_k}{\sqrt{U_{test}^2 + U_{RRT}^2}} \quad (17)$$

- Mandel's h
  (Mandel, 1991)

$$h_{i,j} = \frac{\bar{x}_{j,k} - C_k}{(s_L)_k} \quad (18)$$

Parameter descriptions

Parameter	Description
m	number of replicates in a cell
m_j	number of replicates for lab j
m_{ref}	number of replicates for reference lab
$\bar{x}_{j,k}$	mean value of all replicates for element k in lab j, $\bar{x}_{j,k} = \frac{1}{m_j}\sum_{i=1}^{m} x_{i,j,k}$ (19)
X_k	assigned/reference value; if produced from reference lab analyses: $X_k = \bar{x}_{ref} = \frac{1}{m_{ref}}\sum_{i=1}^{m} x_{i,j,k}$ (20)
p	number of laboratories
$s_{i,j}$	pooled standard deviation of the within-laboratory replicates adjusted for the number of replicates m and the number of labs p
s_{RRT}	standard deviation used in RRT; is selected depending on required informative value, e.g. SD of assigned value, between-lab SD, or as in equation (37)
IQN	normalized interquartile range, $IQN = 0.7313 \cdot IQR$ (21)
IQR	interquartile range, difference between 3. quartile and 1. quartile
u_{test}	participant's own estimate of the standard uncertainty of its result $x_{j,k}$, see par. descr. in box 4
u_{RRT}	standard uncertainty of the assigned value X_k, see par. descr. in box 4
U_{test}	expanded uncertainty of a participant's result $x_{j,k}$, see equation (39)
U_{RRT}	expanded uncertainty of the assigned value X_k, see equation (39)
k_1 k_2	$\left(k_2 + \frac{1}{v}\right) \cdot e^{\left\{-\frac{1}{2}k_2^2\right\}} = \left(-k_1 + \frac{1}{v}\right) \cdot e^{\left\{-\frac{1}{2}k_1^2\right\}}$ (22); $\left(1 - \Phi\left(-\frac{1}{v}\right)\right)^{-1} \left(\Phi(k_2) - \Phi(-k_1)\right) = 1 - \alpha$ (23)
C_k	average for one element in one level, i.e. over all laps $C_k = \frac{1}{p}\sum_{i=1}^{p} \bar{x}_{j,k}$ (24)
$s_{L,k}$	standard deviation between cell mean and level mean for one element, $s_{L,k} = \sqrt{\frac{\sum_{i=1}^{p}(\bar{x}_{j,k} - C_k)^2}{p-1}}$ (25)

Box 3: Significance tests

Equations

***t* test (Rand R. Wilcox, 2009):**

- test hypothesis: $\bar{x}_{j,k}$ is equal to X_k
- *t* statistic:
  - for equal number of replicates m and homoscedasticity:

$$t = \frac{\bar{x}_{j,k} - X_k}{s_p \sqrt{\left(\frac{1}{m_i} + \frac{1}{m_{ref}}\right)} \tag{26}$$

  - for unequal number of replicates m and heteroscedasticity:

$$W = \frac{\bar{x}_{j,k} - X_k}{\sqrt{\frac{s_{j,k}^2}{m_i} + \frac{s_{ref}^2}{m_{ref}}}} \tag{27}$$

- t_{crit}, W is defined by m_j and α
- test hypothesis is rejected for t, W > t_{crit}

u test (Shakhashiro and Toervenyi, 2010):

- test hypothesis: $\bar{x}_{j,k}$ is equal to X_k
- U statistic:

$$U = \frac{|\bar{x}_{j,k} - X_k|}{\sqrt{u_{test}^2 + u_{test}^2}} \tag{28}$$

- $\bar{x}_{j,k}$ is not significantly different, if U < 2.58

ANOVA (Rand R. Wilcox, 2009):

- test hypothesis: more than two means are equal to each other ($\bar{x}_1 = \bar{x}_2 = \cdots = \bar{x}_n$)
- $MS = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{df} = s^2$
- test statistic:

$$F = \frac{MS_{between-group}}{MS_{within-group}} \tag{29}$$

- means differ significantly for $F > F_{crit}$

Tukey's Honestly Significant Difference test (Tukey, 1977):

- test hypothesis: two or more than two means are equal to each other ($\bar{x}_1 = \bar{x}_2 = \cdots = \bar{x}_n$)
- test statistic:

$$HSD = q\sqrt{\frac{MS}{m}} \tag{30}$$

- honestly significant difference is shown, when the pairwise difference between two means is larger than HSD.

Parameter descriptions

Parameter	Description
s_p	mean standard deviation, $s_p = \sqrt{\frac{s_{j,k}^2 + s_{ref}^2}{2}}$ (31)
$s_{j,k}^2$	variance of lab data in one cell and one level, i.e. for one lab and one element, $s_{j,k}^2 = \frac{1}{m_j - 1}\sum_{i=1}^{m}\left(x_{i,j,k} - \bar{x}_{j,k}\right)^2$ (32)
$s_{j,k}$	Standard deviation of lab data in one cell and one level, i.e. for one lab and one element, $s_{j,k} = \sqrt{s_{j,k}^2} = \sqrt{\frac{1}{m_j - 1}\sum_{i=1}^{m}\left(x_{i,j,k} - \bar{x}_{j,k}\right)^2}$ (33)
s_{ref}^2	variance of reference lab data for one element, $s_{ref}^2 = \frac{1}{m_{ref} - 1}\sum_{i=1}^{m}\left(x_{i,ref} - \bar{x}_{ref}\right)^2$ (34)
m_j	number of replicates for lab j
m_{ref}	number of replicates for reference lab
α	probability of type I error, i.e. probability of a false positive
MS	mean square
F_{crit}	value of F distribution, for defined degrees of freedom from samples
df	degrees of freedom
q	value from studentized range distribution table for m and α

Box 4: Parameters of variation and uncertainty

Equations

- Range: see box 1, equation (4)
- Relative range: see box 1, equation (5)
- Standard deviation: see box 1, equation (6)
- Relative standard deviation: see box 1, equation (7)
- MAD: see box 1, equation (8)
- Between-lab standard deviation for element k (Mandel, 1991)

$$s_{L,k} = \sqrt{\frac{\sum_{i=1}^{p}\left(\bar{x}_{j,k}-C_k\right)^2}{p-1}} \quad (35)$$

- Repeatability standard deviation for element k (Mandel, 1991)

$$s_r = \sqrt{\frac{\sum_{j=1}^{p} s_{j,k}^2}{p}} \quad (36)$$

- Reproducibility standard deviation for element k (Horwitz and Albert, 2006)

$$s_R = \sqrt{s_L^2 + s_r^2} \quad (37)$$

- Standard deviation for round robin test (DIN ISO 13528, 2020)

$$s_{RRT} = \sqrt{s_R^2 - s_r^2\left(1 - \frac{1}{m}\right)} \quad (38)$$

- Mandel's k for element k in lab j (Mandel, 1991)

$$k_{j,k} = \frac{s_{j,k}}{s_r} \quad (39)$$

- (combined, expanded) uncertainty (GUM)

$$u_c^2(x_k) = \sum_{i=1}^{N}\left(\frac{\delta f}{\delta x_{i,j,k}}\right)^2 u^2\left(x_{i,j,k}\right) \quad (40)$$

Parameter descriptions

Parameter	Description
$u^2\left(x_{i,j,k}\right)$	standard uncertainty, determined by method A or method B, specified in (JCGM 100:2008, 2008)

4. Detailed results literature review

The following tables present the detailed results of the systematic literature review in (Korf et al., 2022). Tables 1 to 4 show the identified TPP and PPP in peer-reviewed articles as well as reports, conference proceedings, and papers. Tables 5 and 6 summarize the identified purposes, sample materials, and analysis types.

4.1. Performance parameters assessing the trueness TPP)

Table 1: Extracted information for trueness performance parameters in RRT in peer-reviewed articles

reference	deviation from value										significance tests			
	internal+external			external							internal+external			
	absolute deviation	relative deviation	standardized bias	z score	robust z score	z' score	z_U score	zeta score	En score	Mandel's h	t test	U test	Tukey's H.S.D. test	ANOVA
(Dreesen et al., 1979)		●												
(Domalski and Abramowitz, 1983)														
(Nair et al., 1984)														
(Steger et al., 1985)														
(LaFleur and Dodo, 1989)		●												
(Blankenhorn et al., 1992)														
(Desaules et al., 1992)														
(Crosland et al., 1993)														
(Kimbrough and Wakakuwa, 1994)														
(Groot and Hoede, 1994)														
(Quevauviller et al., 1994)														
(Argyraki et al., 1995)				●										
(Quevauviller et al., 1996)														
(Bögershausen et al., 1997)														
(Beck et al., 1997)														
(Kučera et al., 1997)														
(Cook et al., 1997)														
(Hinners et al., 1998)		●									●			
(Maaskant et al., 1998)														
(Kučera et al., 1998)														
(Quevauviller, 1998)														
(Nejedlý et al., 1998)		●												
(Butler and Howe, 1999)														
(Ottner et al., 2000)														
(Balaram, 2000)				●										
(Kleinman et al., 2001)											●			
(Becker et al., 2002)														
(Hall and Oates, 2003)														
(Cools et al., 2004)										●				
(Tahir et al., 2005)	●			●										
(Kreij and Wever, 2005)				●						●				
(Davidson et al., 2006)														
(Tirez et al., 2007)														
(Kalbe et al., 2008)														
(Balzamo et al., 2009)				●										
(Gerboles et al., 2011)				●		●		●						
(Wragg et al., 2011)														
(Ikonomou et al., 2012)													●	●
(Buczko et al., 2012)		●												
(Bürger et al., 2014)														
(Harrington et al., 2014)		●												

(Vittori Antisari et al., 2014)														
(Da Silva Dias et al., 2015)				●	●									
(Nudi et al., 2015)				●										
(Reis et al., 2015)				●										
(Bachmann et al., 2016)							●							
(Geurts et al., 2016)				●										
(Yatkin et al., 2016)		●							●					
(Gartiser et al., 2017)														
(Santoro et al., 2017)											●			
(Jagustyn et al., 2017)														
(Qiao et al., 2017)		●												
(Raven and Self, 2017)	●													
(Qiao et al., 2018)		●		●				●						
(Isobe et al., 2019)		●												
(Kalbe et al., 2019)										●				
(Anca-Couce et al., 2020)														
(Hafner et al., 2020)														
(Pellikka and Kajolinna, 2020)														
(Yatkin et al., 2020)		●		●										
(Damastuti et al., 2020)				●										
(Delvigne et al., 2021)														

Table 2: Extracted information for trueness performance parameters in RRT in reports, conference proceedings, and papers

reference	deviation from value										significance tests			
	internal+external			external							internal+external			
	absolute deviation	relative deviation	standardized bias	z score	robust z score	z’ score	z_U score	zeta score	En score	Mandel’ s h	t test	U test	Tukey’ s H.S.D. test	ANOVA
reports														
(Johnston and Daniel, 1982)		●	●											
(Kennedy et al., 1983)		●												
(Cools et al., 2003)										●				
(Shakhashiro et al., 2006)		●		●										
(Cools et al., 2007)										●				
(Ingham et al., 2007)		●		●										
(Mäkinen et al., 2008)				●										
(Leivuori et al., 2009)				●										
(Shakhashiro and Toervenyi, 2010)		●		●								●		
(Leivuori et al., 2011a)				●										
(Leivuori et al., 2011b)				●										
(Björkjöf et al., 2013)				●										
(Korhonen-Ylönen et al., 2013a)				●										
(Korhonen-Ylönen et al., 2013b)				●										
(Leivuori et al., 2013)				●										
(Koivikko et al., 2016)				●										
(Koivikko et al., 2017)		●		●					●					
(Leivuori et al., 2017)		●		●					●					
(Maunuksela et al., 2018)				●										
(Koivikko et al., 2018)				●					●					
(Koivikko et al., 2019)				●					●					
(Oosterlaken-Buijs, 2019)				●										
(Koivikko et al.)				●					●					
conference proceedings														
(van der Sloot et al.)														
(Saric and Grzunov)		●												
papers														
(Hesbach et al., 2014)		●												

4.2. Performance parameters to assess the precision

Table 3: Extracted information for precision performance parameters in RRT in peer-reviewed articles

reference	parameters of variation										uncer-tainty
	internal+external							external			int.+ext.
	relative range	standard deviation	relative standard deviation	median absolute deviation	precision	relative calculation error	confidence interval	repeatability standard deviation	reproducibility standard deviation	Mandel’ s k	(expanded, combined) uncertainty
(Dreesen et al., 1979)			●		●						
(Domalski and Abramowitz, 1983)		●						●	●		
(Nair et al., 1984)			●								
(Steger et al., 1985)							●				
(LaFleur and Dodo, 1989)											
(Blankenhorn et al., 1992)		●	●					●	●		
(Desaules et al., 1992)											
(Crosland et al., 1993)											
(Kimbrough and Wakakuwa, 1994)		●	●								
(Groot and Hoede, 1994)			●					●	●		
(Quevauviller et al., 1994)		●	●								
(Argyraki et al., 1995)											
(Quevauviller et al., 1996)			●								
(Bögershausen et al., 1997)		●									
(Beck et al., 1997)											
(Kučera et al., 1997)					●		●				
(Cook et al., 1997)		●									
(Hinners et al., 1998)			●								
(Maaskant et al., 1998)							●				●
(Kučera et al., 1998)											●
(Quevauviller, 1998)		●	●								
(Nejedlý et al., 1998)											
(Butler and Howe, 1999)		●									
(Ottner et al., 2000)							●				
(Balaram, 2000)					●						
(Kleinman et al., 2001)			●								
(Becker et al., 2002)			●				●				
(Hall and Oates, 2003)		●									
(Cools et al., 2004)		●						●	●	●	
(Tahir et al., 2005)		●	●	●				●			●
(Kreij and Wever, 2005)								●	●	●	
(Davidson et al., 2006)		●									
(Tirez et al., 2007)							●	●	●		
(Kalbe et al., 2008)											
(Balzamo et al., 2009)											
(Gerboles et al., 2011)		●						●	●		●
(Wragg et al., 2011)		●	●					●	●		
(Ikonomou et al., 2012)		●									
(Buczko et al., 2012)								●	●		●
(Bürger et al., 2014)		●									●
(Harrington et al., 2014)			●		●				●		
(Vittori Antisari et al., 2014)		●					●	●	●		
(Da Silva Dias et al., 2015)		●	●					●	●		
(Nudi et al., 2015)		●	●								
(Reis et al., 2015)		●	●	●							●
(Bachmann et al., 2016)											

(Geurts et al., 2016)			●					●	●		
(Yatkin et al., 2016)		●									●
(Gartiser et al., 2017)			●				●		●		
(Santoro et al., 2017)			●								●
(Jagustyn et al., 2017)		●						●	●		
(Qiao et al., 2017)					●						
(Raven and Self, 2017)											
(Qiao et al., 2018)											●
(Isobe et al., 2019)		●									
(Kalbe et al., 2019)		●	●					●	●	●	
(Anca-Couce et al., 2020)		●									
(Hafner et al., 2020)	●		●			●			●		
(Pellikka and Kajolinna, 2020)		●									
(Yatkin et al., 2020)			●								●
(Damastuti et al., 2020)											
(Delvigne et al., 2021)		●									●

Table 4: Extracted information for precision performance parameters in RRT in reports, conference proceedings, and papers

reference	parameters of variation										uncer-tainty
	internal+external							external			int.+ext.
	relative range	standard deviation	relative standard deviation	median absolute deviation	precision	relative calculation error	confidence interval	repeatability standard deviation	reproducibility standard deviation	Mandel’s k	(expanded, combined) uncertainty
reports											
(Johnston and Daniel, 1982)			●					●	●		
(Kennedy et al., 1983)		●									
(Cools et al., 2003)			●					●	●	●	
(Shakhashiro et al., 2006)					●						●
(Cools et al., 2007)			●					●	●	●	
(Ingham et al., 2007)		●	●					●	●		
(Mäkinen et al., 2008)		●	●								
(Leivuori et al., 2009)		●	●								●
(Shakhashiro and Toervenyi, 2010)											
(Leivuori et al., 2011a)		●	●								●
(Leivuori et al., 2011b)		●	●					●	●		●
(Björkjöf et al., 2013)		●	●								●
(Korhonen-Ylönen et al., 2013a)		●	●								
(Korhonen-Ylönen et al., 2013b)		●	●								
(Leivuori et al., 2013)		●	●								●
(Koivikko et al., 2016)		●	●					●	●		●
(Koivikko et al., 2017)		●	●								●
(Leivuori et al., 2017)		●	●								●
(Maunuksela et al., 2018)		●									●
(Koivikko et al., 2018)		●	●								●
(Koivikko et al., 2019)		●	●								●
(Oosterlaken-Buijs, 2019)								●	●		
(Koivikko et al.)		●									
conference proceedings											
(van der Sloot et al.)		●	●					●	●		
(Saric and Grzunov)		●	●					●			
papers											
(Hesbach et al., 2014)											

4.3. RRT purpose, sample material, and analysis type

Table 5: Extracted information regarding RRT purpose, sample material, and analysis type in peer-reviewed articles

reference	purpose									sample material				analysis type		
	lab assessment	method validation	method assessment	method development	(C)RM testing	establish assigned value(s)	assign consensus value(s)	material characterization	new/unknown sample	solid waste material	mineral material	soil	other material	elemental content	molecule content	other method
(Dreesen et al., 1979)	•		•								•			•		
(Domalski and Abramowitz, 1983)		•	•							•						•
(Nair et al., 1984)		•	•									•				•
(Steger et al., 1985)						•						•		•		
(LaFleur and Dodo, 1989)			•	•						•					•	
(Blankenhorn et al., 1992)	•											•			•	
(Desaules et al., 1992)	•		•									•		•		
(Crosland et al., 1993)		•	•									•		•		•
(Kimbrough and Wakakuwa, 1994)			•									•		•		
(Groot and Hoede, 1994)			•							•			•	•		
(Quevauviller et al., 1994)		•	•								•			•		
(Argyraki et al., 1995)	•		•									•		•		•
(Quevauviller et al., 1996)			•									•		•	•	
(Bögershausen et al., 1997)					•	•							•	•		
(Beck et al., 1997)			•									•		•		•
(Kučera et al., 1997)					•							•		•		
(Cook et al., 1997)			•								•			•		
(Hinners et al., 1998)					•								•	•		•
(Maaskant et al., 1998)					•			•					•	•		
(Kučera et al., 1998)					•							•		•		
(Quevauviller, 1998)		•				•						•		•		•
(Nejedlý et al., 1998)			•										•	•	•	•
(Butler and Howe, 1999)			•	•									•	•		
(Ottner et al., 2000)			•								•					•
(Balaram, 2000)			•									•		•		
(Kleinman et al., 2001)			•									•		•		•
(Becker et al., 2002)		•	•									•			•	
(Hall and Oates, 2003)	•		•								•	•		•		
(Cools et al., 2004)	•											•		•		•
(Tahir et al., 2005)	•												•	•		
(Kreij and Wever, 2005)			•									•		•	•	•
(Davidson et al., 2006)	•		•									•		•		
(Tirez et al., 2007)										•		•		•		
(Kalbe et al., 2008)			•							•	•			•	•	•
(Balzamo et al., 2009)	•											•			•	
(Gerboles et al., 2011)	•												•	•		
(Wragg et al., 2011)			•									•		•		•
(Ikonomou et al., 2012)			•										•		•	
(Buczko et al., 2012)	•											•		•		
(Bürger et al., 2014)	•						•	•			•			•		
(Harrington et al., 2014)		•	•											•		
(Vittori Antisari et al., 2014)					•							•		•		
(Da Silva Dias et al., 2015)	•		•									•		•		•
(Nudi et al., 2015)					•						•	•		•	•	
(Reis et al., 2015)			•								•	•		•		•
(Bachmann et al., 2016)			•							•						•
(Geurts et al., 2016)			•								•			•		
(Yatkin et al., 2016)								•					•	•		
(Gartiser et al., 2017)			•										•			•
(Santoro et al., 2017)			•										•	•		
(Jagustyn et al., 2017)					•				•	•						•
(Qiao et al., 2017)			•									•		•		•

(Raven and Self, 2017)	•										•			•		•
(Qiao et al., 2018)			•								•			•		
(Isobe et al., 2019)			•										•			•
(Kalbe et al., 2019)		•								•		•			•	
(Anca-Couce et al., 2020)			•										•			•
(Hafner et al., 2020)		•	•										•			•
(Pellikka and Kajolinna, 2020)			•										•	•		
(Yatkin et al., 2020)			•										•	•		
(Damastuti et al., 2020)		•											•	•		•
(Delvigne et al., 2021)					•							•		•	•	•

Table 6: Extracted information regarding RRT purpose, sample material, and analysis type in reports, conference proceedings, and papers

reference	purpose									sample material				analysis type		
	lab assessment	method validation	method assessment	method development	(C)RM testing	establish assigned value(s)	assign consensus value(s)	material characterization	new/unknown sample	solid waste material	mineral material	soil	other material	elemental content	molecule content	other method
reports																
(Johnston and Daniel, 1982)			•								•		•	•	•	
(Kennedy et al., 1983)			•									•		•		•
(Cools et al., 2003)	•		•									•		•		
(Shakhashiro et al., 2006)	•									•		•		•		•
(Cools et al., 2007)	•											•		•		
(Ingham et al., 2007)		•									•			•		
(Mäkinen et al., 2008)	•									•				•		
(Leivuori et al., 2009)	•											•		•		
(Shakhashiro and Toervenyi, 2010)	•											•		•		•
(Leivuori et al., 2011a)	•										•			•		
(Leivuori et al., 2011b)	•									•				•		
(Björkjöf et al., 2013)	•									•				•		
(Korhonen-Ylönen et al., 2013a)	•											•			•	
(Korhonen-Ylönen et al., 2013b)	•											•			•	
(Leivuori et al., 2013)	•											•		•		
(Koivikko et al., 2016)	•												•	•		
(Koivikko et al., 2017)	•									•				•		
(Leivuori et al., 2017)	•									•			•	•		
(Maunuksela et al., 2018)	•						•					•			•	•
(Koivikko et al., 2018)	•									•				•		
(Koivikko et al., 2019)	•									•				•		
(Oosterlaken-Buijs, 2019)	•												•	•		
(Koivikko et al.)	•									•				•		•
conference proceedings																
(van der Sloot et al.)			•	•						•				•		
(Saric and Grzunov)			•								•			•		
paper																
(Hesbach et al., 2014)			•										•	•		

References

Anca-Couce, A., Tsekos, C., Retschitzegger, S., Zimbardi, F., Funke, A., Banks, S., Kraia, T., Marques, P., Scharler, R., Jong, W. de, Kienzl, N., 2020. Biomass pyrolysis TGA assessment with an international round robin. Fuel 276, 118002. https://doi.org/10.1016/j.fuel.2020.118002.

Argyraki, A., Ramsey, M.H., Thompson, M., 1995. Proficiency Testing in Sampling: Pilot Study on Contaminated Land. Analyst 120, 2799–2803. https://doi.org/10.1039/AN9952002799.

Bachmann, H.J., Bucheli, T.D., Dieguez-Alonso, A., Fabbri, D., Knicker, H., Schmidt, H.-P., Ulbricht, A., Becker, R., Buscaroli, A., Buerge, D., Cross, A., Dickinson, D., Enders, A., Esteves, V.I., Evangelou, M.W.H., Fellet, G., Friedrich, K., Gasco Guerrero, G., Glaser, B., Hanke, U.M., Hanley, K., Hilber, I., Kalderis, D., Leifeld, J., Masek, O., Mumme, J., Carmona, M.P., Calvelo Pereira, R., Rees, F., Rombolà, A.G., de la Rosa, J. M., Sakrabani, R., Sohi, S., Soja, G., Valagussa, M., Verheijen, F., Zehetner, F., 2016. Toward the Standardization of Biochar Analysis: The COST Action TD1107 Interlaboratory Comparison. Journal of agricultural and food chemistry 64, 513–527. https://doi.org/10.1021/acs.jafc.5b05055.

Balaram, V., 2000. Assessment of the ICP-MS method using the interlaboratory QA study of two Polish soil RMs. Accreditation and Quality Assurance 5, 325–330. https://doi.org/10.1007/s007690000163.

Balzamo, S., Zorzi, P. de, Barbizzi, S., Calabretta, E., Potalivo, M., Rosamilia, S., 2009. Measurement of PAHs in environmental matrices: results from an interlaboratory comparison on the different steps of the measurement procedure. Accred Qual Assur 14, 487–495. https://doi.org/10.1007/s00769-009-0576-x.

Beck, T., Joergensen, R.G., Kandeler, E., Makeschin, F., Nuss, E., Oberholzer, H.R., Scheu, S., 1997. An inter-laboratory comparison of ten different ways of measuring soil microbial biomass C. Soil Biology and Biochemistry 29, 1023–1032. https://doi.org/10.1016/S0038-0717(97)00030-8.

Becker, R., Koch, M., Wachholz, S., Win, T., 2002. Quantification of total petrol hydrocarbons (TPH) in soil by IR-spectrometry and gas chromatography - conclusions from three proficiency testing rounds. Accreditation and Quality Assurance 7, 286–289. https://doi.org/10.1007/s00769-002-0476-9.

Björkjöf, K., Korhonen-Ylönen, K., Kaasalainen, M., Leivuori, M., Väntsi, S., Lanteri, S., Ilmakunnas, M., 2013. Proficiency Test SYKE 10/2012: Leaching testing of solid waste samples. Reports of the Finnish Environment Institute 19 | 2013. Finnish Environment Institute, 66 pp. http://hdl.handle.net/10138/39609.

Blankenhorn, I., Meijer, D., van Delft, R.J., 1992. Inter-laboratory comparison of methods used for analysing polycyclic aromatic hydrocarbons (PAHs) in soil samples. Fresenius Journal of Analytical Chemistry, 497–504. https://doi.org/10.1007/BF00322157.

Bögershausen, W., Cicciarelli, R., Gercken, B., König, E., Krivan, V., Müller-Käfer, R., Pavel, J., Seltner, H., Schelcher, J., 1997. Pure graphite as a reference material for the determination of trace elements - an interlaboratory collaborative study. Fresenius Journal of Analytical Chemistry, 266–273. https://doi.org/10.1007/s002160050151.

Buczko, U., Kuchenbuch, R.O., Übelhör, W., Nätscher, L., 2012. Assessment of sampling and analytical uncertainty of trace element contents in arable field soils. Environmental monitoring and assessment 184, 4517–4538. https://doi.org/10.1007/s10661-011-2282-5.

Bürger, S., Boulyga, S.F., Peńkin, M.V., Bostick, D., Jovanovic, S., Lindvall, R., Rasmussen, G., Riciputi, L., 2014. Quantifying multiple trace elements in uranium ore concentrates: an interlaboratory comparison. J Radioanal Nucl Chem 301, 711–729. https://doi.org/10.1007/s10967-014-3224-9.

Butler, O.T., Howe, A.M., 1999. Development of an international standard for the determination of metals and metalloids in workplace air using ICP-AES: evaluation of sample dissolution procedures through an interlaboratory trial. Journal of Environmental Monitoring (JEM), 23–32. https://doi.org/10.1039/A807526C.

Cook, J.M., Gardner, M.J., Griffiths, A.H., Jessep, M.A., Ravenscroft, J.E., Yates, R., 1997. The comparability of sample digestion techniques for the determination of metals in sediments. Marine Pollution Bulletin 34, 637–644. https://doi.org/10.1016/S0025-326X(96)00186-5.
Cools, N., Delanote, V., Scheldeman, X., Quataert, P., Vos, B. de, Roskams, P., 2004. Quality assurance and quality control in forest soil analyses: a comparison between European soil laboratories. Accreditation and Quality Assurance 9, 688–694. https://doi.org/10.1007/s00769-004-0856-4.
Cools, N., Delanote, V., Vos, B. de, Quataert, P., Roskams, P., Scheldeman, X., 2003. Quality Assurance and Quality Control in Forest Soil Analysis: 3rd FSCC Interlaboratory Comparison. Forest Soil Co-ordinating Centre, Institute for Forestry and Game Management, 303 pp. https://purews.inbo.be/ws/files/275633/184556.pdf.
Cools, N., MIkkelsen, J.H., Vos, B. de, 2007. Quality Assurance and Quality Control in Forest Soil Analysis: 5th FSCC Interlaboratory Comparison. Research Institute for Nature and Forest, 64 pp. https://pureportal.inbo.be/portal/files/7284389/5thFSCC_InterlabComp_report.pdf.
Crosland, A.R., McGrath, S.P., Lane, P.W., 1993. An Interlaboratory Comparison of a Standardised EDTA Extraction Procedure for the Analysis of Available Trace Elements in Two Quality Control Soils. International Journal of Environmental Analytical Chemistry 51, 153–160. https://doi.org/10.1080/03067319308027620.
Da Silva Dias, R., Aparecida de Abreu, C., Ferreira de Abreu, M., Paz González, A., 2015. Statistical Methods for Evaluating Results from Soil Micronutrient Analyses in Interlaboratory Programs. Communications in Soil Science and Plant Analysis 46, 57–71. https://doi.org/10.1080/00103624.2014.988091.
Damastuti, E., Santoso, M., Yusuf, S., Yatu N. S., W., 2020. An interlaboratory comparison of INAA analytical method for coal fly ash elemental characterization. IOP Conf. Series Journal of Physics: Conference Series 1436, 12138. https://doi.org/10.1088/1742-6596/1436/1/012138.
Davidson, C.M., Urquhart, G.J., Ajmone-Marsan, F., Biasioli, M., Da Costa Duarte, A., Díaz-Barrientos, E., Grčman, H., Hossack, I., Hursthouse, A.S., Madrid, L., Rodrigues, S., Zupan, M., 2006. Fractionation of potentially toxic elements in urban soils from five European cities by means of a harmonised sequential extraction procedure. Analytica Chimica Acta 565, 63–72. https://doi.org/10.1016/j.aca.2006.02.014.
Delvigne, C., Guihou, A., Schuessler, J.A., Savage, P., Poitrasson, F., Fischer, S., Hatton, J.E., Hendry, K.R., Bayon, G., Ponzevera, E., Georg, B., Akerman, A., Pokrovsky, O.S., Meunier, J.-D., Deschamps, P., Basile-Doelsch, I., 2021. Silicon Isotope Analyses of Soil and Plant Reference Materials: An Inter-Comparison of Seven Laboratories. Geostand Geoanal Res. https://doi.org/10.1111/ggr.12378.
Desaules, A., LIscher, P., Dahinden, R., Bachmann, H.J., 1992. Comparability of chemical analysis of heavy metals and fluorine in soils: results of an interlaboratory study. Communications in Soil Science and Plant Analysis 23, 363–377. https://doi.org/10.1080/00103629209368595.
Deutsches Institut für Normung e.V., 2018. DIN EN ISO/IEC 17025 General requirements for the competence of testing and calibration laboratories: DIN EN ISO/IEC 17025:2018. Deutsches Institut für Normung, 65 pp.
Deutsches Institut für Normung e.V., 2020. DIN ISO 13528 Statistical methods for use in proficiency testing by interlaboratory comparisons, 207 pp.
DIN EN ISO, 2010. DIN EN ISO/IEC 17043 Conformity assessment - General requirements for proficiency testing, 93 pp.
DIN EN ISO, 2018. DIN EN ISO/IEC 17011 Conformity assessment - Requirements for accreditation bodies accrediting conformity assessment bodies, 63 pp.
DIN EN ISO, 2020. DIN EN ISO/IEC 17000 Conformity assessment - Vocabulary and general principles, 73 pp.

DIN ISO, 1997. DIN ISO 5725-1 Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions, 44 pp.
DIN ISO, 2002. DIN ISO 5725-6 -Accuracy (trueness and precision) of measurement methods and results - Part 6: Use in practice of accuracy values, 91 pp.
DIN ISO, 2002. DIN ISO 5725-5 Accuracy of measurement methods and results - Part 5: Accuracy (trueness and precision) of measurement methods and results - Part 5: Alternative methods for the determination of the precision of a standard measurement method, 123 pp.
DIN ISO, 2002. DIN ISO 5725-2 Accuracy (trueness and precision) of measurement methods and results - Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method, 97 pp.
DIN ISO, 2003. DIN ISO 5725-4 - Accuracy (trueness and precision) of measurement methods and results - Part 4: Basic methods for the determination of the trueness of a standard measurement method, 57 pp.
DIN ISO, 2003. DIN ISO 5725-3 - Accuracy (trueness and precision) of measurement methods and results - Part 3: Intermediate measures of the precision of a standard measurement method, 65 pp.
Domalski, E.S., Abramowitz, S., 1983. Evaluation of data on calorific content values determined during ASTM round robin testing of RDF-3. Resources and Conservation, 3, 233–252. https://doi.org/10.1016/0166-3097(83)90027-5.
Dreesen, D.R., Gladney, E.S., Owens, J.W., 1979. Interlaboratory Comparison of Arsenic, Molybdenum and Selenium Analyses from Uranium Mill Tailings. Journal (Water Pollution Control Federation) 51, 2447–2456.
Gartiser, S., Heisterkamp, I., Schoknecht, U., Burkhardt, M., Ratte, M., Ilvonen, O., Brauer, F., Brückmann, J., Dabrunz, A., Egeler, P., Eisl, A.-M., Feiler, U., Fritz, I., König, S., Lebertz, H., Pandard, P., Pötschke, G., Scheerbaum, D., Schreiber, F., Soldán, P., Weiß, R., Weltens, R., 2017. Results from a round robin test for the ecotoxicological evaluation of construction products using two leaching tests and an aquatic test battery. Chemosphere 175, 138–146. https://doi.org/10.1016/j.chemosphere.2017.01.146.
Gerboles, M., Buzica, D., Brown, R., Yardley, R.E., Hanus-Illnar, A., Salfinger, M., Vallant, B., Adriaenssens, E., Claeys, N., Roekens, E., Sega, K., Jurasović, J., Rychlik, S., Rabinak, E., Tanet, G., Passarella, R., Pedroni, V., Karlsson, V., Alleman, L., Pfeffer, U., Gladtke, D., Olschewski, A., O'Leary, B., O'Dwyer, M., Pockeviciute, D., Biel-Ćwikowska, J., Turšič, J., 2011. Interlaboratory comparison exercise for the determination of As, Cd, Ni and Pb in PM10 in Europe. Atmospheric Environment 45, 3488–3499. https://doi.org/10.1016/j.atmosenv.2010.12.020.
Geurts, R., Spooren, J., Quaghebeur, M., Broos, K., Kenis, C., Debaene, L., 2016. Round robin testing of a percolation column leaching procedure. Waste management (New York, N.Y.) 55, 31–37. https://doi.org/10.1016/j.wasman.2016.06.010.
Groot, G.J. de, Hoede, D., 1994. Validation of Dutch standard leaching tests using NEN-ISO 5725. Environmental Aspects of Construction with Waste Materials 60, 305–314. https://doi.org/10.1016/S0166-1116(08)71467-5.
Hafner, S.D., Fruteau de Laclos, H., Koch, K., Holliger, C., 2020. Improving Inter-Laboratory Reproducibility in Measurement of Biochemical Methane Potential (BMP). Water 12, 1752. https://doi.org/10.3390/w12061752.
Hall, G.E.M., Oates, C.J., 2003. Performance of commercial laboratories in analysis of geochemical samples for gold and the platinum group elements. Geochemistry: Exploration, Environment, Analysis 3, 107–120. https://doi.org/10.1144/1467-787303-006.
Harrington, J.M., Nelson, C.M., Weber, F.X., Bradham, K.D., Levine, K.E., Rice, J., 2014. Evaluation of methods for analysis of lead in air particulates: an intra-laboratory and inter-laboratory

comparison. Environmental science. Processes & impacts 16, 256–261. https://doi.org/10.1039/c3em00486d.

Hesbach, P., Beck, M., Eick, M., Daniels, W.L., Burgers, C., Greiner, A., Hassett, D.J., Heebink, L.V., 2014. Inter-laboratory Comparison of Leaching Methods. U.S. Departoment of Energy, National Energy Technology Laboratory, 28 pp. https://www.researchgate.net/publication/237282890_Interlaboratory_Comparison_of_Leaching_Methods?_iepl%5BgeneralViewId%5D=xd0YlWt8WnFnGDhHm9GMAqlvJox4lwae0cHm&_iepl%5Bcontexts%5D%5B0%5D=searchReact&_iepl%5BviewId%5D=kyxo7fpIB1XftZqtr4DRHEjqnWykThMO99mm&_iepl%5BsearchType%5D=publication&_iepl%5Bdata%5D%5BcountLessEqual20%5D=1&_iepl%5Bdata%5D%5BinteractedWithPosition1%5D=1&_iepl%5Bdata%5D%5BwithoutEnrichment%5D=1&_iepl%5Bposition%5D=1&_iepl%5BrgKey%5D=PB%3A237282890&_iepl%5BtargetEntityId%5D=PB%3A237282890&_iepl%5BinteractionType%5D=publicationTitle.

Hinners, T.A., Hughes, R., Outridge, P.M., Davis, W.J., Simon, K., Woolard, D.R., 1998. Interlaboratory comparison of mass spectrometric methods for lead isotopes and trace elements in NIST SRM 1400 Bone Ash. Journal of Analytical Atomic Spectrometry, 963–970. https://doi.org/10.1039/A803373K.

Ikonomou, M.G., Kelly, B.C., Blair, J.D., Gobas, F. A. P. C., 2012. An interlaboratory comparison study for the determination of dialkyl phthalate esters in environmental and biological samples. Environmental toxicology and chemistry 31, 1948–1956. https://doi.org/10.1002/etc.1912.

Ingham, M.N., Gowing, C.J.B., Harrison, H.M., 2007. The validation of the determination of trace elements by energy dispersive polarised x-ray fluorescence spectrometry. Laboratory Operations Internal Report OR/07/012. British Geological Survey, 121 pp. http://nora.nerc.ac.uk/id/eprint/7196 (accessed 9 October 2020).

Isobe, A., Buenaventura, N.T., Chastain, S., Chavanich, S., Cózar, A., DeLorenzo, M., Hagmann, P., Hinata, H., Kozlovskii, N., Lusher, A.L., Martí, E., Michida, Y., Mu, J., Ohno, M., Potter, G., Ross, P.S., Sagawa, N., Shim, W.J., Song, Y.K., Takada, H., Tokai, T., Torii, T., Uchida, K., Vassillenko, K., Viyakarn, V., Zhang, W., 2019. An interlaboratory comparison exercise for the determination of microplastics in standard sample bottles. Marine Pollution Bulletin 146, 831–837. https://doi.org/10.1016/j.marpolbul.2019.07.033.

Jagustyn, B., Plis, A., Mastalerz, M., Hrabak, J., Ściążko, M., 2017. Investigation of homogeneity and stability of items for proficiency testing of solid recovered fuels (SRF) analysis. Accred Qual Assur 22, 355–360. https://doi.org/10.1007/s00769-017-1283-7.

JCGM - Joint Committee for Guides in Metrology, 2008. Evaluation of measurement data - Guide to the expression of uncertainty in measurement, 134 pp.

Johnston, J.W., Daniel, J.L., 1982. Summary report for the interlaboratory round robin on the MCC-1 static leach test method. U.S: Department of Energy, 185 pp.

Kalbe, U., Berger, W., Eckardt, J., Simon, F.-G., 2008. Evaluation of leaching and extraction procedures for soil and waste. Waste management (New York, N.Y.) 28, 1027–1038. https://doi.org/10.1016/j.wasman.2007.03.008.

Kalbe, U., Lehnik-Habrink, P., Bandow, N., Sauer, A., 2019. Validation of European horizontal methods for the analysis of PAH, PCB and dioxins in sludge, treated biowaste and soil. Environ Sci Eur 31, 704. https://doi.org/10.1186/s12302-019-0211-3.

Kennedy, P.C., Roser, B.P., HUnt, J.L., DAly, B.K., 1983. A New Zealand interlaboratory comparison of analytical data for the CSSC reference soil: N.Z. Soil Bureau Scientific Report 59 44p. Department of Scientific and Industrial Research, 46 pp.

Kimbrough, D.E., Wakakuwa, J., 1994. Interlaboratory Comparison of Instruments Used for the Determination of Elements in Acid Digestate of Solids. Analyst. https://doi.org/10.1039/AN9941900383.

Kleinman, P.J.A., Sharpley, A.N., Gartley, K., Jarrell, W.M., Kuo, S., Menon, R.G., Myers, R., Reddy, K.R., Skogley, E.O., 2001. Interlaboratory comparison of soil phosphorus extracted by various soil test methods. Communications in Soil Science and Plant Analysis 32, 2325–2345. https://doi.org/10.1081/CSS-120000376.
Koivikko, R., Leivuori, M., Kaasalainen, M., Tervonen, K., Lanteri, S., Ilmakunnas, M. Interlaboratory Proficiency Test 04/2020: Leaching behavior test for waste material: Two stage batch leaching test.
Koivikko, R., Leivuori, M., Kaasalainen, M., Tervonen, K., Lanteri, S., Ilmakunnas, M., 2017. Interlaboratory Proficiency Test 12/2016: Leaching test for solid waste samples: Two stage batch leaching test. Reports of the Finnish Environment Institute 12, 112 pp. https://helda.helsinki.fi/handle/10138/181510.
Koivikko, R., Leivuori, M., Kaasalainen, M., Tervonen, K., Lanteri, S., Ilmakunnas, M., 2018. Interlaboratory Proficiency Test 14/2017: Leaching behavior test for solid waste material: Up-flow percolation test. Reports of the Finnish Environment Institute 18, 136 pp. https://helda.helsinki.fi/handle/10138/236326.
Koivikko, R., Leivuori, M., Kaasalainen, M., Tervonen, K., Lanteri, S., Ilmakunnas, M., 2019. Interlaboratory Proficiency Test 13/2018: Leaching beahior test for solid waste material: One stage batch leaching test. Reports of the Finnish Environment Institute, 54 pp. https://helda.helsinki.fi/handle/10138/301217.
Koivikko, R., Leivuori, M., Näykki, T., Sara-Aho, T., Tervonen, K., Lanteri, S., Väisanen, R., Ilmakunnas, M., 2016. Interlaboratory Proficiency Test 08/2015: Metals in waste water and sludge. Reports of the Finnish Environment Institute 11 | 2016. Finnish Environment Institute, 150 pp. http://hdl.handle.net/10138/160817 (accessed 9 October 2020).
Korf, N., Mählitz, P.M., Rotter, V.S., 2022. Round robin tests of secondary raw materials: a systematic review of performance parameters. Reviews in Analytical Chemistry 2022, tbd, tbd.
Korhonen-Ylönen, K., Nuutinen, J., Leivuori, M., Ilmakunnas, M., 2013a. Proficiency Test SYKE 8/2012: Volatile organic compounds in water and soil. Reports of the Finnish Environment Institute 7 | 2013. Finnish Environment Institute, 80 pp. https://helda.helsinki.fi/handle/10138/38618 (accessed 9 October 2020).
Korhonen-Ylönen, K., Nuutinen, J., Leivuori, M., Ilmakunnas, M., 2013b. Proficiency Test SYKE 9/2012: Oil hydrocarbons in water and soil. Reports of the Finnish Environment Institute 8 | 2013. Finnish Environment Institute, 40 pp. http://hdl.handle.net/10138/38619 (accessed 9 October 2020).
Kreij, C. de, Wever, G., 2005. Proficiency Testing of Growing Media, Soil Improvers, Soils, and Nutrient Solutions. Communications in Soil Science and Plant Analysis 36, 81–88. https://doi.org/10.1081/CSS-200042971.
Kučera, J., Sychra, V., Horáková, J., Soukal, L., 1997. Use of INAA in the preparation of a set of soil reference materials with certified values of total element contents. Journal of Radioanalytical and Nuclear Chemistry 215, 147–155. https://doi.org/10.1007/BF02109891.
Kučera, J., Sychra, V., Koubek, J., 1998. A set of four soil reference materials with certified values of total element contents and their extractable fractions. Fresenius Journal of Analytical Chemistry 360, 402–405. https://doi.org/10.1007/s002160050722.
LaFleur, L.E., Dodo, G.H., 1989. An interlaboratory comparison of analytical procedures for the measurement of PCDDs/PCDFs in pulp and paper industry solid wastes. Chemosphere, 1-6, 77–84. https://doi.org/10.1016/0045-6535(89)90107-0.
Leivuori, M., Koivikko, R., Sara-Aho, T., Näykki, T., Tervonen, K., Lanteri, S., Väisanen, R., Ilmakunnas, M., 2017. Interlaboratory Proficiency Test 10/2016: Metals in waste water and recycled material. Reports of the Finnish Environment Institute 8 | 2017. Finnish Environment Institute, 124 pp. http://hdl.handle.net/10138/176910 (accessed 8 October 2020).

Leivuori, M., Korhonen, K., Järvinen, O., Näykki, T., Sara-Aho, T., Tervonen, K., Lanteri, S., Ilmakunnas, M., 2009. SYKE Proficiency Test 4/2009: Metals in waters and soils. Reports of the Finnish Environment Institute 28 | 2009. Finnish Environment Institute, 130 pp. https://helda.helsinki.fi/bitstream/handle/10138/39684/SYKEre_28_2009.pdf?sequence=1.
Leivuori, M., Korhonen, K., Sara-Aho, T., Näykki, T., Järvinen, O., Tervonen, K., Lanteri, S., Ilmakunnas, M., 2011a. SYKE Proficiency Test 3/2010: Metals in waters and sediment. Reports of the Finnish Environment Institute 1 | 2011. Finnish Environment Institute. https://helda.helsinki.fi/bitstream/handle/10138/39997/SYKEre_1_2011.pdf?sequence=4 (accessed 9 October 2020).
Leivuori, M., Korhonen-Ylönen, K., Sara-Aho, T., Näykki, T., Tervonen, K., Lanteri, S., Ilmakunnas, M., 2011b. Proficiency Test SYKE 3/2011: Metals in water and sludge. Reports of the Finnish Environment Institute 22 | 2011. Finnish Environment Institute, 114 pp. https://helda.helsinki.fi/bitstream/handle/10138/39762/SYKEre_22_2011.pdf?sequence=1.
Leivuori, M., Korhonen-Ylönen, K., Sara-Aho, T., Näykki, T., Tervonen, K., Lanteri, S., Ilmakunnas, M., 2013. SYKE Proficiency Test 5/2012: Metals in waters and soil. Reports of the Finnish Environment Institute 1 | 2013. Finnish Environment Institute, 134 pp. http://hdl.handle.net/10138/41753 (accessed 9 October 2020).
Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P.A., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS medicine 6, e1000100. https://doi.org/10.1371/journal.pmed.1000100.
Maaskant, J.F.N., Boekholt, A.H., Jenks, P.J., Rucinski, R.D., 1998. A international interlaboratory study for the production of a sewage sludge certified reference material for routine use in inorganic quality control. Analytical chemistry, 406–409. https://doi.org/10.1007/s002160050723.
Mäkinen, I., Vaajasaari, K., Järvinen, O., Sara-Aho, T., Ivalo, R., Tervonen, K., Ilmakunnas, M., 2008. SYKE Proficiency Test 4/2018: Leading testing of a solid waste - the one stage and the two stage batch leaching test. Reports of the Finnish Environment Institute. Finnish Environment Institute, 51 pp.
Mandel, J., 1991. The validation of measurement through interlaboratory studies. Chemometrics and Intelligent Laboratory Systems 11, 109–119.
Maunuksela, L., Pelkonen, A., Björkjöf, K., Ilmakunnas, M., Kartio, M., Leivuori, M., 2018. Interlaboratory Comparison test 15/2018: Soil improver maturity test. Reports of the Finnish Environment Institute 25, 59 pp. https://www.ruokavirasto.fi/globalassets/laboratoriopalvelut/vertailulaboratoriotoiminta/julkaisuja-pdf/sykere_25_2018.pdf.
Nair, P.S., Logan, T.J., Sharpley, A.N., Sommers, L.E., Tabatabai, M.A., Yuan, T.L., 1984. Interlaboratory Comparison of a Standardized Phosphorus Adsorption Procedure. J. environ. qual. 13, 591–595. https://doi.org/10.2134/jeq1984.00472425001300040016x.
Nejedlý, Z., Campbell, J.L., Teesdale, W.J., Dlouhy, J.F., Dann, T.F., Hoff, R.M., Brook, J.R., Wiebe, H.A., 1998. Inter-laboratory comparison of air particulate monitoring data. Journal of the Air & Waste Management Association (1995) 48, 386–397. https://doi.org/10.1080/10473289.1998.10463698.
Nudi, A.H., Wagener, A., Gabardo, I.T., Lourenço, R.A., Scofield, A., 2015. Interlaboratory Comparison of Quantifying Hydrocarbons and Trace Elements in Sediment Samples from a Tropical Estuary. Journal of the Brazilian Chemical Society 26, 297–309. https://doi.org/10.5935/0103-5053.20140280.

Oosterlaken-Buijs, G.A., 2019. Results of Proficiency Test Total Metals in Polymers. Institute for Interlaboratory Studies Spijkenisse, the Netherlands, 47 pp. http://www.iisnl.com/pdf/iis19P06.pdf (accessed 9 October 2020).
Ottner, F., Gier, S., Kuderna, M., Schwaighofer, B., 2000. Results of an inter-laboratory comparison of methods for quantitative clay analysis. Applied Clay Science, 5-6, 223–243. https://doi.org/10.1016/S0169-1317(00)00015-6.
Pellikka, T., Kajolinna, T., 2020. Emission measurements of heavy metals with the European standard reference methods EN 14385 and EN 13211 - observations from an interlaboratory comparison (ILC) measurements performed at waste-to-energy plant in Finland. Journal of the Air & Waste Management Association (1995), 1–19. https://doi.org/10.1080/10962247.2020.1797926.
Prichard, E., Barwick, V., 2007. Quality Assurance in Analytical Chemistry, 1st ed. Wiley.
Qiao, J., Lagerkvist, P., Rodushkin, I., Salminen-Paatero, S., Roos, P., Lierhagen, S., Jensen, K.A., Engstrom, E., Lahaye, Y., Skipperud, L., 2018. On the application of ICP-MS techniques for measuring uranium and plutonium: a Nordic inter-laboratory comparison exercise. J Radioanal Nucl Chem 315, 565–580. https://doi.org/10.1007/s10967-018-5697-4.
Qiao, J., Salminen-Paatero, S., Rondahl, S., Bourgeaux-Goget, M., Roos, P., Lagerkvist, P., Strålberg, E., Ramebäck, H., 2017. Inter-laboratory exercise with an aim to compare methods for 90Sr and 239,240Pu determination in environmental soil samples. J Radioanal Nucl Chem 314, 813–826. https://doi.org/10.1007/s10967-017-5385-9.
Quevauviller, P., 1998. Operationally defined extraction procedures for soil and sediment analysis I. Standardization. Trends in Analytical Chemistry 17. https://doi.org/10.1016/S0165-9936(97)00119-2.
Quevauviller, P., Lachica, M., Barahona, E., Rauret, G., Ure, A., Gomez, A., Muntau, H., 1996. Interlaboratory comparison of EDTA and DTPA procedures prior to certification of extractable trace elements in calcareous soil. The Science of the total environment 178, 127–132. https://doi.org/10.1016/0048-9697(95)04804-9.
Quevauviller, P., Rauret, G., Muntau, H., Ure, A.M., Rubio, R., López-Sánchez, J.F., Fiedler, H.D., Griepink, B., 1994. Evaluation of a sequential extraction procedure for the determination of extractable trace metal contents in sediments. Fresenius Journal of Analytical Chemistry, 808–814. https://doi.org/10.1007/BF00323110.
Raven, M.D., Self, P.G., 2017. Outcomes of 12 Years of the Reynolds Cup Quantitative Mineral Analysis Round Robin. clays clay miner 65, 122–134. https://doi.org/10.1346/CCMN.2017.064054.
Reis, A.T., Duarte, A.C., Henriques, B., Coelho, C., Lopes, C.B., Mieiro, C.L., Tavares, D.S., Ahmad, I., Coelho, J.P., Rocha, L.S., Cruz, N., Monteiro, R.J.R., Rocha, R., Rodrigues, S., Pereira, E., 2015. An international proficiency test as a tool to evaluate mercury determination in environmental matrices. TrAC Trends in Analytical Chemistry 64, 136–148. https://doi.org/10.1016/j.trac.2014.08.015.
Santoro, A., Held, A., Linsinger, T.P.J., Perez, A., Ricci, M., 2017. Comparison of total and aqua regia extractability of heavy metals in sewage sludge: The case study of a certified reference material. Trends in analytical chemistry : TRAC 89, 34–40. https://doi.org/10.1016/j.trac.2017.01.010.
Saric, M.M., Grzunov, J. Proficiency testing: Experiences from Croatia on the issue of heavy metals determination in marine sediments, in: Proceedings of The First International Proficiency Testing Conference, Sinaia, Romania, 11th-13th October, 2007, pp. 177–182.
Shakhashiro, A., Sansone, U., Trinkl, A., Benesch, T., 2006. Report on the IAEA-CU-2006-01 proficiency test on the determination of radionuclides and trace elements in soil and compost 5. International Atomic Energy Agency, 94 pp.
Shakhashiro, A., Toervenyi, A., 2010. Latin American regional proficiency test on the determination of trace elements and radionuclides in algae, soil and spiked water. Terrestrial Environment

Laboratory, Chemistry Unit, IAEA, 179 pp. https://nucleus.iaea.org/rpst/Documents/report_rla_pt.pdf (accessed 9 October 2020).

Steger, H.F., Bowman, W.S., McKeague, J.A., 1985. Additional Certified Values for CCRMP Soils SO-1 to -4. Geostandards Newsletter 9, 213–218. https://doi.org/10.1111/j.1751-908X.1985.tb00451.x.

Tahir, S.S., Naseem, R., ul Haq, A., Saeed, K., 2005. An inter-laboratory comparison study for the determination of copper and lead from the wastewater of printed circuit board manufacturing industry in Pakistan. Accred Qual Assur 10, 362–368. https://doi.org/10.1007/s00769-005-0022-7.

Tirez, K., Scharf, H., Calzolari, D., Cleven, R., Kisser, M., Lück, D., 2007. Validation of a European standard for the determination of hexavalent chromium in solid material. Journal of Environmental Monitoring (JEM) 9, 749–759. https://doi.org/10.1039/b706724k.

van der Sloot, H.A., van der Wegen, G. J. L., Hoede, D., Groot, G.J. de. Intercomparison of leaching tests for stabilized waste, in: Proceedings of the WASCON '94 Conference, 1-3 June 1994, Maastricht, The Netherlands.

Vittori Antisari, L., Bianchini, G., Dinelli, E., Falsone, G., Gardini, A., Simoni, A., Tassinari, R., Vianello, G., 2014. Critical evaluation of an intercalibration project focused on the definition of new multi-element soil reference materials (AMS-MO1 and AMS-ML1). EQA - Environmental quality 15, 41–64. https://doi.org/10.6092/ISSN.2281-4485/4553.

Wragg, J., Cave, M., Basta, N., Brandon, E., Casteel, S., Denys, S., Gron, C., Oomen, A., Reimer, K., Tack, K., van de Wiele, T., 2011. An inter-laboratory trial of the unified BARGE bioaccessibility method for arsenic, cadmium and lead in soil. The Science of the total environment 409, 4016–4030. https://doi.org/10.1016/j.scitotenv.2011.05.019.

Yatkin, S., Belis, C.A., Gerboles, M., Calzolai, G., Lucarelli, F., Cavalli, F., Trzepla, K., 2016. An interlaboratory comparison study on the measurement of elements in PM10. Atmospheric Environment 125, 61–68. https://doi.org/10.1016/j.atmosenv.2015.10.084.

Yatkin, S., Trzepla, K., Hyslop, N.P., White, W.H., Butler, O., Ancelet, T., Davy, P., Gerboles, M., Kohl, S.D., McWilliams, A., Saucedo, L., van der Haar, M., Jonkers, A., 2020. Comparison of a priori and interlaboratory-measurement-consensus approaches for value assignment of multi-element reference materials on PTFE filters. Microchemical Journal 158, 105225. https://doi.org/10.1016/j.microc.2020.105225.