Journal of Chemical Technology and Metallurgy, 59, 4, 2024, 967-978

EVALUATION OF MEASUREMENT UNCERTAINTY ASSOCIATED WITH THE SAMPLE PREPARATION PROCEDURE

Ava Amideina[1], Petranka Petrova[1], Mitja Kolar[2], Jernej Imperl[2], Petko Mandjukov[1]

[1]South-West University "Neofit Rilski"
Faculty of Natural Sciences and Mathematics
66 Ivan Mihailov Blvd., Blagoevgrad 2700, Bulgaria
[2]University of Ljubljana
Faculty of Chemistry and Chemical Technology
113 Večna pot, SI-1000 Ljubljana, Slovenia
E-mail: pmandjukov@abv.bg

Received 20 September 2023
Accepted 18 January 2024
DOI: 10.59957/jctm.v59.i4.2024.27

ABSTRACT

The measurement uncertainty (MU) in environmental analyses is usually considered as a combination of the additive contributions from sampling, sample preparation and analytical measurement (usually instrumental). The target MU for such studies is usually relatively high due to the natural sample heterogeneity and/or the complicated pretreatment procedure required. In the analytical practice, the contribution of the sample preparation to MU is rarely evaluated. Thus, it remains a hidden part of the one related to the analysis. However, the knowledge for the contribution of sample preparation might provide important information and further possibility for optimization of the entire analytical procedure and reduction of the expanded MU of the analytical result.

From statistical point of view, the separation of uncertainty contributions from the different steps is, generally, not a trivial task. Selection of the proper statistical approach depends on the data structure and quality, variables distribution, etc. In the present study, three different statistical methods for evaluation of the uncertainty contribution of the sample preparation were applied, compared, and discussed. The considered approaches are based on, both, classical and robust analysis of variances (ANOVA) applied to data from instrumental analysis of marine algae and bee honey samples, undergoing microwave digestion. Some general recommendations on the statistical approach selection are revealed based on real experimental data set.

Keywords: uncertainty, sample preparation, ANOVA, robust statistics, range estimation, marine algae, bee honey, environmental analysis.

INTRODUCTION

The environmental analysis, in principle, consists of three main phases: sampling; sample preparation and analytical measurement. Each of these steps contributes to the variability of the result. Some of the specific contributions to MU at different levels might be summarized as following:

- *Sampling* – spatial and/or temporal variability of the sampling target; level of the "at sampling point" representativeness of the acquired sample; its stability; possibility for contamination or analyte losses etc.
- *Sample preparation* – subsampling (acquiring laboratory sample); uncontrollable sample digestion processes; unstable procedural blank; quality of the reagents used; laboratory equipment; ambient conditions etc.
- *Measurement (analysis)* – quality grade of the reference materials used; possible differences in analyte chemical forms (species); calibration; instrumental repeatability etc.

Differentiation of MU contributions requires specially designed experiment with replicate procedures at all levels evaluated. The data obtained from such kind of experiments has clearly hierarchical structure (Fig. 1).

The basic level is the measurement (analysis), since this phase provides, the numerical values being assigned to the measurand [1, 2]. Analytical MU can be evaluated as a pooled standard deviation from the replicates from all digestions [3]. The differences between separate samples and digestions have a minor effect on this estimate. The next level, variance between parallel digestions of the samples, always includes the contribution of the analysis. Therefore, the calculation of standard deviation (classical or robust) from the results obtained from multiple digestions of one sample cannot provide reliable evaluation of the MU of sample preparation. The same holds true for the results from multiple sampling from the same target. The contributions of the sample preparation and measurement are always included in the directly calculated standard deviation between the separate samples.

The variance between targets, comprises the natural distinctions between them, as well as the MUs at all levels mentioned above, including sampling uncertainty. It should be noted that, being obtained from a data belonging to different populations, it has no exact physical meaning. It is simply, an approximate indication about the differences between the targets studied.

In addition, when the study is carried out with a single sampling from different targets, which is the case in the present work, separation of the sampling uncertainty from the one between targets variability is impossible. The variance between samples remains the highest level possible to evaluate and it comprises the sampling MU and the variability between targets. In principle, the variance at the highest level in hierarchy never has a physical meaning, since it is always formed by numerical values belonging to different populations.

The major impacts to the MU of the results from environmental analysis are due to the sampling and sample preparation procedures.

From statistical point of view, the evaluation of MU of sampling and sample preparation are very similar tasks. As being relatively new concept, a higher attention attracts evaluation of MU related to the sampling procedure [4 - 10]. Several tutorials and handbooks on this subject have been published during the last decade [7, 8]. For such evaluations, it is convenient to treat the entire analytical procedure simply as a combination of sampling and analytical measurement. When the sample preparation MU is not an object of study it remains a hidden component of the one related to measurement.

Evaluation of the uncertainty contribution of the sample preparation itself is described in rather limited number of publications [9 - 14]. Normally, the results reported, are based on single statistical approach selected without arguments and presented without particular details on the calculations.

In most of the publications, dedicated both to MU due to sampling and sample preparation, as a statistical method is referred as “Robust ANOVA”. This is rather wide term including large number of different

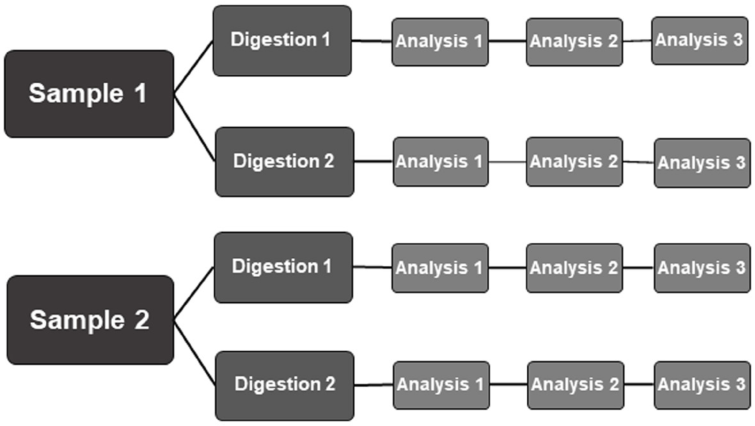

Fig. 1. The experimental data structure allowing separate evaluation of the contributions from various levels to the MU of the analytical result.

statistical approaches. It is also worth to note the lack of comparisons between different algorithms applicable to such problems in the literature.

The separation of MU due to the different sources is necessary for the following purposes:

- General requirements of the international standard ISO/IEC 17025:2017 for analytical method validation, separate evaluation of MU contributions, risk assessment etc. [15];
- Optimization, evaluation and validation of "in house" developed analytical and sample preparation procedures;
- Intra laboratory quality control.

Mathematical section / Algorithms used

The data structure is clearly hierarchical (Fig. 1). In the sequence:

TARGET – SAMPLE – DIGESTION – MEASUREMENT

each level represents a particular population with specific size, type of statistical distribution and values of the parameters. Besides this, every population, in fact, is a sample (in the statistical meaning of the term) taken from the previous one in the sequence. This structure of the data should be considered selecting statistical approaches.

Three algorithms were selected for data processing in the present work: *classical hierarchical (nested) ANOVA*, *range evaluation* and *robust ANOVA* (based on median estimation). As a comparative method the RSC's software – ROBAN was also applied.

The main difference between the algorithms compared is the way of estimation of the most probable value (μ) and dispersion (σ) of the population, based on the limited in size data set (statistical sample). It should be noted that depending on the number and structure of the experimental data points, different statistical approaches might be preferable. In the present work the experiments were planned in a way to obtain balanced data set: the same number of digestions from each sample (2) and the same number of replicates (3) from each digestion.

Hierarchical (nested) ANOVA

The basic requirements to the data set for application of nested ANOVA are generally the same as for the most popular version – the single parameter (one - way) ANOVA [16]:

- Each sample (in statistical meaning) is drawn from a normally distributed population;
- Homoscedasticity - all replicate measurements have uniform variances;
- Within each sample, the digestions/measurements are performed randomly and independently of each other;
- Uncertainty contributions at every level studied are entirely additive.

Considering the presumption for normal distribution of all populations the estimates for $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are mean and standard deviation respectively. Hierarchical ANOVA is far less popular than the single parameter one. Due to this reason is worth briefly to present the correct algorithm. Furthermore, deeper understanding the way of obtaining the results allows to reveal the strong points and limitations of each method, what is impossible when a "black box" type software is used. In Table 1 are presented the basic equations for evaluation of the variance components using nested ANOVA.

The variance, by definition might be expressed as a sum of squared deviations from the most probable value divided by the corresponding degrees of freedom, i.e. squared standard deviation [16, 17]. It is equivalent to the respective squared standard MU. Therefore, starting from the lowest level - analysis:

$$u_a = \sqrt{V_A} \tag{18}$$

where: u_a is standard MU due to analysis of type "A" [1, 2], calculated as a pooled standard deviation from all replicates measured [3].

The variance (V_D) between the parallel digestions within every single sample marks the next level. It consists of two components: V_A and MU due to the sample preparation (u_{SP}). The sample preparation component is present in every single replicate measurement. Considering assumption entirely additive nature of the effects and applying the uncertainty propagation law [2], the V_D can be expressed as:

$$V_D = V_A + nA \times u_{SP}^2 \tag{19}$$

The MU related to sample preparation procedure can be derived from Equation 19 as follows:

Table 1. Summary of hierarchical ANOVA equations.

Description	Equation	No
Means		
Mean value ($\overline{X_{ij}}$) of all replicates for digestion j from sample i	$\overline{X_{ij}} = \frac{1}{nA}\sum_{k=1}^{nA} X_{ijk}$	1
Mean value ($\overline{X}_i$) from all digestions from sample (i)	$\overline{X}_i = \frac{1}{nD}\sum_{j=1}^{nD} \overline{X_{ij}}$	2
Mean value ($\overline{X}$) from all samples (total mean)	$\overline{X} = \frac{1}{nS}\sum_{i=1}^{nS} \overline{X}_i$	3
Sums of squares		
Between replicates	$SS_A = \sum_{i=1}^{nS}\sum_{j=1}^{nD}\sum_{k=1}^{nA}\left(X_{ijk} - \overline{X_{ij}}\right)^2$	4
Between digestions	$SS_D = nA\sum_{i=1}^{nS}\sum_{j=1}^{nD}\left(\overline{X_{ij}} - \overline{X}_i\right)^2$	5
Between samples	$SS_S = nDnA\sum_{i=1}^{nS}\left(\overline{X}_i - \overline{X}\right)^2$	6
Total	$SS_T = \sum_{i=1}^{nS}\sum_{j=1}^{nD}\sum_{k=1}^{nA}\left(X_{ijk} - \overline{X}\right)^2$	7
Degrees of freedom		
Between replicates	$df_A = nSnD(nA - 1)$	8
Between digestions	$df_D = nS(nD - 1)$	9
Between samples	$df_S = nS - 1$	10
Total	$df_T = nSnDnA - 1$	11
Control equations		
Sums of squares	$SS_T = SS_S + SS_D + SS_A$	12
Degrees of freedom	$df_T = df_S + df_D + df_A$	13
Variances		
Between replicates	$V_A = \frac{SS_A}{df_A}$	14
Between digestions	$V_D = \frac{SS_D}{df_D}$	15
Between samples	$V_S = \frac{SS_S}{df_S}$	16
Total	$V_T = \frac{SS_T}{df_T}$	17

Symbols used: nA, nD, nS - numbers of replicate analyses, parallel digestions and samples, respectively; X_{ijk} *– result obtained for replicate k from digestion j, from sample i.*

$$u_{SP} = \sqrt{\frac{V_D - V_A}{nA}} \quad (20)$$

As mentioned earlier, the uncertainty of sampling cannot be evaluated in the present study due to the collection of single samples from each target. It should be noted that a single parameter ANOVA is applicable for u_{SP} evaluation only in the case of multiple digestion of different portions from the same sample or reference material [17].

Range method

The estimate for μ is the midrange of the data set and for σ - the range divided by a corresponding statistical factor depending on the number of data points [3]. The basics of the method are described in the international standard ISO 3085 [18]. Range evaluation belongs to the group of robust statistical methods. When the range method is applied for evaluation of the variance components, in fact, the approach is a simplified version of robust ANOVA.

The method might be applied in two different ways. The direct evaluation of relative MU from the particular source is described in details and practical examples in the EURACHEM guide [7]. For evaluation of absolute MU, modification of the method is required. The modified equations are presented in Table 2.

In most of the cases, the relative uncertainty is required to be declared. Therefore, when is expressed as absolute one, it should be transformed. In such a case, it is most logical to assign the absolute MU to the midrange (or mean) from all data points. The absolute value is a relevant estimate when the measurand values are varying in a relatively narrow range.

Robust ANOVA (based on median)

Robust ANOVA is a large group of statistical approaches based on various techniques and parameter evaluation methods from the robust (distribution free) statistics. Considering the aim of the present work, all methods based on ranks and signs are not applicable. Median and median of absolute deviations (*MAD*) are among the most popular robust estimates of μ and σ, respectively. It is important to notice that *MAD* is not directly equivalent to the standard deviation, hence it cannot be suggested as an MU estimate. As a robust estimate of the dispersion, compatible with standard deviation might be used normalized median of absolute deviations (*MADN*) [17]. It has a meaning of standard measurement uncertainty from type A. The algorithm for separation of the uncertainty contributions is similar to those for the classical hierarchical ANOVA with changed

Table 2. Summary equations for range evaluation of the absolute MU due to different sources.

Description	Equation	No
Ranges / Midranges		
Range (R_{ij}) of all replicates for digestion *j* from sample *i*	$R_{ij} = max(X_{ij1} \dots X_{ijnA}) - min(X_{ij1} \dots X_{ijnA})$	21
Midrange ($\overline{R_{ij}}$) of all replicates for digestion *j* from sample *i*	$\overline{R_{ij}} = \frac{1}{2}\left(max(X_{ij1} \dots X_{ijnA}) + min(X_{ij1} \dots X_{ijnA})\right)$	22
Range (R_i) of all digestions from sample *i*	$R_i = max(\overline{R_{i1}} \dots \overline{R_{inD}}) - min(\overline{R_{i1}} \dots \overline{R_{inD}})$	23
Standard measurement uncertainties		
MU associated with analysis	$u_A = \frac{1}{k_{st}(nS \times nD)} \sum_{i=1}^{nS} \sum_{j=1}^{nD} R_{ij}$	24
MU associated with sample preparation	$u_D = \sqrt{\left(\frac{1}{k_{st} \times nS} \sum_{i=1}^{nS} R_i\right)^2 - \frac{u_A^2}{nA}}$	25

Symbols used: k_{st} - statistical factor (k_{st} = 1.128 and k_{st} = 1.693 for number of data points 2 and 3 respectively).

Table 3. Summary equations for robust ANOVA evaluation of the absolute MU due to different sources.

Description	Equation	No
Median		
Median ($\widetilde{X_{ij}}$) of all replicates for digestion *j* from sample *i*	$\widetilde{X_{ij}} = median\ (X_{ij1} \dots X_{ijnA})$	26
Median ($\widetilde{X_i}$) of all digestions from sample (*i*)	$\widetilde{X_i} = median\ (\widetilde{X_{i1}} \dots \widetilde{X_{inD}})$	27
Median ($\tilde{X}$) from all samples (total median)	$\tilde{X} = median(\widetilde{X_1} \dots \widetilde{X_{nS}})$	28
Absolute deviations and MU		
Between replicates	$d_{ijk} = \lvert X_{ijk} - \widetilde{X_{ij}} \rvert$ $(i = 1 \dots nS; j = 1 \dots nD; k = 1 \dots nA)$	29
MU (u_A) due to analysis (*MADN*)	$u_A = \frac{median(d_{111} \dots d_{nSnDnA})}{0.6745}$	30
Between digestions	$d_{ij} = \lvert \widetilde{X_{ij}} - \widetilde{X_i} \rvert$	31
MU (u_{A+SP}) due to sample preparation including u_A	$u_{A+SP} = \frac{median(d_{11} \dots d_{nSnD})}{0.6745}$	32
MU (u_{SP}) due to sample preparation (only)	$u_{SP} = \sqrt{(u_{SP+A})^2 - (u_A)^2}$	33

estimates and respective modifications (Table 3).

Contrary the range method and nested ANOVA, Eq. 33 does not contain factor related to the number of measurement replicates (*nA*). The reason is the way of median calculation. The median in principle is not directly affected by the number of data points from which is derived.

The algorithms presented in Tables 1 - 3 can be easily realized as a templates in the MS EXCEL software.

EXPERIMENTAL

Case of study

The results from analysis of marine algae and bee honey were used to test the selected statistical approaches. In the present study, single samples were taken from different targets applying judgmental (expert decision based) sampling strategy. That makes impossible evaluation of the MU related to sampling procedure. It cannot be separated from the variation between targets. However, the aim of the study is MU contribution of the sample preparation.

From all samples, two laboratory taken portions were digested independently and each final solution obtained was measured in three replicates (Fig. 1).

The chemical elements measured were Al, Ca, Co, Cr, Cu, Fe, K, Li, Mg, Mn, Na, P, Rb, Sr, Ti, V and Zn for marine algae and Al, Ca, Fe, K, Mg, Mn, Na and Zn for bee honey samples. These were the elements presenting in the samples at concentration levels above the corresponding limits of quantification.

Measurand definition – Concentrations of various analytes (as listed above) expressed in mg kg^{-1} in air-dry material, determined using ICP-OES after microwave sample digestion.

Analysis description

Reagents

All reagents used were of analytical grade. Nitric acid (65 %, LabExpert), hydrogen peroxide (30 %, Fluka) and hydrochloric acid (37 %, Fluka) were used for digestion of both types of samples. A multi-element stock standard solution: Periodic table mix 1 for ICP 10 mg L^{-1} (Sigma Aldrich), was used as a primary standard for calibration. The working calibration standards were prepared by subsequent dilution with 1 % (v/v) HNO_3 in volumetric vessels.

Sample digestion

An identical sample digestion procedure using high-pressure laboratory microwave oven (Milestone Ethos UP, Italy), was applied to both types of samples studied.

A sample portions of ~ 0.5 g were measured using analytical balance Entris 224i-1S, Sartorius, Germany and transferred in 50 mL PTFE vessels for subsequent microwave digestion. A mixture consisting of 8 mL concentrated HNO_3 and 2 mL H_2O_2 was added to each vessel. The microwave digestion was performed by following temperature program: 30 min ramp time to 210°C followed by 20 min hold time at the same temperature (both at power of 1800 W). After completing the program and the subsequent cooling, the digested samples were transferred in 50 mL volumetric flasks. Two mL concentrated HCl were added to each sample before the final dilution. A procedural blank was prepared and treated together with each batch of digested samples. It contained the same amounts of reagents and undergo identical procedure.

ICP-OES measurements

The determination of all analytes in both types of samples was carried out using Agilent 5100 simultaneous vertical dual view Inductively Coupled Plasma Optical Emission Spectrometer (ICP-OES). The instrumental parameters for ICP-OES measurements of studied elements were as follow: RF incident power (1.2 kW); plasma argon flow rate (12 L min^{-1}); auxiliary argon flow rate (1 L min^{-1}); nebulizer argon flow rate (0.7 L min^{-1}); nebulizer (concentric), spray chamber (cyclonic double pass); axial viewing.

An ultrapure MiliQ grade water (ASTM Type 1, 18.2 MΩ cm specific resistivity at 25°C) was used in all operations. All critically important volumes were measured using volumetric devises of ASTM accuracy class “A”.

RESULTS AND DISCUSSION

Since a gravimetric sample quantification was carried out, normalization of the results was required before the MU contributions evaluation. The best way for such normalization is the calculation of the analyte concentration in the sample for each separate replicate. This provides comparability of all values and guarantees avoiding the normalization coefficients effects on the MU. The statistical approaches described above were applied to the data for analyte content in air - dry samples expressed in mg kg^{-1} (Table 4 A, B).

Depending on the intended use, the MU might be expressed as an absolute or relative value. The relative one is more informative for comparisons involving different analytes and/or samples. Generally, the higher concentration of the elements in samples is always accompanied by higher absolute MU and it propagates to all levels being evaluated. Both hierarchical and robust ANOVA are oriented to evaluation of absolute variances. In order to transform the absolute MU of sample preparation to relative is necessary to relate it to the defined concentration. The total MU evaluation is based on all replicates for all digestions from all samples. The logical solution is to use for such purpose overall mean or median including all values used in calculations. Normally, for high enough number of data points and especially, if the measurand values are in the same order of magnitude, both estimates are not significantly different.

Comparison between algorithms

All statistical approaches applied were validated in two ways: “step by step” checking the procedure for compliance with the algorithm description and by direct comparison with results for the same analyte (data set) obtained using ROBAN software provided by Analytical Method Committee (AMC) in Royal Society of Chemistry (RSC) [3, 19]. This is widely used “as is” software from reliable source. The results are referred as “Classical ANOVA” and “Robust ANOVA”, practically without detailed description of both methods. The comparisons were carried out for limited number of analytes. In all cases a perfect agreement between results from hierarchical ANOVA and ROBAN - Classical ANOVA was observed. This is quite intelligible observation when the data structure is taken into account. The results from ROBAN - Robust ANOVA are very close to the range method which also provides significant robustness. The results from median based robust ANOVA are usually the lowest obtained.

The trend in differences between the u_{SP} values obtained by the different algorithms studied can be summarized as:

Hierarchical ANOVA > Range method > Robust ANOVA (median).

This sequence also demonstrates the decrease in sensitivity of the different approaches to deviations from normality of the distributions.

Hierarchical (nested) ANOVA is a powerful statistical tool allowing besides separation of uncertainty contributions, checking various statistical hypotheses. Nevertheless, its applicability is strongly limited by the requirements for normal distributions of all populations and samples. The method is extremely sensitive to deviations from normality. It should be noted that small number of samples (2 - 3 replicates / digestions), even derived from normally distributed population are showing significant skewness, which itself indicates a deviation from the normality. They also might have significantly different distribution parameter estimates as compared with the population. Generally, the nested ANOVA trends to considerable overestimation of the MU from sample preparation in case of small number of digestions and measurement replicates.

Range method is a method having significant robustness. It is less sensitive to deviations from normality of the distributions than the classical ANOVA. For the small statistical samples the range method provides relatively more reliable evaluations than the other approaches studied.

Robust ANOVA is practically the least sensitive from all three compared approaches to the type of statistical distribution of data or presence of outliers. However, it also requires some minimum number of data points to reveal the strong points of the method. In case of small number of replicates, the method trends to underestimation of the MU due to sample preparation.

Comparison between analytes

For all analytes in the two types of samples determined in the present study estimated values for u_{SP} and u'_{SP} are varying in a wide range (Table 4, Fig. 2).

The u_{SP} values are naturally relevant to the analyte concentration level in the sample. Generally, the higher content of certain analyte in a sample implies higher overall MU, which spreads at all levels of contributions, including sample preparation. However, it is not the only reason for differences between analytes. Each one has individual origin defining its specific distribution in the sample and respectively in subsamples undergoing preparation procedures. It makes the interpretation of the observed results quite complicated. Besides this, the ratio between estimates obtained by algorithms indicates similarity and dissimilarity of the data sets distributions for the observed analytes. Since u_{SP} is a concentration dependent value, it is better to make comparisons based on the relative MU (u'_{SP}).

The highest MU from sample preparation in algae samples was found for the elements typical for the lithosphere or marine sediments: Ca, Al, Sr, Fe, Mn, and Ti. Since these elements are not distinctively biogenic (Al is even phytotoxic), they possibly are leached from external particles incorporated in the samples. Such unrepeatable external contamination is a possible explanation for the excessively high MU from sample preparation estimated by any of the algorithms.

Typically, biogenic elements for plant samples: P and Mg (part of the chlorophyll molecule) are showing moderate u'_{SP} values and similar ratios between the MU from sample preparation values (Fig. 2A) obtained by different algorithms.

Effect of the matrix

The same sample preparation procedure was applied to two different matrices and only the analytical results above the limit of quantification (LOQ) for the corresponding analyte were taken into account. As a

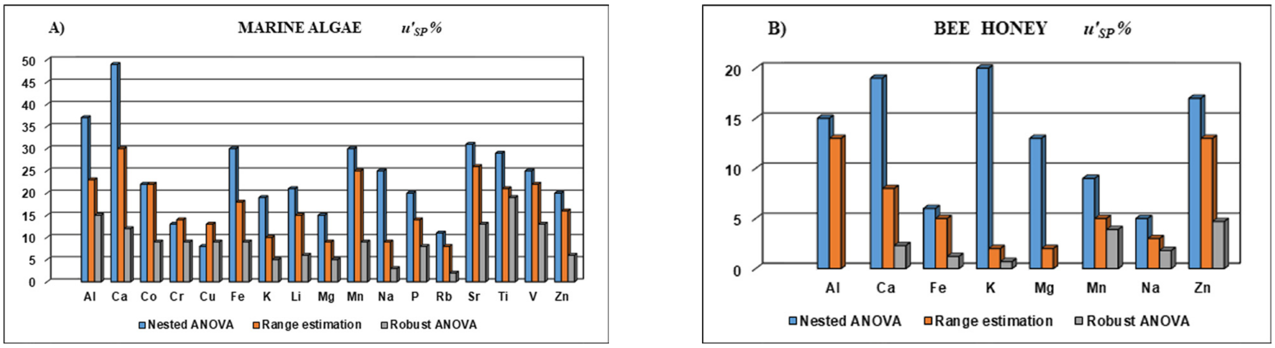

Fig. 2. Comparison of the relative standard measurement uncertainty due to sample preparation (u'_{SP} %) estimated by the three studied methods for all elements in marine algae (A) and bee honey (B).

Table 4. Measurement uncertainty due to sample preparation: absolute (u_{SP}) and relative (u'_{SP}) for marine algae (A) and bee honey (B) samples.

A) Marine algae						
Analyte	Nested ANOVA		Range estimation		Robust ANOVA	
	u_{SP} (mg kg^{-1})	u'_{SP} (%)	u_{SP} (mg kg^{-1})	u'_{SP} (%)	u_{SP} (mg kg^{-1})	u'_{SP} (%)
Al	440	37	280	23	85	15
Ca	11 000	49	6 900	30	1 300	12
Co	2	22	2	22	0.6	9
Cr	1	13	0.7	14	0.3	9
Cu	1	8	0.7	13	0.6	9
Fe	390	30	250	18	67	9
K	5 600	19	2 700	10	900	5
Li	0.3	21	0.2	15	0.05	6
Mg	1 400	15	790	9	300	5
Mn	88	30	62	25	24	9
Na	2 500	25	1 200	9	200	3
P	450	20	340	14	150	8
Rb	3	11	2	8	0.4	2
Sr	65	31	49	26	16	13
Ti	9	29	6	21	3	19
V	1	25	1	22	0.6	13
Zn	6	20	5	16	2	6

B) Bee honey						
Analyte	Nested ANOVA		Range estimation		Robust ANOVA	
	u'_{SP} (mg kg^{-1})	u'_{SP} (%)	u_{SP} (mg kg^{-1})	u'_{SP} (%)	u_{SP} (mg kg^{-1})	u'_{SP} (%)
Al	1.0	15	0.9	13	-	-
Ca	21	19	13	8	2.5	2
Fe	0.5	6	0.4	5	0.1	1
K	220	20	60	2	7.2	1
Mg	7.2	13	2	2	-	-
Mn	0.1	9	0.1	5	0.04	4
Na	4.2	5	5	3	3.3	2
Zn	0.7	17	0.4	13	0.1	5

trend, the MU from sample preparation for marine algae (Table 4 A) is higher than those for the bee honey (Table 4 B). There are two main factors explaining the higher uncertainty component for the algae samples. The first is their lower homogeneity as compared with the bee honey. The inhomogeneity of the subsamples, if not studied separately, remains as a hidden component of MU due to the sample preparation. On the other hand, marine algae samples undergo drying procedure after collection. However, these samples were stored and transported without particular moisture protection. Thus, the sample's status might be suggested as "air - dry". The lack of moisture control creates another sample preparation uncertainty source.

Influence of the analytical data quality

The number of data points is critically important for the quality of evaluation of MU due to sample preparation. In case of two replicates, mean, mid - range and median are practically identical and not enough precise estimates for the most probable value for the data set. Unavoidable imprecision of the estimation of μ reflects to the estimates of σ, decreasing its correctness as well. In such a case, the most reliable estimate of σ (directly related to MU) might be obtained by the range method. Its relatively higher certainty is provided by the statistical factor applied as a correction for insufficient number of data points.

However, not only the number of data points is important for the quality of u_{SP} evaluation. In principle, the lack of bias control e.g. by CRM analysis within every batch of real samples, decreases the reliability of the results. However, it practically does not affect the value of MU of sample preparation.

Completely different is the case with the precision. Repeatability of analytical measurement is quantified by the u_A values. When MU due to analytical determination is too high, the u_{SP} contribution is possible to become statistically insignificant. In such a case the u_{SP} values obtained are either unrealistically low or is impossible to be determined. If the u_A significantly exceeds u_{SP}, the expressions under the square root in equations 20, 25 and 33 becomes close to zero and occasionally negative. Thus, the MU related to sample preparation might be covered by the one from analytical determination.

Another problem related to analytical data quality is the existence of significant trend under repeatability conditions. Normally, it can be caused by insufficient stability of the final solution being analyzed, or instability in the response of the analytical instrument. In both cases the analytical signal becomes "time - dependent". Existence of trend compromises not only the MU components evaluation, but also the results of entire analytical procedure. Side effect from median calculation is the possibility for trend monitoring. When the median is predominantly in the middle of the sequence of replicates it is worth to check for the presence of trend in analytical signal.

Required for the correct application of nested ANOVA normal distribution is completely symmetrical. Small excerpts, even from normally distributed population can never be completely symmetrical. Significant data skewness (asymmetry) in the distribution of the sample makes classical ANOVA not reliable enough. That could be the reason for the highest and most probably overestimated values of u_{SP} obtained. However, it is important to note that increasing number of data points cannot guarantee the lack of skewness.

CONCLUSIONS

The measurement uncertainty associated with the sample preparation was evaluated for samples from marine algae and bee honey undergoing identical sample preparation procedure by application of three statistical approaches: hierarchical (nested) ANOVA, range method and median based robust ANOVA. In most of the cases, the results obtained by different algorithms, for both matrices and for all analytes are significantly different.

Applicability of hierarchical ANOVA for u_{SP} estimation is limited by symmetry of the data distribution. In case of excess skewness, the classical ANOVA variants should be replaced by robust "distribution free" approaches. If the number of replicates/digestions is low, what is the usual case in the analytical practice, the range method is preferable as having significantly higher robustness. Under such conditions the classical ANOVA provides highly overestimated MU values, while the robust (median) approach trends to underestimation. More precise evaluation of the applicability limits of the various statistical approaches for separation of MU from the different levels requires another experimental design and is a subject of further investigations.

The MU related to sample preparation and/or sampling is analyte and sample dependent value, despite the estimation method. The best strategy for its evaluation would be to perform a separate experiment with higher number of digestions and replicate measurements (at least 4 - 5). The evaluated u_{SP} could be assigned to all further analysis involving the same type of samples, analytes, and sample preparation procedure. If the comparison between results shows reasonable agreement between classical and robust ANOVA that would be also evidence for closeness to the normal distribution of all populations and samples.

The MU due to sample preparation might provide important information for further optimization of the sampling /sample preparation procedures as well as for the internal quality control in the analytical laboratory.

Acknowledgments

The authors express their sincere gratitude to the Operational Program "Science and Education for Intelligent Growth" and in particular, the project: "Modernization, digitalization and internationalization of education at the University of Chemical Technology and Metallurgy", which made possible the fruitful collaboration between Ljubljana University, Ljubljana, Slovenia and South - West University "Neofit Rilski", Blagoevgrad, Bulgaria.

REFERENCES

1. JCGM 200:2012, International vocabulary of metrology – Basic and general concepts and associated terms (VIM), 3rd edition.
2. JCGM 100:2008, Evaluation of measurement data - Guide to the expression of uncertainty in measurement (GUM), 1st edition, 2008.
3. S.L.R. Ellison, A. Williams (Eds), Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement, ISBN 978-0-948926-30-3, 3rd edition, 2012, Available at www.eurachem.org
4. M. Thompson (Ed.), AMC technical briefs: What is uncertainty from sampling and why is it important, AMCTB, 16A, Royal Society of Chemistry, 2004, Available at www.rsc.org/amc
5. M. Thompson, Uncertainty of sampling in chemical analysis, Accreditation and Quality Assurance, 3, 1998, 117-121.
6. M.H. Ramsey, M. Thompson, Uncertainty from sampling, in the context of fitness for purpose", Accreditation and Quality Assurance, 12, 2007, 503-513.
7. M.H. Ramsey, S.L.R. Ellison, P. Rostron (Eds.), Eurachem/EUROLAB/ CITAC/Nordtest/AMC Guide: Measurement uncertainty arising from sampling: a guide to methods and approaches, Eurachem (2019), ISBN (978-0-948926-35-8), 2nd edition.
8. B. Magnusson, M. Krysell, E. Sahlin, T. Näykki, Uncertainty From Sampling - A Nordtest Handbook for Sampling Planners on Sapling Quality Assurance And Uncertainty Estimation, 2020, Available at: https://www.nordtest.info/wp/
9. G. Ozay, F. Seyhan, A. Yilmaz, T.B. Whitaker, A.B. Slate, F. Giesbrecht, Sampling Hazelnuts for Aflatoxin: Uncertainty Associated with Sampling, Sample Preparation, and Analysis, Journal of AOAC International, 89, 4, 2006, 1004-1011.
10. T.B Whitaker, A.B. Slate, M. Jacobs, J.M. Hurley, J.G. Adams, F.G. Giesbrecht, Sampling Almonds for Aflatoxin, Part I: Estimation of Uncertainty Associated with Sampling, Sample Preparation, and Analysis, J. AOAC Internat., 89, 4, 2006, 1027-1034.
11. A. Ambrus, E.M. Solymosné, I. Korsós, Estimation of uncertainty of sample preparation for the analysis of pesticide residues, Journal of Environmental Science & Health Part B, 31, 3, 1996, 443-450.
12. W. Korol, G. Bielecka, J. Rubaj, S. Walczyński, Uncertainty from sample preparation in the laboratory on the example of various feeds, Accreditation and Quality Assurance, 20, 2015, 61-66.
13. J.A. Lyn, M.H. Ramsey, R.J. Fussell, R. Wood, Measurement uncertainty from physical sample preparation: estimation including systematic error, The Analyst, 128, 11, 2003, 1391-1398.
14. V. Morgado, C. Palma, R.J.N. Bettencourt da Silva, Monte Carlo bottom-up evaluation of the uncertainty of complex sample preparation: Elemental determination in sediments, Analytica Chimica Acta, 1175, 2021, 338732.
15. ISO/IEC 17025:2017, General requirements for the competence of testing and calibration laboratories, Status confirmed in December 2023.
16. D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens,

S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics. Part A, Data Handling in Science and Technology, 20A, Elsevier, Amsterdam, 1997.

17. J. Miller, J.C. Miller, R.D. Miller., Statistics and chemometrics for analytical chemistry, ISBN-13: 978-1292186719, 7th edition, Coronet Books Inc., 2018.

18. ISO 3085:2019, Iron ores – Experimental methods for checking the precision of sampling, sample preparation and measurement. (Status in December 2023 – published).

19. M. Thompson (Ed.), The Duplicate Method for the estimation of measurement uncertainty arising from sampling, AMCTB, 40, 2009, Available at www.rsc.org/amc