17th IMEKO TC 10 and EUROLAB Virtual Conference
"Global Trends in Testing, Diagnostics & Inspection for 2030"
October 20-22, 2020.

Development of Reference Material of Mercury in Fish: A comparison of different alternatives to homogeneity assessment

Soto, Leonardo[1]; Niño Fabian[1]; Garzón, Diego A [1]; Ahumada, Diego A.[1]

[1]Grupo de Investigación en Metrología Química y Bioanálisis, Subdirección de Metrología Química y Biomedicina. Instituto Nacional de Metrología de Colombia (INM), Bogotá D.C. Colombia. subdirquimica@inm.gov.co, (571) 2542222

***Abstract* – A candidate to certified reference material for mercury in fish was produced from *Calophysus macropterous*, a species of catfish that has been reported with high levels of this contaminant at the Amazonian region. The fish was filleted, freeze-dried and degreased, then sieved, homogenized and packaged; 106 bottles were produced with a net content of 15 g. In the study of homogeneity of the material, ICP-MS in standard mode was used, like a measurement method with microwave-assisted acid digestion. The homogeneity assessment with a random sampling stratified was carried out. The preparation and measurements of the samples were carried out using microwave-assisted digestion and ICP-MS. The current study involved the evaluation of different isotopes (^{199}Hg, ^{200}Hg, ^{201}Hg, ^{202}Hg) and internal standards to correct the analytical signals. Finally, to evaluate the uncertainty due to heterogeneity, different approaches were used. The results showed that there are differences between the classic ANOVA approach and meta-analysis methods, especially when $MS_{between} < MS_{within}$. Furthermore, the most important differences were found between the uncertainties using different combinations of isotopes / internal standards, because in some cases, up to three times higher uncertainties were found.**

Keywords – Homogeneity assessment, Mercury, Fish, ANOVA, Meta-analysis.

I. INTRODUCTION

The fish is one of the main export foods by Latin America; however, the mercury contamination and the dynamic of this element in the ecosystem and its impact on human health has caused difficulties in this market. Consequently, it is one of the most studied environmental contaminants, although in Latin American, the assessment of real status mercury content in the population, food and environmental compartments through monitoring programs by the agents of regulatory authorities to ensure compliance may not be totally realistic in some situations.

This situation may be due to the difficulty of the analysis of this element, low analytical capacity or few tools to ensure the quality of the measurements adequate for the levels of mercury concentration in regional fish. For example, the Amazonian fish has high levels of mercury and, usually, it goes outside the scope of reference methods and commercially certified reference materials. On the other hand, to ensure that mercury measurement results are comparable in space and time, it is necessary for the implementation of measurement traceability through the certified reference materials (CRMs).

The certification of reference materials can be carried out using different approaches; in general, the main sources of uncertainty of the certified value are associated with homogeneity, stability and characterization. Possibly, the most important parameters for consideration in the production of CRMs is that the material is homogenous, because if it is not, the certification process would be impossible to continue.

CRMs are usually prepared in batches for which the property values are determined by measurements on sample representative for the whole batch; therefore, the batch must be as homogeneous as possible. The homogeneity uncertainty ($u_{homogeneity}$), according to the definition of CRM, it must be determined to ensure the uniformity of the property measured in the matrix.

In this context, the objective of this work is to present the results obtained in the assessment of the homogeneity of new reference material obtained from *Calophysus macropterus.* Additionally, this study shows the comparison between different alternatives to estimate the uncertainty associated with the homogeneity of the material.

II. RELATED RESULTS IN THE LITERATURE

The maximum permitted levels of Hg in food are 1 µg

g^{-1} in the USA and 0.5 µg g^{-1} in the EU and China. [1-2] The demand of new CRMs for assessing the accuracy of measurement methods is increasing, and CRMs for metrological traceability of Hg results need to be provide by the National Metrology Institutes. In the last decade, several CRMs have been produced, such as mussels (SRM 1974b), oysters (SRM 1566b), tuna (CRMs 463 and 464), adult trout fillets (SRM 1947), river fish, among others. [3-6]

Other studies reported the preparation and certification of reference material for the total mercury and methyl mercury fraction mass in fish. The reference material was produced from dourada fish (*Brachyplatystoma Flavicans*), the uncertainty component due to homogeneity was 0.62% for total mercury. [7]. NIST had developed two standard reference materials to the monitoring of methyl mercury and total mercury in fish and marine mammals: SRM 1946 Lake Superior Fish Tissue and SRM 1947 Lake Michigan Fish Tissue [8].

The homogeneity assessment, usually it is carried out using the classic ANOVA approach to determine the variation of the measurand between bottles and within bottles. The variation between bottles ($\mu_{homogeneity}$) is calculated using Equation 1. For this analysis, it is assumed that data are homoscedastic, normal, random and independent into units or bottles. The $MS_{between}$ its the variation between units and MS_{within} within units and *n* the number of replicate measurements performed per unit.

$$S_{bb} = \sqrt{\frac{MS_{\text{between}} - MS_{\text{withim}}}{n}} \qquad (1)$$

In general, many studies have found that the $u_{homogeneity}$ is estimated correctly by one-way analysis of variance (ANOVA). However, the uncertainty contribution due to heterogeneity can be hidden by method repeatability and that produces an underestimation [9]. On the other hand, the estimation of uncertainty between bottles by one-way ANOVA it is not possible when the method used has a high repeatability variance, because the argument in equation 1 can become negative, which makes a calculation of variation between bottles impossible.

When the $MS_{within} > MS_{between}$, some solutions have been suggested, either by replacing the estimate for s_{bb} from Equation 1, for example, (i) change or improve the measurement method, (ii) use the method variation, (iii) increase the number of replicas per bottle or unit, (iv) use meta-analysis methods and (v) the use of Bayesian analysis.[9]

III. DESCRIPTION OF THE METHOD

A. Instrumentation

A quadrupole-based NexION 300D ICP-MS instrument with an autosampler (PerkinElmer, Pennsylvania, USA) was used to measure Hg in the homogeneity assessment. The plasma was generated using argon (99.990%), a power of 1300W, gas flow rates of plasma 15 L min^{-1}, auxiliary 0.82 L min^{-1}, and nebulizer 0.6 L min^{-1}. A cyclonic nebulizer, nickel sampler's and skimmer corner were used. The instrumental parameters such as torch position and gas output were optimized before the assay using a tuning solution (1µg kg^{-1} of In, U, Be and Ce) with a relative standard deviation objective of 2% or low. Also, the performance was checked after. Minimum acceptable counts per second are: Be 40 00, In 40 000 and U 30 000.

The acid digestion process of the samples was assistive by microwave, a MultiWave PRO microwave sample preparation system (Anton Paar, Graz, Austria) equipped with a multipurpose rotor with 16 pressure vessels (PTFE-TFM) supported by vessel jacket ceramic was used. The pressure and temperature of the digestion were controlled by a wireless sensor and an IR sensor.

An XP504 balance (Mettler Toledo, Ohio, USA) was used for the weighing of fish. This was calibrated and traceable to the international system through the prototype kilogram of Pt-Ir owned by PTB (Germany). The accuracy of the balance was checked before the assay finds errors less than 0.1%.

B. Reagents and reference materials

The reagents used in the acid digestion of the fish was HNO_3 (69% m/m) from Sigma-Aldrich (St. Louis, USA) and H_2O_2 (30%, SupraPure) from Merck Chemicals (Darmstadt, Germany). The HNO_3 used was purified by double sub-boiling distillation.

Stock solutions of In, U, Rh, and Ge were purchased from Sigma Aldrich; these elements are used as internal standards. Nominal concentrations were 1000 mg/L for Ge, In, and Rh and 10 mg/L for U. Gold stock solution with a nominal concentration of 1000 mg/L in 2% HCl was purchased from Merck.

The Hg standard reference material was purchased from the National Institute of Standards and Technology (NIST), Gaithersburg, (NIST- SRM 3133 - Mercury (Hg) Standard Solution).

C. Reference material preparation

The catfish *Calophysus macropterus* were obtained from local market brought from the Amazon region. The fish weight 40 kg. This fish was used for the preparation of the evaluated material.

The preparation of the material was carried out in several steps: i) Filleted, ii) Lyophilization, iii) Degreasing, iv) Drying, v) ground and sieved, vi) Homogenization and vii) Packing. In the first step, skin, bones and cartilages were removed, and the tissue obtained

was cut and cooled at -80ºC, for lyophilization, which was carried out at 0.07 bar for 72 hours. To carried out the material stability, the lyophilized material was immersed in petroleum benzine at 40ºC and constant agitation for degreasing the fish until the fat content is reduced at 4% (iii). Finally, the solid material was dried in the oven at 40 ºC for three hours (iv).

The fish (dry and degreased) was sieved in the meshes 2.00 mm, 1.7 mm, 1.4 mm, 1 mm, 0.7mm and 0.31 mm, obtaining the majority fraction for the sieve of 1 mm. After that, one fraction was put in a three-axis industrial mixer, and re-homogenized for 78 minutes in cycles of 5 minutes and rest between cycles of 2 minutes. This material was packed with the help of a funnel in 100 mL wide-mouth amber flask (washed with nitric acid 5% for 24 hours and after six washed with ultra-pure water); the final content of the bottles is 15 grams. Each labeled bottle was packed in a food bag and vacuum-sealed in an argon atmosphere and subjected to nuclear radiation to prevent degradation of the material by microorganisms.

D. Sample digestion and measurement

The fish samples fraction from the homogeneity assessment was taken of the prepared material. The digestion procedure was based on AOAC 2015.01, the method changes the relation between HNO_3 and H_2O_2. Around 300 mg of fish was weighed directly into a vessel and it was added with 3 mL HNO_3 and 2 mL H_2O_2. In the first step, the vessel was closed and rested for 12 hours in a pre-digestion stage to minimize mercury losses. The vials were placed in the rotor of the microwave for digestion; the program used was a ramp of 15 minutes to 190°C, maintaining this temperature for another 15 minutes. After this, a cooling stage to 25°C for 2 hours.

Acid extracts are diluted in a quantitative transfer with 25 mL ultra-pure water (conductivity ≤ 18 MΩ cm) and the addition of IS was made in this stage.

The measurement of Hg in the samples was made by ICP-MS, the select isotopes were ^{202}Hg, ^{201}Hg, ^{200}Hg and ^{199}Hg, with a dwell time of 20 ms in standard mode and peak hopping. On each measurement, 60 Sweeps were performed. To reduce the mercury memory effect reported for other authors, a wash time for the ICP-MS was carried out for 60 seconds with HNO_3 5%, and the sample was flushed for 30 seconds.

E. Homogeneity assessment

Two independent experiments were carried out to assess the heterogeneity of the produced batch.:

Experiment 1: The assessment of homogeneity within bottles ($u_{wbottle}$) was made using 7 aliquots from the same bottle. After digestion, the extracts were measured 14 times by ICP-MS. The estimate of $u_{wbottle}$ was made using different combinations of isotopes of mercury and internal standards.

Experiment 2: to assess the homogeneity between-bottles (u_{hom}), 12 bottles were selected from the whole batch following a random stratified sampling scheme. The results of this completely randomized design were analyzed using different approaches

All samples (seven sub-samples per bottle) were analyzed under repeatability conditions and in a randomized order, thus, to be able to distinguish measurement drift from drift originating from the filling sequence.[10]

IV. RESULTS AND DISCUSSIONS

A. Catfish material preparation

The catfish material was prepared according to the above description, 106 units were produced, which were distributed for different studies according to the ISO 35. The yield of the preparation procedure was 3.97%, evaluated as the total mass of the 106 bottles (106 x 15 g) according to the initial mass of catfish. The most significant losses occurred in the filleted and lyophilization followed by degreasing and sieving stages. The moisture content of the material was less than 4% with water activity of 0.14%,

Microbiological monitoring of the candidate to reference material and the unfavorable conditions for the growth of microorganisms, indicating good stability of the material.

Analytical calibration was performed using the NIST-SRM 3133 - Mercury (Hg) Standard Solution.

B. Trend analysis

The first step in the evaluation of data for a homogeneity study is a check of whether any trend in the data can be observed. This evaluation seeks to ensure that the drift errors caused in the performance of the measuring instrument (measurement system) are eliminated. A graphical presentation of the analytical results of the mercury relative responses is shown in Fig 1.

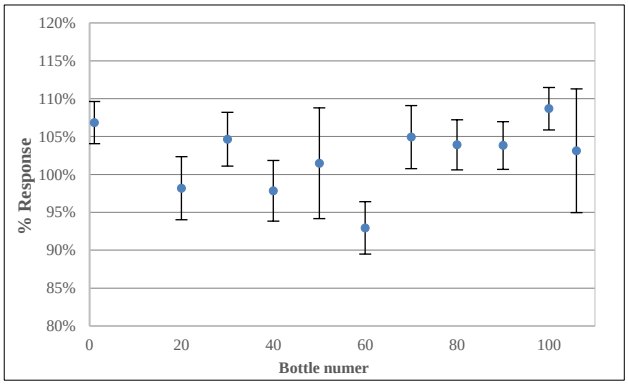

Fig. 1. Packaging trend analysis for material made of Hg in catfish (retaltive response $^{199}Hg/^{72}Ge$)

The linear regression model was applied to the homogeneity measurements. Table 1 summary of results trend studies: (i) measurement studies (instrument's drift) and (ii) material packaging (contamination or degradation). The results show a weak correlation (fitting a simple linear regression model) for response variation between the number of measurement (correlation coefficient, R = 0.44) and the number of bottles.

Table 1. Summary of linear regressions in the trend analysis

	Measurement	**Packaging**
Slope	-0.0008	-0.0018
R^2	0.0101	0.3223
Typical error	0.0308	0.0062
p-value	0.7557	0.1541

Table 1 shows, that p-values were greater than 0.05, which demonstrates the absence of trends. Finally, it is possible to conclude that the digestion step does not generate losses or trends that must take into account.

Before the uncertainty estimation, it is essential to make a distinction between homogeneity uncertainty (u_{hom}), within bottle uncertainty ($u_{wbottle}$), and measurement uncertainty (u_{meas}) for the amount of mercury in the fish. By u_{meas}, it addresses the fact that: (1) the instrument used has an inevitable measurement error and (2) the eventual model inadequacy. By $u_{wbottle}$ is understood as the variation of the material within each of the units, and with u_{hom}, the different sources of the true heterogeneity of the considered batch are meant. This u_{hom} is estimated using extra variance parameters in the model. These sources u_{hom} are assumed to be normally distributed with a zero mean and homogeneous variance.

C. Variation within bottles: $u_{wbottle}$ estimation

Based on the results obtained in this first experiment, the $u_{wbottle}$ was calculated using ANOVA, according to ISO Guide 35. Through this analysis, it is possible to establish u_{meas} and $u_{wbottle}$. Figure 2 shows that u_{meas} for both studies were less than 0.5%, using the four isotopes of mercury.

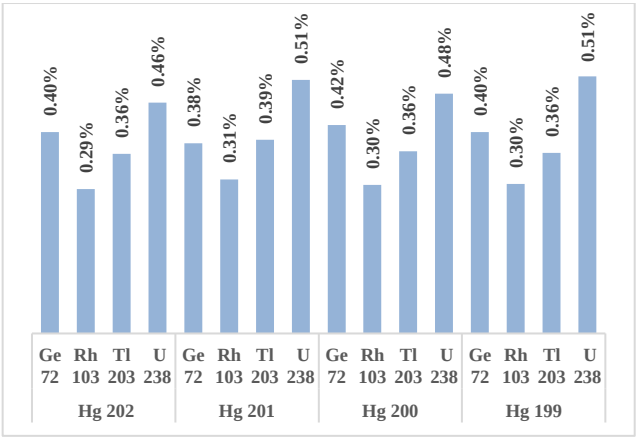

Fig. 2. Relative u_{meas} for material made of Hg in fish using different isotopes and internal standards

It can also be concluded from Fig 2 that when different internal standards were used, no significant changes were observed. In the worst case, the uncertainty for the ^{201}Hg / ^{238}U combination is 0.51%, which is adequate for the purpose.

The obtained results of uncertainty estimation associated to within bottle variation of the total mercury content are summarized in Figure 3. It was found that the candidate to reference material was homogeneous for a minimum sample weight of 300 mg and could therefore be subjected to the next steps in the production process.

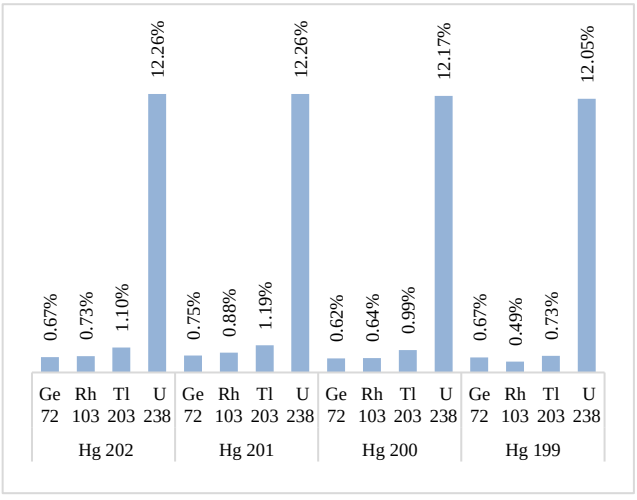

Fig. 3. Uncertainty within the bottles, for the material made of mercury in fish using different isotopes and internal standards.

In Fig 3 is possible to see that homogeneity uncertainty due to the variation within the bottle for the relations calculated using ^{238}U, are much greater than the other internal standards. This phenomenon may be due to some

kind of interference or problems in the digestion process that causes a significant variation of this IS between samples. On the other hand, in Fig 2, it can be seen that the best uncertainty inside bottles was obtained for the relation $^{199}Hg/^{103}Rh$ corroborating that the better isotope of mercury for this measurement is ^{199}Hg.

D. Variation Between bottles: u_{hom} estimation

For the study between bottles, a total of 12 bottles were used, which were selected using a stratified method. The results obtained are shown in Fig 4 using a classical ANOVA method to determine the variation between bottles and the u_{hom}. When comparing Figure 4 with Figure 3, it can be seen that the main component of heterogeneity of this material is the between bottles in homogeneity.

However, all the obtained uncertainties, $u_{wbottle}$ and u_{hom} values, are low and are consistent with the results obtained in other studies. Figure 4 shows the results that sample reference material can be considered homogenous both between bottle and within the bottle.

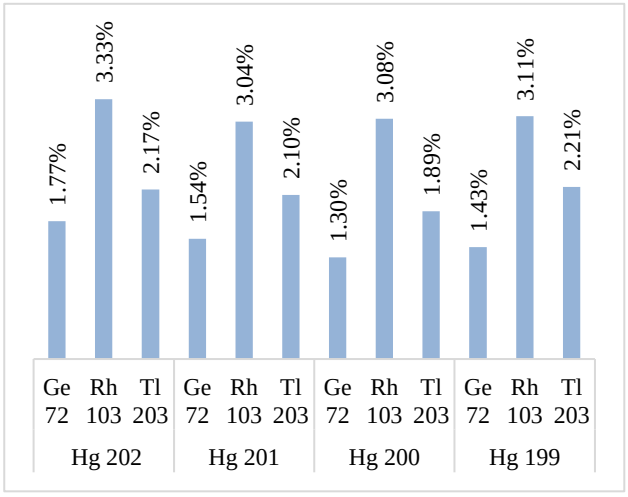

Fig. 4. Uncertainty between-bottles (u_{homo}) for the material made of mercury in fish using different isotopes and internal standards.

Figure 4 shows that the uncertainty between bottles is up to 7 times greater than measurement uncertainty (see Figure 2) and that, as for the uncertainty within the bottle, this indicates that measurement method chosen for homogeneity testing is adequate for the study, because the u_{meas} for the homogeneity testing procedure should be less than one third of the u_{hom}.

These differences, previously described between measurement and homogeneity uncertainties, allows the estimation using equation 1, without having to use any other approximation, as it has a positive difference.

Besides, it can be seen that the best relationship between mercury and internal standard is achieved with ^{200}Hg / ^{72}Ge, which is a subrogated standard; however, another relationship also presented a similar performance, ^{200}Hg / ^{203}Tl, which is an internal standard added in the quantitative transfer. This fact added to the fact that with the other mercury isotopes evaluated indicates that the homogeneity study was well executed.

In conclusion, a mercury material was developed in fish with a homogeneity uncertainty of 1.3%. In comparison with other CRMs marketed at the world level, such as ERM-BB422 of the European Commission, an uncertainty less than double that reported for it was obtained (0.86%) [12].

This comparison indicates that the material has a suitable homogeneity for its use, and that when thoroughly evaluated, it can be used as quality control and other functions of the reference materials, estimating with the classical methods.

E. Considerations in data analysis

To ensure correct analysis of results for the uncertainty estimation associated to the heterogeneity of the batch, the performance of a linear mixed-effects model (ANOVA) was only done afterward a carefully and systematic checking of the assumptions hereafter: first of all, that the experimental error is an independent random variable due to a randomly choice of measurements, secondly that it follows a normal distribution proved by the Shapiro-wilk test, thirdly that its mean is equal to zero reckoned by least squares method and finally, that the design has homogenous variances tested by Levene's test [10].

The Levene's test ($p<0.05$) indicates that the variances are significantly different between the measurements for each bottle (see Table 2). Heteroscedasticity is encountered in many chemical analysis experiments because initially, at measurements, there is a considerable amount of variance in the instrument, analyst and/or the preparation process of the samples, the magnitude of which is mainly determined by the stability of the measurement system and the stability of the samples during the measurement.

F. Meta-analysis (DerSimonian-Laird)

DerSimonian and Laird (DSL) method is a natural extension to the linear mixed-effects models in order to allow the use of random error [10]. Through the random effects and associated distribution, this type of model provides a flexible way to handle two essential characteristics of the repeated measures data: (i) the natural heterogeneity of the batch and (ii) the heteroscedasticity of the data. While in a classical ANOVA all bottles are the same, Der Simonian-Laird is capable of using the variance on each bottle to estimate the batch heterogeneity. Table 2 shows the uncertainty estimation using DSL and ANOVA for some isotopes selected.

Table 2. Summary methods for estimating uncertainty by homogeneity for the material made of Hg in fish.

Isotope/ IS	p-value Leven's test	u_{hom} ANOVA	u_{hom} DSL
$^{199}Hg/^{103}Rh$	0.036	3.11 %	3.02%
$^{200}Hg/^{103}Rh$	0.017	3.33 %	2.12%
$^{201}Hg/^{103}Rh$	0.015	3.04%	2.72%
$^{202}Hg/^{103}Rh$	0.239	3.08%	3.08%

Table 2 shows several aspects such as: (i) as the p-value decreases, there is a greater difference between the uncertainties estimated by DSL and ANOVA; (ii) the u_{homo} of the methods is practically the same when there is no heteroscedasticity. This behaviour can be explained, in principle, as due Der Simonian and Laird method incorporates a test of heterogeneity in the estimation of inter-study variance, so if the *p*-value of the test of heterogeneity is greater than 0.05, the result of the random-effects model is similar to ANOVA results.

V. CONCLUSIONS AND OUTLOOK

Based on the obtained results, it was found that the uncertainties attributed to ICP-MS measurements were compatible measurements to estimate the uncertainty associated with the homogeneity of the material. Further, these experiments also demonstrate how to conduct a homogeneity test in cases where the repetition of measurements is problematic, or the stability of the measurement systems is low.

In this study, the use of the DerSimonian and Laird method was found to be adequate when some of the ANOVA assumptions are not confirmed and thus homogeneity for the Hg analyte is demonstrated, specifically if the data are heteroscedastic. Finally, it is possible to conclude that a reference material for the content of mercury in fish was prepared correctly, stabilized, and bottled, since the material did not present trends in the measurement and bottling studies, and the homogeneity uncertainty obtained was less than 3%.

VI. ACKNOWLEDGMENTS

This research has been conducted with support from Physikalisch-Technische Bundesanstalt (PTB) through the project "Development and implementation of analytical tools for the assurance of quality and traceability in the measurements of toxic elements (Hg) in fish of the Amazon", financed by the Regional Fund for the Quality Infrastructure for biodiversity and climate protection in Latin America and the Caribbean. The authors are indebted to their numerous colleagues who took part in the preparation of the reference material and in the analytical work.

REFERENCES

[1] **Commission Regulation (EC)**: Commission Regulation (EC) No 466/2001 as amended by Regulation (EC) No 221/2002 setting maximum levels for certain contaminants in foodstuffs, *Official Journal of the European Union*, L37, 2001 pp. 4-6**.**

[2] **Food and Agriculture Organization of the United States**: Discussion paper of guideline levels for methylmercury in fish, *codex alimentararius comission*, 2005, p. 4.

[3] **Donais, M.; Saraswati, R;, Mackey, E. et al**.: Certification of three mussel tissue standard reference materials (SRM) for methylmercury and total mercury content, Fresenius *J Anal Chem*, Vol. 358, 1997, pp. 424–430

[4] **Richter, Silke; Schantz, M; Horvat, Milena; Logar, M; Akagi, H; Emons, Hendrik; Levenson, M; Wise, Stephen**: (2001). Certification of the methylmercury content in SRM 2977 Mussel Tissue (organic contaminants and trace elements) and SRM 1566b Oyster Tissue, *Fresenius' journal of analytical* chemistry, Vol. 369, 2001, pp. 364-9

[5] **Quevauviller, Philippe; Drabaek, Iver; Bianchi, Michele**: Certified reference materials (CRMs 463 and 464) for the quality control of methylmercury in tuna fish, *Trends Anal. Chem*, Vol. 15., No. 5, 1996, pp. 160-167.

[6] **Ulrich, Joao; Souza Sarkis, Jorge Eduardo**: Preparation and certification of a reference material for the total mercury and methylmercury mass fractions in fish, Accreditation and Quality Assurance, Vol. 18., 2013, pp. 511-516

[7] **Nedzarek, Arkadiusz; Czerniejewski, Przemysław; Tórz, Agnieszka**: Microelements and macroelements in the body of the invasive Harris mud crab (*Rhithropanopeus harrisii*, Maitland, 1874) from the central coast of the South Baltic Sea, *Environmental Monitoring and Assessment*, Vol 191., 2019, pp. 499

[8] **Davies, W; Davis, W; Christopher, S; Pugh, Rebecca; Donard, O; Krupp, Eva; Point, D.; Horvat, Milena; Gibicar, D; Kljaković-Gašpić, Zorana; Porter, Barbara; Schantz, Michele**: Certification of methylmercury content in two fresh-frozen reference materials: SRM 1947 Lake Michigan fish tissue and SRM 1974b organics in mussel tissue (*Mytilus edulis*), *Analytical and bioanalytical chemistry*, Vol. 387., 2007, pp. 35-41.

[9] **Linsinger, T**: Evaluation of CRM homogeneity in cases of insuf!cient method repeatability: Comparison of Bayesian analysis with substitutes for ANOVA based estimates, *Analytica Chimica Acta: X*, Vol 5, July 2020, DOI: 10.1016/j.acax.2020.100049

[10] **Veroniki, A; Jackson, D; Viechtbauer, W; Bender, R; Bowden, J; Knapp, G; Kuss, O; Higgins, J**: Methods to estimate the between-study variance and its uncertainty in meta-analysis, *Research Synthesis Methods*, Vol 7, 2016,pp. 55-79. DOI: 10.1002/jrsm.1164

[11] **Snell, J; Emteborg, H.; Schimmel, H.**: Certification Report. The certification of the mass fractions of elements in fsh muscle. Certified of reference material ERM –BB422. *European Comission*, 2012, pp. 11-13.