Laboratory PT #2 October 2023

2nd Round of Mercury Laboratory Proficiency Testing

Final Report

1. Objective

The main purpose of this proficiency testing (PT) programme was to evaluate the performance of mercury analyses conducted by the laboratories. It was expected to provide the individual proficiency levels of participating laboratories and the collective mercury monitoring capacity in the region.

2. Proficiency testing provider

This PT was organized by National Institute for Minamata Disease (NIMD) in collaboration with United Nations Environmental Programme, Regional Office for Asia and the Pacific (UNEP-ROAP). Asian Institute of Technology Regional Resource Centre for Asia and the Pacific (AIT RRC.AP) distributed the test item and collected the analytical results. IDEA Consultants, Inc. prepared the test item for this PT.

3. Implementation period

Call for participation:	September – December 2022
Test item distribution:	February – March 2023
Duration of test (analysis):	Until 8 June 2023

4. Participation fee

Free.

5. Test item (sample) and parameter

One (1) dried fish sample was used for analysing total mercury concentration, methylmercury concentration or both (participating laboratories shoes the preferred option).

5.1. Test item preparation

The test item was made from the muscle tissue of a single yellowfin tuna. The tissue was collected, freeze-dried, finely powdered, and sieved (106 μm) to ensure sufficient homogeneity. The test item was also sterilized using gamma-ray.

Approximately 5 g of each test item was packaged in a brown glass bottle. The bottles of test items were sealed in aluminium-lined laminate packs for distribution to the participants.

5.2. Homogeneity testing

The following homogeneity testing of the test item was conducted to ensure that there were no significant differences in the mercury concentrations in the test items between bottles that could affect the result of the PT. Since the test item was collected from a single species and the ratio of total mercury and methylmercury was assumed to be stable between test items, homogeneity testing was conducted by analysing total mercury.

After the preparation of the test item (packed in bottles), ten bottles were selected, and the total mercury analysis (acid digestion - aeration CVAAS measurement) was performed twice for each test item in a bottle.

The homogeneity of the test item was then analysed from the results of the total mercury concentrations. Since the analytical results include the uncertainty due to the (chemical) analytical procedure, homogeneity was judged by the following criterion:

Criterion: $S_s \leqq \sqrt{F_1 \times (0.3 \times \sigma_{ep})^2 + F_2 \times {S_w}^2}$

S_s: relative standard deviation of homogeneity testing

σ_{ep}: (expected) relative standard deviation of the reported results from participants

$w_i^2 = \Sigma (x_{gm}^2 - \bar{x}_g^2) / (m-1)$ x_{gm}: result of m times analysis of the bottle

$S_w^2 = \Sigma w_i^2 / g$ $\bar{x}_g$: average of the result of each bottle

F_1 and F_2 are values which are calculated from the probability distribution. In this homogeneity testing (10 bottles testing), F_1 and F_2 were applied following numbers:

$F_1 = 1.88$ $F_2 = 1.01$

(Even though these values are referred from the Annex B of ISO13528:2015, they are introduced from the random variables of χ^2 distribution and F distribution.)

Analysis results of this homogeneity testing are as follows:

$S_s = 0.0090$

$S_w^2 = 0.000934$

Also, relative standard deviation of the results from the participants (used for evaluation) was as follows:

$\sigma_{ep} = 0.0691$

This standard deviation should be used the value which was used for the performance evaluation for participants. As described in 8.2, the performance of participants was evaluated from the median and normalized interquartile range (NIQR) of the results, thus relative NIQR was used for the confirmation of the criterion.

Therefore, above criterion was judged as follows:

$$\sqrt{F_1 \times (0.3 \times \sigma_{ep})^2 + F_2 \times {S_w}^2}$$
$$= \sqrt{1.88 \times (0.3 \times 0.0691)^2 + 1.01 \times 0.000934}$$
$$= 0.0418 > 0.0090\ (S_s)$$

It was confirmed that the test item was sufficiently homogeneous to evaluate the performance of participants' results.

5.3. Stability testing

To ensure that the concentration of the target parameter (total mercury) was maintained without significant changes during the PT, a following stability testing was conducted after the duration of the analysis. Since the test item was collected from a single species and the ratio of total mercury and methylmercury was assumed to be stable between test items, stability testing was conducted by total mercury analysis.

Ten test items were selected from the stored (not distributed to participants), and total mercury analysis (acid digestion - aeration CVAAS measurement) was performed twice for each test item in a bottle.

The stability of the test item was then analysed by comparing the results before and after the distribution of the test item. The stability of the test item was judged by the following criterion:

Criterion: $|\bar{x} - \bar{y}| \leqq 0.3 \times \sigma_{pt} + 2 \times \sqrt{{u_{(x)}}^2 + {u_{(y)}}^2}$

$\bar{x}$: average of the item before distribution

$\bar{y}$: average of the item after proficiency testing

$u_{(x)}$: uncertainty of the item before distribution

$u_{(y)}$: uncertainty of the item after proficiency testing

σ_{pt}: standard deviation for the proficiency evaluation. In this program, NIQR was applied to evaluation of performance of the participant.

Analysis results of test items before and after the PT are as follows:

$\bar{x} = 4.126$ $\quad u_{(x)} = 0.037$

$\bar{y} = 4.139$ $\quad u_{(y)} = 0.047$

Standard deviation of the result of all participants was as follows:

$\sigma_{pt} = 0.279$

This standard deviation should be used the value which was used for the performance evaluation for participants. As described in 8.2, the performance of participants was evaluated from the

median and normalized interquartile range (NIQR) of the results, thus NIQR was used for the confirmation of the criterion.

Therefore, above criterion was judged as follows:

$$0.3 \times \sigma_{pt} + 2 \times \sqrt{u_{(x)}{}^2 + u_{(y)}{}^2}$$
$$= 0.3 \times 0.279 + 2 \times \sqrt{0.037^2 + 0.047^2}$$
$$= 0.203 > 0.013\ (|\bar{x} - \bar{y}|)$$

It was confirmed that the concentration of total mercury in test item was not changed during the PT.

6. Target parameter

The target parameter of the PT was total mercury and/or methylmercury. Participants could perform analysis and report either or both of total mercury/methylmercury. Participants conducted three total mercury analyses and reported all results. Participants also conducted analysis of moisture in the test item. The result of moisture was used for the analysis of the reported data, however, it was not the target of the PT, and the result of total mercury was not calculated by moisture.

The moisture analysis procedure was instructed to the participants as follows:

1. Take a test item of 100 mg or more and weigh it precisely.
2. Dry the taken test item (100 °C, 2 hours).
3. Weigh the dried test item again and calculate the moisture of the sample from the reduced mass.

It has also been instructed that the sample used for moisture analysis should not be used for total mercury analysis.

7. Participating institutions

This PT was intended for public or university laboratories that perform mercury analysis. It was requested to perform the analysis with a lower detection limit than 0.1 mg/kg on 0.5 g test item.

55 institutions registered in the PT and 48 institutions (total mercury) and 17 institutions (methylmercury) respectively reported the analysis results.

The number of participants for each parameter (total mercury/methylmercury) and status are shown in Table 1.

Table 1 Number of participating laboratories for parameter and step

	Number of laboratories	
Category	Total mercury	Methylmercury
Registered	55	19
Sample received	52	18
Result delivered	48	17

Laboratories are participated from all of UN Regions. Number of laboratories (who reported the results) participated from each region is shown in Table 2.

Table 2 Number of participating laboratories per region

	Number of laboratories	
UN Region	Total mercury	Methylmercury
Africa	6	2
Asia and the Pacific	26	9
Eastern Europe	1	1
Latin America and the Caribbean	7	2
Western Europe and Other Group	8	3

8. Total Mercury Analysis Result

8.1. Basic statistic data of the PT result (total mercury)

The basic statistics of the result of PT are shown in Table 3.

Table 3 Summary of the results of the PT (total mercury)

Statistic data of the results (unit: mg/kg)	
Average:	3.875
Median:	4.046
Standard deviation:	0.791
Minimum	0.003
Maximum	4.707
25 percentiles	3.796
75 percentiles	4.173
Interquartile range (IQR)	0.377
Normalized IQR (NIQR)	0.279
Parameter related to distribution	
Skewness of distribution	-3.547
Kurtosis of distribution	14.88

The distribution of the results from the participants is shown in Fig. 1.

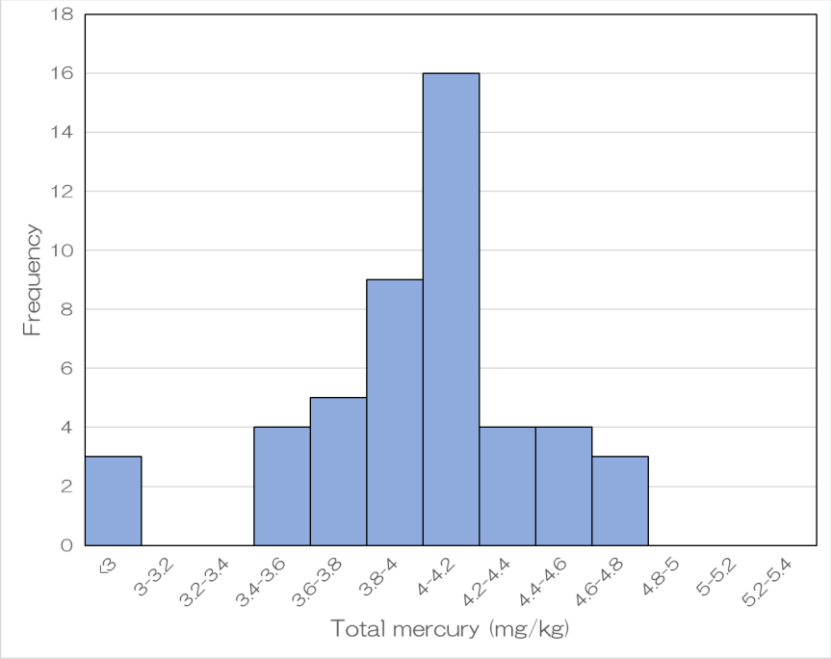


Fig. 1 Histogram of the report data (total mercury)

Some registered participants were unable to receive the test item due to various issues such as customs clearance, etc. There were also participants who were unable to report the analysis results due to problems with the measuring instrument.

These statistical data were calculated from the average of each participant. There was one participant who did not report 3 results (2 results were reported), but all data were used for statistical analysis.

The distribution had a large kurtosis, indicating that a large number of reported values were concentrated around the median, while some other values deviated from it. This means the interquartile range was relatively small compared to the standard deviation.

There were reported values that were much lower than the median. Thus, the skewness values were negative. However, these data did not have a major impact on the performance evaluation because the evaluation was estimated from the median and normalized interquartile range (NIQR). Thus, performance was evaluated based on the data obtained from all reported values without processing outliers.

8.2. Performance evaluation for participants

Median data of all laboratories was applied as agreement value. Performance of the results was evaluated by the robust z score, which was calculated from the median and normalized interquartile range (NIQR).

z score of each participant was calculated from the following equation.

$$z = [(\text{average of reported result}) - (\text{median of all participants})] \,/\, \text{NIQR}$$

Performance of the result is classified by z score as follows:

|z| $\leqq$2: Performance is satisfactory (satisfactory)

2 < |z| < 3: Performance is questionable (caution)

|z| $\geqq$ 3: Performance is unsatisfactory (action)

The results and performances of laboratories are shown in Fig. 2.

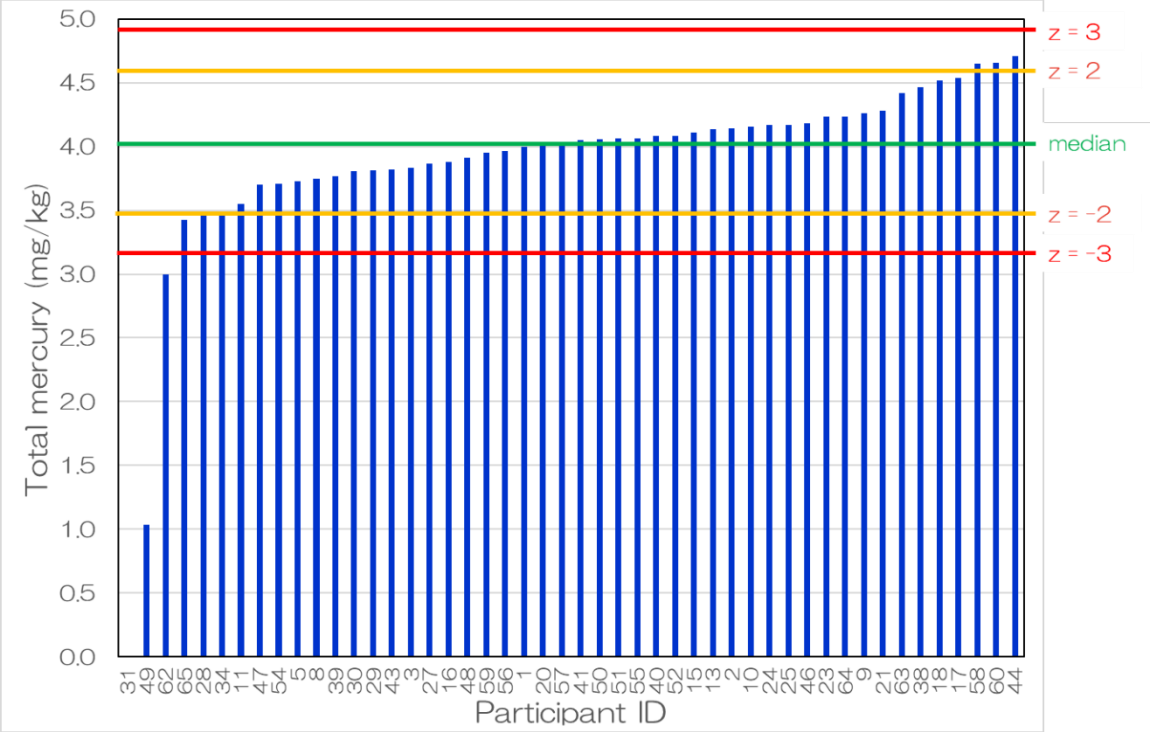


Fig. 2 Result of each participant (total mercury)

The numbers of laboratories disaggregated for each z score range are shown in Table 4.

Table 4 Number of the laboratories in the range of z score (total mercury)

z score	z ≦ -3	-3 <z< -2	-2 ≦z≦ 2	2 < z < 3	z ≧3
n	3	2	40	3	0

As described in 8.1, IQR of reported results was relatively small. Therefore, satisfactory range of the result (mercury concentration) was relatively close and results around 21 % difference from the median was the classified range of unsatisfactory result (absolute value of z-score over 3). Several results had z-scores slightly above 2 in absolute value, but this does not immediately indicate poor analytical procedure or situation.

8.3. Regions of participating laboratories and the performances

Number of laboratories per region where the laboratories are located and the performance of the PT are shown in Fig. 3.

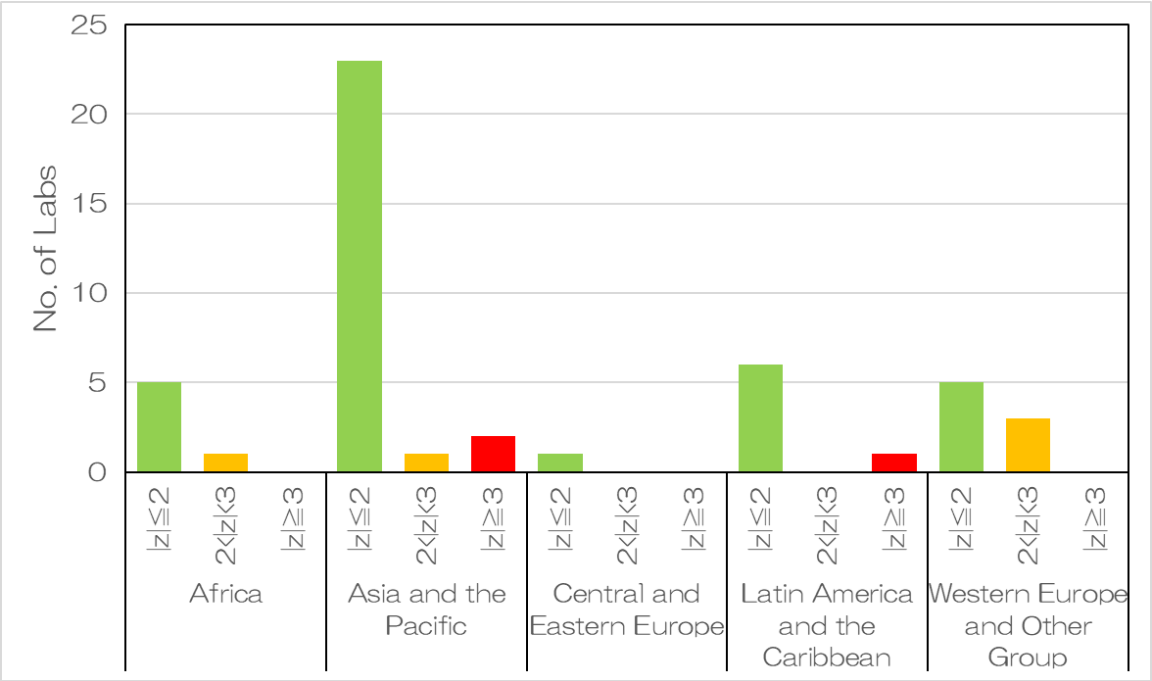


Fig. 3 Number of laboratories per regions and performances (total mercury)

Regional difference was not examined due to the limited number of laboratories participating from regions other than Asia and the Pacific being insufficient for comparison.

8.4. Types of laboratories and the performances

Number of laboratories per type (academic, government, or non-government) and the performance of the PT are shown in Fig. 4.

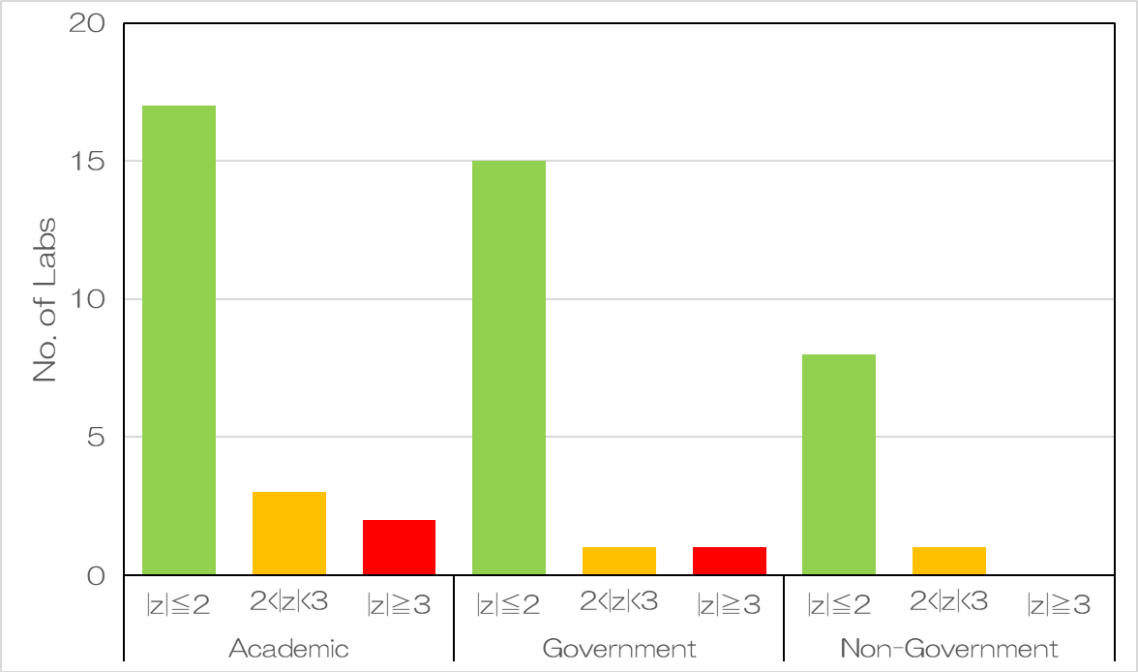


Fig. 4 Number of laboratories per types and performance (total mercury)

Statistical differences among the types of laboratories were not detected from the reported results (Kruskal-Wallis one-way ANOVA on ranks, p=0.77).

8.5. Analysis methods and results

The method of analysis was not specified for the PT and the participants performed analysis by the method that they usually used in their routine analysis, or they were planning to use in the future.

The participants performed analysis of total mercury by the following methods:

- Thermal Decomposition Cold Vapour Atomic Absorption Spectrometry (TDAAS)
- Acid digestion, aeration Cold Vapour Atomic Absorption Spectrometry (CVAAS)
- Acid digestion, Cold Vapour Atomic Fluorescence Spectrometry (CVAFS)
- Acid digestion, Inductively Coupled Plasma Mass Spectrometry (ICP-MS)

The distribution of the results from participants by analysis method is shown in Fig. 5.

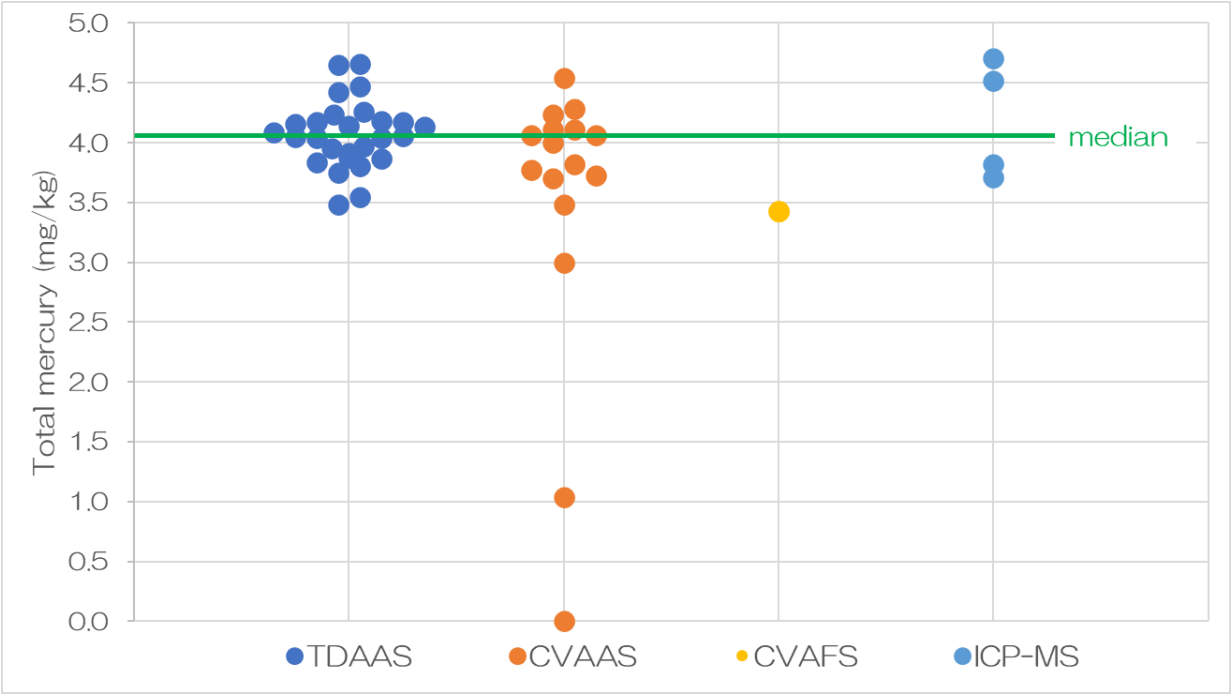


Fig. 5 Distribution of the report data by analysis method (total mercury)

No significant difference was found among the analysis methods employed (Kruskal-Wallis one-way ANOVA on ranks, p=0.14). However, several data measured by CVAAS, which were significantly different from the median. Since it is possible that errors in the data calculation process caused these outliters, which might not be related to the analysis method. On the other hand, TDAAS produced results with no outliers despite having the highest number of results. This can be attributed to the specific instrument designed for mercury analysis, which requires no chemical preparation and straightforward analysis procedure. Therefore, only a few parameters need to be considered to minimise error.

9. Methylmercury Analysis Result

9.1. Basic statistic data of the PT result (Methylmercury)

The basic statistics of the methylmercury result of PT are shown in Table 5.

Table 5 Summary of the results of the PT (methylmercury)

Statistic data of the results (unit: mg/kg)	
Average	3.394
Median	3.557
Standard deviation	0.941
Minimum	0.027
Maximum	4.223
25 percentiles	3.380
75 percentiles	3.829
Interquartile range (IQR)	0.449
Normalized IQR (NIQR)	0.333
Parameter related to distribution	
Skewness of distribution	-3.145
Kurtosis of distribution	11.37

The distribution of the results from the participants is shown in Fig. 6.

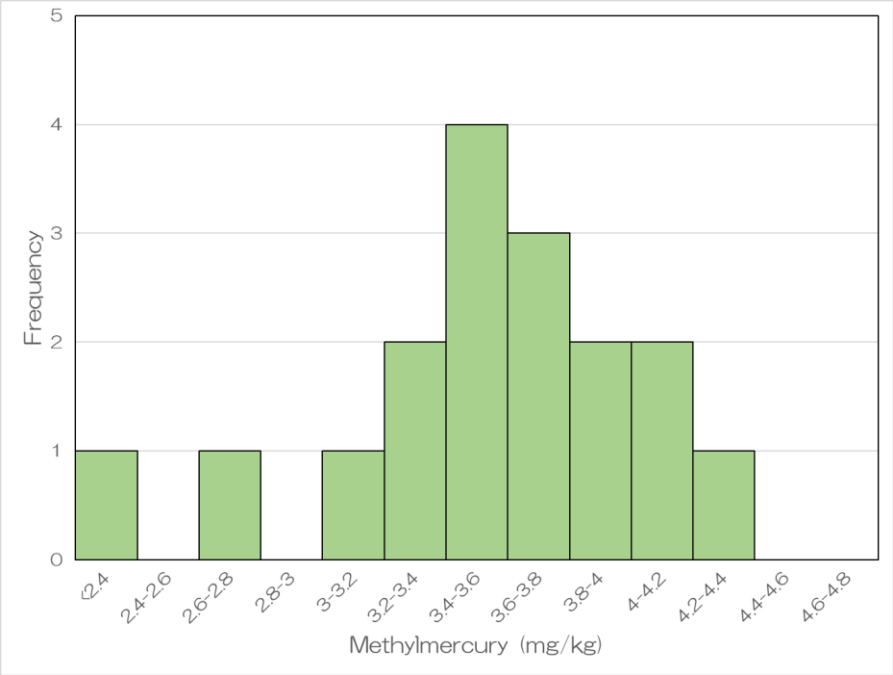


Fig. 6 Histogram of the report data (methylmercury)

Some registered participants did not receive the test item due to customs clearance issue, while one did not report the analysis results.

These statistical data were calculated from the average of each participant. There was a participant who did not report 3 results (2 results were reported); however, all data were used for the statistical analysis.

It was also shown in a large kurtosis of the distribution that relatively many reported values were concentrated in a narrow range around the median, while some reported values deviated from the median. Thus, the results are in a relatively small interquartile range than the standard deviation.

There were reported values that were much lower than the median. Thus, the skewness values were negative. However, these data did not have a major impact on the performance evaluation because the evaluation was estimated from the median and normalized interquartile range (NIQR). Thus, performance was evaluated based on the data obtained from all reported values without processing outliers.

9.2. Performance evaluation for participants

Median data of all laboratories was applied as agreement value. Performance of the results was evaluated by the robust z score, which was calculated from the median and normalized interquartile range (NIQR).

z score of each participant was calculated from the following equation.

$$z = [(\text{average of reported result}) - (\text{median of all participants})] \,/\, \text{NIQR}$$

Performance of the result is classified by z score as follows:

|z| $\leqq$2: Performance is satisfactory (satisfactory)

2 < |z| < 3: Performance is questionable (caution)

|z| $\geqq$ 3: Performance is unsatisfactory (action)

The results and performances of laboratories are shown in Fig. 7.

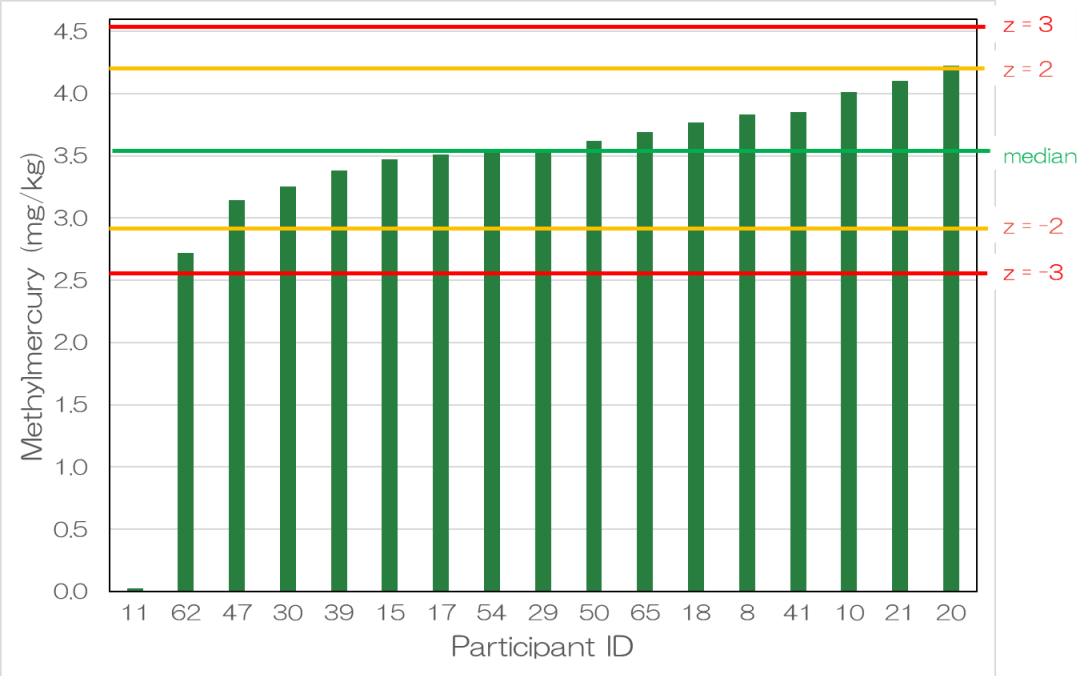


Fig. 7 Result of each participant (methylmercury)

The numbers of laboratories disaggregated for each z score range are shown in Table 6.

Table 6 Number of the laboratories in the range of z score (methylmercury)

z score	z ≦ -3	-3 <z< -2	-2 ≦z≦ 2	2 < z < 3	z ≧3
n	1	1	14	1	0

As described in 9.1, IQR of reported results were relatively small. Therefore, satisfactory range of the result (mercury concentration) was relatively close and results around 28% difference from the median was the classified range of unsatisfactory of the result. (absolute value of z-score over 3). There was a result that had z-scores slightly above 2 in absolute value, but this does not immediately indicate poor analytical procedure or situation.

9.3. Regions of participating laboratories and the performances

Number of laboratories per region where the laboratories are located and the performance of the PT are shown in Fig. 8.

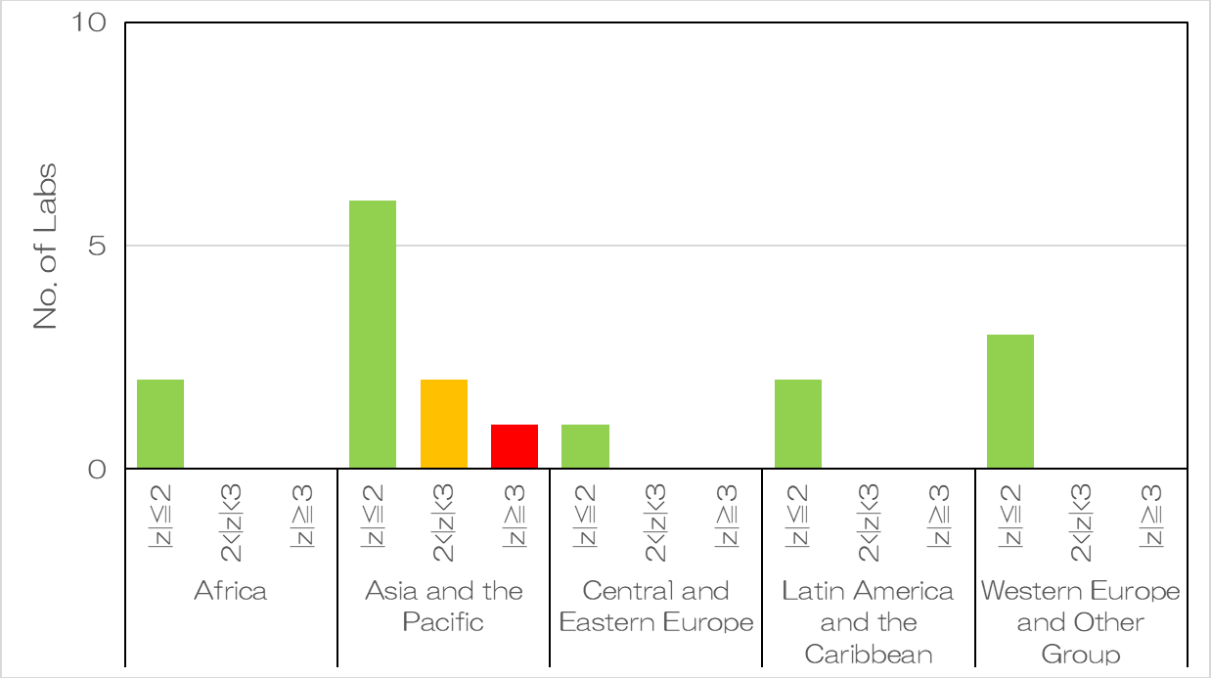


Fig. 8 Number of laboratories per regions and performances

Regional difference was not examined due to the limited numbers of laboratories participating from each region being insufficient for comparison.

9.4. Types of laboratories and the performances

Number of laboratories per type (academic, government, or non-government) and the performance of the PT are shown in Fig. 9.

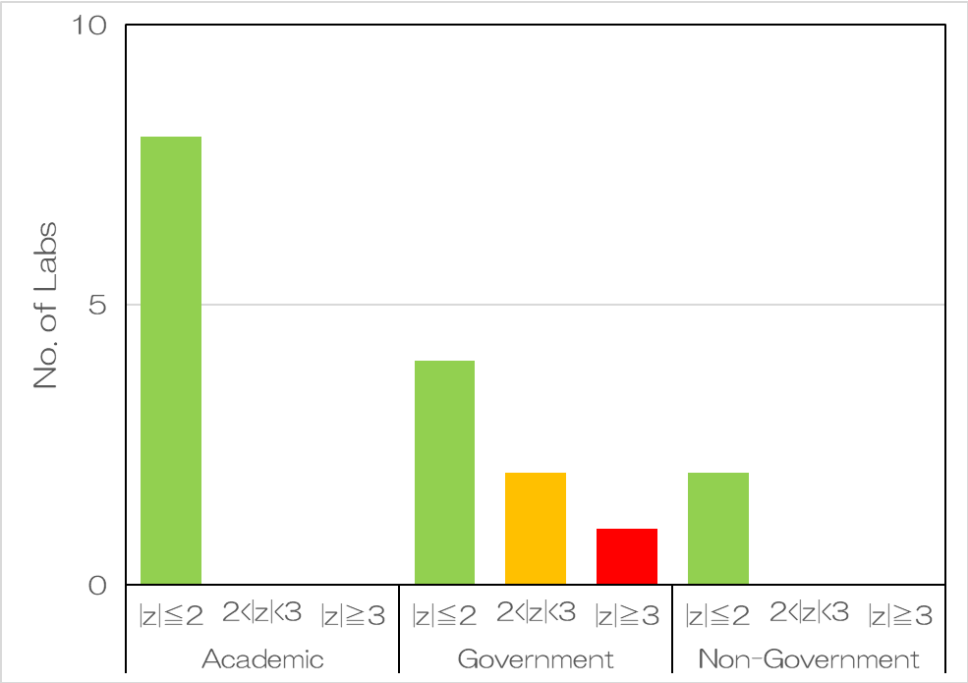


Fig. 9 Number of laboratories per types and performance (methylmercury)

Statistical differences among the types of laboratories were not detected from the reported results (Kruskal-Wallis one-way ANOVA on ranks, p=0.62).

9.5. Analysis methods and results

The method of analysis was not specified for the PT and the participants performed analysis by the method that they usually used in their routine analysis, or they were planning to use in the future.

The participants performed analysis of methylmercury by some type of measuring instruments after the solvent extraction of the target analyte. The measuring instruments used by participants are as follows:

- Gas Chromatograph and Cold Vapour Atomic Fluorescent Spectrometer (GC/CVAFS)
- Gas Chromatograph and Electron Capture Detector (GC/ECD)
- Liquid Chromatograph and Inductively Conducted Plasma Mass Spectrometer (LC/ICP-MS)
- Thermal Decomposition Cold Vapour Atomic Absorption Spectrometer (TDAAS)
- Cold Vapour Atomic Absorption Spectrometer (CVAAS)*
- Inductively Coupled Plasma Mass Spectrometry (ICP-MS)

*: About the report of CVAAS (one report), the method of extraction was not reported and it was unknown.

The distribution of the results from participants by analysis method is shown in Fig. 10.

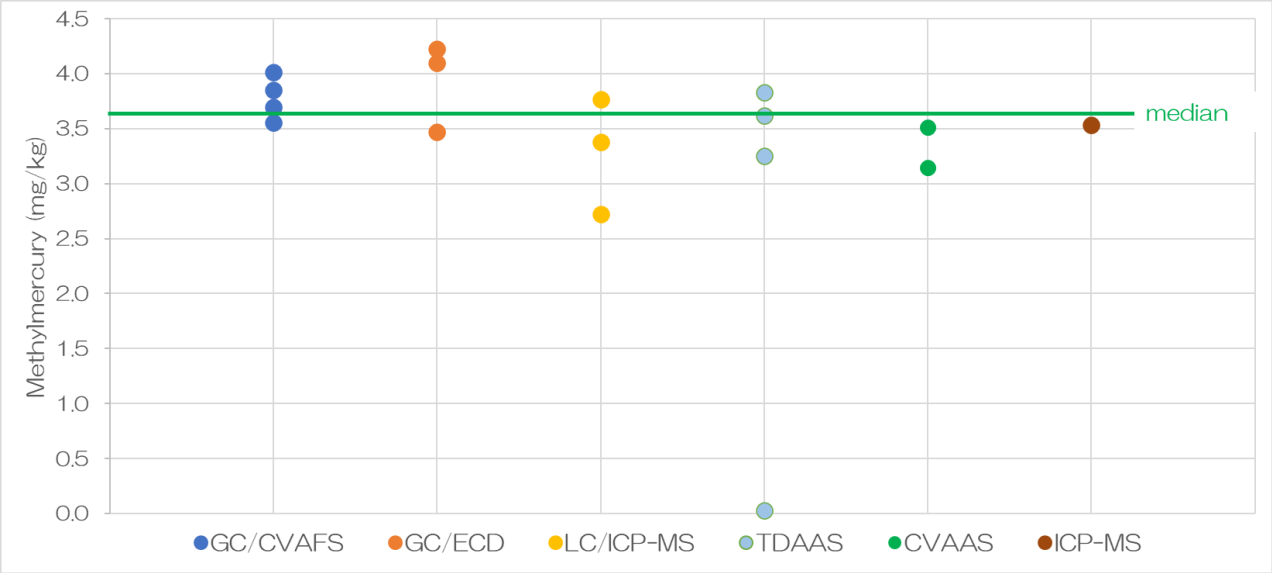


Fig. 10 Distribution of the report data by analysis method (methylmercury)

Noticeable differences were not detected among the analysis methods used. (Kruskal-Wallis one-way ANOVA on ranks, p=0.28).

10. Percentage of methylmercury

The summary of the percentage of methylmercury concentration against total mercury obtained by the same participant is shown in Table 7.

Table 7 Summary of the percentage of methylmercury

	Methylmercury / Total mercury (%)
Average	86.8
Median	90.7
Standard deviation	23.6
Minimum	0.8
Maximum	107.8

The distribution of the percentage of methylmercury concentration is shown in Fig. 11.

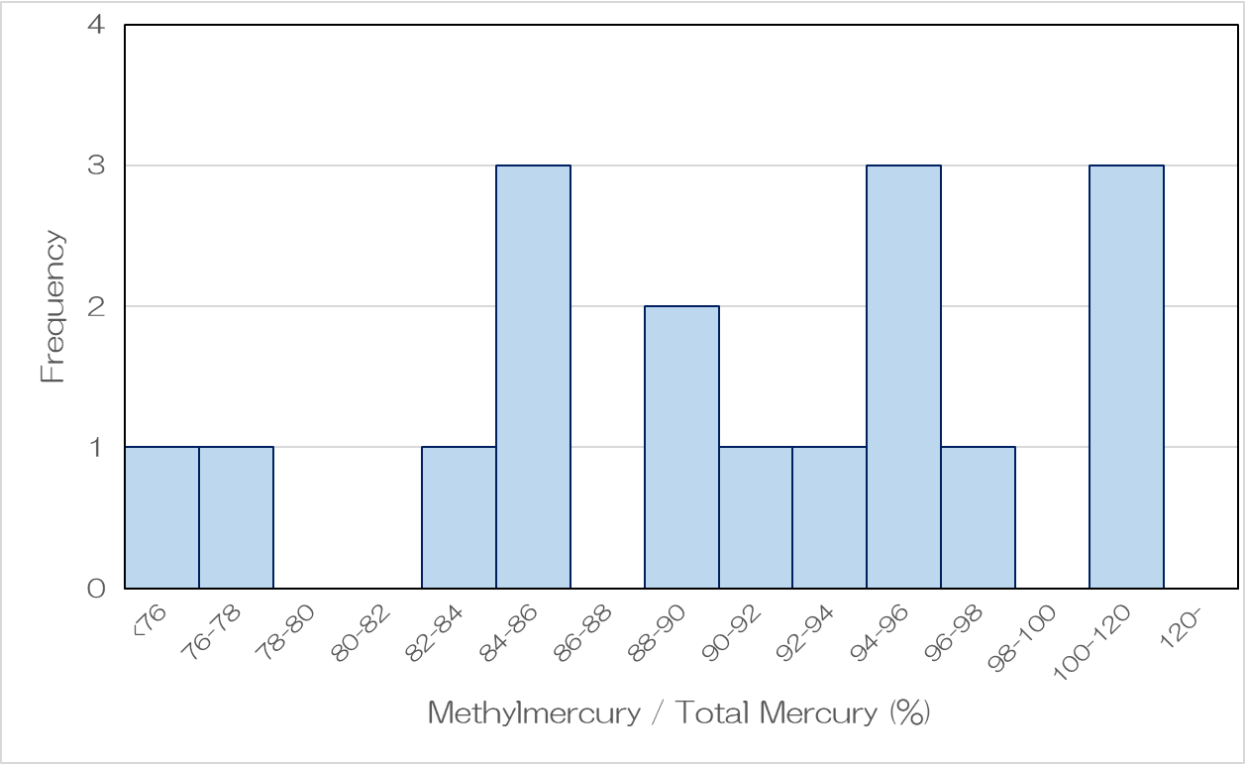
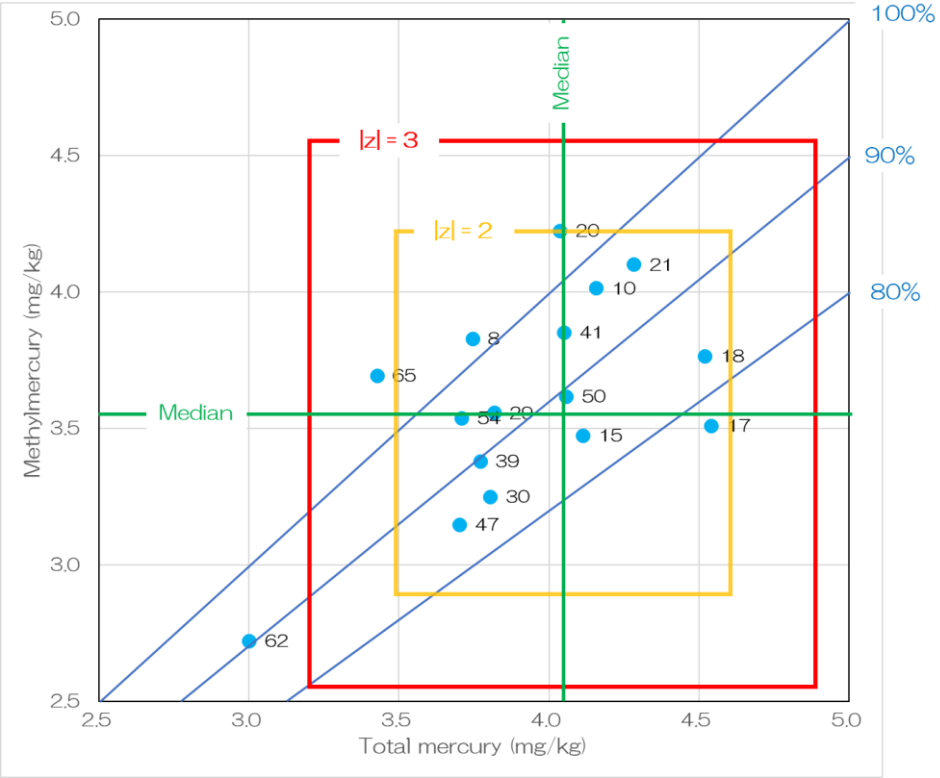


Fig. 11 Histogram of the percentage of methylmercury against total mercury

The plots of the relationship between total mercury and methylmercury concentration are shown in Fig. 12.



*: One outlier data (ID# 11) is not shown in this plot.

Fig. 12 Scatter plot between total mercury and methylmercury

A significant positive correlation between total mercury and methylmercury are detected (r=0.440, one-sided testing, p=0.039). It should be noted that the methodology and measuring instrument used for both analysis are different, so the reason for this correlation is unknown. It is possible that deviations in the commonly used instruments, such as balances may have contributed to this relationship. However, it is unlikely that these instruments would deviate to the same degree as the differences in the participants' results under typical laboratory conditions.

11. Moisture

The basic statistics of the reported moisture of test item is shown in Table 8. There were participants who reported difference moistures for each analytical parameter (total mercury/methylmercury). Therefore, statistics and analysis of moisture were conducted based on each result for the analytical parameter individually.

The basic statistics of moistures are shown in Table 8.

Table 8 Summary of the result of moisture

	Moisture (%)	
	Total mercury	Methylmercury
Average	4.90	4.63
Median	4.45	4.14
Standard deviation	3.76	4.13
Minimum	0.003	0.01
Maximum	19.3	19.0

The distribution of the results of moistures are shown in Fig. 13.

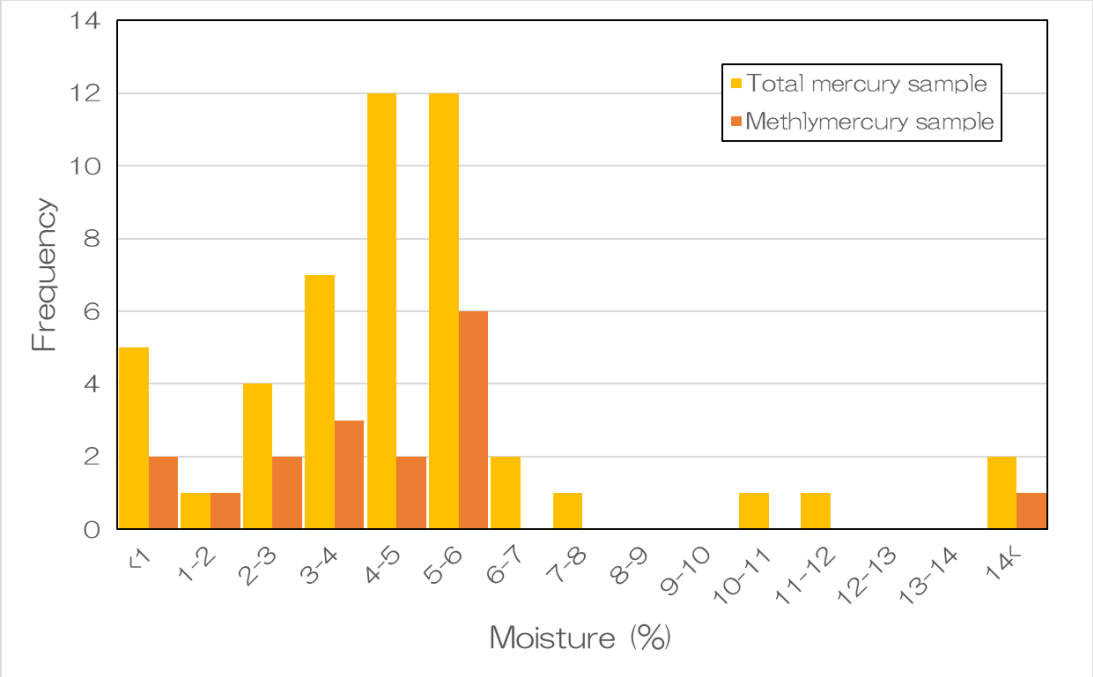


Fig. 13 Histogram of the result of moisture

The plots of the relations of moisture and total mercury concentration are shown in Fig. 14.

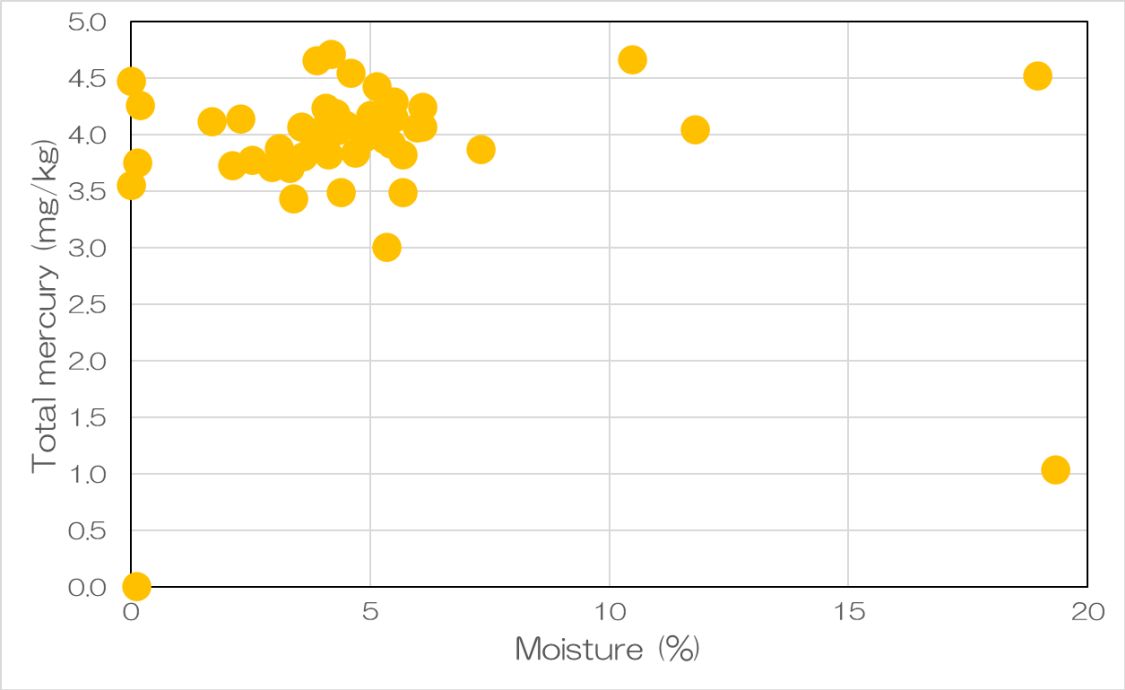


Fig. 14 Total mercury and moisture

The plots of the relations of moisture and methylmercury concentration are shown in Fig. 15.

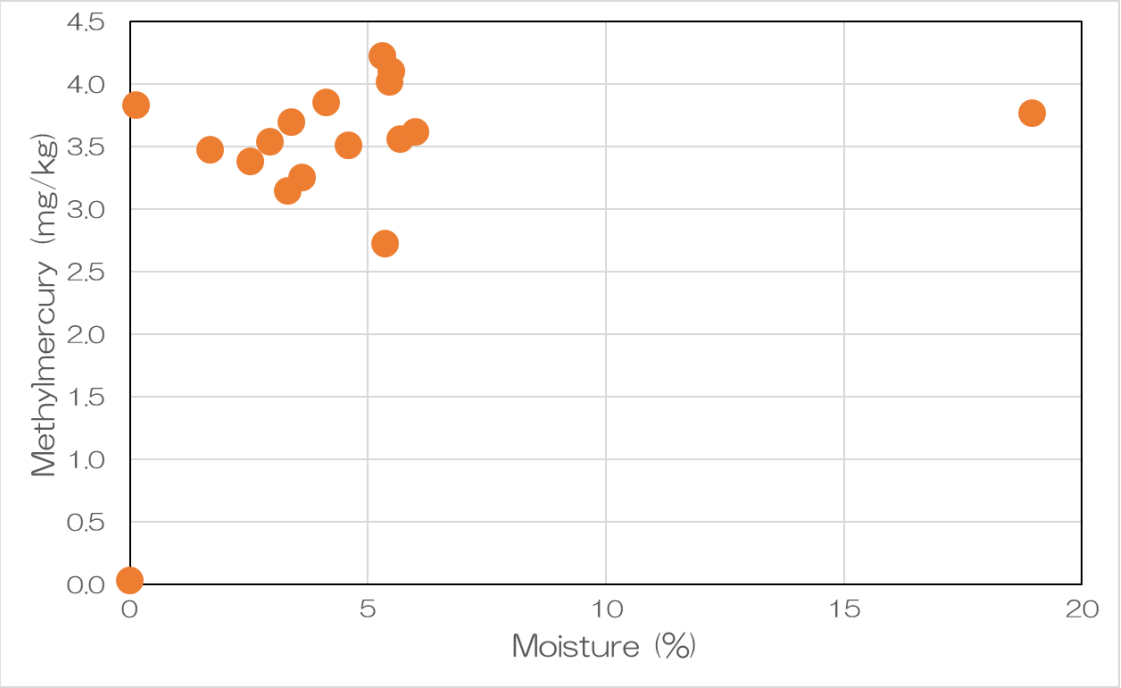


Fig. 15 Methylmercury and moisture

The reported moisture deviation was larger than the mercury deviation of the target parameter analysis (total mercury/methylmercury), but no noticeable relationship between moisture and both parameters of mercury concentration was found. The method used to measure moisture was to subtract the weight measured after drying from the weight measured before drying. Since the moisture was relatively small (median was about 4-5%), it was considered that the uncertainty of the moisture analysis was larger than the change in moisture during the PT analysis period.

12. Conclusion

In both of the results of total mercury and methylmercury analysis, many of the reported results from the participants were concentrated around the median value and IQRs of the results were relatively close. Thus, it was considered that the difference among many mercury analysis laboratories were relatively small. Both of total mercury and methylmercury analysis, some sort of analysis methods were performed among the participating laboratories. Significant differences (bias) among the analytical methods were not detected in both of the target parameters (total mercury / methylmercury).

The range of moisture reported from the participants were larger than the deviation of the target parameters even it was not the target of the PT. Also, the correlation between the mercury and moisture was not detected. Although the result of stability testing indicates that there was not a significant change of testing item during the PT, it can be considered that the difference of moisture affects the results of the PT because the deviations of the analysis results were small.