OSPAR-OIC Intercalibration Study on metals in produced water samples - a QUASIMEME Laboratory Performance Study

Drs. I. Velzeboer, ing. S.J.H. Crum, Drs. C.C. Karman

Final Report C014/09

Institute for Marine Resources and Ecosystem Studies

Wageningen IMARES

Location IJmuiden

Client: Dr. Angus Laurie
Energy Development Unit – Offshore Environment and Decommissioning Branch
Atholl House
86-88 Guilt Street
Aberdeen, AB11 6AR
Scotland

Publication Date: 06-02-2009

- Wageningen ***IMARES*** conducts research providing knowledge necessary for the protection, harvest and usage of marine and coastal areas.
- Wageningen ***IMARES*** is a knowledge and research partner for governmental authorities, private industry and social organisations for which marine habitat and resources are of interest.
- Wageningen ***IMARES*** provides strategic and applied ecological investigation related to ecological and economic developments.

Wageningen IMARES is registered in the Dutch trade record
Amsterdam nr. 34135929,
BTW nr. NL 811383696B04.

A_4_3_2-V6

Contents

Introduction

The Offshore Industry Committee (OIC) of OSPAR discussed in its 2008 meeting the reporting of inputs of metals from offshore installations. INPUT is currently compiling data and information on discharges and emissions to the OSPAR maritime area to be used in the Quality Status Report (QSR). This includes an assessment of the inputs of cadmium, lead, and mercury in produced water. Initial estimates were considered by OIC to be unrepresentative, since many of the analyses recorded values below the analytical detection limits of the techniques used.

Given the urgency of producing reliable information that could be used to prepare estimates of inputs of cadmium, lead, and mercury for the QSR, OIC agreed to conduct a further study, using the assistance of Quasimeme, to ensure quality assurance from sampling to measurement and reporting. The study would involve:

1. a single laboratory undertaking analyses to provide data for the QSR, and
2. an intercalibration exercise that would facilitate improved performance and reliability for ongoing monitoring programs.

This report focuses on the second part of the study proposed by OIC 2008; the intercalibration exercise. The study was carried out by Quasimeme, with Wageningen IMARES as coordinator.

The samples, prepared by IMARES, were measures by nine laboratories from the Netherlands, Denmark, United Kingdom and Norway. Quasimeme took care of the statistical data analysis.

Each laboratory received three samples: one produced water sample without additions, one produced water sample with a low addition of the metals cadmium, lead and mercury and one produced water sample with higher additions of these metals (see Table 2). The produced water was originating from the Wintershall A6A gas production platform on the German continental shelf.

The laboratories were invited to analyse the samples following their regular produced water analysis procedures. Most labs used ICP, either in combination with AES or MS. Three labs analysed mercury using AFS (see Appendix B).

The next chapter describes the material and methods including the sample preparation, analysis instruction, statistical analysis and the reporting. The third chapter shows the results of the intercalibration study, the participating labs are anonymized. The final chapter gives the conclusions and recommendations of the study.

1 Materials and Methods

1.1 Sample preparation

Produced water of a Wintershall A6A gas production platform in the North Sea is used for the OSPAR – OIC Intercalibration study on metals in produced water samples (see *Table 1* for metal concentrations). This produced water is first filtered with a Sartobran 300 filter (Fisher Scientific B.V., product number 087108) to remove bacteria and other particles. Afterwards the produced water is treated with hexane to remove oil and oil like components. To be consistent with standard offshore practices the stripped produced water is spiked with 200ng/mL gold (Gold ICP Standard 1000 mg/L Au CertiPUR, supplied by VWR product number 1.70321.0100) and 20% nitric acid (AnalaR NORMAPUR concentrated nitric acid 69 %, supplied by VWR product number 20425.322), to preserve the test material.

The samples are spiked with low and high concentrations of Cadmium (Cd), Lead (Pb) and Mercury (Hg). As the source material will contain metals, an unspiked sample is also included. The samples are spiked as indicated in *Table 2*.

Table 1. A6A produced water metal concentrations (in µg/L), as reported in the OSPAR one-off monitoring exercise

	Produced water
	Sample material
Cadmium	<0.1
Lead	<1
Mercury	<0.06

Table 2. Spiking concentration of the samples in µg/L

	Unspiked	*Low spiked*	*High spiked*
	QTM001PW	QTM002PW	QTM003PW
Cadmium	0	0.2	10
Lead	0	2	40
Mercury	0	0.15	2

1.2 Analysis instructions

The samples were sent to the participating laboratories in 100mL polyethylene bottles. Each laboratory has therefore received 3 anonymous samples, labelled with a code for identification of the samples. Also a 1mL gold standard was included, which could be used for matrix matching. With the delivery of the samples, each laboratory has received instructions for analysis and reporting (Appendix A).
Before measuring, the samples had to be diluted two times before running on the instrument to reduce the concentration of the acid. Final concentration: Au 100ng/mL and nitric acid 10%. The metals to be measured were Cadmium, Lead and Mercury. The 1000ppm gold standard could be used as an internal standard.

1.3 Statistical analysis

For a detailed overview of the statistical analysis performed on the study results we refer to the QUASIMEME report in Appendix C. In this paragraph the z-score will be explained, since it is the central parameter in the evaluation of the intercalibration study results.

The z-score is a measure of a laboratory's performance in relation to the overall performance of all participating laboratories. The z-score is calculated as follows:

$$z - \text{score} = \frac{\text{Mean from Laboratory - Assigned Value}}{\text{Total Error}}$$

The Assigned Value is calculated using the Cofino model from the values reported by the participating labs, and can be considered the mean value of the dataset. The Total Error is a measure of the variation in the reported values.

Following usual practices e.g. ISO 43, the z-scores can be interpreted as follows for laboratories which take part in Quasimeme to assure the quality of their data for use in international marine monitoring programmes:

IZI < 2 Satisfactory performance
2 < IZI < 3 Questionable performance
IZI > 3 Unsatisfactory performance

IZI > 6 frequently points to gross errors (mistakes with units during reporting, calculation or dilution errors, and so on). The following figure (*Figure 1*) illustrates the interpretation of the z-scores:

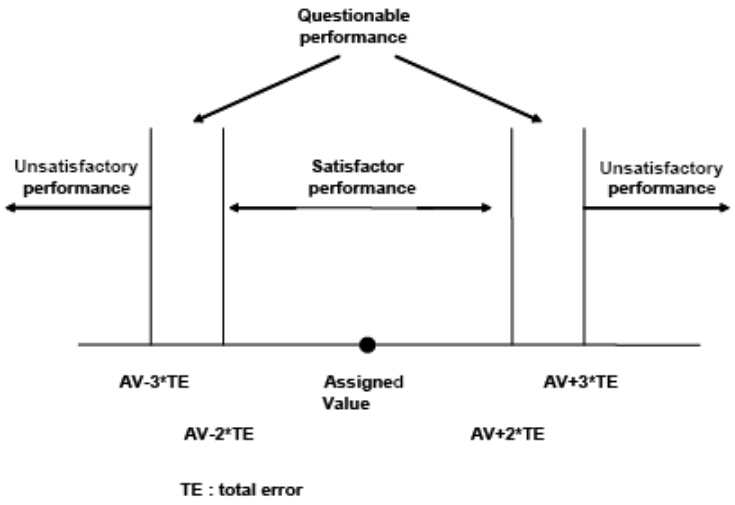


*Figure 1. Illustration of the interpretation of z-scores. AV+2*TE equals IZI=2; AV+3*TE equals IZI=3*

1.4 Reporting

The participating laboratories reported the results to IMARES in an excel-file, which included laboratory, sample-codes, measured concentrations of Cd, Pb and Hg and the method and instrument that were used for the measurements.
A statistical data analysis of the received results has taken place, with the final report of the intercalibration study as result.

2 Results

Raw analysis results and used instruments are reported in Appendix B, the laboratories are anonymized. The statistical data-analysis is performed by Quasimeme. This report is added in Appendix C, DE-17 Metals in produced water, Round 56 – Exercise 852. Figures of that report also shown below (*Figure 2, 3* and *4*). The net recovery of cadmium, lead and mercury respectively, for both the lower and higher spiked samples are shown. The recoveries were calculated by subtracting, for each laboratory, the reported value from the reported blank level. In case a detection level was given, then the reported value was reduced with 50% of this detection level. In addition, the recovery was calculated based on the spiked level. When the detection level was not given, no calculations were made.

In Figures 2, 3 and 4 the recovery of, respectively, cadmium, lead and mercury, is graphically presented. Tables 3 and 4 show the z-scores for, respectively, the various samples and each (anonymized) participating laboratory.

The recovery for cadmium (*Figure 2*) is not satisfactory for all labs. In case of the low spike, only 50% of the participating labs has reported an acceptable result (z – score of $|Z| < 2$). For the high spike sample the results are better since 78% of the labs gain a z – score of $|Z| < 2$ (Table 3 of the Quasimeme report, Appendix C).

Although better than for cadmium, the recovery for lead (*Figure 3*) is also not satisfactory for all labs. In case of the low spike, 63% of the participating labs have an acceptable z – score. For the high spike sample, no z – score is calculated, because it was not possible to set an assigned value (Table 2 and 3 of the Quasimeme report, Appendix C).

The results for mercury show a large variation (*Figure 4*). In case of the low spike, no z-score is calculated, because it was not possible to set an assigned value. For the high spike sample, 44% of the participating labs have a z-score of $|Z| < 2$. (Table 2 and 3 of the Quasimeme report, Appendix C).

It was not possible to set assigned values for cadmium, lead and mercury in the unspiked sample, mercury in the low spiked sample and lead in the high spiked sample. For this reason, no z-scores could be calculated and therefore no graphical output from the Cofino Model of these determinants is shown in Appendix 1 from the Quasimeme report.

Table 3 Summary of Z-scores for each of the samples. '-' means that no Z-scores could be calculated

	% data received	% of Zscores	% of Zscores	% of Zscores	% of Zscores
		|Z|<2	3>|Z|>2	6>|Z|<3	|Z|>6
		Satisfactory	**Questionable**	**Unsatisfactory**	**Extreme**
BLANC					
Cadmium	89	-	-	-	-
Lead	89	-	-	-	-
Mercury	100	-	-	-	-
LOW SPIKE					
Cadmium	89	50		13	25
Lead	89	63		13	13
Mercury	100	-	-	-	-
HIGH SPIKE					
Cadmium	100	78	11	11	
Lead	100	-	-	-	-
Mercury	100	44	11	11	22

Table 4 Ranked Z-scores for all determinants with assigned values

Labcode	Nobs	%
	Submitted	|Z|<2
AJ849	4	100
AJ851	4	100
AJ853	4	75
AJ855	4	75
AJ856	4	75
AJ850	2	50
AJ852	4	25
AJ854	1	0

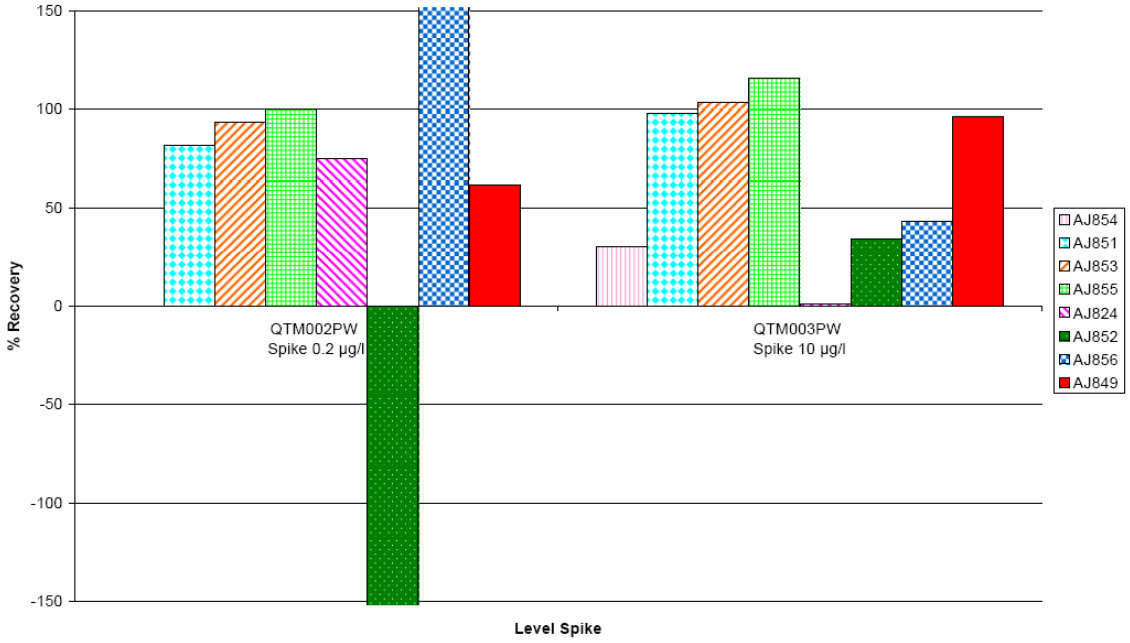


Figure 2. Recovery (net) of cadmium (source: QUASIMEME report, see Appendix C)

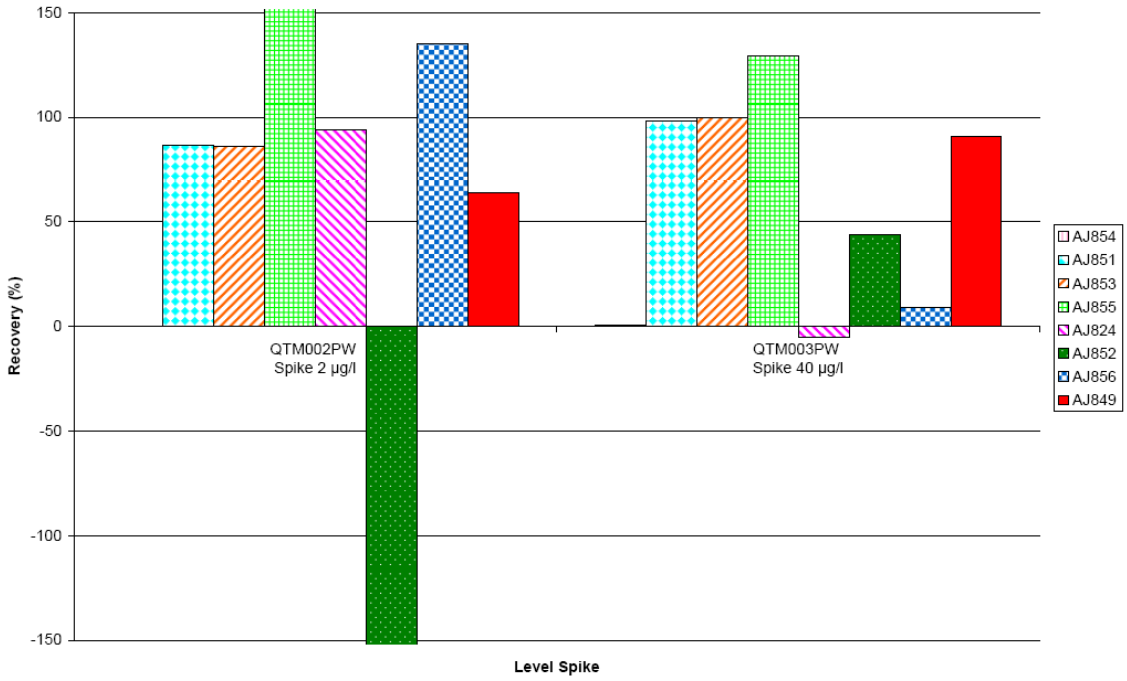


Figure 3. Recovery (net) of lead (source: QUASIMEME report, see Appendix C)

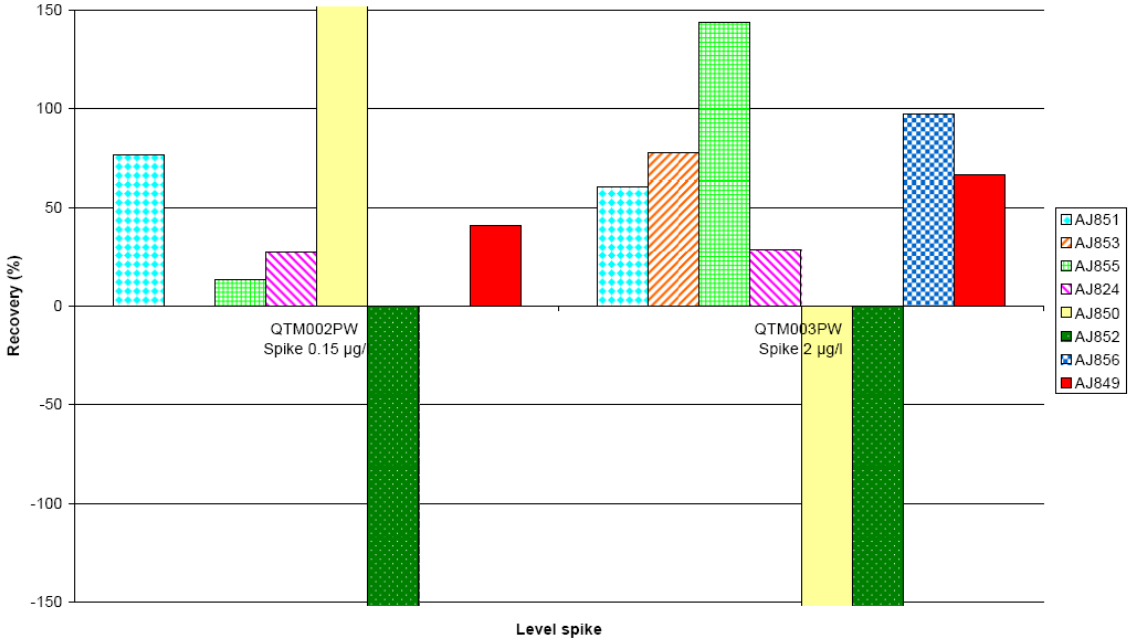


Figure 4. Recovery (net) of mercury (source: QUASIMEME report, see Appendix C)

3 Conclusions and recommendations

From this intercalibration study it can be concluded that there is a large variation between the values reported by the participating labs. This was to be expected, since produced water usually contains relatively high levels of salts (especially chlorine, Cl^-), which can cause interferences in the detection system when measuring with equipment based on ICP (Inductive Coupled Plasma). Next to that, organic components can also cause major interferences. The combination of these unfavourable characteristics makes produced water a difficult matrix to analyse.

Between laboratory differences are high as not each laboratory is capable of dealing appropriately with the disturbances caused by the difficult matrix (produced water). This depends on their procedures for sample preparation and the analytical procedure. The use of different instruments for the actual analysis is likely to introduce between-laboratory variation as well. In this study, the amount of data is not sufficient to comment on preferred detection instruments.

In order to be able to interpret the implications of the large 'between-laboratory' variation, we have projected this variation on the average metal concentrations in produced water, as reported for the OSPAR one-off monitoring exercise. The results are presented in *Table 5*, showing the 25- and 75-percentiles, which effectively means that, when analyzed by the participating laboratories, 50% of the reported values are expected to lie within these boundaries. These between laboratory differences are within a factor of 2-3, but may even be significantly higher (e.g., mercury in produced water from oil platforms).

Overall it can be concluded that metals in produced water are reported with a high variability, which should be taken into account when interpreting reported concentrations and loads to OSPAR by the member states. However, although the overall variability is high, it is still very well possible to produce correct results. Three labs in this study (AJ851, AJ849 and -to a lesser extent- AJ853) reported accurate metal concentrations.

It is, therefore, recommended that the procedures of these labs are studied in more detail in order to identify possible options to reduce the variation in reported metal concentrations.

It is further recommended to include produced water analysis (metals and other components) in the setting of QUASIMEME intercalibration studies, so that the performance of participating laboratories is continuously monitored.

Table 5. Reported mean metal concentrations in produced water from both Gas and Oil producing offshore installations (from the Ospar one-off monitoring exercise) and calculated 25 and 75% confidence intervals using the between lab variation from this intercalibration study

	Mean	Between lab	Percentiles	
	µg/l	VC%	25	75
GAS				
Cadmium	9.6	38	7.2	12.0
Lead	383.5	88	157	610.1
Mercury	1.6	76	0.8	2.5
OIL				
Cadmium	0.1	60	0.1	0.1
Lead	1.2	49	0.8	1.5
Mercury	0.4	146	0.01	0.7

4 Quality Assurance

IMARES utilises an ISO 9001:2000 certified quality management system (certificate number: 08602-2004-AQ-ROT-RvA). This certificate is valid until 15 December 2009. The organisation has been certified since 27 February 2001. The certification was issued by DNV Certification B.V. Furthermore, the chemical laboratory of the Environmental Division has NEN-AND-ISO/IEC 17025:2005 accreditation for test laboratories with number L097. This accreditation is valid until 27 March 2009 and was first issued on 27 March 1997. Accreditation was granted by the Council for Accreditation, with the last inspection being held on 1-4 September 2008.

Justification

Rapport C014/09
Project Number: 439.51028.01

The scientific quality of this report has been peer reviewed by a colleague scientist and the head of the department of Wageningen IMARES.

Approved: Dr. M. Kotterman
Project leader

Signature: 

Date: 06-02-2009

Approved: Drs. J.H.M. Schobben
Head of Department

Signature: 

Date: 06-02-2009

Appendix A. Technical specifications for the OSPAR-OIC Intercalibration study on metals in produced water samples

Description of the samples QTM001PW, QTM002PW and QTM003PW
The samples consist of produced water of a gas production platform in the North Sea. After cleaning and filtration, the produced water is spiked with 200 ng/mL gold and 20% nitric acid, to preserve the test material. Afterwards, the samples are spiked with different concentration of Cadmium (Cd), Lead (Pb) and Mercury (Hg). The samples are sent in three 100 mL polyethylene bottles. A 1 mL 1000 ppm gold standard will be sent with the samples, which can be used as an internal standard.

Study setup
Before measuring, the samples have to be diluted two times before running on the instrument to reduce the concentration of the acid. Final concentration: Au 100 ng/mL and nitric acid 10%. The metals to be measured are Cadmium, Lead and Mercury. The 1000 ppm gold standard can be used as an internal standard.

Reporting and timing
A template for an analysis report for the results will be send by e-mail as soon as possible. Final delivery of the analysis report must be before 19 December. In January (when all laboratory analysis reports have been received) the statistical analysis and reporting will take place, which will result in a final report of the inter calibration study. Raw analysis results will be included in the appendix, but the laboratories will be anonymized. Laboratories will receive a copy of the report in which their own results will be identified.

Appendix B. Raw analysis results

Lab.	Analysis	unit	QTM001PW	QTM002PW	QTM003PW	Instrument
AJ854	Cd Pb Hg	µg/l µg/l µg/l	<2.0 <0.1 <50	<2.0 <0.1 <50	4 0.306 <50	Cd en Pb: ICP-AES Hg: Cold vapor-AFS
AJ851	Cd Pb Hg	µg/l µg/l µg/l	0.168 0.93 < 0.05	0.331 2.66 0.14	9.98 40.3 1.23	FI-ICP-MS
AJ853	Cd Pb Hg	µg/l µg/l µg/l	0.374 0.50 <0.5	0.561 2.22 <0.5	10.7 40.4 1.8	ICP-MS
AJ855	Cd Pb Hg	µg/l µg/l µg/l	<0.1 <1 0.28	0.25 3.81 0.30	11.6 52.3 3.16	ICP-MS
AJ824	Cd Pb Hg	µg/l µg/l µg/l	<0.100 4.30 <0.0200	0.20 6.18 0.051	0.15 2.22 0.575	?
AJ850	Cd Pb Hg	µg/l µg/l µg/l	ND ND 22.1	ND ND 23.5	4.75 12.5 18.3	ICP-AES
AJ852	Cd Pb Hg	µg/l µg/l µg/l	8.13 21 15	3.59 8.2 7	11.56 38.6 7	ICP-MS
AJ856	Cd Pb Hg	µg/l µg/l µg/l	0.2 <1 <0.5	3.4 3.2 <0.5	4.5 4.1 2.2	Cd and Pb: ICP Hg: AFS
AJ849	Cd Pb Hg	µg/l µg/l µg/l	0.107 0.467 <0.02	0.23 1.75 0.0715	9.75 36.9 1.34	Cd and Pb: ICP-SFMS Hg: AFS

Spiked concentrations:

Analysis	unit	QTM001PW	QTM002PW	QTM003PW
Cd	µg/l	0	0.2	10
Pb	µg/l	0	2	40
Hg	µg/l	0	0.15	2

Appendix C. Quasimeme Report: DE-17 Metals in produced water

QUASIMEME

Laboratory Performance Studies

DE-17 Metals in produced water

Round 56 - Exercise 852
October 2008 to March 2009

Report

Issue 1: 23012009

QUASIMEME LABORATORY PERFORMANCE STUDIES

DE-17 Metals in produced water
Round 56 - Exercise 852

Data for exercise 852, DE-17, Metals in produced water, were returned by 9 of the 9 laboratories that participated in this study.

Test Materials

The test materials were prepared at IMARES, IJmuiden, the Netherlands, using produced water collected from a Wintershall gas production platform in the North Sea.

This produced water is first is filtered using a 0.45µm / 0.2µm double membrane filter (Sartobran 300 filter, Fisher Scientific B.V., product number 087108) to remove bacteria and other particles. Afterwards the produced water is shaken with hexane to remove oil and oily components. To be consistent with standard offshore practices the stripped produced water is spiked with 200ng/mL gold (Gold ICP Standard 1000 mg/L Au CertiPUR, supplied by VWR product number 1.70321.0100) and 20% nitric acid (AnalaR NORMAPUR concentrated nitric acid 69 %, supplied by VWR product number 20425.322), to preserve the test material.

The three test materials differed from each other in respect of their metal concentrations. QTM001PW was a unspiked material, while QTM002PW and QTM003PW were spiked with low and high concentrations of Cadmium (Cd), Lead (Pb) and Mercury (Hg).

Each batch of material was prepared in bulk and homogeneity was not tested as it was assumed to be homogeneous.

Data Assessment

All data received from participants are entered into the QUASIMEME database and assessed using a standard procedure to allow direct comparison between participants in each round and between rounds. The approach to the assessment is based on the standard, ISO 13528[1], the IUPAC International Harmonized Protocol for Proficiency Testing (Advanced Draft)[2]. Additions or differences in the assessment from these standards are given or referred to in this report.

In Table 1 all reported data and the spike levels are given. The summary statistics provided in Table 2 are based on Robust Statistics following DIN 38402, the Fast S method and the Cofino Model. However, the assigned value and the laboratory assessment using the z-score are based on the Cofino Model.

Comparison between the robust statistics and the Cofino model continues to be made, and where there are any significant discrepancies between the two methods then further investigative analysis is undertaken. Good agreement has been obtained (ca < 1% difference) for well-behaved measurements. The real differences occurred where there was an effect of methodology on the measurement, e.g. digestion of sediments for trace metal analysis. In these cases the Cofino model is generally able to separate the effects of the method on the results and provide a more reliable estimate of the measurement relating to the method. The standard,

1 ISO 13528:2005. Statistical methods for use in proficiency testing by interlaboratory comparisons.
2 The International Harmonized Protocol for Proficiency Testing of Analytical Chemistry Laboratories. IUPAC Technical Report. Thompson, M., Ellison, S.L.R., Wood, R. Interdivisional Working Party for Harmonization of Quality Assurance Schemes.

ISO 13528, includes statistics for proficiency testing schemes, and uses robust statistics as a basis for the assessment. However, it is generally acknowledged that robust statistics cannot cope with more than 10% extreme values, particularly with a skewed distribution. The Cofino model is able to routinely cope with these types of distribution and provide the best estimate of the consensus value, which may be used as the assigned value.

The Cofino model has been developed for the routine QUASIMEME assessments. From Round 45 the Cofino model uses a Normal Distribution Assumption (NDA). The assigned value is based on the Cofino NDA model without any trimming of the data. This approach includes all data in the evaluation and no subjective truncation or trimming is made. This model has been further developed to include Left Censored Values (LCV)[3]. The development of these models has been fully documented and published.[4, 5, 6] An overview of the assessment with explanation and examples is given in the paper *Assessment Rules for the Evaluation of the QUASIMEME LP Studies Data*[7].

The details of the Cofino Model are provided elsewhere,[6, 7] but in summary the approach is as follows:

- All data included in the assessment
- No data trimmed or downweighted
- Assigned values (AV) based on Cofino NDA model
- LCV[3] are also included, provided certain criteria are met

Tables and Plots

All reported data an the spiked levels are given in Table 1. The dataset is difficult for a good data-assessment. Quite a lot of data are lower than the detection limits of the laboratories. Sometimes laboratories report higher values for the unspiked sample than the spiked samples. The assigned value, total allowable error and descriptive statistics for each determinand are shown in Table 2. Table 3 outlines the percentage of satisfactory data and the limit of determination values submitted for each determinand. Table 4 shows the ranked z-scores of the laboratories that participated in this study. Table 5 gives the constant and proportional errors for each determinand and an overview of indicative values. The performance of the laboratories in this study is illustrated in the z-score histograms. Where the assigned value for a determinand is indicative, the values are plotted as their original reported concentrations. The rules for confirming whether the consensus value should be an assigned value or an indicative value are given in *Assessment Rules for the Evaluation of the QUASIMEME LP Studies Data*[7] with appropriate examples.

In Figure 1, 2 and 3 the nett recovery of cadmium, lead and mercury respectively are shown for both the lower and higher spiked samples. These recoveries were calculated by lowering, for each laboratory, the reported value with the reported blank level. In case the detection level was given than the reported value was lowered with 50% of this detection level. In addition, the recovery was calculated based on the spiked level given in Table 1. When the detection level was not given, no calculations were made. The graphs were cut off at 150% and -150% level as sometimes the recovery was very high or very low (negative).

---

3 *Left Censored Values* is the correct nomenclature for “less than” values

4 Cofino, W.P., Wells, D.E., Ariese, F., van Stokkum, I, Wengener, J. W. and Peerboom, R., *J. Chemometrics and Intelligent Laboratory Systems,* **53**, (2000) 37-55

5 Cofino, W. P., van Stokkum, I.H.M., van Steenwijk, J., and Wells, D E. *Analytica Chimica Acta* **533**, (2005) 31–39.

6 Wells, D.E., Cofino, W.P. and Scurfield, J. A. *The Application of the Cofino Model to Evaluate Laboratory Performance Study Data using the BandWidth Estimator.* FRS Marine Laboratory, Aberdeen, Collaborative Report No. 04/04 (2004)

7 Wells, D.E., and Scurfield, J. A. (2004). *Assessment Rules for the evaluation of the QUASIMEME Laboratory Performance Studies Data – version 2, February 2004.* QUASIMEME Project, FRS Marine Laboratory, 375 Victoria Road, Aberdeen AB11 9DB

Appendix I contains a page of graphical output from the Cofino Model for each determinand, describing the distribution of the data, which may be used in the interpretation and assessment.

Detailed descriptions of each of the plots in Appendix I, with examples are given in the Cofino Model handbook[6]. There are four plots for each determinand.

The upper left plot provides an impression of the probability density function for all data (black) and for the first mode (blue dotted) (PMF1) of the data. Superimposed on these pdf's is a histogram of the individual measurements given in grey color. This plot shows the distribution of the data as a whole, and of the data in the main mode (PMF1) on which the assigned value is based.

The Kilt Plot (Overlap Matrix) (upper right plot) provides an overview of the degree of overlap of each pair of data. It gives a clear indication of the homogeneity (or otherwise) of the data. As a key the white areas indicate maximum overlap of the pdf's and therefore highest agreement (an overlap of one implies that the two laboratories of the pair report exactly the same results), while the black area show the pairs in poor agreement.

The lower left plot is a ranked overview of all data with an error bar of ± 2 s.d. The numerical values are given in blue and the left censored values are given in red.

The ranked z-score plot (lower right) is NOT the FINAL ASSESSMENT. It is based on the Cofino mean of the data, which is normally also the assigned value. However, if there is any adjustment required to the assigned value as a result of the assessment, e.g. use of the nominal concentration or a trimmed value, then the final z-score given in the z-score histograms will reflect these changes. In most cases the two z-score plots will be the same. Any differences between this plot and the final assessment will be indicated in the report.

In addition, combined z-score plots are given per determinand, were the z-scores obtained for the different samples are given next to each other for the individual laboratories.

Appendix II contains a summary of the method codes reported.

The Assigned Values

The Assigned Value is obtained from the main mode of the data using the Cofino Model, and is centered around the highest density of values. Unless otherwise stated, the assigned value is based on this consensus value of *all* data.

Although *all* data are included in the assessment, those values that lie some distance from the Cofino mean (Assigned Value) contribute less to the mean than values which occur at or near the mean. The percentage of data in the main mode (blue area in the upper left Cofino Model plots) that contributes to the Cofino mean, and the Cofino standard deviation of this percentage of data are given in Table 1. The higher the percentage of data, the greater is the overall agreement of the measurements.

The Robust mean and between laboratory CV% are also given in Table 1 for comparison, but these values are not used as a basis for the assigned value or for the laboratory assessment.

The Indicative Values

In some instances it is not possible to set an assigned value, and an indicative value is given.

No assessment of laboratory performance is given where an indicative value is set. An overview of the assessment, with explanation, decision flowcharts and examples, is given in the paper *Assessment Rules for the evaluation of the QUASIMEME Laboratory Performance Studies Data*, available on the QUASIMEME website, www.quasimeme.org. A summary of the categories is given below, and the decisions for each determinand in each matrix are listed in Table 4.

Category 1
For data with the number of numerical observations ≥ 7
An assigned value is based on the Cofino mean when ≥ 33% of values have a z-score of $|Z| < 2$. Where < 33% of the data have $|Z| < 2$ the value is indicative. i.e. at least 33% must be in good agreement.

Category 2
For data with the number of numerical observations > 3 and < 7
An assigned value is based on the Cofino mean when ≥ 70% of values have a z-score of $|Z| < 3$ and a minimum of 4 observations have $|Z| < 2$. Otherwise the value is indicative. i.e. for small datasets, n = 4 to n = 7, there needs to be very good agreement and a maximum of one extreme value before an assigned value can be given.

Category 3
For data with the number of numerical observations < 4
No assigned value is given. Normally the median value is given as an indicative value.

Category 4
For data with the high Total Error% >100% in combination with bad performance, no assigned value is given.

The Z-score Assessment

A z-score [8] is calculated for each participant's data for each matrix / determinand combination which is given an assigned value. The z-score is calculated as follows:

$$z\text{-score} = \frac{\text{Mean from Laboratory} - \text{Assigned Value}}{\text{Total Error}}$$

It is emphasized that in many interlaboratory studies the between-laboratory standard deviation obtained from the statistical evaluation of the study is used as 'total error' in the formula above. In Quasimeme the total error is estimated independently taking the needs of present-day international monitoring programs as starting point. For each determinand in a particular matrix, a proportional error (PE) and a constant error (CE) have been defined. The total error depends on the magnitudes of these errors and on the assigned value:

$$\text{Total Error} = \frac{\text{Assigned Value} \times \text{Proportional Error (\%)}}{100} + 0.5 \times \text{Constant Error}$$

The values for the PE and CE are set by the Scientific Assessment Group and are monitored annually. The values are based on the following criteria:

Consistency of the required standard of performance to enable participating laboratories to monitor their assessment over time.

[8] International Harmonized Protocol for Proficiency Testing of (Chemical) Analytical Laboratories. M Thompson, R Wood, Journal of AOAC International Vol. 76, No. 4, 1993

Achievable targets in relation to the current state of the art and the level of performance needed for national and international monitoring programmes.

The assessment is based on ISO 43 as z-scores. The QUASIMEME model is designed to provide a consistent interpretation over the whole range of concentration of analytes provided, including an assessment where Left Censored Values (LCVs) are reported.

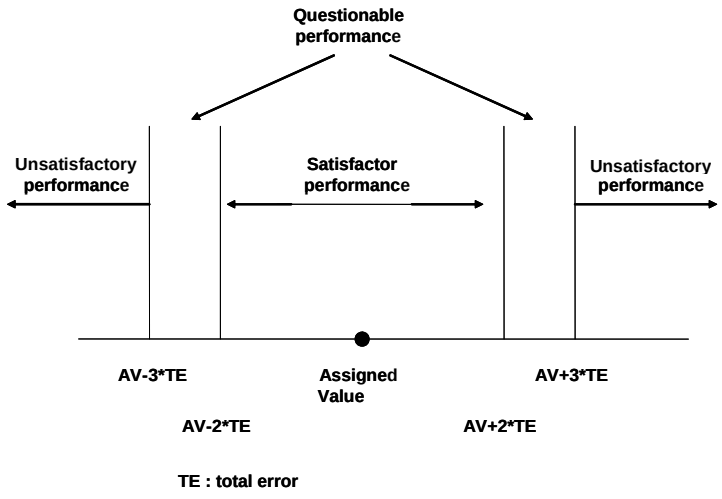
The proportional error is set at 6% for nutrients and for standard solutions, and 12.5% for all other matrices. This applies to all determinands. The constant error has been set for each determinand or determinand group (e.g. chlorinated biphenyls). This value was initially set to reflect the limit of determination, but is at present more closely related to the overall laboratory performance. The magnitude of the CE is set to provide a constant assessment in terms of z-score regardless of concentration. Therefore at low concentrations the level of accuracy required to obtain a satisfactory z-score is less stringent than at a high concentrations.

The performance of the laboratories is examined in detail when the total error exceeds 50% of the consensus concentration. If there is good agreement between the laboratories, i.e. the criteria to set an assigned value are met, the CE may be revised to a lower value reflecting the performance of laboratories for this measurement at lower concentrations. These revisions are undertaken at the time of the assessment and ratified by the Scientific Assessment Group. In making any adjustments to the CE an overall assessment of performance at these lower concentrations over a number of different rounds is reviewed. This provides evidence of a long-term trend of improved performance rather than a single set of data. When the agreement is judged to be insufficient, no assigned value is established. In such cases an indicative value is given.

Following usual practices e.g. ISO 43, the z-scores can be interpreted as follows for laboratories which take part in Quasimeme to assure the quality of their data for use in international marine monitoring programmes:

|Z| < 2 Satisfactory performance
2 < |Z| < 3 Questionable performance
|Z| > 3 Unsatisfactory performance

The following figure illustrates the interpretation of the z-scores:



|z| > 6 frequently points to gross errors (mistakes with units during reporting, calculation or dilution errors, and so on).

It is not possible to calculate a z-score for left censored values (LCV's). Quasimeme provides a simple quality criterion:

LCV/2 < (concentration corresponding to |z|=3) : LCV consistent with assigned value
LCV/2 > (concentration corresponding to |z|=3) : LCV inconsistent with assigned value, i.e. LCV reported by laboratory much higher than numerical values reported by other laboratories.

Z score key:	S – Satisfactory
	Q – Questionable
	U – Unsatisfactory
LCV key:	C – Consistent
	I – Inconsistent
No data:	B - Blank

All details of publications relating to the QUASIMEME assessment are available on the website at www.quasimeme.org.

If you have any comments or requests, please contact:

QUASIMEME Project Office
Wageningen UR
Alterra CWK
P.O. Box 47
6700 AA Wageningen
The Netherlands

Tel.: +31 (0) 317 48 65 46
Fax: +31 (0) 317 41 90 00
E-mail: quasimeme@wur.nl
http://www.quasimeme.org

Table 1. All reported values for metals in produced water

Exercise No.: 852 Round: 56

Group: DE-17 Year: 2008-2009

		Labcode								
Cadmium (µg/L)	**Spike level**	**AJ854**	**AJ851**	**AJ853**	**AJ855**	**AJ824**	**AJ850**	**AJ852**	**AJ856**	**AJ849**
QTM001PW	0	<2.0	0.168	0.374	<0.1	<0.100	ND	8.13	0.2	0.107
QTM002PW	0.2	<2.0	0.331	0.561	0.25	0.20	ND	3.59	3.4	0.23
QTM003PW	10	4	9.98	10.7	11.6	0.15	4.75	11.56	4.5	9.75
Lead (µg/L)										
QTM001PW	0	<0.1	0.93	0.50	<1	4.30	ND	21	<1	0.467
QTM002PW	2	<0.1	2.66	2.22	3.81	6.18	ND	8.2	3.2	1.75
QTM003PW	40	0.306	40.3	40.4	52.3	2.22	12.5	38.6	4.1	36.9
Mercury (µg/L)										
QTM001PW	0	<50	< 0.05	<0.5	0.28	<0.0200	22.1	15	<0.5	<0.02
QTM002PW	0.15	<50	0.14	<0.5	0.30	0.051	23.5	7	<0.5	0.0715
QTM003PW	2	<50	1.23	1.8	3.16	0.575	18.3	7	2.2	1.34

Figure 1. Recovery (nett) of cadmium in produced water of Exercise 852

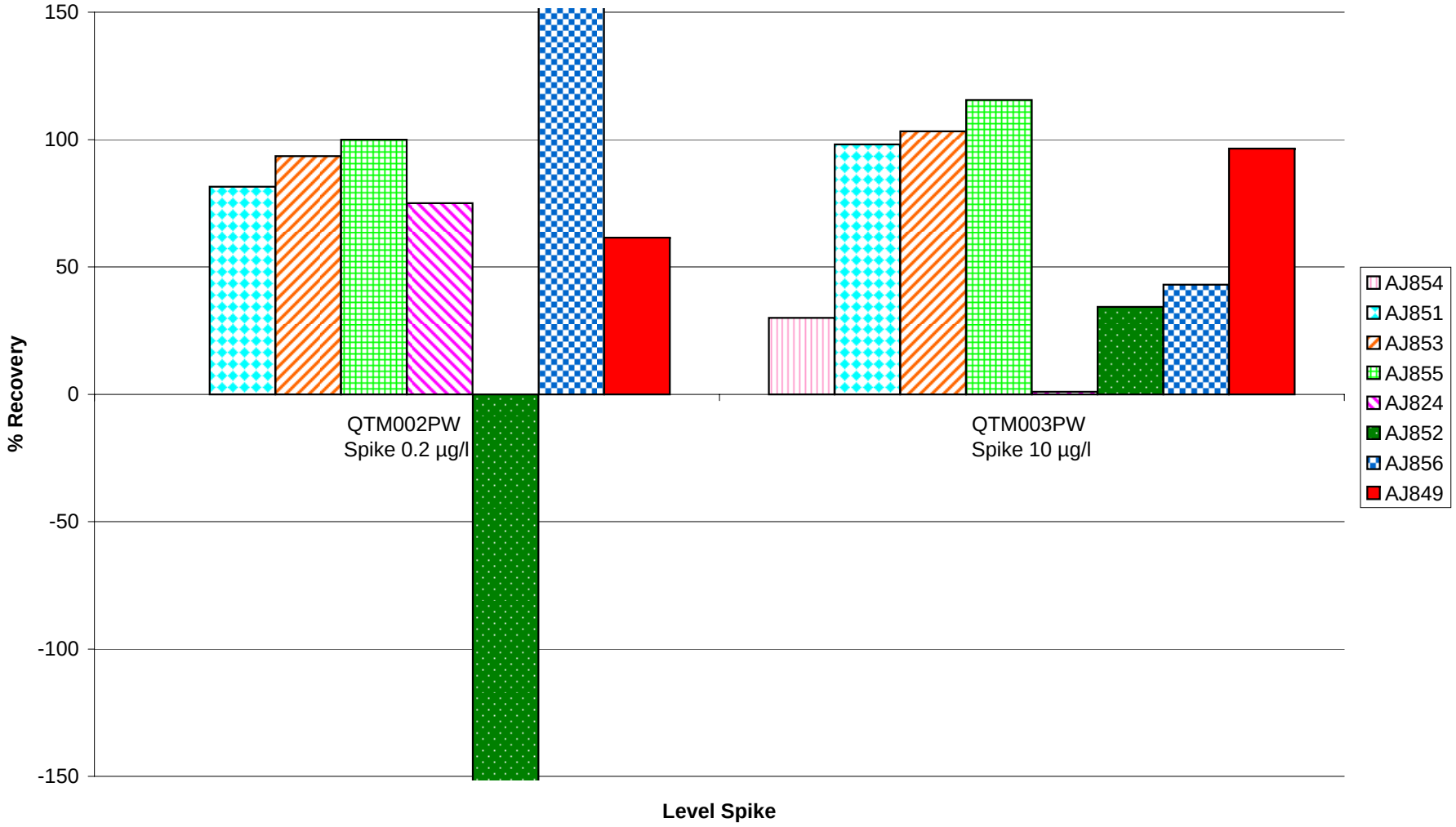


Figure 2: Recovery (nett) of lead in produced water of Exercise 852

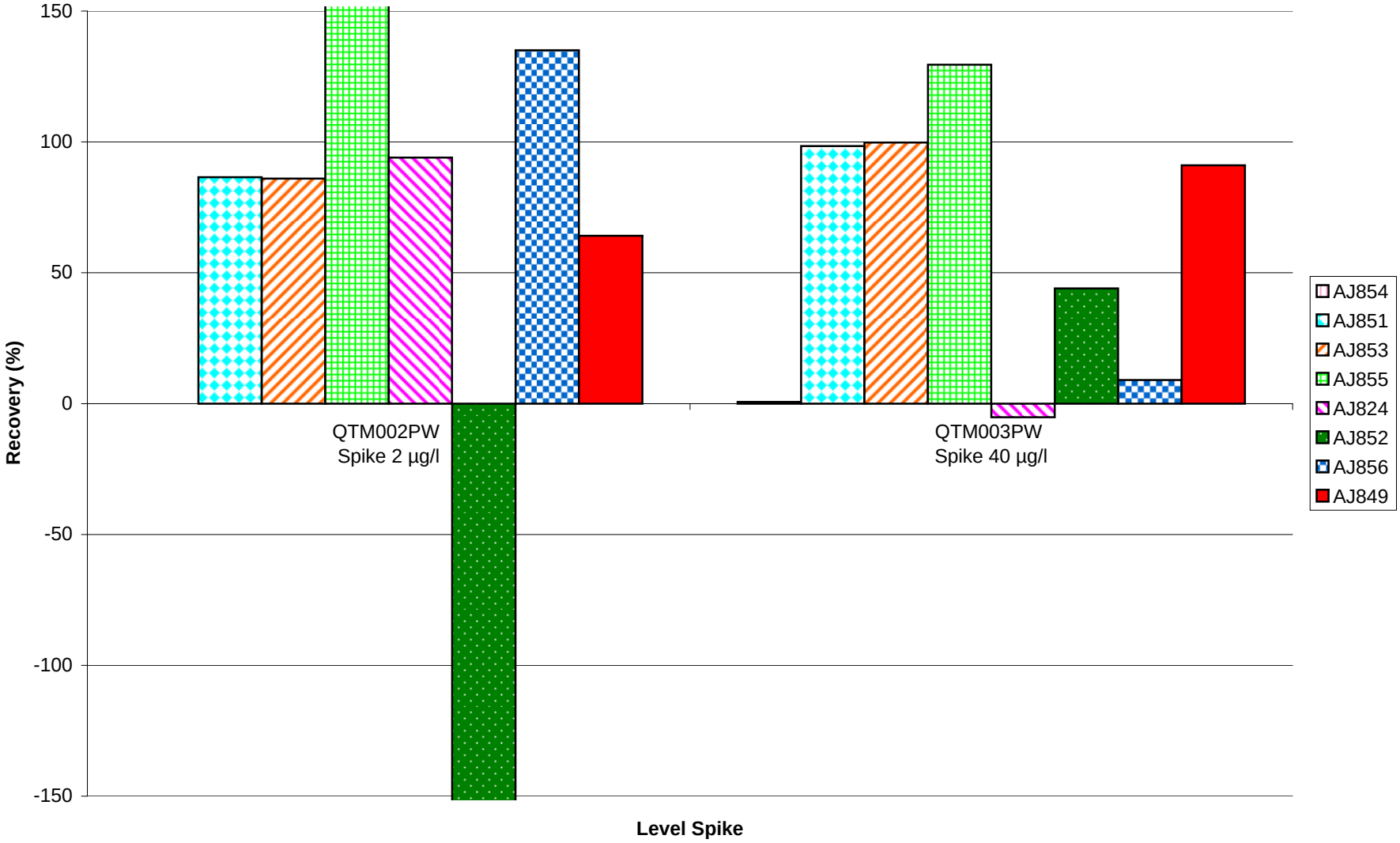


Figure 3: Recovery (nett) of mercury in produced water of Exercise 852

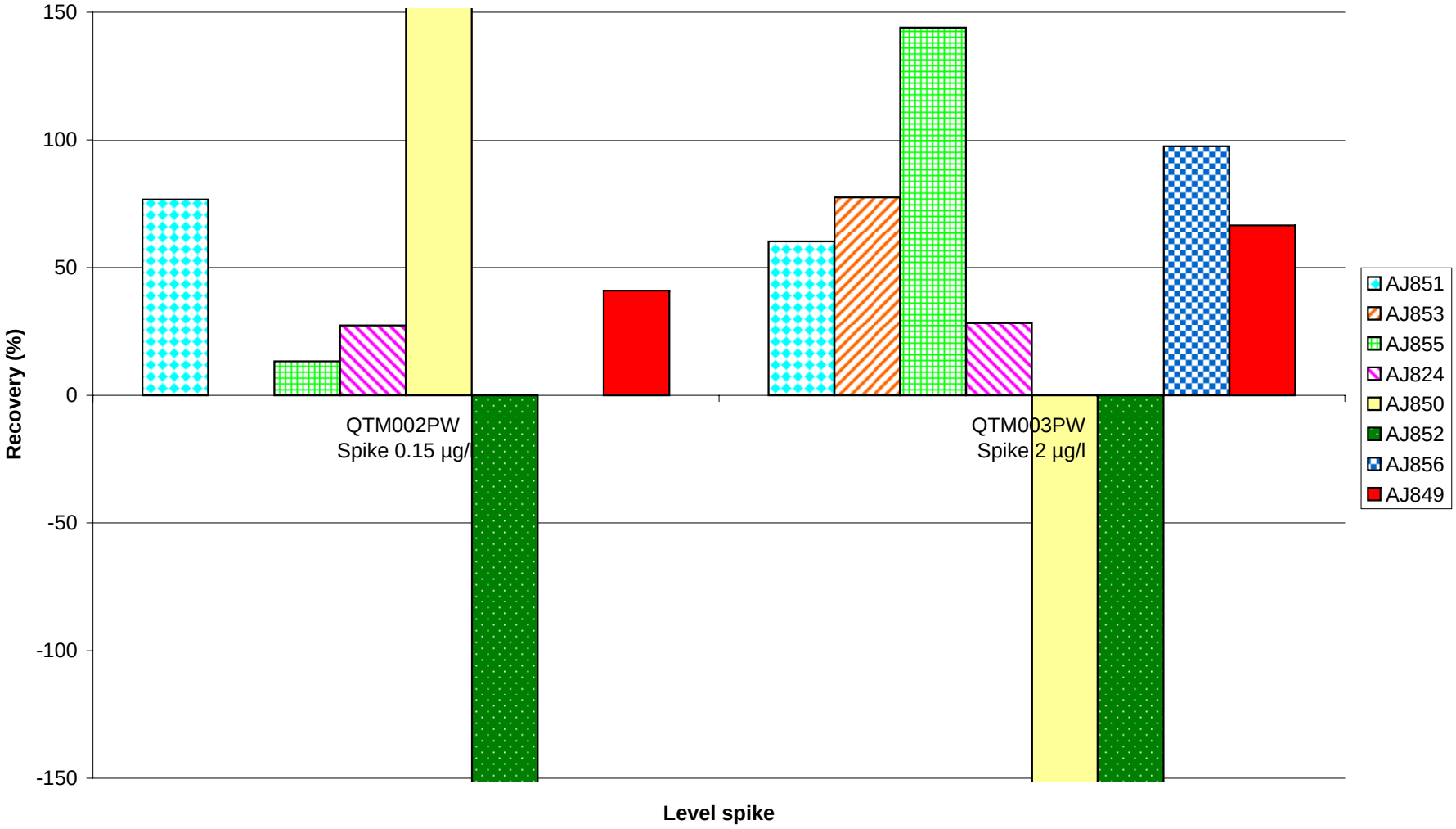


Table 2 Summary Statistics for QUASIMEME Participants

Exercise No. 852 Round 56
Group DE17 Year 2008-2009
Total Number of laboratories 9

Matrix/ Determinand	Assigned Value	Units	Total Error%	NObs Numerical	NObs LCV	Median Value	Basis for AV	Skewness	Model mean	Model Between Lab CV%	Model percentage in PMF1	DIN38402 Mean	DIN38402 Between Lab CV%	FastS Mean	FastS Between Lab CV%		
QTM001PW																	
Cadmium		*µg/l*		*5*	*3*	*0.20*	*NDA*	*1.50*	*0.14*	*86.0*	*56.8*	*0.21*	*97.2*	*0.16*	*62.7*	0	FALSE
Lead		*µg/l*		*5*	*3*	*0.93*	*NDA*	*1.39*	*0.60*	*94.2*	*49.2*	*1.11*	*92.5*	*0.61*	*90.4*	0	FALSE
Mercury		*µg/l*		*3*	*6*	*15.0*	*NDA*	*-0.40*	*0.71*	*519*	*36.8*	*12.5*	*160*	*18.6*	*40.9*	0	FALSE
																	TRUE
																	TRUE
QTM002PW																	TRUE
Cadmium	0.289	µg/l	25.9	7	1	0.33	NDA	0.93	0.29	60.1	62.0	0.31	87.2	0.25	58.7	1	FALSE
Lead	2.928	µg/l	25.2	7	1	3.20	NDA	0.89	2.93	49.1	59.3	3.84	55.7	2.67	49.3	1	FALSE
Mercury		*µg/l*		*6*	*3*	*0.22*	*NDA*	*1.50*	*0.14*	*146*	*62.5*	*0.14*	*279*	*0.11*	*171*	0	FALSE
																	TRUE
																	TRUE
QTM003PW																	TRUE
Cadmium	8.937	µg/l	25.0	9	0	9.75	NDA	-0.50	8.94	37.7	62.9	7.83	35.9	10.7	24.9	1	FALSE
Lead		*µg/l*		*9*	*0*	*36.9*	*NDA*	*-0.15*	*26.4*	*87.6*	*79.1*	*25.6*	*52.8*	*41.1*	*44.1*	0	FALSE
Mercury	1.685	µg/l	25.0	8	1	2.00	NDA	1.83	1.69	75.7	67.0	2.22	90.9	1.49	84.7	1	FALSE

RTWUR 09/09/2008

Entries in italics are given as indicative values only
NObs = Total number of observations reported

Table 3 Summary of Z scores and Left Censored Values (LCVs)

Exercise No.	852	Round	56
Group	DE17	Year	2008-2009
Total Number of laboratories			9

Matrix/	% of the	% of Zscores	% of Zscores	% of Zscores	% of Zscores	% Consistent	% Inconsistent	Minimum	Maximum
Determinand	data received	|Z|<2	3>|Z|>2	6>|Z|>3	|Z|>6	LCV	LCV	LCV	LCV
		Satisfactory	Questionable	Unsatisfactory	Extreme				
QTM001PW									
Cadmium	89							0.10	2.00
Lead	89							0.10	1.00
Mercury	100							0.02	50.00
QTM002PW									
Cadmium	89	50		13	25		13	2.00	2.00
Lead	89	63		13	13	13		0.10	0.10
Mercury	100							0.50	50.00
QTM003PW									
Cadmium	100	78	11	11					
Lead	100								
Mercury	100	44	11	11	22		11	50.00	50.00

RTWUR 09/09/2008

Units of measurement for LCVs ('Less than') are given in Table 1

Table 4 Ranked Z scores for all determinands with assigned values

Exercise No.	852	Round	56	
Group	DE17	Year	2008-2009	
Total Number of laboratories			9	

Labcode	NObs |Z|<2	Possible % |Z|<2	Labcode	NObs Submitted	Actual % Z|<2
AJ849	4	100	AJ849	4	100
AJ851	4	100	AJ851	4	100
AJ853	3	75	AJ853	4	75
AJ855	3	75	AJ855	4	75
AJ856	3	75	AJ856	4	75
AJ824	1	25	AJ850	2	50
AJ850	1	25	AJ824	4	25
AJ852	1	25	AJ852	4	25
AJ854	0	0	AJ854	1	0

NObs (|Z| < 2)	Total number of satisfactory observations (|Z| < 2) from each laboratory.
% (|Z| < 2) Possible	Total number of satisfactory observations as a % of the total number of determinands with assigned values
NObs submitted	Number of datasets submitted by each laboratory for each determinand with an assigned value.
% (|Z| < 2) Actual	% of observations submitted that were satisfactory (|Z| < 2)

Table 5 Constant and Proportional Errors and Criteria for Indicative Values

Exercise No.	852	Round	56
Group	DE17	Year	2008-2009
Total Number of laborator	9		

Matrix/ Determinand	Proportional Error	Constant Error	Indicative Category
QTM001PW			
Cadmium	*25*	*0.005*	*2*
Lead	*25*	*0.01*	*2*
Mercury	*25*	*0.001*	*3*
QTM002PW			
Cadmium	25	0.005	
Lead	25	0.01	
Mercury	*25*	*0.001*	*2*
QTM003PW			
Cadmium	25	0.005	
Lead	*25*	*0.01*	*1*
Mercury	25	0.001	

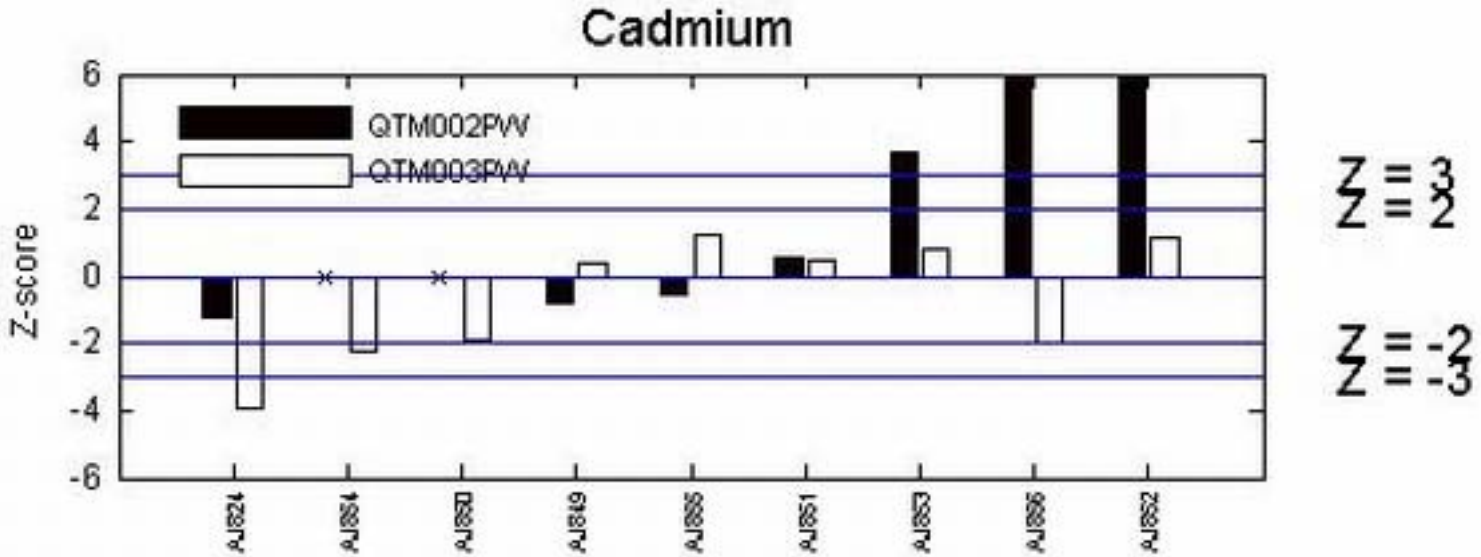
RTWUR 09/09/2008

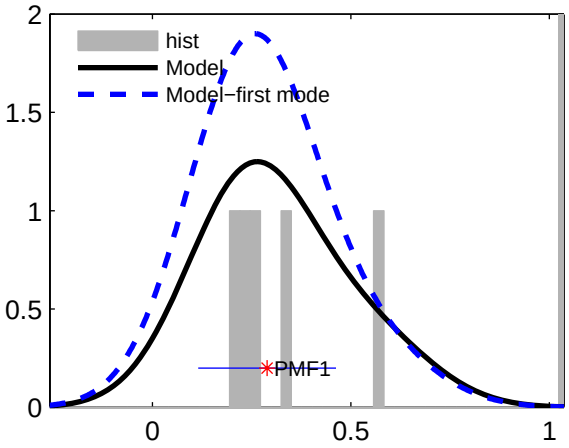
Indicative values are shaded grey

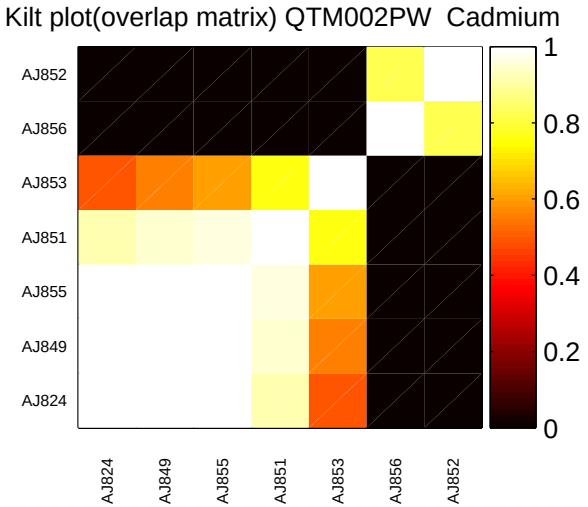
Category 1: NObs (num) >=7, AV requires more than 33% |Z|<2
and a minimum of 4 observations with |Z|<2
Category 2: 3<NObs (num) <7, AV requires >70% of data have |Z|<3
and a minimum of 4 observations with |Z|<2

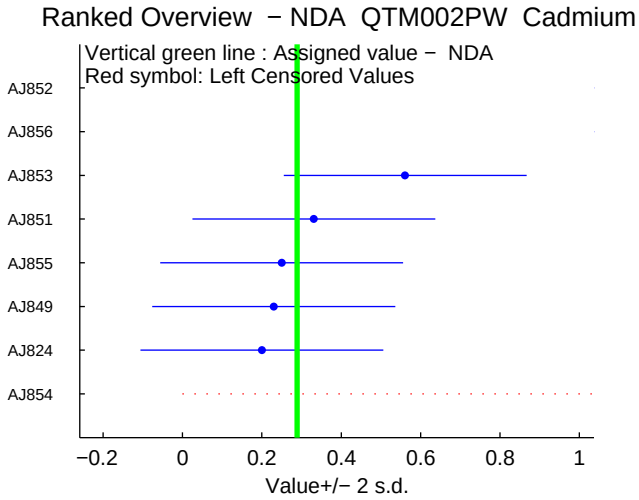
Category 3: NObs (num)<4, No assigned value
Category 4: Total Error greater than 100%
Category 5: Judgement of QPO

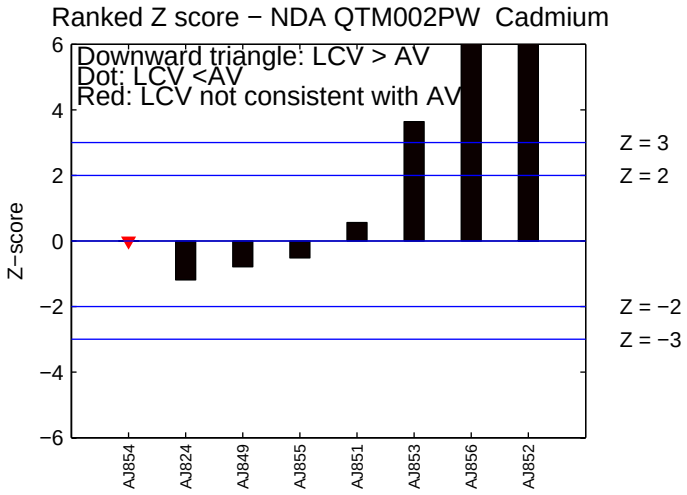
Appendix I
Summary Plots

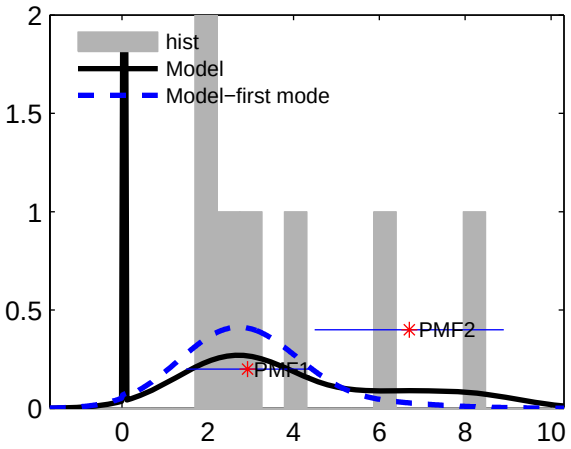

Cadmium
QTM002PVV
QTM003PVV
Z-score
6
4
2
0
-2
-4
-6
Z = 3
Z = 2
Z = -2
Z = -3

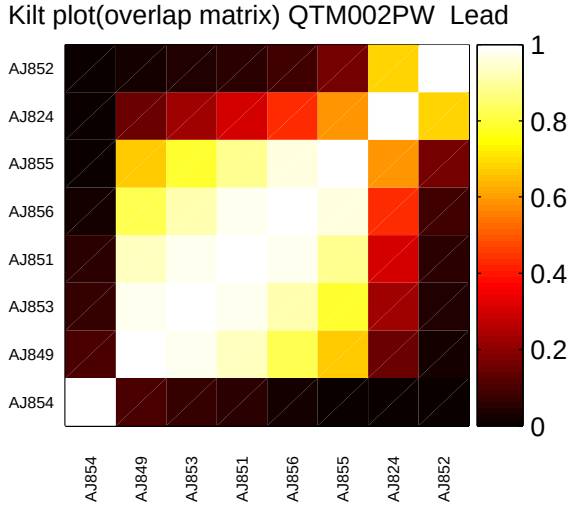

hist
Model
Model−first mode
PMF1
2
1.5
1
0.5
0
0
0.5
1

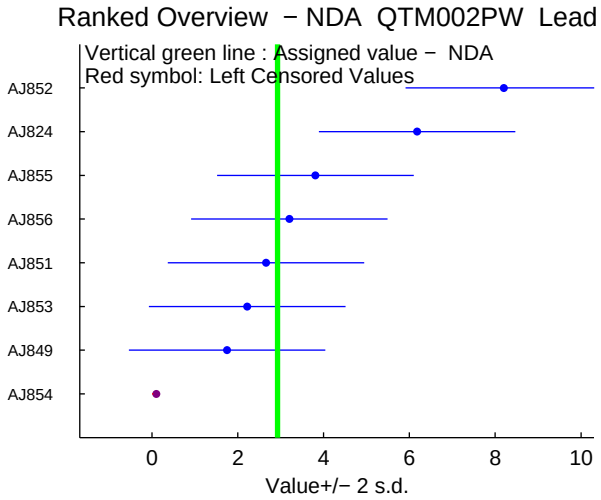

Kilt plot(overlap matrix) QTM002PW Cadmium
AJ852
AJ856
AJ853
AJ851
AJ855
AJ849
AJ824
1
0.8
0.6
0.4
0.2
0

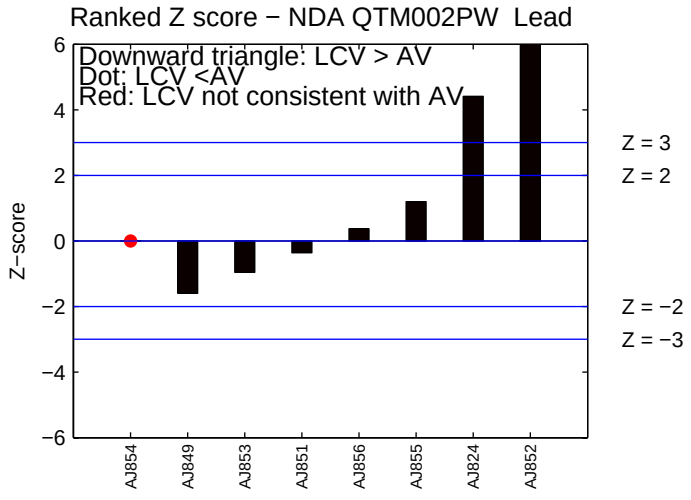

Ranked Overview − NDA QTM002PW Cadmium
Vertical green line : Assigned value − NDA
Red symbol: Left Censored Values
AJ852
AJ856
AJ853
AJ851
AJ855
AJ849
AJ824
AJ854
−0.2
0
0.2
0.4
0.6
0.8
1
Value+/− 2 s.d.

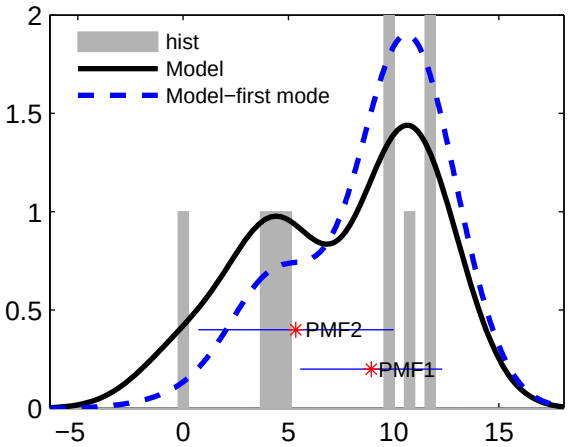

Ranked Z score − NDA QTM002PW Cadmium
Downward triangle: LCV > AV
Dot: LCV <AV
Red: LCV not consistent with AV
Z−score
6
4
2
0
−2
−4
−6
Z = 3
Z = 2
Z = −2
Z = −3
AJ854
AJ824
AJ849
AJ855
AJ851
AJ853
AJ856
AJ852

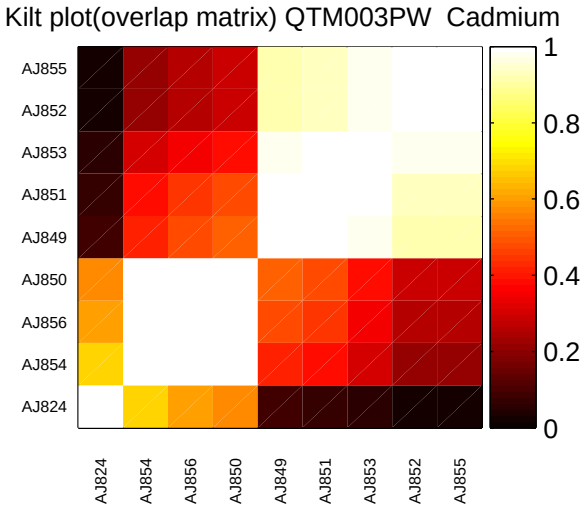

hist
Model
Model−first mode
PMF2
PMF1
2
1.5
1
0.5
0
0
2
4
6
8
10

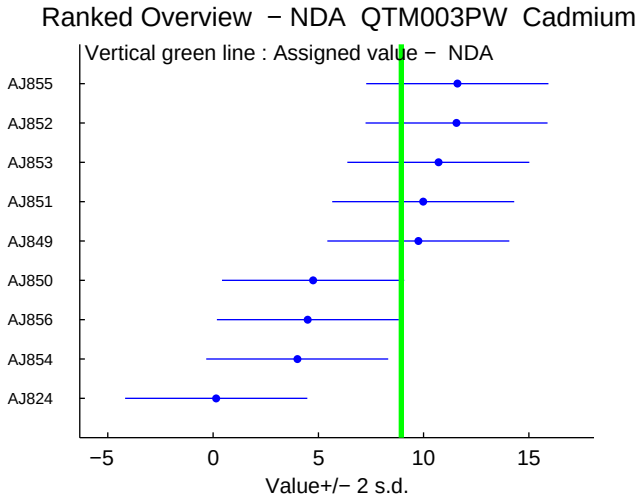

Kilt plot(overlap matrix) QTM002PW Lead
AJ852
AJ824
AJ855
AJ856
AJ851
AJ853
AJ849
AJ854
1
0.8
0.6
0.4
0.2
0
AJ854
AJ849
AJ853
AJ851
AJ856
AJ855
AJ824
AJ852

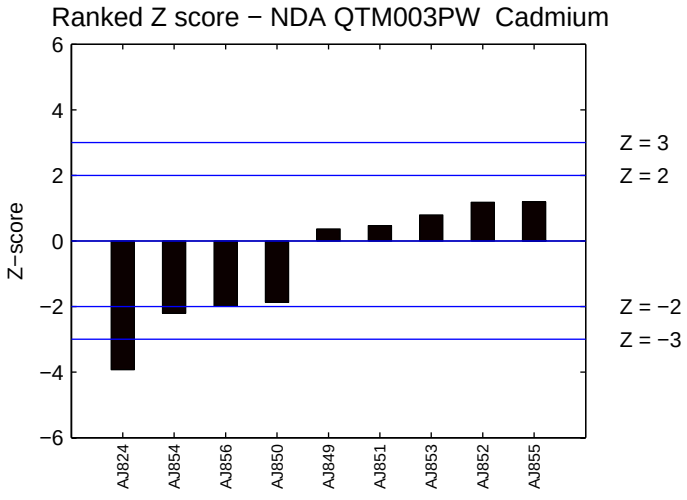

Ranked Overview − NDA QTM002PW Lead
Vertical green line : Assigned value − NDA
Red symbol: Left Censored Values
AJ852
AJ824
AJ855
AJ856
AJ851
AJ853
AJ849
AJ854
0
2
4
6
8
10
Value+/− 2 s.d.

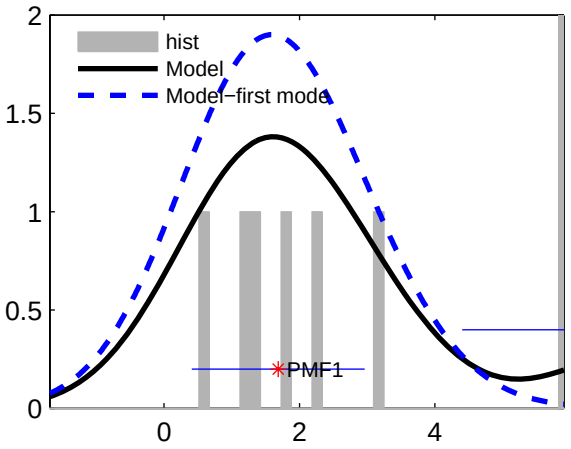

Ranked Z score − NDA QTM002PW Lead
Downward triangle: LCV > AV
Dot: LCV <AV
Red: LCV not consistent with AV
Z−score
6
4
2
0
−2
−4
−6
Z = 3
Z = 2
Z = −2
Z = −3
AJ854
AJ849
AJ853
AJ851
AJ856
AJ855
AJ824
AJ852

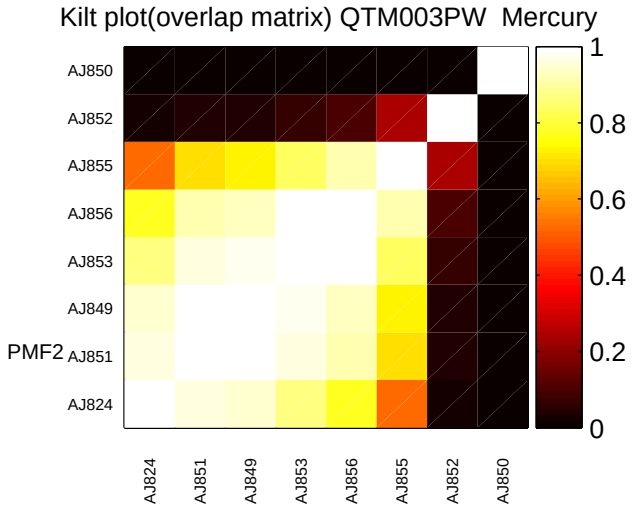

hist
Model
Model−first mode
PMF2
PMF1
2
1.5
1
0.5
0
−5
0
5
10
15

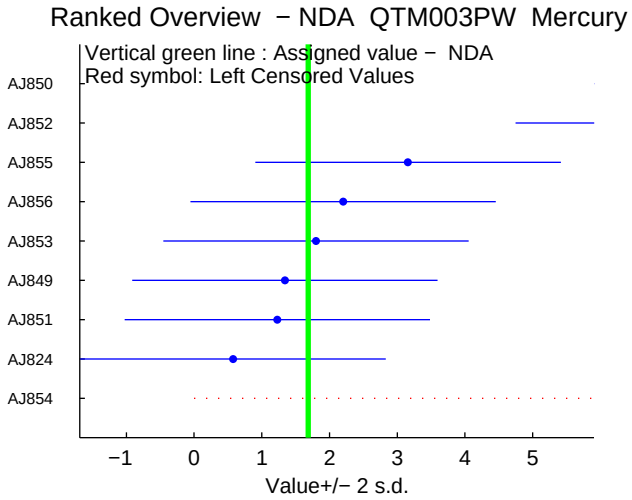

Kilt plot(overlap matrix) QTM003PW Cadmium
AJ855
AJ852
AJ853
AJ851
AJ849
AJ850
AJ856
AJ854
AJ824
1
0.8
0.6
0.4
0.2
0
AJ824
AJ854
AJ856
AJ850
AJ849
AJ851
AJ853
AJ852
AJ855

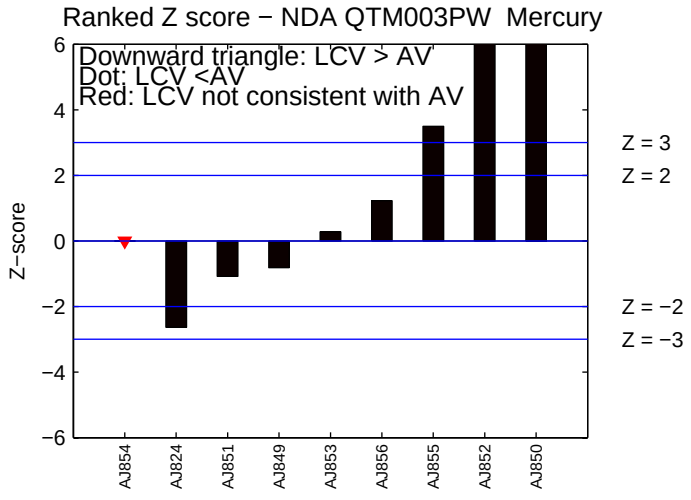

Ranked Overview − NDA QTM003PW Cadmium
Vertical green line : Assigned value − NDA
AJ855
AJ852
AJ853
AJ851
AJ849
AJ850
AJ856
AJ854
AJ824
−5
0
5
10
15
Value+/− 2 s.d.


Ranked Z score − NDA QTM003PW Cadmium
Z−score
6
4
2
0
−2
−4
−6
Z = 3
Z = 2
Z = −2
Z = −3
AJ824
AJ854
AJ856
AJ850
AJ849
AJ851
AJ853
AJ852
AJ855


hist
Model
Model−first mode
PMF1
2
1.5
1
0.5
0
0
2
4


Kilt plot(overlap matrix) QTM003PW Mercury
AJ850
AJ852
AJ855
AJ856
AJ853
AJ849
PMF2 AJ851
AJ824
AJ824
AJ851
AJ849
AJ853
AJ856
AJ855
AJ852
AJ850
1
0.8
0.6
0.4
0.2
0


Ranked Overview − NDA QTM003PW Mercury
Vertical green line : Assigned value − NDA
Red symbol: Left Censored Values
AJ850
AJ852
AJ855
AJ856
AJ853
AJ849
AJ851
AJ824
AJ854
−1
0
1
2
3
4
5
Value+/− 2 s.d.


Ranked Z score − NDA QTM003PW Mercury
Downward triangle: LCV > AV
Dot: LCV <AV
Red: LCV not consistent with AV
Z−score
6
4
2
0
−2
−4
−6
Z = 3
Z = 2
Z = −2
Z = −3
AJ854
AJ824
AJ851
AJ849
AJ853
AJ856
AJ855
AJ852
AJ850