

Australian Government

Department of Industry, Science and Resources National Measurement Institute

Proficiency Test Final Report AQA 22-20 Hydrocarbons in River Water

April 2023

© Commonwealth of Australia 2023.

Unless otherwise noted, the Commonwealth owns the copyright (and any other intellectual property rights, if any) in this publication.

All material in this publication is provided under a Creative Commons Attribution 4.0 International Licence (CC BY 4.0), with the exception of:

- the Commonwealth Coat of Arms;
- the logo of the Department of Industry, Science and Resources;
- photographs of our staff and premises; and
- content supplied by third parties.

Creative Commons Attribution 4.0 International Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. A summary of the licence terms is available at:

<u>creativecommons.org/licenses/by/4.0/</u>. Further details are available on the Creative Commons website, at: <u>creativecommons.org/licenses/by/4.0/legalcode</u>.

You may not copy, distribute, transmit or adapt any material in this publication in any way that suggests that this department or the Commonwealth endorses you or any of your services or products.

Attribution

Material contained in this publication is to be attributed to this department as:

© Commonwealth of Australia, Department of Industry, Science and Resources, Proficiency Test Final Report AQA 22-20 Hydrocarbons in River Water, 2023.

Third party copyright

Wherever a third party holds copyright in material contained in this publication, the copyright remains with that party. Their permission may be required to use the material.

This department has made all reasonable efforts to:

- clearly label material where the copyright is owned by a third party;
- ensure that the copyright owner has consented to this material being contained in this publication.

Using the Commonwealth Coat of Arms

The terms of use for the Coat of Arms are available on the Department of Prime Minister and Cabinet's website, at <u>www.pmc.gov.au/resource-centre/government/commonwealth-coat-arms-information-and-guidelines</u>

ACKNOWLEDGMENTS

This study was conducted by the National Measurement Institute (NMI). Support funding was provided by the Australian Government Department of Industry, Science and Resources.

I would like to thank the management and staff of the participating laboratories for supporting the study. It is only through widespread participation that we can provide an effective service to laboratories.

The assistance of the following NMI staff members in the planning, conduct and reporting of the study is acknowledged.

Jenny Xu Geoff Morschel Hamish Lenton Isaac Schipp Luke Baker

Raluca Iavetz Manager, Chemical Reference Values 105 Delhi Road, North Ryde, NSW 2113, Australia Phone: +61 2 9449 0178 Email: raluca.iavetz@measurement.gov.au



NATA Accredited for compliance with ISO/IEC 17043

TABLE OF CONTENTS

SUMMARY		1
1 INTRODU	JCTION	3
1.1 NMI Pr	oficiency Testing Program	3
1.2 Study /	Aims	3
1.3 Study (Conduct	3
2 STUDY II	NFORMATION	4
2.1 Study	Γimetable	4
2.2 Particip	pation	4
2.3 Selecti	on of Hydrocarbon Analytes	4
2.4 Test M	aterial Preparation	4
2.5 Homog	eneity and Stability of Test Materials	5
2.6 Test M	aterial Storage, Dispatch and Receipt	6
2.7 Instruc	tions to Participants	6
2.8 Interim	Report	6
3 PARTICII	PANT LABORATORY INFORMATION	7
3.1 Particip	oants' Test Methods	7
3.2 Basis c	of Participants' Measurement Uncertainty Estimates	7
3.3 Particip	oants' Comments	8
4 PRESEN	TATION OF RESULTS AND STATISTICAL ANALYSIS	10
4.1 Results	s Summary	10
4.2 Outliers	s and Gross Errors	10
4.3 Assign	ed Value	10
4.4 Robust	Average and Robust Between-Laboratory Coefficient of Variation	10
4.5 Perform	nance Coefficient of Variation	10
4.6 Target	Standard Deviation for Proficiency Assessment	11
4.7 z-Score	e	11
4.8 <i>E</i> _n -Sco	re	11
4.9 Tracea	bility and Measurement Uncertainty	11
5 TABLES	AND FIGURES	12
6 DISCUSS	SION OF RESULTS	60
6.1 Assign	ed Value	60
6.2 Measu	rement Uncertainty Reported by Participants	61
6.3 z-Score	e	61
6.4 <i>E_n</i> -Sco	re	66
6.5 False N	Jegatives	67
6.6 Report	ing of Additional Analytes	67
6.7 Particip	oants' Analytical Methods	68
6.8 Certifie	d Reference Materials	71
6.9 Summa	ary of Participants' Performance	72
6.10 Compa	rison with Previous Studies	75
7 REFERE	NCES	78
APPENDIX 1	SAMPLE PREPARATION	79
APPENDIX 2	ASSESSMENT OF HOMOGENEITY AND STABILITY	80

A2.1 Homoge	neity	80
A2.2 Stability		81
APPENDIX 3	ROBUST AVERAGE AND ASSOCIATED UNCERTAINTY, z-SC	ORE AND
En-SCORE CAL	CULATIONS	84
A3.1 Robust A	verage and Associated Uncertainty	84
A3.2 z-Score	and <i>E_n</i> -Score Calculation	84
APPENDIX 4	PARTICIPANTS' TEST METHODS	85
APPENDIX 5	ACRONYMS AND ABBREVIATIONS	88

THIS PAGE IS INTENTIONALLY BLANK

SUMMARY

AQA 22-20 Hydrocarbons in River Water commenced in October 2022. Nineteen laboratories registered to participate, and all participants submitted results.

The sample set consisted of four water samples. Samples were prepared in the Sydney NMI laboratory using surface water from Browns Waterhole in Sydney. Participants measured total recoverable hydrocarbons (TRH) in Sample S1, volatile hydrocarbons (C6 to C10), benzene, toluene, ethylbenzene and xylenes (BTEX) in Sample S2, and polycyclic aromatic hydrocarbons (PAHs) in Samples S3 and S4.

The assigned values for all scored analytes were the robust averages of participants' results. The associated uncertainties were estimated from the robust standard deviations of participants' results.

Traceability: The consensus of participants' results is not traceable to any external reference, so although expressed in SI units, metrological traceability has not been established.

The outcomes of the study were assessed against the aims as follows:

• Compare the performances of participants and assess their accuracy in the identification and measurement of petroleum hydrocarbon pollutants in river water.

Laboratories 2, 5, 7, 14, 15, 16 and 18 reported results for all 18 scored analytes.

For Sample S1, one participant reported hydrocarbon ranges outside of the recommended National Environment Protection Measure (NEPM) ranges.

Four participants did not report results for analytes that they tested for and were present in the samples (total of five results). One participant reported numeric results for analytes that were not spiked into the samples (total of two results).

Of 315 *z*-scores, 277 (88%) returned a score of $|z| \le 2.0$, indicating a satisfactory performance.

Of 315 E_n -scores, 272 (86%) returned a score of $|E_n| \le 1.0$, indicating agreement of the participant's result with the assigned value within their respective expanded uncertainties.

Laboratories 5, 7 and 14 returned satisfactory *z*-scores and E_n -scores for all 18 scored analytes.

• Develop the practical application of traceability and measurement uncertainty, and provide participants with information that will be useful in assessing their uncertainty estimates.

Of 391 numeric results, 383 (98%) were reported with an associated expanded measurement uncertainty. Reported expanded uncertainties were within the range of 2.8% to 52% relative.

Participants reported a wide range of procedures for estimating their measurement uncertainty.

• Evaluate the laboratories' test methods.

For TRH analysis, participants used liquid-liquid extraction (LLE) with various extraction solvents. All participants used gas chromatography (GC) coupled with flame ionisation detection (FID).

For BTEX analysis, the majority of participants used purge-and-trap GC coupled to either mass spectrometry (MS) or tandem mass spectrometry (MS/MS), though three participants used headspace GC-MS instead. Three participants also reported an additional LLE step.

For PAHs in this study, participants were given the option of either analysing $1 \times 500 \text{ mL}$ bottle or $3 \times 100 \text{ mL}$ bottles. Fifteen participants requested the $1 \times 500 \text{ mL}$ option, while four participants requested the $3 \times 100 \text{ mL}$ option. For analysis, one participant used solid phase extraction, while all other participants used LLE. A variety of extraction solvents were used. One participant used GC-FID, while all other participants used GC-MS(/MS).

• *Produce materials that can be used in method validation and as control samples.*

The test samples of this proficiency study are homogeneous and are well characterised. Surplus samples are available for purchase from NMI and can be used for quality control and method validation purposes.

1 INTRODUCTION

1.1 NMI Proficiency Testing Program

The National Measurement Institute (NMI) is responsible for Australia's national measurement infrastructure, providing a range of services including a chemical proficiency testing program.

Proficiency testing (PT) is the 'evaluation of participant performance against pre-established criteria by means of interlaboratory comparison'.¹ NMI PT studies target chemical testing in areas of high public significance such as trade, environment, law enforcement and food safety. NMI offers studies in:

- pesticide residues in fruit, vegetables and herbs, water and soil;
- petroleum hydrocarbons in water and soil;
- per- and polyfluoroalkyl substances in water, soil, biota and food;
- inorganic analytes in water, soil, filters, food and pharmaceuticals;
- controlled drug assay, drugs in wipes and clandestine laboratory; and
- allergens in food.

1.2 Study Aims

The aims of the study were to:

- compare the performances of participants and assess their accuracy in the identification and measurement of petroleum hydrocarbon pollutants in river water;
- develop the practical application of traceability and measurement uncertainty, and provide participants with information that will be useful in assessing their uncertainty estimates;
- evaluate the laboratories' test methods; and
- produce materials that can be used in method validation and as control samples.

The choice of the test method was left to the participating laboratories.

1.3 Study Conduct

The conduct of NMI proficiency tests is described in the NMI Study Protocol for Proficiency Testing.² The statistical methods used are described in the NMI Chemical Proficiency Testing Statistical Manual.³ These documents have been prepared with reference to ISO/IEC 17043,¹ and The International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories.⁴

NMI is accredited by the National Association of Testing Authorities, Australia (NATA) to ISO/IEC 17043 as a provider of proficiency testing schemes.¹ This study falls within the scope of NMI's accreditation.

2 STUDY INFORMATION

2.1 Study Timetable

The timetable of this study was:

Invitations sent	24/10/2022
Samples sent	22/11/2022
Results due	13/01/2023
Interim report	31/01/2023

An extended results reporting period was given due to sample delivery delays and to account for end-of-year shut down periods. The interim report was delayed to accommodate exceptional circumstances affecting a participant.

2.2 Participation

Nineteen laboratories registered to participate, and all participants were assigned a confidential laboratory code number for this study. All participants submitted results.

2.3 Selection of Hydrocarbon Analytes

The hydrocarbons and their concentrations in this study were typical of those encountered by environmental testing laboratories monitoring water to assess the impact of transport fuels in the environment, or the contamination from industry that entails the use of wood, petroleum or coal to generate heat and power. Investigation levels for the hydrocarbons studied are set out in the National Environmental Protection (Assessment of Site Contamination) Measure, Schedule B1 *Guideline on Investigation Levels for Soil and Groundwater.*⁵

For Sample S1, participants were requested to measure semi-volatile hydrocarbons (>C10-C40) and total recoverable hydrocarbons (TRH). For Sample S2, participants were requested to measure volatile hydrocarbons (C6-C10), and benzene, toluene, ethylbenzene and xylenes (BTEX). Participants were also provided with a list of potential poly-aromatic hydrocarbons (PAHs) spiked into Samples S3 and S4 (Table 1).

Naphthalene	Fluorene	Benz[a]anthracene	Benzo[a]pyrene
Acenaphthylene	Phenanthrene	Chrysene	Indeno[1,2,3-cd]pyrene
Acenaphthene	Fluoranthene	Benzo[b]fluoranthene	Dibenz[<i>a</i> , <i>h</i>]anthracene
Anthracene	Pyrene	Benzo[k]fluoranthene	Benzo[g,h,i]perylene

Table 1 Possible Spiked PAHs in Samples S3 and S4

2.4 Test Material Preparation

Four test samples were prepared by spiking water from a local river. Sample S1 was spiked with diesel fuel, Sample S2 was spiked with unleaded petrol and diesel fuel, and Samples S3 and S4 were spiked with differing amounts of anthracene, benzo[a]pyrene, chrysene, fluoranthene, fluorene, phenanthrene and pyrene.

The spiked values in each sample is presented in Table 2. Additional information on sample preparation is given in Appendix 1.

Table 2 Spiked Values of Test Samples

Sample	Analyte	Spiked Value (µg/L)	Uncertainty* (µg/L)
S1	TRH	2750	140

Sample	Analyte	Spiked Value (µg/L)	Uncertainty* (µg/L)
	Benzene	50.1	2.5
	Toluene	110	6
S2	Ethylbenzene	9.29	0.46
	Xylenes	69.7	3.5
	Total BTEX	239	12
	Anthracene	0.993	0.050
	Benzo[a]pyrene	3.48	0.17
	Chrysene	3.51	0.18
S 3	Fluoranthene	5.03	0.25
	Fluorene	14.0	0.7
	Phenanthrene	11.5	0.6
	Pyrene	11.9	0.6
	Anthracene	8.99	0.45
	Benzo[a]pyrene	0.701	0.035
	Chrysene	11.0	0.5
S 4	Fluoranthene	4.95	0.25
	Fluorene	3.01	0.15
	Phenanthrene	1.49	0.07
	Pyrene	6.03	0.30

* Estimated expanded uncertainty at approximately 95% confidence using a coverage factor of 2. Stability was not considered and so the expanded uncertainty is related to the concentration of the analyte at the time of spiking.

2.5 Homogeneity and Stability of Test Materials

No homogeneity or stability testing was conducted before the samples were sent. The samples were prepared, packaged and stored using a process that has been demonstrated to produce sufficiently homogeneous and stable samples in previous NMI hydrocarbons in river water PT studies. Additionally, the storage stability of petroleum hydrocarbons spiked into water samples has been previously established.⁶

Participants' results gave no reason to question the homogeneity or stability of these samples (Appendix 2).

To further assess possible instability, the results returned by participants were compared to the spiked values.

For Sample S1 TRH, the robust average was 59% of the spiked value. This is similar to values observed in previous NMI hydrocarbons in river water PT studies, and as there was a reasonable consensus between participants' results, an assigned value was set.

For Sample S2 BTEX, the robust averages were between 78% and 100% of the spiked values, providing good support for the stability of these analytes.

For Samples S3 and S4 PAHs, the robust averages of scored analytes were between 56% and 89% of the spiked values, and these were similar to values observed in previous NMI hydrocarbons in river water PT studies. Assigned values were set where there was a reasonable consensus between participants' results.

2.6 Test Material Storage, Dispatch and Receipt

The test samples were stored at approximately 4 °C prior to dispatch. Samples were dispatched on 22 November 2022.

The following items were also sent to participants:

- a covering letter which included a description of the test samples and instructions for participants; and
- a form for participants to confirm the receipt and condition of the test samples.

An Excel spreadsheet for the electronic reporting of results was emailed to participants.

2.7 Instructions to Participants

Participants were instructed as follows:

- Quantitatively analyse the samples using your routine test method.
- Do not test for volatile hydrocarbons (C6-C10) or BTEX components in Sample S1.
- Participants need not test for all listed analytes.
- For each analyte in each sample, report a single result in units of µg/L expressed as if reporting to a client, applying the limit of reporting of the method used for analysis. This is the figure that will be used in all statistical analysis in the study report.
- For each analyte in each sample, report the associated expanded uncertainty in units of $\mu g/L$ (e.g. 2000 ± 200 $\mu g/L$), if determined.
- Report results for the following:
 - Sample S1: Semi-volatile hydrocarbons (>C10-C40) and Total Recoverable Hydrocarbons (TRH). Use your laboratory's chosen quantitation range, and indicate what this range is. Australian NEPM fractions >C10-C16, >C16-C34, >C34-C40 are encouraged. The concentration range is between 200 – 10000 µg/L.
 - $\circ~$ Sample S2: Volatile Hydrocarbons (C6-C10), Benzene, Toluene, Ethylbenzene, Total Xylenes and Total BTEX. Individual BTEX components concentration is between 0.2 800 μ g/L.
 - \circ Samples S3 and S4: PAHs. The concentration range is between $0.05 50 \,\mu g/L$.
- Give details of your methodology and basis of uncertainty estimate as requested by the results sheet emailed to you.
- Return the completed results sheet by 12 December 2022 by email to proficiency@measurement.gov.au.

The results due date was extended to 13 January 2023 due to customs clearance delays to some international participants, and to then account for end-of-year shut down periods. An additional extension was given to one participant due to exceptional circumstances.

2.8 Interim Report

An interim report was emailed to all participants on 31 January 2023.

The interim report was delayed to allow the participant given the additional extension to report their results.

3 PARTICIPANT LABORATORY INFORMATION

3.1 Participants' Test Methods

Participants were requested to provide information about their test methods. Responses are presented in Appendix 4.

3.2 Basis of Participants' Measurement Uncertainty Estimates

Participants were requested to provide information about their basis of measurement uncertainty (MU). Responses are presented in Table 3.

Lab. Approach to Estimating		Information Sources	for MU Estimation*	Guide Document
Code	MU	Precision	Method Bias	MU
1	Top Down - precision and estimates of the method and laboratory bias	Control samples - SS Duplicate analysis Instrument calibration	Laboratory bias from PT studies Instrument calibration Recoveries of SS Standard purity	Eurachem/CITAC Guide
2	Top Down - precision and estimates of the method and laboratory bias	Control samples - CRM Duplicate analysis	CRM	ISO/GUM
3	Top Down - precision and estimates of the method and laboratory bias	Control samples - SS Duplicate analysis Instrument calibration	Instrument calibration Recoveries of SS Standard purity	Eurachem/CITAC Guide
4	Top Down - precision and estimates of the method and laboratory bias	Control samples - SS Duplicate analysis	Recoveries of SS	
5	Top Down - precision and estimates of the method and laboratory bias	Control samples - CRM Duplicate analysis Instrument calibration	CRM Instrument calibration Recoveries of SS	NATA General Accreditation Guidance Estimating and Reporting Measurement Uncertainty of Chemical Test Results
6	Top Down - precision and estimates of the method and laboratory bias	Control samples - CRM Duplicate analysis Instrument calibration	CRM Instrument calibration Recoveries of SS	ISO/GUM
7	Top Down - precision and estimates of the method and laboratory bias	Control samples - SS	Recoveries of SS	ISO/GUM
8	Top Down - precision and estimates of the method and laboratory bias	Control samples - SS Duplicate analysis Instrument calibration	Instrument calibration Recoveries of SS Standard purity	NMI Uncertainty Course
	Ton Down provision and	Standard deviation from PT studies only		
9	estimates of the method and laboratory bias	Control samples - SS Duplicate analysis	CRM Instrument calibration Recoveries of SS	ISO/GUM

Table 3 Basis of Uncertainty Estimate

Lab. Approach to Estimating		Information Sources	for MU Estimation*	Guide Document
Code	MU	Precision	Method Bias	MU
11	Top Down - precision and estimates of the method and laboratory bias	Control samples - CRM	CRM	Eurachem/CITAC Guide
12	Based on historical data	Duplicate analysis Instrument calibration	Instrument calibration Standard purity	Eurachem/CITAC Guide
13	Top Down - precision and estimates of the method and laboratory bias	Control samples	Recoveries of SS	ISO/GUM
14	Standard deviation of replicate analyses multiplied by 2 or 3	Control samples - SS Duplicate analysis	CRM Instrument calibration Recoveries of SS Standard purity	NATA GAG Estimating and Reporting MU
		Standard deviation from PT studies only		
15	Control Chart	Control samples - Lab Control Sample	Recoveries of SS	
16	Bottom Up (ISO/GUM, fish bone/cause and effect diagram)	Duplicate analysis Instrument calibration	Instrument calibration Recoveries of SS	Eurachem/CITAC Guide
17	Top Down - precision and estimates of the method and laboratory bias	Control samples - SS Duplicate analysis	Instrument calibration Recoveries of SS Standard purity	NMI Uncertainty Course
18	Top Down - precision and estimates of the method and laboratory bias	Control samples - CRM	CRM Instrument calibration Recoveries of SS	ISO/GUM
19	Top Down - precision and estimates of the method and laboratory bias	Control samples - RM Duplicate analysis	CRM	ISO/GUM

* CRM = Certified Reference Material; RM = Reference Material; SS = Spiked Samples

3.3 Participants' Comments

Participants were invited to make any comments or suggestions on the samples, this study, or possible future studies. Such feedback may be useful in improving future studies. Participants' comments are presented in Table 4.

Table 4 Participants' Comments

Lab. Code	Sample	Participant's Comments	Study Coordinator's Response
1	S1	Sample broke	Samples are packaged with bubble-wrap and packaging filler to minimise potential damage during transit. Participants are requested to return the sample receipt notification form once they receive the samples. Any problems with the samples should be reported as soon as possible, and NMI will replace the sample(s) free of charge if deemed necessary.
7	S2	Sample S2 received with headspace in vials.	

Lab. Code	Sample	Participant's Comments	Study Coordinator's Response
14	S 2	A very small amount of what appeared to be algae was observed in the sample. Since that material was at the bottom of the vial the sample was used without any treatment, other than subsampling from one of the vials to prepare dilutions. The second vial was analysed directly (unopened, not used for preparation of dilutions). Results reported from least dilute sample which was within the highest calibrant (50 μ g/L), which was a x 10 dilution for results > 50 μ g/L.	In the samples remaining at NMI, small brown particles can be seen in the samples, which the study coordinator believes is what the participant is referring to. These particles dissolve upon shaking the samples. The original river water matrix was also filtered and autoclaved prior to preparing these samples, and therefore it is believed that the particles observed are not algae. Participants' results were in general very close to the spiked values for this sample, and therefore these particles did not appear to have affected participants' results.

4 PRESENTATION OF RESULTS AND STATISTICAL ANALYSIS

4.1 Results Summary

Participant results are listed in Tables 5 to 29 with summary statistics: robust average, median, mean, number of numeric results (N), maximum (Max), minimum (Min), robust standard deviation (Robust SD) and robust coefficient of variation (Robust CV). Bar charts of results and performance scores are presented in Figures 2 to 24. An example chart with interpretation guide is shown in Figure 1.



Figure 1 Guide to Presentation of Results

4.2 Outliers and Gross Errors

Outliers were results less than 50% and greater than 150% of the robust average, and these were removed before the calculation of the assigned value.^{3,4} Gross errors were obvious blunders, e.g. results reported with incorrect units or basis, and such results were removed for the calculation of all summary statistics.^{3,4}

4.3 Assigned Value

The assigned value is defined as the 'value attributed to a particular property of a proficiency test item'.¹ In this PT study, the property is the concentration of the analytes in the samples. Assigned values were the robust averages of participants' results, and the expanded uncertainties were estimated from the associated robust SDs (Appendix 3).

4.4 Robust Average and Robust Between-Laboratory Coefficient of Variation

The robust averages and associated expanded MUs, and robust between-laboratory CVs (a measure of the variability of participants' results) were calculated using the procedure described in ISO 13528.⁷

4.5 Performance Coefficient of Variation

The performance coefficient of variation (PCV) is a fixed measure of the between-laboratory variation that in the judgement of the study coordinator would be expected from participants, given the analyte concentrations. The PCV is not the CV of participants' results; it is set by the study coordinator and is based on the concentration of the analytes and experience from previous studies, and is supported by mathematical models such as the Thompson-Horwitz equation.⁸ By setting a fixed and realistic value for the PCV, a participant's performance does not depend on other participants' performances and can be compared from study to study.

4.6 Target Standard Deviation for Proficiency Assessment

The target standard deviation for proficiency assessment (σ) is the product of the assigned value (*X*) and the PCV, as presented in Equation 1.

$$\sigma = X \times PCV \qquad Equation 1$$

4.7 *z-*Score

For each participant result, a *z*-score is calculated according to Equation 2.

$$z = \frac{(\chi - X)}{\sigma} \qquad Equation 2$$

where:

z is z-score

- χ is a participant's result
- *X* is the assigned value
- σ is the target standard deviation for proficiency assessment from Equation 1

For the absolute value of a *z*-score (|z|):

- $|z| \le 2.0$ is satisfactory;
- 2.0 < |z| < 3.0 is questionable; and
- $|z| \ge 3.0$ is unsatisfactory.

To account for potential low bias in consensus value due to inefficient methodologies, scores may be adjusted for a 'maximum acceptable result'. Additional information is given in Section 6.3.

4.8 *E_n*-Score

The E_n -score is complementary to the *z*-score in assessment of laboratory performance. E_n -score includes measurement uncertainty and is calculated according to Equation 3.

$$E_n = \frac{(\chi - X)}{\sqrt{U_{\chi}^2 + U_X^2}} \qquad Equation 3$$

where:

 E_n is E_n -score

- χ is a participant's result
- *X* is the assigned value
- U_{χ} is the expanded uncertainty of the participant's result
- U_X is the expanded uncertainty of the assigned value

For the absolute value of an E_n -score ($|E_n|$):

- $|E_n| \le 1.0$ is satisfactory; and
- $|E_n| > 1.0$ is unsatisfactory.

4.9 Traceability and Measurement Uncertainty

Laboratories accredited to ISO/IEC 17025 must establish and demonstrate the traceability and measurement uncertainty associated with their test results.⁹

Guidelines for quantifying uncertainty in analytical measurement are described in the Eurachem/CITAC Guide.¹⁰

5 TABLES AND FIGURES

Table 5

Sample Details

Sample No.	S1
Matrix	River Water
Analyte	>C10-C16
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	NR	NR		
2*	2000	600	7.20	1.91
3*	1300	390	2.93	1.15
4	NT	NT		
5	853	256	0.20	0.11
6*	381	115	-2.68	-2.32
7	690	210	-0.79	-0.50
8	820	205	0.00	0.00
9	NT	NT		
10	670	230	-0.91	-0.55
11	651	NR	-1.03	-1.13
12	NS	NS		
13	990	248	1.04	0.59
14	1000	500	1.10	0.34
15	635	167.8	-1.13	-0.82
16	1210	360	2.38	1.00
17	918	275	0.60	0.31
18	635	190	-1.13	-0.76
19*	212	85	-3.71	-3.53

* Outlier, see Section 4.2

Statistics

Assigned Value	820	150
Spike Value	Not Spiked	
Robust Average	830	230
Median	820	170
Mean	860	
Ν	15	
Max	2000	
Min	212	
Robust SD	350	
Robust CV	42%	











Table 6

Sample Details

Sample No.	S1
Matrix	River Water
Analyte	>C16-C34
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	NR	NR		
2*	1600	480	3.60	1.32
3	1000	300	0.38	0.21
4	NT	NT		
5	1194	358	1.42	0.67
6*	383	115	-2.94	-2.78
7	841	255	-0.48	-0.30
8	690	173	-1.29	-1.02
9	NT	NT		
10	620	309	-1.67	-0.89
11*	410	NR	-2.80	-3.25
12	NS	NS		
13	1086	239	0.84	0.54
14	1100	550	0.91	0.30
15	1038	274.3	0.58	0.34
16	820	250	-0.59	-0.37
17	800	240	-0.70	-0.45
18	1040	312	0.59	0.31
19	<500	500		

* Outlier, see Section 4.2

Statistics

Assigned Value	930	160
Spike Value	Not Spiked	
Robust Average	890	220
Median	920	170
Mean	900	
Ν	14	
Max	1600	
Min	383	
Robust SD	320	
Robust CV	37%	









Figure 3

Sample No.	S1
Matrix	River Water
Analyte	>C34-C40
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty
1	NR	NR
2	<50	NR
3	<500	NR
4	NT	NT
5	< 100	30
6	<100	NR
7	< 100	30
8	<100	NR
9	NT	NT
10	<100	63
11	<100	NR
12	NS	NS
13	<50	10
14	< 50	25
15	<400	105.7
16	<200	NR
17	<50	NR
18	<100	NR
19	<500	500

Statistics

Insufficient data to calculate statistics.

Table 8 Additional hydrocarbon ranges to those defined in the NEPM, $^{\rm 5}$ reported by participants for Sample S1

Lab. Code	Range	Result (µg/L)	Uncertainty (µg/L)
4	C10-C14	2000	540
	C15-C36	650	208

Table 9

Sample Details

Sample No.	S1
Matrix	River Water
Analyte	TRH
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1*	2650	795	2.00▼	1.00▼
2*	3700	1110	7.93	1.74
3	2300	NR	2.00▼	1.00▼
4*	400	200	-5.09	-3.42
5	2047	614	1.41	0.52
6*	764	229	-3.65	-2.35
7	1531	460	-0.63	-0.28
8	1510	378	-0.71	-0.36
9	1000	300	-2.72	-1.57
10	1290	322	-1.58	-0.88
11	1061	NR	-2.48	-1.97
12	NS	NS		
13	2076	519	1.52	0.63
14	2100	1100	1.62	0.36
15	1673	442.2	-0.07	-0.03
16	2030	NR	1.34	1.06
17	1718	NR	0.11	0.09
18	1675	502	-0.06	-0.03
19*	212	NR	-5.83	-4.62

* Outlier, see Section 4.2; ▼ Adjusted Score, see Section 6.3

Statistics

Assigned Value	1690	320
Spike Value	2750	140
Robust Average	1620	460
Max Acceptable	3300	
Result		
Median	1670	360
Mean	1650	
Ν	18	
Max	3700	
Min	212	
Robust SD	780	
Robust CV	48%	

If a participant did not report a TRH value, the TRH result was calculated by the study coordinator by summing the individual hydrocarbon ranges reported, and no estimate of the uncertainty of the TRH result was made.











Sample No.	S2
Matrix	River Water
Analyte	C6-C10
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty
1	NT	NT
2	420	96
3	200	60
4	290	87
5	368	147
6	573	171.9
7	372.8	82
8	560	170
9	NT	NT
10	320	72
11	464	NR
12	NS	NS
13	304	97
14	NT	NT
15	NT	NT
16	361	110
17	442	133
18	410	123
19	420	170

Statistics

Assigned Value	Not Set	
Spike Value	Not Spiked	
Robust Average	393	68
Median	391	60
Mean	393	
Ν	14	
Max	573	
Min	200	
Robust SD	100	
Robust CV	26%	



Table 11

Sample Details

Sample No.	S2
Matrix	River Water
Analyte	Benzene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	59.1	14.8	1.24	0.61
2	46	2.3	-0.51	-0.86
3	30	10	-2.65	-1.85
4	41	12	-1.18	-0.70
5	50.7	15.2	0.12	0.06
6	47	9.4	-0.37	-0.28
7	54.04	11.89	0.57	0.34
8	62	15.5	1.63	0.76
9	45.5	13.65	-0.58	-0.30
10	46	9.1	-0.51	-0.39
11	48.3	8.5	-0.20	-0.16
12	NS	NS		
13	52	16	0.29	0.13
14	52	13	0.29	0.16
15	49.15	15.18	-0.09	-0.04
16	45.2	13.6	-0.62	-0.33
17	64	19	1.90	0.73
18	51.1	15.3	0.17	0.08
19	50.1	20	0.04	0.01
Statistics				
			7	

Assigned Value	49.8	3.8
Spike Value	50.1	2.5
Robust Average	49.8	3.8
Median	49.6	3.2
Mean	49.6	
Ν	18	
Мах	64	
Min	30	
Robust SD	6.5	
Robust CV	13%	





Sample No.	S2
Matrix	River Water
Analyte	Toluene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	100.1	23	0.05	0.03
2	120	6	1.39	1.91
3	60	20	-2.64	-1.79
4	81	24	-1.23	-0.71
5	110	33	0.72	0.31
6	110	22	0.72	0.45
7	116.8	25.7	1.17	0.64
8	102	26	0.18	0.10
9	83	24.9	-1.09	-0.62
10	93	18.9	-0.42	-0.30
11	102.5	20	0.21	0.15
12	NS	NS		
13	105	31	0.38	0.18
14	100	25	0.05	0.03
15	99.91	35.70	0.04	0.02
16	86.5	26	-0.86	-0.47
17	148	44	3.27	1.08
18	94.1	28.2	-0.35	-0.18
19	84.9	34	-0.97	-0.41

Statistics

Assigned Value	99.3	9.0
Spike Value	110	6
Robust Average	99.3	9.0
Median	100	9
Mean	99.8	
Ν	18	
Мах	148	
Min	60	
Robust SD	15	
Robust CV	15%	





Figure 7

Sample No.	S2
Matrix	River Water
Analyte	Ethylbenzene
Unit	μg/L

17

11

7

1.2

13%

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	9.8	2.7	0.37	0.18
2	10	0.5	0.51	0.79
3	<10	NR		
4	7.0	2.1	-1.64	-1.03
5	8.7	2.6	-0.42	-0.22
6	11	2.2	1.23	0.74
7	9.61	2.83	0.23	0.11
8	10	2.5	0.51	0.27
9	7.55	2.265	-1.25	-0.73
10	9	1.7	-0.21	-0.16
11	9.58	1.9	0.21	0.14
12	NS	NS		
13	8	3	-0.93	-0.42
14	11	2.8	1.23	0.59
15	10.5	2.96	0.87	0.40
16	8.62	2.59	-0.48	-0.25
17	9.5	3	0.15	0.07
18	8.92	2.68	-0.27	-0.13
19	8.61	3.4	-0.49	-0.20
Statistics				
Assigned Value	9.29	0.74]	
Spike Value	9.29	0.46		
Robust Average	9.29	0.74		
Median	9.50	0.72		
Mean	9.26			

Ν

Max

Min

Robust SD

Robust CV



Figure 8

Sample No.	S2
Matrix	River Water
Analyte	Xylenes
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	55.9	14.5	0.04	0.02
2	59	3	0.41	0.73
3*	20	6	-4.27	-5.09
4	44	13	-1.39	-0.86
5	53.4	16	-0.26	-0.13
6	68	13.6	1.49	0.88
7	61.39	13.51	0.69	0.41
8	58	15	0.29	0.16
9	40	12	-1.87	-1.25
10	53	9.9	-0.31	-0.25
11	58	7.5	0.29	0.29
12	NS	NS		
13	51	17	-0.55	-0.26
14	64	16	1.01	0.51
15	55.37	14.74	-0.03	-0.02
16	50.2	15.1	-0.65	-0.35
17	59	18	0.41	0.19
18	53.7	16.1	-0.23	-0.12
19	55.9	22	0.04	0.01

* Outlier, see Section 4.2

Statistics

Assigned Value	55.6	3.6
Spike Value	69.7	3.5
Robust Average	54.7	4.3
Median	55.6	2.9
Mean	53.3	
Ν	18	
Max	68	
Min	20	
Robust SD	7.3	
Robust CV	13%	



Sample No.	S2
Matrix	River Water
Analyte	Total BTEX
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	224.9	67.5	0.24	0.11
2	230	15	0.40	0.63
3	NR	NR		
4	172	52	-1.38	-0.84
5	222	66.8	0.15	0.07
6	236	70.8	0.58	0.26
7	241.8	60.45	0.76	0.40
8	225	68	0.25	0.12
9	176.05	52.815	-1.26	-0.75
10	201	40	-0.49	-0.38
11	218.38	44	0.04	0.03
12	NS	NS		
13	234	70	0.52	0.24
14	230	58	0.40	0.22
15	214.93	68.58	-0.06	-0.03
16	190	57	-0.83	-0.46
17	280.5	84	1.95	0.75
18	207.8	62.3	-0.28	-0.14
19	200	80	-0.52	-0.21
Statistics				

Assigned Value	217	14
Spike Value	239	12
Robust Average	217	14
Median	222	13
Mean	218	
Ν	17	
Мах	280.5	
Min	172	
Robust SD	23	
Robust CV	11%	


Figure 10

Sample No.	S3
Matrix	River Water
Analyte	Anthracene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	<0.2	NR		
2	0.6	0.03	-1.90	-1.39
3	<1	NR		
4	0.68	0.20	-1.27	-0.61
5	1	0.30	1.27	0.46
6	<1	NR		
7	1.08	0.38	1.90	0.58
8	<1	NR		
9**	9.3	2.79	67.14	3.03
10	<1	0.36		
11	<1	NR		
12	0.77	0.28	-0.56	-0.21
13	<1	0.40		
14	0.82	0.41	-0.16	-0.05
15*	1.34	0.42	3.97	1.10
16	0.8	0.2	-0.32	-0.15
17	<1	NR		
18	1.10	0.33	2.06	0.70
19	0.7	0.28	-1.11	-0.43

* Outlier, ** Gross Error, see Section 4.2

Assigned Value	0.84	0.17
Spike Value	0.993	0.050
Robust Average	0.88	0.19
Median	0.81	0.19
Mean	0.89	
Ν	10	
Мах	1.34	
Min	0.6	
Robust SD	0.24	
Robust CV	27%	





Figure 11

Sample No.	S3
Matrix	River Water
Analyte	Benzo[<i>a</i>]pyrene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty
1	1.4	0.5
2	1.3	0.26
3	2	0.6
4	1.3	0.38
5	2.4	0.72
6	<1	NR
7	1.39	0.49
8	1.58	0.395
9	0.885	0.2655
10	1	0.38
11	1.93	0.44
12	1.4	0.68
13	2.04	0.59
14	1.9	0.95
15	1.83	0.69
16	1.2	0.4
17	1	0.3
18	1.49	0.45
19	1.18	0.47
Statistics		
Assigned Value	Not Set	
Spike Value	3.48	0.17
Robust Average	1.50	0.26
Median	1.40	0.27
Mean	1.51	
Ν	18	

2.4

0.885

0.45 30%

Мах

Min

Robust SD

Robust CV



Sample No.	S3
Matrix	River Water
Analyte	Chrysene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty		
1	1.4	0.4		
2	1.8	0.05		
3	3	1		
4	1.9	0.58		
5	2.61	0.78		
6	1.2	0.4		
7	2.27	0.79		
8	2.1	1.1		
9	0.615	0.1845		
10	1	0.35		
11	2.64	0.66		
12	1.55	0.56		
13	<1	0.27		
14	2.5	1.3		
15	2.19	0.79		
16	2.9	0.9		
17	1.2	0.4		
18	2.18	0.65		
19	1.33	0.53		
Statistics				
Assigned Value	Not Set			
Spike Value	3.51	0.18		
Robust Average	1.92	0.45		
Median	2.00	0.53		
Mean	1.91			
Ν	18			
Max	3			

0.615

0.76

40%

Min

Robust SD

Robust CV



Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Fluoranthene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	3.8	1.1	-0.09	-0.04
2	2.5	0.19	-2.34	-3.47
3	4	1	0.26	0.14
4	3.6	1.1	-0.43	-0.22
5	4.31	1.29	0.80	0.34
6	3.8	1	-0.09	-0.05
7	4.04	1.41	0.33	0.13
8	3.54	0.9	-0.54	-0.32
9	4.15	1.245	0.52	0.23
10	3.5	1.18	-0.61	-0.29
11	5.52	1.6	2.00▼	1.00▼
12	3.45	0.9	-0.69	-0.42
13	3.75	1.05	-0.17	-0.09
14	4.5	2.3	1.13	0.28
15	3.72	1.40	-0.23	-0.09
16	3.3	1	-0.95	-0.52
17	5	1.5	1.99	0.75
18	4.31	1.29	0.80	0.34
19	2.75	1.1	-1.90	-0.96

▼ Adjusted Score, see Section 6.3

Assigned Value	3.85	0.34
Spike Value	5.03	0.25
Robust Average	3.85	0.34
Max Acceptable	6.2	
Result		
Median	3.80	0.30
Mean	3.87	
Ν	19	
Max	5.52	
Min	2.5	
Robust SD	0.59	
Robust CV	15%	









Sample Details

Sample No.	S3
Matrix	River Water
Analyte	Fluorene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	10.2	3.3	1.15	0.42
2	4.3	0.43	-3.37	-3.00
3	9	3	0.23	0.09
4	5.9	1.8	-2.15	-1.23
5	10.02	3.01	1.01	0.40
6	10.2	3	1.15	0.45
7	11.28	3.95	1.98	0.62
8	8.2	2.5	-0.38	-0.17
9*	1.1	0.33	-5.82	-5.28
10	5.6	2.14	-2.38	-1.21
11	12.7	3.2	2.00▼	1.00▼
12	9.2	4.1	0.38	0.12
13	7.85	2.75	-0.65	-0.28
14	7.4	3.7	-1.00	-0.33
15	11.8	3.69	2.00▼	0.79
16	7.9	2.4	-0.61	-0.29
17	9	2.7	0.23	0.10
18	10.1	3.0	1.07	0.42
19	6.16	2.5	-1.95	-0.89

* Outlier, see Section 4.2; ▼ Adjusted Score, see Section 6.3

Assigned Value	8.7	1.4
Spike Value	14.0	0.7
Robust Average	8.5	1.6
Max Acceptable Result	16.7	
Median	9.0	1.0
Mean	8.3	
Ν	19	
Max	12.7	
Min	1.1	
Robust SD	2.7	
Robust CV	32%	











Sample No.	S3
Matrix	River Water
Analyte	Phenanthrene
Unit	µg/L

Participant Results

Lab. Code	Result	Uncertainty	Z	En
1	10	3.2	1.00	0.38
2	6.1	0.61	-1.99	-2.07
3	9	3	0.23	0.09
4	6.6	2.0	-1.61	-0.92
5	9.19	2.76	0.38	0.16
6	11.3	3.4	1.99	0.73
7	10.79	3.78	1.60	0.53
8	9.1	2.3	0.31	0.16
9	8.75	2.625	0.04	0.02
10	6.5	2.19	-1.69	-0.90
11	12.2	2.9	2.00▼	1.00▼
12	9	2.7	0.23	0.10
13	7.59	2.50	-0.85	-0.41
14	7.8	3.9	-0.69	-0.22
15	9.69	3.57	0.76	0.27
16	7.8	2.3	-0.69	-0.35
17	9.5	2.9	0.61	0.26
18	9.86	2.96	0.89	0.37
19	5.87	2.3	-2.17	-1.11

▼ Adjusted Score, see Section 6.3

Assigned Value	8.7	1.1
Spike Value	11.5	0.6
Robust Average	8.7	1.1
Max Acceptable Result	14.1	
Median	9.0	1.0
Mean	8.77	
Ν	19	
Max	12.2	
Min	5.87	
Robust SD	1.9	
Robust CV	22%	



Laboratory





Figure 16

Sample No.	S3
Matrix	River Water
Analyte	Pyrene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	8.4	2.4	-0.88	-0.51
2	6.1	1.2	-2.46	-2.64
3	10	3	0.23	0.11
4	8.9	2.7	-0.53	-0.28
5	10.75	3.23	0.74	0.33
6	10	3	0.23	0.11
7	10.45	3.66	0.54	0.21
8	9.86	3.9	0.13	0.05
9	9.95	2.985	0.19	0.09
10	9.9	4.1	0.16	0.06
11	12.8	3.6	2.16	0.86
12	9	3.7	-0.46	-0.18
13	8.99	2.25	-0.47	-0.29
14	10	5	0.23	0.07
15	9.95	3.21	0.19	0.09
16	9.1	2.7	-0.39	-0.21
17	12	3.6	1.61	0.64
18	9.86	2.96	0.13	0.06
19	6.92	2.8	-1.90	-0.96
Statistics				

Assigned Value	9.67	0.62
Spike Value	11.9	0.6
Robust Average	9.67	0.62
Median	9.90	0.68
Mean	9.63	
Ν	19	
Мах	12.8	
Min	6.1	
Robust SD	1.1	
Robust CV	11%	



Figure 17

Sample No.	S4
Matrix	River Water
Analyte	Anthracene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	5.6	1.7	-0.35	-0.17
2	4	0.2	-2.15	-2.62
3	6	2	0.10	0.04
4	5.4	1.6	-0.58	-0.29
5	6.49	1.95	0.65	0.28
6	8.3	2.5	2.00▼	0.92
7	7.27	2.54	1.53	0.52
8	5.47	1.4	-0.50	-0.28
9*	1.2	0.36	-5.31	-5.98
10	4.2	1.47	-1.93	-1.05
11	7.02	1.9	1.25	0.55
12	5.2	1.9	-0.80	-0.35
13	6.51	2.60	0.68	0.22
14	6.1	3.1	0.21	0.06
15	5.62	1.76	-0.33	-0.15
16	5.6	1.7	-0.35	-0.17
17	6.5	2	0.67	0.28
18	7.17	2.15	1.42	0.56
19	4.38	1.8	-1.73	-0.79

* Outlier, see Section 4.2; ▼ Adjusted Score, see Section 6.3

Assigned Value	5.91	0.70
Spike Value	8.99	0.45
Robust Average	5.80	0.74
Max Acceptable Result	11	
Median	5.62	0.75
Mean	5.69	
Ν	19	
Max	8.3	
Min	1.2	
Robust SD	1.3	
Robust CV	22%	









Sample No.	S4
Matrix	River Water
Analyte	Benzo[<i>a</i>]pyrene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty
1	0.34	0.12
2	0.2	0.04
3	<1	NR
4	<0.5	NR
5	< 1	0.3
6	<1	NR
7	< 1	0.3
8	<1	NR
9	0.225	0.0675
10	<0.5	0.19
11	<1	NR
12	0.3	0.15
13	<1	0.29
14	0.43	0.22
15	<1	0.3
16	0.2	0.1
17	<1	NR
18	< 1	NR
19	0.25	0.1

Assigned Value	Not Set	
Spike Value	0.701	0.035
Robust Average	0.275	0.084
Median	0.250	0.070
Mean	0.278	
Ν	7	
Max	0.43	
Min	0.2	
Robust SD	0.089	
Robust CV	32%	



Sample Details

Sample No.	S4
Matrix	River Water
Analyte	Chrysene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	Z	En
1	7	2	0.30	0.13
2	5.8	0.29	-0.90	-0.79
3	9	3	2.00▼	0.72
4	6.4	1.9	-0.30	-0.14
5	8.59	2.58	1.88	0.67
6*	3	1	-3.68	-2.49
7	6.15	2.15	-0.55	-0.23
8	6.74	2	0.04	0.02
9*	2.45	0.735	-4.23	-3.21
10*	2.7	0.9	-3.98	-2.81
11	8.69	2.2	1.98	0.81
12	6	2.2	-0.70	-0.28
13	<1	0.27		
14	7.8	3.9	1.09	0.27
15	6.88	2.49	0.18	0.07
16*	10.9	3.3	2.00▼	1.00▼
17	4.5	2.3	-2.19	-0.86
18	5.21	1.56	-1.48	-0.78
19	4.89	2	-1.80	-0.79

* Outlier, see Section 4.2; ▼ Adjusted Score, see Section 6.3

Assigned Value	6.7	1.1
Spike Value	11.0	0.5
Robust Average	6.2	1.4
Max Acceptable	13.0	
Result		
Median	6.3	1.3
Mean	6.3	
Ν	18	
Max	10.9	
Min	2.45	
Robust SD	2.4	
Robust CV	39%	











Sample Details

Sample No.	S4
Matrix	River Water
Analyte	Fluoranthene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En
1	4.5	1.3	1.55	0.62
2	2.3	0.17	-2.47	-2.98
3	3	1	-1.19	-0.60
4	3.3	0.98	-0.64	-0.33
5	3.82	1.15	0.31	0.14
6	4.2	1.3	1.00	0.40
7	3.7	1.30	0.09	0.04
8	3.05	0.8	-1.10	-0.66
9	4.4	1.32	1.37	0.54
10	3	1.01	-1.19	-0.59
11	4.52	1.3	1.59	0.64
12	3.25	0.85	-0.73	-0.42
13	4.03	1.13	0.69	0.32
14	4.4	2.2	1.37	0.33
15	3.74	1.41	0.16	0.06
16	3.1	0.9	-1.00	-0.55
17	4.3	1.3	1.19	0.48
18	3.74	1.12	0.16	0.08
19	2.68	1.1	-1.77	-0.82
Statistics				
Assigned Value	3.65	0.42]	

Assigned Value	3.65	0.42
Spike Value	4.95	0.25
Robust Average	3.65	0.42
Median	3.74	0.54
Mean	3.63	
Ν	19	
Max	4.52	
Min	2.3	
Robust SD	0.73	
Robust CV	20%	



Laboratory





Sample No.	S4
Matrix	River Water
Analyte	Fluorene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty
1	2.4	0.8
2	0.7	0.07
3	1	0.5
4	1.2	0.35
5	1.84	0.552
6	2.3	0.7
7	2.11	0.74
8	1.7	0.4
9**	7.45	2.235
10	1	0.39
11	2.12	0.53
12	1.72	0.76
13	1.59	0.56
14	1.5	0.75
15	2.35	0.74
16	1.3	0.4
17	1.7	0.5
18	2.01	0.60
19	1.19	0.48

** Gross Error, see Section 4.2

Assigned Value	Not Set	
Spike Value	3.01	0.15
Robust Average	1.66	0.33
Median	1.70	0.36
Mean	1.65	
Ν	18	
Max	2.4	
Min	0.7	
Robust SD	0.56	
Robust CV	34%	



Sample Details

Sample No.	S4
Matrix	River Water
Analyte	Phenanthrene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	Z	En
1	1.5	0.5	2.00▼	0.79
2	0.7	0.07	-2.39	-2.36
3	1	0.5	-0.55	-0.17
4	0.91	0.27	-1.10	-0.58
5	1.1	0.33	0.06	0.03
6	1.5	0.5	2.00▼	0.79
7	1.33	0.47	1.47	0.49
8	1	0.3	-0.55	-0.27
9*	1.7	0.51	2.00▼	1.00▼
10	<1	0.35		
11	1.26	0.3	1.04	0.51
12	1.07	0.33	-0.12	-0.06
13	1.02	0.34	-0.43	-0.19
14	0.96	0.48	-0.80	-0.26
15	1.23	0.45	0.86	0.30
16	0.9	0.3	-1.16	-0.57
17	1.1	0.3	0.06	0.03
18	1.26	0.38	1.04	0.42
19	0.8	0.32	-1.77	-0.82

* Outlier, see Section 4.2; ▼ Adjusted Score, see Section 6.3

Assigned Value	1.09	0.15
Spike Value	1.49	0.07
Robust Average	1.12	0.16
Max Acceptable	1.82	
Result		
Median	1.09	0.15
Mean	1.13	
Ν	18	
Max	1.7	
Min	0.7	
Robust SD	0.27	
Robust CV	24%	











Sample No.	S4
Matrix	River Water
Analyte	Pyrene
Unit	μg/L

Participant Results

Lab. Code	Result	Uncertainty	z	En	
1	5.1	1.4	0.54	0.26	
2	2.8	0.56	-2.71	-2.81	
3	5	1.5	0.40	0.18	
4	4.6	1.4	-0.17	-0.08	
5	4.94	1.48	0.31	0.14	
6	5.1	1.5	0.54	0.25	
7	4.95	1.73	0.32	0.13	
8	3.85	1	-1.23	-0.81	
9	5.4	1.62	0.96	0.41	
10	4.6	2.17	-0.17	-0.05	
11	5.52	1.6	1.13	0.49	
12	4	1.7	-1.02	-0.41	
13	5.01	1.25	0.41	0.22	
14	5.4	2.7	0.96	0.25	
15	4.67	1.51	-0.07	-0.03	
16	3.9	1.2	-1.16	-0.65	
17	5	1.5	0.40	0.18	
18	5.25	1.58	0.75	0.33	
19	3.56	1.4	-1.64	-0.80	
Statistics					
Assigned Value	4.72	0.39			

Assigned Value	4.72	0.39
Spike Value	6.03	0.30
Robust Average	4.72	0.39
Median	4.95	0.30
Mean	4.67	
Ν	19	
Max	5.52	
Min	2.8	
Robust SD	0.68	
Robust CV	15%	



Figure 24

6 DISCUSSION OF RESULTS

6.1 Assigned Value

The robust average of participants' results was used as the assigned value for all scored analytes. The robust averages and associated expanded uncertainties were calculated using the procedure described in ISO 13528.⁷ Results less than 50% and greater than 150% of the robust average were removed before the calculation of the assigned value.^{3,4} The calculation of the expanded uncertainty for robust averages is presented in Appendix 3, using Sample S2 benzene as an example.

Traceability: The consensus of participants' results is not traceable to any external reference, so although expressed in SI units, metrological traceability has not been established.

No assigned value was set for Sample S1 >C34-C40 as there were no numeric results reported. No assigned values were set for Sample S3 benzo(a)pyrene and chrysene, and Sample S4 benzo(a)pyrene and fluorene, as reported numeric results were too variable, and there were relatively low recoveries of the spiked value. Sample S2 C6-C10 range was also not scored; historically this has been due to its volatile nature and therefore data has been provided for information only, though there was reasonable consensus between participants' results for this analyte in this study.

A comparison of the assigned values (or robust average if no assigned value was set) and the spiked values is presented in Table 30. Similar ratios of assigned value to spiked value have been observed in previous NMI hydrocarbons in river water PT studies, and in this study assigned values were set if there was a reasonable consensus of participants' results.

Sample	Analyte	Assigned Value (Robust Average) (µg/L)	Spiked Value (µg/L)	Assigned Value (Robust Average) / Spiked Value (%)	
S 1	TRH	1690	2750	61	
S2	Benzene	49.8	50.1	99	
	Toluene	99.3	110	90	
	Ethylbenzene	9.29	9.29	100	
	Xylenes	55.6	69.7	80	
	Total BTEX	217	239	91	
\$3	Anthracene	0.84	0.993	85	
	Benzo[a]pyrene	(1.50)	3.48	(43)	
	Chrysene	(1.92)	3.51	(55)	
	Fluoranthene	3.85	5.03	77	
	Fluorene	8.7	14.0	62	
	Phenanthrene	8.7	11.5	76	
	Pyrene	9.67	11.9	81	
S4	Anthracene	5.91	8.99	66	
	Benzo[a]pyrene	(0.275)	0.701	(39)	
	Chrysene	6.7	11.0	61	
	Fluoranthene	3.65	4.95	74	

Table 30 Comparison of Assigned Value (or Robust Average) and Spiked Value

Sample	Analyte	Assigned Value (Robust Average) (µg/L)	Spiked Value (µg/L)	Assigned Value (<i>Robust</i> Average) / Spiked Value (%)
	Fluorene	(1.66)	3.01	(55)
	Phenanthrene	1.09	1.49	73
	Pyrene	4.72	6.03	78

6.2 Measurement Uncertainty Reported by Participants

Participants were asked to report the expanded uncertainty estimates associated with their results and the basis of this uncertainty. It is a requirement of ISO/IEC 17025 that laboratories have procedures to estimate the uncertainty of chemical measurements and to report this uncertainty in specific circumstances, including when the client's instruction so requires.⁹

Of 391 numeric results submitted for analytes of interest in this study, 383 (98%) were reported with an associated uncertainty. Participants used a wide variety of procedures to estimate their expanded MU (Table 3). Two participants reported using the NATA General Accreditation Guidance Estimating and Reporting MU as their guide; NATA no longer publishes this document.¹¹

Laboratory **11** did not report uncertainties for the hydrocarbon ranges in Samples S1 and S2. This participant reported being accredited to ISO/IEC 17025. Laboratories **3**, **11**, **16**, **17** and **19** Sample S1 TRH result had no uncertainty as the result was calculated by the study coordinator by summing the individual hydrocarbons ranges reported.

The magnitude of reported uncertainties was within the range of 2.8% to 52% relative. In general, an expanded measurement uncertainty of less than 10% relative is likely to be unrealistically small for the routine measurement of a hydrocarbon pollutant in water, while an expanded uncertainty of over 50% is likely to be too large and not fit-for-purpose. Of 383 MUs, 11 were below 10% relative, while seven were greater than 50% relative.

Uncertainties associated with results returning a satisfactory *z*-score but an unsatisfactory E_n -score may have been underestimated.

An estimate of uncertainty expressed as a value should not be attached to a non-value result.¹⁰ Laboratories **5**, **7**, **10**, **13**, **14**, **15** and **19** attached an uncertainty to at least one of their non-value results.

In some cases, results were reported with an inappropriate number of significant figures. Including too many significant figures may inaccurately reflect the measurement precision. The recommended format is to write uncertainty to no more than two significant figures and then to write the result with the corresponding number of decimal places. For example, instead of $10.02 \pm 3.01 \mu g/L$, it is better to report this as $10.0 \pm 3.0 \mu g/L$.¹⁰

6.3 *z-*Score

Target SDs equivalent to 15% and 20% PCV were used to calculate *z*-scores. CVs predicted by the Thompson-Horwitz equation,⁸ the between-laboratory CVs obtained and target SDs (as PCV) for this study are presented for comparison in Table 31.

Sample	Analyte	Assigned Value (Robust Average) (µg/L)	Thompson-Horwitz CV ^a (%)	Between-Laboratory CV ^b (%)	Target SD (as PCV) (%)
S1	>C10-C16	820	16	24	20
	>C16-C34	930	16	23	20
	TRH	1690	15	27	15
	C6-C10	(393)	18	26	Not Set
	Benzene	49.8	22	13	15
52	Toluene	99.3	22	15	15
52	Ethylbenzene	9.29	22	13	15
	Xylenes	55.6	22	11	15
	Total BTEX	217	20	11	15
	Anthracene	0.84	22	24	15
	Benzo[a]pyrene	(1.50)	22	30	Not Set
	Chrysene	(1.92)	22	40	Not Set
S 3	Fluoranthene	3.85	22	15	15
	Fluorene	8.7	22	28	15
	Phenanthrene	8.7	22	22	15
	Pyrene	9.67	22	11	15
S4	Anthracene	5.91	22	20	15
	Benzo[a]pyrene	(0.275)	22	32	Not Set
	Chrysene	6.7	22	24	15
	Fluoranthene	3.65	22	20	15
	Fluorene	(1.66)	22	34	Not Set
	Phenanthrene	1.09	22	22	15
	Pyrene	4.72	22	15	15

Table 31 Comparison of Thompson-Horwitz CV, Between-Laboratory CV and Target SD

^a Calculated from the assigned value (robust average).

^b Robust between-laboratory CV (outliers removed where applicable).

To account for possible low bias in the consensus values due to laboratories using inefficient extraction or analytical techniques, a total of 12 *z*-scores were adjusted across the following analytes: Sample S1 TRH, Sample S3 fluoranthene, fluorene and phenanthrene, and Sample S4 anthracene, chrysene and phenanthrene. For these analytes, a maximum acceptable result was set to two target SDs more than the spiked value, and results lower than the maximum acceptable result but with a *z*-score greater than 2.0 had their *z*-score adjusted to 2.0. This ensured that laboratories reporting results close to the spiked value were not penalised. *z*-Scores for results higher than the maximum acceptable result were not adjusted, and *z*-scores less than 2.0 were left unaltered.

Of 315 results for which *z*-scores were calculated, 277 (88%) returned a satisfactory score of $|z| \le 2.0$, indicating a satisfactory performance.

Laboratories 2, 5, 7, 14, 15, 16 and 18 reported results for all 18 scored analytes. Of these participants, Laboratories 5, 7 and 14 returned satisfactory *z*-scores for all scored analytes.

Satisfactory *z*-scores were achieved for all scored analytes reported by Laboratories 8 (17), 13 (16) and 1 (15).

Laboratory **12** analysed Samples S3 and S4 only (PAHs), and returned satisfactory *z*-scores for all ten scored analytes.

The dispersal of participants' *z*-scores is presented by laboratory in Figure 25 and by analyte in Figure 26.





Figure 27 presents participants' *z*-scores for Sample S1 (TRH) only. Participants with a trend of *z*-scores below the zero line possibly had an inefficient extraction process for TRH. As the ratio of the assigned value to the spiked value was 61% for TRH, participants reporting results with higher satisfactory *z*-scores may have more efficient extraction methodologies.



Figure 28 presents participants' *z*-scores for Sample S2 (BTEX) only. A trend of *z*-scores on one side of the zero line may indicate laboratory bias for BTEX analytes.



Figure 29 presents participants' *z*-scores for Samples S3 and S4 (PAHs) only. Participants with a trend of *z*-scores below the zero line may have an inefficient extraction process for PAHs. As the ratio of the assigned values to the spiked values ranged from 61% to 85%, results with higher satisfactory *z*-scores may correspond to the more efficient extraction of PAHs.

PAHs results from Laboratory **9** were very varied; Sample S3 fluorene and Sample S4 anthracene and chrysene were significantly lower than the assigned value (13% to 37% relative), while Sample S3 anthracene was significantly higher than the assigned value (1107% relative). The results reported by this participant for the other PAHs returned satisfactory *z*-scores.



Scatter plots of *z*-scores for anthracene, fluoranthene, phenanthrene and pyrene in Samples S3 and S4 are presented in Figures 30 to 33. Scores are predominantly in the upper right and lower left quadrants, indicating that laboratory bias is the major contributor to the variability of results. Points close to the diagonal axis demonstrate excellent repeatability while points close to the zero demonstrate excellent repeatability and accuracy.



Figure 32 z-Score Scatter Plot – Phenanthrene



6.4 *En*-Score

 E_n -scores can be interpreted in conjunction with *z*-scores, as an unsatisfactory E_n -score can either be caused by issues with measurement, or uncertainty, or both. If a participant did not report any uncertainty with a result, an expanded uncertainty of zero (0) was used to calculate the E_n -score. For results whose *z*-scores were adjusted as discussed in Section 6.3 *z*-Score, any E_n -scores greater than 1.0 were set to 1.0.

Of 315 results for which E_n -scores were calculated, 272 (86%) returned a satisfactory score of $|E_n| \le 1.0$, indicating agreement of the participant's result with the assigned value within their respective expanded uncertainties.

Laboratories 5, 7, 14 and 18 returned satisfactory E_n -scores for all 18 scored analytes.

Satisfactory E_n -scores were achieved for all scored analytes reported by Laboratories **13** (16) and **1** (15).
Laboratory **12** analysed Samples S3 and S4 only (PAHs), and returned satisfactory E_n -scores for all ten scored analytes.





6.5 False Negatives

Table 32 presents false negative results. These are analytes present in the samples which a participant tested for, but did not report a numeric result; for example, when participants reported a 'less than' result (< x) when the assigned value was higher than their limit of reporting (LOR), or did not report anything. For analytes where no assigned value was set, results have only been considered to be false negatives where the robust average and spiked value were significantly higher than the participants' LOR, or if no value was reported.

Lab. Code	Sample	Analyte	Assigned Value (Robust Average) (µg/L)	Spiked Value (µg/L)	Result (µg/L)
1	S3	Anthracene	0.84	0.993	<0.2
6	S3	Benzo[a]pyrene	(1.50)	3.48	<1
10	S4	Phenanthrene	1.09	1.49	<1
12	S3	Chrysene	(1.92)	3.51	<1
15	S4	Chrysene	6.7	11.0	<1

Table	32	False	Negative	s
-------	----	-------	----------	---

6.6 Reporting of Additional Analytes

One participant reported additional analytes that were not spiked into the test samples by the study coordinator. These results are presented in Table 33.

Lab. Code	Sample	Analyte	Result (µg/L)	Uncertainty (µg/L)
12	S 3	Benz[a]anthracene	2.15	0.60
13	S4	Benz[a]anthracene	8.56	2.40

Table 33 Non-Spiked Analytes Reported by Participants

6.7 Participants' Analytical Methods

TRH (Sample S1)

Five participants reported taking the whole sample for analysis, while the other participants reported sample test portions ranging from 35 mL to 460 mL. In this study there was no evident correlation between the results obtained and the reported sample volume used (Figure 35).



Figure 35 Sample S1 TRH Results vs Sample Volume

All participants reporting methodologies used liquid-liquid extraction (LLE), with dichloromethane (DCM), hexane (HEX), or a mixture of pentane (PENT) and HEX as the solvent. Four participants reported an additional clean-up step (either silica, Na₂SO₄ or filtration). All participants used gas chromatography (GC) coupled with flame ionisation detection (FID). The most common methodology employed was LLE with DCM, with no clean-up, and using GC-FID for analysis. A summary of results compared to methodology is presented in Figure 36.



Figure 36 Sample S1 TRH Results vs Methodology

BTEX (Sample S2)

Ten participants reported taking the whole sample (40 mL) for analysis, while the other participants reported sample test portions ranging from 5 mL to 20 mL. In this study there was no evident correlation between the results obtained and the reported sample volume used (Figure 37).



Figure 37 Sample S2 Total BTEX Results vs Sample Volume

For BTEX analysis, participants used either purge-and-trap (P&T) GC coupled to mass spectrometry (MS) or tandem mass spectrometry (MS/MS), or headspace (HS) GC-MS. Three participants reported LLE as part of their preparation, using either DCM or methanol (MeOH) for extraction. The most common methodology was P&T GC-MS. A summary of results compared to methodology is presented in Figure 38.



PAHs (Samples S3 and S4)

Results flagged as gross errors have been excluded from the following discussion.

For this study, participants were given the option of either analysing 1 x 500 mL bottle (15 participants) or 3 x 100 mL bottles (4 participants), depending on which suited their laboratory's method. Participants reported test portions ranging from 35 mL to 500 mL. *z*-Scores obtained as compared to the sample volume used is presented in Figure 39. Heavier PAHs such as benzo[*a*]pyrene or chrysene may be adsorbing to the container, causing higher variability for such analytes. Participants may need to shake or sonicate the sample prior to sampling, and/or rinse the sample bottle.





The majority of participants used LLE, except for one participant who used solid-phase extraction (SPE). DCM, HEX, a mixture of HEX/PENT and a mixture of DCM and ethyl acetate (EtOAc) were reported as extraction solvents. Most participants used GC-MS(/MS) for analysis, except for one participant who used GC-FID. Two laboratories reported a clean-up step (Na₂SO₄ or filtration). The most common methodology employed for PAHs was LLE with DCM, no clean-up, and using GC-MS for analysis. A summary of results compared to methodology is presented in Figure 40. As the ratio of assigned values to spiked values ranged from 61% to 85%, results with higher satisfactory *z*-scores may correspond to the more efficient extraction of PAHs.



Figure 40 Samples S3 and S4 PAHs z-Scores vs Methodology

6.8 Certified Reference Materials

Participants were requested to report whether certified standards or matrix reference materials had been used as part of the quality assurance for the analysis.

Eleven participants reported using certified standards, one participant reported using matrix reference materials, and one participant reported using both certified standards and matrix reference materials. The following were reported:

- NMI (MX015)
- AccuStandard (DRH-004S-R1-5X)
- o2si
- Sigma-Aldrich (68281, UST127, CRM47930)
- Restek (30095, 31011)
- ISO 17034 standards

These materials may or may not meet the internationally recognised definition of a CRM:

'reference material, accompanied by documentation issued by an authoritative body and providing one or more specified property values with associated uncertainties and traceabilities, using valid procedures'¹²

6.9 Summary of Participants' Performance

Summaries of participants' results and performances for scored analytes in this PT study are presented in Tables 34 and 35, and Figure 41. Table 34 Summary of Participants' Samples S1 and S2 Results for Scored Analytes*

Lab. Code	S1 >C10-C16	S1 >C16-C34	S1 TRH	S2 Benzene	S2 Toluene	S2 Ethylbenzene	S2 Xylenes	S2 Total BTEX
AV	820	930	1690	49.8	99.3	9.29	55.6	217
SV	-	-	2750	50.1	110	9.29	69.7	239
1	NR	NR	2650	59.1	100.1	9.8	55.9	224.9
2	2000	1600	3700	46	120	10	59	230
3	1300	1000	2300	30	60	<10	20	NR
4	NT	NT	400	41	81	7.0	44	172
5	853	1194	2047	50.7	110	8.7	53.4	222
6	381	383	764	47	110	11	68	236
7	690	841	1531	54.04	116.8	9.61	61.39	241.8
8	820	690	1510	62	102	10	58	225
9	NT	NT	1000	45.5	83	7.55	40	176.05
10	670	620	1290	46	93	9	53	201
11	651	410	1061	48.3	102.5	9.58	58	218.38
12	NS	NS	NS	NS	NS	NS	NS	NS
13	990	1086	2076	52	105	8	51	234
14	1000	1100	2100	52	100	11	64	230
15	635	1038	1673	49.15	99.91	10.5	55.37	214.93
16	1210	820	2030	45.2	86.5	8.62	50.2	190
17	918	800	1718	64	148	9.5	59	280.5
18	635	1040	1675	51.1	94.1	8.92	53.7	207.8
19	212	<500	212	50.1	84.9	8.61	55.9	200

* All values are in $\mu g/L$. Shaded cells are results which returned a questionable or unsatisfactory *z*-score. AV = Assigned Value, SV = Spiked Value.

Lab. Code	S3 Anthracene	S3 Fluoranthene	S3 Fluorene	S3 Phenanthrene	S3 Pyrene	S4 Anthracene	S4 Chrysene	S4 Fluoranthene	S4 Phenanthrene	S4 Pyrene
AV	0.84	3.85	8.7	8.7	9.67	5.91	6.7	3.65	1.09	4.72
SV	0.993	5.03	14.0	11.5	11.9	8.99	11.0	4.95	1.49	6.03
1	<0.2	3.8	10.2	10	8.4	5.6	7	4.5	1.5	5.1
2	0.6	2.5	4.3	6.1	6.1	4	5.8	2.3	0.7	2.8
3	<1	4	9	9	10	6	9	3	1	5
4	0.68	3.6	5.9	6.6	8.9	5.4	6.4	3.3	0.91	4.6
5	1	4.31	10.02	9.19	10.75	6.49	8.59	3.82	1.1	4.94
6	<1	3.8	10.2	11.3	10	8.3	3	4.2	1.5	5.1
7	1.08	4.04	11.28	10.79	10.45	7.27	6.15	3.7	1.33	4.95
8	<1	3.54	8.2	9.1	9.86	5.47	6.74	3.05	1	3.85
9	9.3	4.15	1.1	8.75	9.95	1.2	2.45	4.4	1.7	5.4
10	<1	3.5	5.6	6.5	9.9	4.2	2.7	3	<1	4.6
11	<1	5.52	12.7	12.2	12.8	7.02	8.69	4.52	1.26	5.52
12	0.77	3.45	9.2	9	9	5.2	6	3.25	1.07	4
13	<1	3.75	7.85	7.59	8.99	6.51	<1	4.03	1.02	5.01
14	0.82	4.5	7.4	7.8	10	6.1	7.8	4.4	0.96	5.4
15	1.34	3.72	11.8	9.69	9.95	5.62	6.88	3.74	1.23	4.67
16	0.8	3.3	7.9	7.8	9.1	5.6	10.9	3.1	0.9	3.9
17	<1	5	9	9.5	12	6.5	4.5	4.3	1.1	5
18	1.10	4.31	10.1	9.86	9.86	7.17	5.21	3.74	1.26	5.25
19	0.7	2.75	6.16	5.87	6.92	4.38	4.89	2.68	0.8	3.56

Table 35 Summary of Participants' Samples S3 and S4 Results for Scored Analytes*

* All values are in $\mu g/L$. Shaded cells are results which returned a questionable or unsatisfactory *z*-score. AV = Assigned Value, SV = Spiked Value.



Figure 41 Summary of Participants' Performance

6.10 Comparison with Previous Studies

To enable direct comparison with previous studies, the target SD (as PCV) used to calculate *z*-scores has been kept constant between PT studies.

TRH

A summary of *z*-scores and E_n -scores, presented as a percentage of the total number of scores for each study, obtained by participants for TRH in river water over the last 10 studies where this was scored (2014–2022) is presented in Figure 42. Over this period, the average proportion of satisfactory scores was 74% for *z*-scores and 64% for E_n -scores.



Total BTEX

A summary of *z*-scores and E_n -scores, presented as a percentage of the total number of scores for each study, obtained by participants for total BTEX in river water over the last 10 studies where this was scored (2014–2022) is presented in Figure 43. Over this period, the proportion of satisfactory scores has remained high, with an average proportion of 94% and 88% for *z*-scores and E_n -scores respectively. In particular, all *z*-scores for three of the last four studies were satisfactory, and in this study all E_n -scores were also satisfactory.





PAHs

NMI has included PAHs in river water as PT samples since 2015. A summary of *z*-scores and E_n -scores (presented as a percentage of the total number of scores for each study) obtained by participants for PAHs in water over the last 8 studies (2015–2022) is presented in Figure 44. Over this period, the proportion of satisfactory scores has remained fairly consistent, with an average proportion of 86% and 81% for *z*-scores and E_n -scores respectively.



A plot of the assigned value, expressed as a percentage of the spiked value, for PAHs in river water since 2015 is presented in Figure 45 for scored analytes. In this study, anthracene was spiked in both Samples S3 and S4; the recovery for Sample S3 (spiked at a lower concentration) was significantly higher than for Sample S4. The recovery for chrysene for this study was lower than for previous studies. For fluoranthene, fluorene, phenanthrene and pyrene, the recoveries were similar to the average across all studies.



Figure 45 Ratio of Assigned Value to Spiked Value for PAHs in River Water PT Studies

Measurement Uncertainties

As discussed in Section 6.2, it is a requirement of ISO/IEC 17025 that laboratories report their uncertainties. Figure 46 presents a summary of the relative uncertainties as reported by participants over the last 10 studies (2014–2022). Over this time period, 89% of participants reported that they were accredited to ISO/IEC 17025. The vast majority (96%) of numeric results were reported with uncertainties. Additionally, most results (89%) were reported with relative uncertainties between 10% and 50%; in this study a greater proportion of results had relative uncertainties within this range as compared to the previous few studies.



Figure 46 Summary of Participants' Relative Uncertainties for NMI Hydrocarbons in River Water PT Studies

7 REFERENCES

Please note that for all undated references, the latest edition of the referenced document (including any amendments) applies.

- [1] ISO/IEC 17043, Conformity assessment General requirements for proficiency testing.
- [2] NMI, 2021, *Study Protocol for Proficiency Testing*, viewed February 2022, https://www.industry.gov.au/sites/default/files/2020-10/cpt_study_protocol.pdf>
- [3] NMI, 2022, *Statistical Manual*, viewed February 2022, https://www.industry.gov.au/sites/default/files/2019-07/cpt_statistical_manual.pdf
- [4] Thompson, M., Ellison, S.L.R., & Wood, R., 2006, 'The International Harmonized Protocol For The Proficiency Testing Of Analytical Chemistry Laboratories', *Pure Appl. Chem.*, vol 78, pp 145–196.
- [5] National Environment Protection Council, National Environment Protection (Assessment of Site Contamination) Measure 1999 Amendment Measure 2013, viewed February 2022, https://www.legislation.gov.au/Details/F2013C00288>
- [6] La Greca, B., 1996, 'Storage Stability Study: Petrol BTEX Residues in Water', *ACSL Public Interest Project*, AGAL.
- [7] ISO 13528, Statistical methods for use in proficiency testing by interlaboratory comparison.
- [8] Thompson, M., 2000, 'Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing', *Analyst*, vol 125, pp 385–386.
- [9] AS ISO/IEC 17025, General requirements for the competence of testing and calibration laboratories.
- [10] Eurachem/CITAC Guide CG 4, QUAM:2012.P1, Quantifying Uncertainty in Analytical Measurement, 3rd edition, viewed February 2022, http://www.eurachem.org/images/stories/Guides/pdf/QUAM2012_P1.pdf>
- [11] NATA, 2020, *Update to Measurement Uncertainty resources*, viewed February 2022, https://nata.com.au/news/update-to-measurement-uncertainty-resources/>
- [12] JCGM 200:2012, International vocabulary of metrology Basic and general concepts and associated terms (VIM), 3rd edition.

APPENDIX 1 SAMPLE PREPARATION

A1.1 Diesel Fuel and River Water Preparation

Diesel fuel was purchased from a local retail outlet and treated to remove volatiles. Approximately 500 mL of diesel fuel was placed in a heated (80 °C) open container and sparged with nitrogen. Treatment continued until the GC-FID chromatogram indicated that essentially all the hydrocarbons eluting before C_{10} had been removed. This same treated diesel fuel has been used in previous NMI Hydrocarbon PT studies.

Water was sampled from Browns Waterhole in Sydney. The water was filtered under vacuum through an Advantec 150 mm glass fibre filter. After filtration, the water was placed in Schott bottles and autoclaved.

A1.2 Test Sample Preparation

Sample S1

A diesel spiking solution was prepared by weighing a portion of the treated diesel fuel into a 500 mL volumetric flask and making to volume with methanol. Amber glass bottles of approximately 500 mL capacity were rinsed with acetone and dried. The cleaned bottles were placed in an air-conditioned room overnight. Filtered autoclaved water (498.5 \pm 0.2 g, or 500 mL at 25 °C) was weighed into the bottles. Methanol/diesel spiking solution (1.20 mL) was added to each bottle using a Hamilton dispenser. The bottles were immediately capped and inverted to mix the solution. Each bottle was then labelled and shrink-wrapped.

Sample S2

Filtered autoclaved water (41.88 ± 0.05 g, or 42 mL at 25 °C) was weighed into Agilent vials. A composite spike solution was prepared by adding aliquots of diesel and unleaded petrol to methanol. Two of the BTEX compounds was fortified with additional laboratory solvent. The composite spiking solution was made up to volume with methanol. Composite spiking solution (1.0 mL) was added to each vial. Each vial was capped after spiking, and then labelled and shrink-wrapped.

Samples S3 and S4

The spiking solutions were prepared by dissolving each standard material in dichloromethane. Diluted spiking solutions were prepared using acetone. The autoclaved water was placed in a stainless steel container. After spiking the water was stirred using a top-driven impeller stirrer for at least 2 hours. The samples were then dispensed into 500 mL amber glass bottles which were labelled and shrink-wrapped.

Between preparation and dispatch all samples were stored in a cool room at 4 °C.

APPENDIX 2 ASSESSMENT OF HOMOGENEITY AND STABILITY

A2.1 Homogeneity

No homogeneity testing was completed for this study as the samples were prepared using a process previously demonstrated to produce sufficiently homogeneous samples. The results of this study also gave no reason to question the samples' homogeneity. Comparisons of *z*-scores to bottle number analysed by participants are presented in Figures 47 to 58 for scored analytes (gross errors have been removed). Results have only been included when the participant was sent one sample set. For Sample S2 (where participants were provided with 2 bottles per sample set) and Samples S3 and S4 PAHs (for participants requesting the 3 x 100 mL bottles option), all containers have been graphed with the grey dotted line indicating results from the one participant.





A2.2 Stability

No stability testing was conducted for this study as the samples were prepared, stored and dispatched using a process previously demonstrated to produce sufficiently stable samples. The results of this study also gave no reason to question the samples' transportation stability. Comparisons of *z*-scores to days in transit are presented in Figures 59 to 70 for scored analytes (gross errors have been removed).





APPENDIX 3 ROBUST AVERAGE AND ASSOCIATED UNCERTAINTY, *z*-SCORE AND E_{rr} -SCORE CALCULATIONS

A3.1 Robust Average and Associated Uncertainty

Robust averages were calculated using the procedure described in ISO 13528.⁷ The associated uncertainties were estimated as according to Equation 4.

$$u_{rob\ av} = 1.25 \times \frac{S_{rob\ av}}{\sqrt{p}}$$
 Equation 4

where:

Urob av	is the standard uncertainty of the robust average
$S_{rob av}$	is the standard deviation of the robust average
р	is the number of results

The expanded uncertainty $(U_{rob av})$ is the standard uncertainty multiplied by a coverage factor of 2 at approximately 95% confidence level.

A worked example is set out below in Table 36.

Table 36 Uncertainty of the Robust Average for Sample S2 Benzene

Number of results (p)	18
Robust Average	49.8 µg/L
Srob av	6.5 μg/L
$u_{rob\ av}$	1.9 μg/L
k	2
$U_{rob\ av}$	3.8 μg/L

Therefore, the robust average for benzene in Sample S2 is $49.8 \pm 3.8 \,\mu$ g/L.

A3.2 *z*-Score and *E*_n-Score Calculation

For each participant's result, a *z*-score and E_n -score are calculated according to Equations 2 and 3 respectively (Section 4).

A worked example is set out below in Table 37.

Table 37 *z*-Score and *E_n*-Score for Sample S1 >C10-C16 Result Reported by Laboratory 2

Participant Result (µg/L)	Assigned Value (µg/L)	Target Standard Deviation	z-Score	<i>E</i> _n -Score
2000 ± 600	820 ± 150	20% as PCV, or: $0.2 \times 820 = 164 \ \mu g/L$	$z = \frac{2000 - 820}{164} = 7.20$	$E_n = \frac{2000 - 820}{\sqrt{600^2 + 150^2}} = 1.91$

APPENDIX 4 PARTICIPANTS' TEST METHODS

Participants were requested to provide information about their test methods. Responses are presented in Tables 38 to 40. Some responses may be modified so that the participant cannot be identified.

Lab. Code	Sample Volume (mL)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
1	460	Liquid-Liquid	DCM	Silica	GC-FID	In house
2	500	Liquid-Liquid	DCM	None	GC-FID	USEPA 8260
3						
4						
5	35	Liquid-Liquid	DCM	None	GC-FID	USEPA 8015
6	35	Liquid-Liquid	DCM	Sodium Sulphate	GC-FID	
7	35	Liquid-Liquid	DCM		GC-FID	USEPA 8260
8	200	Liquid-Liquid	DCM	None	GC-FID	USEPA 8015B
9	40	Liquid-Liquid	Hexane	None	GC-FID	USEPA 8015D
10	500	Liquid-Liquid	DCM	Filtration	GC-FID	USEPA SW-846 Method 3510C
11	500	Liquid-Liquid	DCM	None	GC-FID	In-House method TRH C10-16 by Purge&trap
12				NS		
13	500	Liquid-Liquid	DCM	None	GC-FID	in-house method
14	498	Liquid-Liquid	DCM	None	GC-FID	In House Method >C10-C40
15	35	Liquid-Liquid	Hexane	Silica	GC-FID	USEPA 3510
16		Liquid-Liquid	Hexane/Pentane		GC-FID	In house
17	200	Liquid-Liquid	DCM	None	GC-FID	USEPA 8015B
18	35	Liquid-Liquid	DCM		GC-FID	USEPA 8015
19	80	Liquid-Liquid	DCM	None	GC-FID	USEPA 8270

 $Table \ 38 \ Methodology-Sample \ S1 \ TRH$

Lab. Code	Sample Volume (mL)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
1	5				P&T GC-MS	USEPA 8260
2	40	N/A	N/A	None	P&T GC-MS	USEPA 8260
3	10	headspace	N/A	None	Headspace GC-MS	
4	44	Liquid-Liquid	Methanol	None	P&T GCMS	USEPA 8260
5	40	Liquid-Liquid	DCM	None	P&T GC-MS	USEPA 8260
6	40	Liquid-Liquid	None	None	P&T GC-MS	In house
7	43	none	none	none	P&T GC-MS	USEPA 8260
8	40	Purge & Trap	None	None	P&T GC-MS	USEPA 8260
9	10	Headspace	None	None	Headspace GC-MS	USEPA 5021A
10	5	Purge and Trap	NA	None	P&T GC-MS	USEPA Method 5030B
11	40	Purge&trap	NA	None	GCMS	In-house method
12			NS	5		
13	20	Purge and Trap	NA	None	P&T GC-MS	in-house method
14	5	P&T GC-MS	Nil	None	P&T GC-MS	In-house method
15	43	Purge&Trap	N/A	None	P&T GC-MS	USEPA 8260
16	12				Headspace GC-MS	In house
17	40	Purge and Trap	None	None	P&T GC-MS/MS	USEPA 8260
18	5	NA	NA		P&T GC-MS	USEPA 8260
19	44	NA	N/A	None	P&T GC-MS	USEPA 8260

Table 39 Methodology – Sample S2 BTEX

Lab. Code	Sample Volume (mL)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
1	500	SPE	DCM:EtOAc		GC-MS/MS	USEPA 8270
2	100	Liquid-Liquid	DCM	None	GC-MS/MS	USEPA 8270
3						
4	50	Liquid-Liquid	DCM	None	GC-MS	USEPA 8270
5	35	Liquid-Liquid	DCM	None	GC-MS/MS	USEPA 8270
6	35	Liquid-Liquid	DCM	Sodium Sulphate	GC-QQQ	
7	35	Liquid-Liquid	DCM		GC-MS/MS	USEPA 8260
8	200	Liquid-Liquid	DCM	None	GC-MS	USEPA 8015B
9	40	Liquid-Liquid	Hexane	None	GC-MS	USEPA 8270D
10	100	Liquid-Liquid	DCM	Filtration	GC-MS	USEPA SW-846 Method 3510C
11	500	Liquid-Liquid	DCM	None	GCMS	In-house method
12	250	Liquid-Liquid	DCM	None	GC-MS	USEPA 8270
13	100	Liquid-Liquid	DCM	None	GC-MS	in-house method
14	515	Liquid-Liquid	DCM	None	GC-MS	In-house method
15	35	Liquid-Liquid	DCM	None	GC-MS/MS	USEPA 8270
16	50	Liquid-Liquid	Hexane/Pentane		GC-MS	In house
17	200	Liquid-Liquid	DCM	None	GC-MS	USEPA 8015B
18	35	Liquid-Liquid	DCM		GC-FID	USEPA 8270
19	81 & 82	Liquid-Liquid	DCM	None	GC-MS	USEPA 8270

Table 40 Methodology – Samples S3 and S4 PAHs

APPENDIX 5 ACRONYMS AND ABBREVIATIONS

AV	Assigned Value
BTEX	Benzene, Toluene, Ethylbenzene, Xylenes
CITAC	Cooperation on International Traceability in Analytical Chemistry
CRM	Certified Reference Material
CV	Coefficient of Variation
DCM	Dichloromethane
EtOAc	Ethyl Acetate
FID	Flame Ionisation Detection
GAG	General Accreditation Guidance (NATA)
GC	Gas Chromatography
GUM	Guide to the expression of Uncertainty in Measurement
HEX	Hexane
HS	Headspace
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
LLE	Liquid-Liquid Extraction
LOR	Limit Of Reporting
Max	Maximum
Md	Median
MeOH	Methanol
Min	Minimum
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MU	Measurement Uncertainty
Ν	Number of numeric results
NATA	National Association of Testing Authorities, Australia
NEPM	National Environmental Protection Measure
NMI	National Measurement Institute, Australia
NR	Not Reported
NS	Not Supplied
NT	Not Tested
P&T	Purge and Trap
РАН	Polycyclic Aromatic Hydrocarbon
PCV	Performance Coefficient of Variation

PENT	Pentane
PT	Proficiency Testing
RA	Robust Average
RM	Reference Material
SD	Standard Deviation
SI	International System of Units
SPE	Solid Phase Extraction
SS	Spiked Samples
SV	Spiked Value (or formulated concentration of a PT sample)
TRH	Total Recoverable Hydrocarbons
USEPA	United States Environmental Protection Agency

END OF REPORT