Australian Government
Department of Industry, Science and Resources

National Measurement Institute

Proficiency Test Final Report AQA 24-04 Hydrocarbons in Soil

July 2024

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

Jasmine Duong

Luminita Antin

Luke Baker

I would also like to thank Nathan Camilleri from SGS Australia for reviewing this report.

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

Accredited for compliance with ISO/IEC 17043

TABLE OF CONTENTS

THIS PAGE IS INTENTIONALLY BLANK

SUMMARY

AQA 24-04 Hydrocarbons in Soil commenced in February 2024. Twenty-two laboratories enrolled to participate, and all participants submitted results.

Three test samples were prepared at the NMI laboratory in Sydney using soil from various sources. Participants were asked to report Total Recoverable Hydrocarbons (TRH) in Sample S1, benzene, toluene, ethylbenzene and xylenes (BTEX) and volatile hydrocarbons (C6 to C10) in Sample S2, and polycyclic aromatic hydrocarbons (PAHs) in Sample S3. The assigned values were the robust averages of participants' results for all scored analytes. 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 hydrocarbon pollutants in soil.*

Laboratories **1**, **2**, **5**, **6**, **7**, **8**, **15**, **16**, **19**, **21** and **22** reported numeric results for all 18 scored analytes.

Laboratories **3** and **13** did not report numeric results for analytes that they tested for and were present in the test samples (total of five results).

Of 359 z-scores, 327 (91%) returned $|z| \leq 2.0$, indicating an acceptable performance.

Of 356 E_n-scores, 301 (85%) returned $|E_n| < 1.0$, indicating agreement of the participant's result with the assigned value within their respective uncertainties.

Laboratories **6**, **7** and **16** returned acceptable z-scores and E_n-scores for all 18 scored analytes.

- *Evaluate participants' methods for the measurement of hydrocarbon pollutants in soil.*

For TRH, the most common methodology used was solid-liquid extraction (SLE) with dichloromethane/acetone, and using GC-FID for analysis. For BTEX, the most common methodology used was SLE with methanol, and using purge and trap GC-MS for analysis. For PAHs, the most common methodology used was SLE with dichloromethane/acetone, and using GC-MS for analysis.

There was no significant trend observed between methodology used and results obtained.

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

Of 396 numeric results, 375 results (95%) were reported with an associated expanded measurement uncertainty. The magnitude of the reported expanded uncertainties was within the range 6.4% to 58% of the reported result.

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

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

NMI also has certified reference material MX015 Hydrocarbon-Contaminated Soil available for purchase.

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 comparisons'.[1] NMI PT studies target chemical testing in areas of high public significance such as trade, environment, law enforcement and food safety. NMI offers PT studies in:

- pesticide residues in fruit, vegetables and herbs, soil and water;
- petroleum hydrocarbons in soil and water;
- per- and polyfluoroalkyl substances in soil, water, food and biota;
- inorganic analytes in soil, water, 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 hydrocarbon pollutants in soil;
- evaluate participants' methods for the measurement of hydrocarbon pollutants in soil;
- develop the practical application of traceability and measurement uncertainty, and provide participants with information that will be useful in assessing their uncertainty estimates; 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 and The International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories.[1,4]

NMI is accredited by the National Association of Testing Authorities, Australia (NATA) to ISO/IEC 17043:2023 as a provider of proficiency testing schemes.[1] This study is within the scope of NMI's accreditation.

2 STUDY INFORMATION

2.1 Study Timetable

The timetable of the study was:

Invitations sent	19/02/2024
Samples sent	18/03/2024
Results due	22/04/2024
Interim Report	30/04/2024
Preliminary Report	3/05/2024

2.2 Participation and Laboratory Code

Twenty-two laboratories enrolled to participate, and all participants were assigned a confidential laboratory code number for this study. All participants submitted results.

2.3 Selection of Analytes

The hydrocarbons in this study, and their spiked levels, were typical of those encountered by environmental testing laboratories. Investigation levels for the hydrocarbons studied are set out in the National Environment Protection (Assessment of Site Contamination) Measure (NEPM) Schedule B1 *Guideline on Investigation Levels for Soil and Groundwater*.[5]

Sample S1 assessed total recoverable hydrocarbons (TRH), Sample S2 assessed volatile hydrocarbons, and benzene, toluene, ethylbenzene and xylenes (BTEX), and Sample S3 assessed polycyclic aromatic hydrocarbons (PAHs). A list of potential PAHs for Sample S3 is presented in Table 1; this list was also provided to participants.

Table 1 List of Possible PAHs for Samples S3 and S4

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

2.4 Test Material Preparation

Two soils were used as the starting materials in this study: soil from a local suburban garden for Sample S1, and topsoil purchased from a local supplier for Samples S2 and S3.

Sample S1 (TRH) was prepared by spiking the soil with treated diesel fuel and commercially purchased hydraulic oil.

Sample S2 (BTEX) was prepared by spiking the soil with unleaded petrol, treated diesel fuel and benzene.

Sample S3 (PAHs) was prepared by spiking the soil with various PAHs.

The spiked values are presented in Table 2.

Table 2 Spiked Values of Samples

Sample	Analyte	Spiked Value (mg/kg)	Uncertainty (mg/kg)[a]
S1[b]	>C10-C16	1080	50
	>C16-C34	1700	80

Sample	Analyte	Spiked Value (mg/kg)	Uncertainty (mg/kg)[a]
	>C34-C40	239	12
	TRH	3020	150
S2[b]	Benzene	224	11
	Toluene	723	36
	Ethylbenzene	138	7
	Xylenes	617	31
	Total BTEX	1700	90
S3	Acenaphthene	3.19	0.16
	Acenaphthylene	2.78	0.14
	Anthracene	2.01	0.10
	Benz[*a*]anthracene	0.788	0.039
	Benzo[*a*]pyrene	1.50	0.07
	Chrysene	0.596	0.030
	Fluoranthene	0.403	0.020
	Fluorene	0.495	0.025
	Phenanthrene	0.400	0.020
	Pyrene	0.996	0.050

[a] 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 at the time of spiking.

[b] Samples S1 and S2 were spiked with diesel and petrol. Spiked values of individual analytes have been estimated based on analysis of the diesel and petrol.

Further information on the preparation of the samples is given in Appendix 1.

2.5 Homogeneity and Stability of Test Materials

No homogeneity or stability testing was conducted for this PT study. The samples were prepared, packaged, stored and dispatched using a process that has been demonstrated to produce homogeneous and stable samples in previous NMI Hydrocarbons in Soil PT studies. The storage stability of petroleum hydrocarbons in soil has also been previously established.[6]

Participants' results did not give reason to question the homogeneity or transport stability of the samples (Appendix 2). To further assess possible instability, participants' results were compared to the spiked values (Section 6.1). For TRH, assigned values were within 87% to 102% of the spiked value, providing good support for its stability. Assigned values for scored BTEX and PAHs were within the ranges of 51% to 64% and 48% to 92% of the spiked values respectively, which is similar to ratios observed in previous NMI Hydrocarbons in Soil PT studies, and an assigned value was set for analytes in this study if there was a reasonable consensus of participants' results.

2.6 Sample Storage, Dispatch and Receipt

Prior to dispatch, Samples S1 and S3 were stored in a refrigerator at approximately 4 °C, and Sample S2 was stored in a freezer at approximately -20 °C. The samples were packaged in insulated polystyrene foam boxes with cooler bricks and dispatched by courier on 18 March 2024.

The following items were also sent to participants:

- a 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.
- Report results on as received basis in units of mg/kg for the following:
    - Sample S1: Semi-volatile hydrocarbons (>C10-C40) and TRH. Use your laboratory's chosen quantitation range, and indicate what this range is. Results will be assessed using Australian NEPM fractions >C10-C16, >C16-C34, >C34-C40 and TRH. The concentration range is between 1000 – 20000 mg/kg.
    - Sample S2: Volatile Hydrocarbons (C6-C10), Benzene, Toluene, Ethylbenzene, Xylenes and Total BTEX. Individual BTEX components concentration is between 50 – 5000 mg/kg.
    - Sample S3: PAHs. The concentration range is between 0.05 – 50 mg/kg.
- Report results as you would report to a client, i.e. corrected for recovery or not according to your standard procedure, and applying the limit of reporting of the method used for analysis (no limit of reporting has been set for this study).
- For each analyte, report the associated expanded uncertainty in units of mg/kg (e.g. 2000 ± 200 mg/kg).
- Report any listed analyte not tested as NT as the result.
- Report the basis of your uncertainty estimates as requested in the results sheet (e.g. uncertainty budget, repeatability precision, long term result variability).
- Please complete the method details as requested in the results sheet.
- Return the completed results sheet by email (proficiency@measurement.gov.au).
- Please return results by 15 April 2024. Late results may not be included in the study report.

The results due date was extended to 22 April 2024 for all participants, accounting for the Easter public holiday period.

2.8 Interim Report and Preliminary Report

An Interim Report was emailed to all participants on 30 April 2024.

A Preliminary Report was emailed to all participants on 3 May 2024. This report included a summary of the results reported by laboratories, assigned values, performance coefficient of variations, *z*-scores and E_n-scores for each analyte in this study. No data from the Preliminary Report has been changed in the present Final Report.

3 PARTICIPANT LABORATORY INFORMATION

3.1 Test Methods Reported by Participants

Participants were requested to provide information about their test methods. Responses received 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 received are presented in Table 3. Some responses may be modified so that the participant cannot be identified.

Table 3 Basis of Expanded Uncertainty Estimate

Lab. Code	Analyte	Approach to Estimating MU	Information Sources for MU Estimation*		Guide Document for Estimating MU
			Precision	Method Bias	
1	All	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples - SS		NATA Technical Note 33
2	All	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - CRM		ISO/GUM
3	All	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples - CRM Duplicate analysis	CRM Recoveries of SS	ISO/GUM
4	TRH/PAHs	Coverage factor not reported			
5	All	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - SS	Recoveries of SS	Eurachem/CITAC Guide
6	All	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples	Recoveries of SS	ISO/GUM
7	All	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples Duplicate analysis Instrument calibration	Laboratory bias from PT studies CRM Instrument calibration Recoveries of SS	Eurachem/CITAC Guide

Lab. Code	Analyte	Approach to Estimating MU	Information Sources for MU Estimation*		Guide Document for Estimating MU
			Precision	Method Bias	
8	All	Top Down - reproducibility (standard deviation) from PT studies used directly Coverage factor not reported	Duplicate analysis	CRM Instrument calibration	
9	All	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - RM Duplicate analysis Instrument calibration	Laboratory bias from PT studies CRM Instrument calibration	Eurachem/CITAC Guide In house Macro MU Calculation Pack based on QC Data
10	TRH/PAHs	Standard deviation of replicate analyses multiplied by 2 or 3 k = 2	Control samples - SS Duplicate analysis Instrument calibration	Instrument calibration Recoveries of SS Standard purity	Eurachem/CITAC Guide
11	All	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples - CRM	CRM Instrument calibration Recoveries of SS	
12	All	Standard uncertainty based on historical data Coverage factor not reported	Duplicate analysis Instrument calibration	CRM Instrument calibration Standard purity	Eurachem/CITAC Guide
13	All	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples - SS Duplicate analysis	Instrument calibration Recoveries of SS	Eurachem/CITAC Guide
14	All	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples - CRM Duplicate analysis Instrument calibration	CRM Instrument calibration	Eurachem/CITAC Guide
15	All	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - SS		NATA Technical Note 33
16	All	Top Down - precision and estimates of the method and laboratory bias k = 2	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

Lab. Code	Analyte	Approach to Estimating MU	Information Sources for MU Estimation*		Guide Document for Estimating MU
			Precision	Method Bias	
17	All	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - SS		ISO/GUM
18	All	Bottom Up (ISO/GUM, fish bone/cause and effect diagram) Coverage factor not reported	Control samples Duplicate analysis Instrument calibration	Laboratory bias from PT studies CRM Instrument calibration Recoveries of SS	Eurachem/CITAC Guide
19	All	Coverage factor not reported			
20	All	Top Down - precision and estimates of the method and laboratory bias Coverage factor not reported	Control samples - CRM	CRM Recoveries of SS	Eurachem/CITAC Guide
21	All	Top Down - precision and estimates of the method and laboratory bias	Control samples Duplicate analysis Instrument calibration	Instrument calibration Recoveries of SS Standard purity	Eurachem/CITAC Guide
	TRH/PAHs	k = 2			
	BTEX	Coverage factor not reported			
22	All	Top Down - precision and estimates of the method and laboratory bias k = 2	Control samples Duplicate analysis	CRM Recoveries of SS	ISO/GUM

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

3.3 Participants' Comments

Participants were invited to comment on this study or future studies. Such feedback may be useful in improving future studies. Participants' comments received are presented in Table 4. Some comments may be modified so that the participant cannot be identified.

Table 4 Participants' Comments

Lab. Code	Sample	Participant's Comments	Study Coordinator's Response
18	All	Date analysed should be by test, and sometimes get analysed several times by different operators	Thank you for your feedback, we will look into adding additional date analysed fields in future results sheets.

4 PRESENTATION OF RESULTS AND STATISTICAL ANALYSIS

4.1 Results Summary

Participant results are listed in Tables 5 to 25 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), as well as other estimates of analyte mass fraction. Bar charts of results and performance scores are presented in Figures 2 to 21. An example chart with interpretation guide is shown in Figure 1.

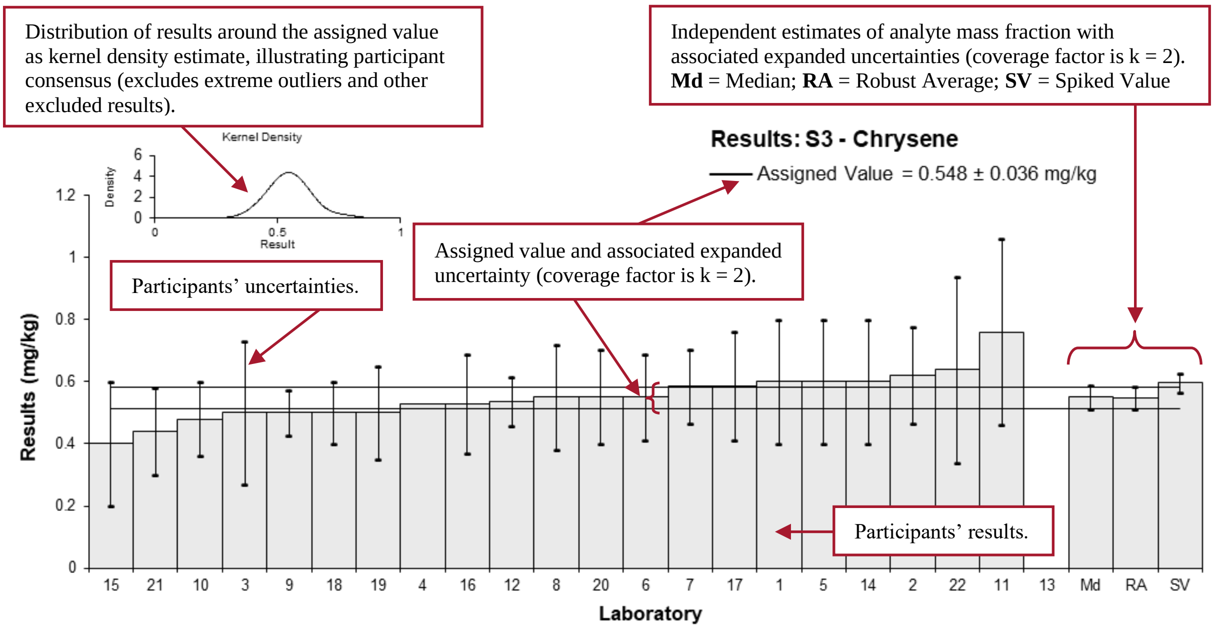


Figure 1 Guide to Presentation of Results

4.2 Outliers, Extreme Outliers and Other Excluded Results

Outliers were any result less than 50% and greater than 150% of the robust average, and these were removed before the calculation of the assigned value.[3,4] Extreme outliers were any obvious blunders, e.g. results with incorrect units, or for a different analyte or sample, and such results were removed before the calculation of all summary statistics.[3]

All results reported by Laboratory **21** for Sample S1, and Laboratory **19** for Sample S2, were consistently higher and lower respectively than the robust average by approximately the same factor for all analytes. This is an indication of laboratory or method bias. To avoid unfair scoring, these results were excluded from the calculation of all summary statistics, including the robust average calculations, as these would otherwise bias the assigned values.

4.3 Assigned Value

The assigned value is defined as the 'value attributed to a particular property or characteristic of a proficiency test item'.[1] In this PT study, the property is the mass fraction 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 CVs (a measure of the variability of 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 levels of analytes present. The PCV is not the CV of participants' results; it is set by the study coordinator and is based on the mass fraction 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' performance 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.

$$\sigma = X \times PCV \qquad \textit{Equation 1}$$

4.7 z-Score

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

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

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

To account for potential low bias in consensus value due to inefficient methodologies, scores may be adjusted for a 'maximum acceptable result' (see also Section 6.3).

For the absolute value of a *z*-score:

- $|z| \leq 2.0$ is acceptable;
- $2.0 < |z| < 3.0$ is questionable; and
- $|z| \geq 3.0$ is unacceptable.

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.

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

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:

- $|E_n| < 1.0$ is acceptable; and
- $|E_n| \geq 1.0$ is unacceptable.

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	Soil
Analyte	>C10-C16
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	840	300	-0.69	-0.31
2	692.39	173.1	-1.74	-1.28
3	1190	357	1.80	0.69
4	1138	NR	1.43	2.45
5	830	200	-0.76	-0.50
6	923.7	277	-0.09	-0.05
7	1028	298	0.65	0.29
8	1070.3	321.1	0.95	0.40
9	760	95	-1.26	-1.41
10	730	210	-1.47	-0.92
11	919	300	-0.13	-0.06
12	NR	NR		
13	1100	330	1.16	0.48
14	930	142	-0.05	-0.04
15	910	200	-0.19	-0.12
16	992	298	0.39	0.18
17	828.5	207.1	-0.77	-0.49
18	880	241	-0.41	-0.22
19	1010	303	0.52	0.23
20	960	214	0.16	0.10
21**	2671	750	12.34	2.30
22	1014	405	0.55	0.19

** Excluded Result, see Section 4.2

Statistics

Assigned Value	937	82
Spike Value	1080	50
Robust Average	937	82
Median	927	76
Mean	937	
N	20	
Max	1190	
Min	692.39	
Robust SD	150	
Robust CV	16%	

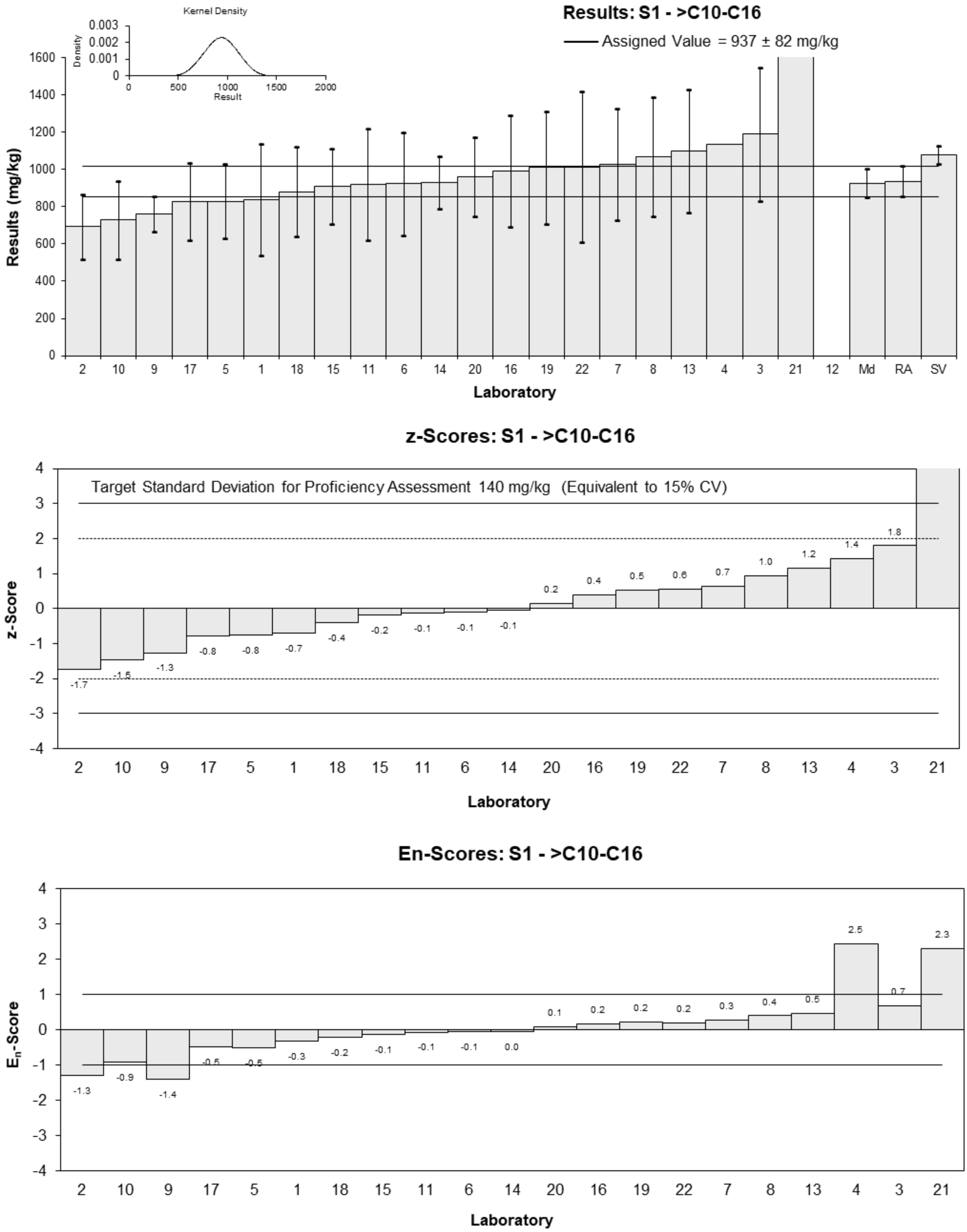

Results: S1 - >C10-C16
Assigned Value = 937 ± 82 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
2 10 9 17 5 1 18 15 11 6 14 20 16 19 22 7 8 13 4 3 21 12 Md RA SV
z-Scores: S1 - >C10-C16
Target Standard Deviation for Proficiency Assessment 140 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
En-Scores: S1 - >C10-C16
E_n-Score
Laboratory

Figure 2

Table 6

Sample Details

Sample No.	S1
Matrix	Soil
Analyte	>C16-C34
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	1500	400	-0.89	-0.55
2	1596	399	-0.52	-0.32
3	1900	570	0.66	0.29
4	2037	NR	1.18	2.36
5	1900	300	0.66	0.52
6	1601.8	481	-0.49	-0.26
7	2187	634	1.76	0.71
8	1644.7	493.4	-0.33	-0.17
9	1740	257	0.04	0.03
10	1400	330	-1.27	-0.93
11	1640	500	-0.35	-0.17
12	NR	NR		
13	2100	630	1.43	0.58
14	1860	312	0.50	0.38
15	1800	400	0.27	0.17
16	1765	530	0.13	0.06
17	1377	344.3	-1.36	-0.96
18	1590	414	-0.54	-0.32
19	1750	525	0.08	0.04
20	1640	212	-0.35	-0.36
21**	4301	1250	9.91	2.05
22	1766	706	0.14	0.05

** Excluded Result, see Section 4.2

Statistics

Assigned Value	1730	130
Spike Value	1700	80
Robust Average	1730	130
Median	1750	120
Mean	1740	
N	20	
Max	2187	
Min	1377	
Robust SD	230	
Robust CV	13%	

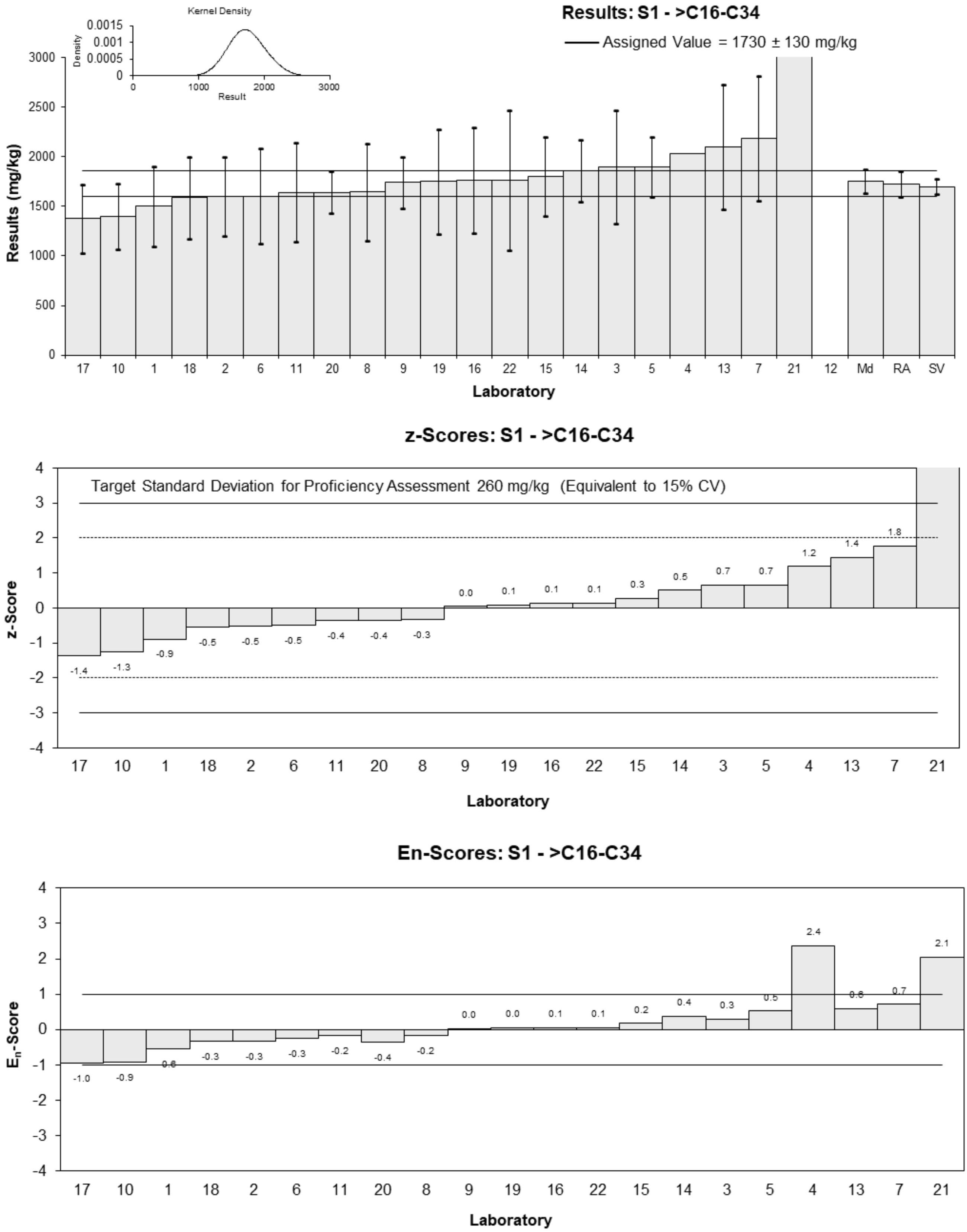

Results: S1 - >C16-C34
Assigned Value = 1730 ± 130 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
z-Scores: S1 - >C16-C34
Target Standard Deviation for Proficiency Assessment 260 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
En-Scores: S1 - >C16-C34
En-Score
Laboratory

Figure 3

Table 7

Sample Details

Sample No.	S1
Matrix	Soil
Analyte	>C34-C40
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	300	100	1.60	0.56
2	243	60.8	0.03	0.01
3*	400	120	4.35	1.28
4	NR	NR		
5	300	100	1.60	0.56
6	263.8	79	0.60	0.26
7	200	50	-1.16	-0.73
8	242.8	72.8	0.02	0.01
9	220	48	-0.61	-0.39
10	170	41	-1.98	-1.43
11	281	90	1.07	0.41
12	NR	NR		
13	270	81	0.77	0.33
14	270	104	0.77	0.26
15	320	100	2.15	0.75
16	246	74	0.11	0.05
17	198	49.5	-1.21	-0.77
18	210	65	-0.88	-0.45
19	170	51	-1.98	-1.23
20	240	51	-0.06	-0.03
21**	707	210	12.81	2.19
22	208	120	-0.94	-0.28

* Outlier, ** Excluded Result, see Section 4.2

Statistics

Assigned Value	242	29
Spike Value	239	12
Robust Average	246	31
Median	243	30
Mean	250	
N	19	
Max	400	
Min	170	
Robust SD	53	
Robust CV	22%	

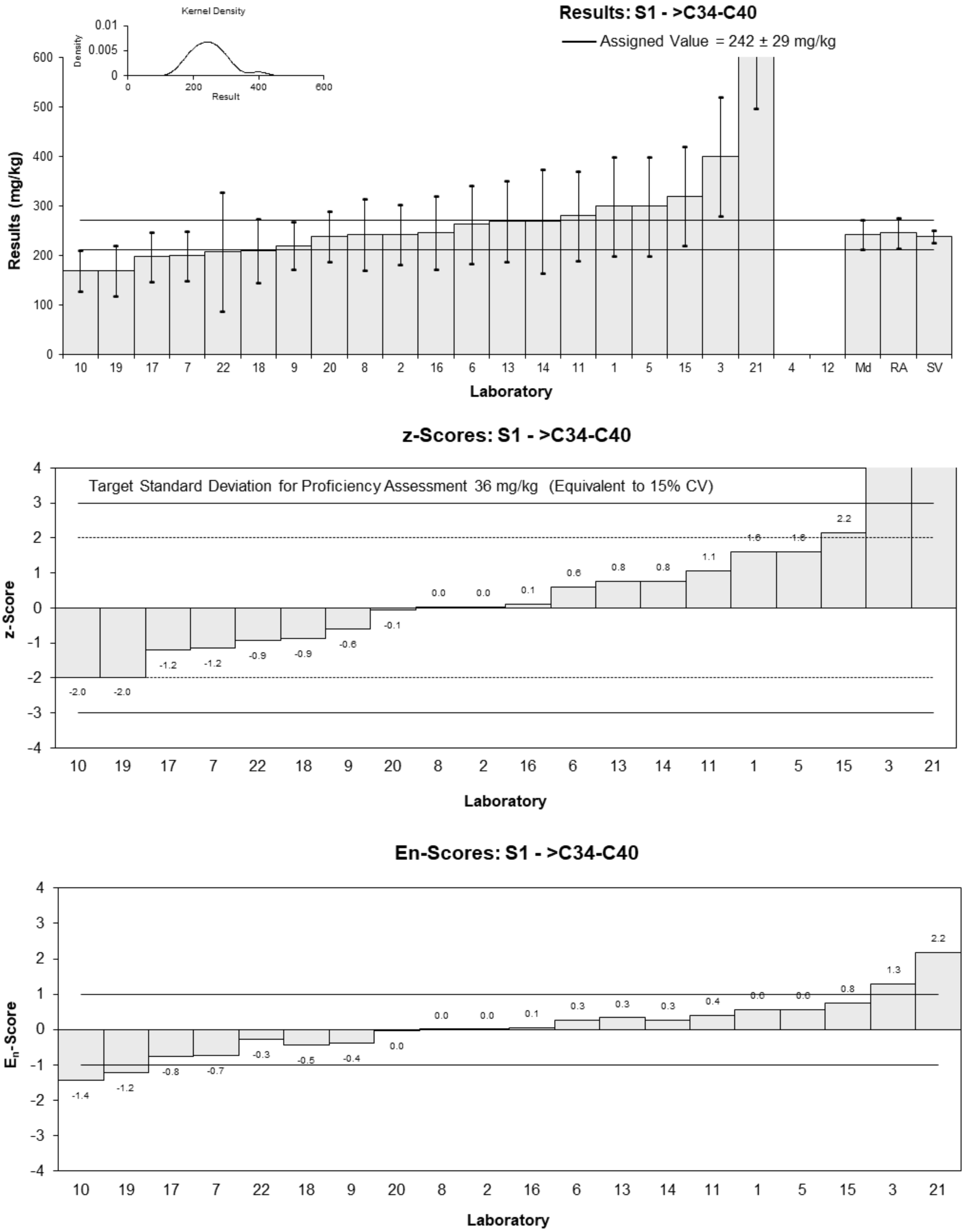

Results: S1 - >C34-C40
Assigned Value = 242 ± 29 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
10 19 17 7 22 18 9 20 8 2 16 6 13 14 11 1 5 15 3 21 4 12 Md RA SV
z-Scores: S1 - >C34-C40
Target Standard Deviation for Proficiency Assessment 36 mg/kg (Equivalent to 15% CV)
z-Score
-2.0 -2.0 -1.2 -1.2 -0.9 -0.9 -0.6 -0.1 0.0 0.0 0.1 0.6 0.8 0.8 1.1 1.6 1.6 2.2
Laboratory
10 19 17 7 22 18 9 20 8 2 16 6 13 14 11 1 5 15 3 21
En-Scores: S1 - >C34-C40
En-Score
-1.4 -1.2 -0.8 -0.7 -0.3 -0.5 -0.4 0.0 0.0 0.0 0.1 0.3 0.3 0.3 0.4 0.6 0.6 0.8 1.3 2.2
Laboratory
10 19 17 7 22 18 9 20 8 2 16 6 13 14 11 1 5 15 3 21

Figure 4

Table 8 Non-NEPM Hydrocarbon Ranges Reported by Participants for Sample S1[5]

Lab. Code	Range	Result (mg/kg)	Uncertainty (mg/kg)
4	>C34-C38	139	NR
12	C7-C9	<20	6.7
	C10-C14	273	60
	C15-C36	2360	340

THIS PAGE IS INTENTIONALLY BLANK

Table 9

Sample Details

Sample No.	S1
Matrix	Soil
Analyte	TRH
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	2640	NR	-0.62	-1.35
2	2531.39	633	-0.87	-0.57
3	3490	1047	1.33	0.54
4	3314	NR	0.93	2.02
5	3030	NR	0.27	0.60
6	2789	837	-0.28	-0.14
7	3415	854	1.16	0.58
8	2957	887	0.11	0.05
9	2720	NR	-0.44	-0.95
10	2300	390	-1.40	-1.39
11	2804	900	-0.24	-0.11
12	2640	350	-0.62	-0.67
13	3500	1100	1.35	0.53
14	3060	513	0.34	0.27
15	3030	NR	0.27	0.60
16	3003	901	0.21	0.10
17	2404	601	-1.16	-0.80
18	2680	NR	-0.53	-1.15
19	2930	880	0.05	0.02
20	2840	628	-0.16	-0.11
21**	7680	2380	10.93	2.00
22	2988	NR	0.18	0.39

** Excluded Result, see Section 4.2

Statistics

Assigned Value	2910	200
Spike Value	3020	150
Robust Average	2910	200
Median	2930	170
Mean	2910	
N	21	
Max	3500	
Min	2300	
Robust SD	360	
Robust CV	12%	

Laboratories **1**, **5**, **15**, **18** and **22** did not report a TRH value. The study coordinator summed the individual hydrocarbon ranges, and no estimate of the uncertainty of the TRH result was made.

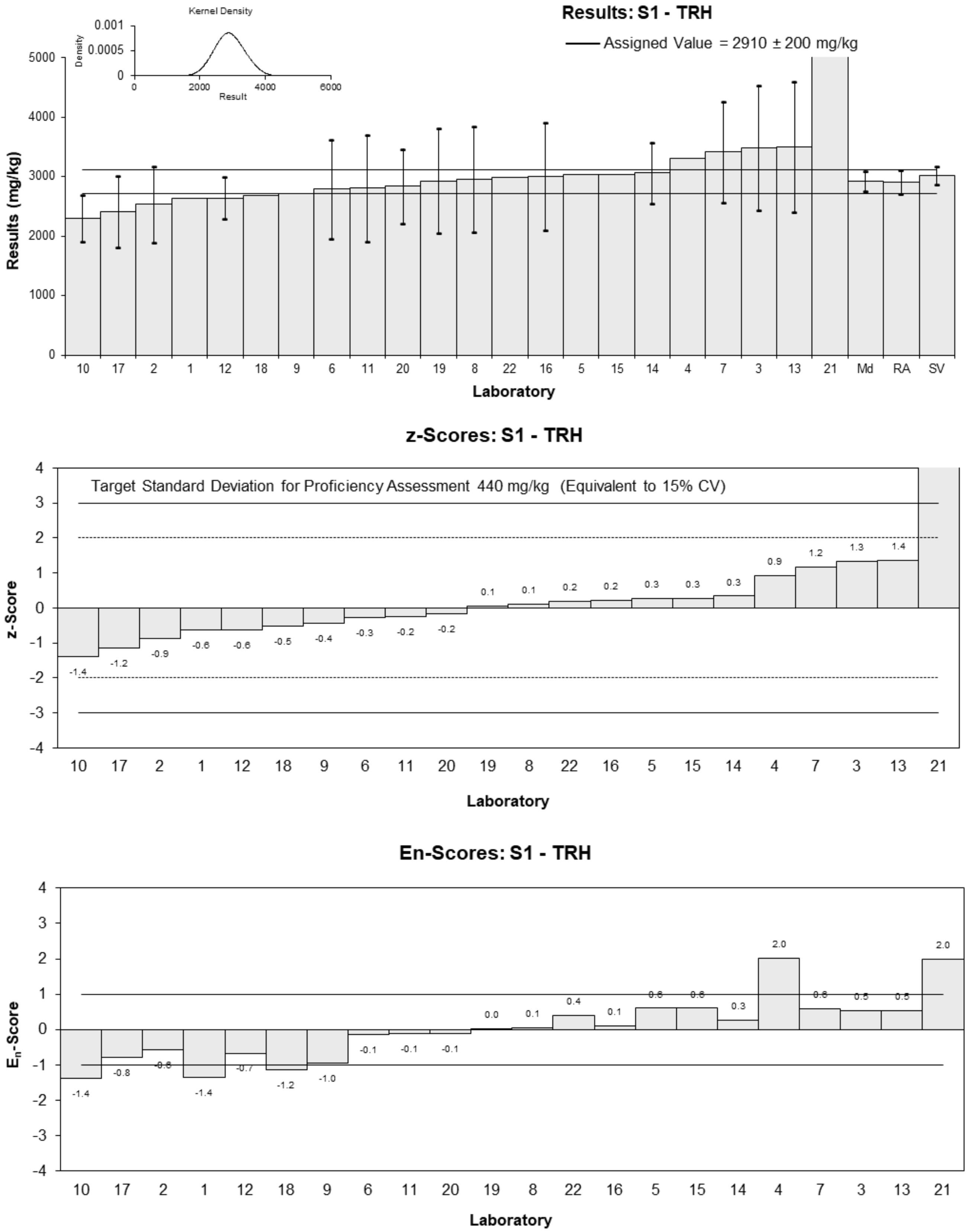

Kernel Density
Density
Result
Results: S1 - TRH
Assigned Value = 2910 ± 200 mg/kg
Results (mg/kg)
Laboratory
10 17 2 1 12 18 9 6 11 20 19 8 22 16 5 15 14 4 7 3 13 21 Md RA SV
z-Scores: S1 - TRH
Target Standard Deviation for Proficiency Assessment 440 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
10 17 2 1 12 18 9 6 11 20 19 8 22 16 5 15 14 4 7 3 13 21
En-Scores: S1 - TRH
E_n-Score
Laboratory
10 17 2 1 12 18 9 6 11 20 19 8 22 16 5 15 14 4 7 3 13 21

Figure 5

Table 10

Sample Details

Sample No.	S2
Matrix	Soil
Analyte	C6-C10
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty
1	2900	900
2	NT	NT
3**	22645	7925
4	NT	NT
5	3800	800
6	2450	638
7	2935	587
8**	4.65	1.4
9	3500	855
10	NR	NR
11	2620	800
12	NT	NT
13	1300	390
14	2530	572
15	3500	800
16	2667	800
17	3366	1009
18	3030	830.7
19**	1400	420
20	2390	559
21	NR	NR
22	1863	745

** Extreme Outlier or Excluded Result, see Section 4.2

Statistics

Assigned Value	Not Set	
Spike Value	Not Spiked	
Robust Average	2810	460
Median	2780	360
Mean	2780	
N	14	
Max	3800	
Min	1300	
Robust SD	680	
Robust CV	24%	

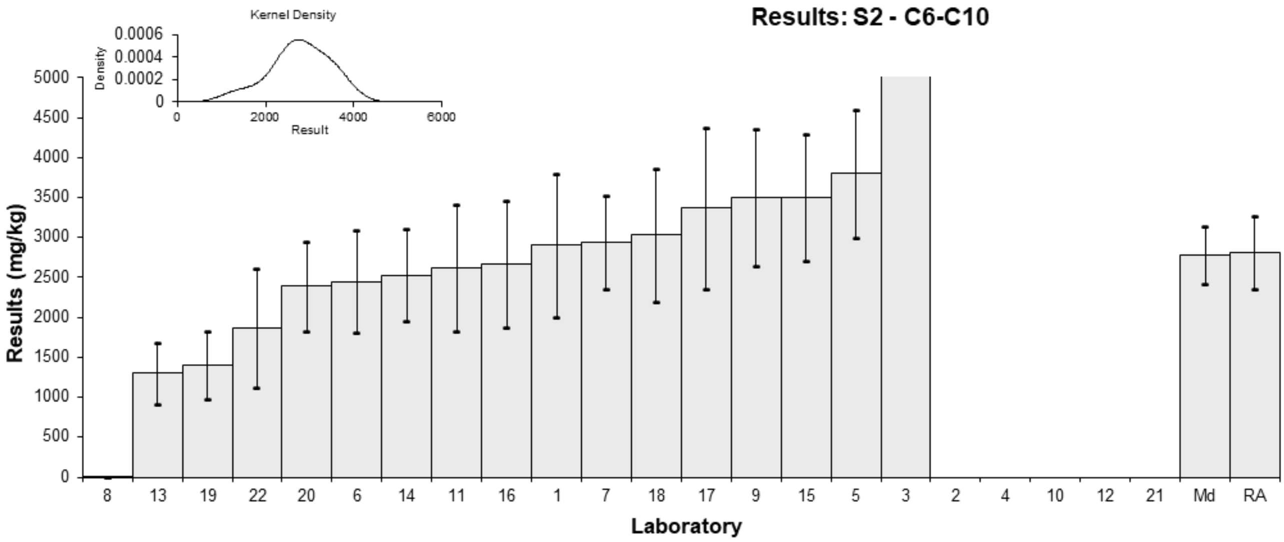

Kernel Density
0.0006
0.0004
0.0002
0
Density
0
2000
4000
6000
Result
Results: S2 - C6-C10
Results (mg/kg)
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
8
13
19
22
20
6
14
11
16
1
7
18
17
9
15
5
3
2
4
10
12
21
Md
RA
Laboratory

Figure 6

Table 11

Sample Details

Sample No.	S2
Matrix	Soil
Analyte	Benzene
Unit	mg/kg

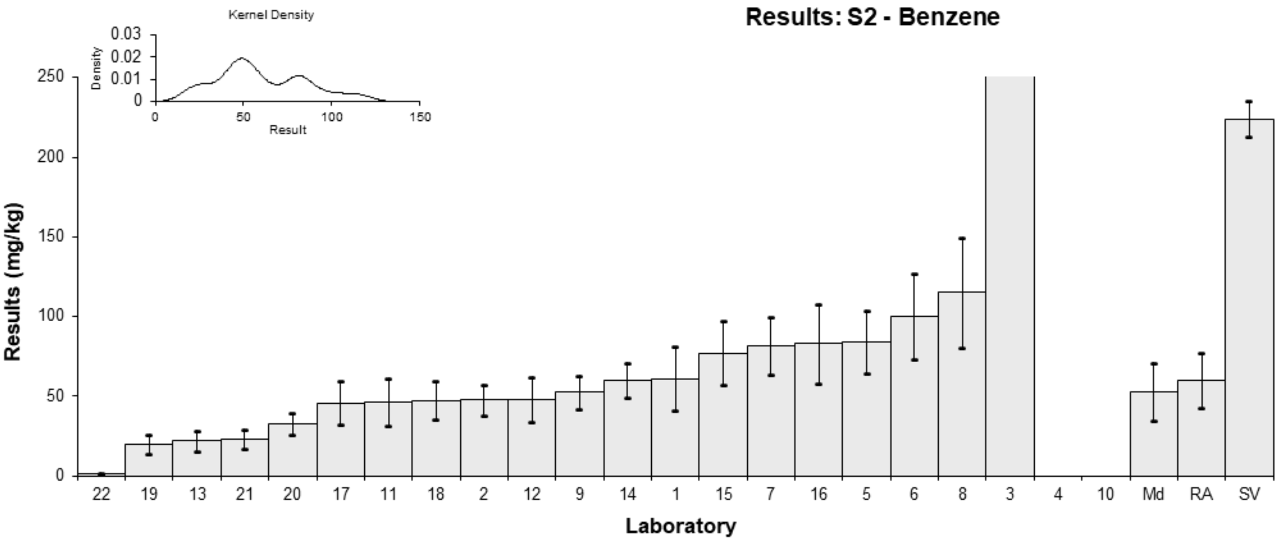
Participant Results

Lab. Code	Result	Uncertainty
1	61	20
2	47.71	9.55
3**	392	137
4	NT	NT
5	84	20
6	99.9	26.9
7	81.4	17.908
8	115	34.5
9	52.6	10.5
10	NR	NR
11	46.5	15
12	48	14
13	22	6.6
14	60.0	11
15	77	20
16	82.9	25
17	45.53	13.66
18	47.2	12.1
19**	20	6
20	32.4	6.7
21	23	6.4
22**	1.53	0.6

** Extreme Outlier or Excluded Result, see Section 4.2

Statistics

Assigned Value	Not Set	
Spike Value	224	11
Robust Average	60	17
Median	53	18
Mean	60	
N	17	
Max	115	
Min	22	
Robust SD	28	
Robust CV	47%	



Kernel Density
Density
0.03
0.02
0.01
0
0
50
100
150
Result
Results: S2 - Benzene
Results (mg/kg)
250
200
150
100
50
0
22
19
13
21
20
17
11
18
2
12
9
14
1
15
7
16
5
6
8
3
4
10
Md
RA
SV
Laboratory

Figure 7

Table 12

Sample Details

Sample No.	S2
Matrix	Soil
Analyte	Toluene
Unit	mg/kg

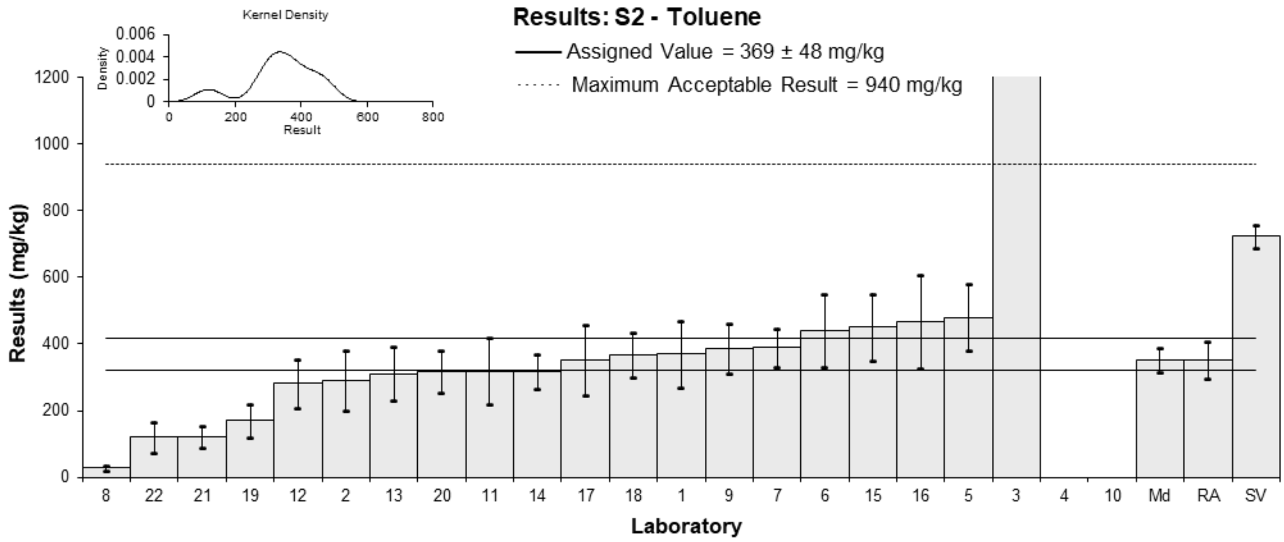
Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	370	100	0.02	0.01
2	290.95	91.59	-1.41	-0.75
3**	2687	940	41.88	2.46
4	NT	NT		
5	480	100	2.00▼	
6	439	109	1.26	0.59
7	389	58.35	0.36	0.26
8**	28	8.4	-6.16	-7.00
9	386	74	0.31	0.19
10	NR	NR		
11	318	100	-0.92	-0.46
12	281	74	-1.59	-1.00
13	310	81	-1.07	-0.63
14	318	53	-0.92	-0.71
15	450	100	1.46	0.73
16	468	140	1.79	0.67
17	351.1	105.3	-0.32	-0.15
18	367	66.8	-0.04	-0.02
19**	170	51	-3.60	-2.84
20	317	63	-0.94	-0.66
21*	119.9	32.4	-4.50	-4.30
22*	119.5	47	-4.51	-3.71

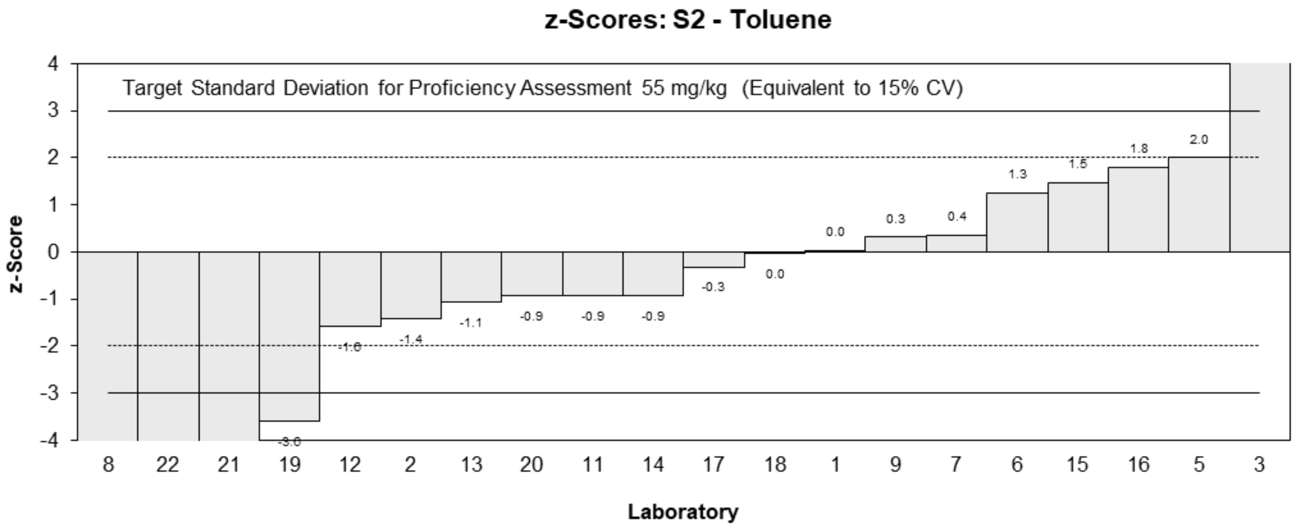
* Outlier, ** Extreme Outlier or Excluded Result, see Section 4.2; ▼ Adjusted Score, see Section 6.3

Statistics

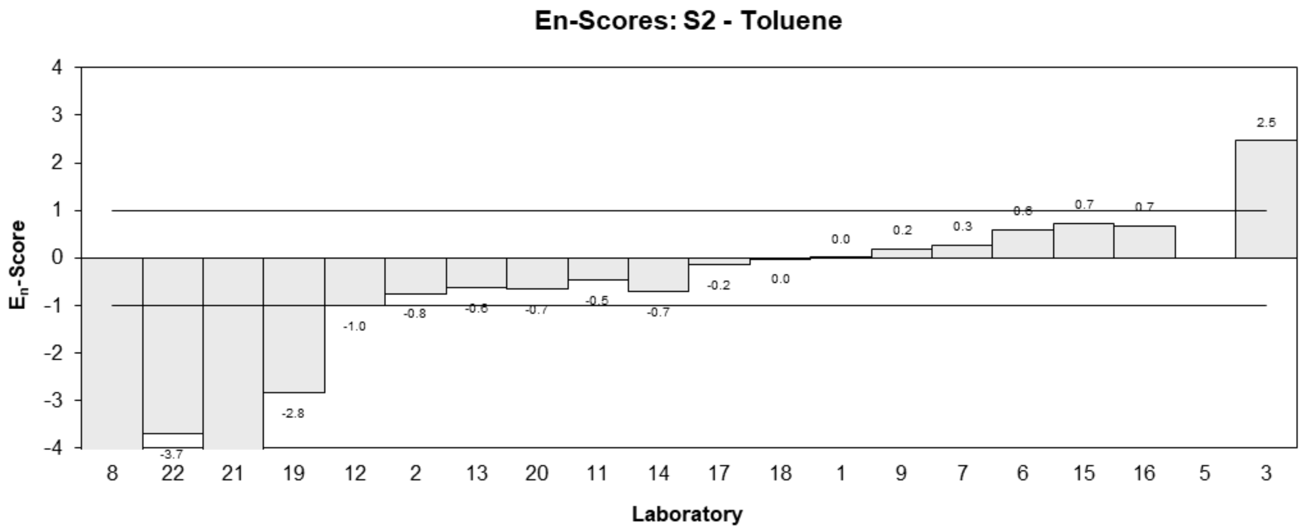
Assigned Value	369	48
Spike Value	723	36
Robust Average	351	55
Max Acceptable Result	940	
Median	351	37
Mean	340	
N	17	
Max	480	
Min	119.5	
Robust SD	90	
Robust CV	26%	



Results: S2 - Toluene
Assigned Value = 369 ± 48 mg/kg
Maximum Acceptable Result = 940 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
8 22 21 19 12 2 13 20 11 14 17 18 1 9 7 6 15 16 5 3 4 10 Md RA SV



z-Scores: S2 - Toluene
Target Standard Deviation for Proficiency Assessment 55 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
-3.6 -1.6 -1.4 -1.1 -0.9 -0.9 -0.9 -0.3 0.0 0.0 0.3 0.4 1.3 1.5 1.8 2.0
8 22 21 19 12 2 13 20 11 14 17 18 1 9 7 6 15 16 5 3



En-Scores: S2 - Toluene
En-Score
Laboratory
-3.7 -2.8 -1.0 -0.8 -0.6 -0.7 -0.5 -0.7 -0.2 0.0 0.0 0.2 0.3 0.6 0.7 0.7 2.5
8 22 21 19 12 2 13 20 11 14 17 18 1 9 7 6 15 16 5 3

Figure 8

Table 13

Sample Details

Sample No.	S2
Matrix	Soil
Analyte	Ethylbenzene
Unit	mg/kg

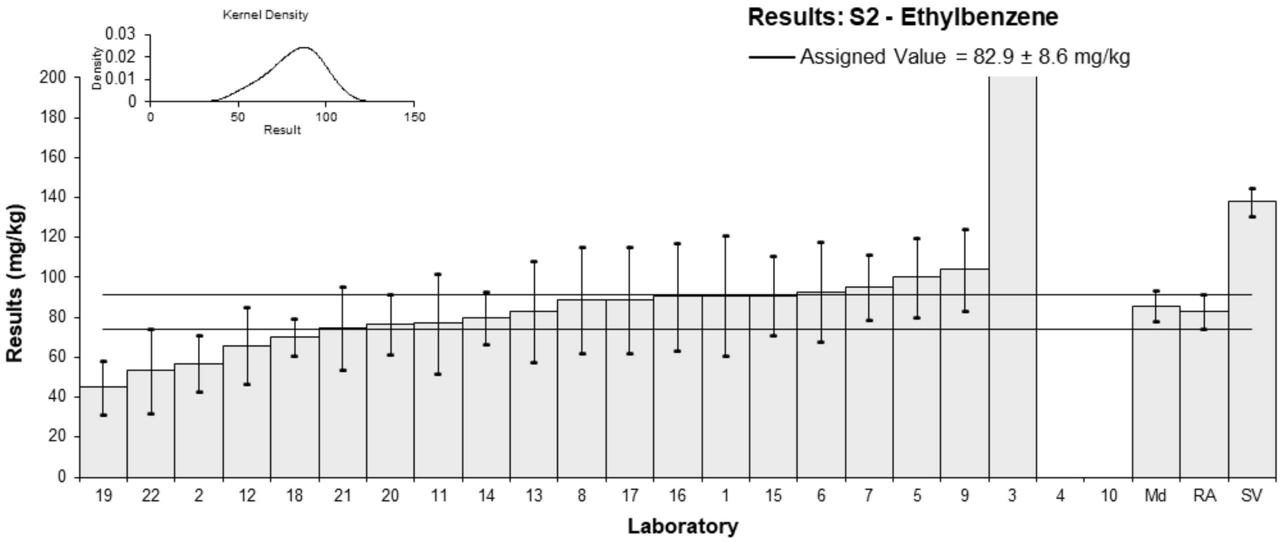
Participant Results

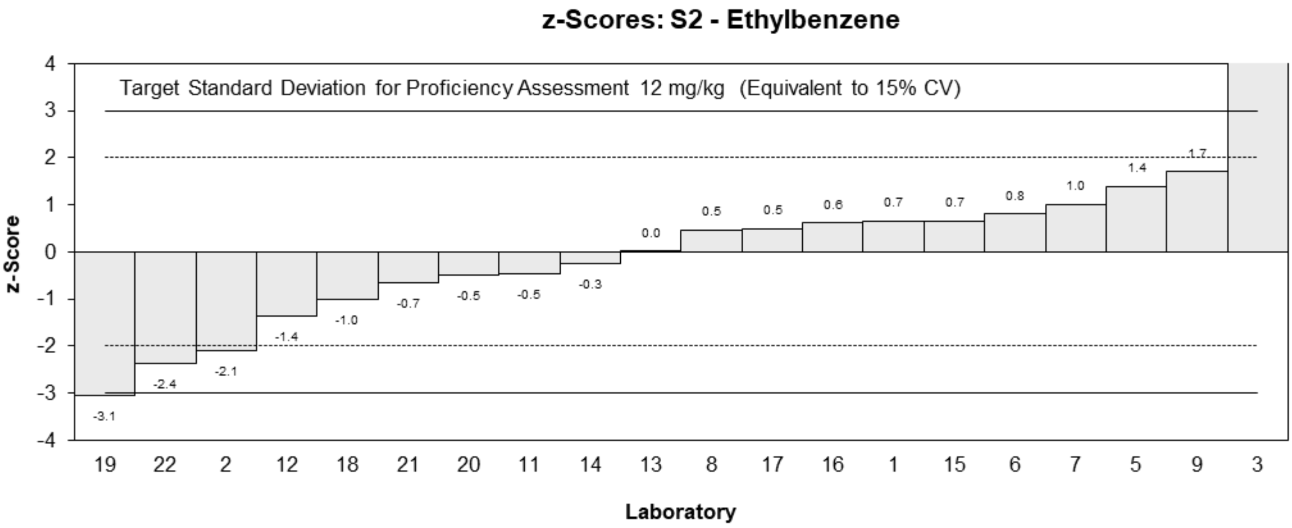
Lab. Code	Result	Uncertainty	z	E_n
1	91	30	0.65	0.26
2	56.93	14.23	-2.09	-1.56
3**	726	254	51.72	2.53
4	NT	NT		
5	100	20	1.38	0.79
6	92.8	25.1	0.80	0.37
7	95.3	16.201	1.00	0.68
8	88.5	26.6	0.45	0.20
9	104	20.5	1.70	0.95
10	NR	NR		
11	77.2	25	-0.46	-0.22
12	66	19	-1.36	-0.81
13	83	25	0.01	0.00
14	79.7	13.1	-0.26	-0.20
15	91	20	0.65	0.37
16	90.5	27	0.61	0.27
17	88.95	26.69	0.49	0.22
18	70.2	9.4	-1.02	-1.00
19**	45	13.5	-3.05	-2.37
20	76.7	15	-0.50	-0.36
21	74.6	20.9	-0.67	-0.37
22	53.3	21	-2.38	-1.30

** Extreme Outlier or Excluded Result, see Section 4.2

Statistics

Assigned Value	82.9	8.6
Spike Value	138	7
Robust Average	82.9	8.6
Median	85.8	7.7
Mean	82.2	
N	18	
Max	104	
Min	53.3	
Robust SD	15	
Robust CV	18%	


Kernel Density
Results: S2 - Ethylbenzene
Assigned Value = 82.9 ± 8.6 mg/kg
Results (mg/kg)
Laboratory
19 22 2 12 18 21 20 11 14 13 8 17 16 1 15 6 7 5 9 3 4 10 Md RA SV


z-Scores: S2 - Ethylbenzene
Target Standard Deviation for Proficiency Assessment 12 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
19 22 2 12 18 21 20 11 14 13 8 17 16 1 15 6 7 5 9 3

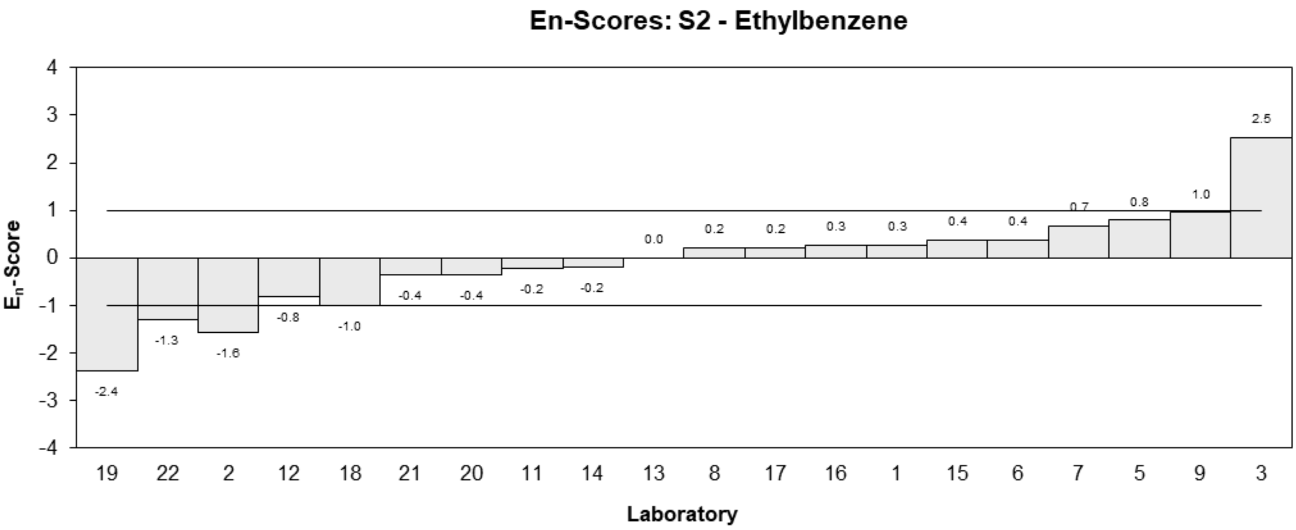

En-Scores: S2 - Ethylbenzene
En-Score
Laboratory
19 22 2 12 18 21 20 11 14 13 8 17 16 1 15 6 7 5 9 3

Figure 9

Table 14

Sample Details

Sample No.	S2
Matrix	Soil
Analyte	Xylenes
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	400	100	0.08	0.05
2*	187.85	98.90	-3.50	-2.02
3**	3255	1139	48.27	2.51
4	NT	NT		
5	450	100	0.93	0.53
6	410	98	0.25	0.15
7	414	66.24	0.32	0.27
8**	39	11.7	-6.01	-12.10
9	456	87	1.03	0.67
10	NR	NR		
11	378	120	-0.29	-0.14
12	296	92	-1.67	-1.03
13	370	110	-0.42	-0.22
14	354	62	-0.69	-0.61
15	420	100	0.42	0.24
16	417	125	0.37	0.17
17	408.4	122.5	0.23	0.11
18	436	87.2	0.69	0.45
19**	182	55	-3.59	-3.48
20	364	83.7	-0.52	-0.35
21	388	97	-0.12	-0.07
22	230	92	-2.78	-1.72

* Outlier, ** Extreme Outlier or Excluded Result, see Section 4.2

Statistics

Assigned Value	395	27
Spike Value	617	31
Robust Average	387	33
Median	400	27
Mean	375	
N	17	
Max	456	
Min	187.85	
Robust SD	54	
Robust CV	14%	

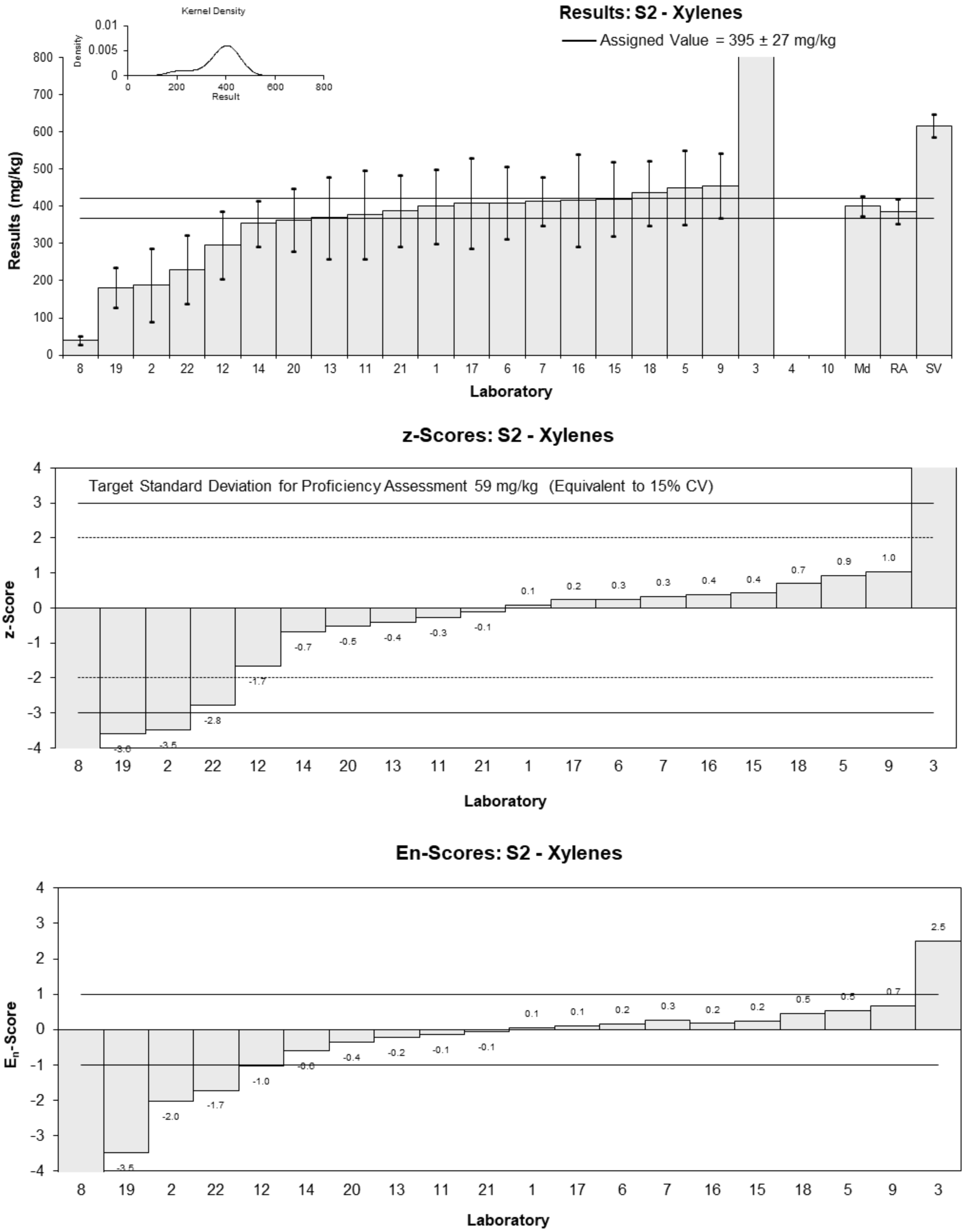

Results: S2 - Xylenes
Assigned Value = 395 ± 27 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
8 19 2 22 12 14 20 13 11 21 1 17 6 7 16 15 18 5 9 3 4 10 Md RA SV
z-Scores: S2 - Xylenes
Target Standard Deviation for Proficiency Assessment 59 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
8 19 2 22 12 14 20 13 11 21 1 17 6 7 16 15 18 5 9 3
En-Scores: S2 - Xylenes
E_n-Score
Laboratory
8 19 2 22 12 14 20 13 11 21 1 17 6 7 16 15 18 5 9 3

Figure 10

Table 15

Sample Details

Sample No.	S2
Matrix	Soil
Analyte	Total BTEX
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	920	300	0.38	0.16
2	583.44	214.27	-2.20	-1.19
3**	7060	2471	47.43	2.50
4	NT	NT		
5	1100	300	1.76	0.72
6	1042	260	1.32	0.61
7	979.7	196	0.84	0.49
8*	275.15	82.5	-4.56	-4.33
9	999	NR	0.99	1.17
10	NR	NR		
11	819	250	-0.39	-0.19
12	691	199	-1.37	-0.79
13	750	230	-0.92	-0.47
14	812	143	-0.44	-0.32
15	1000	200	1.00	0.57
16	1058	318	1.44	0.56
17	893.9	268.2	0.18	0.08
18	920	184	0.38	0.23
19**	417	125	-3.47	-2.72
20	790	181	-0.61	-0.38
21	606	182	-2.02	-1.24
22*	404	161	-3.57	-2.39

* Outlier, ** Extreme Outlier or Excluded Result, see Section 4.2

Statistics

Assigned Value	870	110
Spike Value	1700	90
Robust Average	830	130
Median	860	130
Mean	810	
N	18	
Max	1100	
Min	275.15	
Robust SD	220	
Robust CV	26%	

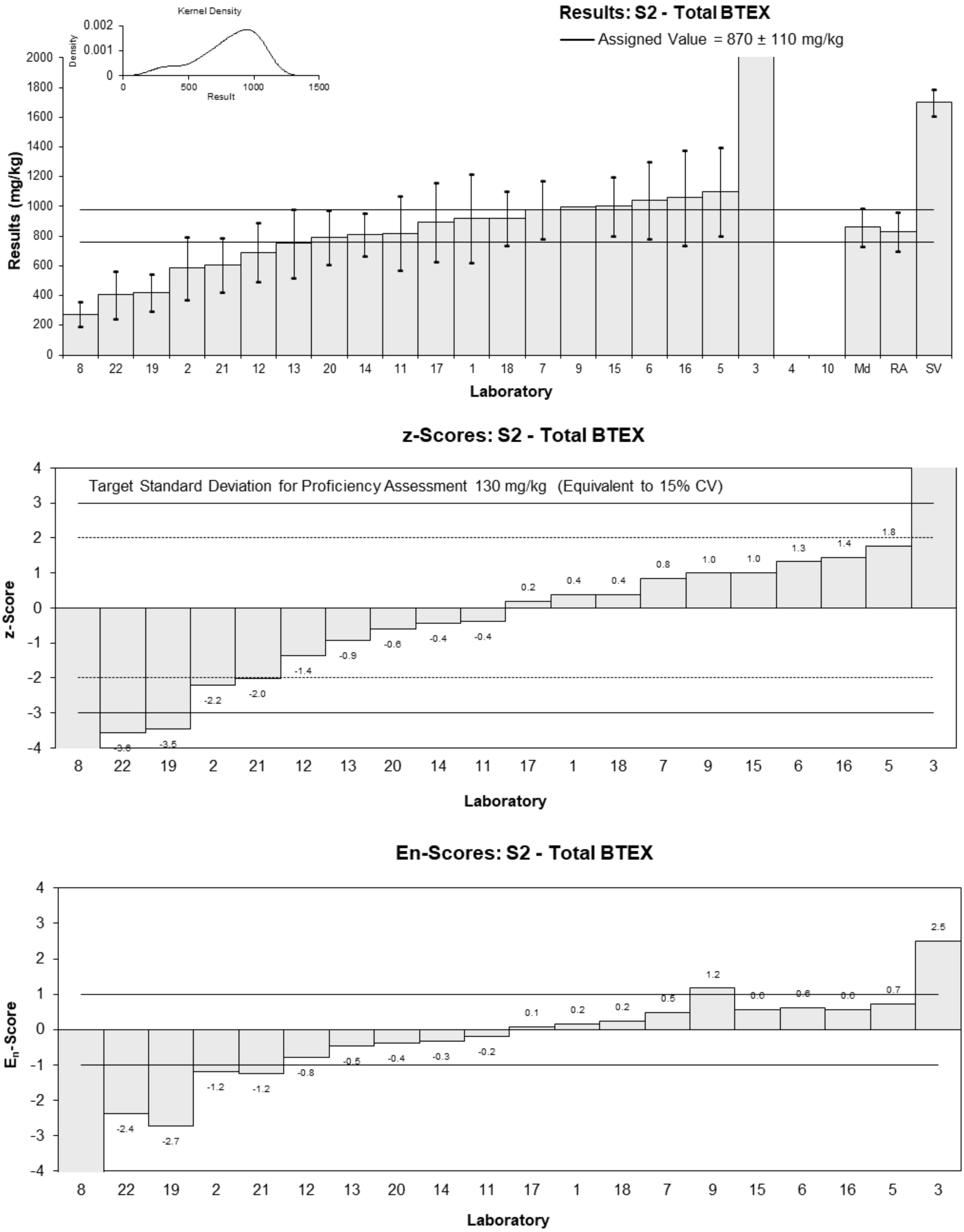

Results: S2 - Total BTEX
Assigned Value = 870 ± 110 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
z-Scores: S2 - Total BTEX
Target Standard Deviation for Proficiency Assessment 130 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
En-Scores: S2 - Total BTEX
En-Score
Laboratory

Figure 11

Table 16

Sample Details

Sample No.	S3
Matrix	Soil
Analyte	Acenaphthene
Unit	mg/kg

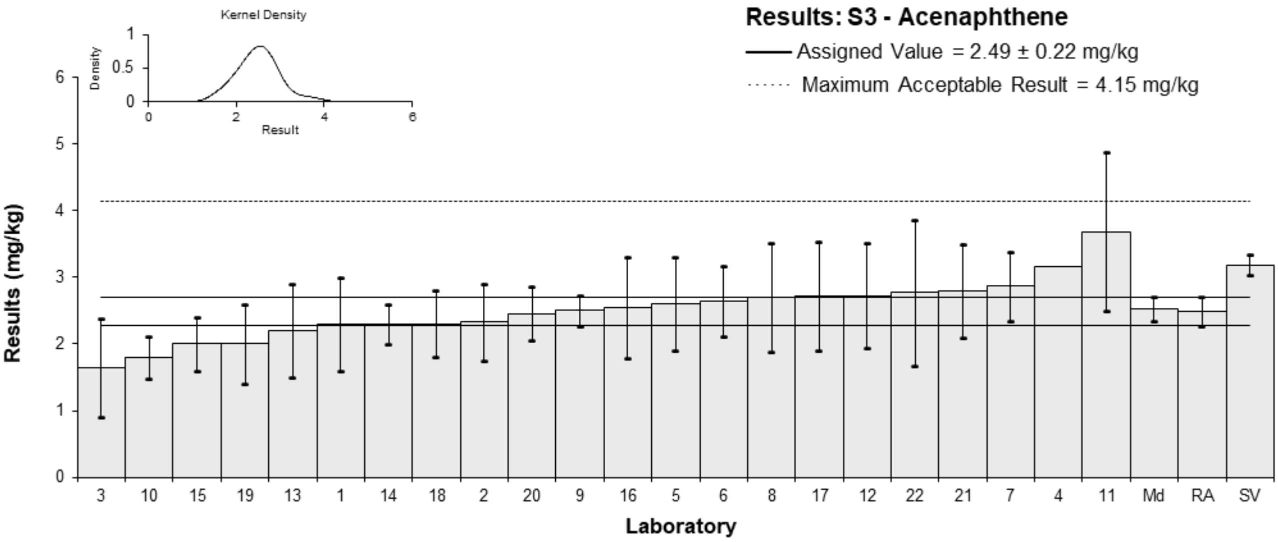
Participant Results

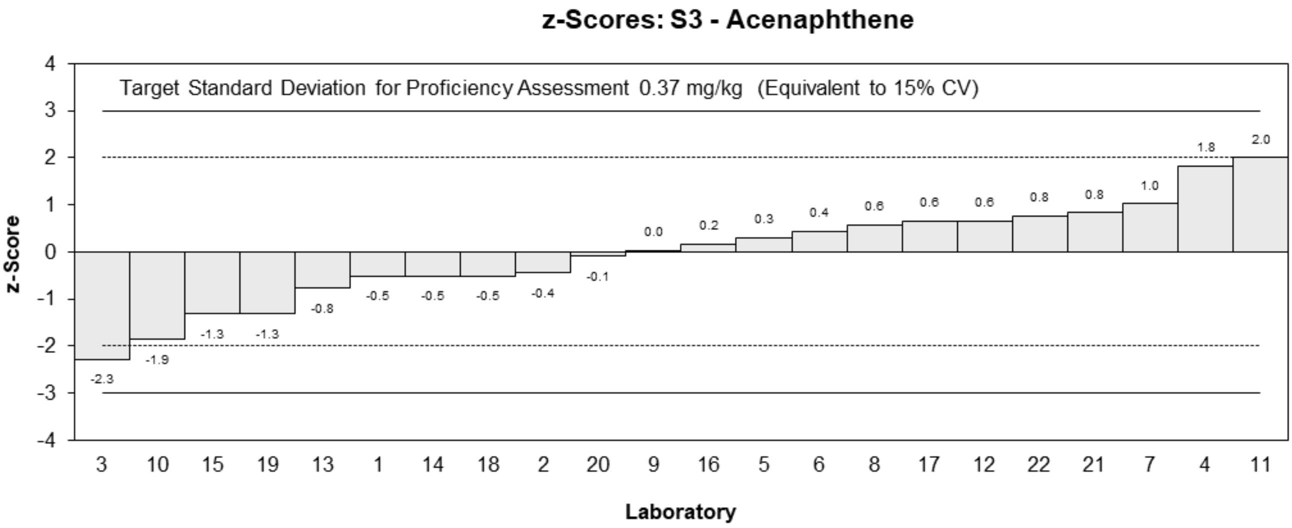
Lab. Code	Result	Uncertainty	z	E_n
1	2.3	0.7	-0.51	-0.26
2	2.33	0.58	-0.43	-0.26
3	1.64	0.74	-2.28	-1.10
4	3.17	NR	1.82	3.09
5	2.6	0.7	0.29	0.15
6	2.651	0.530	0.43	0.28
7	2.87	0.52	1.02	0.67
8	2.7	0.81	0.56	0.25
9	2.5	0.23	0.03	0.03
10	1.80	0.31	-1.85	-1.82
11	3.69	1.2	2.00▼	
12	2.73	0.78	0.64	0.30
13	2.2	0.7	-0.78	-0.40
14	2.3	0.3	-0.51	-0.51
15	2	0.4	-1.31	-1.07
16	2.55	0.76	0.16	0.08
17	2.7281	0.8184	0.64	0.28
18	2.3	0.5	-0.51	-0.35
19	2	0.6	-1.31	-0.77
20	2.46	0.405	-0.08	-0.07
21	2.8	0.7	0.83	0.42
22	2.77	1.1	0.75	0.25

▼ Adjusted Score, see Section 6.3

Statistics

Assigned Value	2.49	0.22
Spike Value	3.19	0.16
Robust Average	2.49	0.22
Max Acceptable Result	4.15	
Median	2.53	0.18
Mean	2.50	
N	22	
Max	3.69	
Min	1.64	
Robust SD	0.41	
Robust CV	16%	


Kernel Density
Density
Result
Results: S3 - Acenaphthene
Assigned Value = 2.49 ± 0.22 mg/kg
Maximum Acceptable Result = 4.15 mg/kg
Results (mg/kg)
3 10 15 19 13 1 14 18 2 20 9 16 5 6 8 17 12 22 21 7 4 11 Md RA SV
Laboratory


z-Scores: S3 - Acenaphthene
Target Standard Deviation for Proficiency Assessment 0.37 mg/kg (Equivalent to 15% CV)
z-Score
-2.3 -1.9 -1.3 -1.3 -0.8 -0.5 -0.5 -0.5 -0.4 -0.1 0.0 0.2 0.3 0.4 0.6 0.6 0.6 0.8 0.8 1.0 1.8 2.0
3 10 15 19 13 1 14 18 2 20 9 16 5 6 8 17 12 22 21 7 4 11
Laboratory

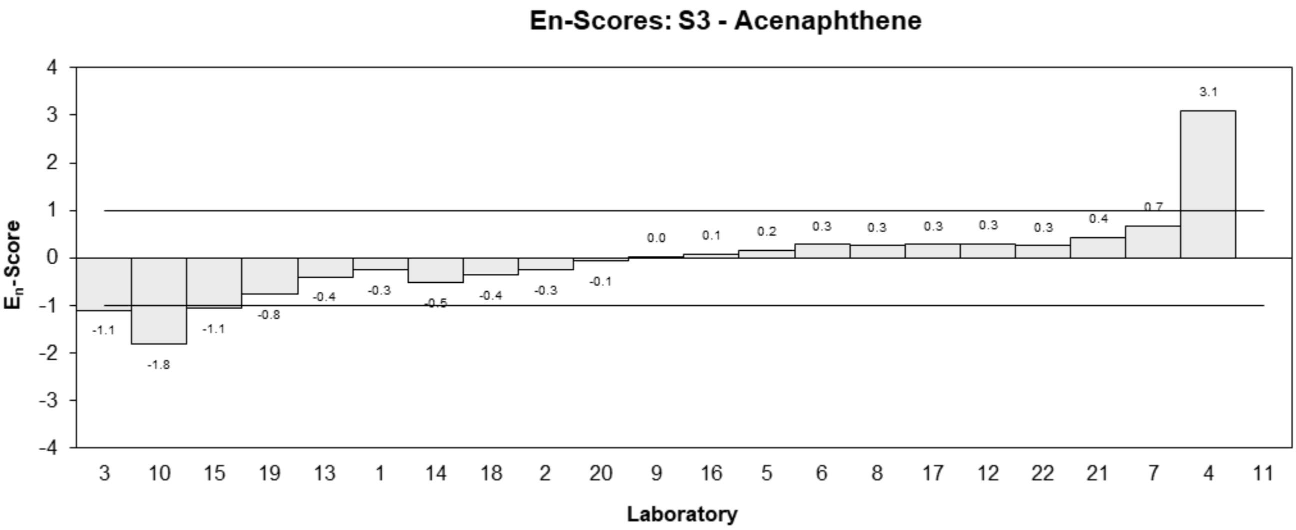

En-Scores: S3 - Acenaphthene
En-Score
-1.1 -1.8 -1.1 -0.8 -0.4 -0.3 -0.5 -0.4 -0.3 -0.1 0.0 0.1 0.2 0.3 0.3 0.3 0.3 0.3 0.4 0.7 3.1
3 10 15 19 13 1 14 18 2 20 9 16 5 6 8 17 12 22 21 7 4 11
Laboratory

Figure 12

Table 17

Sample Details

Sample No.	S3
Matrix	Soil
Analyte	Acenaphthylene
Unit	mg/kg

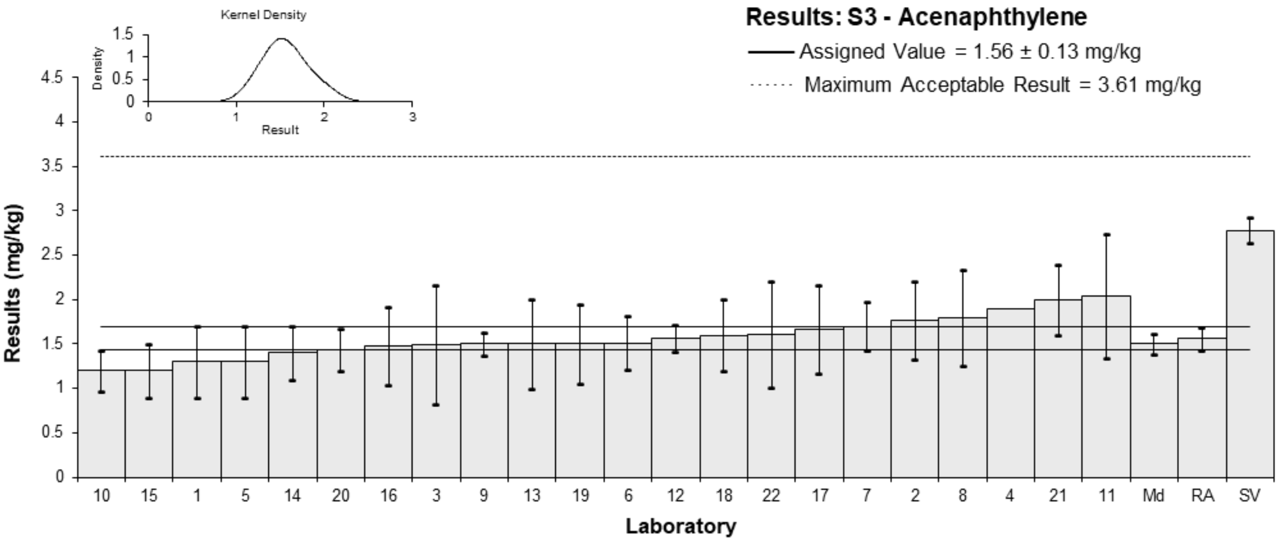
Participant Results

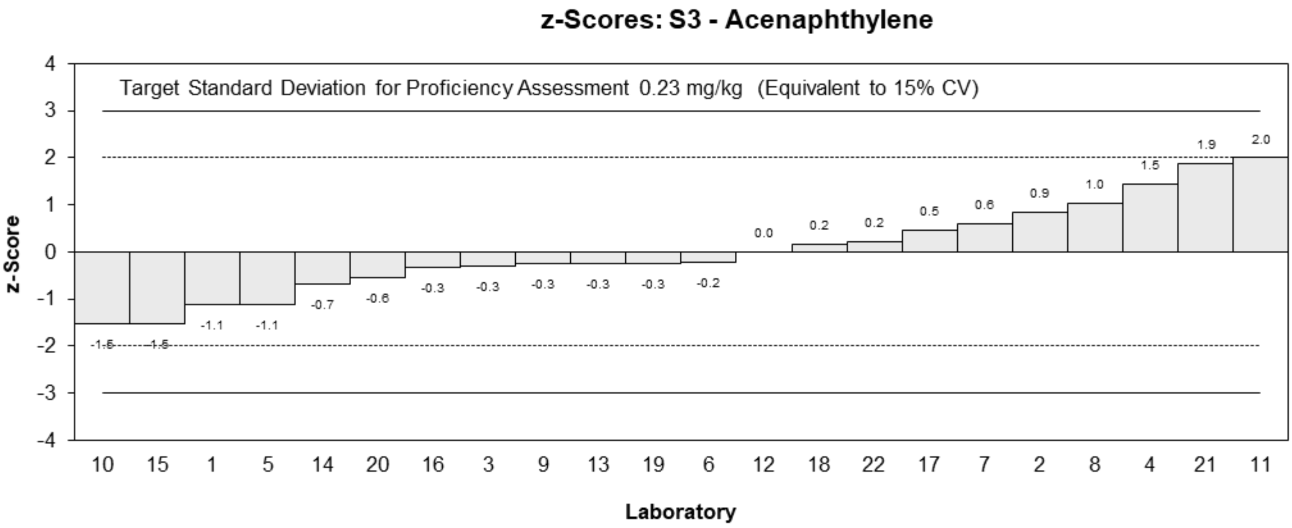
Lab. Code	Result	Uncertainty	z	E_n
1	1.3	0.4	-1.11	-0.62
2	1.76	0.44	0.85	0.44
3	1.49	0.67	-0.30	-0.10
4	1.90	NR	1.45	2.62
5	1.3	0.4	-1.11	-0.62
6	1.509	0.302	-0.22	-0.16
7	1.7	0.27	0.60	0.47
8	1.8	0.54	1.03	0.43
9	1.5	0.13	-0.26	-0.33
10	1.20	0.23	-1.54	-1.36
11	2.04	0.7	2.00▼	
12	1.56	0.15	0.00	0.00
13	1.5	0.5	-0.26	-0.12
14	1.4	0.3	-0.68	-0.49
15	1.2	0.3	-1.54	-1.10
16	1.48	0.44	-0.34	-0.17
17	1.6673	0.5002	0.46	0.21
18	1.6	0.4	0.17	0.10
19	1.5	0.45	-0.26	-0.13
20	1.43	0.236	-0.56	-0.48
21	2	0.4	1.88	1.05
22	1.61	0.6	0.21	0.08

▼ Adjusted Score, see Section 6.3

Statistics

Assigned Value	1.56	0.13
Spike Value	2.78	0.14
Robust Average	1.56	0.13
Max Acceptable Result	3.61	
Median	1.50	0.11
Mean	1.57	
N	22	
Max	2.04	
Min	1.2	
Robust SD	0.24	
Robust CV	16%	


Results: S3 - Acenaphthylene
Assigned Value = 1.56 ± 0.13 mg/kg
Maximum Acceptable Result = 3.61 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
10 15 1 5 14 20 16 3 9 13 19 6 12 18 22 17 7 2 8 4 21 11 Md RA SV


z-Scores: S3 - Acenaphthylene
Target Standard Deviation for Proficiency Assessment 0.23 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
10 15 1 5 14 20 16 3 9 13 19 6 12 18 22 17 7 2 8 4 21 11

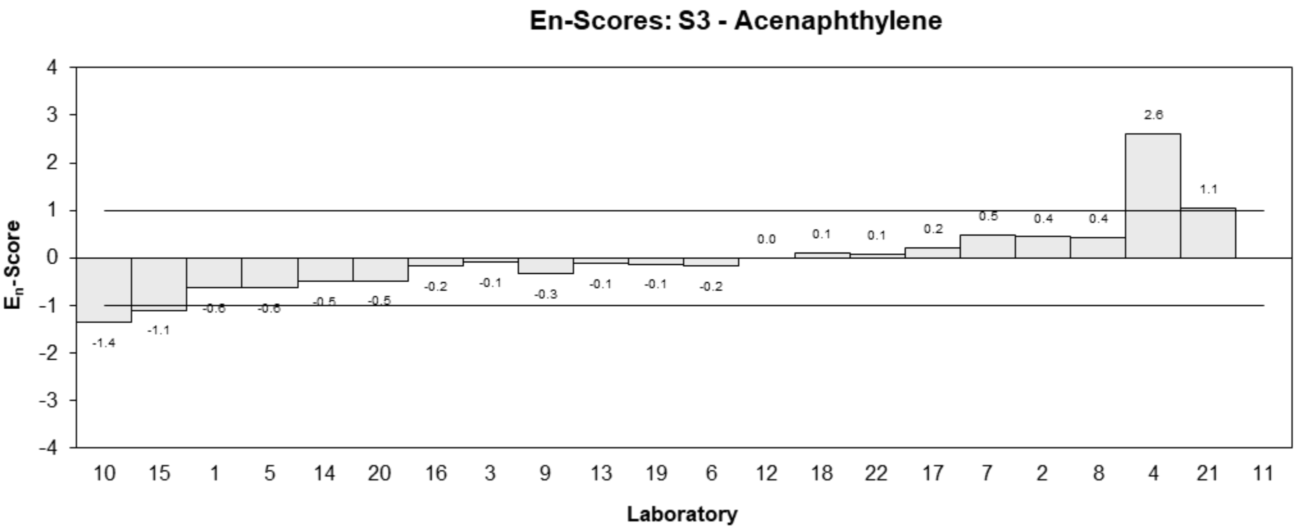

En-Scores: S3 - Acenaphthylene
En-Score
Laboratory
10 15 1 5 14 20 16 3 9 13 19 6 12 18 22 17 7 2 8 4 21 11

Figure 13

Table 18

Sample Details

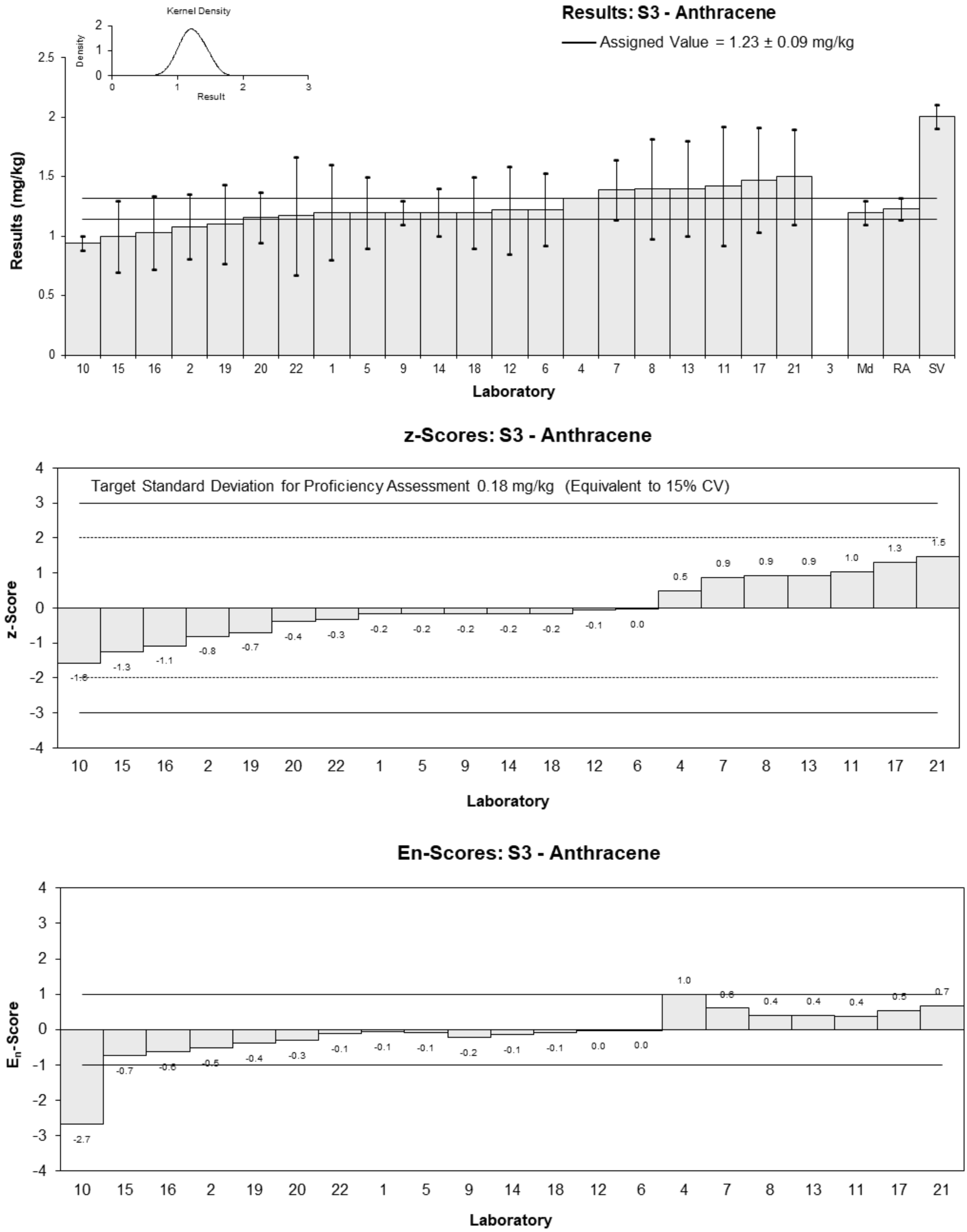
Sample No.	S3
Matrix	Soil
Analyte	Anthracene
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	1.2	0.4	-0.16	-0.07
2	1.08	0.27	-0.81	-0.53
3	<0.01	NR		
4	1.32	NR	0.49	1.00
5	1.2	0.3	-0.16	-0.10
6	1.224	0.306	-0.03	-0.02
7	1.39	0.25	0.87	0.60
8	1.4	0.42	0.92	0.40
9	1.2	0.1	-0.16	-0.22
10	0.94	0.06	-1.57	-2.68
11	1.42	0.5	1.03	0.37
12	1.22	0.37	-0.05	-0.03
13	1.4	0.4	0.92	0.41
14	1.2	0.2	-0.16	-0.14
15	1	0.3	-1.25	-0.73
16	1.03	0.31	-1.08	-0.62
17	1.4711	0.4413	1.31	0.54
18	1.2	0.3	-0.16	-0.10
19	1.1	0.33	-0.70	-0.38
20	1.16	0.212	-0.38	-0.30
21	1.5	0.4	1.46	0.66
22	1.17	0.5	-0.33	-0.12

Statistics

Assigned Value	1.23	0.09
Spike Value	2.01	0.10
Robust Average	1.23	0.09
Median	1.20	0.10
Mean	1.23	
N	21	
Max	1.5	
Min	0.94	
Robust SD	0.17	
Robust CV	14%	



Kernel Density
Density
Result
Results: S3 - Anthracene
Assigned Value = 1.23 ± 0.09 mg/kg
Results (mg/kg)
10 15 16 2 19 20 22 1 5 9 14 18 12 6 4 7 8 13 11 17 21 3 Md RA SV
Laboratory
z-Scores: S3 - Anthracene
Target Standard Deviation for Proficiency Assessment 0.18 mg/kg (Equivalent to 15% CV)
z-Score
-1.6 -1.3 -1.1 -0.8 -0.7 -0.4 -0.3 -0.2 -0.2 -0.2 -0.2 -0.2 -0.1 0.0 0.5 0.9 0.9 0.9 1.0 1.3 1.5
10 15 16 2 19 20 22 1 5 9 14 18 12 6 4 7 8 13 11 17 21
Laboratory
En-Scores: S3 - Anthracene
En-Score
-2.7 -0.7 -0.6 -0.5 -0.4 -0.3 -0.1 -0.1 -0.1 -0.2 -0.1 -0.1 0.0 0.0 1.0 0.6 0.4 0.4 0.4 0.5 0.7
10 15 16 2 19 20 22 1 5 9 14 18 12 6 4 7 8 13 11 17 21
Laboratory

Figure 14

Table 19

Sample Details

Sample No.	S3
Matrix	Soil
Analyte	Benz[*a*]anthracene
Unit	mg/kg

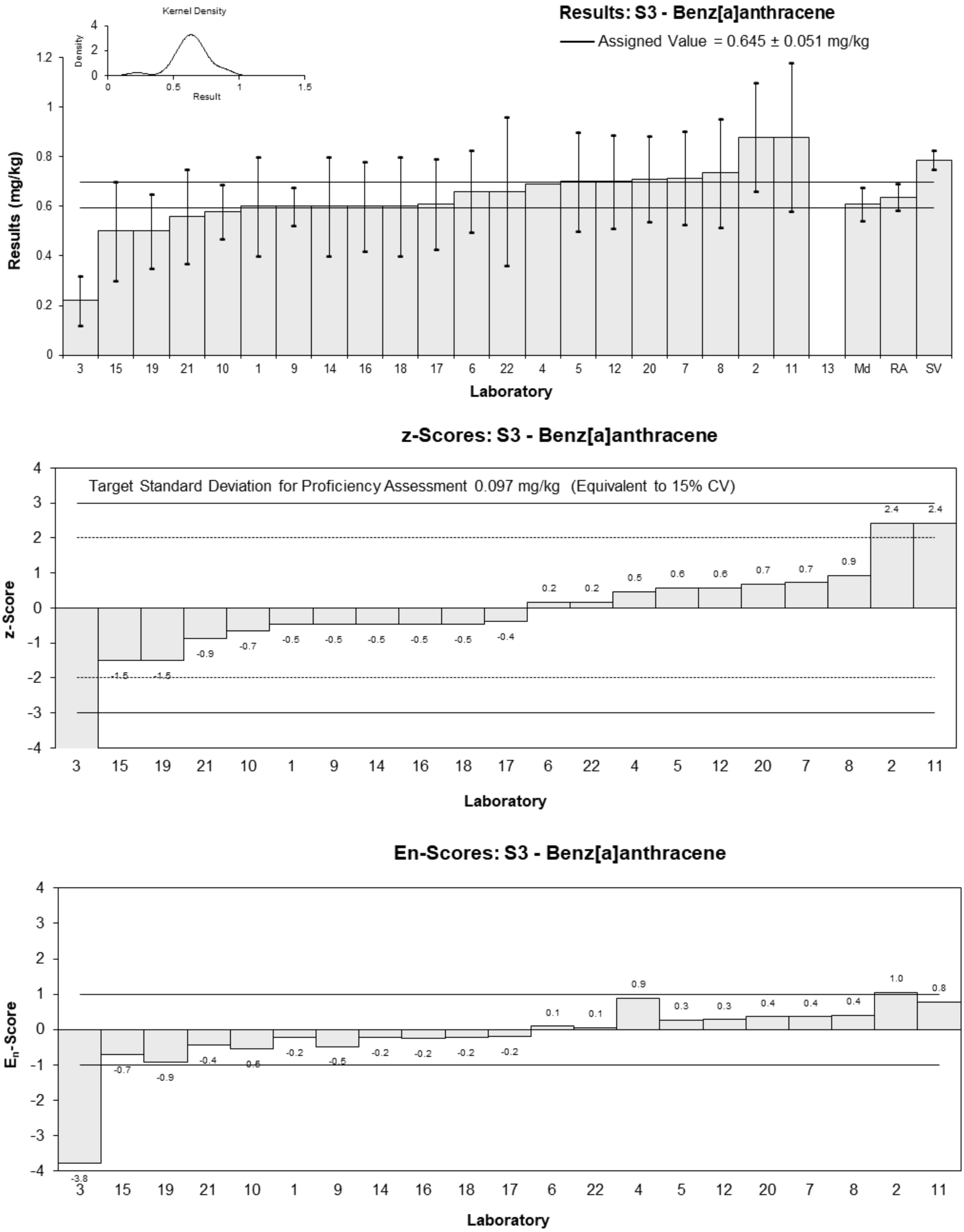
Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	0.6	0.2	-0.47	-0.22
2	0.88	0.22	2.43	1.04
3*	0.22	0.1	-4.39	-3.79
4	0.69	NR	0.47	0.88
5	0.7	0.2	0.57	0.27
6	0.66	0.165	0.16	0.09
7	0.715	0.19	0.72	0.36
8	0.735	0.22	0.93	0.40
9	0.6	0.078	-0.47	-0.48
10	0.58	0.11	-0.67	-0.54
11	0.88	0.3	2.43	0.77
12	0.7	0.19	0.57	0.28
13	<0.5	NR		
14	0.6	0.2	-0.47	-0.22
15	0.5	0.2	-1.50	-0.70
16	0.60	0.18	-0.47	-0.24
17	0.6085	0.1826	-0.38	-0.19
18	0.6	0.2	-0.47	-0.22
19	0.5	0.15	-1.50	-0.92
20	0.71	0.173	0.67	0.36
21	0.56	0.19	-0.88	-0.43
22	0.66	0.3	0.16	0.05

* Outlier, see Section 4.2

Statistics

Assigned Value	0.645	0.051
Spike Value	0.788	0.039
Robust Average	0.637	0.054
Median	0.609	0.066
Mean	0.633	
N	21	
Max	0.88	
Min	0.22	
Robust SD	0.100	
Robust CV	16%	



Results: S3 - Benz[a]anthracene
Assigned Value = 0.645 ± 0.051 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
3 15 19 21 10 1 9 14 16 18 17 6 22 4 5 12 20 7 8 2 11 13 Md RA SV
z-Scores: S3 - Benz[a]anthracene
Target Standard Deviation for Proficiency Assessment 0.097 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
-1.5 -1.5 -0.9 -0.7 -0.5 -0.5 -0.5 -0.5 -0.5 -0.4 0.2 0.2 0.5 0.6 0.6 0.7 0.7 0.9 2.4 2.4
En-Scores: S3 - Benz[a]anthracene
En-Score
Laboratory
-3.8 -0.7 -0.9 -0.4 -0.5 -0.2 -0.5 -0.2 -0.2 -0.2 -0.2 0.1 0.1 0.9 0.3 0.3 0.4 0.4 0.4 1.0 0.8

Figure 15

Table 20

Sample Details

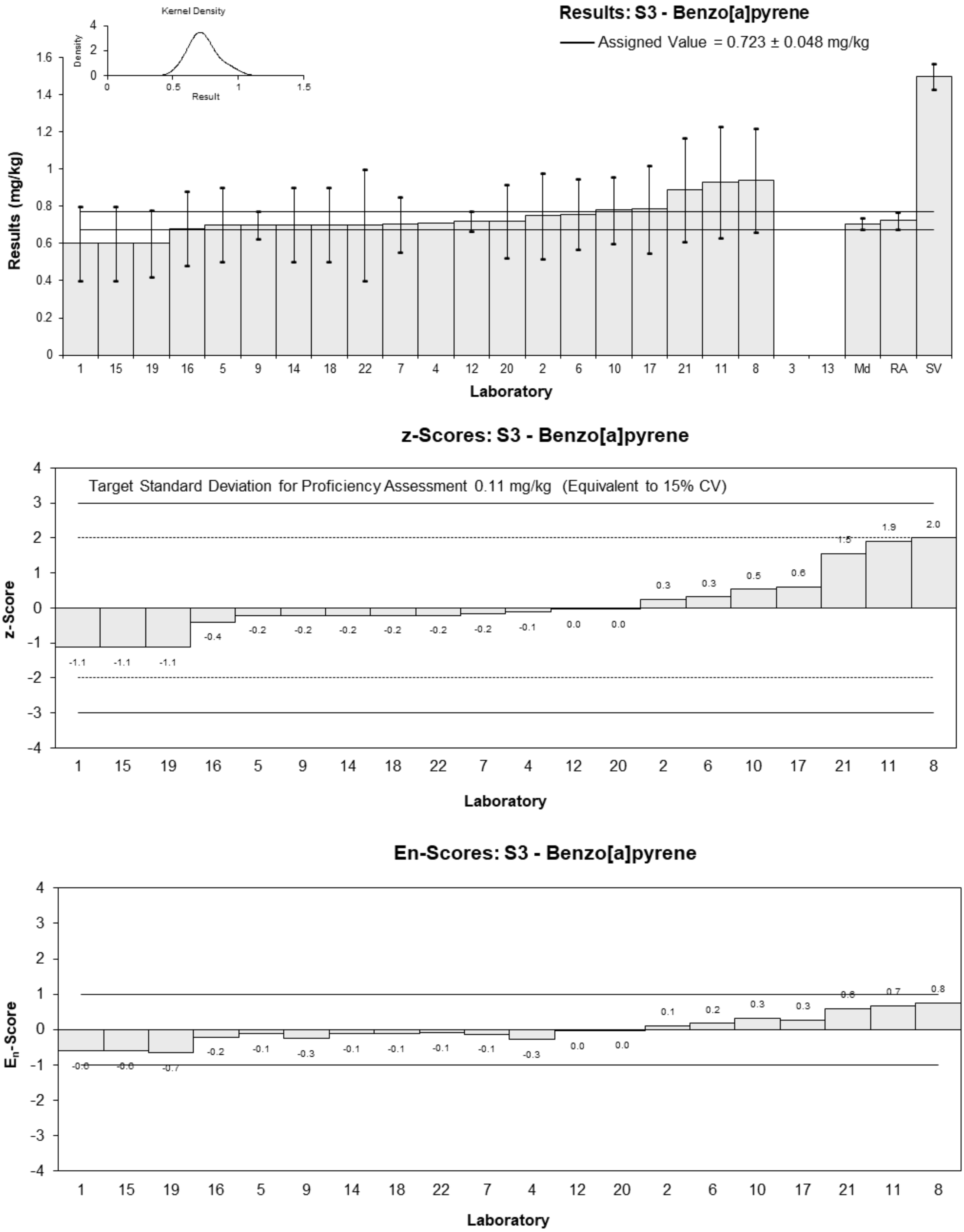
Sample No.	S3
Matrix	Soil
Analyte	Benzo[*a*]pyrene
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	0.6	0.2	-1.13	-0.60
2	0.75	0.23	0.25	0.11
3	<0.01	NR		
4	0.71	NR	-0.12	-0.27
5	0.7	0.2	-0.21	-0.11
6	0.757	0.189	0.31	0.17
7	0.703	0.15	-0.18	-0.13
8	0.94	0.28	2.00	0.76
9	0.7	0.074	-0.21	-0.26
10	0.78	0.18	0.53	0.31
11	0.93	0.3	1.91	0.68
12	0.72	0.056	-0.03	-0.04
13	<0.5	NR		
14	0.7	0.2	-0.21	-0.11
15	0.6	0.2	-1.13	-0.60
16	0.68	0.20	-0.40	-0.21
17	0.7867	0.236	0.59	0.26
18	0.7	0.2	-0.21	-0.11
19	0.6	0.18	-1.13	-0.66
20	0.72	0.197	-0.03	-0.01
21	0.89	0.28	1.54	0.59
22	0.7	0.3	-0.21	-0.08

Statistics

Assigned Value	0.723	0.048
Spike Value	1.50	0.07
Robust Average	0.723	0.048
Median	0.707	0.029
Mean	0.733	
N	20	
Max	0.94	
Min	0.6	
Robust SD	0.086	
Robust CV	12%	



Kernel Density
Density
Result
Results: S3 - Benzo[a]pyrene
Assigned Value = 0.723 ± 0.048 mg/kg
Results (mg/kg)
Laboratory
1 15 19 16 5 9 14 18 22 7 4 12 20 2 6 10 17 21 11 8 3 13 Md RA SV
z-Scores: S3 - Benzo[a]pyrene
Target Standard Deviation for Proficiency Assessment 0.11 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
1 15 19 16 5 9 14 18 22 7 4 12 20 2 6 10 17 21 11 8
En-Scores: S3 - Benzo[a]pyrene
En-Score
Laboratory
1 15 19 16 5 9 14 18 22 7 4 12 20 2 6 10 17 21 11 8

Figure 16

Table 21

Sample Details

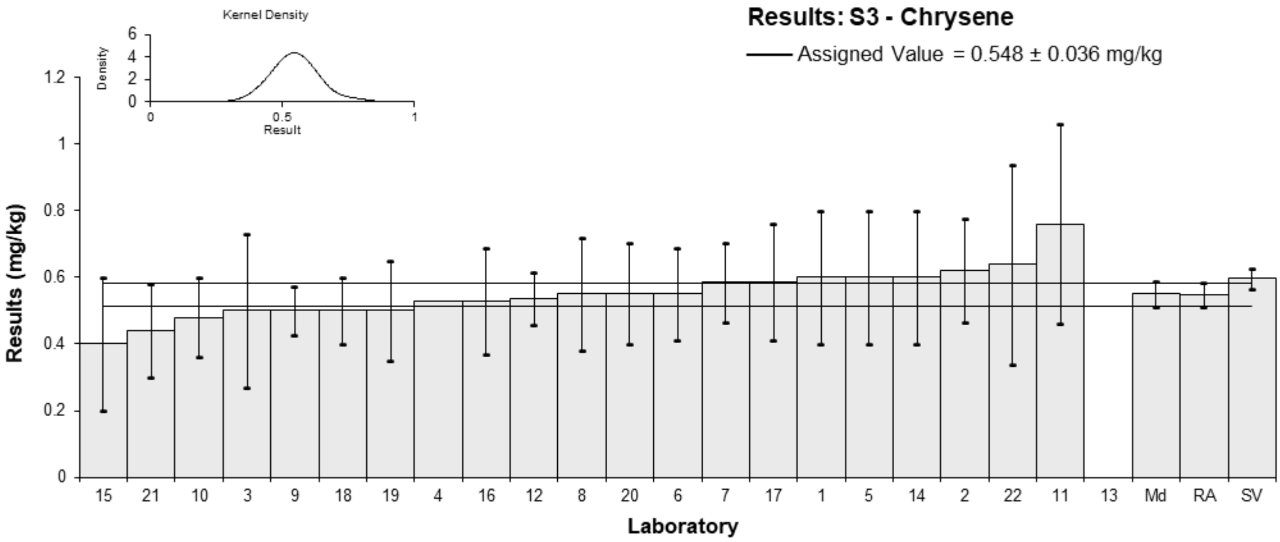
Sample No.	S3
Matrix	Soil
Analyte	Chrysene
Unit	mg/kg

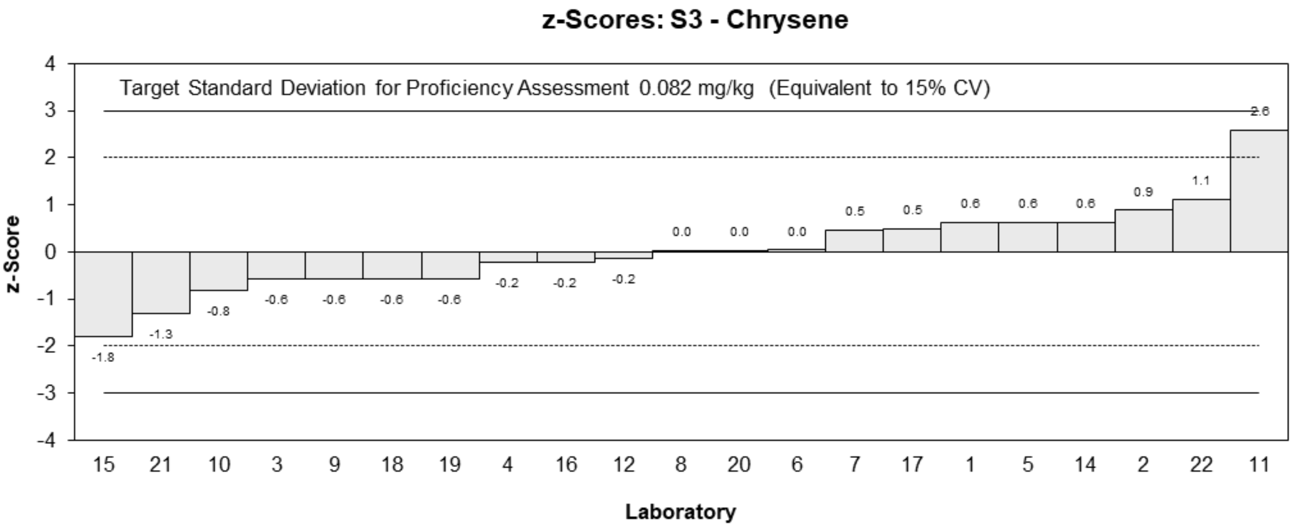
Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	0.6	0.2	0.63	0.26
2	0.62	0.155	0.88	0.45
3	0.5	0.23	-0.58	-0.21
4	0.53	NR	-0.22	-0.50
5	0.6	0.2	0.63	0.26
6	0.551	0.138	0.04	0.02
7	0.585	0.12	0.45	0.30
8	0.55	0.17	0.02	0.01
9	0.5	0.074	-0.58	-0.58
10	0.48	0.12	-0.83	-0.54
11	0.76	0.3	2.58	0.70
12	0.536	0.08	-0.15	-0.14
13	<0.5	NR		
14	0.6	0.2	0.63	0.26
15	0.4	0.2	-1.80	-0.73
16	0.53	0.16	-0.22	-0.11
17	0.5871	0.1761	0.48	0.22
18	0.5	0.1	-0.58	-0.45
19	0.5	0.15	-0.58	-0.31
20	0.55	0.152	0.02	0.01
21	0.44	0.14	-1.31	-0.75
22	0.64	0.3	1.12	0.30

Statistics

Assigned Value	0.548	0.036
Spike Value	0.596	0.030
Robust Average	0.548	0.036
Median	0.550	0.040
Mean	0.550	
N	21	
Max	0.76	
Min	0.4	
Robust SD	0.067	
Robust CV	12%	


Results: S3 - Chrysene
Assigned Value = 0.548 ± 0.036 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
15 21 10 3 9 18 19 4 16 12 8 20 6 7 17 1 5 14 2 22 11 13 Md RA SV


z-Scores: S3 - Chrysene
Target Standard Deviation for Proficiency Assessment 0.082 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
15 21 10 3 9 18 19 4 16 12 8 20 6 7 17 1 5 14 2 22 11
-1.8 -1.3 -0.8 -0.6 -0.6 -0.6 -0.6 -0.2 -0.2 -0.2 0.0 0.0 0.0 0.5 0.5 0.6 0.6 0.6 0.9 1.1 2.6

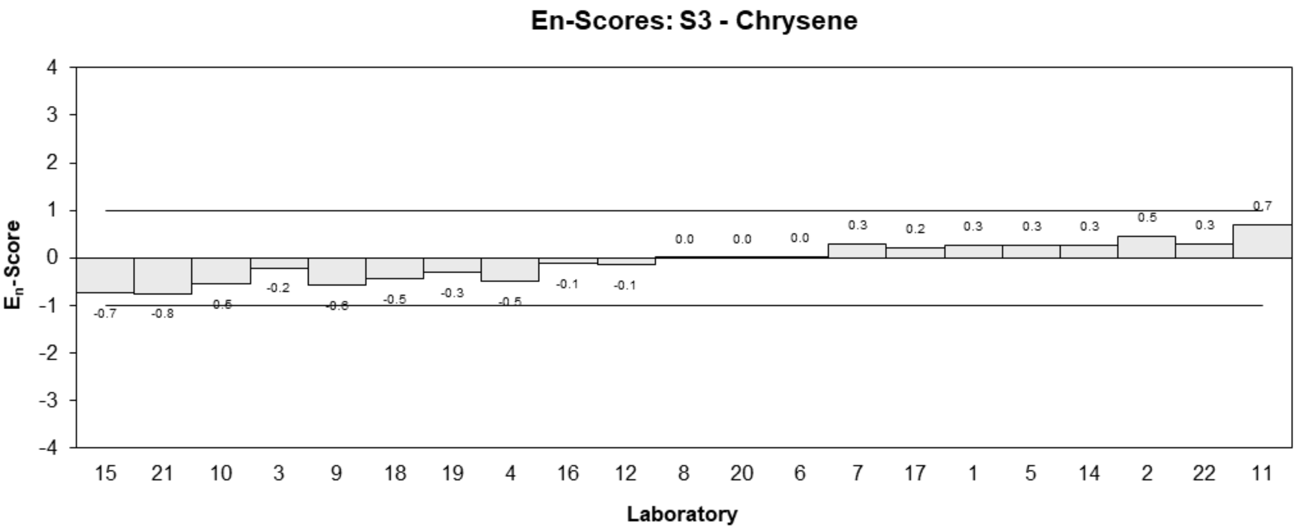

En-Scores: S3 - Chrysene
E_n-Score
Laboratory
15 21 10 3 9 18 19 4 16 12 8 20 6 7 17 1 5 14 2 22 11
-0.7 -0.8 -0.5 -0.2 -0.6 -0.5 -0.3 -0.5 -0.1 -0.1 0.0 0.0 0.0 0.3 0.2 0.3 0.3 0.3 0.5 0.3 0.7

Figure 17

Table 22

Sample Details

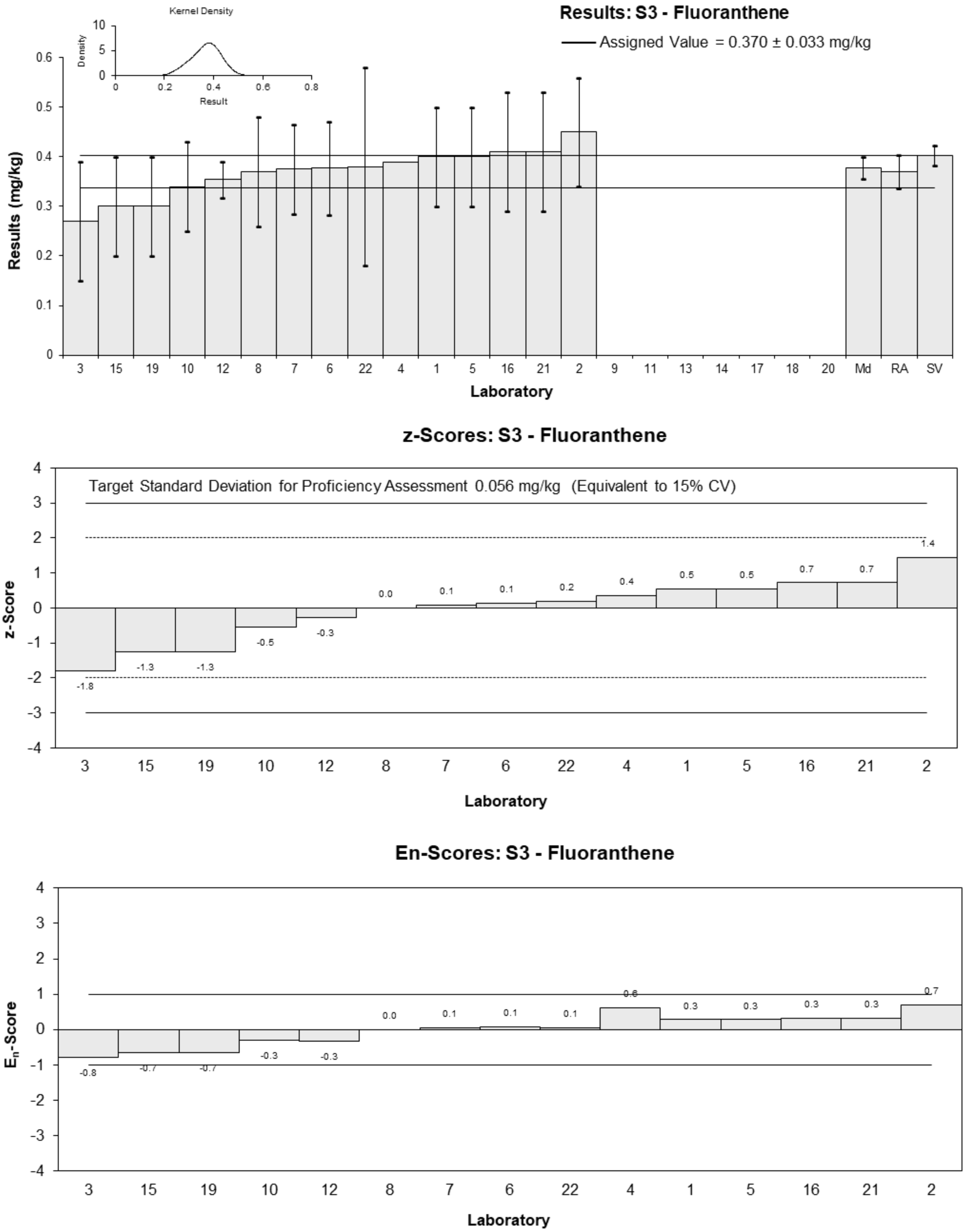
Sample No.	S3
Matrix	Soil
Analyte	Fluoranthene
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	0.4	0.1	0.54	0.28
2	0.45	0.11	1.44	0.70
3	0.27	0.12	-1.80	-0.80
4	0.39	NR	0.36	0.61
5	0.4	0.1	0.54	0.28
6	0.377	0.094	0.13	0.07
7	0.375	0.09	0.09	0.05
8	0.37	0.11	0.00	0.00
9	<0.5	NR		
10	0.34	0.09	-0.54	-0.31
11	<0.5	NR		
12	0.354	0.036	-0.29	-0.33
13	<0.5	NR		
14	<0.5	NR		
15	0.3	0.1	-1.26	-0.66
16	0.41	0.12	0.72	0.32
17	<0.5	NR		
18	<0.5	0.1		
19	0.3	0.1	-1.26	-0.66
20	<0.50	NR		
21	0.41	0.12	0.72	0.32
22	0.38	0.2	0.18	0.05

Statistics

Assigned Value	0.370	0.033
Spike Value	0.403	0.020
Robust Average	0.370	0.033
Median	0.377	0.022
Mean	0.368	
N	15	
Max	0.45	
Min	0.27	
Robust SD	0.051	
Robust CV	14%	



Kernel Density
Density
Result
Results: S3 - Fluoranthene
Assigned Value = 0.370 ± 0.033 mg/kg
Results (mg/kg)
Laboratory
3 15 19 10 12 8 7 6 22 4 1 5 16 21 2 9 11 13 14 17 18 20 Md RA SV
z-Scores: S3 - Fluoranthene
Target Standard Deviation for Proficiency Assessment 0.056 mg/kg (Equivalent to 15% CV)
z-Score
-1.8 -1.3 -1.3 -0.5 -0.3 0.0 0.1 0.1 0.2 0.4 0.5 0.5 0.7 0.7 1.4
Laboratory
3 15 19 10 12 8 7 6 22 4 1 5 16 21 2
En-Scores: S3 - Fluoranthene
E_n-Score
-0.8 -0.7 -0.7 -0.3 -0.3 0.0 0.1 0.1 0.1 0.6 0.3 0.3 0.3 0.3 0.7
Laboratory
3 15 19 10 12 8 7 6 22 4 1 5 16 21 2

Figure 18

Table 23

Sample Details

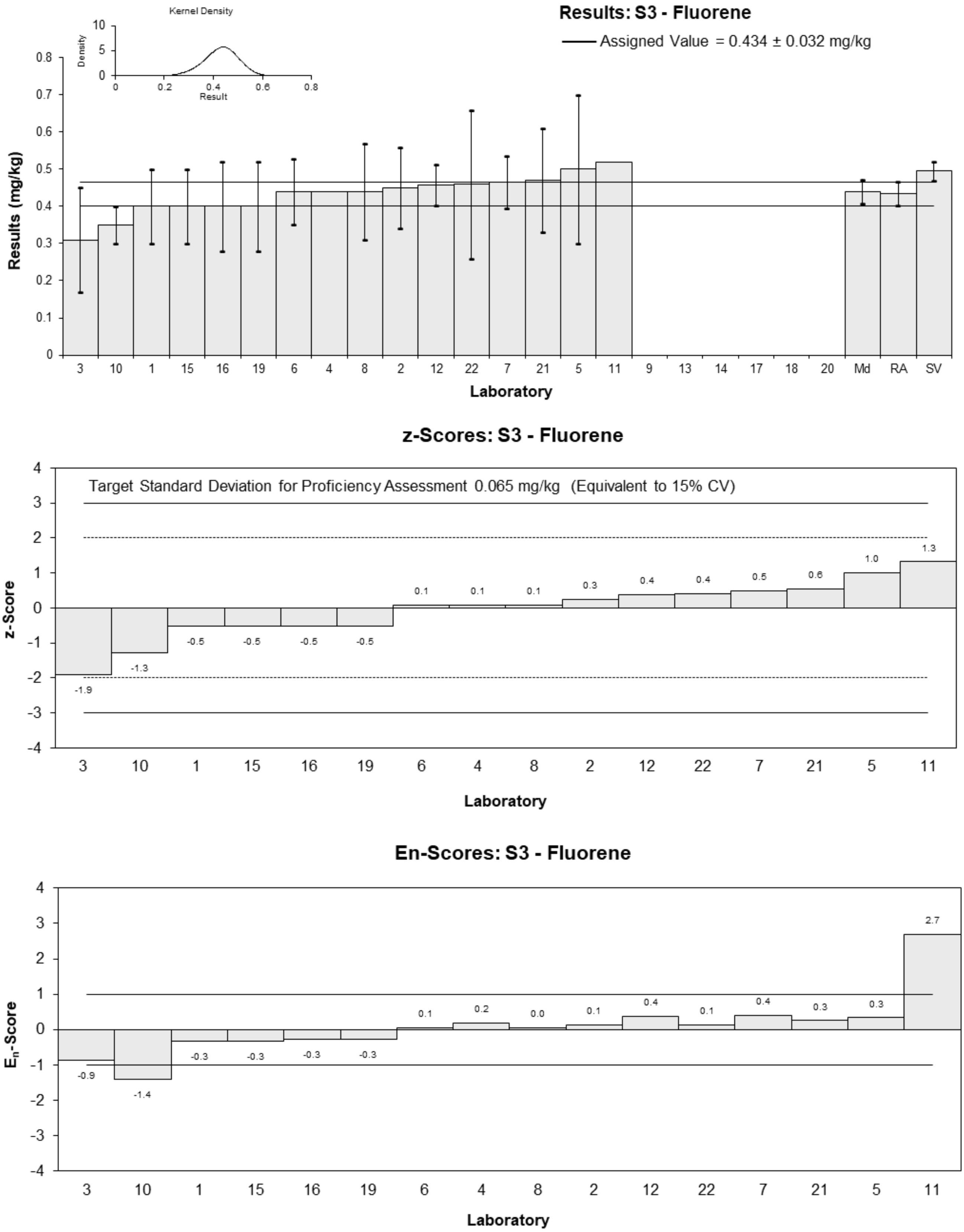
Sample No.	S3
Matrix	Soil
Analyte	Fluorene
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	0.4	0.1	-0.52	-0.32
2	0.45	0.11	0.25	0.14
3	0.31	0.14	-1.90	-0.86
4	0.44	NR	0.09	0.19
5	0.5	0.2	1.01	0.33
6	0.439	0.088	0.08	0.05
7	0.465	0.07	0.48	0.40
8	0.44	0.13	0.09	0.04
9	<0.5	NR		
10	0.35	0.05	-1.29	-1.42
11	0.52	NR	1.32	2.69
12	0.458	0.055	0.37	0.38
13	<0.5	NR		
14	<0.5	NR		
15	0.4	0.1	-0.52	-0.32
16	0.40	0.12	-0.52	-0.27
17	<0.5	NR		
18	<0.5	0.1		
19	0.4	0.12	-0.52	-0.27
20	<0.50	NR		
21	0.47	0.14	0.55	0.25
22	0.46	0.2	0.40	0.13

Statistics

Assigned Value	0.434	0.032
Spike Value	0.495	0.025
Robust Average	0.434	0.032
Median	0.440	0.032
Mean	0.431	
N	16	
Max	0.52	
Min	0.31	
Robust SD	0.051	
Robust CV	12%	



Kernel Density
Density
Result
Results: S3 - Fluorene
Assigned Value = 0.434 ± 0.032 mg/kg
Results (mg/kg)
Laboratory
3 10 1 15 16 19 6 4 8 2 12 22 7 21 5 11 9 13 14 17 18 20 Md RA SV
z-Scores: S3 - Fluorene
Target Standard Deviation for Proficiency Assessment 0.065 mg/kg (Equivalent to 15% CV)
z-Score
-1.9 -1.3 -0.5 -0.5 -0.5 -0.5 0.1 0.1 0.1 0.3 0.4 0.4 0.5 0.6 1.0 1.3
3 10 1 15 16 19 6 4 8 2 12 22 7 21 5 11
Laboratory
En-Scores: S3 - Fluorene
En-Score
-0.9 -1.4 -0.3 -0.3 -0.3 -0.3 0.1 0.2 0.0 0.1 0.4 0.1 0.4 0.3 0.3 2.7
3 10 1 15 16 19 6 4 8 2 12 22 7 21 5 11
Laboratory

Figure 19

Table 24

Sample Details

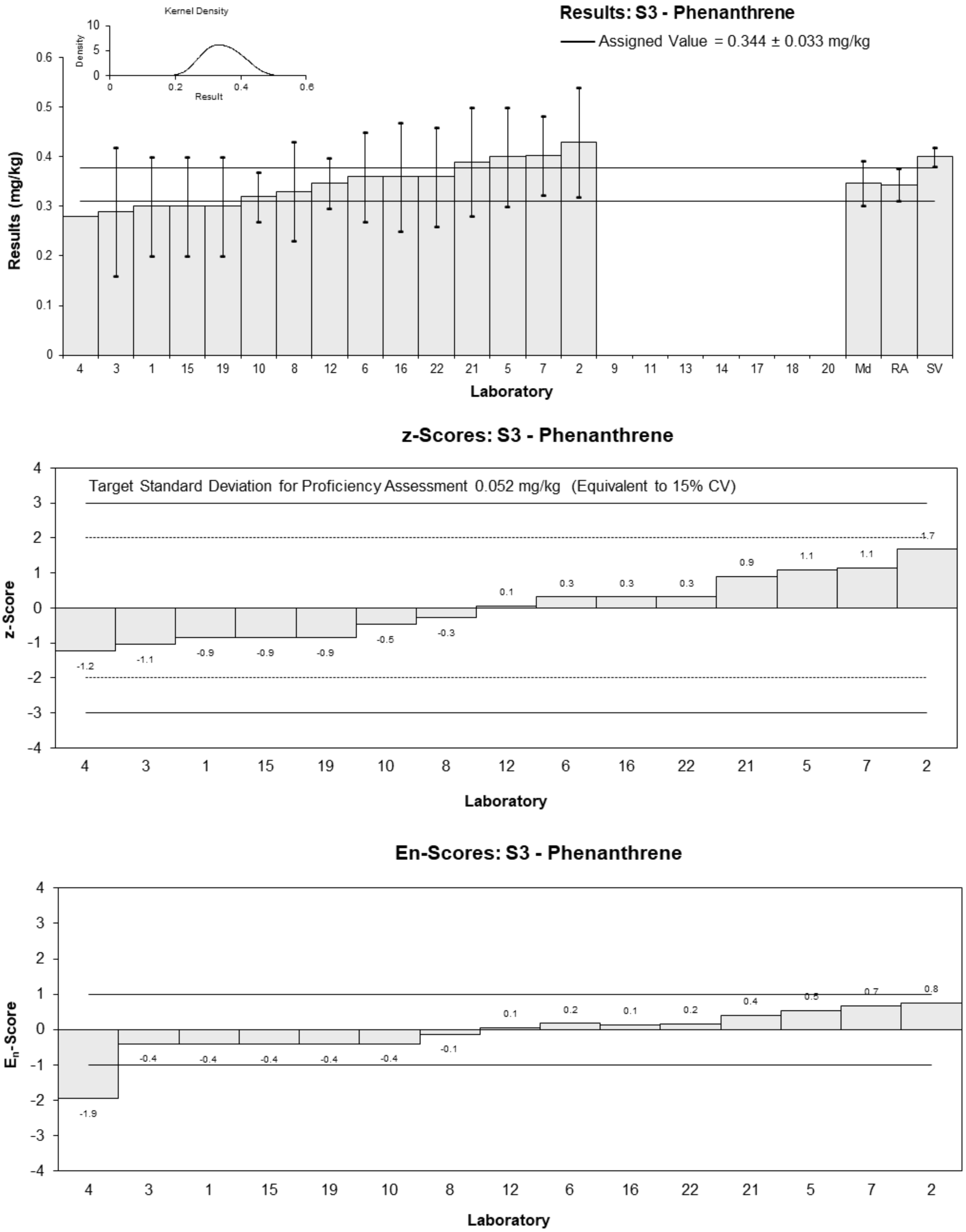
Sample No.	S3
Matrix	Soil
Analyte	Phenanthrene
Unit	mg/kg

Participant Results

Lab. Code	Result	Uncertainty	z	E_n
1	0.3	0.1	-0.85	-0.42
2	0.43	0.11	1.67	0.75
3	0.29	0.13	-1.05	-0.40
4	0.28	NR	-1.24	-1.94
5	0.4	0.1	1.09	0.53
6	0.36	0.090	0.31	0.17
7	0.403	0.08	1.14	0.68
8	0.33	0.10	-0.27	-0.13
9	<0.5	NR		
10	0.32	0.05	-0.47	-0.40
11	<0.5	NR		
12	0.347	0.05	0.06	0.05
13	<0.5	NR		
14	<0.5	NR		
15	0.3	0.1	-0.85	-0.42
16	0.36	0.11	0.31	0.14
17	<0.5	NR		
18	<0.5	0.03		
19	0.3	0.1	-0.85	-0.42
20	<0.50	NR		
21	0.39	0.11	0.89	0.40
22	0.36	0.1	0.31	0.15

Statistics

Assigned Value	0.344	0.033
Spike Value	0.400	0.020
Robust Average	0.344	0.033
Median	0.347	0.045
Mean	0.345	
N	15	
Max	0.43	
Min	0.28	
Robust SD	0.052	
Robust CV	15%	



Results: S3 - Phenanthrene
Assigned Value = 0.344 ± 0.033 mg/kg
Kernel Density
Density
Result
Results (mg/kg)
Laboratory
4
3
1
15
19
10
8
12
6
16
22
21
5
7
2
9
11
13
14
17
18
20
Md
RA
SV
z-Scores: S3 - Phenanthrene
Target Standard Deviation for Proficiency Assessment 0.052 mg/kg (Equivalent to 15% CV)
z-Score
Laboratory
En-Scores: S3 - Phenanthrene
En-Score
Laboratory

Figure 20

Table 25

Sample Details

Sample No.	S3
Matrix	Soil
Analyte	Pyrene
Unit	mg/kg

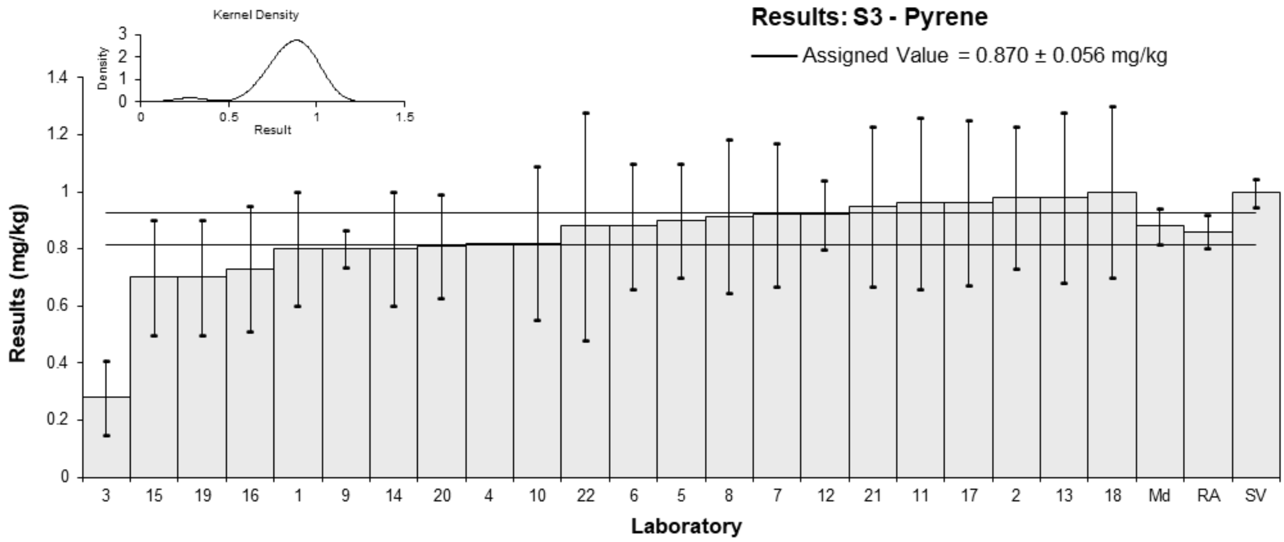
Participant Results

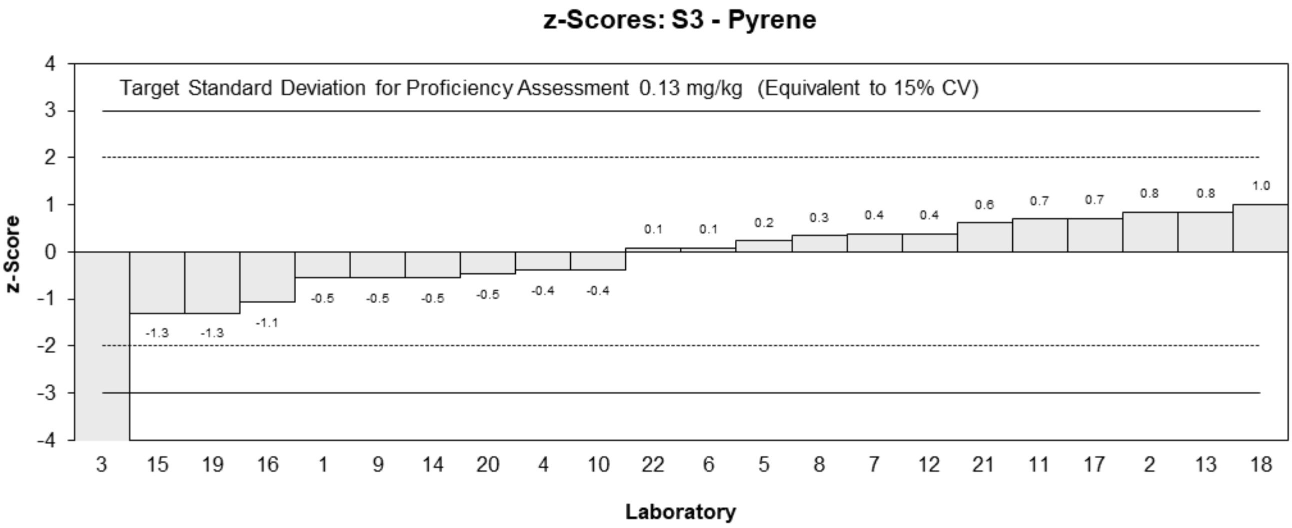
Lab. Code	Result	Uncertainty	z	E_n
1	0.8	0.2	-0.54	-0.34
2	0.98	0.25	0.84	0.43
3*	0.28	0.13	-4.52	-4.17
4	0.82	NR	-0.38	-0.89
5	0.9	0.2	0.23	0.14
6	0.881	0.220	0.08	0.05
7	0.92	0.25	0.38	0.20
8	0.915	0.27	0.34	0.16
9	0.8	0.066	-0.54	-0.81
10	0.82	0.27	-0.38	-0.18
11	0.96	0.3	0.69	0.29
12	0.92	0.12	0.38	0.38
13	0.98	0.3	0.84	0.36
14	0.8	0.2	-0.54	-0.34
15	0.7	0.2	-1.30	-0.82
16	0.73	0.22	-1.07	-0.62
17	0.9624	0.2887	0.71	0.31
18	1	0.3	1.00	0.43
19	0.7	0.2	-1.30	-0.82
20	0.81	0.182	-0.46	-0.32
21	0.95	0.28	0.61	0.28
22	0.88	0.4	0.08	0.02

* Outlier, see Section 4.2

Statistics

Assigned Value	0.870	0.056
Spike Value	0.996	0.050
Robust Average	0.860	0.059
Median	0.881	0.063
Mean	0.841	
N	22	
Max	1	
Min	0.28	
Robust SD	0.11	
Robust CV	13%	


Kernel Density
Density
Result
Results: S3 - Pyrene
Assigned Value = 0.870 ± 0.056 mg/kg
Results (mg/kg)
3 15 19 16 1 9 14 20 4 10 22 6 5 8 7 12 21 11 17 2 13 18 Md RA SV
Laboratory


z-Scores: S3 - Pyrene
Target Standard Deviation for Proficiency Assessment 0.13 mg/kg (Equivalent to 15% CV)
z-Score
-1.3 -1.3 -1.1 -0.5 -0.5 -0.5 -0.5 -0.4 -0.4 0.1 0.1 0.2 0.3 0.4 0.4 0.6 0.7 0.7 0.8 0.8 1.0
3 15 19 16 1 9 14 20 4 10 22 6 5 8 7 12 21 11 17 2 13 18
Laboratory

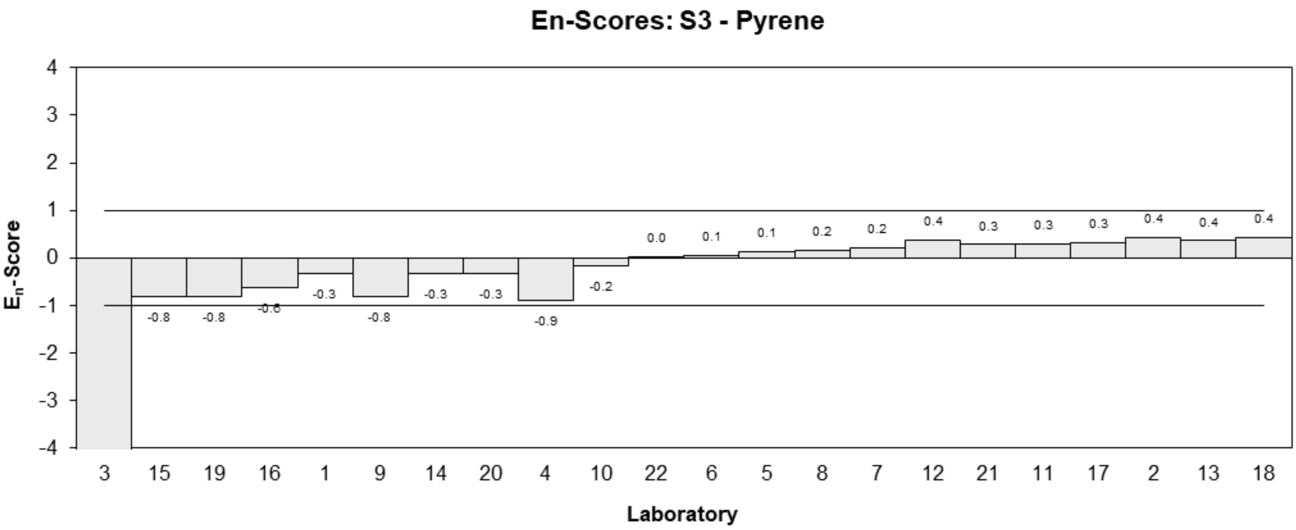

En-Scores: S3 - Pyrene
En-Score
-0.8 -0.8 -0.6 -0.3 -0.8 -0.3 -0.3 -0.9 -0.2 0.0 0.1 0.1 0.2 0.2 0.4 0.3 0.3 0.3 0.4 0.4 0.4
3 15 19 16 1 9 14 20 4 10 22 6 5 8 7 12 21 11 17 2 13 18
Laboratory

Figure 21

6 DISCUSSION OF RESULTS

6.1 Assigned Value

The assigned values for all scored analytes were the robust averages of participants' results. If there were results less than 50% or greater than 150% of the robust average, these were excluded from the calculation of each assigned value.[3,4] The robust averages and associated expanded uncertainties were calculated using the procedure described in ISO 13528.[7] The calculation of the expanded uncertainty for robust averages is presented in Appendix 3, using Sample S3 chrysene 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 S2 benzene, as numeric results reported for this analyte were highly variable; this may be due to the volatility of this compound. Sample S2 C6-C10 was also not scored because of its volatile nature and results have been provided for information only, though participants' results in this study were in reasonable consensus with each other.

For some of the analytes spiked in this study, a proportion of the analyte may be strongly bound to the soil, and so may not be readily extracted and measured. What laboratories measure may be described as 'extractable analyte', and the result may be influenced by the efficiency of the extraction process used. Therefore, the assigned value for some analytes may instead be the best estimate of the amount of 'extractable analyte'.

A comparison of the assigned values (or robust averages if no assigned value was set) and the spiked values is presented in Table 26. The assigned values for TRH were within the range of 87% to 102% of the spiked values, showing good consensus between the spiked and assigned values. The assigned values for scored BTEX and PAHs were within the ranges of 51% to 64% and 48% to 92% of the spiked values respectively. Similar ratios have been observed in previous PT studies, and an assigned value was set if there was a reasonable consensus of participants' results.

Table 26 Comparison of Assigned Value (or Robust Average) and Spiked Value

Sample	Analyte	Assigned Value *(Robust Average)* (mg/kg)	Spiked Value* (mg/kg)	Assigned Value *(Robust Average)* / Spiked Value (%)
S1	>C10-C16	937	1080	87
	>C16-C34	1730	1700	102
	>C34-C40	242	239	101
	TRH	2910	3020	96
S2	Benzene	*(60)*	224	*(27)*
	Toluene	369	723	51
	Ethylbenzene	82.9	138	60
	Xylenes	395	617	64
	Total BTEX	870	1700	51
S3	Acenaphthene	2.49	3.19	78
	Acenaphthylene	1.56	2.78	56
	Anthracene	1.23	2.01	61
	Benz[*a*]anthracene	0.645	0.788	82

Sample	Analyte	Assigned Value *(Robust Average)* (mg/kg)	Spiked Value* (mg/kg)	Assigned Value *(Robust Average)* / Spiked Value (%)
	Benzo[*a*]pyrene	0.723	1.50	48
	Chrysene	0.548	0.596	92
	Fluoranthene	0.370	0.403	92
	Fluorene	0.434	0.495	88
	Phenanthrene	0.344	0.400	86
	Pyrene	0.870	0.996	87

* Samples S1 and S2 were spiked with diesel and petrol. Spiked values of individual analytes have been estimated based on analysis of the diesel and petrol.

6.2 Measurement Uncertainty Reported by Participants

Participants were asked to report estimates of the expanded uncertainty associated with their results. 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 396 numeric results, 375 results (95%) were reported with an associated expanded MU. Participants used a wide variety of procedures to estimate their uncertainty (Table 3). Some participants reported using NATA MU documentation as their guide; NATA no longer publishes these documents.[11]

Laboratory **4** did not provide uncertainties for any of their reported results. Laboratory **9** did not provide uncertainties for Sample S1 TRH and Sample S2 Total BTEX results (they did provide uncertainties for individual components). Laboratory **11** did not provide an uncertainty for one result (Sample S3 fluorene). These three participants all reported being accredited to ISO/IEC 17025.

Sample S1 TRH results from Laboratories **1**, **5**, **15**, **18** and **22** had no uncertainties, as the results were calculated by the study coordinator by summing the individual hydrocarbon range results reported, and no estimates of the uncertainty were made.

The magnitude of the reported expanded uncertainties was within the range 6.4% to 58% of the reported value. In general, an expanded uncertainty of less than 15% relative may be unrealistically small for the routine measurement of a hydrocarbon pollutant in soil, while an expanded uncertainty of over 50% may be too large and not fit-for-purpose. Of the 375 expanded MUs reported, 22 were less than 15% relative while three were greater than 50% relative.

Participants were also requested to report the coverage factor associated with their uncertainties (Table 3). Ten participants reported a coverage factor of k = 2.

Uncertainties associated with results returning an acceptable *z*-score but an unacceptable E_n-score may have been underestimated.

Laboratory **18** attached estimates of the expanded MU for results reported as less than their limit of reporting (LOR). An estimate of uncertainty expressed as a value cannot be attached to a result expressed as a range.[10]

In some cases, the results were reported with an inappropriate number of significant figures. Including too many significant figures may inaccurately reflect the precision of measurements. 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 2.7281 ± 0.8184 mg/kg, it is better to report this result as 2.73 ± 0.82 mg/kg.[10]

6.3 z-Score

Target SDs equivalent to 15% CV were used to calculate *z*-scores. CVs predicted by the Thompson-Horwitz equation,[8] the between-laboratory CVs and target SDs (as PCV) in this study are presented for comparison in Table 27.

Table 27 Comparison of Thompson-Horwitz CVs, Between-Laboratory CVs, Target SDs

Sample	Analyte	Assigned Value *(Robust Average)* (mg/kg)	Thompson-Horwitz CV[a] (%)	Between-Laboratory CV[b] (%)	Target SD (as PCV) (%)
S1	>C10-C16	937	5.7	16	15
	>C16-C34	1730	5.2	13	15
	>C34-C40	242	7	21	15
	TRH	2910	4.8	12	15
S2	C6-C10	*(2810)*	4.8	24	Not Set
	Benzene	*(60)*	8.6	47	Not Set
	Toluene	369	6.6	20	15
	Ethylbenzene	82.9	8.2	18	15
	Xylenes	395	6.5	11	15
	Total BTEX	870	5.8	20	15
S3	Acenaphthene	2.49	14	16	15
	Acenaphthylene	1.56	15	16	15
	Anthracene	1.23	16	14	15
	Benz[*a*]anthracene	0.645	17	14	15
	Benzo[*a*]pyrene	0.723	17	12	15
	Chrysene	0.548	18	12	15
	Fluoranthene	0.370	19	14	15
	Fluorene	0.434	18	12	15
	Phenanthrene	0.344	19	15	15
	Pyrene	0.870	16	12	15

[a] Calculated from the assigned value (robust average).

[b] Robust between-laboratory CV with outliers removed, if applicable.

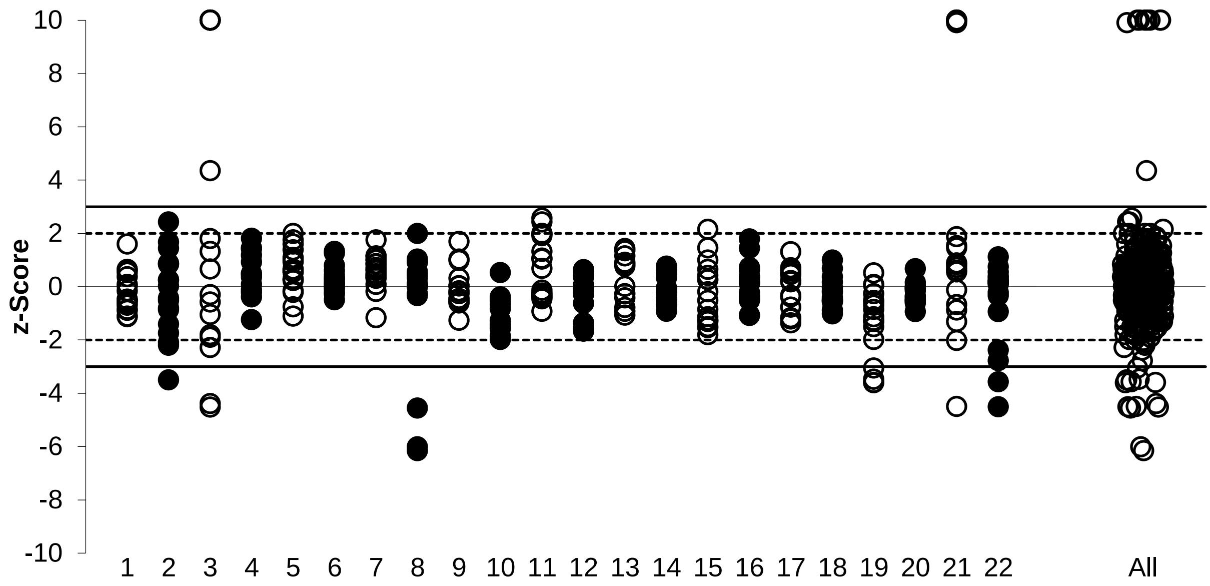
To account for possible low bias in the consensus values due to participants using inefficient analytical or extraction techniques, a total of three *z*-scores were adjusted across the following: Sample S2 toluene, and Sample S3 acenaphthene and acenaphthylene. A maximum acceptable result was set as the spiked value plus two target SDs of the spiked value. 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 participants reporting results close to the spiked value were not penalised. *z*-Scores for results higher than the maximum acceptable result and *z*-scores less than 2.0 were left unaltered.

Of 359 results for which *z*-scores were calculated, 327 (91%) returned a score of $|z| \leq 2.0$, indicating an acceptable performance.

Laboratories **1**, **2**, **5**, **6**, **7**, **8**, **15**, **16**, **19**, **21** and **22** reported numeric results for all 18 scored analytes. Of these participants, Laboratories **1**, **5**, **6**, **7** and **16** returned acceptable *z*-scores for all 18 scored analytes.

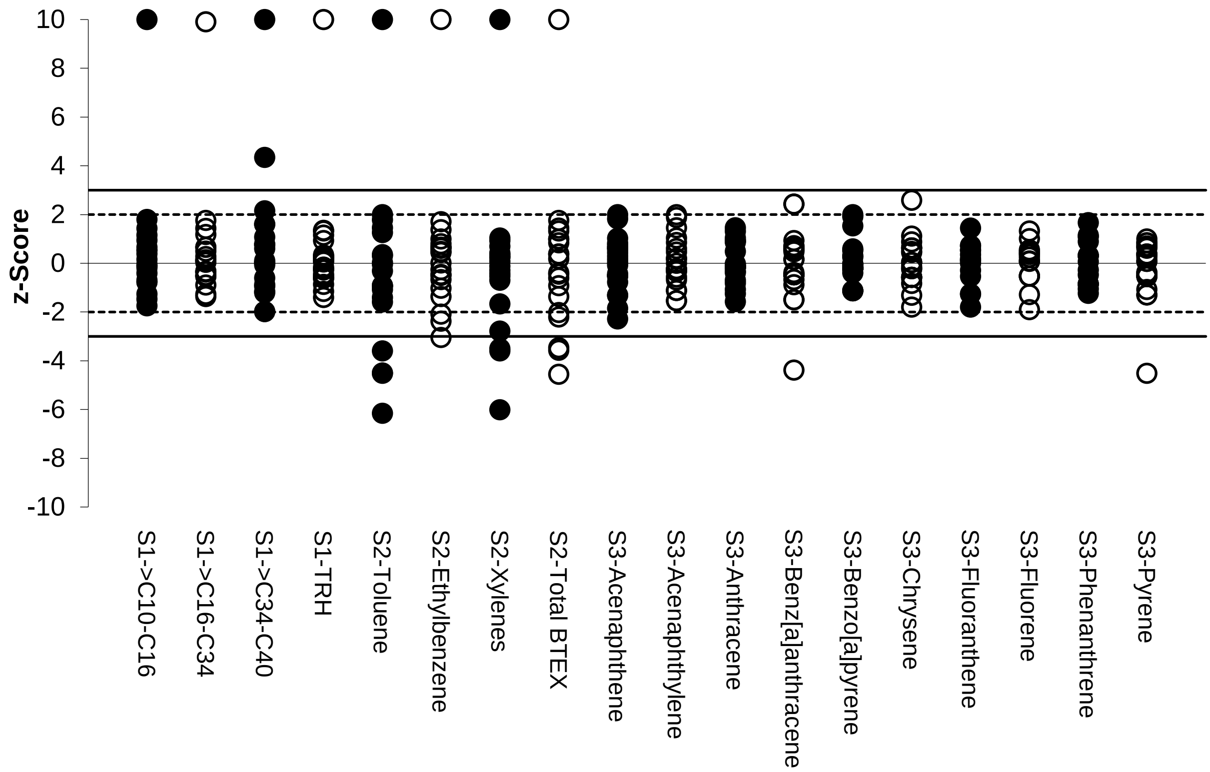
Nine participants received acceptable *z*-scores for all scored analytes that they reported results for: Laboratories **9** (15), **12** (15), **14** (15), **17** (15), **18** (15), **20** (15), **13** (12), **10** (14) and **4** (13).

The dispersal of participants' *z*-scores is presented graphically by laboratory in Figure 22 and by analyte in Figure 23.



z-Scores greater than 10.0 have been plotted at 10.0.

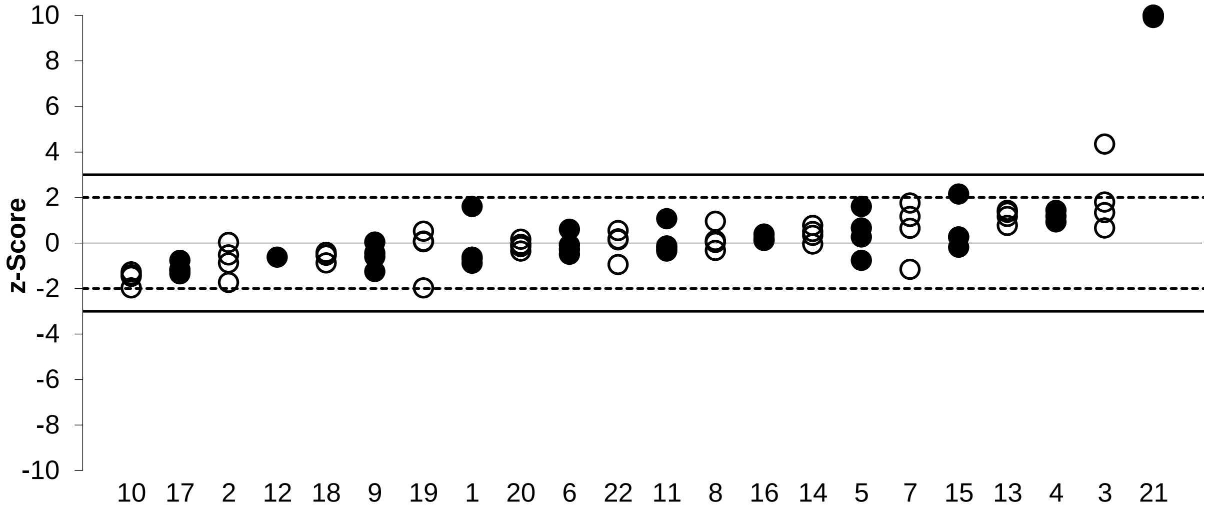
Figure 22 *z*-Score Dispersal by Laboratory



z-Scores greater than 10.0 have been plotted at 10.0.

Figure 23 *z*-Score Dispersal by Analyte

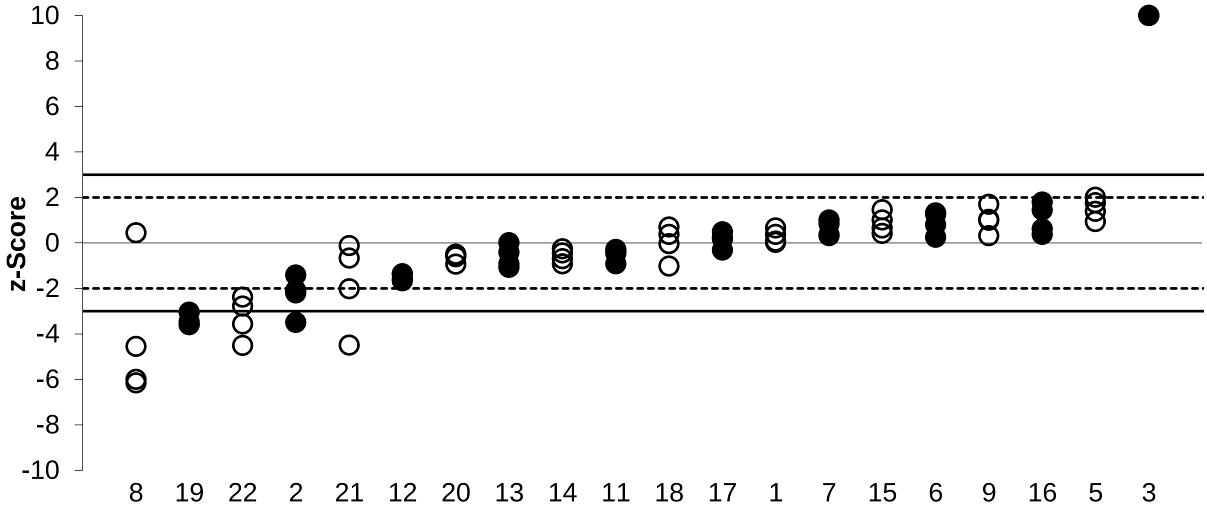
Participants' *z*-scores for Sample S1 TRH only are presented in Figure 24. A trend of questionable or unacceptable *z*-scores on one side of the zero line may indicate laboratory bias for TRH measurements.



z-Scores greater than 10.0 have been plotted at 10.0.

Figure 24 Sample S1 TRH *z*-Score Dispersal by Laboratory

Participants' *z*-scores for Sample S2 BTEX only are presented in Figure 25. A trend of questionable or unacceptable *z*-scores on one side of the zero line may indicate laboratory bias for BTEX measurements; in particular, laboratories whose results consistently return questionable or unacceptable *z*-scores below the zero line may have an inefficient extraction process for BTEX. As the ratio of the assigned value to the spiked value was 51% for Total BTEX, participants reporting results with higher acceptable *z*-scores may have more efficient methodologies.



z-Scores greater than 10.0 have been plotted at 10.0.

Figure 25 Sample S2 BTEX *z*-Score Dispersal by Laboratory

Participants' *z*-scores for Sample S3 PAHs only are presented in Figure 26. A trend of questionable or unacceptable *z*-scores on one side of the zero line may indicate laboratory bias for PAHs measurements; in particular, laboratories whose results consistently return questionable or unacceptable *z*-scores below the zero line may have an inefficient extraction process for PAHs. As the ratios of the assigned values to the spiked values ranged from 48% to 92%, participants reporting results with higher acceptable *z*-scores may have more efficient methodologies.

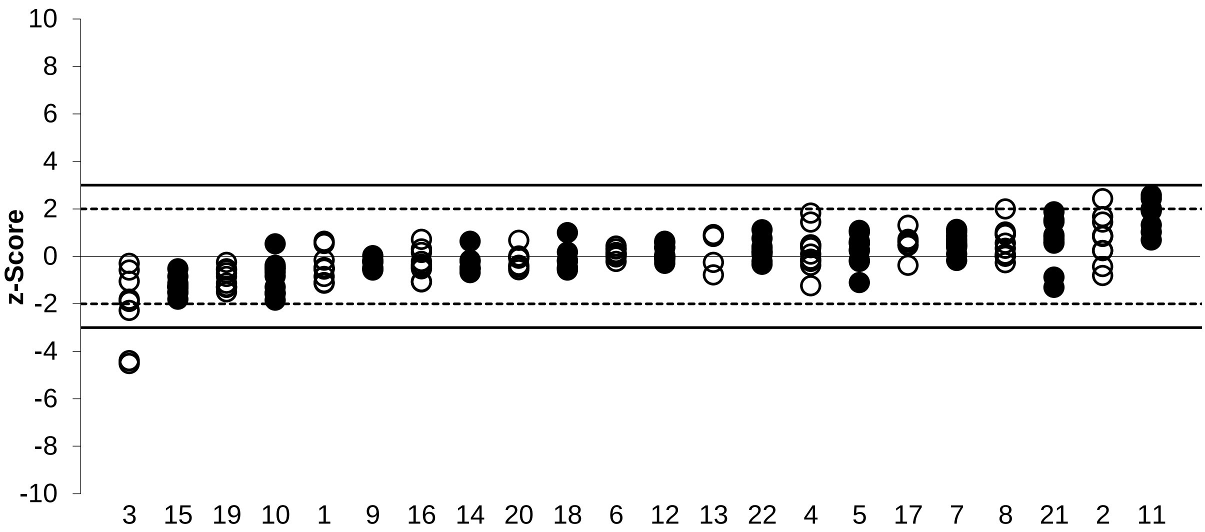


Figure 26 Sample S3 PAHs *z*-Score Dispersal by Laboratory

6.4 E_n-Score

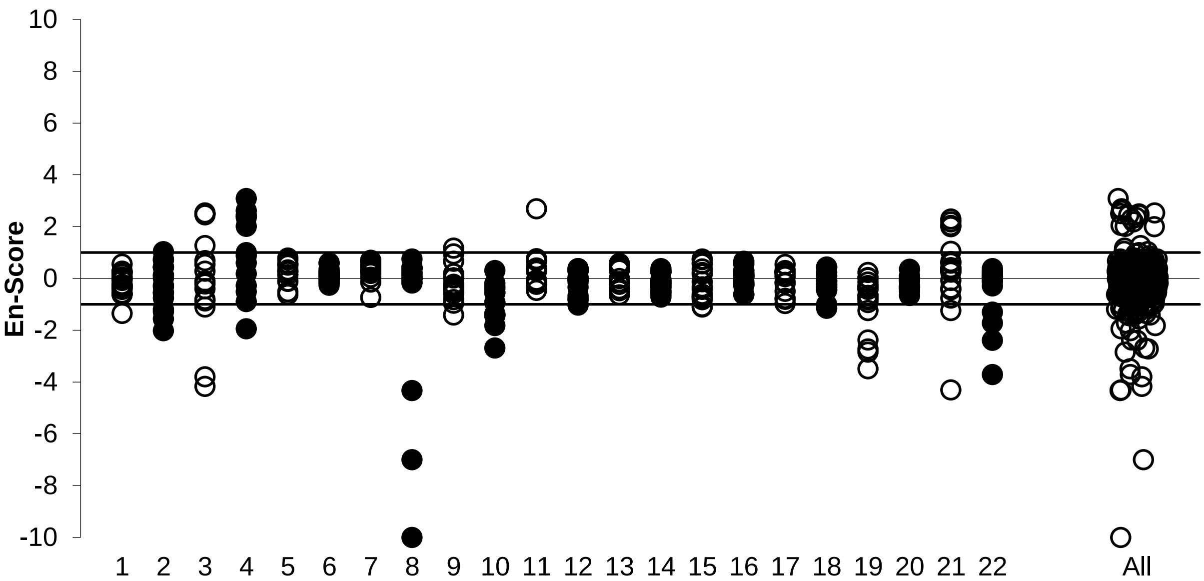
Where a laboratory did not report an expanded uncertainty with a result, an expanded uncertainty of zero (0) was used to calculate the E_n-score. For results for which *z*-scores were adjusted as discussed in Section 6.3 z-Score, no E_n-score has been reported.

Of 356 results for which E_n-scores were calculated, 301 (85%) returned an acceptable score of $|E_n| < 1.0$, indicating agreement of the participant's result with the assigned value within their respective uncertainties.

Laboratories **6**, **7** and **16** returned acceptable E_n-scores for all 18 scored analytes. Laboratory **5** also reported numeric results for all analytes, however as one of their result's *z*-score was adjusted as described above, E_n-scores were only calculated for 17 of their results; this participant returned acceptable E_n-scores for all 17 results.

Four participants received acceptable E_n-scores for all scored analytes they reported results for: Laboratories **14** (15), **17** (15), **20** (15) and **13** (12).

The dispersal of participants' E_n-scores is presented graphically by laboratory in Figure 27.



E_n-Scores lower than -10.0 have been plotted at -10.0.

Figure 27 E_n-Score Dispersal by Laboratory

6.5 False Negatives

Table 28 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, participants reporting a 'less-than' result (< *x*) when the assigned value was higher than their LOR, or laboratories that did not report anything.

Table 28 False Negatives

Lab. Code	Sample	Analyte	Assigned Value (mg/kg)	Spiked Value (mg/kg)	Result (mg/kg)
3	S3	Anthracene	1.23	2.01	<0.01
		Benzo[*a*]pyrene	0.723	1.5	<0.01
13	S3	Benz[*a*]anthracene	0.645	0.788	<0.5
		Benzo[a]pyrene	0.723	1.5	<0.5
		Chrysene	0.548	0.596	<0.5

6.6 Participants' Analytical Methods

A variety of analytical methods were used by participants in this study (Appendix 4).

Results that were removed from all statistical calculations in Section 5 have also been removed from all discussion in this section. Where charts refer to n = *x*, this corresponds to *x* number of participants using that methodology.

TRH

Sample S1 was a 50 g soil sample. Participants used a sample size between 4 g and 10 g for TRH analysis, with most participants using 10 g. A plot of results against sample mass used for analysis is presented in Figure 28.

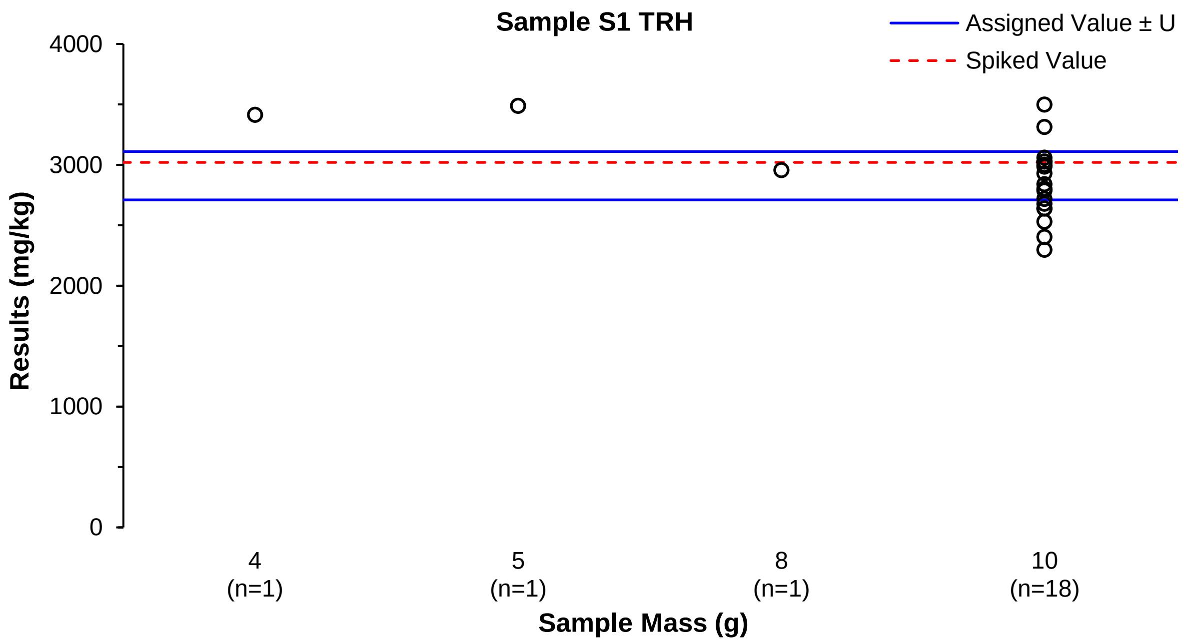


Figure 28 Sample S1 TRH Results vs Sample Mass Used for Analysis

Participants reported using either solid-liquid extraction (SLE) or sonication, with dichloromethane (DCM), acetone (ACE), hexane (HEX), or combinations of these as the extraction solvent(s). Four participants reported a silica clean-up step. All participants used gas chromatography (GC) coupled to flame ionisation detection (FID) for analysis.

A plot of results against methodology for Sample S1 TRH is presented in Figure 29. Methodologies are listed in order of reported extraction technique, extraction solvent(s), clean-up and instrument. The most common methodology used to analyse TRH in this study was SLE with DCM/ACE as the extraction solvent, with no clean-up and using GC-FID for analysis.

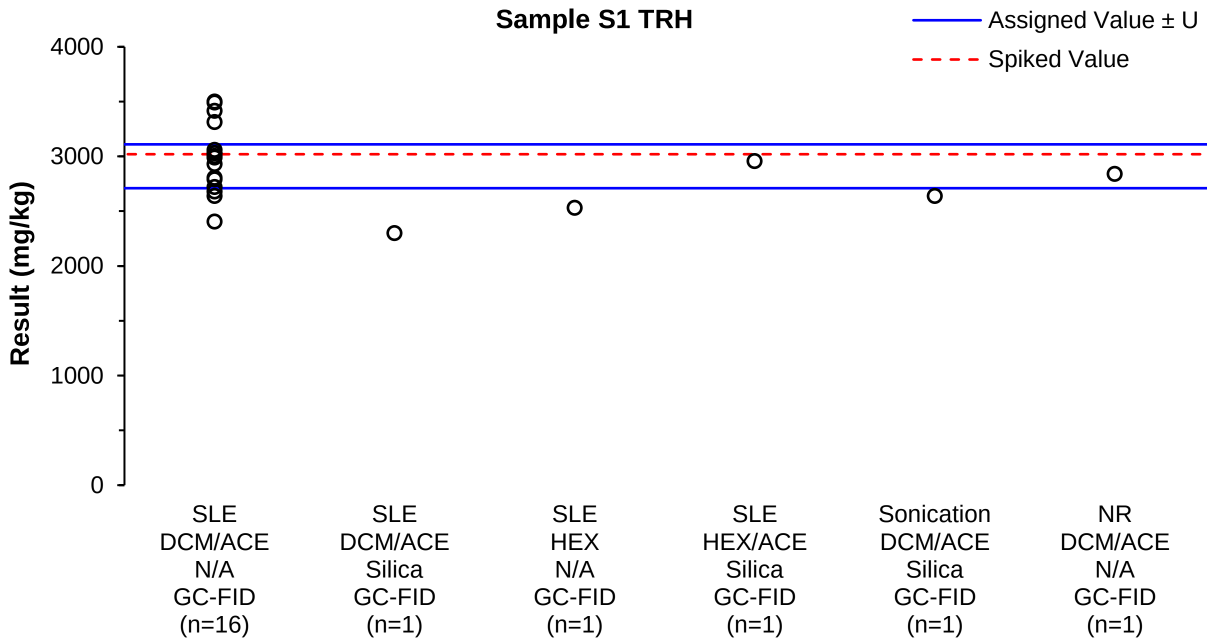


Figure 29 Sample S1 TRH Results vs Methodology

BTEX

Sample S2 was a 50 g soil sample. Participants used a sample size between 1.1 g and 14 g for BTEX analysis, with most participants using either 5 g or 10 g. A plot of results against sample mass used for analysis is presented in Figure 30.

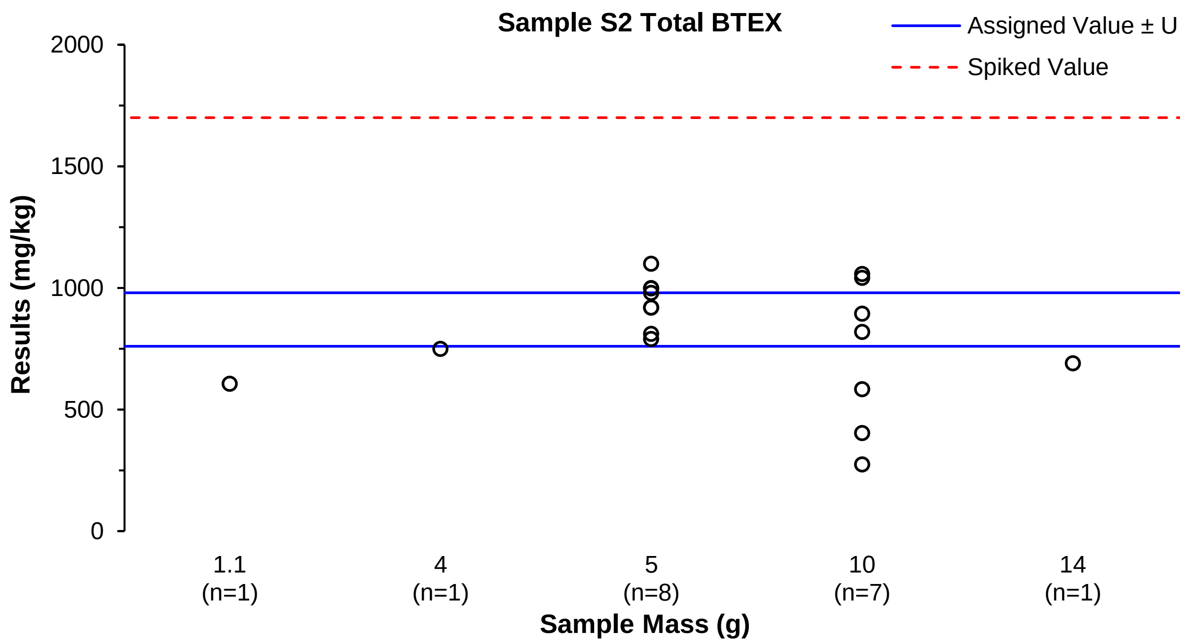


Figure 30 Sample S2 Total BTEX Results vs Sample Mass Used for Analysis

Extraction techniques reported by participants included SLE and sonication. All participants reporting numeric results for this sample reported using methanol (MeOH) as their extraction solvent, except for one participant who reported using DCM/ACE. No participant reported a clean-up step. Two of those participants used headspace (HS) GC coupled to mass spectrometry (MS) or tandem mass spectrometry (MS/MS), while all other participants used purge and trap (P&T) GC-MS(/MS).

A plot of results and methodology for Total BTEX in Sample S2 is presented in Figure 31. Methodologies are listed in order of extraction technique, extraction solvent(s), clean-up and instrument. The most common methodology used to analyse BTEX in this study was SLE with MeOH, using P&T GC-MS for analysis.

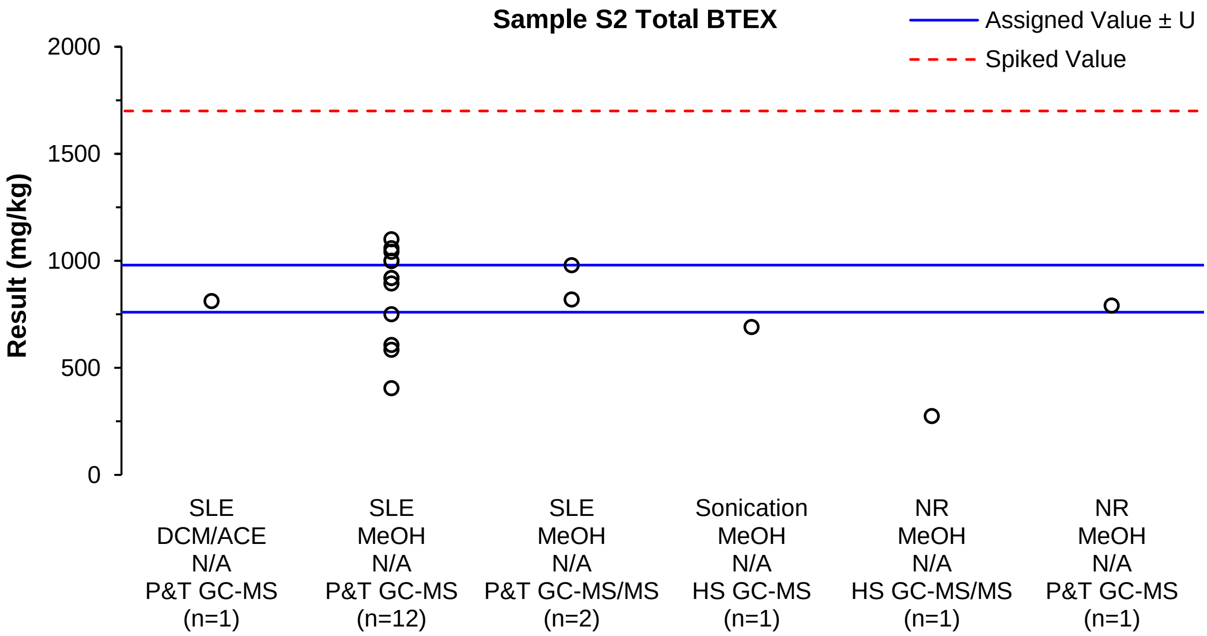


Figure 31 Sample S2 Total BTEX Results vs Methodology

PAHs

Sample S3 was a 50 g soil sample. Participants used a sample size between 2 g and 10 g for PAHs analysis, with the majority of participants using 10 g. A plot of *z*-scores against sample mass used for analysis is presented in Figure 32.

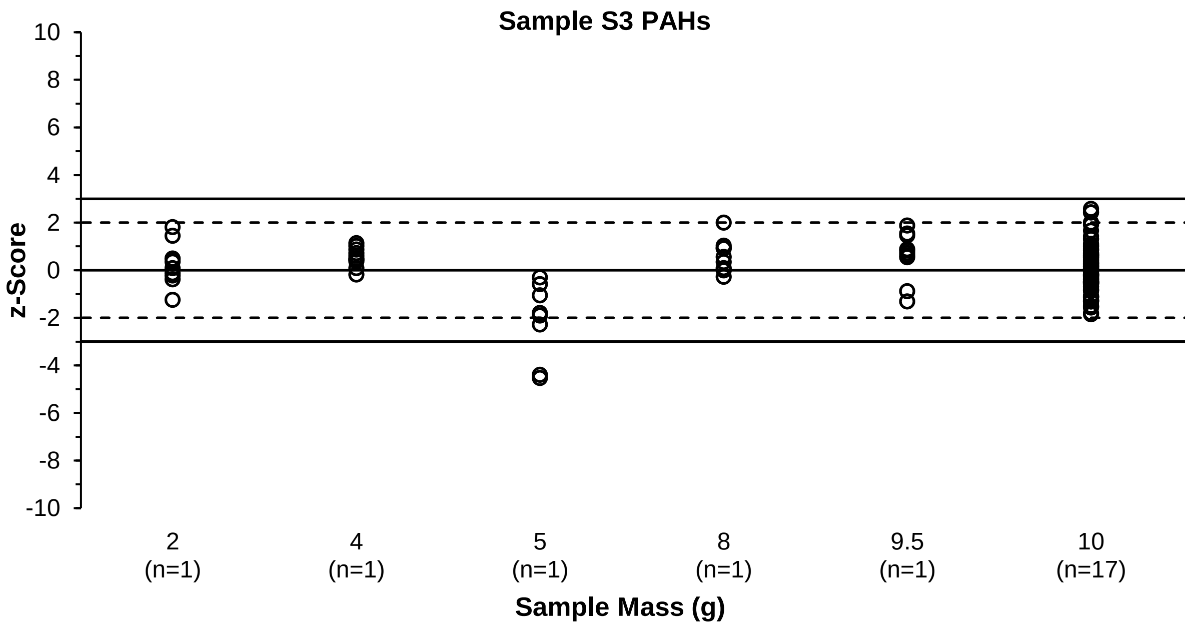


Figure 32 Sample S3 PAHs *z*-Scores vs Sample Mass Used for Analysis

Participants reported using either SLE or sonication, with DCM, ACE, HEX, ethyl acetate (EtOAc), or combinations of these as the extraction solvent(s). One participant reported using a silica clean-up step. All participants used GC-MS(/MS) for analysis.

A plot of *z*-scores obtained and methodology used for the PAHs in Sample S3 is presented in Figure 33. Methodologies are listed in order of extraction technique, extraction solvent(s), clean-up and instrument. The most common methodology used to analyse PAHs in this study was SLE with DCM/ACE as the extraction solvent, with no clean-up and using GC-MS for analysis.

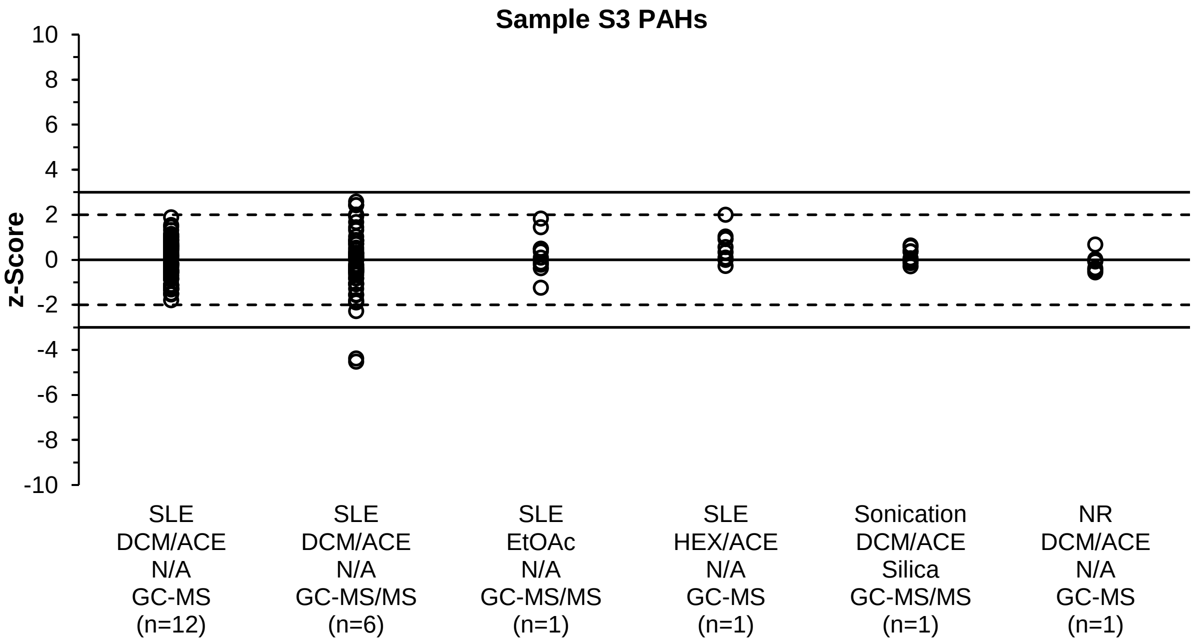


Figure 33 Sample S3 PAHs *z*-Scores vs Methodology

6.7 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 their analysis.

Ten participants reported using certified standards, three participants reported using matrix reference materials, and two participants reported using both. The following were reported by participants:

- NMI MX015
- Accustandard, e.g. DRH-004S-R1-5X
- o2si
- Sigma Aldrich, e.g. CRM358, CRM143
- ISO/IEC 17034 traceable standards
- ISO/IEC 17025 compliant standards

These materials may or may not meet the internationally recognised definition of a Certified Reference Material:

> '***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]

6.8 Summary of Participants' Results and Performances

Summaries of participants' results and performances for scored analytes in this PT study are presented in Tables 29 and 30, and Figure 34.

Table 29 Summary of Participants' Results (Samples S1 and S2)*

Lab. Code	Sample S1				Sample S2			
	>C10-C16	>C16-C34	>C34-C40	TRH	Toluene	Ethylbenzene	Xylenes	Total BTEX
AV	937	1730	242	2910	369	82.9	395	870
SV	1080	1700	239	3020	723	138	617	1700
1	840	1500	300	2640	370	91	400	920
2	692.39	1596	243	2531.39	290.95	56.93	187.85	583.44
3	1190	1900	400	3490	2687	726	3255	7060
4	1138	2037	NR	3314	NT	NT	NT	NT
5	830	1900	300	3030	480	100	450	1100
6	923.7	1601.8	263.8	2789	439	92.8	410	1042
7	1028	2187	200	3415	389	95.3	414	979.7
8	1070.3	1644.7	242.8	2957	28	88.5	39	275.15
9	760	1740	220	2720	386	104	456	999
10	730	1400	170	2300	NR	NR	NR	NR
11	919	1640	281	2804	318	77.2	378	819
12	NR	NR	NR	2640	281	66	296	691
13	1100	2100	270	3500	310	83	370	750
14	930	1860	270	3060	318	79.7	354	812
15	910	1800	320	3030	450	91	420	1000
16	992	1765	246	3003	468	90.5	417	1058
17	828.5	1377	198	2404	351.1	88.95	408.4	893.9

Lab. Code	Sample S1				Sample S2			
	>C10-C16	>C16-C34	>C34-C40	TRH	Toluene	Ethylbenzene	Xylenes	Total BTEX
18	880	1590	210	2680	367	70.2	436	920
19	1010	1750	170	2930	170	45	182	417
20	960	1640	240	2840	317	76.7	364	790
21	2671	4301	707	7680	119.9	74.6	388	606
22	1014	1766	208	2988	119.5	53.3	230	404

* All values are in mg/kg. Shaded cells are results which returned a questionable or unacceptable *z*-score. AV = Assigned Value; SV = Spiked Value.

Table 30 Summary of Participants' Results (Sample S3)*

Lab. Code	Sample S3									
	Acenaphthene	Acenaphthylene	Anthracene	Benz[*a*]anthracene	Benzo[*a*]pyrene	Chrysene	Fluoranthene	Fluorene	Phenanthrene	Pyrene
AV	2.49	1.56	1.23	0.645	0.723	0.548	0.370	0.434	0.344	0.870
SV	3.19	2.78	2.01	0.788	1.50	0.596	0.403	0.495	0.400	0.996
1	2.3	1.3	1.2	0.6	0.6	0.6	0.4	0.4	0.3	0.8
2	2.33	1.76	1.08	0.88	0.75	0.62	0.45	0.45	0.43	0.98
3	1.64	1.49	<0.01	0.22	<0.01	0.5	0.27	0.31	0.29	0.28
4	3.17	1.90	1.32	0.69	0.71	0.53	0.39	0.44	0.28	0.82
5	2.6	1.3	1.2	0.7	0.7	0.6	0.4	0.5	0.4	0.9
6	2.651	1.509	1.224	0.66	0.757	0.551	0.377	0.439	0.36	0.881
7	2.87	1.7	1.39	0.715	0.703	0.585	0.375	0.465	0.403	0.92
8	2.7	1.8	1.4	0.735	0.94	0.55	0.37	0.44	0.33	0.915
9	2.5	1.5	1.2	0.6	0.7	0.5	<0.5	<0.5	<0.5	0.8
10	1.80	1.20	0.94	0.58	0.78	0.48	0.34	0.35	0.32	0.82
11	3.69	2.04	1.42	0.88	0.93	0.76	<0.5	0.52	<0.5	0.96

Lab. Code	Sample S3									
	Acenaphthene	Acenaphthylene	Anthracene	Benz[*a*]anthracene	Benzo[*a*]pyrene	Chrysene	Fluoranthene	Fluorene	Phenanthrene	Pyrene
12	2.73	1.56	1.22	0.7	0.72	0.536	0.354	0.458	0.347	0.92
13	2.2	1.5	1.4	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5	0.98
14	2.3	1.4	1.2	0.6	0.7	0.6	<0.5	<0.5	<0.5	0.8
15	2	1.2	1	0.5	0.6	0.4	0.3	0.4	0.3	0.7
16	2.55	1.48	1.03	0.60	0.68	0.53	0.41	0.40	0.36	0.73
17	2.7281	1.6673	1.4711	0.6085	0.7867	0.5871	<0.5	<0.5	<0.5	0.9624
18	2.3	1.6	1.2	0.6	0.7	0.5	<0.5	<0.5	<0.5	1
19	2	1.5	1.1	0.5	0.6	0.5	0.3	0.4	0.3	0.7
20	2.46	1.43	1.16	0.71	0.72	0.55	<0.50	<0.50	<0.50	0.81
21	2.8	2	1.5	0.56	0.89	0.44	0.41	0.47	0.39	0.95
22	2.77	1.61	1.17	0.66	0.7	0.64	0.38	0.46	0.36	0.88

* All values are in mg/kg. Shaded cells are results which returned a questionable or unacceptable *z*-score. AV = Assigned Value; SV = Spiked Value.

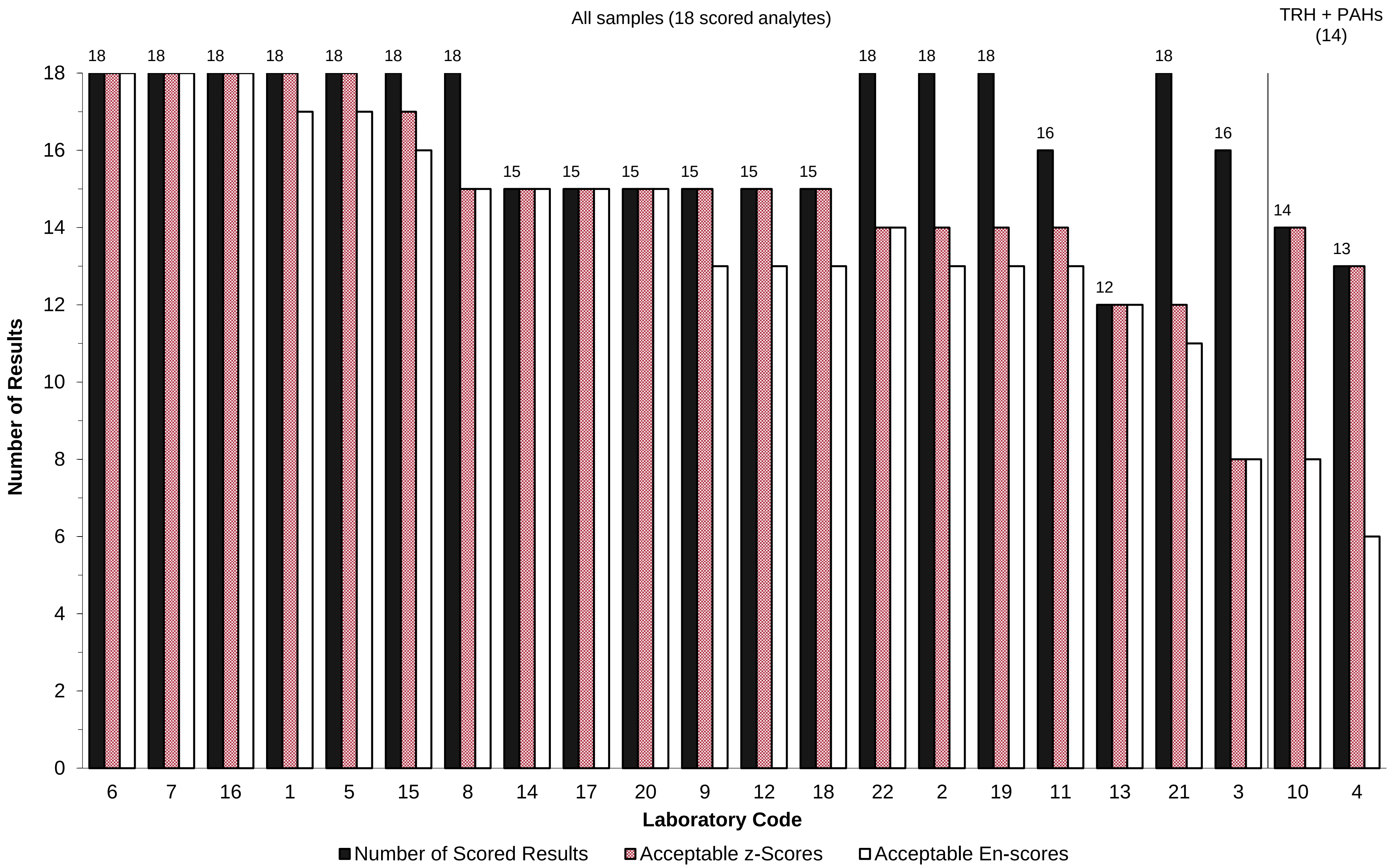


Figure 34 Summary of Participants' Performance

6.9 Comparison with Previous Hydrocarbons in Soil PT Studies

To enable direct comparison with results from previous Hydrocarbons in Soil PT studies, the target SD used to calculate *z*-scores has been kept constant at 15% PCV.

Individual performance history reports are emailed to each participant at the end of each study; the consideration of *z*-scores for an analyte over time provides much more useful information than a single *z*-score. Over time, laboratories should expect at least 95% of their scores to lie within the range $|z| \leq 2.0$. Scores in the range $2.0 < |z| < 3.0$ can occasionally occur, however, these should be interpreted in conjunction with the other scores obtained by that laboratory. For example, a trend of *z*-scores on one side of the zero line is an indication of method or laboratory bias.

TRH

A summary of the acceptable performance (presented as a percentage of the total number of scores) obtained by participants for TRH in soil over the last 10 studies (2015 – 2024) is presented in Figure 35. Over this period, the average proportion of acceptable *z*-scores was 93%, and the average proportion of acceptable E_n-scores was 76%.

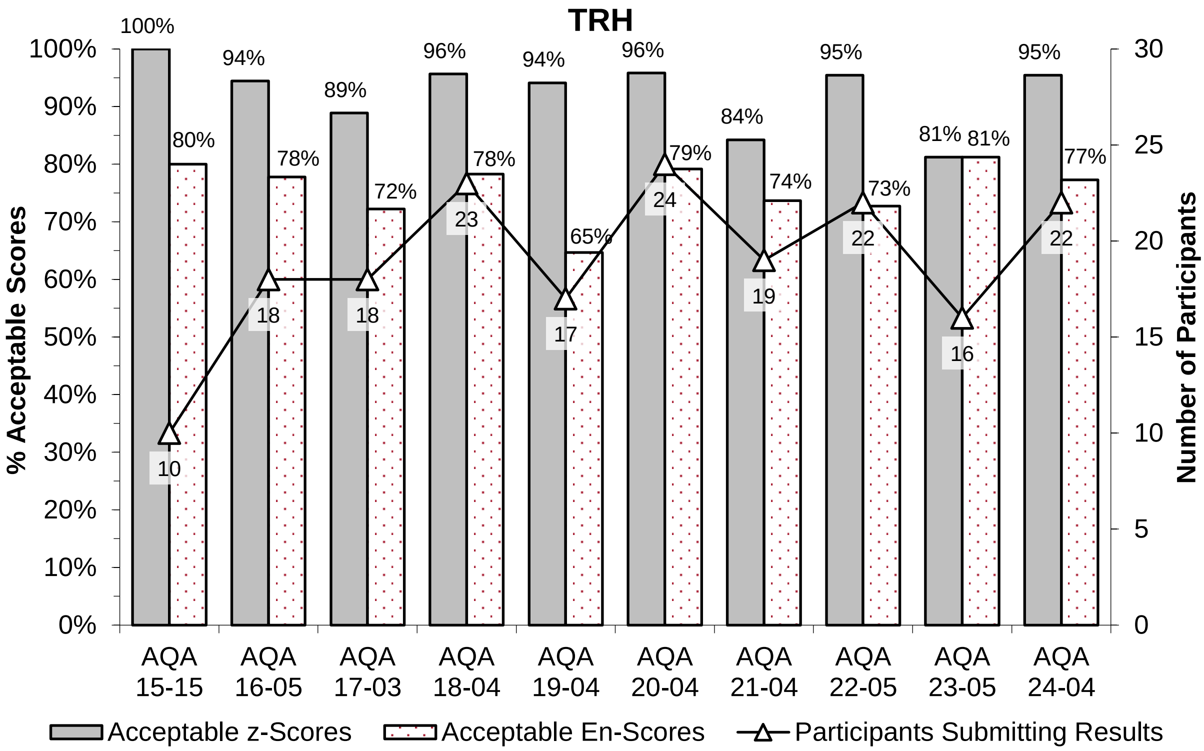


Figure 35 Participants' Performance for TRH in Hydrocarbons in Soil PT Studies

Total BTEX

A summary of the acceptable performance (presented as a percentage of the total number of scores) obtained by participants for Total BTEX in soil over the last 10 studies (2015 – 2024) is presented in Figure 36. Over this period, the average proportion of acceptable *z*-scores was 85%, and the average proportion of acceptable E_n-scores was 81%. The spiked and assigned values were similar to the immediate previous study, AQA 23-05, however participants returned a lower proportion of acceptable *z*-scores and E_n-scores in this study.

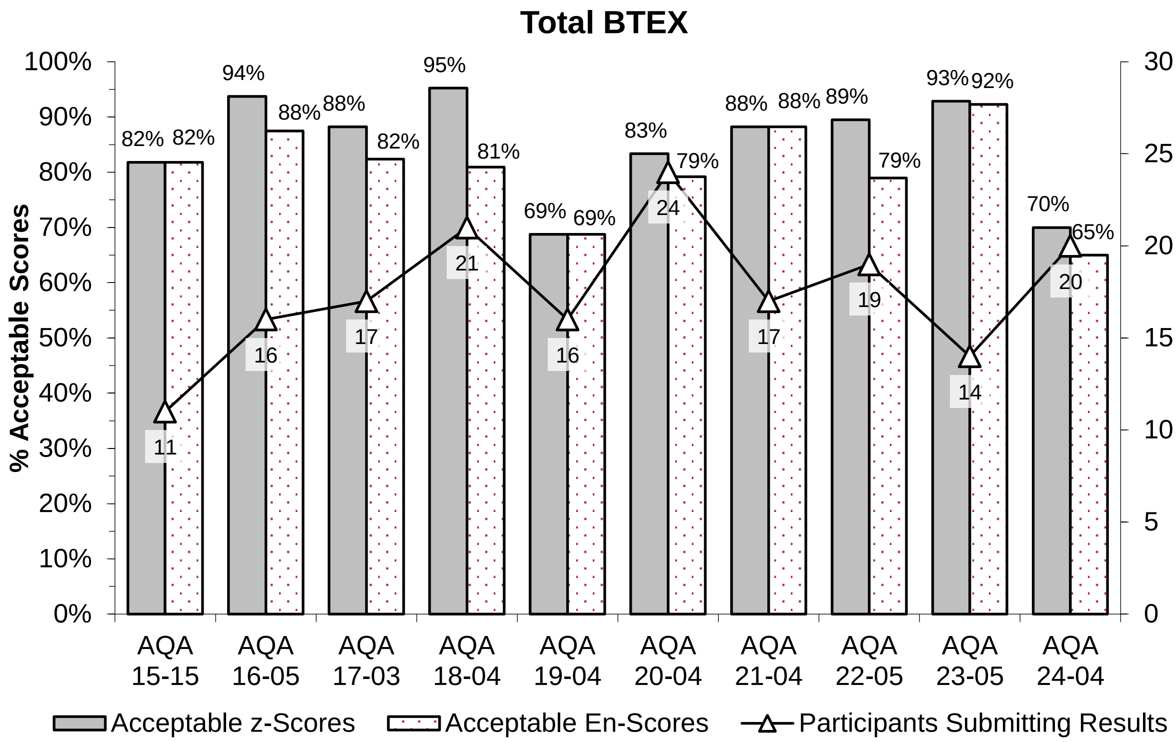


Figure 36 Participants' Performance for Total BTEX in Hydrocarbons in Soil PT Studies

PAHs

PAHs in soil was first introduced in NMI PT studies in 2016. A summary of the acceptable performance (presented as a percentage of the total number of scores) obtained by participants for PAHs in soil over the last 9 studies (2016 – 2024) is presented in Figure 37. Over this period, the average proportion of acceptable *z*-scores was 91%, and the average proportion of acceptable E_n-scores was 87%.

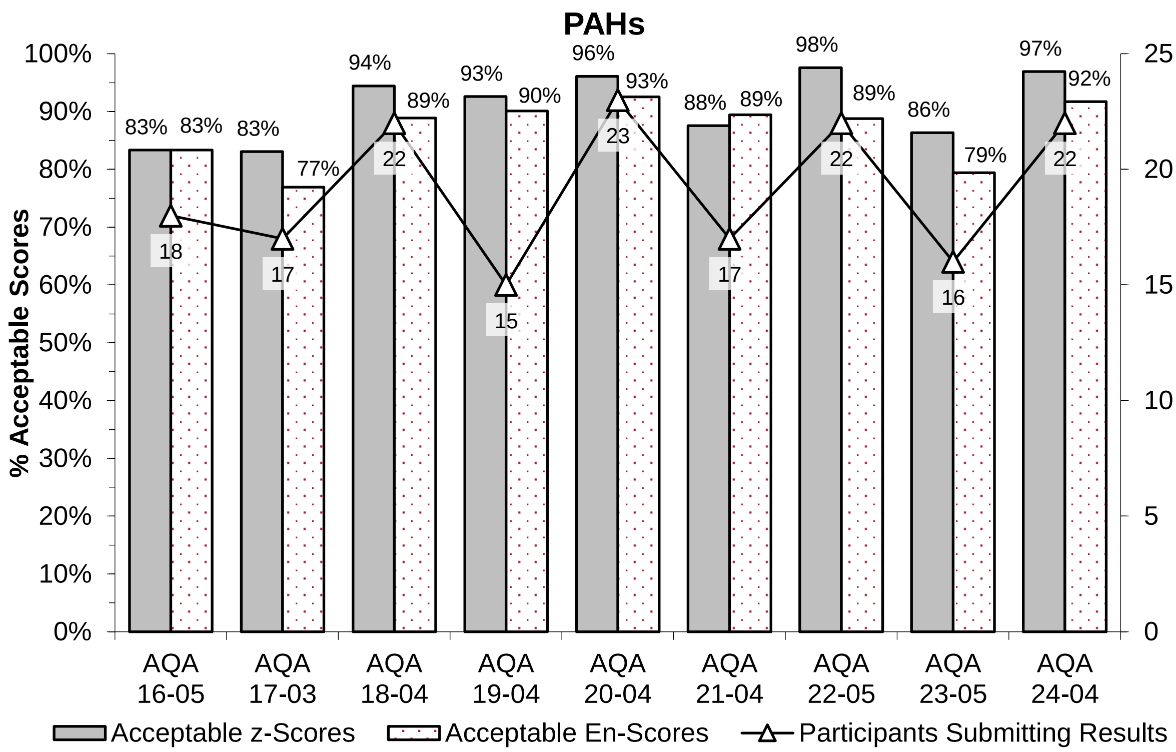


Figure 37 Participants' Performance for PAHs in Hydrocarbons in Soil PT Studies

A plot of the robust average expressed as a percentage of the spiked value for PAHs in topsoil since 2016 is presented in Figure 38. Results from samples with other soil matrices have not been included as it has been previously seen that the nature of the soil matrix can substantially affect the recovery of some analytes.[13]

This was the first study to include acenaphthene and acenaphthylene as spiked analytes.

For all spiked PAHs in this study, the robust averages were lower than the spiked values, consistent with previous studies. Throughout NMI Hydrocarbons in Soil PT studies, anthracene and benzo[*a*]pyrene have consistently had lower recoveries, averaging 50% and 46% respectively for the robust average to spiked value. In this study, it was seen that acenaphthylene also had a relatively low robust average to spiked value ratio of 56%. The average of robust averages to spiked values for the other PAHs were higher, ranging from 78% to 86%.

For this study, the ratio of robust averages to spiked values were either similar to or higher than the average of previous studies.

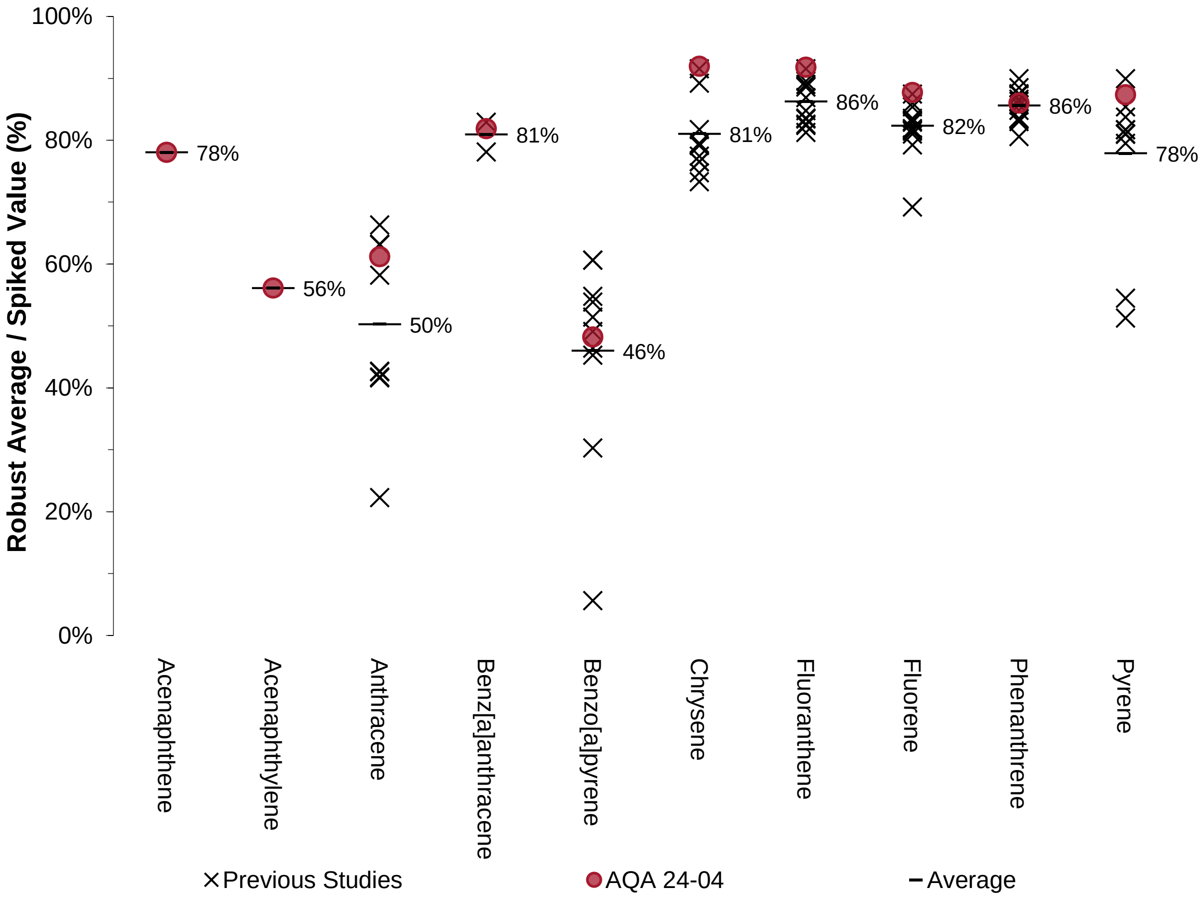


Figure 38 Recoveries of PAHs in Topsoil for Hydrocarbons in Soil PT Studies

As discussed in Section 6.2, it is a requirement of ISO/IEC 17025 that laboratories report their uncertainties.[9] Figure 39 presents a summary of the relative uncertainties as reported by participants over the last 10 studies (2015–2024). Over this time period, the vast majority of numeric results were reported with uncertainties (95%), with on average 89% of participants reporting that they were accredited to ISO/IEC 17025. Most participants over this time period reported relative expanded uncertainties between 15% and 50%, however around 15% of relative uncertainties were outside this range, and may have been unrealistically small or too large and not fit-for-purpose.

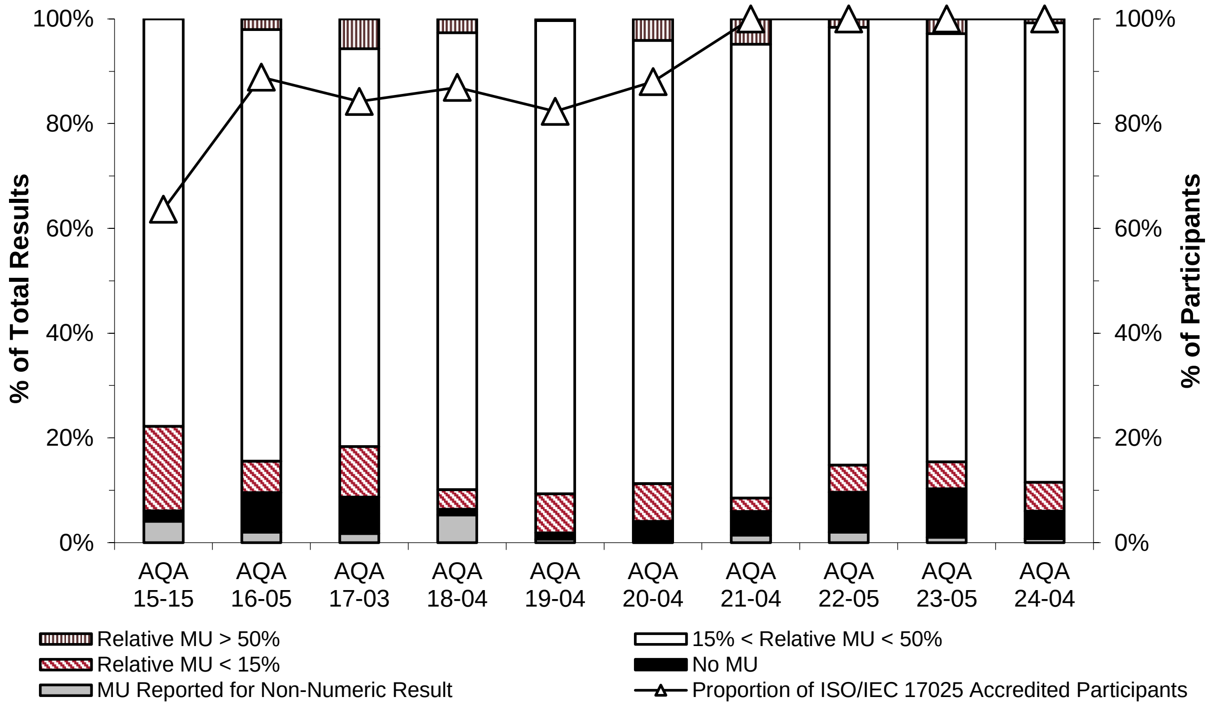


Figure 39 Summary of Participants' Relative Uncertainties for NMI Hydrocarbons in Soil PT Studies

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 the competence of proficiency testing providers*.

[2] NMI, 2024, *Study Protocol for Proficiency Testing*, viewed May 2024, <https://www.industry.gov.au/sites/default/files/2020-10/cpt_study_protocol.pdf>.

[3] NMI, 2024, *Chemical Proficiency Testing Statistical Manual*, viewed May 2024, <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 (Assessment of Site Contamination) Measure 1999 as amended 2013, viewed May 2024, <https://www.legislation.gov.au/F2008B00713/latest/text>

[6] Worrall, R.D., 1996, 'Total Petroleum Hydrocarbons in Soil: Storage Stability Study', 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] 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 May 2024, <http://www.eurachem.org/images/stories/Guides/pdf/QUAM2012_P1.pdf>.

[11] NATA, 2020, *Update to Measurement Uncertainty Resources*, viewed May 2024, <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.

[13] NMI, 2018, *Proficiency Test Report AQA 18-04 Hydrocarbons in Soil*.

APPENDIX 1 SAMPLE PREPARATION

A1.1 Diesel Fuel 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 was used in previous NMI Hydrocarbon PTs.

A1.2 Test Sample Preparation

Two soils were used as the starting matrices in this study. The first soil was topsoil from a local suburban garden, which had been dried at 120 °C for at least two hours and then sieved, collecting the 180 μm to 355 μm fraction; this soil was used to prepare Sample S1. The second soil was topsoil purchased from a local supplier described as Menangle topsoil, which had been dried at 120 °C for at least two hours and then sieved, collecting the 180 μm to 355 μm fraction; this soil was used to prepare Samples S2 and S3.

Sample S1: Into a stainless steel pot, dried and sieved soil (2300.3 g) was placed. Dichloromethane was added to moisten the soil. An aliquot of sparged diesel (6.715 g) was added by transferring 8 aliquots of 1.000 mL of diesel into the drum. The mass of diesel was calculated using a density of 0.83936 g/mL. In addition, 3.0 mL of PENRITE INDUS PRO HYDRAULIC 68 was added. The mixture was thoroughly stirred and the solvent was allowed to evaporate. The mixture was divided into 50 g portions using a Retsch PT 100 sample divider and packed into screw-capped glass jars, labelled and stored in a refrigerator.

Sample S2: Dried, sieved soil (4008.2 g) was placed in a stainless steel drum with a clamp-locked lid. The drum and soil were cooled in a freezer overnight. Diesel (five aliquots of 0.84 mL corresponding to 3.53 g using ρ = 0.83936 g/mL) was added to the soil. Benzene (650 µL) was added to the soil using a positive displacement pipette, followed by unleaded petrol (31.003 g). The drum was sealed and vigorously shaken. The sealed drum was then packed into another large drum and surrounded by cold gel-packs. The drums were then tumbled for 60 minutes on a hoop mixer. The soil was scooped into glass jars, tapped, topped up to minimise the vapour space and sealed. The process of filling the jars was conducted with the drum in an open freezer to minimise the loss of volatiles. The jars were labelled, sealed with Parafilm, shrink-wrapped and stored in a freezer.

Sample S3: Dried, sieved soil (1084.0 g) was placed in a 3 L round bottom flask. Dichloromethane was then added to the soil to allow it to be suspended. Using a Gilson pipette, aliquots of the PAH standard solutions were added to the round bottom flask. The quantity of each standard was calculated using the target final mass of soil after the dilution of the contents of the round bottom flask. To minimise the creation of dust, approximately 20 mL of reagent grade water was added to the flask. The flask was shaken to mix. The solvent was then evaporated using a Büchi rotary evaporator. The bath temperature was set at ambient and gently increased to no more than 40 °C during the evaporation, the condenser temperature at 5 °C and less than 20 kPa of vacuum. After evaporating the dichloromethane, the soil was transferred to a V-mixer and diluted with 1116.2 g of clean soil. The total soil mass was 2200.2 g. The V-mixer was tumbled for two hours. After mixing the soil was divided using a Retsch PT100 sample divider into fifty samples of at least 50 g, placed in screw-capped glass jars, labelled and placed in a refrigerator.

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 results to bottle number for scored analytes are presented in Figures 40 to 51 (solid blue lines correspond to the assigned value ± U for each analyte; results have not been included here if they were excluded from all statistical calculations in Section 5, or if that participant was sent more than one container for that sample). No significant fill order trend was observed.

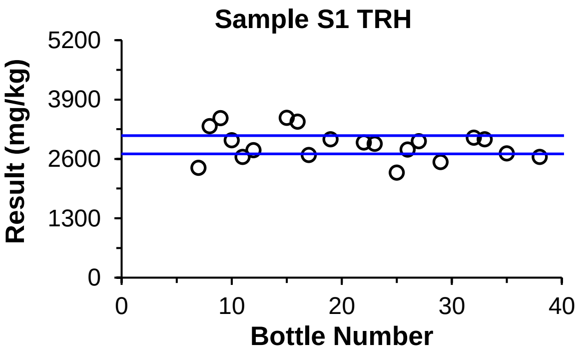


Figure 40 S1 TRH Results vs Bottle Number

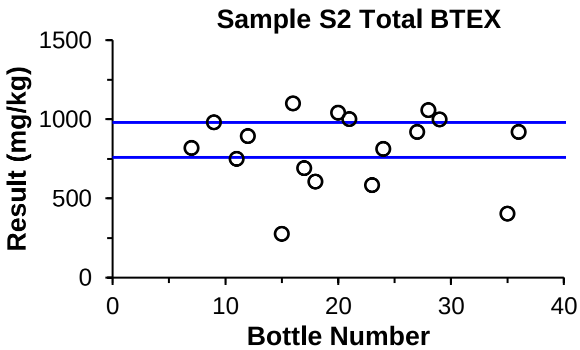


Figure 41 S2 Total BTEX Results vs Bottle Number

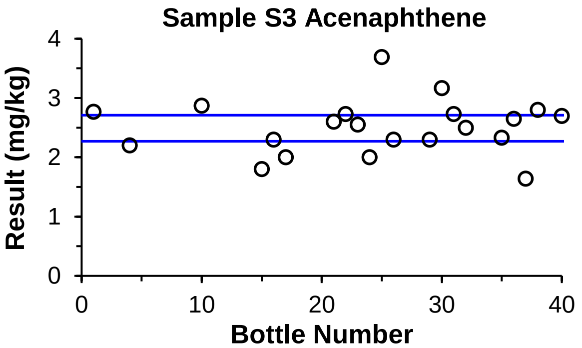


Figure 42 S3 Acenaphthene Results vs Bottle Number

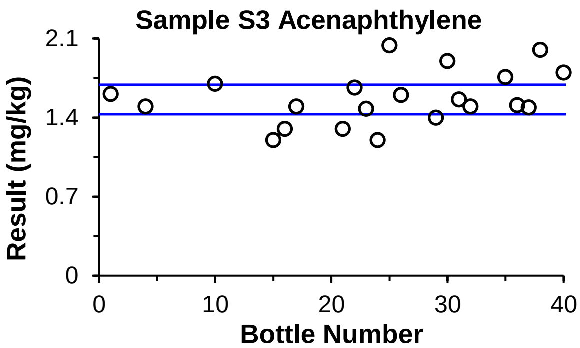


Figure 43 S3 Acenaphthylene Results vs Bottle Number

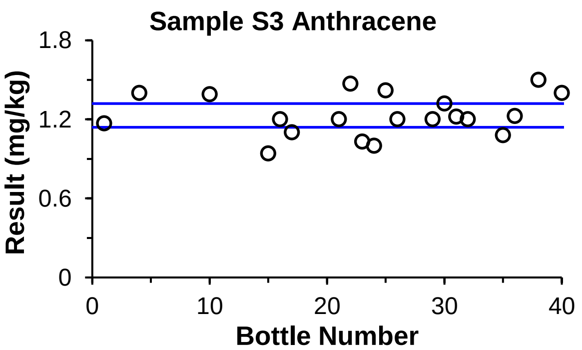


Figure 44 S3 Anthracene Results vs Bottle Number

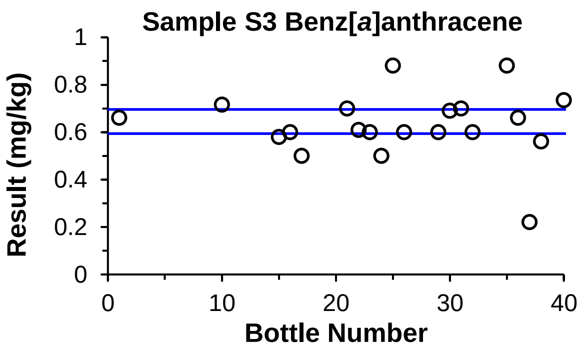


Figure 45 S3 Benz[*a*]anthracene Results vs Bottle Number

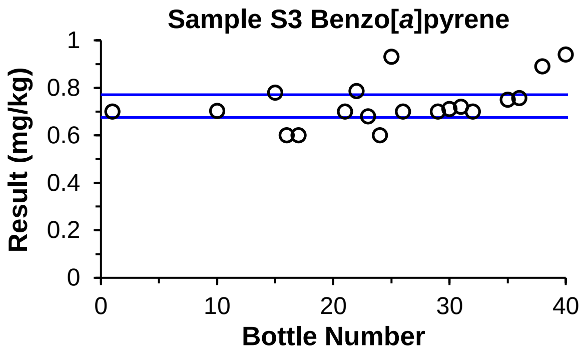


Figure 46 S3 Benzo[*a*]pyrene Results vs Bottle Number

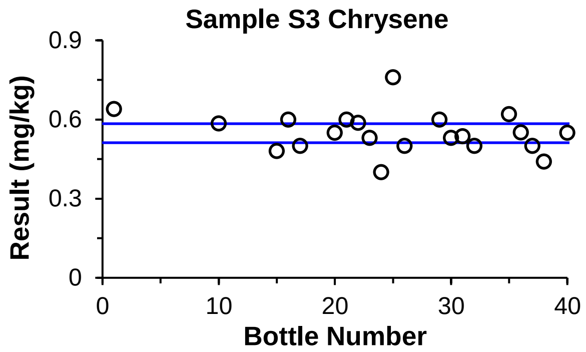


Figure 47 S3 Chrysene Results vs Bottle Number

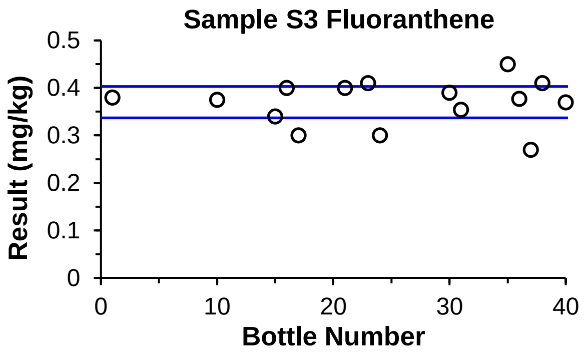


Figure 48 S3 Fluoranthene Results vs Bottle Number

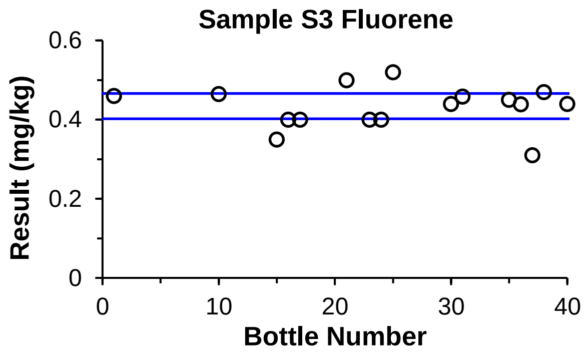


Figure 49 S3 Fluorene Results vs Bottle Number

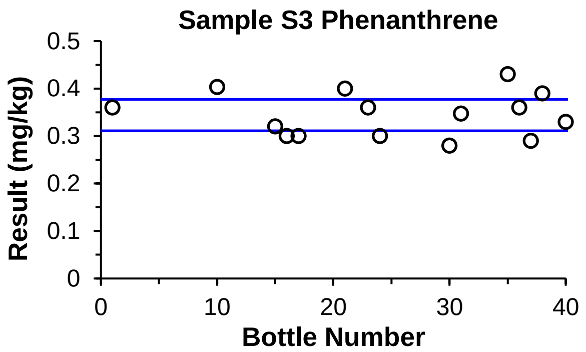


Figure 50 S3 Phenanthrene Results vs Bottle Number

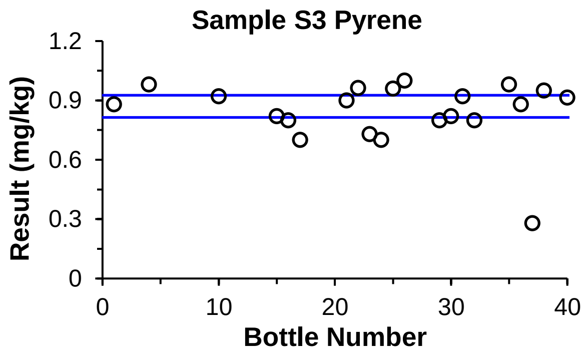


Figure 51 S3 Pyrene Results vs Bottle Number

A2.2 Stability

No stability testing was completed for this study as the samples were prepared, stored and dispatched using a process previously demonstrated to produce sufficiently stable samples for similar analytes and matrices over a similar time frame. After preparation and before dispatch, Samples S1 and S3 were stored in a refrigerator at approximately 4 °C, and Sample S2 was stored in a freezer at approximately -20 °C. For dispatch, samples were packaged into insulated polystyrene foam boxes with cooler bricks.

The results of this study also gave no reason to question the samples' transportation stability. Comparisons of results to days spent in transit for scored analytes are presented in Figures 52 to 63 (solid blue lines correspond to the assigned value ± U for each analyte; results have not been included here if they were excluded from all statistical calculations in Section 5). No significant trend was observed.

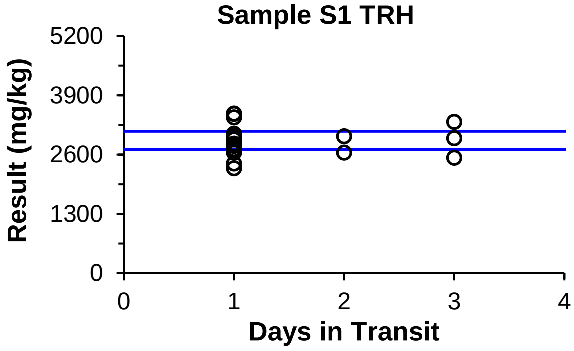


Figure 52 S1 TRH Results vs Transit Days

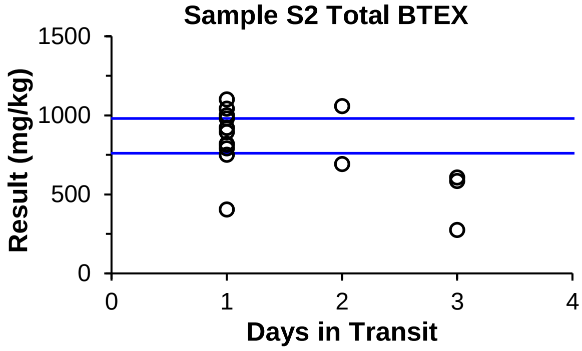


Figure 53 S2 Total BTEX Results vs Transit Days

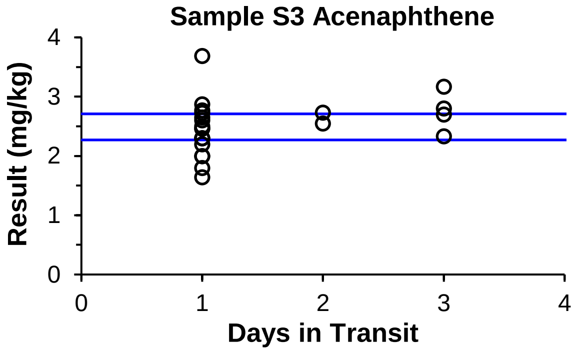


Figure 54 S3 Acenaphthene Results vs Transit Days

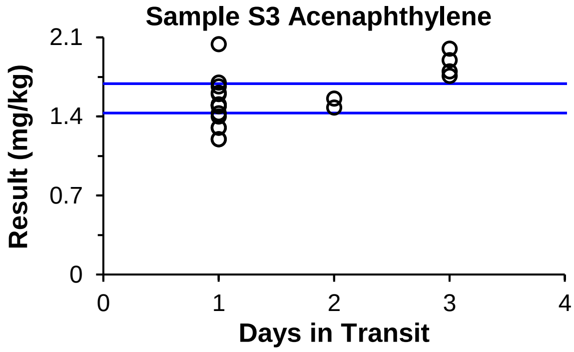


Figure 55 S3 Acenaphthylene Results vs Transit Days

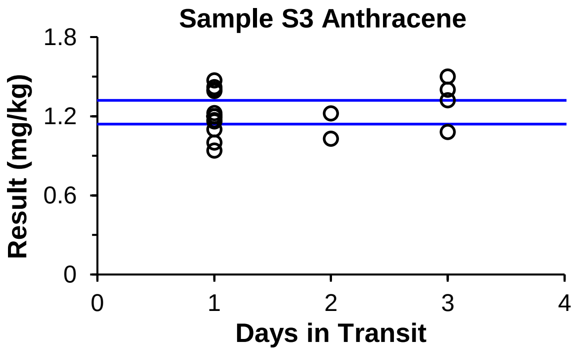


Figure 56 S3 Anthracene Results vs Transit Days

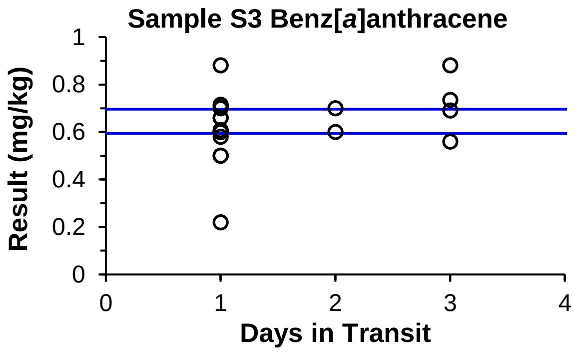


Figure 57 S3 Benz[*a*]anthracene Results vs Transit Days

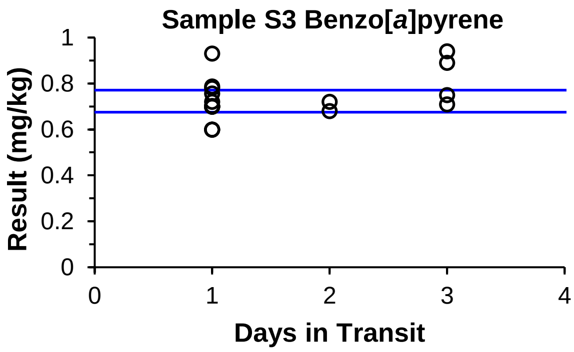


Figure 58 S3 Benzo[*a*]pyrene Results vs Transit Days

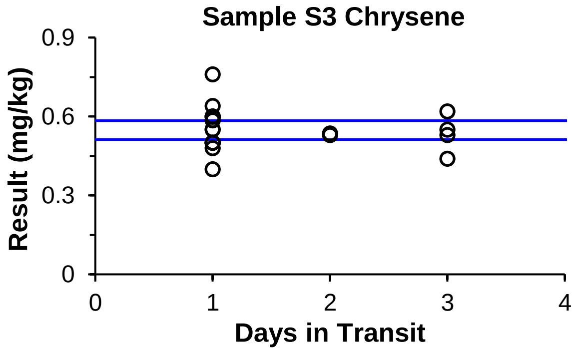


Figure 59 S3 Chrysene Results vs Transit Days

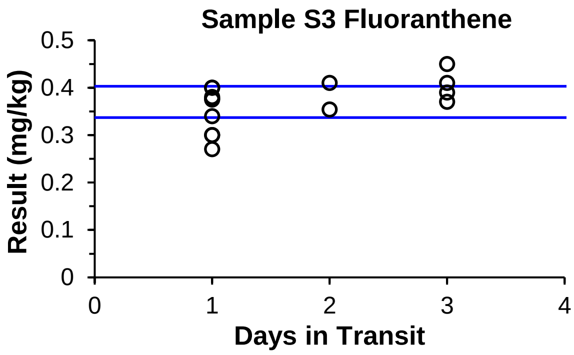


Figure 60 S3 Fluoranthene Results vs Transit Days

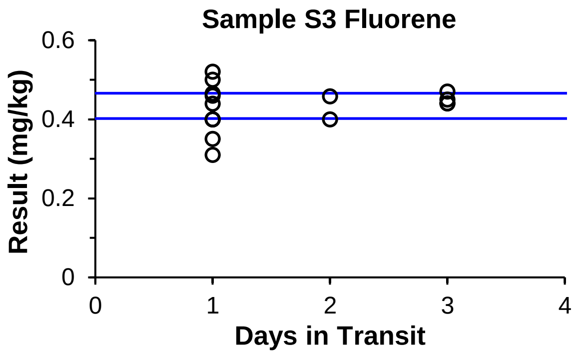


Figure 61 S3 Fluorene Results vs Transit Days

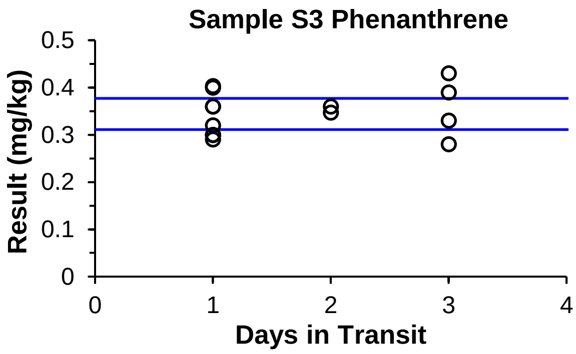


Figure 62 S3 Phenanthrene Results vs Transit Days

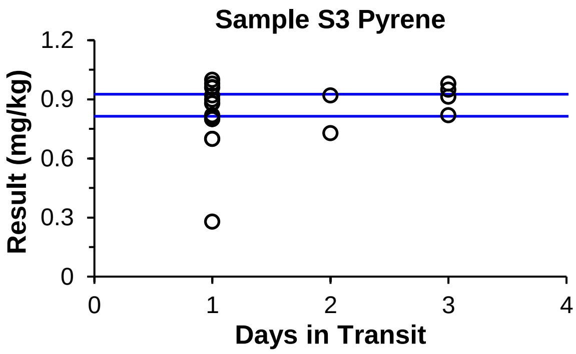


Figure 63 S3 Pyrene Results vs Transit Days

APPENDIX 3 ROBUST AVERAGE AND ASSOCIATED UNCERTAINTY, z-SCORE AND E_n-SCORE CALCULATIONS

A3.1 Robust Average and Associated Uncertainty

The robust average was calculated using the procedure described in ISO 13528.[7] The associated uncertainty was estimated as according to Equation 4.

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

where:

$u_{rob\ av}$ is the standard uncertainty of the robust average

$S_{rob\ av}$ is the standard deviation of the robust average

p 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 31.

Table 31 Uncertainty of the Robust Average for Sample S3 Chrysene

No. Results (p)	21
Robust Average	0.548 mg/kg
$S_{rob\ av}$	0.067 mg/kg
$u_{rob\ av}$	0.018 mg/kg
k	2
$U_{rob\ av}$	0.036 mg/kg

Therefore, the robust average for Sample S3 chrysene is 0.548 ± 0.036 mg/kg.

A3.2 z-Score and E_n-Score Calculations

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 32.

Table 32 *z*-Score and E_n-Score Calculation for Sample S1 >C10-C16 Result Reported by Laboratory 1

Participant Result (mg/kg)	Assigned Value (mg/kg)	Target SD	*z*-Score	E_n-Score
840 ± 300	937 ± 82	15% as PCV, or: 0.15 × 937 = 140.55 mg/kg	$z = \frac{840 - 937}{140.55} = -0.69$	$E_n = \frac{840 - 937}{\sqrt{300^2 + 82^2}} = -0.31$

APPENDIX 4 TEST METHODS REPORTED BY PARTICIPANTS

Participants were requested to provide information about their test methods. Responses are presented in Tables 33 to 35. Some responses may be modified so that the participant cannot be identified.

Table 33 Test Methods Sample S1 TRH

Lab. Code	Sample Mass (g)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
1	10	Solid-Liquid	DCM:Acetone	None	GC-FID	USEPA 3510
2	10	Solid-Liquid	Hexane		GC-FID	USEPA 8260
3	5	Solid-Liquid	50:50 DCM/Acetone	N/A	GC-FID	USEPA 8260
4	10	Solid-Liquid	DCM/Acetone		GC-FID	In-house
5	10	Solid-Liquid	DCM:Acetone	None	GC-FID	USEPA 3510
6	10	Solid-Liquid	DCM/Acetone (50:50)	None	GC-FID	USEPA 8260
7	4	Solid-Liquid	DCM/Acetone	NIL	GC-FID	In house
8	8	Solid-Liquid	Hexane:Acetone	Silica	GC-FID	
9	10	Solid-Liquid	1:1 DCM: ACETONE		GC-FID	USEPA SW 846 - 8015
10	10	Solid-Liquid	DCM:Acetone	Silica	GC-FID	NEPM Schedule B3
11	10	Solid-Liquid	DCM/Acetone		GC-FID	
12	10	Sonication	DCM:Acetone 1:1	Silica	GC-FID	USEPA 8015
13	10	Solid-Liquid	Acetone:DCM		GC-FID	2010 NEPM Schedule B3
14	10	Solid-Liquid	Acetone:DCM		GC-FID	USEPA SW846 8015
15	10	Solid-Liquid	DCM:Acetone	None	GC-FID	USEPA 3510
16	10	Solid-Liquid	1:1 DCM/ACETONE	N/A	GC-FID	USEPA 8015
17	10	Solid-Liquid	DCM:ACE	N/A	GC-FID	In house
18	10	Solid-Liquid	DCM:Acetone	NA	GC-FID	USEPA 8260
19	10	Solid-Liquid	DCM/Acetone	None	GC-FID	USEPA 8270C
20	10	30	DCM/Acetone	n/a	GC-FID	USEPA 8260
21	5.5	Solid-Liquid	DCM	Silica	GC-FID	In house
22	10	Solid-Liquid	DCM/Acetone	None	GC-FID	USEPA 8270

Table 34 Test Methods Sample S2 BTEX

Lab. Code	Sample Mass (g)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
1	5	Solid-Liquid	Methanol	None	P&T GC-MS	USEPA 8260
2	10	Solid-Liquid	Methanol		P&T GC-MS	USEPA 8260
3	5	Solid-Liquid	Methanol	N/A	P&T GC-MS	USEPA 8260
4						

Lab. Code	Sample Mass (g)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
5	5	Solid-Liquid	Methanol	None	P&T GC-MS	USEPA 8260
6	10	Solid-Liquid	MeOH	None	P&T GC-MS	USEPA 8260
7	5	Solid-Liquid	METHANOL	NIL	P&T GC-MS/MS	In house
8	10		Methanol	N/A	Headspace GC-MS/MS	
9	5	Solid-Liquid	METHANOL		P&T GC-MS	USEPA 8260
10	2	Solid-Liquid	matrix modifier solution	n.a.	Headspace GC-MS	USEPA 5021A
11	10	Solid-Liquid	Methanol		P&T GC-MS/MS	
12	14	Sonication	Methanol	Nil	Headspace GC-MS	USEPA 8260 & 5021
13	4	Solid-Liquid	Methanol		P&T GC-MS	USEPA 8260
14	5	Solid-Liquid	Acetone:DCM		P&T GC-MS	USEPA SW846 8260
15	5	Solid-Liquid	Methanol	None	P&T GC-MS	USEPA 8260
16	10	Solid-Liquid	METHANOL	N/A	P&T GC-MS	USEPA 8260
17	10	Solid-Liquid	MeOH	N/A	P&T GC-MS	In house
18	5	Solid-Liquid	Methanol	NA	P&T GC-MS	USEPA 8260
19	10	Solid-Liquid	Methanol	None	P&T GC-MS/MS	USEPA 8260B
20	5	10	Methanol	n/a	P&T GC-MS	USEPA 8260
21	1.1	Solid-Liquid	MeOH	None	P&T GC-MS	USEPA 8260
22	10	Solid-Liquid	MeOH	None	P&T GC-MS	USEPA 8260

Table 35 Test Methods Sample S3 PAHs

Lab. Code	Sample Mass (g)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
1	10	Solid-Liquid	DCM:Acetone	None	GC-MS	USEPA 8270
2	10	Solid-Liquid	DCM:Acetone		GC-MS/MS	USEPA 8260
3	5	Solid-Liquid	50:50 DCM/Acetone	N/A	GC-MS/MS	USEPA 8270
4	2	Solid-Liquid	Ethyl acetate		GC-MS/MS	In-house
5	10	Solid-Liquid	DCM:Acetone	None	GC-MS	USEPA 8270
6	10	Solid-Liquid	DCM /Acetone (50:50)	None	GC-MS/MS	USEPA 8260
7	4	Solid-Liquid	DCM/Acetone	NIL	GC-MS	In house
8	8	Solid-Liquid	Hexane:Acetone	N/A	GC-MS	In-house
9	10	Solid-Liquid	1:1 DCM: ACETONE		GC-MS	USEPA 8270
10	10	Solid-Liquid	DCM:Acetone	n.a.	GC-MS/MS	USEPA610
11	10	Solid-Liquid	DCM/Acetone		GC-MS/MS	

Lab. Code	Sample Mass (g)	Extraction Details	Extraction Solvent	Clean-Up	Measurement Instrument	Method Reference
12	10	Sonication	DCM:Acetone 1:1	Silica	GC-MS/MS	USEPA 8270
13	10	Solid-Liquid	Acetone:DCM		GC-MS	USEPA 8270
14	10	Solid-Liquid	Acetone:DCM		GC-MS	USEPA SW846 8270
15	10	Solid-Liquid	DCM:Acetone	None	GC-MS	USEPA 8270
16	10	Solid-Liquid	1:1 DCM/ACETONE	N/A	GC-MS/MS	USEPA 8270
17	10	Solid-Liquid	DCM:ACE	N/A	GC-MS	In house
18	10	Solid-Liquid	DCM:Acetone	NA	GC-MS	USEPA 8260
19	10	Solid-Liquid	DCM/Acetone	None	GC-MS	USEPA8270C
20	10	30	DCM/Acetone	n/a	GC-MS	USEPA 8270
21	9.5	Solid-Liquid	DCM:acetone	None	GC-MS	USEPA 8270
22	10	Solid-Liquid	DCM/Acetone	None	GC-MS	USEPA 8270

APPENDIX 5 ACRONYMS AND ABBREVIATIONS

ACE	Acetone
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
GC	Gas Chromatography
GUM	Guide to the expression of Uncertainty in Measurement
HEX	Hexane
HS	Headspace (GC)
IEC	International Electrotechnical Commission
ISO	International Standards Organization
k	Coverage Factor
LOR	Limit Of Reporting
Max	Maximum
Md	Median
MeOH	Methanol
Min	Minimum
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MU	Measurement Uncertainty
N	Number of numeric results
NATA	National Association of Testing Authorities, Australia
NEPM	National Environmental Protection Measure
NMI	National Measurement Institute, Australia
NR	Not Reported
NT	Not Tested
P&T	Purge and Trap (GC)
PAH	Polycyclic Aromatic Hydrocarbon
PCV	Performance Coefficient of Variation
PT	Proficiency Testing

RA	Robust Average
RM	Reference Material
SD	Standard Deviation
SI	International System of Units
SLE	Solid-Liquid Extraction
SS	Spiked Samples
SV	Spiked Value, or formulated concentration of a PT sample
TRH	Total Recoverable Hydrocarbons
U	Expanded Uncertainty
US EPA	United States Environmental Protection Agency

END OF REPORT