INSTITUTE FOR DEFENSE ANALYSES

# Statistical Methods in the Multi-Laboratory Validation of a PFAS Measurement Method

Allyson M. Buytendyk
Katherine I. Fisher
Tyler J. Pleasant
Sara Runkel
Jay S. Shah
John J. Silk

August 2023



IDA Product 3000051

INSTITUTE FOR DEFENSE ANALYSES
730 East Glebe Road
Alexandria, Virginia 22305-3086

The Institute for Defense Analyses is a nonprofit corporation that operates three Federally Funded Research and Development Centers. Its mission is to answer the most challenging U.S. security and science policy questions with objective analysis, leveraging extraordinary scientific, technical, and analytic expertise.


**About This Publication**

This work was conducted by the IDA Systems and Analyses Center under contract HQ0034-19-D-0001, project AM-2-5056, "Validation of PFAS Measurement Methods," for the Executive Director, Strategic Environmental Research and Development Program (SERDP) and Environmental Security Technology Certification Program (ESTCP), under the Deputy Assistant Secretary of Defense, Environment. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

**Acknowledgments**

The authors would like to thank Leon Hirsch, Noah Plymale, and Allison King of the IDA Systems and Analyses Center for their technical review of this report.


**For More Information**


Allyson M. Buytendyk, Project Leader
abuytend@ida.org, 703-845-6806

Leonard J. Buckley, Director, Science and Technology Division
lbuckley@ida.org, 703-578-2800


**Copyright Notice**

INSTITUTE FOR DEFENSE ANALYSES

IDA Product 3000051

# Statistical Methods in the Multi-Laboratory Validation of a PFAS Measurement Method

Allyson M. Buytendyk
Katherine I. Fisher
Tyler J. Pleasant
Sara Runkel
Jay S. Shah
John J. Silk

# Executive Summary

Per- and polyfluoroalkyl substances (PFAS) are persistent and mobile pollutants that have drawn the attention of the scientific community and regulatory agencies concerning the potential health impacts exposure to these man-made chemicals have on humans. The Strategic Environmental Research and Development Program (SERDP) and the Environmental Security and Technology Certification Program (ESTCP), with the Environmental Protection Agency (EPA) Office of Water, Engineering and Analysis Division, developed a new method—EPA Draft Method 1633—for measuring trace contamination of 40 different PFAS in eight diverse environmental matrices: groundwater (GW), surface water (SW), wastewater (WW), soils, sediment, landfill leachate, fish tissue, and biosolids (i.e., municipal wastewater treatment plant residuals). This method uses liquid chromatography tandem mass spectrometry to quantify PFAS analytes using isotopically labeled compounds. Using a validated laboratory procedure (i.e., analytical method) to quantify PFAS provides consistent and reliable measurements that offer confidence when comparing data across different samples of the same environmental matrix type.

SERDP/ESTCP sponsored the Institute for Defense Analyses (IDA) to conduct statistical analyses, in the joint Department of Defense (DoD) and EPA multi-laboratory validation (MLV) study of EPA Draft Method 1633, to ensure objective and unbiased results. SERDP/ESTCP's study plan for the PFAS MLV closely follows the process outlined in the EPA Alternate Test Procedure (ATP) guidance, EPA 821-B-18-001, which describes the tests and procedures for developing quality control (QC) acceptance criteria from the data generated in a study. The ATP specifies the statistical formulas based on the number of labs analyzing each sample. The PFAS MLV study includes 10 participating laboratories and 3 types of datasets: initial calibration (ICAL), initial demonstration of capability (IDC), and environmental matrix samples.

IDA's role in the PFAS MLV study is to calculate statistical values using the lab-generated data to summarize the overall performance of the method. The IDA-calculated values will inform the QC acceptance criteria that the EPA establishes for the method. IDA has analyzed the ICAL, aqueous IDC, and three types of environmental aqueous matrices (WW, SW, and GW) datasets provided by the sponsor. IDA used the statistical formulas outlined in the MLV/EPA's ATP for most analysis tests and identified alternative calculations in instances when a discrepancy between the PFAS MLV dataset and formulas occurred (see table below). As an additional measure, IDA was blind to the lab names and

locations of the environmental samples and not part of the validation and verification process of the datasets. The table below provides an overview of each analysis test in the MLV, the associated performance metric defined by the EPA for a test and the range of IDA calculated values summarizing the performance of the 40 PFAS "target" analytes, and the 24 isotopically labeled compounds called extracted internal standards.

This report documents the formulas IDA used in the statistical analyses and provides some high-level observations about the aqueous datasets. A digital appendix with the summary statistic data tables generated by IDA for the ICAL, aqueous IDC, and WW, SW, and GW matrices accompanies this document.

**Overview of the PFAS MLV Datasets and Analyses**

| PFAS MLV Dataset | Analysis Test in MLV | MLV Data Allowed Use of ATP Formula? | Performance Metric in EPA's ATP | Target Analyte Performance[1] | Extracted Internal Standard Performance[1] |
|---|---|---|---|---|---|
| **ICAL** | Calibration Linearity | No | pooled percent relative standard deviation (RSD) | 7.31% to 13.8% | 4.11% to 12.1% |
| **Aqueous IDC** | Method Detection Limit (MDL) | Yes | pooled MDL | 0.315 to 9.89 ng/L | N/A |
| | Limit of Quantitation Verification | N/A | N/A | N/A | N/A |
| | Initial Precision and Recovery | Yes | mean percent recovery | 95.0% to 109% | 69.1% to 98.1% |
| | | | percent RSD | 3.35% to 11.5% | 5.36% to 17.2% |
| **Matrix Samples** | Ongoing Precision and Recovery | Yes | mean percent recovery | 89.0% to 109% | 53.2% to 101% |
| | | | percent RSD | 7.29% to 15.9% | 7.18% to 26.6% |
| | Low-Limit Ongoing Precision and Recovery | Yes | mean percent recovery | 88.3% to 113% | 50.8% to 108% |
| | | | percent RSD | 8.22% to 14.3% | 8.03% to 21.8% |
| | Matrix Spike WW[2] | No | percent RSD | 8.94% to 68.0% | N/A |
| | Matrix Spike SW | | | 6.50% to 104% | N/A |
| | Matrix Spike GW[2] | | | 3.71% to 54.4% | N/A |

[1] Nine labs reported values in most datasets.

[2] Only eight labs reported values.

# Contents


1. Background....1
2. Statistical Methods....5
   - A. Initial Calibration (ICAL) Dataset....5
     1. Calibration Linearity....5
   - B. Initial Demonstration of Capability (IDC) Dataset....6
     1. Method Detection Limit (MDL)....6
     2. Limit of Quantitation Verification (LOQVER)....8
     3. Initial Precision and Recovery (IPR)....9
   - C. Environmental Matrix Dataset....11
     1. Ongoing Precision and Recovery (OPR) and Low-Level Ongoing Precision and Recovery (LLOPR)....11
     2. Matrix Spike Recovery....12
3. Discussion....13
   - A. Calibration Linearity....13
   - B. Method Detection Limit (MDL)....17
   - C. Limit of Quantitation Verification (LOQVER)....19
   - D. Initial Precision and Recovery (IPR)....21
   - E. Matrix Spike Recovery....26
   - F. Ongoing Precision and Recovery (OPR)....27
   - G. Low-Level Ongoing Precision and Recovery (LLOPR)....30
4. Summary....33

Appendix A. PFAS MLV Analytes....A-1
Appendix B. Interpretation of Box and Whisker Plots....B-1
Appendix C. Data Overview Plots....C-1
Appendix D. List of Tables in the Digital Appendix....D-1
Illustrations....E-1
References....F-1
Abbreviations....G-1

# 1. Background

Per- and polyfluoroalkyl substances (PFAS)[1] are a persistent and mobile pollutant that have drawn the attention of the scientific community and regulatory agencies due to concerns about the potential health impact exposure to these man-made chemicals have on humans. Once championed for their heat-, oil- and water-resistant properties, attributed to a molecular structure with a short, strong bond between carbon and fluorine atoms, these substances now come as a detriment to the environment. PFAS do not easily break down and can migrate into soil, water, and air. Because of the widespread use of PFAS across the United States, including at many military installations, these chemicals are present in various regulatory environmental matrices including water, sediments, soils, and fish tissue.[2] Analysis of environmental samples help elucidate which PFAS are present and at what quantities to understand the extent of the contamination and inform decisions about cleanup for an area.

Using a validated laboratory procedure (i.e., analytical method) to quantify PFAS provides consistent and reliable measurements that offer confidence when comparing data across different samples of the same environmental matrix type. Validation of an analytical method is a process that demonstrates the method is appropriate for its intended purpose. Analytical methods can also establish performance metrics for regulatory compliance. These metrics may include accuracy, precision, specificity, linearity, range, limit of detection (LOD), limit of quantitation (LOQ), and robustness.[3]

Since PFAS are a joint concern of the Department of Defense (DoD) and the Environmental Protection Agency (EPA), together they developed an analytical measurement method using liquid chromatography-tandem mass spectrometry (LC-MS/MS). This method offers broad applications to fulfill regulatory compliances under the

---

1 PFAS are a large group of synthetic chemicals used across the globe in consumer goods (e.g., cookware, clothing, cosmetics) and industrial applications specifically, aqueous film-forming foam used by the DoD to extinguish hazardous fires. "What are PFAS," ATSDR (Agency for Toxic Substances and Disease Registry), November 1, 2022, https://www.atsdr.cdc.gov/pfas/health-effects/overview.html; "Per- and Polyfluoroalkyl Substances (PFAS): A Nation Issue that Requires National Solutions," Department of Defense, Environmental Cleanup and Compliance, https://www.acq.osd.mil/eie/eer/ecc/pfas/.

2 "PFAS Explained," U.S. Environmental Protection Agency, April 10, 2023, https://www.epa.gov/pfas/pfas-explained.

3 M. Thompson, S. Ellison, and R. Wood, "Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC Technical Report)," *Pure and Applied Chemistry* 74 (5) (2002): 835–855, https://doi.org/10.1351/pac200274050835.

Clean Water Act[4] as it includes 40 PFAS in 8 diverse environmental matrices: groundwater (GW), surface water (SW), wastewater (WW), soils, sediment, landfill leachate, fish tissue, and biosolids (i.e., municipal wastewater treatment plant residuals). The Strategic Environmental Research and Development Program (SERDP) and the Environmental Security and Technology Certification Program (ESTCP) are currently sponsoring a validation study for this PFAS method.[5]

The study plan for the multi-lab validation (MLV) closely follows the process outlined in an EPA document for new methods for organic and inorganic analytes used in Clean Water Act programs.[6] The EPA's Alternate Test Procedure (ATP) provides guidance for developing performance-based quality control (QC) criteria using statistical results from the data collected in the study. The EPA's ATP also includes the overall procedures for the statistical analyses and the formulas for computing the acceptance criteria as part of the evaluation of new analytical methods for approval and inclusion in the Code of Federal Regulation (40 CFR Part 136).[7] Methods that complete the laboratory validation process following specific guidance and approved by the EPA are made available to support regulatory or guidance activities.

In 2022, SERDP/ESTCP sponsored the Institute for Defense Analyses (IDA) as the independent organization to conduct the statistical analyses in the MLV of the PFAS measurement method to ensure the results were objective and unbiased.[8] The MLV study design comprises 10 laboratories that generate 3 types of datasets: initial calibration (ICAL), initial demonstration of capability (IDC), and environmental matrix samples. Prior to delivery to IDA, the datasets undergo several reviews by the sponsor, the data manager, and data validators. Additionally, the sponsor anonymized all lab names and environmental locations in the dataset as an additional measure so IDA was blind to those identities. IDA has analyzed the ICAL, aqueous IDC, and samples of three environmental matrices (WW,

---

4 "Summary of the Clean Water Act," U.S. Environmental Protection Agency, June 22, 2023, https://www.epa.gov/laws-regulations/summary-clean-water-act.

5 EPA Draft Method 1633. Historically, EPA published draft methods on the Clean Water Act Methods website after completing the single-laboratory validation report. "CWA Analytical Methods for Per- and Polyfluorinated Alkyl Substances (PFAS)," U.S. Environmental Protection Agency, July 28, 2023, https://www.epa.gov/cwa-methods/cwa-analytical-methods-and-polyfluorinated-alkyl-substances-pfas.

6 SERDP/ESTCP, *Study Plan for Multi-Laboratory Validation of Draft EPA Method 1633 – PFAS in Aqueous, Solid, Biosolids, and Tissue Samples by LC-MS/MS*, March 2022.

7 U.S. Environmental Protection Agency, *Protocol for Review and Validation of New Methods for Regulated Organic and Inorganic Analytes in Wastewater Under EPA's Alternative Test Procedure Program,* EPA 821-B-18-001 (Washington, DC: Environmental Protection Agency, February 2018), https://www.epa.gov/sites/default/files/2018-03/documents/chemical-atp-protocol_feb-2018.pdf.

8 IDA also supported SERDP/ESTCP in the single-laboratory validation. A. M. Buytendyk, S. C. Runkel, and S. M. Cazares, "Data Compilation in Support of Single Laboratory Validation of a Novel Per- and Polyfluoroalkyl Substances (PFAS) Detection Method for Environmental Matrices," IDA Document D-22794 (Alexandria, VA: Institute for Defense Analyses, 2021).

SW, and GW) datasets. For each dataset, IDA inspected and evaluated the analysis metrics in the MLV/EPA's ATP, and identified alternative calculations in instances with a discrepancy between the dataset and formulas. In this report, Chapter 2 documents the formulas IDA used in the statistical analyses and highlights instances where those differ from the EPA's ATP. Chapter 3 discusses high-level observations about the datasets and each statistical test. Chapter 4 provides a short summary about the datasets and overall analyses.

# 2. Statistical Methods

IDA's role in the PFAS MLV study is to calculate9 the statistical values for each dataset type that summarizes the overall performance of the method for each test. These calculated values inform the QC acceptance criteria that the EPA will establish for the method. The EPA's ATP specifies three tiers of statistical formulas based on the number of labs analyzing each sample where Tier 3 requires a minimum of nine labs.[10] This chapter summarizes the statistical methods including the formulas IDA used to analyze the datasets received by the sponsor in the PFAS MLV study.

## A. Initial Calibration (ICAL) Dataset

### 1. Calibration Linearity

Calibration establishes the relationship between the amount of an analyte (e.g., concentration) to an instrument response (e.g., signal area) by fitting a curve to data corresponding to the instrument measurements made at known analyte values. Calibration linearity refers to there being a linear relationship between the analyte concentration and the value predicted by an instrument using the calibration curve. A linear calibration curve is not required for the relationship between the actual concentration and predicted concentration to be linear, only that the calibration curve is monotonic and accurately relates the concentration to the measured signal. Internal standards or a known quantity of other compounds are often added to the sample to compare the instrument response between the standard and the analyte to determine how much of the analyte is present. When a calibration curve is proportional, a response factor (RF) expresses the ratio of the signal area to the amount (e.g., mass) of analyte compared to the signal-to-mass ratio of the standard.[11]

The metric in the EPA ATP for determining the performance of a calibration curve based on a straight line through the origin is the percent relative standard deviation

---

9 IDA performs calculations on the dataset using coded scripts in Python version 3.7.8, rounds statistical values based on the number of significant figures reported in the dataset, and delivers the outputs as CSV files to the sponsor.

10 EPA, *Protocol for Review and Validation of New Methods,* G-22.

11 $Response\ Factor\ (RF) = \frac{Area_{analyte} Mass_{standard}}{Area_{standard} Mass_{analyte}}$

(RSD).[12] The percent RSD is the standard deviation divided by the mean of all the RFs multiplied by 100, for an analyte for each lab. The RSD limit is the QC acceptance criterion for the linearity test and is determined by combining or "pooling" the percent RSD from each individual lab.[13]

The PFAS MLV ICAL dataset for the linearity test includes three RSD values from each lab, for an analyte or internal standard. These three RSD values correspond to the three calibration tests performed by a lab. This dataset does not contain the necessary measured RF values to calculate an individual lab's overall percent RSD nor a pooled percent RSD for an analyte using the calculations as described in the EPA ATP. The International Union of Pure and Applied Chemistry (IUPAC) provides an alternative formula for combining the RSD of multiple series of measurements to calculate a pooled percent RSD for the PFAS MLV (Equation 1).

**Equation 1: Pooled Percent RSD[14]**

$$RSD_{pooled} = \sqrt{\frac{\sum(n_i-1)RSD_i^2}{\sum(n_i-1)}};$$

**where n = number of RF values, RSDi = relative standard deviation of ith RF values.**

## B. Initial Demonstration of Capability (IDC) Dataset

### 1. Method Detection Limit (MDL)

The MDL is the lowest analyte concentration that a method can detect reliably and provides an exact procedure to evaluate the limit of detection (LOD)[15] for an analytical method.[16] The Code of Federal Regulations (CFR) defines MDL as "the minimum measured concentration of a substance that can be reported with 99% confidence that the

---

[12] Relative standard deviation (RSD) is also known as coefficient of variance (CV).

[13] EPA, *Protocol for Review and Validation of New Methods,* G-23.

[14] International Union of Pure and Applied Chemistry (IUPAC), *Compendium of Chemical Terminology*, 2nd ed., compiled by A. D. McNaught and A. Wilkinson (Blackwell Scientific Publications, Oxford, 1997), https://doi.org/10.1351/goldbook; "Assignment and Presentation of Uncertainties of the Numerical Results of Thermodynamic Measurements," *Pure and Applied* Chemistry 53 (9) (1981): 1805–1826, http://dx.doi.org/10.1351/pac198153091805.

[15] The limit of detection is the lowest analyte concentration producing a response detectable above the noise level of the system, typically three times the noise level.

[16] L. H. Keith, W. Crummett, J. Deegan, R. A. Libby, J. K. Taylor, and G. Wentler, "Principles of environmental analysis," *Analytical Chemistry* 55 (14) (1983): 2210–2218, https://doi/10.1021/ac00264a003; J. A. Glaser, D. L. Forest, G. D. McKee, S. A. Quave, and W. L. Budde, "Trace analyses for wastewaters," *Environmental Science & Technology* 15 (12) (1981): 1426–1435, https://doi.org/10.1021/es00094a002.

measured concentration is distinguishable from method results."[17] The process for determining the MDL described in the EPA ATP involves analyzing seven samples of the matrix containing a known concentration or "spike" of the analyte and seven without the analyte or "blank" samples where the samples were taken through all steps of the method. The method limit (also known as the LOD) is the QC acceptance criterion and is found using a pooled MDL, from each of the individual lab's MDL.[18]

The PFAS MLV IDC dataset for the MDL test contains seven spiked sample concentration measurements and at least seven blank sample measurements, for most labs, for an analyte. The CFR and the EPA outlines the calculations for the individual lab's MDL (Equations 2-4),19 for an analyte, as follows:

1. Find the MDL for the spiked samples, using the reported concentration values, for an analyte:

**Equation 2: MDL Spiked Samples for Lab j ($MDL_{s,j}$)**

$$MDL_{s,j} = S_{s,j} \cdot t_{(n-1,1-\propto=0.99)};$$

**where $S_{s,j}$ = sample standard deviation, of spiked sample measured concentrations for lab j, $t_{(n-1,1-\propto=0.99)}$ = student's t-value for the one tailed test at the 99% confidence level with n-1 degrees of freedom.**

2. Find the MDL for the blank samples, using the reported results, for an analyte:
    a. If none of the blank samples give a numerical result, the MDL for the blank samples does not apply.
    b. If some (but not all) of the blank samples give a numerical result, the MDL for the blank samples is the maximum value.
    c. If all of the blank samples give a numerical result, the MDL for the blank samples is:

---

17 Code of Federal Regulations (CFR), Title 40, Part 136, Appendix B.

18 EPA, *Protocol for Review and Validation of New Methods,* G-23.

19 40 CFR Part 136, Appendix B; EPA, *Protocol for Review and Validation of New Methods,* G-9.

**Equation 3: MDL Blank Samples for Lab j ($MDL_{b,j}$)**

$$MDL_{b,j} = \bar{X}_j + S_{b,j} \cdot t_{(n-1,1-\propto=0.99)};$$

where $\bar{X}_j$ = mean measured concentration of the blank samples for lab j, $S_b$ = sample standard deviation, of the blank samples measured concentration for lab j, $t_{(n-1,1-\propto=0.99)}$ = student's t-value for the one tailed test at the 99% confidence level with n-1 degrees of freedom.

3. Determine the MDL by comparing the calculated $MDL_s$ and $MDL_b$ values:

**Equation 4: MDL**

$$MDL_j = \max\{MDL_{s,j}, MDL_{b,j}\};$$

where $MDL_{s,j}$ = the MDL for the spiked samples for lab j, $MDL_{b,j}$ = the MDL for the blank samples for lab j.

After finding each individual lab's MDL for an analyte, Equation 5 shows the calculation for a pooled MDL for the PFAS MLV.

**Equation 5: Pooled MDL[20]**

$$MDL_{pooled} = \sqrt{\sum_{j=1}^{m} \frac{n_j}{N} \left( \frac{MDL_j}{t_{(n_j,1-\alpha=0.99)}} \right)^2} \; t_{(N,1-\alpha=0.99)};$$

where m = number of labs, $MDL_j$ = method detection limit for the *jth* lab, $n_j$ = number of replicates for the *jth* lab, N = total number of replicates, $t_{(n,1-\alpha=.99)}$ = student's t-value for the one tailed test at the 99% confidence level with n degrees of freedom.

## 2. Limit of Quantitation Verification (LOQVER)

The limit of quantitation (LOQ) is the lowest concentration level of an analyte that produces a quantitative result with a specific degree of confidence.[21] The DoD Quality Systems Manual (QSM) for Environmental Laboratories, referenced by the MLV study

[20] EPA, *Protocol for Review and Validation of New Methods,* G-22.

[21] Keith, et al., "Principles of environmental analysis," 2210–2218; The LOQ is commonly defined as ten times the noise level.

plan, describes the LOQ verification (LOQVER) procedure for analyzing four to seven samples to establish precision and relative bias for each laboratory near the LOQ.[22]

The PFAS MLV IDC dataset for the LOQVER test includes a single spiked sample concentration measurement for an analyte or internal standard from most of the labs. This dataset does not contain the necessary measured concentration data to calculate an individual lab's precision because the standard deviation of single data point is undefined.[23] Equation 6 displays the bias calculation for each lab, using the data for analytes and internals standards, for the PFAS MLV.

**Equation 6: LOQ Percent Bias[24]**

$$LOQ_{bias,j} = \frac{\text{spike concentration} - \bar{X}_j}{\text{spike concentration}} \cdot 100;$$

where $\bar{X}_j$ = mean of the measured sample concentrations for lab j.

### 3. Initial Precision and Recovery (IPR)

The IPR test demonstrates whether a lab's capability to produce results are acceptable before the labs analyze the environmental samples.[25] Precision characterizes the variability that occurs in a series of experiments under similar conditions and therefore measures the reproducibility of a result.[26] Sources of random error contributing to the variability or scatter in the result include differences in the reagents and instruments used as well as different analysts conducting the experiment across labs in a study. The precision obtained for a single lab over a period of time expresses the within-lab reproducibility and the precision from results across different laboratories indicates the between-lab reproducibility. Recovery shows how the instrument response to an analyte in a sample compares to the response expected based on the calibration model.

The two metrics in the EPA ATP for determining the performance of the labs IPR are the mean percent recovery of the spiked sample measurements and a combined standard

---

22 Department of Defense, Department of Energy (DoD, DOE), *DoD Quality Systems Manual Version 5.4*, Module 4, Section 1.5.2 (Washington, DC: DoD, DOE, 2021), 77–78, https://www.denix.osd.mil/edqw/denix-files/sites/43/2021/10/QSM-Version-5.4-FINAL.pdf.

23 IDA explored the possibility of including data from other tests (e.g., the spiked samples in the MDL test), however, the spike concentrations were less than the LOQ for some of the labs where the LOQVER test specified the spike concentrations at 1-2 times the LOQ.

24 DoD, DOE, *DoD QSM Version 5.4*, 77.

25 EPA, *Protocol for Review and Validation of New Methods,* G-6.

26 M. J. Green, "Peer Reviewed: A Practical Guide to Analytical Method Validation," *Analytical Chemistry* 68 (9) (1983): 305A–309A, https://doi.org/10.1021/ac961912f; "LC-MS Method Validation," University of Tartu, https://sisu.ut.ee/lcms_method_validation.

deviation that includes the within- and between-lab standard deviations.[27] The upper- and lower-percent recovery limits are the QC acceptance criteria for recovery, which are constructed using the overall mean and a combined standard deviation of the within- and between-lab standard deviations. The percent RSD of the percent recovery is the QC acceptance criterion for precision, where the within-lab standard deviation is divided by the overall percent recovery mean multiplied by 100.

The PFAS MLV IDC dataset for the IPR test contains four spiked sample concentration measurements and the corresponding percent recoveries for a given analyte or internal standard, and lab. The EPA ATP outlines the calculations for an analyte in Equations 7–10:[28]

**Equation 7: Between Lab Standard Deviation ($s_b$)**

$$s_b = \sqrt{\frac{\sum_{j=1}^{m}\left(\overline{X_j} - \overline{X}\right)^2}{\mathrm{m} - 1}}$$

where m = the number of labs, $\overline{X}$ = overall mean of the percent recovery from all labs, $\overline{X_j}$ = the mean percent recovery for the jth lab.

**Equation 8: Within Lab Standard Deviation ($s_w$)**

$$s_w = \sqrt{\frac{\sum_{j=1}^{m}\left(s_j\right)^2}{\mathrm{m}}}$$

where m = the number of labs, sj = the variance of the percent recovery values for the jth lab.

**Equation 9: IPR Combined Standard Deviation ($s_{IPR}$)**

$$s_{IPR} = \sqrt{\left(1 + \frac{1}{m}\right)s_b^2 + \left(\frac{1}{4} - \frac{1}{n}\right)s_w^2};$$

where m = the number of labs, n = the number of data points per lab, $s_b$ = the between lab standard deviation, $s_w$ = the within lab standard deviation.

[27] EPA, *Protocol for Review and Validation of New Methods,* G-25–26.

[28] EPA, *Protocol for Review and Validation of New Methods,* G-25–26.

**Equation 10: RSD**

$$RSD = \frac{s_w}{\bar{X}} \cdot 100;$$

where $s_w$ = the within lab standard deviation, $\bar{X}$ = mean percent recovery across all labs.

## C. Environmental Matrix Dataset

### 1. Ongoing Precision and Recovery (OPR) and Low-Level Ongoing Precision and Recovery (LLOPR)

Both the OPR and LLOPR tests are done throughout the environmental matrix analyses and provide assurance the results produced by the labs are consistent and reproducible throughout the study. The OPR test, sometimes referred to as a QC check, demonstrates the labs' routine performance with known amounts of analytes (similar or identical to the IPR samples). The LLOPR test verifies the LOQ with samples spiked at low concentrations.

The metric in the EPA ATP for determining the performance of the labs' OPR is the mean percent recovery of the spiked sample measurements and a combined standard deviation that includes the within- and between-lab standard deviations.[29] The upper- and lower-percent recovery limits are the QC acceptance criteria for recovery, which are constructed using the overall mean and combined standard deviation.

The PFAS MLV environmental matrix datasets for the OPR and LLOPR tests contain several spiked sample concentration measurements and the corresponding percent recoveries for analytes and internal standards, for most labs. The EPA ATP outlines the same calculations for finding the between- and within-lab standard deviations in Equations 7 and 8[30] to compute the combined standard deviation for the OPR (Equation 11), for analytes in an environmental matrix dataset.

---

[29] EPA, *Protocol for Review and Validation of New Methods,* G-26.

[30] The calculation for the within-lab standard deviations (Equation 8) excludes instances where a lab reports a single spiked sample concentration measurement as the standard deviation of a single value is undefined.

**Equation 11: OPR Combined Standard Deviation ($s_{OPR}$)[31]**

$$s_{OPR} = \sqrt{\left(1 + \frac{1}{m}\right) s_b^2 + \left(1 - \frac{1}{n}\right) s_w^2}\ ;$$

**where m = the number of labs, n = the number of data points per lab, $s_b$ = the between-lab standard deviation, $s_w$ = the within-lab standard deviation.**

Equation 10 is also the formula to calculate the RSD in the OPR test. Similarly, the calculations for the LLOPR test follow those for the OPR using Equations 7, 8, 10, and 11.

## 2. Matrix Spike Recovery

The matrix spike recovery tests whether the environmental matrix (e.g., WW, SW, GW) surrounding the analyte interferes in the sample preparation or instrument response affecting the ability to accurately quantify the analyte in a field sample. Structural analogs and stable isotopically labeled compounds[32] both have similar properties to the analyte and provide one technique to determine possible matrix effects. The EPA ATP describes another procedure for determining the method performance of a matrix in instances where an isotopic analog of an analyte is not available to use as an internal standard.[33] The metric defined is the relative percent difference between matrix spike and matrix spike duplicates.[34]

The PFAS MLV method is an isotopic dilution method where isotopically labeled compounds are spiked into the field samples, although, not all analytes in the study have an isotopic analog. The environmental matrix datasets for the matrix spike test contain concentration measurements from spiked field samples and the corresponding percent recoveries, for analytes and internal standards, for most labs. Although most labs made triplicate measurements of the analytes for each matrix sample, the dataset did include information to associate the matrix spike measurement with the corresponding isotopic standard measurement to calculate the relative percent difference. The calculations for the matrix spike test instead include those in Equations 7 and 8 to determine $s_b$ and $s_w$ as well as Equation 10 to find the RSD for the matrix test.[35]

---

[31] EPA, *Protocol for Review and Validation of New Methods,* G-26.

[32] Compounds where an atom in the molecule is replaced by a different stable (non-radioactive) isotope of that atom (e.g., deuterium is an isotope of hydrogen).

[33] EPA, *Protocol for Review and Validation of New Methods,* G-27.

[34] EPA, *Protocol for Review and Validation of New Methods,* G-27.

[35] The WW and GW datasets ended up with only results from eight laboratories; however, IDA still analyzed these datasets as outlined at Tier 3.

# 3. Discussion

The PFAS method includes 40 "target" analytes prevalent in environmental matrices that are quantified with standard or isotopically labeled compounds added to the samples. There are also 24 extracted internal standard (EIS) compounds and 7 non-extracted internal standard (NIS) compounds added to the samples.[36] EIS are isotopically labeled PFAS compounds that are added to samples prior to any preparation steps. NIS are isotopically labeled PFAS compounds added just before analyzing the prepared samples in the LC-MS/MS instrument.

In all of the datasets IDA received, the sponsor excluded one of the labs (Lab 8) for not performing the method correctly, which left nine labs for most of the datasets. One lab was missing from the WW (Lab 10) and GW (Lab 9) datasets leaving only eight labs. IDA still followed the EPA's Tier 3 formulas for the matrix spike samples as IDA's code was developed prior to receiving the matrix datasets and it allowed for comparison across datasets by using the same formula. Additionally, IDA did not include any data qualified or flagged with the letter "U," meaning the analyte was not detected or detected at a concentration less than the MDL. Appendix C provides summary figures of the lab results for each of the 40 PFAS across the ICAL, aqueous IDC, and WW, SW, GW environmental matrices datasets.

## A. Calibration Linearity

A linear calibration curve is where the instrument response is linearly proportional to the amount of analyte in the sample meaning the measured instrument signal at known amounts (or concentrations) of the analyte fits the equation of line (i.e., $y=mx + b$) for a range of concentrations. To assess the calibration linearity for the MLV dataset, the calibration curve is a straight line through the origin (zero response at zero concentration where $b=0$) and proportional with the response factor/ratio. The percent RSD of the average ratio of the instrument response to the analyte amount or RF for an analyte compared to a standard expresses the overall amount of deviation from a straight line where each point in the calibration has equal weight (i.e., measurements at low concentration have the same impact as high concentrations). The typical acceptance criterion for a linear calibration in

[36] See Appendix A for the list of target PFAS analytes, EIS PFAS compounds, and NIS PFAS compounds. For all tests, IDA calculated values for the target analytes, and for most tests the EIS compounds. IDA only calculated values for the percent recoveries of the NIS compounds for all the aqueous matrix samples as the values were not populated for all the labs for some of the datasets.

analytical chemistry is a percent RSD of less than 15% or 20%, which is consistent with other research fields that use percent RSD as a metric of performance.[37] The MLV study plan also allowed reporting the percent RSD or the relative standard error (RSE) for calibration linearity.[38]

The PFAS MLV ICAL dataset includes, for each target analyte or EIS compound, three sets of calibration values, including number of calibration points, average RF, standard deviation, and RSD values from nine labs. Lab 3 also reported RSE values for nine of their calibrations, most of which had a percent RSD of 19.5% or above. Missing from the ICAL dataset were the individual RF values and the corresponding concentration values so IDA could not independently verify the calibration models nor the RSD/RSE values. IDA also observed several inconsistencies between the reported standard deviation values calculated by a third party compared to the lab reported RSDs and mean RFs for each calibration.

The sponsor provided another dataset with the concentrations for each calibration sample (CS) used by each lab. The number of CS reported was inconsistent with the number of calibration points reported, with the ICAL average RF and RSD values for some labs adding more ambiguity to the ICAL dataset. The heatmap in Figure 1 is a visualization of all the percent RSDs of the RFs reported for each analyte using a Z-score to better compare the individual laboratory measurements. A Z-score is a measure of how many standard deviations below or above a value is from the population mean. In this heatmap, blue shades indicate a lab's reported value is below the mean for an analyte, yellow shades depict a value is above the mean, and black or dark shades represent the value is close to the mean score (i.e., a Z-score of zero). Equation 12 shows how to compute a Z-score of a measurement value X using the mean and standard deviation for all the average RF values for a given analyte. Along the x-axis from left to right, the first 40 PFAS are the target analytes followed by the 24 PFAS EIS compounds. Most of the average RSDs reported are within 3 standard deviations from the mean for each analyte (i.e., down a column in Figure 1). The EIS compound $^{13}C_2$-PFTeDA had the highest reported RSD of 34% followed by the target analytes (7:3FTCA, NFDHA, PFMBA) and EIS compounds ($D_3$-NMeFOSAA and $^{13}C_2$-PFDoA). The median percent RSD varied from 5.35% to 11.8% for target analytes and 2.17% to 11.0% for EIS compounds.

---

[37] R. Burrows and J. Parr, "Evaluating the Goodness of Instrument Calibration for Chromatography Procedures," *LCGC Supplements* 38 (11) (2020): 35–38, https://www.chromatographyonline.com/view/evaluating-the-goodness-of-instrument-calibration-for-chromatography-procedures. Generally, data with a percent RSD greater than 30% indicates a larger spread in data and could be related to an issue with the performance of the method or instrument.

[38] RSE is the standard deviation of the mean divided by the square root of the sample size and multiplied by 100.

**Equation 12: Z-score**

$$Zscore = \frac{X - \bar{X}}{\sigma_X};$$

**where X = data point in the set, $\bar{X}$ = mean of all values in a set, $\sigma_X$ = sample standard deviation of the set.**

The pooled percent RSD is a way to compile percent RSDs in a series of measurements with different means and the statistical value is most meaningful when those measurements are performed under similar conditions, like a method validation, to estimate the overall precision. The computed pooled percent RSD values[39] (Equation 1) for the PFAS MLV ICAL dataset span between 7.31% to 13.8% for the target analytes and 4.11% to 12.1% for the EIS compounds. Likely due to the RSD values having outliers primarily on the high end, the pooled percent RSD for this dataset is almost always larger than the median percent RSD (by an average of 16.8%). One possible factor affecting the linearity of a calibration model in the method could be from LC-MS/MS instrument components, such as the ionization source or detector.[40]

[39] The reference for the pooled percent RSD value cited the Bartlett Test as an option to test whether a series of measurements had the same precision or standard deviation prior to calculating a pooled value. IDA did not pursue the Bartlett Test with the PFAS MLV ICAL dataset as several inconsistencies were observed between the reported standard deviation, mean RFs, and RSDs values.

[40] "LC-MS Method Validation: 3.1. Linearity," University of Tartu, https://sisu.ut.ee/lcms_method_validation/31-linearity.

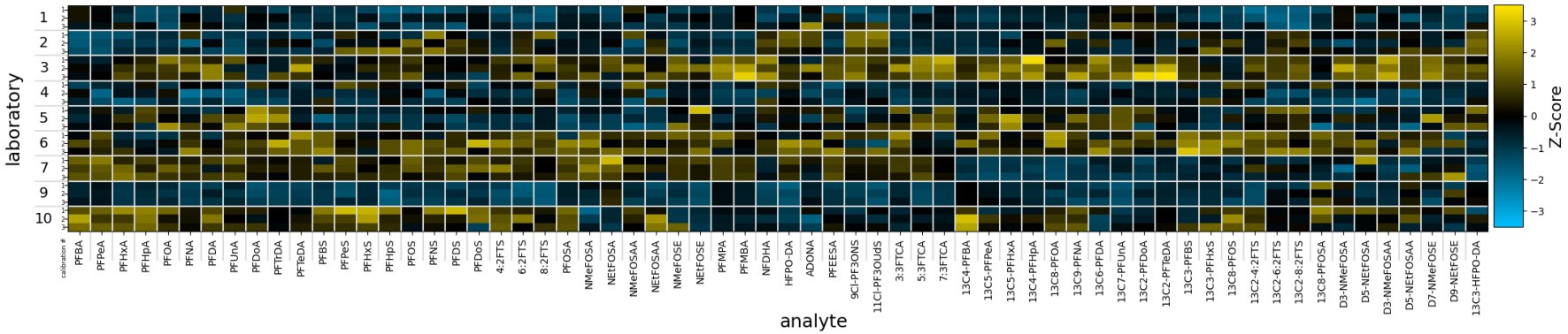


**Figure 1. Z-score of the RSDs of the RFs reported for each of the three calibration tests conducted by every lab, for 40 target PFAS analytes and 24 EIS PFAS compounds.**

## B. Method Detection Limit (MDL)

The detection limit for an analyte is an important value to establish for an analytical method. Detection limits can also be contentious especially in low-level analyses for substances that are toxic or pose harm to the environment as regulators use values to assess risk and compliance. There are several different "detection" definitions, which can cause confusion.[41] The MDL test quantifies the lowest reliable concentration of an analyte when processing a blank or sample through the complete analytical method.[42] The MDL is theoretically derived as an error distribution associated with the operational characteristics of the method. The pooled percent MDL is a statistical value from a series of measurements performed under similar conditions by multiple laboratories.

The PFAS MLV IDC dataset for the MDL test contained seven spiked sample concentration measurements and at least seven blank sample measurements, for nine labs, for the target analytes. The scatterplot in Figure 2 is a visualization of the individual lab MDL values (dash) calculated using Equation 2 through Equation 4 and the pooled MDL value (triangle) calculated using Equation 5, for each analyte across all nine labs. The computed pooled MDL values for the PFAS MLV Aqueous IDC dataset span between 0.315 to 9.89 ng/L for the target analytes.

---

41 Keith, et al., "Principles of environmental analysis."

42 *ACS Reagent Chemicals*, Part 1: Introduction and Definitions (Washington, DC: ACS Publications, 2017), https://pubs.acs.org/doi/book/10.1021/acsreagents; Glaser, et al., "Trace analyses for wastewaters," 1426–1435.

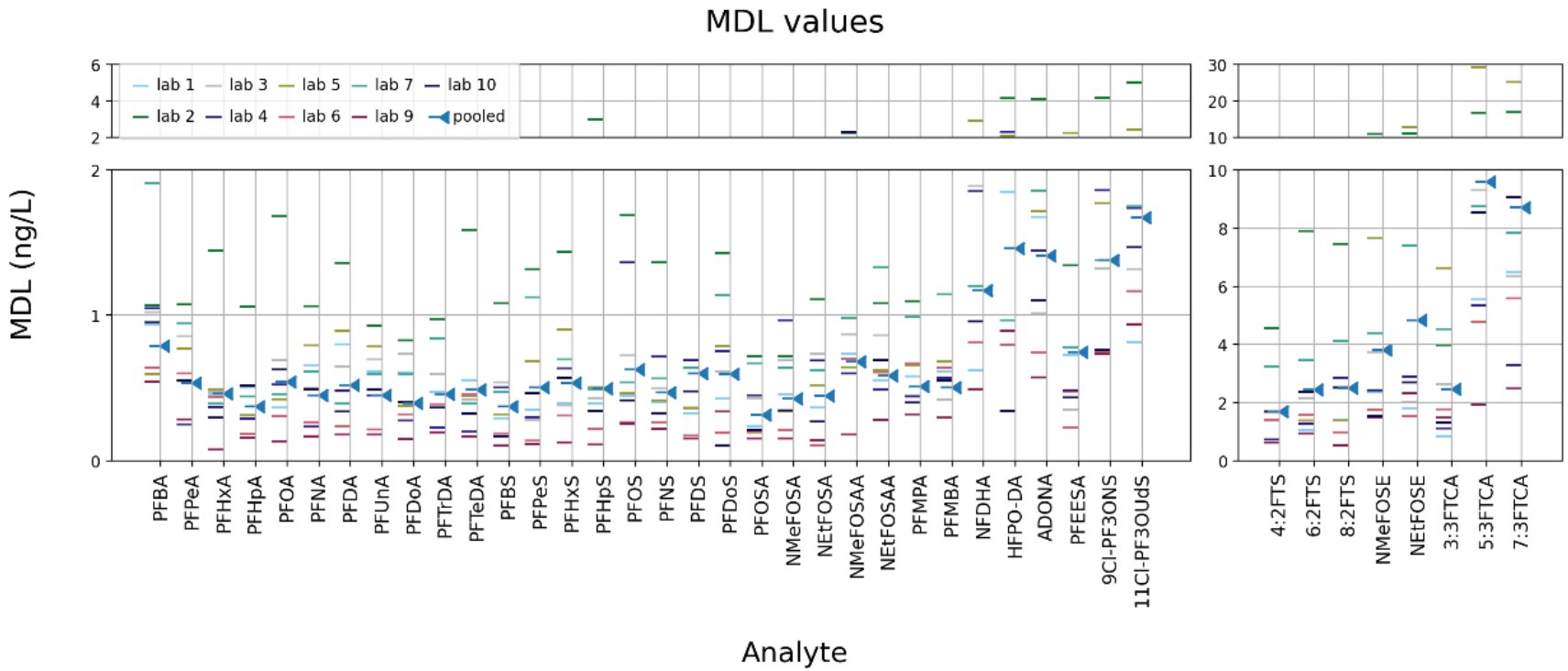


**Figure 2. MDL values calculated for nine labs that went into the computed pooled MDL for 40 target PFAS analytes. The fluorotelomer sulfonic acids (4:2FTS, 6:2FTS, and 8:2FTS), perfluorooctane sulfonamide ethanols (NMeFOSE and NEtFOSE), and fluorotelomer carboxylic acids (3:3FTCA, 5:3FTCA, and 7:3FTCA) are displayed on separate axes to avoid visual suppression of smaller MDL values. A break in the y-axes of both plots avoids visual suppression of reported MDL values within each group.**

## C. Limit of Quantitation Verification (LOQVER)

The LOQVER test quantifies the precision and relative bias for each lab when measuring analytes or internal standards at 1 to 2 times the LOQ. The PFAS MLV dataset for the LOQVER test is sparse with only 18 values for each analyte or internal standard with seven of the nine labs only reporting a single value and Labs 2 and 6 reporting multiple values. IDA explored whether the spiked sample MDL dataset was suitable to include in the LOQVER analysis to calculate the precision for each of the labs. The spike concentrations used by several labs for the spiked MDL measurements were less than the LOQ and did not meet the PFAS MLV study requirement so the spiked sample MDL dataset was not appropriate to use for all labs to calculate the labs' precision. IDA did calculate the percent relative bias for each lab shown in Figure 3. Relative bias is an estimate of systematic error; however, the calculations for most labs in the LOQVER test are based on a mean of one data point and not a well-represented estimate of the systematic error for those labs.

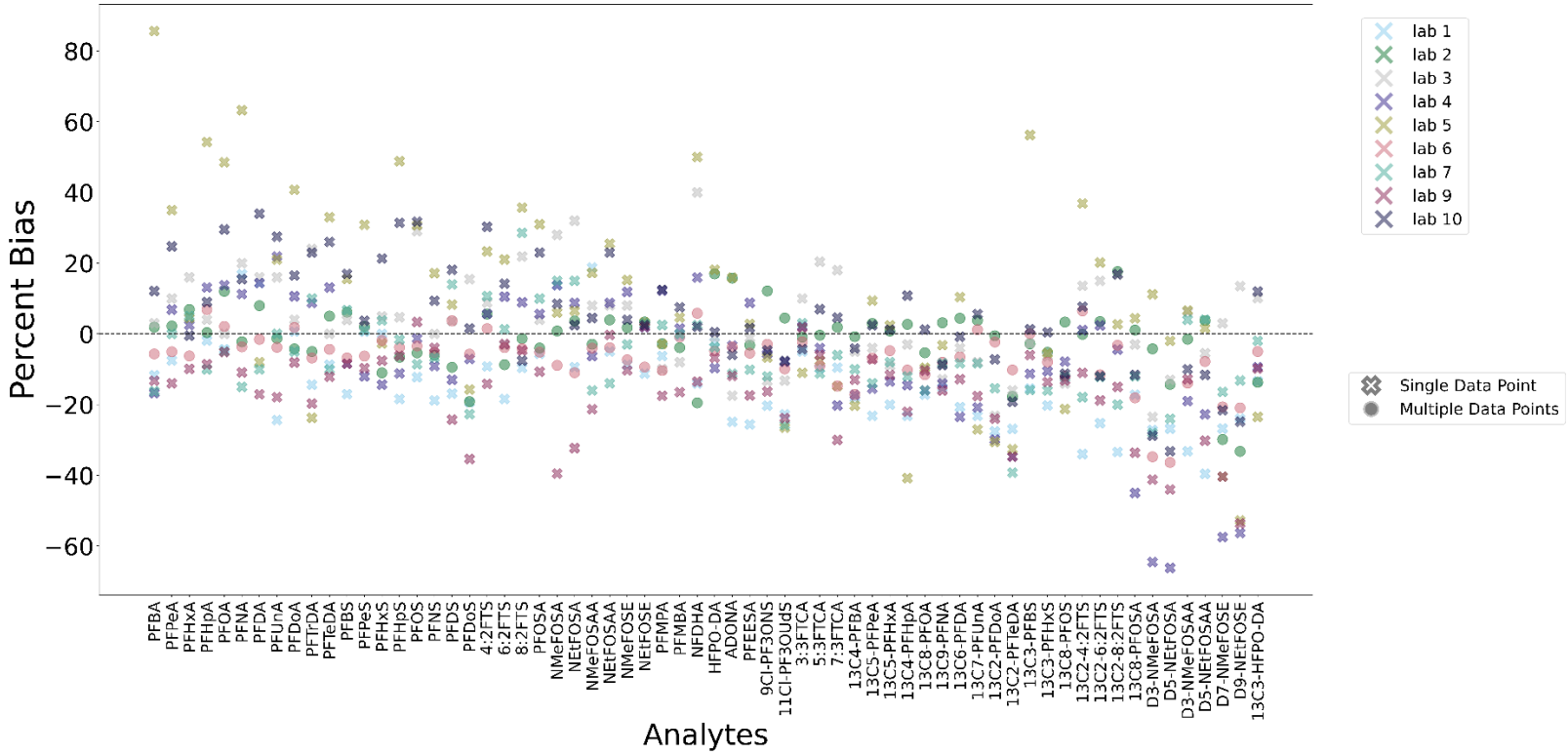


**Figure 3. Percent bias of the concentrations measured in the aqueous IDC dataset reported by nine labs, for 40 PFAS target analytes and 24 EIS compounds.**

## D. Initial Precision and Recovery (IPR)

The IPR test establishes the variability within each lab and the reproducibility of a result between labs prior to the labs using the method with the environmental samples. The PFAS MLV dataset for the IPR includes each lab reporting four results, corresponding to measurements of the four aliquots of reference matrix spiked with analytes and standards for every lab.[43] The boxplot in Figure 4 is a visualization of the percent recoveries of the four reported measurements made by all nine labs for each target and EIS compounds in the aqueous IPR dataset and shows the spread in values across labs. A more detailed explanation of the box and whisker plot is in Appendix B. The "X" markers indicate data points outside the range defined in the box and whiskers and the colors show the lab reporting that value. The overall mean percent recovery for the target analytes range from 95.0% to 109% and EIS compounds ranged from 69.1% to 98.1%. The perfluorooctane sulfonamides (FOSA) and the perfluorooctane sulfonamide ethanols (FOSE) EIS compounds have the lowest average and median recovery values, yet the corresponding target analytes recovery values center around 100%.

[43] SERDP/ESTCP, *Study Plan for Multi-Laboratory Validation*.

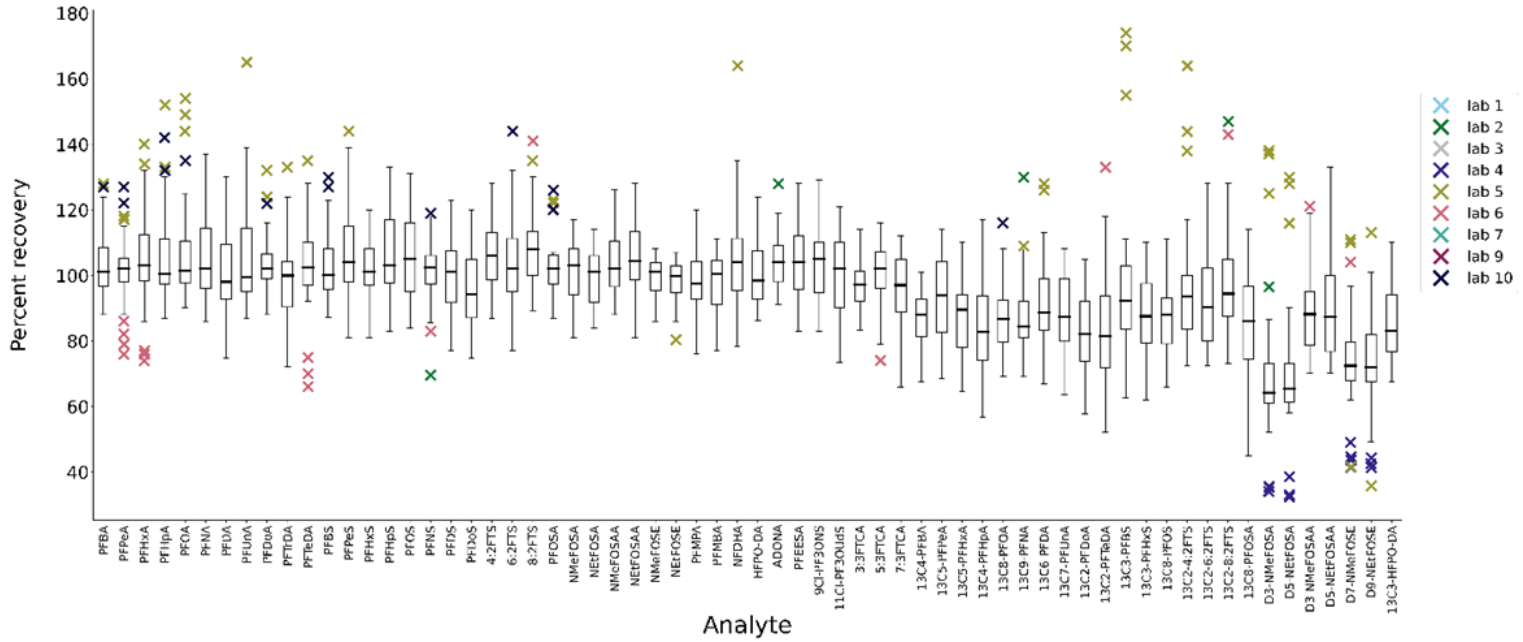


**Figure 4. Percent recovery of the four measurements in the aqueous IPR dataset reported by all nine labs, for 40 PFAS target analytes and 24 PFAS EIS compounds.**

The blue piecewise function (Equation 13) in Figure 5 is an empirical model known as the Horwitz curve that generalizes the relationship between the reproducibility between labs[44] for concentrations ranging from 10% to 10 parts per billion (ppb).[45] The curve was derived from studying thousands of results from interlaboratory analyses of analytes including food, pharmaceuticals, and pesticides.[46] Thompson later proposed a piecewise function in Equation 13 as the extremes of the original Horwitz curve tend to overestimate the variability.[47] In food analyses, the Horwitz curve has been used as a performance criterion and the x-axis is normally displayed with concentration units decreasing to the right.[48] An important disclaimer is the Horwitz curve provides a comparison to reproducibility results from other collaborative studies at similar concentration levels and is not an estimate of uncertainty or evaluation of performance.

**Equation 13: Modified Horwitz Curve**

$$\text{percent RSD}_{\text{H}} = \begin{cases} 22\%, \text{ if } c < 1.2 \times 10^{-7} \\ (2\ \%)c^{-0.1505}, \text{ if } 1.2 \times 10^{-7} \leq c \leq 0.138 \\ (1\ \%)c^{-0.5}, \text{ if } c > 0.138 \end{cases}$$

**where c = dimensionless fraction of concentrations (e.g., ppb).**

---

44 The RSD of reproducibility, or $RSD_R$, is defined as the between-lab precision, which is the sum of the within-laboratory precision, $s_r$, and the "pure" between laboratory precision, $s_L$, expressed as variances. W. Horwitz and R. Albert, "The Horwitz ratio (HorRat): A useful index of method performance with respect to precision," *Journal of Association of Official Analytical Chemists (AOAC) International* 89, (4) (2006): 1095–1109, https://doi.org/10.1093/jaoac/89.4.1095.

45 D. L. Massart, J. Smeyers-Verbeke, and Y. V. Heyden, "Benchmarking Analytical Methods Horwitz Curve," *LCGC Europe*, 18 (10) (2005): 528–531, https://www.chromatographyonline.com/view/benchmarking-analytical-methods-horwitz-curve.

46 W. Horwitz, L. R. Kamps, and K.W. Boyer, "Quality Assurance in the Analysis of Foods for Trace Constituents," *Journal of Association of Official Analytical Chemists (AOAC) International* 63, (6) (1980): 1344–1354, https://doi.org/10.1093/jaoac/63.6.1344; Royal Society of Chemistry, "The amazing Horwitz function," *AMC Technical Brief* no. 17, (2004), ed. M. Thompson, https://www.rsc.org/images/horwitz-function-technical-brief-17_tcm18-214859.pdf.

47 M. Thompson, "Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing," *Analyst* 125, (2000): 385–386, https://doi.org/10.1039/B000282H.

48 Horwitz and Albert, "The Horwitz ratio (HorRat),"1095–1108.

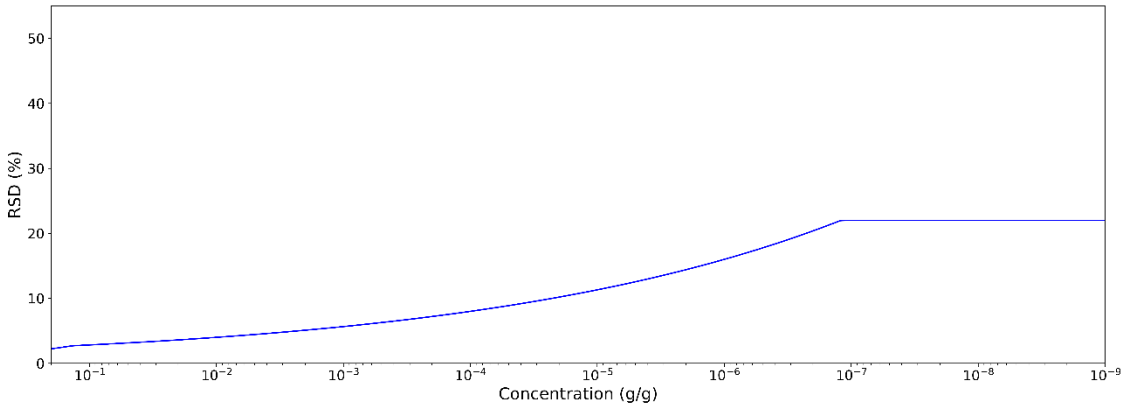


**Figure 5. Plot of the modified Horwitz curve depicting the relationship between concentration and percent RSD.**

The scatter plot in Figure 6 offers a visualization of the within-lab variabilities of the four measurements made by each lab (the standard deviation values for each lab used in Equation 8, for every analyte) plotted as the percent RSD verses concentrations. Most points are below 20% percent RSD with the exception of several analytes from Lab 5. The computed overall percent RSD values (Equation 10) ranged from 3.35% to 11.5% for the target analytes and 5.36% to 17.2% for the EIS compounds.

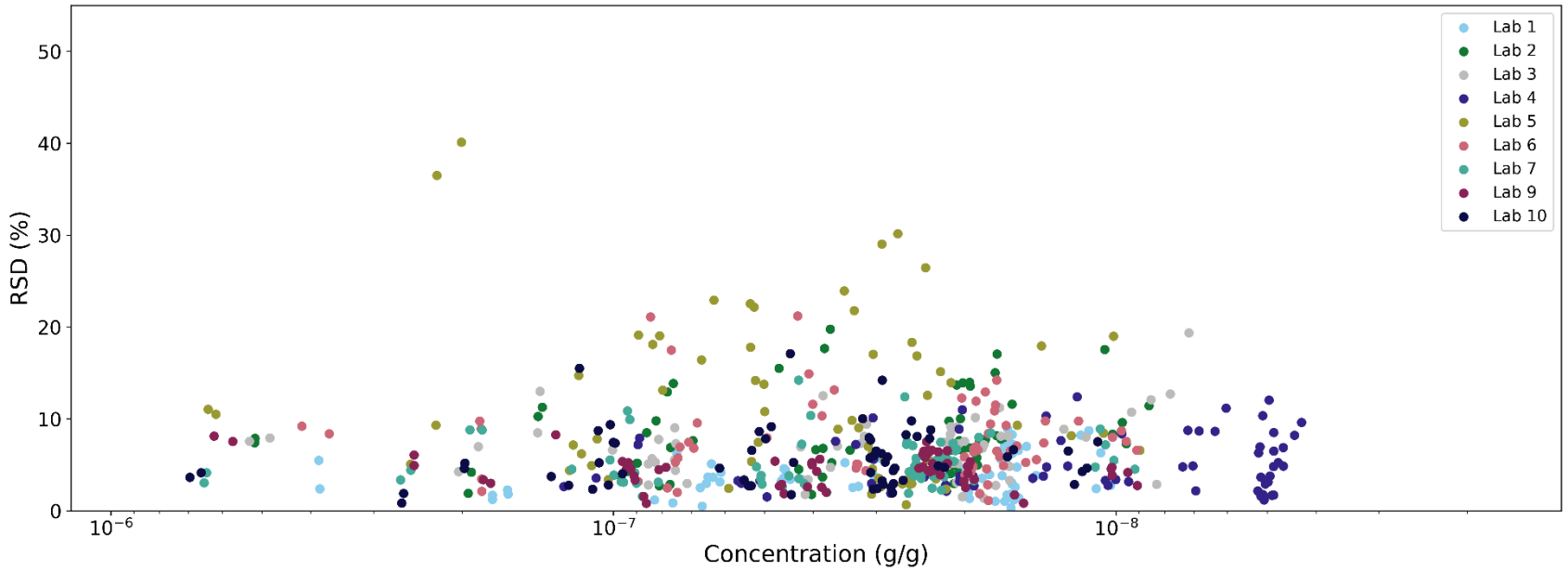


**Figure 6. Percent RSD of the IPR percent recovery values reported in the aqueous IDC dataset, for each of the 40 PFAS target analytes and 24 PFAS EIS compounds, for the nine labs.**

## E. Matrix Spike Recovery

The matrix spike recovery test explores how the method performs with real-world environmental matrices. The aqueous matrices—WW, SW, GW—are sourced from various locations in the environment. Specific information about the source of each aqueous matrix was not provided to IDA. An independent lab added or spiked the 40 target analytes into replicate samples of all the aqueous matrices. Each of the labs received a set of six samples containing the target analytes and an unspiked sample (i.e., aqueous matrix without modifications), for each different type of aqueous matrix, totaling to 91 aqueous study samples. The laboratories were responsible for adding the EIS and NIS compounds to the samples. The number of labs in each of the aqueous matrix datasets IDA analyzed varied either because a lab chose not to participate or the sponsor deemed the results did not qualify (e.g., lab did not follow method correctly). Table 1 includes details about the number of matrices, samples, and labs reporting results for the aqueous matrices.

**Table 1. Number of Real-World Aqueous Matrices and Laboratories Reporting Results**

| Aqueous Matrices | WW[1] | SW[2] | GW[3] |
|---|---|---|---|
| **# of Matrices** | 7 | 3 | 3 |
| **# of Study Samples** | 49 | 21 | 21 |
| **# of Labs Reporting Results** | 8 | 9 | 7 to 8 |

[1] Labs 8 and 10 were not included.

[2] Lab 8 was not included.

[3] Labs 8 and 9 were not included. Lab 1 had results for 2 out of 3 samples.

IDA calculated the percent RSD (Equation 10) for each of the three aqueous matrix datasets for the EIS compounds and target analyte. The percent RSD for the EIS compounds included the recoveries from the unspiked samples in addition to the spiked samples. The percent RSD for the target analytes were from the six spiked samples; those results are shown in Figure 7. The computed overall percent RSD values for the matrix samples ranged from 8.94% to 68.0 % (WW), 6.50% to 104% (SW), and 3.71% to 54.4% (GW). Target analytes PFHxS and PFOS (SW, Lab 1) and 4:2FTS (WW, Lab 6) each had a single reported percent recovery greater than 1000% for a sample, which likely contributed to the high computed overall RSD value for that matrix. Additionally, one of the GW samples did not have any results for PFOS and PFHxS and another GW sample only had one lab result reported for PFHxS. About 10% of the computed percent RSDs are greater than 30% with most of the values associated with the perfluorooctane sulfonamidoacetic acids (FOSAA).

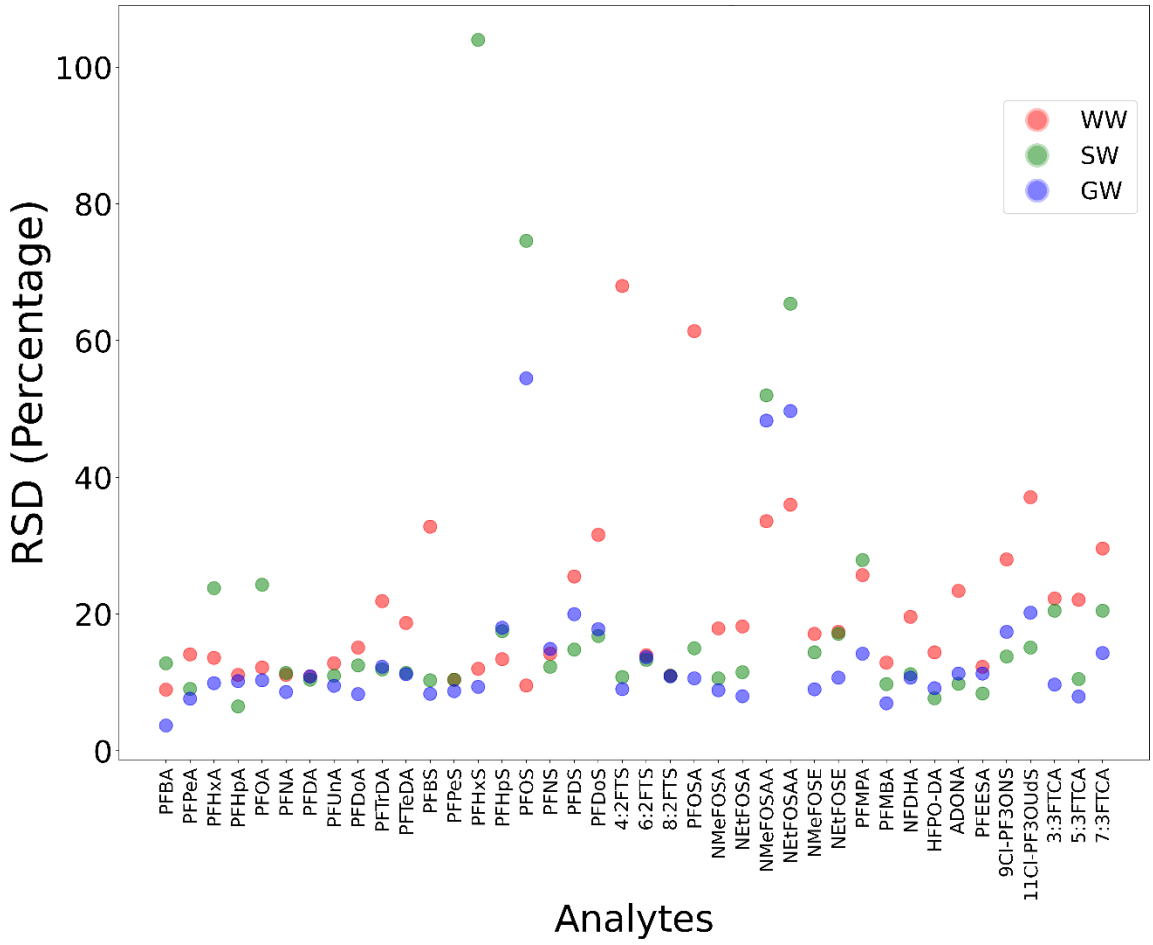


**Figure 7. Percent RSD calculated from all the results across labs in each aqueous matrix spike dataset for each of the 40 PFAS analytes.**

## F. Ongoing Precision and Recovery (OPR)

The OPR test helps to assure the results produced by the labs when analyzing the matrix spike samples remain within the specified precision and recovery limits for the method. Labs use an aliquot of a method blank[49] spiked with analytes and standards. The OPR test also shows the variability within each lab and the reproducibility of a result between labs for the method across the aqueous environmental matrix spike samples.

The PFAS MLV OPR dataset associated with all the aqueous matrices includes nine labs. IDA received each aqueous matrix dataset with OPR values separately and computed the mean percent recovery values and the overall RSD for each matrix. Some labs performed the SW and GW matrix samples together and reported OPR measurements associated with both datasets. Later, IDA received a combined aqueous matrix dataset with

[49] Reagent water that is treated exactly as a sample, including exposure to all glassware, equipment, solvents, reagents, internal standards, and labeled compounds that are used with samples. SERDP/ESTCP, *Study Plan for Multi-Laboratory Validation*.

all the values for the WW, SW, and GW results with the OPR measurements and an additional column to indicate if values were associated with a single matrix (e.g., WW, SW, GW) or more than one dataset (e.g., SW and GW).

The boxplot in Figure 8 is a visualization of the percent recoveries of the reported OPR measurements in the combined aqueous matrix dataset with nine labs for each target analyte and EIS compound. The number of OPR results reported by each lab varied from 3 to 11. Because there were more WW samples, more of the results are associated with the WW dataset compared to the SW and GW datasets (Table 1). Lab 10 did not report OPR values associated with WW datasets while Lab 9 only reported OPR values related to the WW dataset. Lab 6 had the lowest percent recovery values overall, and Lab 9 had the highest percent recoveries for many of the target analytes as seen by the "X" markers in Figure 8. The overall mean percent recovery for the target analytes range from 89.0% to 109% and EIS compounds ranged from 53.2% to 101%. Similar to the IPR results, the FOSA and FOSE EIS compounds have the lowest average and median recovery values, yet the corresponding target analytes recovery values center around 100%.

The scatter plot in Figure 9 shows each lab's within-lab variability (the standard deviation values for each lab used in Equation 8) in their OPR measurements, for every analyte, where the percent RSD is plotted as a function of the concentration. Most points are below 20% percent RSD with the exception of a small number of analytes across several labs. The computed overall percent RSD values (Equation 10) ranged from 7.29% to 15.9% for the target analytes and 7.18% to 26.6% for the EIS compounds.

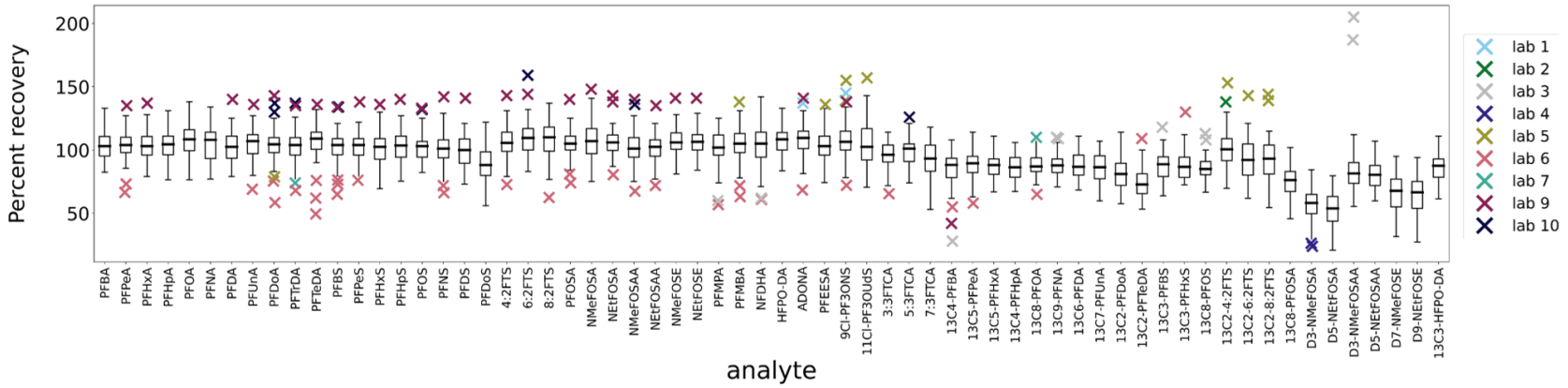


**Figure 8. Percent recovery of the OPR measurements in the WW, SW, GW matrix spike datasets reported by all nine labs, for 40 PFAS target analytes and 24 PFAS EIS compounds.**

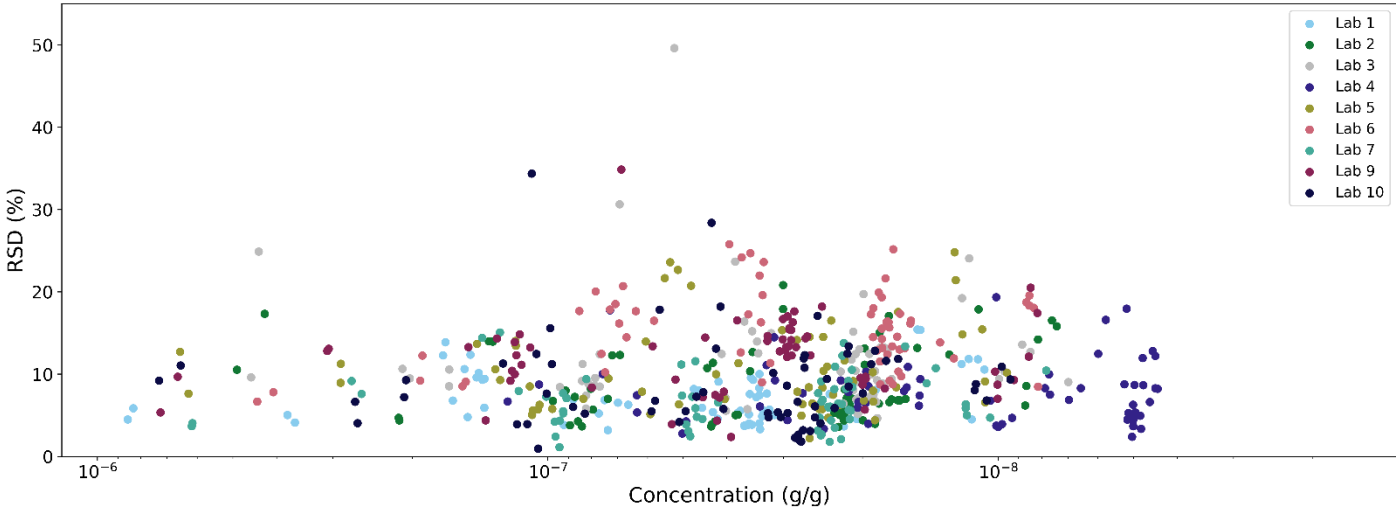


**Figure 9. Percent RSD of the OPR percent recovery values reported in the WW, SW, GW matrix spike datasets calculated for every lab, for each of the 40 PFAS target analytes and 24 PFAS EIS compounds.**

## G. Low-Level Ongoing Precision and Recovery (LLOPR)

The LLOPR test verifies the LOQ with samples spiked at low concentrations. Similar to the OPR test, the percent recovery measurements show the variability within each lab and the reproducibility of a result between labs for the method across the aqueous environmental matrix spike samples.

The PFAS MLV LLOPR dataset associated with all the aqueous matrices includes nine labs. Similar to the OPR dataset, IDA received each aqueous matrix dataset with LLOPR values separately and computed the mean percent recovery values and the overall RSD for each matrix. Later, IDA received a combined aqueous matrix dataset with all the values for the WW, SW, and GW results with the OPR measurements and an additional column to indicate if values were associated with a single matrix (e.g., WW, SW, GW) or more than one dataset (e.g., SW and GW).

The boxplot in Figure 10 is a visualization of the percent recoveries of the reported LLOPR measurements made by nine labs for the target and EIS compounds in the aqueous WW, SW, GW. Because there were more WW samples, more of the results are associated with the WW dataset compared to the SW and GW datasets (Table 1). The number of LLOPR results reported by each lab varied from 3 to 11 values. Lab 10 did not report LLOPR values associated with the WW dataset while Lab 9 only reported LLOPR values related to the WW dataset. Figure 10 shows Lab 9 reported the highest percent recoveries for some of the target analytes and Lab 5 reported the highest percent recoveries for some EIS compounds and a few target analytes. The overall mean percent recovery for the target analytes range from 88.3% to 113% and EIS compounds ranged from 50.8% to 108%. The range of LLOPR recovery values is slightly larger compared to the range of recovery values for the IPR and OPR. Similar to the IPR and OPR results, the FOSA and FOSE EIS compounds have the lowest average and median recovery values yet the corresponding target analytes recovery values center around 100%.

The scatter plot in Figure 11 shows each lab's within-lab variability (the standard deviation values for each lab used in Equation 8) in their LLOPR measurements, for every analyte, where the percent RSD is plotted as a function of the concentration. Most points are below 20% percent RSD with the exception of a small number of analytes across several labs. The computed overall percent RSD values (Equation 10) ranged from 8.22% to 14.3% for the target analytes and 8.03% to 21.8% for the EIS compounds.

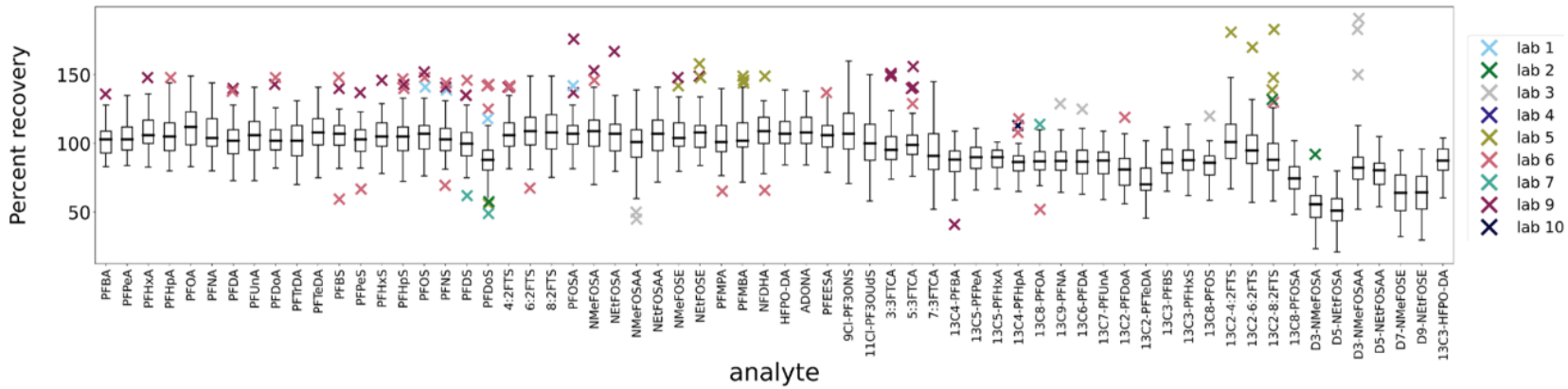


**Figure 10. Percent recovery of the LLOPR measurements in the WW, SW, GW matrix spike datasets reported by nine labs, for 40 PFAS target analytes and 24 PFAS EIS compounds.**

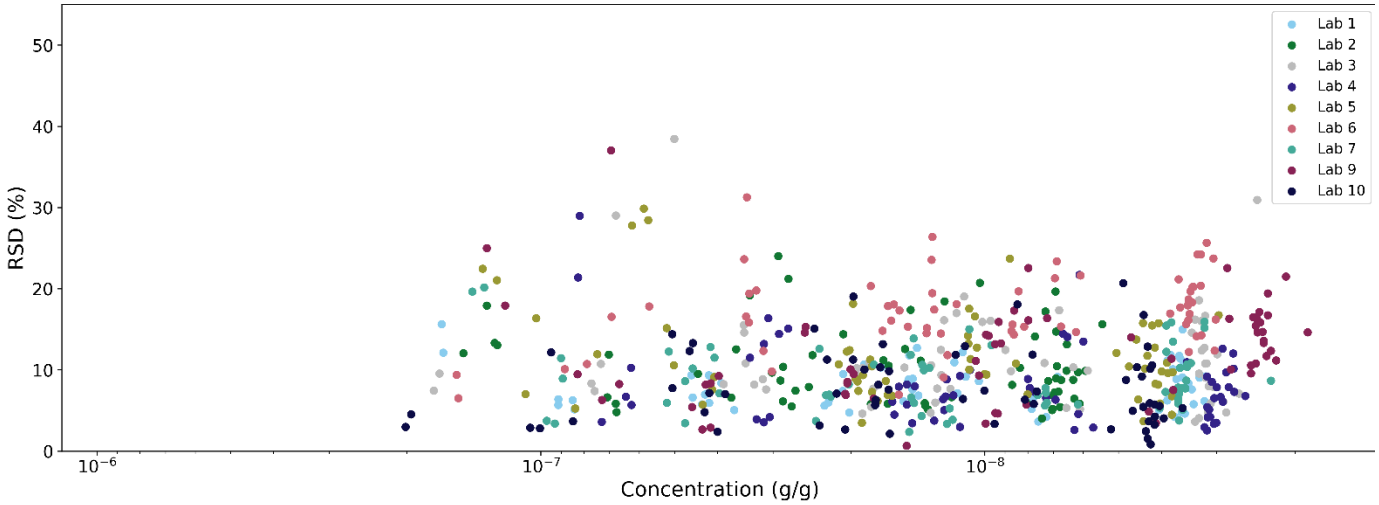


**Figure 11. Percent RSD of the LLOPR percent recovery values reported in the WW, SW, GW matrix spike datasets calculated for every lab, for each of the 40 PFAS target analytes and 24 PFAS EIS compounds.**

# 4. Summary

Method validation is a process that demonstrates that the results generated by conducting the method are reproducible and reliable for the intended purpose. A validated method to quantify PFAS is important in identifying which analytes are present in an area and to set a baseline for future monitoring. IDA analyzed five datasets provided by the sponsor in the PFAS MLV including the ICAL, aqueous IDC, and samples from three environmental matrices: WW, SW, and GW. IDA was blind to the lab identities and sample locations, and was not part of the validation/verification process of the datasets. IDA inspected and evaluated the analysis metrics in the MLV/EPA's ATP, and recommended alternative calculations in instances with a discrepancy between the dataset and formulas. IDA then calculated statistical values for the overall method performance measures including: calibration linearity, LOQVER, IPR, OPR, LLOPR, and matrix spike recovery. These values were provided to the sponsor with the intent that the statistical values will inform the QC acceptance criteria set by the EPA for the method.

Table 2 summarizes the PFAS MLV datasets, validation tests, performance measures, and range of values that IDA computed. The overall mean percent recovery and percent RSD values for the target analytes and the EIS compounds were fairly consistent across the IPR, OPR, and LLOPR tests, respectively. The percent RSD values for the matrix spike recoveries of the target analytes were much broader which is likely due to a smaller number of samples for the SW and GW matrices and specific target analytes being problematic. Additional details about the specific analyses include:

- The ICAL dataset included summary statistics for each of the labs' calibration trials and separately reported calibration concentration values from each lab. IDA was unable to independently verify the summary statistics and identified several instances where the reported average RF and standard deviation did not align with the RSD value. IDA supplied an alternative statistical formula for calculating the pooled RSD value for the method as the dataset was not structured for using the formula outlined in the EPA's ATP.
- The MDL dataset was structured to use the EPA's ATP statistical formulas and IDA calculated an aqueous pooled MDL value.
- The LOQVER test was not described in the EPA's ATP, rather the MLV study plan cited a DoD reference for calculating precision and bias. The LOQVER dataset was comprised of mostly single data point values for each lab and was

not structured to calculate a lab's precision. IDA calculated each lab's percent relative bias but was unable to evaluate the degree of systematic error.

- The IPR dataset was structured to use the EPA's ATP statistical formulas and IDA calculated a mean percent recovery value and an overall percent RSD.
- The matrix spike dataset comprised of percent recovery values of target analytes spiked prior to the delivery of the matrix samples to the labs. The EPA ATP procedure for determining the method performance of a matrix is not for an isotopic dilution method like this PFAS method. IDA calculated the overall percent RSD for each matrix.
- The OPR and LLOPR were structured to use the EPA's ATP statistical formulas. Some labs performed the SW and GW tests together and reported OPR and LLOPR measurements associated with both datasets. IDA calculated a mean percent recovery value and an overall percent RSD for each WW, SW, and GW matrix and across all aqueous matrices (e.g., WW, SW, GW) combined.

**Table 2. Summary of PFAS MLV Statistical Analyses for the Aqueous Datasets**

| PFAS MLV Dataset | Analysis Test in MLV | MLV Data Allowed Use of ATP Formula? | Performance Metric in EPA's ATP | Target Analyte Performance[1] | Extracted Internal Standard Performance[1] |
|---|---|---|---|---|---|
| **ICAL** | Calibration Linearity | No | pooled percent relative standard deviation (RSD) | 7.31% to 13.8% | 4.11% to 12.1% |
| **Aqueous IDC** | Method Detection Limit (MDL) | Yes | pooled MDL | 0.315 to 9.89 ng/L | N/A |
| | Limit of Quantitation Verification | N/A | N/A | N/A | N/A |
| | Initial Precision and Recovery | Yes | mean percent recovery | 95.0% to 109% | 69.1% to 98.1% |
| | | | percent RSD | 3.35% to 11.5% | 5.36% to 17.2% |
| **Matrix Samples** | Ongoing Precision and Recovery | Yes | mean percent recovery | 89.0% to 109% | 53.2% to 101% |
| | | | percent RSD | 7.29% to 15.9% | 7.18% to 26.6% |
| | Low-Limit Ongoing Precision and Recovery | Yes | mean percent recovery | 88.3% to 113% | 50.8% to 108% |
| | | | percent RSD | 8.22% to 14.3% | 8.03% to 21.8% |

| PFAS MLV Dataset | Analysis Test in MLV | MLV Data Allowed Use of ATP Formula? | Performance Metric in EPA's ATP | Target Analyte Performance[1] | Extracted Internal Standard Performance[1] |
|---|---|---|---|---|---|
| | Matrix Spike WW[2] | No | percent RSD | 8.94% to 68.0% | N/A |
| | Matrix Spike SW | | | 6.50% to 104% | N/A |
| | Matrix Spike GW[2] | | | 3.71% to 54.4% | N/A |

[1] Nine labs reported values in most datasets.

[2] Only eight labs reported values.

# Appendix A.
# PFAS MLV Analytes

**Table A-1. List of PFAS Analytes and Standards in MLV**

| Classification Type | PFAS Acronym | Quantification Reference |
|---|---|---|
| **Target Analyte** | | |
| **perfluoroalkyl carboxylic acids** | **PFBA** | $^{13}C_4$-PFBA |
| | **PFPeA** | $^{13}C_5$-PFPeA |
| | **PFHxA** | $^{13}C_5$-PFHxA |
| | **PFHpA** | $^{13}C_4$-PFHpA |
| | **PFOA** | $^{13}C_8$-PFOA |
| | **PFNA** | $^{13}C_9$-PFNA |
| | **PFDA** | $^{13}C_6$-PFDA |
| | **PFUnA** | $^{13}C_7$-PFUnA |
| | **PFDoA** | $^{13}C_2$-PFDoA |
| | **PFTrDA** | avg. $^{13}C_2$-PFTeDA and $^{13}C_2$-PFDoA |
| | **PFTeDA** | $^{13}C_2$-PFTeDA |
| **perfluoroalkyl sulfonic acids** | **PFBS** | $^{13}C_3$-PFBS |
| | **PFPeS** | $^{13}C_3$-PFHxS |
| | **PFHxS** | $^{13}C_3$-PFHxS |
| | **PFHpS** | $^{13}C_8$-PFOS |
| | **PFOS** | $^{13}C_8$-PFOS |
| | **PFNS** | $^{13}C_8$-PFOS |
| | **PFDS** | $^{13}C_8$-PFOS |
| | **PFDoS** | $^{13}C_8$-PFOS |
| **fluorotelomer sulfonic acids** | **4:2FTS** | $^{13}C_2$-4:2FTS |
| | **6:2FTS** | $^{13}C_2$-6:2FTS |
| | **8:2FTS** | $^{13}C_2$-8:2FTS |
| **perfluorooctane sulfonamides** | **PFOSA** | $^{13}C_8$-PFOSA |
| | **NMeFOSA** | $D_3$-NMeFOSA |
| | **NEtFOSA** | $D_5$-NEtFOSA |
| **perfluorooctane sulfonamidoacetic acids** | **NMeFOSAA** | $D_3$-NMeFOSAA |
| | **NEtFOSAA** | $D_5$-NEtFOSAA |
| **perfluorooctane sulfonamide ethanols** | **NMeFOSE** | $D_7$-NMeFOSE |
| | **NEtFOSE** | $D_9$-NEtFOSE |

| Classification Type | PFAS Acronym | Quantification Reference |
|---|---|---|
| **per-and-polyfluoroether carboxylic acids** | **PFMPA** | $^{13}C_5$-PFPeA |
| | **PFMBA** | $^{13}C_5$-PFPeA |
| | **NFDHA** | $^{13}C_5$-PFHxA |
| | **HFPO-DA** | $^{13}C_3$-HFPO-DA |
| | **ADONA** | $^{13}C_3$-HFPO-DA |
| **ether sulfonic acids** | **PFEESA** | $^{13}C_5$-PFHxA |
| | **9Cl-PF3ONS** | $^{13}C_3$-HFPO-DA |
| | **11Cl-PF3OUdS** | $^{13}C_3$-HFPO-DA |
| **fluorotelomer carboxylic acids** | **3:3FTCA** | $^{13}C_5$-PFPeA |
| | **5:3FTCA** | $^{13}C_5$-PFHxA |
| | **7:3FTCA** | $^{13}C_5$-PFHxA |
| **EIS compounds** | | |
| **perfluoroalkyl carboxylic acids** | **$^{13}C_4$-PFBA** | $^{13}C_3$-PFBA |
| | **$^{13}C_5$-PFPeA** | $^{13}C_2$-PFHxA |
| | **$^{13}C_5$-PFHxA** | $^{13}C_2$-PFHxA |
| | **$^{13}C_4$-PFHpA** | $^{13}C_2$-PFHxA |
| | **$^{13}C_8$-PFOA** | $^{13}C_4$-PFOA |
| | **$^{13}C_9$-PFNA** | $^{13}C_5$-PFNA |
| | **$^{13}C_6$-PFDA** | $^{13}C_2$-PFDA |
| | **$^{13}C_7$-PFUnA** | $^{13}C_2$-PFDA |
| | **$^{13}C_2$-PFDoA** | $^{13}C_2$-PFDA |
| | **$^{13}C_2$-PFTeDA** | $^{13}C_2$-PFDA |
| **perfluoroalkyl sulfonic acids** | **$^{13}C_3$-PFBS** | $^{18}O_2$-PFHxS |
| | **$^{13}C_3$-PFHxS** | $^{18}O_2$-PFHxS |
| | **$^{13}C_8$-PFOS** | $^{13}C_4$-PFOS |
| | **$^{13}C_2$-4:2FTS** | $^{18}O_2$-PFHxS |
| | **$^{13}C_2$-6:2FTS** | $^{18}O_2$-PFHxS |
| | **$^{13}C_2$-8:2FTS** | $^{18}O_2$-PFHxS |
| **perfluorooctane sulfonamides** | **$^{13}C_8$-PFOSA** | $^{13}C_4$-PFOS |
| | **$D_3$-NMeFOSA** | $^{13}C_4$-PFOS |
| | **$D_5$-NEtFOSA** | $^{13}C_4$-PFOS |
| **perfluorooctane sulfonamidoacetic acids** | **$D_3$-NMeFOSAA** | $^{13}C_4$-PFOS |
| | **$D_5$-NEtFOSAA** | $^{13}C_4$-PFOS |
| **perfluorooctane sulfonamide ethanols** | **$D_7$-NMeFOSE** | $^{13}C_4$-PFOS |
| | **$D_9$-NEtFOSE** | $^{13}C_4$-PFOS |
| **per-and-polyfluoroether carboxylic acids** | **$^{13}C_3$-HFPO-DA** | $^{13}C_2$-PFHxA |

| Classification Type | PFAS Acronym | Quantification Reference |
|---|---|---|
| **NIS compounds** | | |
| **perfluoroalkyl carboxylic acids** | **$^{13}C_3$-PFBA** | N/A |
| | **$^{13}C_2$-PFHxA** | N/A |
| | **$^{13}C_4$-PFOA** | N/A |
| | **$^{13}C_5$-PFNA** | N/A |
| | **$^{13}C_2$-PFDA** | N/A |
| **perfluoroalkyl sulfonic acids** | **$^{18}O_2$-PFHxS** | N/A |
| | **$^{13}C_4$-PFOS** | N/A |

# Appendix B.
# Interpretation of Box and Whisker Plots

Box and whisker plots graphically present the data without making assumptions about the distribution (Figure B-1). The plots also show the spread and skewness in values across a grouping of data. The center line in the box is the median. The top and bottom of the box covers where half of the data are found from the 25th to 75th percentiles, with the length of the box defining the interquartile range. The two whisker lines outside of the box indicate the maximum and minimum of the dataset up to the interquartile (IQR) range (the range defined in the box) multiplied by 1.5. The circles are data points outside the range defined in the whiskers. [50]

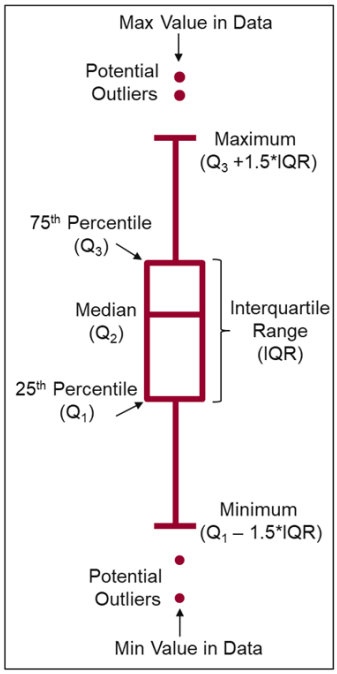


**Figure B-1. Box and whisker plot description.**

[50] "Box Plot with Minitab," Lean Sigma Corporation, December 22, 2015, https://www.leansigmacorporation.com/box-plot-with-minitab/?nab=1.

# Appendix C.
# Data Overview Plots

IDA analyzed five datasets in the PFAS MLV including the ICAL, aqueous IDC, and samples from three environmental matrices: WW, SW, and GW. Each environmental matrix included a total of six samples spiked with the 40 PFAS analytes where three samples were spiked at a "low" concentration and three samples were spiked at a "high" concentration. The following plots provide a visualization combining data across the datasets to show the calibration concentrations, calculated MDL values, and spiked and measured concentration values for each of the 40 target analytes across the nine labs (Figures C-1–C-4).

Each plot includes the following:

- 'X' data points indicate the calibration concentrations reported by the labs.
- Blue line is the pooled MDL value calculated by IDA.
- Light blue shading is the lowest and highest lab MDL value calculated by IDA.
- Green line is the low spike concentration of the analyte added in each of the aqueous matrix samples (e.g., WW, SW, GW) reported in the dataset.
- Purple line is the high spike concentration of the analyte added in each of the aqueous matrix samples (e.g., WW, SW, GW) reported in the datasets.
- Color matched boxplots are the labs measured spike concentrations corrected for any measured analyte in the unspiked sample in the aqueous matrix samples (e.g., WW, SW, GW), reported in the datasets.

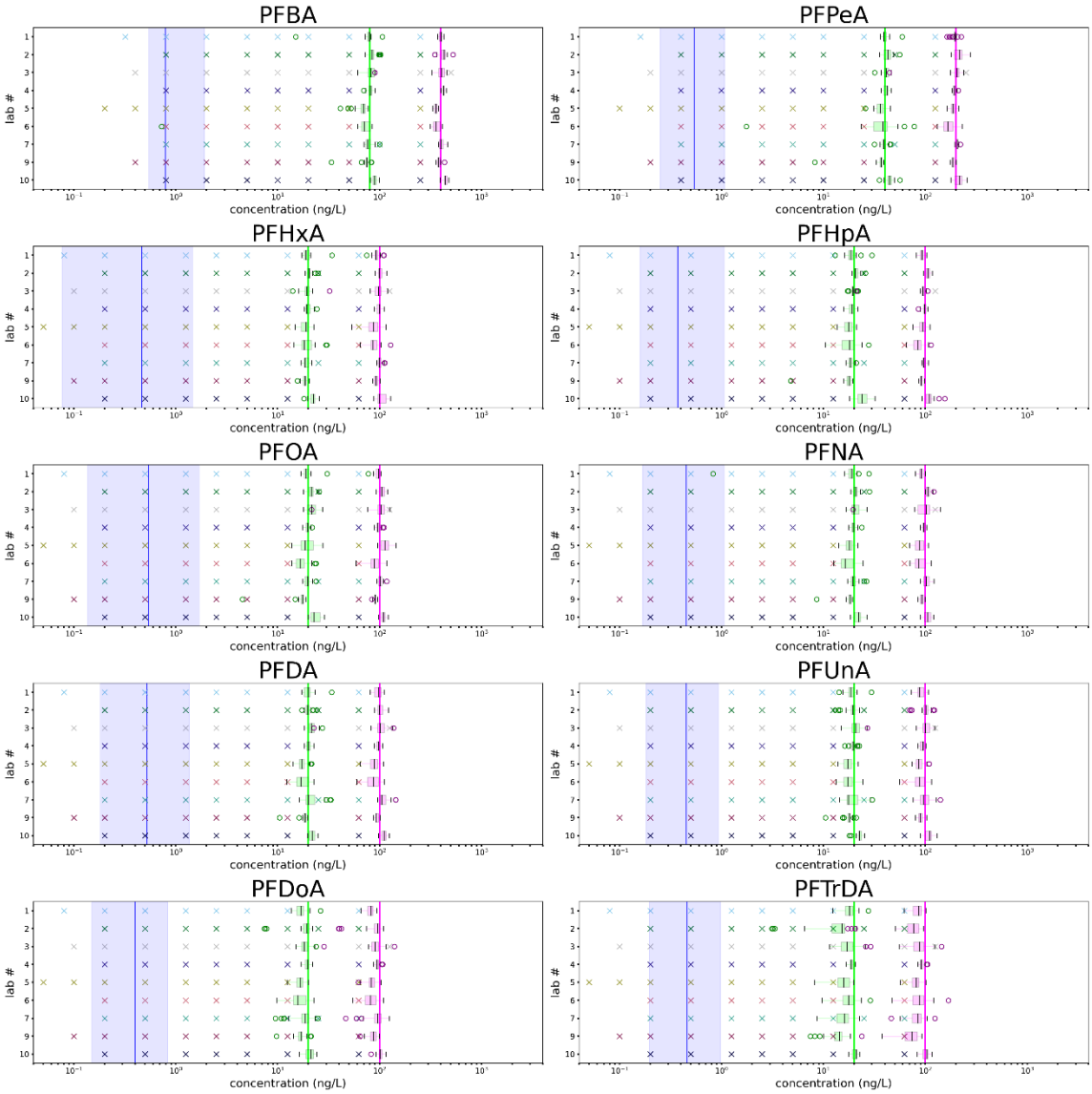


**Figure C-1. Plot 1 of 4 depicting the calibration concentrations, calculated MDL values, and spiked and measured concentration values of 10 target analytes across the 9 labs.**

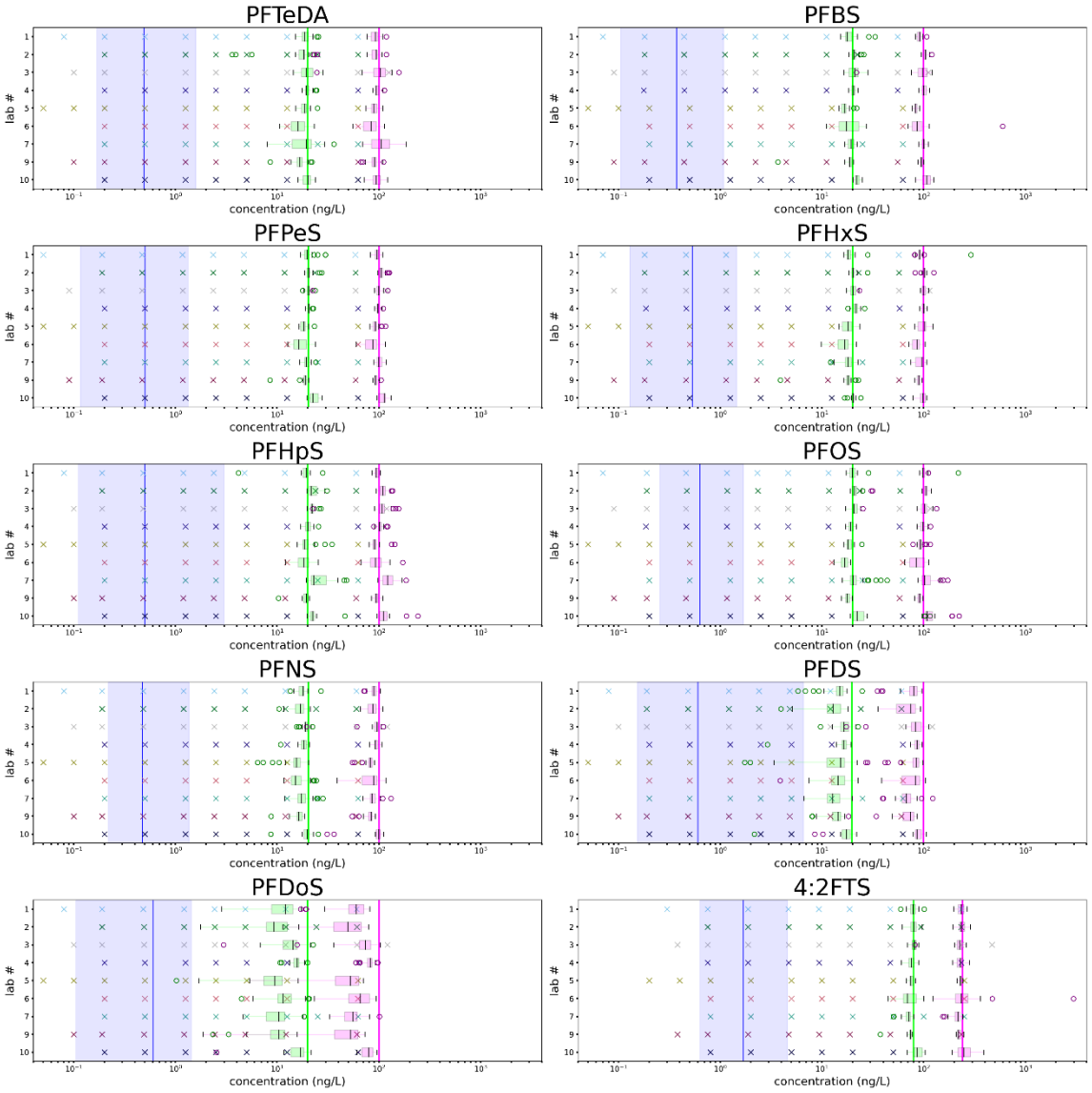


**Figure C-2. Plot 2 of 4 depicting the calibration concentrations, calculated MDL values, and spiked and measured concentration values of 10 target analytes across the 9 labs.**

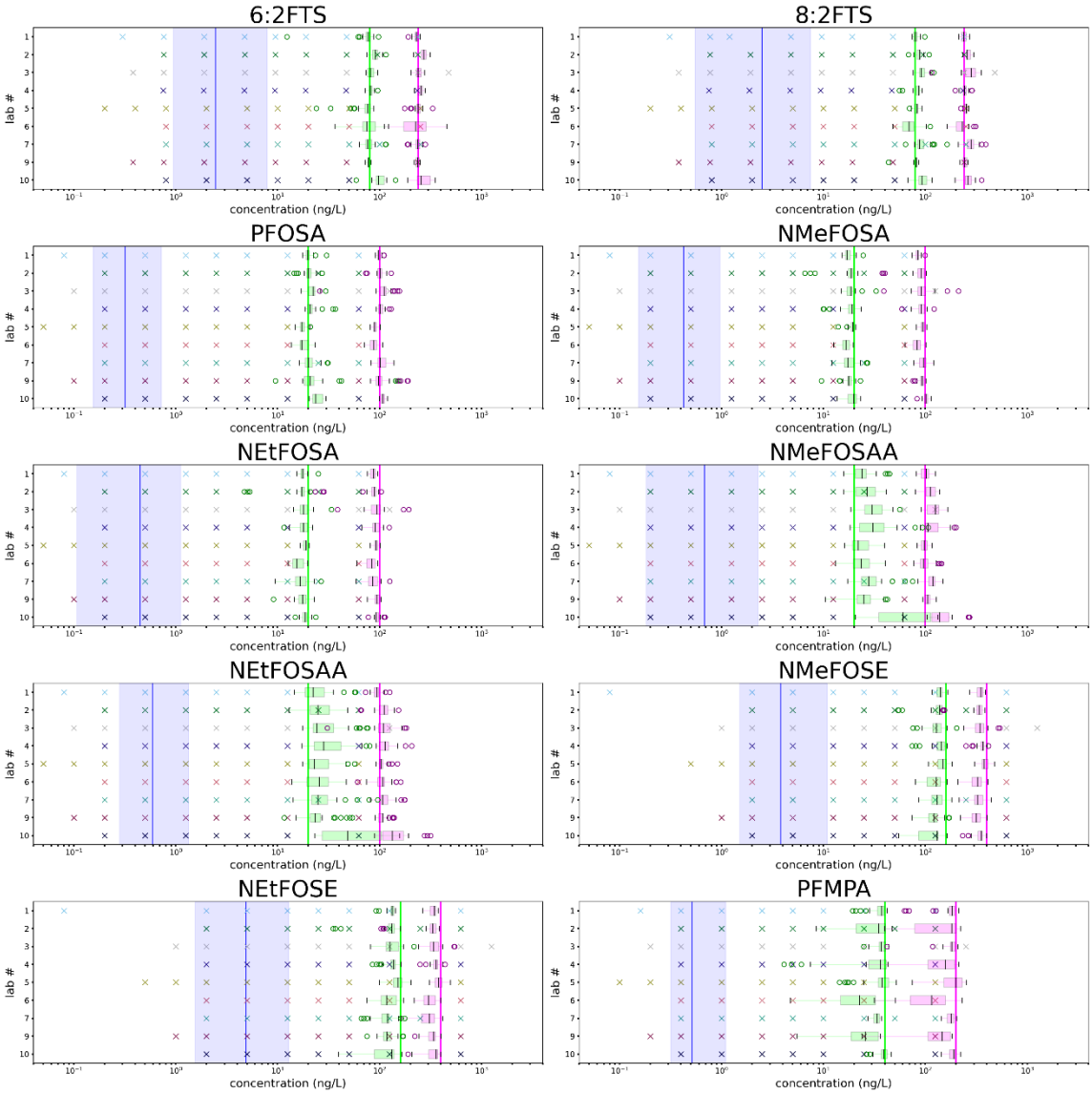


**Figure C-3. Plot 3 of 4 depicting the calibration concentrations, calculated MDL values, and spiked and measured concentration values of 10 target analytes across the 9 labs.**

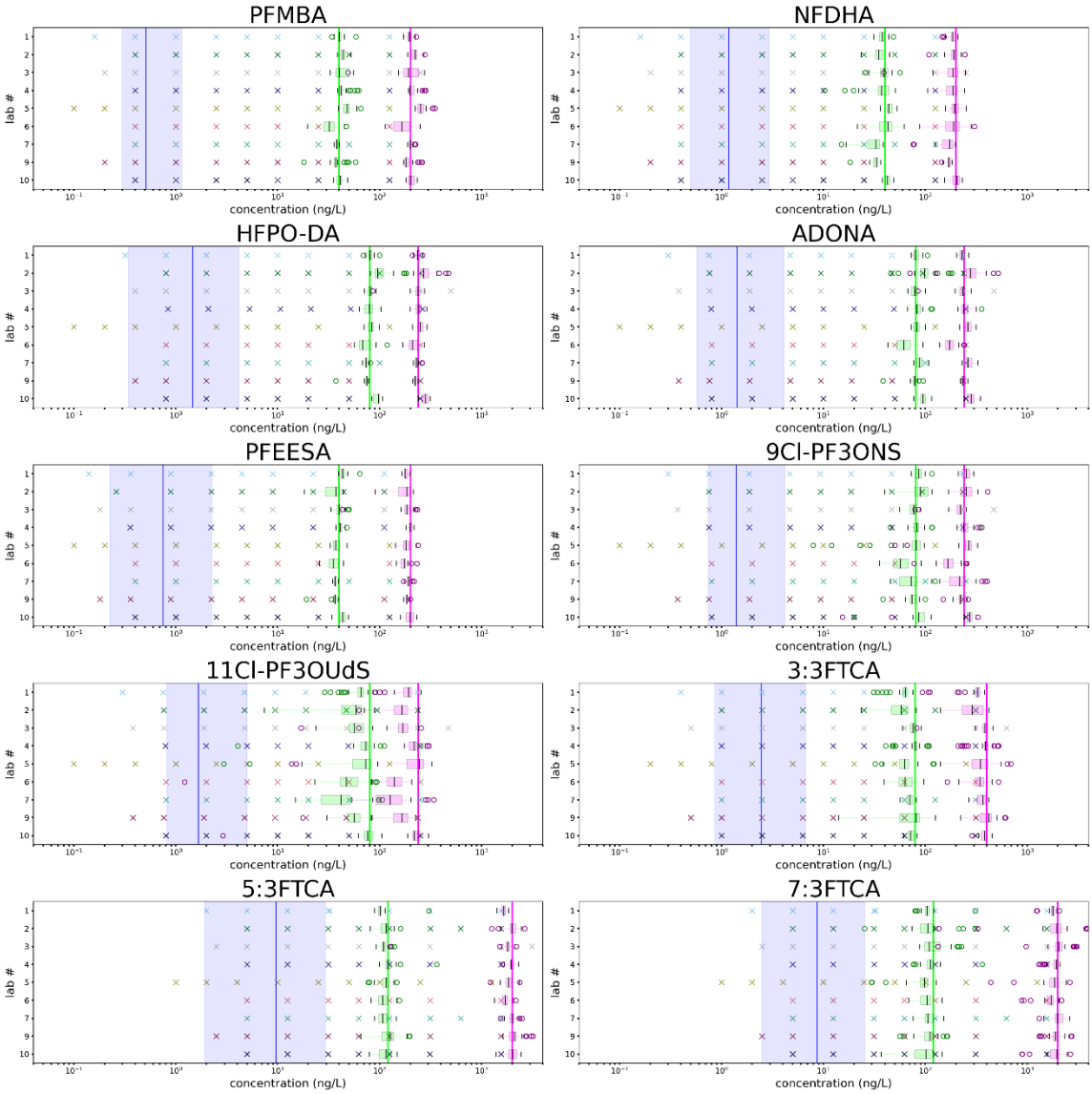


**Figure C-4. Plot 4 of 4 depicting the calibration concentrations, calculated MDL values, and spiked and measured concentration values of 10 target analytes across the 9 labs.**

# Appendix D.
# List of Tables in the Digital Appendix

Table D-1 is a list of the IDA-generated tables as CSV files for each of the listed PFAS MLV datasets included in the digital appendix. Accompanied with each table is a TXT file that includes the MLV dataset version and description of the data in each column as well as the formula or citation to any statistical equations.

**Table D-1. List of IDA Generated Tables Corresponding to the PFAS MLV Datasets**

| Dataset | Version | IDA Table File |
|---|---|---|
| **ICAL** | ICAL Concentrations_08182022.xlsx<br>ICAL Average RF_05182023.xlsx | ICAL_calibration_V0_220907_093746.csv<br>AverageRF_ICAL_results_V4_230519_091739.csv |
| **IDC** | RW_DBexport_V1_20230426.csv | MDL_results_V1_230503_215159.csv<br>LOQVER_results_V1_230503_215921.csv<br>IPR_results_V1_230503_215140.csv |
| **WW** | WW_DBexport_V7_20230328.csv | LLOPR_results_V4_230406_212723.csv<br>OPR_results_V4_230406_212237.csv<br>Matrix_EIS_results_V4_230406_212819.csv<br>Matrix_sample_results_V4_230406_211329.csv<br>Matrix_compiled_results_V4_230406_211329.csv<br>MB_results_V4_230406_212853.csv |
| **SW** | SW_DBexport_V4_20230407.csv | LLOPR_results_V0_230411_080130.csv<br>OPR_results_V0_230411_080146.csv<br>Matrix_EIS_results_V0_230411_080212.csv<br>Matrix_sample_results_V0_230411_080232.csv<br>Matrix_compiled_results_V0_230411_080232.csv<br>MB_results_V0_230411_080058.csv |
| **GW** | GW_DBexport_V6_20230417b.csv | LLOPR_results_V0_230421_074935.csv<br>OPR_results_V0_230420_183700.csv<br>Matrix_EIS_results_V0_230420_175829.csv<br>Matrix_sample_results_V0_230421_153930.csv<br>Matrix_compiled_results_V0_230421_153930.csv<br>MB_results_V0_230420_183436.csv |
| **All Aqueous** | WW_SW_GW_EXPORT_20230605.csv | LLOPR_results_V1_230607_124655.csv<br>OPR_results_V1_230607_124749.csv<br>Matrix_EIS_results_V1_230607_124828.csv<br>Matrix_NIS_results_V1_230607_124909.csv |

# Illustrations

## Tables

Table 1. Number of Real-World Aqueous Matrices and Laboratories Reporting Results....26
Table 2. Summary of PFAS MLV Statistical Analyses for the Aqueous Datasets ....34
Table A-1. List of PFAS Analytes and Standards in MLV .... A-1
Table D-1. List of IDA Generated Tables Corresponding to the PFAS MLV Datasets .... D-1

## Figures

Figure 1. Z-score of the RSDs of the RFs reported for each of the three calibration tests conducted by every lab, for 40 target PFAS analytes and 24 EIS PFAS compounds. ....16
Figure 2. MDL values calculated for nine labs that went into the computed pooled MDL for 40 target PFAS analytes. The fluorotelomer sulfonic acids (4:2FTS, 6:2FTS, and 8:2FTS), perfluorooctane sulfonamide ethanols (NMeFOSE and NEtFOSE), and fluorotelomer carboxylic acids (3:3FTCA, 5:3FTCA, and 7:3FTCA) are displayed on separate axes to avoid visual suppression of smaller MDL values. A break in the y-axes of both plots avoids visual suppression of reported MDL values within each group....18
Figure 3. Percent bias of the concentrations measured in the aqueous IDC dataset reported by nine labs, for 40 PFAS target analytes and 24 EIS compounds. ....20
Figure 4. Percent recovery of the four measurements in the aqueous IPR dataset reported by all nine labs, for 40 PFAS target analytes and 24 PFAS EIS compounds. ....22
Figure 5. Plot of the modified Horwitz curve depicting the relationship between concentration and percent RSD. ....24
Figure 6. Percent RSD of the IPR percent recovery values reported in the aqueous IDC dataset, for each of the 40 PFAS target analytes and 24 PFAS EIS compounds, for the nine labs. ....25
Figure 7. Percent RSD calculated from all the results across labs in each aqueous matrix spike dataset for each of the 40 PFAS analytes. ....27
Figure 8. Percent recovery of the OPR measurements in the WW, SW, GW matrix spike datasets reported by all nine labs, for 40 PFAS target analytes and 24 PFAS EIS compounds. ....29
Figure 9. Percent RSD of the OPR percent recovery values reported in the WW, SW, GW matrix spike datasets calculated for every lab, for each of the 40 PFAS target analytes and 24 PFAS EIS compounds. ....29

Figure 10. Percent recovery of the LLOPR measurements in the WW, SW, GW matrix spike datasets reported by nine labs, for 40 PFAS target analytes and 24 PFAS EIS compounds. ....31
Figure 11. Percent RSD of the LLOPR percent recovery values reported in the WW, SW, GW matrix spike datasets calculated for every lab, for each of the 40 PFAS target analytes and 24 PFAS EIS compounds. ....31
Figure B-1. Box and whisker plot description. ....B-1
Figure C-1. Plot 1 of 4 depicting the calibration concentrations, calculated MDL values, and spiked and measured concentration values of 10 target analytes across the 9 labs. ....C-2
Figure C-2. Plot 2 of 4 depicting the calibration concentrations, calculated MDL values, and spiked and measured concentration values of 10 target analytes across the 9 labs. ....C-3
Figure C-3. Plot 3 of 4 depicting the calibration concentrations, calculated MDL values, and spiked and measured concentration values of 10 target analytes across the 9 labs. ....C-4
Figure C-4. Plot 4 of 4 depicting the calibration concentrations, calculated MDL values, and spiked and measured concentration values of 10 target analytes across the 9 labs. ....C-5

# References


*ACS Reagent Chemicals*, Part 1: Introduction and Definitions. Washington, DC: ACS Publications, 2017. https://pubs.acs.org/doi/book/10.1021/acsreagents.

Agency for Toxic Substances and Disease Registry (ATSDR). "What are PFAS." November 1, 2022. https://www.atsdr.cdc.gov/pfas/health-effects/overview.html.

"Assignment and Presentation of Uncertainties of the Numerical Results of Thermodynamic Measurements." *Pure and Applied Chemistry* 53 (9) (1981): 1805–1826. http://dx.doi.org/10.1351/pac198153091805.

Burrows, R., and J. Parr. "Evaluating the Goodness of Instrument Calibration for Chromatography Procedures." *LCGC Supplements* 38 (11) (2020): 35–38. https://www.chromatographyonline.com/view/evaluating-the-goodness-of-instrument-calibration-for-chromatography-procedures.

Buytendyk, A, M., S. C. Runkel, and S. M. Cazares. "Data Compilation in Support of Single Laboratory Validation of a Novel Per- and Polyfluoroalkyl Substances (PFAS) Detection Method for Environmental Matrices." IDA Document D-22794. Alexandria, VA: Institute for Defense Analyses, 2021.

Code of Federal Regulations (CFR), Title 40, Part 136, Appendix B.

Department of Defense, Department of Energy. *DoD Quality Systems Manual Version 5.4*, Module 4, Section 1.5.2. Washington, DC: Department of Defense, Department of Energy, 2021: 77–78. https://www.denix.osd.mil/edqw/denix-files/sites/43/2021/10/QSM-Version-5.4-FINAL.pdf.

Department of Defense, Environmental Cleanup and Compliance. "Per- and Polyfluoroalkyl Substances (PFAS): A Nation Issue that Requires National Solutions." https://www.acq.osd.mil/eie/eer/ecc/pfas/.

Glaser, J. A., D. L. Forest, G. D. McKee, S. A. Quave, and W. L. Budde. "Trace analyses for wastewaters." *Environmental Science & Technology* 15 (12) (1981): 1426–1435. https://doi.org/10.1021/es00094a002.

Green, M. J. "Peer Reviewed: A Practical Guide to Analytical Method Validation." *Analytical Chemistry* 68 (9) (1983): 305A–309A. https://doi.org/10.1021/ac961912f.

Horwitz, W., and R. Albert. "The Horwitz ratio (HorRat): A useful index of method performance with respect to precision." *Journal of Association of Official Analytical Chemists (AOAC) International* 89 (4) (2006): 1095–1109. https://doi.org/10.1093/jaoac/89.4.1095.

Horwitz, W., L. R. Kamps, and K.W. Boyer. "Quality Assurance in the Analysis of Foods for Trace Constituents." *Journal of Association of Official Analytical*

*Chemists (AOAC) International* 63 (6) (1980): 1344–1354. https://doi.org/10.1093/jaoac/63.6.1344.

International Union of Pure and Applied Chemistry (IUPAC). *Compendium of Chemical Terminology*, 2nd ed., compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford, 1997. https://doi.org/10.1351/goldbook.

Keith, L. H., W. Crummett, J. Deegan, R. A. Libby, J. K. Taylor, and G. Wentler. "Principles of environmental analysis." *Analytical Chemistry* 55 (14) (1983): 2210–2218. https://doi/10.1021/ac00264a003.

Lean Sigma Corporation. "Box Plot with Minitab." December 22, 2015., https://www.leansigmacorporation.com/box-plot-with-minitab/?nab=1.

Massart, D. L., J. Smeyers-Verbeke, and Y. V. Heyden. "Benchmarking Analytical Methods Horwitz Curve." *LCGC Europe* 18 (10) (2005): 528–531. https://www.chromatographyonline.com/view/benchmarking-analytical-methods-horwitz-curve.

Royal Society of Chemistry. "The amazing Horwitz function." *AMC Technical Brief* no. 17, (2004), ed. M. Thompson. https://www.rsc.org/images/horwitz-function-technical-brief-17_tcm18-214859.pdf.

SERDP/ESTCP. *Study Plan for Multi-Laboratory Validation of Draft EPA Method 1633 – PFAS in Aqueous, Solid, Biosolids, and Tissue Samples by LC-MS/MS*. March 2022.

Thompson, M. "Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing." *Analyst* 125 (2000): 385–386. https://doi.org/10.1039/B000282H.

Thompson, M, S. Ellison, and R. Wood. "Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC Technical Report)." *Pure and Applied Chemistry* 74 (5) (2002): 835–855. https://doi.org/10.1351/pac200274050835.

University of Tartu. "LC-MS Method Validation." https://sisu.ut.ee/lcms_method_validation.

University of Tartu. "LC-MS Method Validation: 3.1. Linearity." https://sisu.ut.ee/lcms_method_validation/31-linearity.

U.S. Environmental Protection Agency. "CWA Analytical Methods for Per- and Polyfluorinated Alkyl Substances (PFAS)." July 28, 2023. https://www.epa.gov/cwa-methods/cwa-analytical-methods-and-polyfluorinated-alkyl-substances-pfas.

U.S. Environmental Protection Agency. "PFAS Explained." April 10, 2023. https://www.epa.gov/pfas/pfas-explained.

U.S. Environmental Protection Agency. *Protocol for Review and Validation of new Methods for Regulated Organic and Inorganic Analytes in Wastewater Under EPA's Alternative Test Procedure Program.* EPA 821-B-18-001. Washington, DC: Environmental Protection Agency, February 2018.

https://www.epa.gov/sites/default/files/2018-03/documents/chemical-atp-protocol_feb-2018.pdf.

U.S. Environmental Protection Agency. “Summary of the Clean Water Act.” June 22, 2023. https://www.epa.gov/laws-regulations/summary-clean-water-act.

# Abbreviations

| | |
|---|---|
| ATP | Alternate Test Procedure |
| CFR | Code of Federal Regulations |
| CS | calibration sample |
| DoD | Department of Defense |
| EIS | extracted internal standard |
| EPA | Environmental Protection Agency |
| ESTCP | Environmental Security Technology Certification Program |
| FOSA | perfluorooctane sulfonamides |
| FOSE | perfluorooctane sulfonamide ethanols |
| GW | groundwater |
| ICAL | initial calibration |
| IDC | initial demonstration of capability |
| IPR | initial precision and recovery |
| IQR | interquartile range |
| IUPAC | International Union of Pure and Applied Chemistry |
| LC-MS/MS | liquid chromatography-tandem mass spectrometry |
| LOD | limit of detection |
| LOQ | limit of quantitation |
| LOQVER | limit of quantitation verification |
| LLOPR | low-level ongoing precision and recovery |
| MDL | method detection limit |
| ML | method limit |
| MLV | multi-lab validation |
| NIS | non-extracted internal standard |
| OPR | ongoing precision and recovery |
| PFAS | per- and polyfluoroalkyl substances |
| ppb | parts per billion |
| QC | quality control |
| SW | surface water |
| RF | response factor |
| RSD | relative standard deviation |
| SERDP | Strategic Environmental Research and Development Program |
| WW | waste water |

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| August 2023 | FINAL | |

**4. TITLE AND SUBTITLE**

Statistical Methods in the Multi-Laboratory Validation of a PFAS Measurement Method

**5a. CONTRACT NUMBER**

HQ0034-19-D-0001

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

Buytendyk, Allyson M.
Fisher, Katherine I.
Pleasant, Tyler J.
Runkel, Sara
Shah, Jay S.
Silk, John J.

**5d. PROJECT NUMBER**

AM-2-5056

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Institute for Defense Analyses
Systems and Analyses Center
730 East Glebe Road
Alexandra, VA 22305-3086

**8. PERFORMING ORGANIZATION REPORT NUMBER**

IDA Product 3000051

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

SERDP and ESTCP
4800 Mark Center Drive
Suite 16F16
Alexandria, VA 22350-3605

**10. SPONSOR/MONITOR'S ACRONYM(S)**

SERDP/ESTCP

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION/AVAILABILITY STATEMENT**

Distribution Statement A. Approved for public release; distribution is unlimited.

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

The Institute for Defense Analyses (IDA) is conducting the statistical analyses in the Department of Defense (DoD)-sponsored multi-laboratory validation (MLV) study of the per- and polyfluoroalkyl substances (PFAS) measurement method using liquid chromatography tandem mass spectrometry (LC-MS/MS) for eight different environmental matrices. The results of the statistical analyses inform the quality control (QC) acceptance criteria for the method to be established by the Environmental Protection Agency (EPA). This report documents the statistical analyses used to characterize the overall performance of the method in the MLV study, highlights instances where statistical formulas used in the analysis differ from the EPA guidance for new method approval, and provides high-level observations about the environmental samples, PFAS, and labs across the aqueous datasets.

**15. SUBJECT TERMS**

Department of Defense (DoD); multi-laboratory validation (MLV); per- and polyfluoroalkyl substances (PFAS); liquid chromatography tandem mass spectrometry (LC-MS/MS); quality control (QC); Environmental Protection Agency (EPA)

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Spangler, Kimberly Y. |
| Unclassified | Unclassified | Unclassified | Same as Report | 56 | 19b. TELEPHONE NUMBER *(Include area code)* 973-724-3156 |

Reset

**Standard Form 298** (Rev. 8/98)
Prescribed by ANSI Std. Z39.18