United States Environmental Protection Agency

Office of Water

www.epa.gov January 2024

Multi-laboratory Validation Study Report for Method 1621: Determination of Adsorbable Organic Fluorine (AOF) in Aqueous Matrices by Combustion Ion Chromatography (CIC)

U.S. Environmental Protection Agency
Office of Water (4303T)
Office of Science and Technology
Engineering and Analysis Division
1200 Pennsylvania Avenue, NW
Washington, DC 20460

EPA 821-R-24-003

Acknowledgements

This report was prepared by S. Bekah Burket and Adrian Hanley of the Engineering and Analysis Division within The EPA's Office of Water, with assistance from Mirna Alpizar and Harry McCarty of General Dynamics Information Technology (GDIT).

The EPA acknowledges the support of the several organizations and individuals in the development and validation of the draft method for the detection of adsorbable organic fluorine in aqueous samples, including the members of the EPA's workgroup, the developers of the original procedure, the organizations that provided the bulk samples of wastewater, and the EPA's support contractor staff who oversaw the day-to-day operations during the study and assisted the EPA in the preparation of this report. At a minimum, that includes the following:

S. Bekah Burket	EPA OW/OST/EAD	Mike Martin	Hampton Roads Sanitation District
Adrian Hanley	EPA OW/OST/EAD	Louis Logan	Massachusetts Water Resources Authority
Jenifer Jones	EPA ORD/CESER	Steve Rhode	Massachusetts Water Resources Authority
Jody Shoemaker	EPA ORD/CESER	Mahesh Pujari	Los Angeles City Sanitation
Brian Pike	Pace Analytical Services, LLC	Mirna Alpizar	General Dynamics Information Technology
Nick Nigro	Pace Analytical Services, LLC	Harry McCarty	General Dynamics Information Technology
Joel Kelly	Pace Analytical Services, LLC	Kenneth Miller	General Dynamics Information Technology
Tom Patten	Pace Analytical Services, LLC, and member of ASTM D19	William Lipps	Shimadzu Scientific Instrument, Inc., and member of ASTM D19
Ron MacGillivray	Delaware River Basin Commission	Takuro Kato	Mitsubishi Chemical, and member of ASTM D19
Craig Forbes	Hampton Roads Sanitation District	Jayesh Gandhi	Metrohm USA, Inc., and member of ASTM D19
Blaine Hoffman	Hampton Roads Sanitation District		

The EPA also thanks the staff at the 11 laboratories in Table 2-1 that participated in the validation study.

Disclaimer

Mention of company names, trade names, or commercial products in this report does not constitute endorsement or recommendation for use.

Contact

Questions or comments regarding this report should be addressed to:

CWA Methods Team, Engineering and Analysis Division (4303T)
Office of Science and Technology
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue Washington, DC 20460

https://www.epa.gov/cwa-methods
https://www.epa.gov/cwa-methods/forms/contact-us-about-cwa-analytical-methods

Executive Summary

The goal of this project was to complete a multi-laboratory validation study of a method for determination of adsorbable organic fluorine (AOF) in aqueous samples. AOF is a "method-defined parameter" (MDP) meaning that the measurement result is defined solely by the method used to determine the analyte. In the case of AOF, Method 1621, estimates an aggregate concentration of any organofluorine compounds in the sample that are retained on the granular activated carbon (GAC) sorbent and subsequently measured by combustion ion chromatography (CIC).

Because there are no methods for the detection of AOF in aqueous samples that are approved for use in Clean Water Act compliance monitoring, in 2020 the US Environmental Protection Agency (EPA) convened a workgroup of the EPA, laboratory, and utility staff, supported by contractors. The workgroup selected a draft ASTM International standard with the primary intended use for the analysis of wastewater for compliance monitoring. The method was validated in a single-laboratory method validation (SLV) study. Based on the results of the SLV, EPA Method 1621 was published, a multi-laboratory method validation (MLV) study was performed in 2023. The results of the MLV study are discussed in this report.

The MLV study of Method 1621 met all of the EPA's goals. The study generated precision and recovery data for aqueous matrices. Out of the 429 matrix spike results gathered during this study, 96% of the results had recoveries between 50 and 150 percent. Only 3% of the spiked samples had recoveries below 50% and 1% were above 150%. This level of performance is suitable for an aggregate method.

However, because organofluorines are ubiquitous, the method is prone to contamination from the various transfer lines and valves in the adsorption unit, air contamination, and background contamination from the capping material on the granular activated carbon (GAC) columns or from the carbon itself. Additionally, results may be biased based on the composition of organofluorines within a sample. This method is useful to broadly assess organofluorine contamination in aqueous matrices. The method detection limit studies demonstrated that the method can be sensitive down to 1.5 µg F^-/L when using stringent instrument cleaning protocols and GAC columns containing low fluoride background. These points are emphasized in the method.

There also is a need for testing for organofluorines in biosolids, soils, sediments, and fish tissue, and portions of this method may form the basis for future studies, as technology advances and sample preparation techniques evolve.

Table of Contents

List of Tables

List of Figures

1. Introduction

The goal of this project was to complete a multi-laboratory validation (MLV) study of EPA Method 1621, which is a procedure for the measurement of adsorbable organic fluorine (AOF). The method is designed to estimate the concentration of organic fluorine in aqueous environmental samples arising from per- and polyfluoroalkyl substances (PFAS), as well as from non-PFAS compounds such as pesticides and pharmaceuticals.

Background

Fluorine is the thirteenth most abundant element in the earth's crust; however, it is mostly found in inorganic forms such as fluoride salts. Only a handful of the over 3,000 naturally occurring organohalogens contain one or more fluorine atoms covalently bonded to a carbon atom. Most organofluorine chemicals found in the environment are man-made. Of the man-made organofluorines, PFAS, fluorinated pharmaceuticals, and fluorinated pesticides are the most widespread. PFAS are of particular concern due to their persistence in the environment.

PFAS include thousands of man-made chemicals that have been in production since the 1940s and are found in a variety of consumer products such as cookware, food packaging, and water-repellent fabrics. Some of the most common legacy sources of PFAS were from the manufacture of non-stick materials and largely consisted of perfluorinated compounds such as perfluorooctanesulfonic acid (PFOS) and perfluorooctanoic acid (PFOA). Even though voluntary efforts to phase out those compounds began in 2008, they are persistent in the environment and resistant to typical environmental degradation processes. Since the phase out of PFOS and PFOA, a large variety of other polyfluorinated alkyl substances are now in common use as alternatives to PFOS and PFOA.

Many PFAS are soluble in water, and as a result, are highly mobile in the environment. PFAS are extensively distributed across all trophic levels and are found in soil, air, and groundwater at sites across the globe. Most PFAS do not readily breakdown in the environment, and many are known to bioaccumulate in aquatic and terrestrial biota, with some compounds bioaccumulating more than others.

EPA Workgroup

Various organizations and regulatory authorities at state, federal, and international levels are taking action to address the release of PFAS to the environment. However, developing analytical methods for each individual PFAS compound is impractical, if not impossible. Therefore, this project pursued validation of a procedure to determine an aggregate measure of organofluorine substances in aqueous matrices as a proxy for PFAS and other fluorinated pesticides and pharmaceuticals, known as AOF. The EPA assembled a workgroup in 2020 led by the Engineering and Analysis Division, which is part of the Office of Science and Technology within the Office of Water (OW/OST/EAD), to spearhead efforts for the single-laboratory validation (SLV) and multi-laboratory validation (MLV) studies.

Method Selection

For the SLV study, the workgroup identified a draft standard from ASTM International (Reference 1) which uses combustion ion chromatography (CIC) as the best starting point for the development of an EPA method. EAD, in conjunction with members of the ASTM D19 Committee, prepared a hybrid study plan/quality assurance project plan to develop and validate an EPA procedure in a single laboratory that would be based on the original ASTM draft standard.

The SLV study took place during 2022 and tested many different PFAS and non-PFAS compounds. The SLV study produced consistent data and many of the procedures from the draft ASTM standard were

incorporated into the draft of Method 1621. In 2023, a new hybrid study plan/quality assurance project plan was prepared for an MLV study of the draft method that focused mainly on PFHxS as the spike compound for the quality control samples and the real-world matrices. In addition, a secondary smaller study was done using one of the nine study samples spiked with PFBA, PFOS, and a mixed PFAS standard.

Summary of the Results of the Single-laboratory Studies

Two SLV studies were performed: one by Pace Analytical® Services, LLC, in their IDEA Laboratory, and a second SLV study at EPA ORD in Cincinnati, using a CIC instrument from a different vendor. As noted in the report on those two studies (EPA 2022), the SLV studies of the draft method met the following EPA criteria.

1. *The method provides an aggregate response for adsorbable organofluorine from single compounds with chain lengths between C_4 and C_8 as well as non-PFAS fluorinated compounds using combustion ion chromatography.*

The studies generated initial precision and recovery data for aqueous matrices. IPR recoveries between 80 – 120% were achieved in most of the 30 reagent water IPR samples analyzed during this study (25 out of 30 results, or 83%). The single-laboratory validation matrix spike results demonstrated that the method was capable of quantifying AOF from various environmental aqueous sources with recoveries between 50 and 150%. However, the SLV report indicated potential bias for certain organofluorines.

2. *The method is sensitive enough to be used as a screening method.*

Due to the ubiquitous occurrence of PFAS, almost all method blanks contained some organofluorine. Blank results suggested AOF baseline contamination originated from the capping material used in the adsorption columns and that it could not be removed with a nitrate pre-wash. High blank levels were observed in certain cases. Thus, the SLV data demonstrated that the method was sensitive; however, unless strict cleaning protocols are followed, the method can be subject to significant blank contamination. Nonetheless, the method can reliably estimate the concentrations of organofluorines at low part-per-billion levels.

3. *Can be implemented at a typical mid-sized full-service environmental laboratory.*

Not all laboratories own an adsorption unit or a combustion ion chromatography unit; therefore, the ease of implementation is unknown, but there is little doubt that a typical full-service laboratory could implement this procedure. The procedure has many similarities to other EPA organohalogen methods (e.g., Methods 1650 and 9020B). All the standards for the method are commercially available from one or more vendors. Also, two laboratories performed similar single-laboratory validation studies, the Pace IDEA laboratory, as detailed in the SLV report, and EPA's ORD Laboratory in Cincinnati, OH, as summarized in Appendix C of the SLV report. Both laboratories achieved similar method performance and results. This indicates that the method can likely be implemented in a typical mid-sized, full-service environmental laboratory.

Study Objectives

The main objective of the MLV study was to develop data to characterize the performance of a new method for adsorbable organofluorine compounds. As such, EPA's goals were to:

- Obtain data from matrices that are representative of the method's intended use
- Obtain data from laboratories that are representative of those likely to use the approved method, but that were not directly involved in the method development
- Obtain feedback from laboratory users on the specifics of the draft method
- Use study data to characterize performance of the method
- Develop statistically derived quality control (QC) acceptance criteria that will reflect method performance capabilities in real-world situations

The design of the MLV study is described in a formal study plan that is included as Appendix B to this report. The design is based on the specifications in the EPA's *Protocol for Review and Validation of New Methods for Regulated Organic and Inorganic Analytes in Wastewater Under EPA's Alternate Test Procedure Program* (USEPA 2018). Briefly, the design involved:

- Eleven laboratories, with a goal of complete wastewater data from at least six laboratories
- Nine wastewater samples from a variety of sources
- Determination of retention times for fluoride and separation from chloride peak
- Multi-point calibration of the target analyte
- Initial demonstration of capability (IDC) by each laboratory
- Determination of method detection limits (MDLs) by each laboratory
- Analyses of matrix spike and matrix spike duplicate (MS/MSD) samples prepared from each of the nine wastewater samples

Results Contained in this Report

The data tables contained in the body of this report are summaries of the data.

2. Identification and Selection of Laboratories

After completing the single-laboratory studies at the beginning of 2022, the EPA and GDIT identified potential participants in the multi-laboratory study. By May 2022, through a combination of established relationships, review of an EPA database of laboratory capabilities, internet searches, and telephone calls, 16 potential participants were identified, including commercial environmental laboratories, state laboratories, and utility laboratories. The EPA identified potential Regional and emergency response laboratories within the Agency as well.

Between April and June 2022, through emails and telephone calls, the EPA put together a list of potential participants and ultimately targeted 15 laboratories, including those who would be contracted and those who could volunteer. GDIT developed a contractual statement of work (SOW) covering all aspects of the study, sent a formal solicitation to nine commercial/research laboratories, and received bids from six of the labs. Some of the potential participants were not able to participate because of time and/or staff constraints. The selected paid laboratories comprised five commercial laboratories and one research laboratory, each of whom received a purchase order for participation. The EPA arranged for five volunteer participants and used the GDIT SOW as the basis for a memorandum of understanding with each of the volunteer laboratories. The list of the original eleven participants is provided in Table 2-1.

Table 2-1. List of Participating Laboratories

Bureau Veritas, Canada	Pace Analytical Services, Massachusetts
Enthalpy Analytical	SGS AXYS, Canada
Eurofins, Lancaster, PA	Thermo Fisher, IC Applications Lab
EPA ORD, Cincinnati, OH	Trace Elemental Instruments
EPA Robert S. Kerr Environmental Research Laboratory, Ada, OK	University of North Carolina at Charlotte
Mandel Scientific Inc.	

On November 18, 2022, immediately prior to the start of the study, the EPA held a kick-off call with all the participating laboratories to discuss the specifics of the study. A summary of the call, including answers to any questions raised by the participants, was circulated to all the laboratories after the call. The EPA subsequently held conference calls from February 2023 through October 2023. Because not all participants were able to attend every call, the discussion and any critical points were summarized and circulated by email after each call.

Unfortunately, despite the EPA's best efforts, not all eleven laboratories were able to complete the study. Two laboratories dropped out during the initial phase of the study due to instrument issues. Where practical, the data from these laboratories were considered for use in the study. In the end, nine laboratories provided full data sets for all aspects of the study. Having data from nine laboratories met the EPA's study design goals of acquiring data from at least six laboratories for the wastewater matrices.

Other than the list of laboratories in Table 2-1 above, the remainder of this report does not associate specific results with a named laboratory. Rather, each laboratory that completed any portion of the study was randomly assigned an identifying number between 1 and 11.

3. Study Samples Selection

The wastewater samples used in the study were selected to meet the specifications in the EPA's New Method Protocol (USEPA 2018), namely, that at least one of the wastewater matrix types should have one of the characteristics below:

- Total suspended solids (TSS) greater than 40 mg/L
- Total dissolved solids (TDS) greater than 100 mg/L
- Oil and grease (O&G) greater than 20 mg/L
- Conductivity as NaCl, greater than 120 mg/L
- Hardness as $CaCO_3$, greater than 140 mg/L

GDIT obtained large volumes of nine real-world wastewaters from three major wastewater treatment operations: Hampton Roads Sanitation District (HRSD), Los Angeles Sanitation and Environment (LASAN), and Massachusetts Water Resources Authority (MWRA), including samples of wastewater effluents and influents, as well as samples of aqueous matrices from indirect dischargers to those systems. Table 3-1 lists the samples provided.

Table 3-1. Study Sample Matrix Sources

Sample Identification	Industry Type	Sample Identification	Industry Type
Sample #1	POTW-1	Sample #6	Bus Washing Station
Sample #2	Dairy Effluent	Sample #7	Pharmaceutical Effluent
Sample #3	Hospital Effluent	Sample #8	Industrial Effluent
Sample #4	Metal Finisher	Sample #9	POTW-3
Sample #5	POTW-2		

Due to the total sample volume required for this project, each of the samples used for the study was collected in bulk in multiple 10-L LDPE cubitainers. After collection, the bulk samples were shipped in coolers with ice to GDIT's sample repository at Microbac Laboratories in Baltimore, MD, and stored in a walk-in refrigerator at 0 - 6 °C.

Study Samples Characterization

Following collection of the 10-L bulk samples that were shipped to GDIT's repository, smaller aliquots of the wastewater sources were collected and preserved following the requirements in 40 CFR Part 136, Table II. Samples for AOF were collected in 125-mL HDPE, wide-mouth bottles, and were left unpreserved. Aliquots of each of the nine wastewater samples were shipped on ice directly to Eurofins Lancaster Laboratories Environment Testing (ELLE) and tested for the five water quality characteristics listed above, as well as for Total Organic Carbon (TOC), anions, pH, and AOF.

I. Water Quality Parameters

At least two of the sample characteristics criteria were met by each of the study samples, and the EPA deemed the samples suitable for use in the study. A summary of the sample characteristics is provided in Table 3-2. The values in bold font indicate that the sample met the requirements for that parameter.

Table 3-2. Water Quality Characteristics of the Study Samples

Sample #	TSS (mg/L)	TDS (mg/L)	O&G (mg/L)	Conductivity (µmhos/cm)	Hardness (mg/L)
1	9.4	**410**	ND	**930**	140
2	**210**	**300**	**115**	**630**	42
3	**160**	**480**	19.6	**910**	ND
4	30	**670**	ND	**980**	30
5	5.1	**1600**	2.2	**3000**	**200**
6	**790**	**1700**	**27.9**	**3200**	110
7	4.1	**300**	1.5	**400**	13
8	4.5	**930**	ND	**1800**	130
9	13	**940**	1.8	**1800**	**300**

ND = Not detected

II. Method Interferent Compounds Levels

The pH and TOC levels were determined for each of the study samples. At low pH levels, inorganic fluoride in the sample can cause high bias for AOF results, while high levels of organic carbon can lead to negative interferences by inhibiting quantitative adsorption of halogen bound organic substances to the activated carbon. Note that TOC was measured in these samples, instead of dissolved organic carbon, because filtering the samples in the field was considered impractical, especially given that the samples were collected by volunteers.

The nitrate wash employed in EPA Method 1621 has been shown to remove levels of inorganic fluoride up to 8 mg/L; therefore, the concentration of inorganic fluoride was also determined for each of the study samples, as well as the concentration of chloride, which can interfere with the fluoride peak in ion chromatography when found at concentrations above 500 mg/L. With the exceptions of Sample #5 and Sample #6, which had concentrations of chloride above 500 mg/L and required dilution prior to analysis, the other known interferents in the bulk samples were within the method-specified limits. Results for these parameters are summarized in Table 3-3 below.

Table 3-3. Results for TOC, pH, Inorganic Fluoride, and Chloride

Sample #	TOC (mg/L)	pH (SU)	Inorganic F (mg/L)	Chloride (mg/L)
1	10	7.5	0.62	130
2	140	5.9	0.72	46
3	110	7.1	1.1	84
4	13	9.5	2.6	47
5	10	7.5	0.57	**840**
6	80	6.9	1.1	**970**
7	32	7.2	0.55	58
8	26	8.2	5.3	350
9	18	7.0	1.1	290

Chlorine was not detected in any of the study samples; therefore, none of the bulk samples needed to be dechlorinated.

III. AOF Reconnaissance Results

As noted in Section 1, the study design called for analyses of MS/MSD pairs for each study sample. In order to provide information to each participating laboratory about the concentrations at which to spike those MS/MSD pairs of each sample, the EPA sent single aliquots of each study sample to ELLE for determination of the background AOF levels. The "reconnaissance" results from those analyses were

used by the EPA and GDIT to develop spiking concentrations for the MS/MSD aliquots of each of the study samples. Reconnaissance results for each study sample are summarized in Table 3-4 below. ELLE diluted all the study samples by a factor of 2 to prevent any possible damage to their instrument; therefore, all the detection limits for the reconnaissance results are 2.0 µg F^-/L.

Table 3-4. AOF Reconnaissance Results, µg F^-/L

Sample #	Result	Sample #	Result
1	4.4	6	*ND*
2	*ND*	7	*ND*
3	10	8	*ND*
4	6.9	9	2.8
5	2.4		

ND = Not detected

Preparation and Shipping of Study Samples Aliquots

The sample aliquoting and shipping activities took place during the week of June 12, 2023. GDIT personnel retrieved the 10-L LDPE cubitainers from the sample repository and homogenized each study sample by placing the bulk volume in a large open container and thoroughly mixing it for 10 minutes. The samples were aliquoted as 100-mL volumes using HDPE pipettes, into 125-mL HDPE, wide-mouth containers of the same type and lot number as the ones used for the reconnaissance analysis, and the samples were stored at 0 - 6 °C until they were shipped to the participating laboratories (within at most two days of preparation).

Due to their high levels of chloride, Sample #5 and Sample #6 were diluted 2x during the homogenization process using PFAS-free reagent water obtained from one of the participating laboratories. For these two samples, one of the 10-L cubitainers for each sample was homogenized with equal amounts of PFAS-free water, and the study samples aliquoted from the diluted bulk sample.

GDIT shipped the study samples to the participating laboratories in coolers containing enough ice to maintain the samples at a temperature < 6 °C for several days. Once received by the laboratories, the samples were stored at 0 – 6 °C until the time of analysis. In a few situations where a laboratory did not have enough refrigerator space, the laboratory was allowed to store the samples in a freezer at ≤ -20 °C. In those cases, the samples were allowed to fully thaw before any spiking solution was added to the sample for analysis.

4. Calibration and Calibration Verification

In EPA Method 1621, the ion chromatograph (IC) is calibrated using a sodium fluoride standard that is combusted, instead of directly injected, in order to normalize the results for any possible loss of analyte that may be caused by sample combustion. However, the calibration standards do not go through the carbon adsorption process. The calibration curve is based on the concentration of fluoride ion in the standard (i.e., µg F^-/L), rather than the mass or concentration of a specific fluorine-containing compound.

Two brands of IC instruments, Dionex Integrion HPIC and Metrohm 930 IC Compact Flex, were used during this study to encompass the instruments that most laboratories would be using during typical sample analyses. To minimize potential variability in the results that might be caused by differences in IC column types, the laboratories were required to use an IC column model from the same vendor as their instrument. Therefore, laboratories that used the Dionex IC, were required to use a Dionex IonPac AS24, 2x250 mm IC column, while laboratories that used the Metrohm IC were required to use a MetroSep A, Supp 5, 4x150 mm IC column. GDIT provided the appropriate IC columns to those laboratories that were using a different IC column from the ones required for the study.

Initial Calibration

The IC instruments were initially calibrated using a series of six standards designated as CS-1 to CS-6, with a concentration range from 1.0 to 50.0 µg F^-/L. Table 4-1 lists the concentrations of those calibration standards.

Table 4-1. Initial Calibration Standards, µg F^-/L

Analyte	CS-1	CS-2	CS-3	CS-4	CS-5	CS-6
Fluoride	1.0	2.0	5.0	10.0	25.0	50.0

All eleven laboratories completed the calibration portion of the MLV study. Each of their calibrations was evaluated using four different model fits to see which would be the best approach to calibration. The four models were:

- Linear fit with 1/x concentration weighting,
- Linear fit with $1/x^2$ concentration weighting,
- Quadratic fit with 1/x concentration weighting, and
- Quadratic fit with $1/x^2$ concentration weighting.

The calibration models were not forced through zero. Each laboratory was asked to calculate the percent relative standard error (%RSE) to assess the fit of their calibration curve for each model, using the equation below. The software for some of the instruments could not calculate the %RSE based on the equation used by the EPA. In those instances, the laboratories calculated the %RSE using a spreadsheet provided by GDIT with the correct equation.

$$\%RSE = 100 \times \sqrt{\sum_{i=1}^{n} \frac{\left[\frac{x_i' - x_i}{x_i}\right]^2}{n - p}}$$

where:

x_i = Nominal concentration (true value) of each calibration standard
x'_i = Measured concentration of each calibration standard
n = Number of standard levels in the curve
p = Type of curve (2 = linear, 3 = quadratic)

Table 4-2 summarizes the %RSEs of the initial calibrations performed by the participating laboratories. Laboratory 7 added one extra calibration point and extended their initial calibration up to 100 µg F^-/L. The extended calibration for Laboratory 7 gave them slightly lower %RSEs for all four fits when compared to the %RSEs calculated when their curve was cut off at 50 µg F^-/L. Their lower %RSEs were comparable to the %RSEs from the remaining 10 laboratories.

Table 4-2. Initial Calibration % Relative Standard Error

Calibration Fit	%RSE		
	Mean (n=11)	Min	Max
Linear with 1/x weighting	10.3	2.5	22.8
Linear with $1/x^2$ weighting	7.1	2.1	12.8
Quadratic with 1/x weighting	5.3	0.07	15.1
Quadratic with $1/x^2$ weighting	4.1	0.22	11.8

Figure 4-1 shows the %RSE for each of the laboratories for each of the calibration fits. The highest %RSEs for the linear fits were observed for Laboratories 1, 3, 4, 7, and 8. Laboratories 4 and 7 had a %RSE > 10% for the quadratic with 1/x weighting calibration fit, while the rest of the laboratories had a %RSE well below 10% for the same fit. As the fit was moved to quadratic with $1/x^2$ weighting, only Laboratory 4's %RSE remained above 10%. With the exception of Laboratory 4, most calibrations showed better model fit (e.g., lower %RSE) when using a quadratic with $1/x^2$ weighting fit.

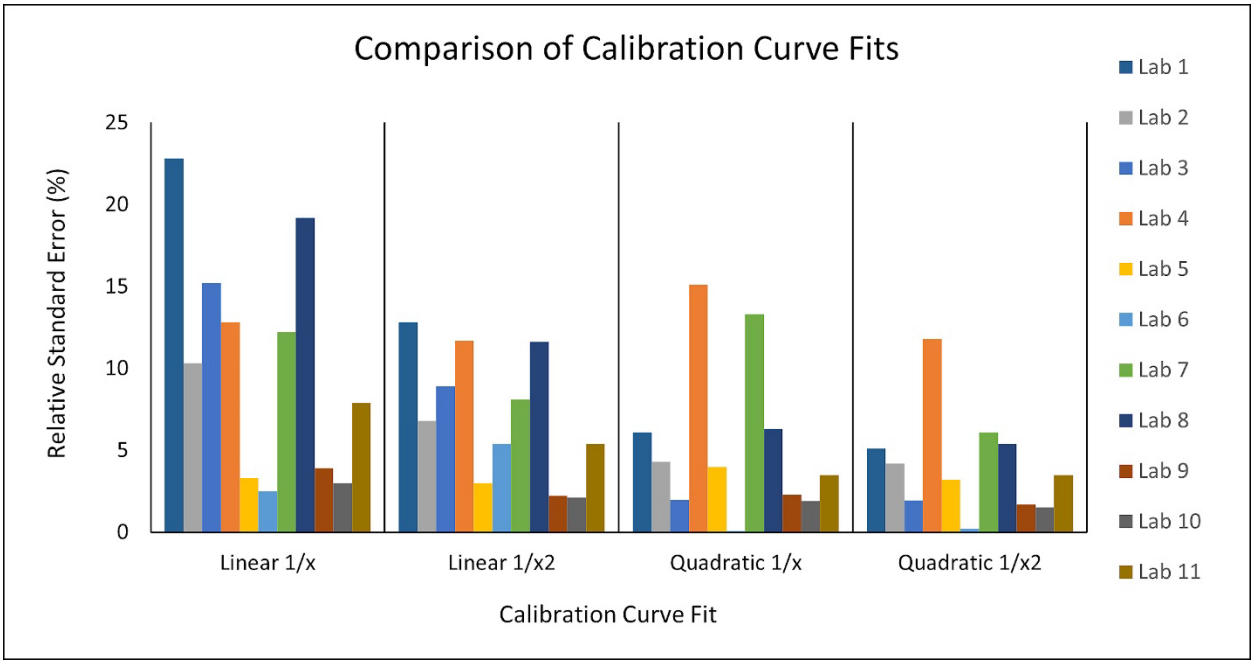

Figure 4-1. Comparison of % Relative Standard Error Across Calibration Models

Although %RSEs for the linear fit with 1/x concentration weighting were < 20% for 10 of the 11 laboratories, visual examination showed that the highest point of the calibration curve did not fall along the curve, whereas that point did fall along the curve for the quadratic fit with $1/x^2$ concentration weighting, as shown in the example in Figure 4-2 below. Based on the %RSE results and visual inspection of the curves, GDIT and the EPA chose a quadratic with $1/x^2$ weighting fit for the analysis of samples for this study and relayed that requirement to each of the laboratories in the study.

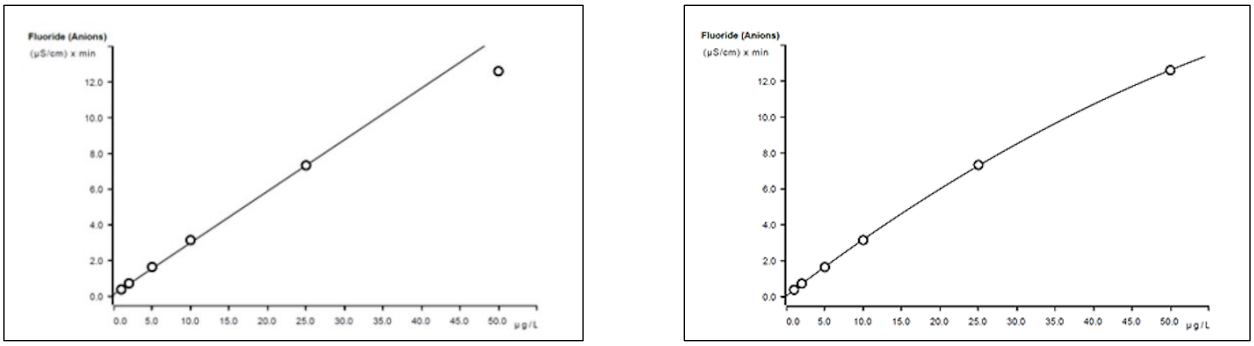

Figure 4-2. Example of linear $1/x^2$ fit vs. quadratic $1/x^2$ fit

In addition to requiring the quadratic calibration fit, prior to analyzing study samples, EPA required the laboratories to extend their calibration range to 100 µg F^-/L. The discussion for the extended calibration results is included with Section 8 of this report.

Calibration Verification

A calibration verification (CV) standard was analyzed at the beginning and at the end of each analytical batch. As with the ICAL, the CV standards were combusted, but not adsorbed onto the carbon columns. Verification was performed at two concentrations, 2 µg F^-/L and 25 µg F^-/L (200 ng and 2,500 ng). A total of 118 calibration verification standards were analyzed during the study. The recoveries ranged from 80.2% to 135.6%, with an average recovery of 101.3%. Table 4-3 summarizes the recoveries observed for both concentrations. The highest recovery for the low-level CV was observed for Laboratory #8, while the highest recovery for the mid-level CV was observed for Laboratory #1.

Table 4-3. Observed Recoveries for Calibration Verifications per Concentration Level

Calibration Verification	Conc. (µg F^-/L)	# Results	Recovery (%)		
			Mean	Min	Max
Low-level	2	59	100.98	85.5	135.6
Mid-level	25	59	101.58	80.2	122.1

GDIT performed a *t*-test that confirmed that the mean recoveries for the two spiking levels were not significantly different, while an *F*-test indicated that the variance in the mid-level recoveries was half that seen for the low-level recoveries. Table 4-4 summarizes the observed failures for the calibration verification using two recovery ranges for each concentration level. For the recovery range of 90 – 110%, four recoveries for the low-level and two recoveries for the mid-level failed the lower limit (LL), while seven recoveries for the low-level and three recoveries for the mid-level failed the upper limit (UL). The observed failures are equivalent to an 18.6% failure rate for the low-level CV and 8.5% failure rate for the mid-level CV. When the recovery range is increased to 80 – 120%, only one recovery for each concentration level falls outside the upper limits, which is equivalent to a 1.7% failure rate. The two failing CV results are from Laboratories #1 and #8. Based on these failure rates, the interim recovery of 80 – 120% in the draft method will be retained and will apply to both CV concentrations.

Table 4-4. Observed Calibration Verification Recovery Failures for Two Potential Acceptance Criteria per Concentration Level

Calibration Verification	Conc. (µg F^-/L)	Criteria 1		Criteria 2	
		LL 90%	UL 110%	LL 80%	UL 120%
Low-level	2	4	7	0	1
Mid-level	25	2	3	0	1

5. Method Detection Limit

As part of the MLV study, the EPA required that the laboratories determine their method detection limit (MDL) for AOF. For the purpose of the study, the MDL was determined using the revised MDL procedure *Definition and Procedure for the Determination of the Method Detection Limit*, Revision 2, December 2016, EPA 821-R-16-006, and also included at 40 CFR Part 136, Appendix B.

The revised procedure defines the MDL as:

> *"… the minimum measured concentration of a substance that can be reported with 99% confidence that the measured concentration is distinguishable from method blank results."*

The procedure consists of two parts: determination of the MDL based on method blanks (MDL_b), and determination of the MDL based on spiked samples (MDL_s). Both MDL_b and MDL_s are determined in a reference matrix, using at least seven replicates prepared and analyzed on three non-consecutive days.

The MDL_b is calculated as:

$$MDL_b = \overline{X} + t_{(n-1,1-\propto=0.99)}S_b$$

where:

$\bar{X}$ = Mean of the method blank results (use zero in place of the mean if the mean is negative)

$t_{(n-1,\ 1-\alpha\ =\ 0.99)}$ = Student's *t*-value appropriate for the single-tailed 99th percentile *t* statistic and a standard deviation estimate with n-1 degrees of freedom.

S_b = Sample standard deviation of the replicate method blank sample analyses.

Note: The equation above is used when all the method blanks for an individual analyte give numerical results. If some (but not all) of the method blank results give numerical results, then the MDL_b is set to be equal to the highest method blank result.

The MDL_s is calculated as:

$$MDL_s = t_{(n-1,\ \ 1-\propto=0.99)}S_s$$

where:

$t_{(n-1,\ 1-a\ =\ 0.99)}$ = Student's *t*-value appropriate for a single-tailed 99th percentile *t* statistic and a standard deviation estimate with n-1 degrees of freedom.

S_s = Sample standard deviation of the replicate spiked sample analyses.

For MDL determinations in this study, the reference matrix used was PFAS-free reagent water. After both an MDL_b and MDL_s were calculated, each laboratory set their initial MDL to the greater of the MDL_b and MDL_s values.

AOF MDL Determination

Method 1621 tests for the aggregate concentration of organic fluorine compounds which has been adsorbed onto the granular activated carbon. Because there are numerous PFAS and non-PFAS organofluorine compounds found in the environment, it would be impossible to determine an MDL that would include every compound. During the SLV study, the EPA ORD investigated the recoveries of various individual PFAS analytes and found that PFHxS showed consistent recoveries ranging from 99%

to 106%. Therefore, GDIT and the EPA specified PFHxS as the compound to be used by all of the MLV laboratories for their MDL studies.

One of the first steps in the EPA MDL procedure is for the laboratory to make an estimate of the MDL. The laboratories were directed to consider the limitations of the instrument, more specifically the limitation presented by the observed GAC background, when estimating their initial MDL. GDIT and the EPA directed the laboratories to set the estimated initial MDL as three times the standard deviation of the initial background observed for the GAC columns. The spike levels were then established to be approximately between 2 and 10 times the estimated MDL, at 5.0 µg F^-/L.

Due to a lack of reproducibility of their calibration verifications and after several attempts at instrument troubleshooting, Laboratory 4 had to drop out from the study after the initial calibration. The remaining 10 laboratories proceeded with the MDL study by preparing and analyzing at least seven replicate method blanks and seven replicate spiked samples. Laboratory 6 analyzed nine method blanks and nine spiked sample replicates while Laboratory 8 analyzed eight method blanks and eight spiked sample replicates. This approach yielded 73 results for blanks and 73 results for spiked samples.

Laboratory 1 was the only laboratory that did not have a numerical value for each of the seven blank replicates, and therefore the highest blank value was used by that laboratory for their MDL_b. The initial MDL estimates for six of the laboratories were established from the MDL_b, due to the high background levels present in the GAC columns. GDIT determined the pooled MDL by using the average blank concentration of all the labs plus the left-tailed inverse of the Student's *t* distribution times the pooled highest standard deviation of each laboratory from either the blanks or the spiked samples. The results are summarized in Table 5-1 below.

Table 5-1. Pooled MDL Results (µg F^-/L)

# Labs	Pooled MDL_s	Max. MDL_s	# MDL_b	Max. MDL_b
10	1.5	2.9	6	3.2

Because many organofluorine compounds are prone to surface adsorption, analytical results between different manufacturer's adsorption apparatus may not be comparable due to differences in the numbers and areas of surfaces (e.g., tubing, syringes, etc.) with which the sample comes in contact during sample loading onto the GAC. For that reason, GDIT and the EPA compared the MDL study results for the two adsorption instruments used in the study, Analytik-Jena APU sim and the Mitsubishi TXA-04 units, to determine if there was a consistent difference that could be attributed to the instrument configuration. A Student's *t*-test and an Analysis of Variance (ANOVA) were used to compare the MDL_s and MDL_b for each system and both tests showed no statistical difference between the adsorption units.

6. Precision and Recovery

As part of the method validation effort, the EPA required that each laboratory perform an initial precision and recovery (IPR) study. Each IPR study consisted of a set of four reagent water aliquots, spiked at approximately 25 µg F^-/L with PFHxS. The laboratories calculated and reported the mean concentration, mean recovery, the standard deviations of the recoveries, and the relative standard deviations (RSDs) of the recoveries for each set of four IPR aliquots.

Each laboratory also included one ongoing precision and recovery (OPR) sample with each batch of study samples prepared and analyzed. The OPR aliquots were spiked at the same concentration as the IPR aliquots. Each laboratory calculated and reported the recovery of the spiked analytes in each OPR aliquot.

One of the EPA's objectives in conducting the multi-laboratory validation study was to generate data from which to derive multi-laboratory QC acceptance criteria for the various performance tests in the method, or to evaluate the ability of the method to meet commonly applied acceptance criteria for some performance tests. In this study, GDIT calculated QC acceptance criteria for IPR and OPR tests based on the procedures and equations in Appendix G of the *Protocol for Review and Validation of New Methods for Regulated Organic and Inorganic Analytes in Wastewater Under EPA's Alternate Test Procedure Program* (USEPA 2018), with modifications to account for the actual number of laboratories, samples, and replicates in the study. In order to yield a more complete dataset, the IPR and OPR data were combined when calculating the criteria, using different formulae to generate IPR- and OPR-specific criteria that account for how they would be evaluated in practice (i.e., mean recovery and RSD of four IPR replicates, and OPR recovery evaluated on an individual basis). Briefly:

- The QC acceptance criterion for recovery in the IPR test was calculated by constructing a prediction interval around the mean percent recovery, using the Student's *t* value, with the degrees of freedom determined using the Satterthwaite estimation procedure (Satterthwaite, 1946), using the calculated between- and within-laboratory variance components, weighted based on future IPR usage (assuming means of four replicates per laboratory).
- The maximum acceptable RSD for the four IPR aliquots was calculated as an upper confidence limit around the observed RSD of the results from all the laboratories. The RSD_{IPR} (computed as s_w divided by $\overline{X}$) is multiplied by the square root of a 95th percentile *F* value with 3 degrees of freedom in the numerator and n_T-*m* degrees of freedom in the denominator, where *m* = the number of laboratories, and n_T is the number of data points across all laboratories.
- The QC acceptance criterion for recovery in the OPR test was calculated by constructing a prediction interval around the mean percent recovery, using the Student's *t* value, with the degrees of freedom determined using the Satterthwaite estimation procedure, using the calculated between- and within-laboratory variance components, weighted based on future OPR usage (assuming a single replicate per laboratory).

The calculated IPR and OPR QC acceptance criteria for AOF are presented in Table 6-1. The statistically determined recovery ranges for the IPR and OPR reflect the variability within each laboratory, as well as the variability across the laboratories that completed that portion of the study. Ten laboratories finished the IPR part of the study; however, only nine laboratories finished the study sample analysis portion and provided OPR results. The recovery ranges combined the results for both the IPR and the OPR results.

Table 6-1. Calculated IPR and OPR QC Acceptance Criteria, (%)

		IPR		
# Labs	# Results	Range for Mean	Max RSD	OPR Range
10	97	79.7 – 114.3	14.7	74.7 – 119.4

The IPR is assessed based on the mean of the four IPR replicates; therefore, results for individual IPR aliquots may fall outside the established range. All mean recoveries fell within the calculated IPR mean range with the exception of Laboratory 7, which had a mean recovery of 117.1%. The calculated maximum RSD was well below the method interim requirement of 20%. Three OPR recoveries fell outside the calculated OPR range, two recoveries for Laboratory 1 and one recovery for Laboratory 10.

Table 6-2 summarizes the observed failure rates for the OPR using four possible recovery ranges based on rounding up the calculated range above to multiples of 5% or 10%.

Table 6-2. Observed OPR Recovery Failures for Four Potential Acceptance Criteria

	Observed Failures							
	Criteria 1		Criteria 2		Criteria 3		Criteria 4	
Total # Results	LL 85%	UL 115%	LL 80%	UL 120%	LL 75%	UL 125%	LL 70%	UL 130%
57	5	4	2	3	0	3	0	2

The criteria in Table 6-2 are arranged with the tightest criteria on the left, and increasingly wider criteria to the right. For an acceptance criterion of 85 – 115%, nine recoveries from six different laboratories, or 15.8% of the results, fell outside the limits. For an acceptance criterion of 80 – 120%, five recoveries from three laboratories, or 8.8% of the results, fell outside the limits.

For an acceptance criterion of 75 – 125%, only the three recoveries from Laboratories 1 and 10 fell outside the range, reducing the failure rate to 5.3%. For an acceptance criterion of 70 – 130%, two recoveries, one from Laboratory 1 and one from Laboratory 10, or 3.5% of the results, fell outside the range.

Based on the results, the EPA established an acceptance limit of 80 – 120% for the mean IPR recovery, with an RSD < 20%, and retained the interim limit of 70 – 130% for the OPR.

7. Method Blanks

Each laboratory analyzed two method blanks with each analytical batch of study samples (one at the beginning of the analytical sequence and one at the end). The result from one closing method blank from Laboratory 11 had to be dropped because the bottom GAC column was lost during instrumental analysis, yielding a total of 114 blank results from nine laboratories. The descriptive statistics for those 114 blanks are summarized in Table 7-1 below and do not include any method blanks associated with the IPR and MDL studies.

The draft of EPA Method 1621 that was used for the method validation study stated that the method blanks should not exceed 5 µg F-/L, and preferably, should be less than 3 µg F-/L. For that reason, GDIT evaluated the method blanks using both criteria. The results showed that 107 of 114 method blanks were < 3 µg F^-/L, five results were between 3 and 5 µg F^-/L, and two method blanks were > 5 µg F^-/L limit. Two of the blanks that were between 3 and 5 µg F^-/L were from Laboratory 2, two from Laboratory 8, and one was from Laboratory 10. The two blanks with results above 5 µg F^-/L were from Laboratory 2. The contamination in the blanks from Laboratory 2 may be due to the fact that the laboratory did not store the GAC columns segregated from possible fluoride sources at the start of the project. A new lot of GAC columns was provided to the laboratory for the remainder of the study. The two high method blanks affected the results for all the unspiked samples for Laboratory 2, with the exception of Sample #3, which was prepared and analyzed in a different batch. Because the affected sample results were all < 5x the associated blank values, there are no means by which to ascertain whether the presence of the analyte may be attributed to contamination only. However, because the unspiked sample results were comparable to the results of the unspiked samples for the other laboratories, none of the results were excluded from the study due to contamination. Higher blank values were consistently seen from Laboratories 2, 8, and 9 for different GAC column lot numbers, compared to the other six laboratories.

Table 7-1. Method Blank Summary (µg F^-/L)

# Blanks Analyzed	Minimum	Maximum	Mean	Std Dev	Median	# Below 3 µg F^-/L	# Between 3 and 5 µg F^-/L	# Above 5 µg F^-/L
114	0.0	10.45	1.37	1.34	0.97	107	5	2

GDIT calculated a QC acceptance criterion for the method blanks of < 4.0 µg F^-/L using the mean plus two standard deviations of the blank measurements. Of the 114 method blank results in the study, 97% fell below the calculated limit.

8. Matrix Spike Results

Generating the most useful matrix spike data requires knowledge of the background levels of analytes in the unspiked samples. Spiking levels should not be unrealistically high, nor should they be so low that the spiked amount is difficult to discern from the background concentration in the original sample. Rather than waiting for each participating laboratory to analyze all the unspiked samples, and then develop a customized spiking scheme based on those results, the EPA and GDIT used the reconnaissance analysis results discussed in Section 3 to develop sample-specific spiking levels for all the samples in the study. The EPA determined that the low spike level should be between 2x and 10x the background AOF level of each of the study samples, and the high spike level should be at least double the concentration of the low spike level. This spiking scheme design gave the possibility of collecting up to 154 data points at the low end of the calibration range, up to 264 data points around the mid-point of the calibration, and up to 110 data points at the high end of the calibration range.

The results discussed in this report include 124 data points at the low end of the calibration range, 215 data points around the mid-point, and 90 data points at the high end of the calibration, for a total of 429 spike recoveries.

GDIT directed the laboratories to analyze the full sample volume that was provided and were instructed not to dilute the samples prior to analysis. The laboratories calculated all results based on a nominal 100-mL sample volume and not the individual volumes in each sample bottle.

Three study samples, #2, #3, and #6, had TSS levels above 100 mg/L and required the use of quartz wool plugs as pre-filters for both the unspiked and spiked analyses. The quartz wool plugs were combusted separately from the top and bottom GAC columns. Results for the quartz wool were added to results from the top and bottom GAC columns and are reported in the summary tables below. The laboratories also were required to combust quartz wool blanks with each analytical batch that contained the samples requiring quartz wool pre-filters. The quartz wool blank consisted of a small plug of quartz wool placed inside an empty GAC glass column (to make the amount of glass wool for the blank as close as possible to the amount used for the samples), rinsed with 25 mL of *0.01M* sodium nitrate followed by 20 mL of reagent water. The quartz wool blank columns were not connected in tandem to the blank GAC columns to prevent the transfer of any possible contamination from the quartz wool to the blank GAC columns and cause a possible bias when subtracting the method blank from samples on the same preparation batch that did not require the use of quartz wool. The quartz wool blank result was added to the result of the first method blank (top and bottom GAC columns) from the preparation batch and the summed result was subtracted from the total result for those three samples to compensate for any background fluorine that may have been present in the quartz wool.

Extended Calibration Results

Some of the samples had background AOF concentrations high enough that the high spike level would exceed the calibration range set at the beginning of the study; therefore, the laboratories were required to perform an extended calibration, as shown in Table 8-1.

Table 8-1. Extended Calibration Standards (µg F^-/L)

Analyte	CS-1	CS-2	CS-3	CS-4	CS-5	CS-6	CS-7*	CS-8*
Fluorine	1.0	2.0	5.0	10.0	25.0	50.0	75.0	100.0

*New standards added to the curve

The individual %RSEs were all below 20% with the exception of Laboratory 8. This laboratory's extended curve had a %RSE of 20.7%. Because their lowest point was consistently biased high, the laboratory was allowed to drop the lowest point of their curve (CS-1) after several failed attempts at

achieving a %RSE < 20%. This change allowed their %RSE to decrease to 6.1% making the maximum observed %RSE of the quadratic $1/x^2$ fit for the extended calibration in any of the laboratories below 10%. A summary of the %RSE results is presented in Table 8-2.

Table 8-2. Initial Calibration % Relative Standard Error Summary

Calibration Fit	Mean %RSE	Min %RSE	Max %RSE
Quadratic $1/x^2$	4.5	1.1	9.5

Unspiked Sample Results

GDIT directed the laboratories to analyze the study samples unspiked, and to use those results to calculate their spiked sample recoveries. The results for the unspiked sample analyses for the nine laboratories, plus the original reconnaissance results, are summarized in Table 8-3.

Table 8-3. Unspiked Study Sample Results, µg F^-/L

Sample #	Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	Lab 10	Lab 11	Recon
1	11.6	13.8	5.8	3.6	9.8	2.8	2.1	4.1	4.1	4.4
2	10.1	*ND*	4.1	*ND*	*ND*	*ND*	*ND*	*ND*	*ND*	*ND*
3	16.6	14.2	13.1	9.5	12.4	8.7	8.5	8.7	11.9	10.0
4	15.1	12.8	19.6	7.4	8.4	9.9	6.1	7.3	3.8	6.9
5	2.4	*ND*	*ND*	*ND*	*ND*	*ND*	*ND*	3.4	*ND*	2.4
6	9.8	16.1	11.8	*4.2*	5.0	*5.5*	*ND*	4.9	*ND*	*ND*
7	8.6	*ND*	*ND*	*ND*	*ND*	*ND*	*ND*	*ND*	*ND*	*ND*
8	7.2	*ND*	2.4	*ND*	*ND*	*ND*	*ND*	*ND*	*ND*	*ND*
9	*4.7*	*ND*	3.2	*2.3*	2.8	*ND*	*ND*	6.9	*ND*	2.8

ND = Not detected at the laboratory's MDL

Results for the unspiked samples from most of the laboratories were similar to those found during the reconnaissance analysis. In general, Laboratory 1 had the highest results for the samples. Laboratories 1, 2, and 3 had the highest results for Samples #4 and #6.

Matrix Spike Results

GDIT directed the laboratories to analyze matrix spike aliquots for each of the nine real-world wastewater matrices. The matrix spike analyses were done in two parts. For Part 1, the laboratories were asked to spike separate aliquots of each of the nine study samples at two levels using linear PFHxS, and each spike level was prepared and analyzed in duplicate, i.e., a matrix spike (MS) and matrix spike duplicate (MSD). For Part 2, the EPA selected Sample #7 and GDIT directed the laboratories to prepare three sets of MS/MSD aliquots spiked at two levels, using PFBA, PFOS, and a mixed PFAS standard. The mixed PFAS solution contained 30 PFAS compounds with carbon chains ranging from C4 to C14 and had eleven perfluoroalkyl carboxylic acids, seven perfluoroalkyl sulfonates (linear and branched), three fluorotelomer sulfonates, two ether sulfonic acids, two perfluorooctane sulfonamidoacetic acids (linear and branched), two per- and polyfluoroether carboxylic acids, and three perfluoroalkane sulfonamides. The results from both parts are discussed in detail below.

Part 1 Study Samples Spiked with PFHxS

GDIT calculated the mean recovery of each spiked sample across nine laboratories, along with the minimum and maximum observed values. The values are presented in Table 8-4.

Table 8-4. Percent Recoveries for All Sample Spiked with PFHxS

Sample #	# of Results	Nominal Spike Conc (µg F^-/L)	% Recovery			RSD (%)
			Mean	Min	Max	
1	18	30	92.4	74.9	108.0	12.4
	18	60	96.9	89.1	110.8	5.8
2	18	10	102.8	45.2	181.5	35.8
	18	30	95.7	70.0	113.7	11.2
3	18	30	85.7	72.8	99.8	9.5
	18	60	83.8	56.3	101.2	14.7
4	18	30	95.7	55.4	139.5	21.5
	18	60	93.0	78.0	103.5	7.1
5	18	30	98.1	70.6	113.8	10.3
	18	60	99.5	92.3	109.1	5.6
6	16	10	101.6	46.1	175.4	32.6
	18	30	94.4	63.2	112.3	15.3
7	18	10	101.3	71.9	146.5	15.8
	18	30	97.7	85.6	107.5	7.5
8	18	10	99.1	82.4	123.6	9.3
	17	30	99.7	91.2	114.2	5.7
9	18	30	97.0	79.9	122.5	10.7
	18	60	98.4	85.4	117.2	8.2

For Samples #6 and #8, GDIT dropped three results, either due to sample loss during instrumental analysis or known contamination. All the mean recoveries in Table 8-4 fall within the range of 60 to 120%. Of the 321 individual recoveries, seven were below 60%, and 13 were over 120%, which is equivalent to 6.2% failure rate.

The within-sample RSDs were < 20% for most of the samples, with the exception of Samples #2, #4, and #6, at the lowest level spike. The variability was driven mainly by results from Laboratories 1 and 8. The overall RSD for the entire set of spikes was 16.4%.

GDIT performed an ANOVA test which showed that there was a statistical difference between the spike-level groups. To determine which of the groups was different, GDIT performed a Tukey HSD test at the 5% significance level. The Tukey HSD test confirmed that the recoveries at the lower end of the calibration curve (10 µg F^-/L) were statistically different from the recoveries at the higher end of the curve.

GDIT used the results of the MS/MSD analyses performed in this part of the study to calculate the MS/MSD QC acceptance criteria presented in Table 8-5. One point from the MS/MSD results was dropped from the calculation of the acceptance criteria because the corresponding aliquot from that pair was lost due to instrument failure; therefore, there were 320 data points used in this calculation. Of the 320 total points used in the statistics, nine results failed the lower recovery limit, and seven results failed the higher recovery limit of the calculated range, which is equivalent to 5% of the data.

Table 8-5. Calculated MS/MSD QC Acceptance Criteria, (%)

# labs	Total # of Points	Recovery Range	Maximum RPD
9	320	65.0 – 127.5	34.5

Table 8-6 summarizes the observed failures for the matrix spikes using two acceptance criteria. For the recovery range of 60 – 120%, seven recoveries failed the lower limit (LL), and 13 recoveries failed the upper limit (UL). When the recovery range is increased to 50 – 150%, two of the recoveries fall below 50% and four recoveries, from Laboratories 7 and 11, are above 150%, which corresponds to 1.9% of the data.

Table 8-6. Observed Matrix Spike Recovery Failures for Two Potential Acceptance Criteria

# Results	Criteria 1		Criteria 2	
	LL 60%	UL 120%	LL 50%	UL 150%
320	7	13	2	4

Based on the failure rates for the two potential recovery ranges, the EPA established a recovery range for MS/MSD of 50 – 150% and a Relative Percent Difference (RPD) limit of ≤30% for EPA Method 1621.

Part 2 Sample #7 Spiked with Various PFAS Compounds

Six sets of MS/MSD aliquots were prepared from Sample #7 using three different spiking solutions and two spiking levels that matched those used with PFHxS in Part 1 of the study for this sample. Results for nine laboratories are presented in Table 8-7 below.

Table 8-7. Recoveries for Different Standards Spiked into Sample #7

Spiking Standard	# of Results	Nominal Spike Conc. (µg F^-/L)	% Recovery			RSD (%)
			Mean	Min	Max	
PFBA	18	10	64.1	41.0	86.9	23.9
	18	30	63.6	22.8	107.8	27.5
PFOS	18	10	93.6	41.1	145.0	23.6
	18	30	84.5	33.3	102.3	23.8
Mixed PFAS	18	10	83.1	36.9	107.7	22.6
	18	30	83.1	53.7	95.2	11.7

The within-sample RSDs were > 20% for PFBA and PFOS and for the low-level mixed PFAS spike, indicating higher variability among the laboratories. The variability was mainly driven by recoveries from Laboratories 1, 2, 7 and 9.

The variability within each laboratory's spiked duplicates is presented in Table 8-8. Some laboratories had better control of their replicates, while others had wildly different results between their duplicates. Laboratories 1, 7 and 9 had the highest variability in their MS/MSD results, regardless of the spike level or standard used. This may be due to a lack of familiarity of these laboratories with the method and/or instrumentation. Additionally, percent breakthrough was above 50% for PFBA for Laboratory 1. Laboratory 2 had variability in the low-level aliquots spiked with PFOS, with 76% and 121% recoveries for the MS/MSD pair, resulting in an RPD of 45.9% for that pair.

Table 8-8. Relative Percent Difference (RPD) for Duplicate Spikes per Laboratory

Standard	Nominal Conc. (µg F^-/L)	RPD (%)								
		Lab 1	Lab 2	Lab 3	Lab 5	Lab 7	Lab 8	Lab 9	Lab 10	Lab 11
PFBA	10	53.9	7.4	8.4	10.5	1.5	10.9	5.4	2.4	11.8
	30	6.9	13.4	3.8	3.7	48.4	1.3	79.4	3.7	0.7
PFOS	10	48.7	45.9	4.9	2.1	12.2	2.3	42.4	9.2	15.7
	30	76.3	2.8	3.0	3.4	95.7	20.7	4.9	1.8	18.9
Mixed PFAS	10	35.8	4.4	14.5	1.3	21.8	4.7	5.6	0.3	12.4
	30	5.3	3.3	1.8	0.3	2.8	1.9	34.7	0.0	9.2

Table 8-9 summarizes the observed failures for the Phase 2 matrix spikes using two different acceptance criteria for each individual spiked standard. For the recovery range of 60 – 120%, 21 recoveries failed the lower limit (LL), and two failed the upper limit (UL). When the recovery range is widened to 50 – 150%, 11 recoveries fall below 50% and none are above 150%.

Table 8-9. Observed Recovery Failures for Two Potential Acceptance Criteria per Spiked Standard

Standard	# Results	Criteria 1		Criteria 2	
		LL 60%	UL 120%	LL 50%	UL 150%
PFBA	36	15	0	7	0
PFOS	36	3	2	3	0
Mixed PFAS	36	3	0	1	0

GDIT performed an ANOVA test which showed that there were differences between the recoveries for the three different spiking standards. Further, GDIT performed a Tukey HSD test which confirmed the recoveries for PFBA were different from the recoveries for the other two standards. The low recoveries observed for PFBA may be due to the fact that short-chain PFAS (four carbons or less) tend to be poorly retained on GAC. Low recoveries for PFBA also were observed during the single-laboratory method validation study. The percent breakthrough for this compound was consistently above 50%, confirming its low adsorption on GAC. There were three recoveries < 50% for PFOS, for Laboratories 1, 7, and 9. For Laboratory 7, the low recovery may be due to the lab accidentally spiking that sample at the low-spike level, based on the other results observed for PFOS from this lab.

9. Calibration with an Organofluorine Standard

Different fluorine-containing compounds have different combustion efficiencies. Thus, there was a concern that using inorganic standards such as sodium fluoride for instrument calibration may under- or overestimate the concentrations of organofluorine compounds because those compounds have a wide range of combustion efficiencies. To determine if the type of standard, inorganic vs. organic, used in the calibration may cause such an estimation bias for AOF, Laboratory 3 was directed to perform a special study to compare if AOF results may be biased based on the type of fluoride standard used for instrument calibration.

Note that the use of inorganic fluoride allows for wider calibration ranges to be prepared using a combustion volume of 0.2 mL (lower combustion volumes may be used if the working calibration standards have higher concentrations). In contrast, because the organic fluoride standards have a concentration of 50 µg/mL (approximately 28 µg F^-/mL), the analyst must combust 0.35 mL of the standard for the 100 µg F^-/L calibration point. Using such higher volumes of liquid in a combustion boat designed to combust solids can lead to splattering within the instrument, which can cause internal contamination. Also, being a simple fluoride salt, sodium fluoride is less costly and thus a more economical option for instrument calibration than a PFAS standard.

Comparison of Calibrations Using Organofluorine Standards

In the special study, Laboratory 3 calibrated their IC instrument with PFHxS, using a series of six calibration standards designated as CS-1 to CS-6. For ease of comparison with the calibration performed using sodium fluoride, the PFHxS standards were prepared with as close to the same fluoride concentrations as were used in that initial calibration. The PFHxS standards were combusted directly without carbon adsorption, as was done with the sodium fluoride standard. Because the inorganic fluoride calibration showed a quadratic model was a better fit, the organofluorine calibration used a quadratic fit, with $1/x^2$ weighing, not forced through zero. The %RSE was calculated to assess the model fit, as was done with the inorganic initial calibration (Section 4.0). Table 9-1 compares the recoveries and %RSE of both calibrations from Laboratory 3.

Table 9-1. Comparison of Inorganic vs. Organic Standard Calibrations

Calibration Std	Concentration (µg F^-/L)	Recovery (%)	
		NaF	PFHxS
CS-1	1.0	99.0	97.8
CS-2	2.0	102.7	105.0
CS-3	5.0	98.9	99.9
CS-4	10.0	98.8	98.5
CS-5	25.0	100.7	96.7
CS-6	50.0	99.9	102.6
	%RSE	2.0	4.1

The model fit of both calibrations was well within the %RSE requirement of ≤ 20% and no systematic differences were observed between the two calibrations.

Combustion Efficiencies of Organofluorine Standards

During the single-laboratory validation study, the combustion efficiencies of several organofluorine standards were tested against an inorganic fluoride calibration. For the multi-laboratory validation, Laboratory 3 was asked to perform direct combustion, without carbon adsorption, of four of the standards

used in the study, tested against an organofluorine calibration curve, to see if there was any difference in the combustion efficiencies.

Table 9-2. Combustion Efficiencies of Standards by Direct Combustion

	Inorganic Calibration		Organic Calibration	
Standard	Mass (ng F^-)	% Recovery	Mass (ng F^-)	% Recovery
NaF	1000	99.5	1000	95.9
PFBA	302	48.6	304	98.3
PFOS	653	108.8	667	93.4
Mixed PFAS	742	94.6	745	99.2

Based on the % Recovery data in Table 9.2 for the organofluorine calibration, direct combustion of PFBA had an efficiency two-fold higher than with an inorganic fluoride calibration (i.e., 98.3% versus 48.6%).

Results for Matrix Spikes

Sample #7 was selected by the EPA to be used for this special study. Aliquots of the sample were spiked with the mixed PFAS standard at the same two spiking levels used in Section 8. One pair of MS and MSD aliquots was analyzed for each spike level. The mean recovery and RPD of the MS/MSD for each spike level are provided in Table 9-3. The table also compares the recoveries for this sample between both types of calibration standards. (The result for the inorganic calibration is the average of the MS/MSD from Laboratory 3 only and not the average across all the participating laboratories.)

Table 9-3. Recoveries of Mixed PFAS Standard, Inorganic vs. Organic Calibrations

Nominal Spike Conc (µg F^-/L)	Inorganic Calibration		Organic Calibration	
	Mean Recovery (%)	RPD (%)	Mean Recovery (%)	RPD (%)
10	100.5	14.5	93.0	3.2
30	91.6	1.8	91.5	3.1

GDIT performed statistical analyses using a *t*-test and an ANOVA that demonstrated that there was no statistical difference in the recoveries when the instrument is calibrated using an inorganic vs. an organic fluoride standard when analyzing samples composed of a mixture of PFAS compounds. But given the different affinities of the GAC for short-chain and long-chain PFAS during the adsorption process, the loss of some short-chain PFAS from the mixed standard might not be immediately apparent in the MS/MSD recovery data. Therefore the Method 1621 requirement to use sodium fluoride for the instrument calibration will not be changed. However, further evaluation of a calibration with organofluorine standards may be warranted and may be included in a future revision of Method 1621.

10. Column Breakthrough

Breakthrough occurs when the sorbate leaves the top column due to either low adsorption of the compound to the adsorbent bed, or by exceeding the capacity of the adsorbent bed, and is captured to some extent on the bottom column. Because GAC columns can adsorb other compounds besides organofluorines, high levels of co-adsorbed non-fluorine-containing contaminants can cause breakthrough of organofluorine analytes on the GAC by overloading the bed. Also, short-chain PFAS, like PFBA, are prone to lower recoveries, due to poor retention on the carbon. Therefore, GDIT asked the laboratories to calculate the percent breakthrough between the top and bottom columns for each of the replicate analyses using the equation below.

$$\%\, Breakthrough = \frac{(M_2 - B_2) \times 100}{[(M_1 - B_1) + (M_2 - B_2)]}$$

where,

M_1 = Mass measured for the first column, µg F^-
M_2 = Mass measured for the second column, µg F^-
B_1 = Mass measured for first column of the initial method blank, µg F^-
B_2 = Mass measured for second column of the initial method blank, µg F^-

Percent breakthroughs above 50% were observed in 18 of the 81 unspiked sample results and in 23 of the 36 results for Sample #7 spiked with PFBA. Of the 18 results showing high column breakthrough for the unspiked sample results, six were for unspiked samples above the laboratory's MDL, while the remaining 12 results were for unspiked samples with results below the laboratory's MDL. Because sample results below the laboratory's MDL are considered non-detects, having the % breakthrough above the limit is not considered an issue for these samples. The average column breakthrough was 19.5%, with a maximum breakthrough of 125.2%, which was seen in Sample #6 unspiked for Laboratory 1.

Table 10-1 summarizes the observed percent breakthrough failure rate for the 475 spiked and unspiked samples with detectable AOF, using three potential acceptance criteria. When the criterion was set at ≤ 20%, 165 results failed the limit. When the criterion was widened to ≤ 30%, 86 results failed the limit, and when the criterion was set at ≤ 50%, 29 results failed the limit, with 23 of those results being the spikes for PFBA.

Table 10-1. Observed Failure Rates for Three Potential % Breakthrough Acceptance Criteria

Criterion	# Results	# Failures	Failure Rate (%)
≤ 20%	475	165	34.7
≤ 30%	475	86	18.1
≤ 50%	475	29	4.2

The percent breakthrough observed in the study was ≤ 50% for 94% of the 475 detected results across the nine wastewater samples tested, which could lead some to suggest that column breakthrough issues may not warrant combusting the top and bottom GAC columns separately. However, separate combustion of the top and bottom GAC may still be necessary for certain types of wastewater samples, especially when the sample source is unfamiliar to the laboratory. In situations where the laboratory consistently analyzes the same sample sources for the same client, and the percent breakthroughs for those samples are consistently below 50%, separate combustion of the top and bottom GAC columns may not be necessary, but may require periodic checks. Based on the results from this study, EPA Method 1621 will include recommendations for column breakthrough for samples with concentrations below the laboratory's quantitation limit and for samples with baseline historical data.

11. Data Review and Validation

The laboratories submitted the results for all analyses in this study as electronic data deliverables (EDDs) in Excel format and supported by raw data and reporting forms provided in PDF format that were equivalent to a hardcopy data package. Separate data submissions were provided for each technical directive (e.g., each task) in the study. GDIT reviewed both the EDDs and the supporting raw data for completeness and data quality and evaluated the data based on the preliminary method performance criteria described in the draft method. The EPA will establish formal method performance criteria based on the statistical analyses of the data from this multi-laboratory validation study. The data review process was patterned after that used by GDIT for various other Office of Water studies.

As part of the review process, GDIT added data qualifiers and comments to the EDD when applicable, and the file was saved with a new file name that indicated the data had been reviewed. This approach preserves the original laboratory submission. While some data were qualified as part of this review process, the results were not rejected based on the preliminary acceptance criteria in the method, and the qualifiers applied by GDIT are intended to make the data user aware of potential data quality issues. The exception to this is in cases where the result did not meet qualitative identification criteria (e.g., retention time outside the window).

GDIT evaluated the MDL, IPR, reconnaissance (unspiked real-world samples), MS/MSD, OPR, and method blank results in both the EDD and PDF raw data, while initial calibration and calibration verification results were not provided in the EDD and therefore only the PDF raw data were evaluated. GDIT independently calculated all of the results to ensure that they were within 1% of the reported values. This 1% allowance accounts for rounding differences between the data available on the instrument and the values with fewer decimal places that are typically reported in the hard copy. The data review steps taken by GDIT are briefly described in the sections below.

Completeness check – The data report narratives in the data package were reviewed and any quality control or performance related issues were noted. The data were verified to be consistent with the narrative and appropriate validation qualifiers were applied. EDD elements and results were checked for completeness and consistency with the raw data. Elements checked included the EPA sample identifier, analysis date and time, laboratory qualifiers, found concentrations, sample sizes (volume), dilution factors, spiked amounts, percent recoveries, percent breakthrough, preparation date, and concentration units. Hardcopy and EDD data for all samples and blanks were checked to ensure that no data were missing or inconsistent. Hardcopy data were checked to ensure that all chromatograms and quantitation reports were available for all analyses and that all samples were reported in the sample pretreatment and sample preparation worksheet records.

Initial Calibration – The initial calibrations were checked to ensure that at least 5 calibration standards were analyzed and that they covered the concentration range of 100 – 5,000 ng F^- (1.0– 50 µg F^-/L). The model fit (%RSE) for the initial calibration standards was checked to ensure that it was ≤ 20%. The %RSE calculations were independently performed by GDIT to ensure that they were within 1% of the reported values. For the matrix sample analyses, the calibration was extended to 100 – 10,000 ng F^- (1.0 – 100 µg F^-/L) and that extended calibration also was reviewed and validated.

Calibration Verification (CV) – All samples and blanks were checked to ensure they were bracketed by a CV every 10 samples. The observed recoveries for the CV were within 80 – 120%. The reported concentration calculations were independently performed by GDIT to ensure they were within 1% of the reported values.

Instrument Sensitivity – Results between the MDL and the laboratory's quantitation or nominal reporting limit were checked to confirm they were flagged by the laboratory. Any flagged results were

checked to confirm the concentration warranted the qualifier (e.g., was between the MDL and the quantitation limit). All reported non-detects were checked to confirm no signal was present, or that they were detects below the MDL value.

Blank Contamination – Analyte concentrations reported in the method blank were compared to the results for the associated samples. If the concentration in a sample was greater than 10 times the concentration in the method blank, then the sample result was considered to be unaffected and the validation flag "B, RNAF" was applied to indicate that the blank contamination was at a low enough level to not significantly affect sample results.

OPR Recovery – The percent recoveries for the OPR were checked to ensure they were within preliminary control limits of 70 – 130%. Independent calculations of the percent recovery were performed to ensure they were within 1% of the reported values.

Matrix Spike Recovery – Real-world matrices were used to prepare matrix spike samples and recoveries were checked to ensure they were within the preliminary control limits of 50 – 150%. Percent recoveries for a few samples were recalculated to ensure they were within 1% of the reported values. High MS recoveries indicated a positive interference or a high bias. Isolated instances of high recovery are not uncommon, and patterns across multiple MS samples are more of a concern.

Quantification Check – The reported results in each sample were independently calculated by GDIT, using equations provided in the draft method to ensure they were within 1% percent of the reported values.

Column Breakthrough Check – The percent column breakthrough was checked for each analysis to ensure it was below the preliminary control limit of 50%. Independent calculations were performed to ensure they were within 1% of the reported values.

12. Conclusions

The Office of Water's Engineering and Analysis Division completed the multi-laboratory validation study of EPA Method 1621 for adsorbable organic fluorine in aqueous samples. The multi-laboratory validation study achieved all its intended goals, as outlined below.

1. Obtain data from matrices that are representative of the method's intended use

As described in Section 3, the wastewater matrices were a diverse selection of wastewaters from multiple parts of the country with different physical parameters, as demonstrated in Table 3-2. The aqueous matrices chosen are typical of what might be analyzed by a laboratory performing NPDES compliance monitoring and included some pretreatment samples that would be more challenging than a typical NPDES compliance final effluent.

2. Obtain data from laboratories that are representative of those likely to use the approved method, but that were not directly involved in its development

The majority of the laboratories that participated in this method validation study were commercial laboratories that routinely perform NPDES compliance monitoring analyses. University, EPA, and in-house vendor laboratories also participated. Because commercial laboratories are those most likely to use this method for NPDES compliance monitoring, the use of those laboratories in this study meets this goal.

3. Obtain feedback from laboratory users on the specifics of the draft method

Participant laboratories were encouraged to provide feedback on the method, and most did. Where appropriate, the EPA has revised the method in response to such feedback.

4. Use study data to characterize performance of the method

All the data collected during this study were reviewed and evaluated to characterize the performance of this method, as summarized in detail in Sections 4 through 9. This includes data on calibration, initial precision and recovery, method detection limits, and performance in real-world matrices.

5. Develop statistically derived QC acceptance criteria that will reflect method performance capabilities in real-world situations

Sections 6 through 8 contain statistically derived QC limits that were calculated from the data collected during this study. The laboratories that participated are representative of the real-world laboratories that would potentially run this method, and the matrices are typical of matrices that a laboratory using this method would analyze.

13. References

ASTM, WK68866, *New Test Method for Determination of Adsorbable Organic Fluorine in Waters and Wastewaters by Adsorption on Activated Carbon Followed by Combustion Ion Chromatography*, (no date), www.astm.org

Jones, J., S. Burket, A. Hanley, and J. Shoemaker, 2022, *Development of a Standardized Adsorbable Organofluorine Screening Method for Wastewaters with Detection by Combustion Ion Chromatography*. RSC Publishing, Cambridge, UK, 14(36):3501-3511.

Satterthwaite, F. E., 1946, *An Approximate Distribution of Estimates of Variance Components*, Biometrics Bulletin, 2: 110–114.

USEPA, 2018, *Protocol for Review and Validation of New Methods for Regulated Organic and Inorganic Analytes in Wastewater Under EPA's Alternate Test Procedure Program,* EPA 821-B-16-001, Office of Water, Office of Science and Technology, Engineering and Analysis, Washington, DC, February 2018.

USEPA, 2022, Draft Method 1621*, Screening Method for the Determination of Adsorbable Organic Fluorine (AOF) in Aqueous Matrices by Combustion Ion Chromatography (CIC)*, EPA 821-D-22-002, Office of Water, Office of Science and Technology, Engineering and Analysis, Washington, DC, April 2022.

USEPA, 2022, *Report on the Single-laboratory Validation of Clean Water Act Method 1621 for Adsorbable Organic Fluoride (AOF)*, EPA 820-R-22-003, Office of Water, Office of Science and Technology, Engineering and Analysis, Washington, DC, April 2022.

Appendix A
Study Plan

(This appendix attempts to preserve the pagination and numbering of the original document)

Study Plan/QAPP for the Multi-Laboratory Validation Study for Draft EPA Method 1621 - Adsorbable Organic Fluorine

Prepared for

S. Bekah Burket
U.S. Environmental Protection Agency
Office of Water
Office of Science and Technology
Engineering and Analysis Division (4303T)
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Prepared by

General Dynamics Information Technology, Inc.
3170 Fairview Park Dr.
Falls Church, VA 22042

Prepared under

EPA Scientific and Technical Support Contract EP-C-17-024

September 7, 2022

Approvals

Brian D'Amico, EPA, EAD TASB Chief	Adrian Hanley, EPA, WACOR
S. Bekah Burket, EPA, Alt WACOR/Project Manager	Lemuel Walker, EPA, QA Coordinator
Lynn Walters, GDIT, Program Manager	Harry McCarty, GDIT, Work Assignment Manager
Mirna Alpizar, GDIT, Study Manager	Marguerite Jones, GDIT QA Manager

Acknowledgments

This study plan was prepared under the direction of S. Bekah Burket of the Engineering and Analysis Division (EAD) within the U.S. Environmental Protection Agency (EPA) Office of Water by General Dynamics Information Technology, Inc. (GDIT) under contract to the EPA Office of Water EAD. We thank the ASTM International working group members for their comments during development of this plan and the ASTM D19 Committee for providing the draft standard.

Disclaimer

This document has been reviewed and approved by the Engineering and Analytical Support Branch of EAD. Mention of company names, trade names, or commercial products does not constitute endorsement or recommendation for use.

TABLE OF CONTENTS

Page

DISTRIBUTION LIST

Brian D'Amico, EPA EAD
Adrian Hanley, EPA EAD
S. Bekah Burket, EPA EAD
Lemuel Walker, EPA EAD
Lynn Walters, GDIT
Harry McCarty, GDIT
Mirna Alpizar, GDIT
Marguerite Jones, GDIT
Emily Surpin, GDIT
ASTM D19 workgroup members
EPA AOF workgroup members
GDIT Technical Staff (Study Coordinator, Data Reviewer, Statistician)

Study Plan/QAPP for the Multi-Laboratory Validation Study for Draft Method 1621 - Adsorbable Organic Fluorine

Section 1 Background

Various organizations and regulatory authorities at state, federal, and international levels are taking action to address the release of per- and polyfluoroalkyl substances (PFAS) to the environment. Some of the most common legacy sources of PFAS were from the manufacture of non-stick materials and largely consisted of perfluorinated compounds such as perfluorooctane sulfonate (PFOS) and perfluorooctanoic acid (PFOA). Voluntary efforts to phase out those compounds began in 2008, but a large variety of other PFAS are now in common use as alternatives to PFOS and PFOA. There are other significant sources of PFAS, including aqueous film-forming foams (AFFF) used in firefighting, that contain many more fluorine-containing compounds. In addition to PFAS, fluorine is broadly used in the pharmaceutical and pesticide industries because of the stability of the carbon-fluorine bond. The combination of PFAS and non-PFAS fluorinated compounds are referred to hereafter as organofluorines.

From September 2021 through March 2022, the EPA Office of Water conducted a single-laboratory validation (SLV) study of a screening method for the aggregate measurement of organofluorines based on draft ASTM standard WK68866. Briefly, the organofluorines from aqueous environmental samples are adsorbed on granular activated carbon (GAC) and analyzed by combustion ion chromatography (CIC). EPA posted the procedure that resulted from the SLV as Draft Method 1621 (Ref 8.1).

This document is a hybrid study plan and quality assurance project plan (QAPP) for the multi-laboratory validation of Draft Method 1621. The format and content of the document are consistent with the EPA Office of Water's Hybrid Quality Assurance/Study Plan Format for Method Development and Method Validation Studies, Revision 1, July 2020. As such, it should be considered a "living document" that is intended to be updated as decisions are made by EAD and the study progresses.

The study will be conducted in the following seven phases, as described in Sections 4.1 through 4.7 below.

- Phase 1: Solicitation of Laboratories
- Phase 2: Procurement of Standards, Supplies, and Study Samples
- Phase 3: Initial Calibration and Demonstration of Capability
- Phase 4: Analysis of Study Samples
- Phase 5: Special Study for Calibration with Organic Fluoride Standard
- Phase 6: Data Verification and Validation
- Phase 7: Development of QC Acceptance Criteria

Section 2 Study Objectives and Schedule

The goals of the multi-laboratory validation study are to:

- Obtain data from matrices that are representative of the method's intended use
- Obtain data from laboratories that are representative of those likely to use the method
- Obtain feedback from laboratory users on the specifics of the draft method
- Use the study data to characterize performance of the method
- Develop statistically derived quality control (QC) acceptance criteria that will reflect method performance capabilities in real-world situations

In addition to the overall objective described above, EAD has two general quality objectives for this study:

1. Except where otherwise directed, all validation data must be generated according to the analytical and quality assurance/quality control (QA/QC) procedures specified in this study plan and the draft procedure. Alternatively, the data must be the result of pre-approved and documented changes to the procedure.
2. All data produced must be capable of being verified for accuracy and inconsistencies by an independent reviewer of the analytical data package.

To meet these quality objectives, EPA and GDIT will employ the following QA/QC strategies:

- All project activities will be performed in accordance with this study plan.
- The participating laboratories (either contracted or volunteer) must have demonstrated experience in performing work of a similar nature, preferably experience in EPA Method 1650 (AOX) or similar sorption and combustion procedures in water samples and ion chromatography and must have a comprehensive QA program in place and operating throughout their study operations. The laboratory will be required to follow all QC procedures defined in Section 5 of this Study Plan. (These requirements will be included in the laboratory statement of work or memorandum of understanding.)
- The study report and final draft method will be reviewed by GDIT, EPA, and the EPA workgroup to ensure the QC requirements meet the study objectives.

Cumulatively, these requirements are intended to ensure that the data produced in this study are of appropriate and documented quality.

Approval of this Study/QA Plan by EAD Engineering & Analytical Support Branch will be obtained prior to any of the following activities taking place. The planned schedule for the study is as follows, but may change due to circumstances beyond GDIT's or EPA's control:

Schedule	**Activity**
Month 1	• Solicit and contract services from commercial environmental laboratories • Obtain agreement for participation and a signed memorandum of understanding from each volunteer laboratory • Procure standards, selected supplies, and study samples for the participating laboratories
Month 2	• Prepare and ship sample collection kits to the major wastewater treatment facilities that will be collecting the samples for use in the study. • Collection and shipment of real-world wastewater samples from major wastewater treatment facilities. The samples will be shipped from the collectors to GDIT's facility in Falls Church, Va., for aliquoting. • Analysis of water quality parameters for study samples by a contracted laboratory (Sections 4.2.2 and 4.4) • Initial demonstration of capability by all laboratories.
Month 3	• Ship study samples to the participating analytical laboratories • Laboratory analyses and delivery of the data to GDIT
Month 4	• Conduct data quality review of analytical and QC data • Develop project-specific databases for storage and retrieval of analytical data
Month 5	• Compile data files for each target analyte for statistical analysis • Prepare summary project information and graphics for final reporting

Section 3 Study Management

The study will be managed by S. Bekah Burket, the Alternate Work Assignment Contracting Officer's Representative (Alt WACOR) and EAD Project Manager, under the supervision of Adrian Hanley, the EPA WACOR, who will provide technical direction to GDIT. The EAD Project Manager will be responsible for recruiting and providing formal direction to any EPA laboratories that may participate in the study (see Section 4.1). Lemuel Walker, EAD Quality Assurance Coordinator, will provide QA oversight for EPA.

Day-to-day management and coordination of study activities, including procurement of services from and oversight of commercial (paid) environmental laboratories, and monitoring of laboratory progress, both paid and voluntary, will be performed by GDIT Study Manager, Mirna Alpizar, under the supervision of the GDIT Work Assignment Manager, Harry McCarty, and the GDIT Program Manager, Lynn Walters, and in accordance with EAD guidance. Ms. Alpizar will be supported by a team of GDIT technical staff (i.e., study coordinator, data reviewer, statistician), who will assist with details of study implementation, such as identification of qualified laboratories, coordinating with GDIT's purchasing department, coordinating shipment of supplies and study samples, data review and data analysis. Marguerite Jones, GDIT Quality Assurance Manager, will provide QA oversight for GDIT, with assistance from GDIT's QA Coordinator, Emily Surpin. The organization chart in Figure 1 below illustrates the relationship of these parties.

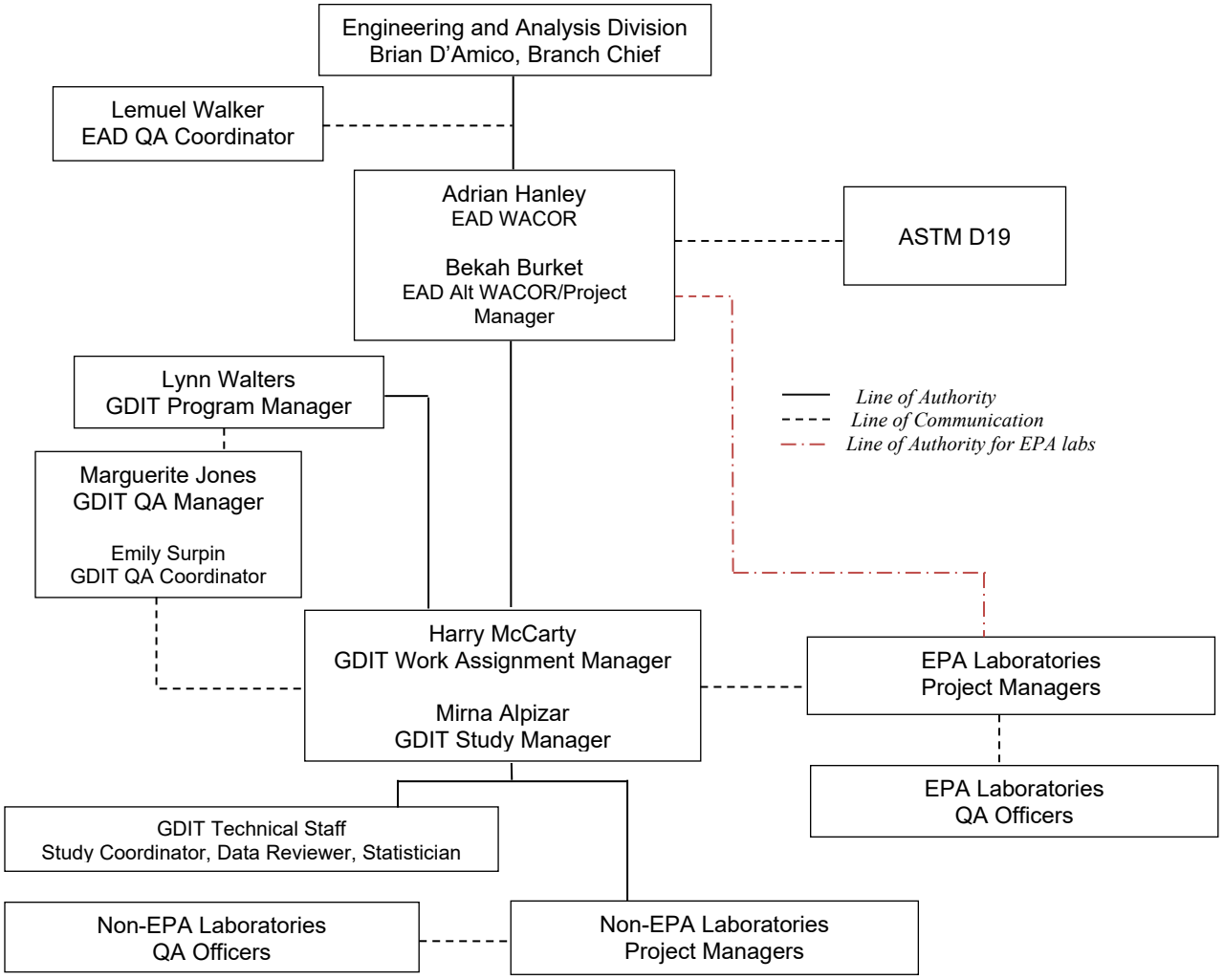

Figure 1. Organization Chart

The laboratories performing the study will be identified in the first phase. In keeping with the approach described in ASTM Standard D2777 (Ref. 8.2), EPA will solicit participation from the largest number of laboratories practical, recognizing the possibility that some participants may drop out or otherwise fail to provide usable data. Currently, EPA plans to include at least 10 laboratories in the study, if practical. Part of the rationale for the number of participants is to gain additional support for promulgation of the final method from the commercial laboratory community. However, the final number of laboratories also will be determined by EPA based on factors such as cost, availability of instrumentation, and EPA's ability to enlist voluntary (unpaid) participation of suitable vendor and/or EPA Regional laboratories.

For this study, the participating laboratories must own and be familiar with the operation of a CIC system and will be responsible for ensuring proper operation of the system for the entirety of this study. The laboratories will, with oversight from the GDIT Study Manager, perform various studies needed to optimize the method. Once the initial tests are complete and findings reviewed and approved by the GDIT Study Manager and EAD Project Manager, the laboratories will prepare and analyze real-world samples for this study.

The laboratories will submit their analytical results to GDIT. GDIT staff, under EAD direction, will review and evaluate all the analytical data and assist EAD in drawing conclusions from the results. GDIT staff will prepare a draft study report that summarizes the results and conclusions for EAD review. Throughout each phase listed in Section 4, GDIT and EPA will also share all data generated during the study with ASTM Committee D19.

GDIT will be responsible for providing the method-required reference standards to each laboratory. To do so, GDIT staff will procure the standards from one or more commercial vendors and coordinate shipment of the standards directly from the vendor(s) to each participating laboratory.

Section 4 Technical Approach

The study will be performed in seven phases. Phase 1 (Section 4.1) involves identification and solicitation of appropriate laboratories with the necessary analytical instrumentation. Phase 2 (Section 4.2) involves the procurement of standards, supplies, and study samples. Phase 3 (Section 4.3) consists of each laboratory performing initial calibration and demonstration of capability. Phase 4 (Section 4.4) involves the evaluation of real-world wastewater study samples. Phase 5 (Section 4.5) involves a special study for calibration with organic fluoride standard. Phase 6 (Section 4.6) involves data verification and validation. Phase 7 (Section 4.7) involves the development of QC acceptance criteria.

The testing under Phase 3 through Phase 7 will be performed using a set of two GAC columns, placed in series on the absorption module. Each column will be combusted separately, and the percent breakthrough calculated for each sample, using Equation 1.

Equation 1 Percent Breakthrough

$$\%\, Breakthrough = \frac{(C_2 - B_2) \times 100}{[(C_1 - B_1) + (C_2 - B_2)]}$$

where:
C_1 = Measured µg F^- for the first column
C_2 = Measured µg F^- for the second column
B_1 = µg F^- for first column from reagent water blank
B_2 = µg F^- for second column from reagent water blank

4.1 Phase 1 – Solicitation of Laboratories

The focus of Phase 1 is to identify and solicit at least 10 laboratories to participate in the study. As noted earlier, some of the laboratories will be contracted by GDIT and others may participate as volunteers. The contracted laboratories are likely to be commercial environmental laboratories, whereas volunteer participants may be EPA Regional laboratories or other organizations that are unlikely to be able to accept payment for their participation.

Participating laboratories will:

- Possess a CIC instrument and the necessary sample processing equipment
- Have experience running the sample preparation process for EPA Method 1650 (AOX) or similar sorption and combustion procedures
- Have staff available who are experienced in method development and evaluation
- Have management support for the project
- Have a comprehensive QA program in place and operating throughout their study operations
- Follow all QC procedures defined in Section 5 of this Study Plan

GDIT and EPA will develop a broad list of likely participants and contact them in advance of a formal solicitation to determine their potential interest. Once the list of potential participants has been established, GDIT staff will competitively solicit bids using government-approved procurement procedures and an EAD-approved statement of work (SOW), or equivalent documentation that details the requirements for sample preparation, storage, shipment, analysis, QA/QC, and reporting. The SOW also will stipulate that the laboratory must have a comprehensive laboratory QA program in place and operating at all times during performance under the SOW and this program must be consistent with EPA *Guidance for Developing Quality Systems for Environmental Programs* (Ref. 8.3) and the general laboratory procedures specified in the *Handbook for Analytical Quality Control in Water and Wastewater Laboratories* (Ref. 8.4). The GDIT Study Manager will also work with EPA to develop suitable mechanisms to engage any volunteer laboratories identified by EPA. Such mechanisms may involve a memorandum of understanding (MOU) and/or a voluntary participation agreement form previously developed by EPA for similar studies.

Regardless of the nature of a laboratory's participation, contracted or volunteer, the same study requirements will apply and will be described in a study-specific statement of work and study-specific instructions.

4.2 Phase 2 – Procurement of Standards, Supplies, and Study Samples

Phase 2 of the study involves procuring sufficient quantities of the analytical standards needed to perform the method, GAC columns, column holders, and the samples that will be analyzed in the study.

4.2.1 Standards and Selected Supplies

Providing the same standards, GAC columns, and column holders to each laboratory removes these as sources of variability and provides an incentive for both contracted and volunteer laboratories to participate.

After identifying likely commercial sources for the standards and GAC columns, GDIT will use government-approved procurement procedures and an EPA-approved SOW (or equivalent) to obtain sufficient quantities of the needed materials and have them shipped directly to the participating laboratories. GDIT anticipates that these materials will need to be procured from multiple existing commercial sources.

Once the GDIT purchase orders are in place, GDIT staff will work with the vendors to schedule and direct the shipments of materials to each participating laboratory. GDIT Study Manager will notify each laboratory of impending shipments, track each shipment from the vendor to the laboratory, and confirm condition of the materials on receipt with each laboratory. GDIT Study Manager will work with the vendors and laboratories to resolve any issues or discrepancies and will communicate with EPA regularly.

4.2.2 Study Samples

The focus of the study is on the analysis of real-world aqueous environmental matrices. EPA and GDIT will work with municipal, state, and EPA Regional contacts to obtain sufficient volumes of up to ten real-world wastewaters to be used in the study. The wastewater samples will include effluents from a publicly owned treatment works (POTW) and wastewaters from specific industrial discharges if they can be obtained in sufficient quantities. At least one of the wastewater matrix types should have one of the following characteristics, such that each criterion below is represented by at least one wastewater:

- Total suspended solids (TSS) greater than 40 mg/L
- Total dissolved solids (TDS) greater than 100 mg/L
- Oil and grease greater than 20 mg/L
- NaCl greater than 120 mg/L
- $CaCO_3$ greater than 140 mg/L

The study samples will be collected in bulk quantities by water treatment facilities and shipped to GDIT. GDIT staff will then prepare and ship aliquots of each sample to each of the laboratories participating in the study. Prior to shipping study samples, GDIT staff will coordinate schedules with laboratory staff. GDIT Study Manager also will confirm condition of the samples on receipt, work with the laboratories to resolve any issues or discrepancies, and communicate regularly with EPA.

4.3 Phase 3 – Initial Calibration and Demonstration of Capability

Prior to analyzing any of the study samples, each laboratory will perform an initial multi-point calibration and conduct an initial demonstration of capability. Before proceeding with later phases of the effort, EAD, GDIT, and the laboratories will discuss the results from Phase 3. To further reduce analytical variation, the laboratories will be required to use the same ion chromatography (IC) columns based on their IC instrument manufacturer. For example, if more than one laboratory uses a Thermo-Fisher IC instrument, then these laboratories will be required to use the same IC column model, mobile phase, and gradient.

4.3.1 Initial Calibration

Each laboratory will perform an initial calibration using a sodium fluoride standard as specified in Draft Method 1621. The calibration must contain at least 5 calibration levels and fall within the applicable range of 1 µg F^-/L to 100 µg F^-/L.

The calibration model to be used will be external standard calibration by combustion. A simple linear regression model will be used, not forced through the origin. The acceptability of the calibration curve shall be determined using the relative standard error (RSE). The goal of this study is that the RSE be at or below 20% in order to establish linearity. However, the final method acceptance criteria will be derived from the results of the study. The RSE is to be calculated using Equation 2 below.

Equation 2. Relative Standard Error

$$RSE = 100 \times \sqrt{\sum_{i=1}^{n} \frac{\left[\frac{x_i' - x_i}{x_i}\right]^2}{n - p}}$$

where:
x_i = Nominal concentration (true value) of each calibration standard
x'_i = Measured concentration of each calibration standard
n = Number of standard levels in the curve
p = Type of curve (1 = average, 2 = linear, 3 = quadratic)

4.3.2 Method Detection Limit Study

Following successful demonstration of the calibration range and linearity, the laboratory will determine the Method Detection Limit (MDL) using the revised MDL procedure published in 2017 at Code of Federal Regulations (CFR), Title 40, Part 136, Appendix B (Ref. 8.5). The laboratory will prepare and analyze at least seven replicate reagent water samples spiked with sodium perfluorohexanesulfonate (PFHxS) and an equal number of unspiked reagent water samples over three non-consecutive days and determine the MDL_s (the MDL based on spiked samples) and MDL_b (the MDL based on method blanks), respectively. The spiking level will be based on the guidance in the revised MDL procedure, the results for the initial calibration, and other information, such as recovery data from the SLV study.

Given the prevalence of low levels of PFAS components in sampling equipment and instrumentation, the MDL for the procedure may be based on the MDL_b, rather than the MDL_s. Several iterations of the MDL study may be required. EAD, GDIT, and the laboratories will discuss the results from the MDL study before proceeding with later phases of this effort.

4.3.3 IPR Determination

Each laboratory will perform an initial precision and recovery (IPR) study to provide sufficient data for development of performance specifications. The IPR consists of four replicate samples of reagent water spiked with PFHxS at around the midpoint of the calibration range and carried through the entire analytical process (sample preparation and analysis).

Each laboratory will calculate the percent recovery (% R) using Equation 3 below and report it to three significant figures (e.g., 102%).

Equation 3. Percent Recovery

$$\%\, R = \frac{C_s}{C_n} \times 100$$

where:
C_s = Measured concentration of the spiked sample aliquot
C_n = Nominal (theoretical) concentration of the spiked aliquot

Each laboratory will also calculate the relative standard deviation (RSD) from the results of the four replicates using Equation 4 below and report it to two significant figures.

Equation 4. Relative Standard Deviation

$$RSD = \frac{SD}{C_{avg}} \times 100$$

where:
SD = Standard deviation of C_s for the four replicates
C_{avg} = Average measured concentration for the four replicates

IPR results generated during Phase 3 will be used to develop the final IPR performance criteria (% R and RSD) during Phase 7 of the study.

4.4 Phase 4 - Analyses of Study Samples

Pre-shipment Reconnaissance Analyses

Because some study participant laboratories (e.g., instrument vendors, university research laboratories) may not have the capability to perform the sample pretreatment checks required in the draft method, a single laboratory will be selected to perform reconnaissance of the study samples by analyzing for pH, chlorine, inorganic fluoride, and background AOF levels, prior to shipment of study samples to the participating laboratories. EPA and GDIT will use the reconnaissance analyses results to establish the spike levels to be used by the laboratories and to determine if any of the samples will need pH adjustment or dilution prior to analysis.

Study Samples and Aliquots

Each participating laboratory will receive up to ten samples, consisting of:

- ten (10) 125-mL sample aliquots for up to nine of the aqueous matrices, and
- twenty-six (26) 125-mL sample aliquots of one aqueous matrix

The number of containers ensures the laboratories will have enough volume in the event reanalysis is required or an accidental spill happens during sample shipment.

Parts 1, 2, and 3 Analyses

The sample analyses in this study will be divided into three parts, as described below and outlined in Table 1. The study design will yield up to 100 unspiked sample results and up to 580 spiked sample results if 10 laboratories participate in the study and all 10 complete all required analyses.

- For Part 1, each laboratory will be required to analyze each of the study samples unspiked for the laboratory-specific AOF background. Each laboratory will use their unspiked sample results to adjust the concentration of their spiked results and calculate the percent recoveries for the spiked sample results from Part 2 and Part 3 of the study.

- For Part 2, each laboratory will spike a matrix spike/matrix spike duplicate (MS/MSD) pair for each sample at two spiking levels using PFHxS.

- For Part 3, each laboratory will use one sample, selected by EPA, to prepare separate MS/MSD pairs spiked with the following standards: PFOS, PFBA, and a mixed PFAS standard, and spiked at two spike levels (i.e., six MS/MSD pairs total).

Table 1 Study Analytical Design

Study Part	Matrix ID	ID	Spiked Analyte	# Replicates per Lab	Nominal Value*
Part #1	1 – 10	Unspiked	None	1 per Matrix	None
Part #2	#1	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
	#2	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
	#3	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
	#4	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
	#5	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
	#6	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
	#7	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
	#8	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
	#9	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
	#10	MS/MSD-1	PFHxS	2	Spike Level Low
		MS/MSD-2	PFHxS	2	Spike Level High
Part #3	# X To be selected by EPA	MS/MSD-1	PFOS	2	Spike Level Low
		MS/MSD-2	PFOS	2	Spike Level High
		MS/MSD-1	PFBA	2	Spike Level Low
		MS/MSD-2	PFBA	2	Spike Level High
		MS/MSD-1	Mixed PFAS	2	Spike Level Low
		MS/MSD-2	Mixed PFAS	2	Spike Level High

*Spike concentrations will be selected after analysis of the reconnaissance AOF background concentrations. The laboratories will be informed what nominal concentration to use for each spike level.

Determination of MS/MSD Precision and Recovery

The laboratories will calculate the percent recovery and precision of the MS/MSD samples using Equation 5 and Equation 6 below. Percent recoveries and the relative percent difference (RPD) will be reported to two significant figures.

Equation 5. Percent Recovery for MS/MSD Samples

$$\%\,R = \frac{C_s - C_i}{C_n} \times 100$$

where:
C_s = Measured concentration of the spiked sample aliquot
C_i = Measured concentration of the unspiked sample aliquot
C_n = Nominal (theoretical) concentration of the spiked aliquot

Equation 6. Relative Percent Difference

$$RPD = \frac{|R_2 - R_1|}{(R_1 + R_2)/2}$$

where:
R_1 = Recovery for MS
R_2 = Recovery for MSD

4.5 Phase 5 – Special Study for Calibration with Organic Fluoride Standard

One laboratory will be directed to perform a special study for comparison of instrument calibration using inorganic fluoride vs. organic fluoride. That laboratory will perform an initial calibration as in Phase 3 but using PFHxS instead of sodium fluoride. All the acceptance requirements spelled out in Section 4.3.1 will apply to this calibration as well.

That laboratory will also analyze one sample, selected by EPA, spiked at two concentrations using a mixed PFAS standard. The laboratory will analyze one MS/MSD pair for each spike level.

4.6 Phase 6 – Data Verification and Validation

All the results from laboratories participating in the study will be reviewed and validated by GDIT staff, relative to the study's goals. Every data submission will be checked for completeness (e.g., verify that all study samples were analyzed, and all required results were submitted) and to determine if the supporting documentation indicates that the laboratory followed the method and study-specific instructions.

Each laboratory will be required to submit the raw data (e.g., instrument printouts and copies of sample preparation records) that supports the study results. GDIT staff will examine all the raw data, perform spot checks of a percentage of the calculations from each laboratory, and ensure that the reported results can be traced back through all steps in the analytical process. If any issues are identified, GDIT Study Manager will work with the laboratory to clarify the situation, obtain any missing information, and document the resolution. GDIT Study Manager will advise EPA of the status of the review efforts on a regular basis.

Because this is a method validation effort, there are no *a priori* quality control acceptance criteria, and data from the study will not be excluded from consideration simply because they appear to fail the draft method performance criteria which were based on the SLV study. Every effort will be made to retain as many results as practical. GDIT staff will flag results from samples with obvious documented failures (e.g., adsorption unit contamination) for exclusion from use in developing method performance information and will document the rationale for such exclusions in the project files and/or the project database. However, in the absence of evidence of such failures, all the results will be included in the initial data set. GDIT and EPA will use statistical tests to determine if results for specific laboratories or samples may be outliers that should be removed from use in developing QC acceptance criteria. Suspected outliers will be examined in detail by the GDIT Study Manager and the laboratory before they are excluded from use in developing method performance summaries.

4.7 Phase 7 – Development of QC Acceptance Criteria

The last major phase of the study will be to develop statistically based QC acceptance criteria and summarize method performance in real-world samples. The overall procedures used for that process are described in Section 7 of this study plan.

Section 5 Quality Assurance and Quality Control Procedures

For this study, the laboratories will perform the following traditional QC checks found in Draft EPA Method 1621:

- Initial calibration (ICAL), 5-point minimum

- Initial precision and recovery (IPR)
- Calibration verification (CV)
- Method blanks carried through the entire procedure
- Ongoing Precision and Recovery (OPR)

Because the study will analyze many pairs of spiked samples, there is no requirement to prepare additional routine MS/MSD samples with the study sample batches.

See Table 2 below for frequency and study requirements for each of the QC checks.

Table 2 QC Checks

QC Check	Frequency	Acceptance Criteria	Study Requirements
Initial calibration (ICAL), 5-point minimum	Once at the beginning of the study and as needed throughout the study	% RSE of the calibration curve should be ≤ 20%. Report all results as generated.	ICAL data will be provided by each laboratory and compared to typical calibration criteria
Initial demonstration of capability (IPR and MDL)	Once	IPR % recovery and % RSD criteria will be established after review of study data. Report all results as generated	Each laboratory will provide IPR data. Study data will be used to develop criteria
Calibration verification (CV)	Beginning and at the end of the analytical batch. Initial calibration can be used in place of initial CV if samples are analyzed within 24 hours of initial calibration.	CV % recovery should be within ± 20% of the mid-level standard from the initial calibration. Report all results as generated.	CV data will be provided by each laboratory and compared to typical calibration criteria
Method blank	Two per batch of 10 field samples or fewer	To be determined based on study data. Report all results as generated.	Perform as specified in the procedure
Ongoing precision and recovery (OPR)	One per preparation batch of 10 field samples or fewer	To be determined based on study data. Report all results as generated. Study data will be used to develop criteria.	Perform as specified in the draft method to demonstrate performance at the OPR concentration.
Matrix spike/ Matrix spike duplicate (MS/MSD)	One pair per spike level per study sample. See Table 1.	To be determined based on study data. Report all results as generated.	MS/MSD results will be used to demonstrate method performance. RPD between the MS and MSD analyses will be used to develop acceptance criteria for duplicate unspiked analyses.
Qualitative Identification Criteria	Fluoride peak in all analyses	RT within ± 0.2 min of initial CV or ICAL RT if the ICAL was performed with the batch of samples	Perform as specified in procedure

It is each of the laboratories' responsibility to maintain their instrumentation and to ensure that all study samples are analyzed on a properly calibrated instrument. If the calibration linearity is outside the nominal criteria, the laboratory will take standard measures to attempt to correct the problem before any study samples are analyzed. The laboratories are also responsible for inspecting all study samples and standards to ensure they meet all the study requirements. If typical measures do not correct the problem or if study schedules will be impacted due to necessary repairs or replacement of study samples or standards, the laboratory will notify the GDIT Study Manager to indicate the impact on study schedules, the laboratory's plans to resolve the problem(s), and if any study samples will need to be reanalyzed.

Each laboratory will report the results from the QC operations, either in electronic format, or if necessary, in hard copy. GDIT staff will compile the QC results in a database specific to this project (See Section 6).

Section 6 Data Reporting and Management

6.1 Laboratory Reporting Requirements

Each laboratory will be required to submit summary-level data. The summary-level data must be submitted in PDF and electronic format and must include:

- GDIT-assigned sample number
- Analyte identification
- Laboratory sample number (if applicable)
- Type of sample (e.g., method blank, OPR, sample)
- Dilution factor (if applicable)
- Spike levels nominal concentration
- Measured concentration
- Reporting units (µg F^-/L)
- Sample preparation date
- Time of analysis (hour and minute)
- Analytical batch ID (if applicable)
- % Breakthrough

Each laboratory will be required to report supporting raw data for all analyses performed and maintain their raw data for a period of seven (7) years and provide them upon request (at additional cost negotiated as necessary).

Raw data shall include:

- Calibration data summary
- Chromatogram and quantitation report for analysis of each sample, blank, calibration standard, and calibration verification standard and all manual worksheets pertaining to sample or QC data or the calculations thereof. Chromatograms must contain the following header information: laboratory-assigned identifier (lab file ID), date and time of analysis, and instrument ID. Chromatograms should be labeled with the name of the analyte, either directly out from the peak or on a printout of retention times. Quantitation reports must contain determined concentrations, area or peak responses, and RTs.
- Sample preparation bench sheets
- Copies of laboratory notebooks showing weights, volumes, and calculations for preparation of standards
- Certificates of analysis for all standards used
- Other data that will allow verification of the calculations performed, such that a third party can recalculate the final results from the raw data.

The laboratory will adhere to the following rules when reporting data:

- All reports and documentation, including instrument printouts and other raw data, must be sequentially paginated, clearly labeled with the laboratory name, and labeled to provide sufficient identification for method blanks, calibration, etc., necessary to link the raw data with associated summary reports.

- Results from all analyses, including failed experiments, must be reported, including calibration data and any dilutions or re-analyses performed. The laboratory must also include an explanation for any re-analyses performed and identify which of the analyses the laboratory considers to be the most appropriate for use.
- Results of all measurements must be reported to two significant figures in µg F^-/L to facilitate review and evaluation.
- The terms "zero" and "trace" are not to be used; the term "not detected" (ND) is to be used for each measurement for which no signal is produced or if method-specified qualitative identification criteria are not met.
- Every value must be reported, even if the value is negative. If the value is below the lowest calibration standard, a "J" flag must be applied to this value.
- Results must be reported for all study samples, including QC samples. Although sample results are blank corrected, the results for both method blanks must also be reported with the associated sample data.

In addition, the laboratory will be required to submit a written "narrative report" with each data package. The narrative report will contain detailed descriptions of any difficulties encountered in the generation of the analytical results and QC data and any attempts to resolve the difficulties. It also will contain a detailed description of any necessary modifications to the draft AOF procedure.

Each laboratory must have a comprehensive data management plan in place that is consistent with the principles set forth in the Good Automated Laboratory Practices (Ref. 8.6), or with commonly employed data management procedures approved by The NELAC Institute (TNI). This data management plan must be in place and in use at all times during the performance of this study.

6.2 GDIT Data Management and Reporting

GDIT will store all study records and submitted data (hard copy and electronic) in an organized fashion on a secure local area network that is backed up nightly and as needed in hardcopy files in their secure office facility, throughout the duration of their contract. Cumulatively, these files will include the following documents and records:

- This study plan (including all submitted draft versions, comments, and revisions)
- Documentation of the procedures used to assess the competency of laboratories participating in this study
- Documents and records associated with the solicitation and award of participant laboratories, including the SOWs or equivalent that describe participant laboratory requirements
- Documents and records associated with the procurement of standards and study samples, including SOWs or equivalent that describe the process used to collect and produce study samples
- The name, address, phone number and primary contact at the standards vendor and each participating laboratory
- Copies of all written correspondence (excluding emails) with laboratory staff, sampling personnel, and EPA staff regarding the study
- A log (or other record) that documents verbal communication with laboratory staff, sample coordinators, sampling personnel, and EPA staff regarding study status or problems
- Records concerning sample shipment and receipt
- All analytical data resulting from this study
- All laboratory comments on the method resulting from this study
- Records of all GDIT data review assessments and statistical analyses submitted to EPA

- All draft and final reports submitted to EPA pertaining to this study

Section 7 Evaluation of Method Performance

EPA's overall goal is to develop method performance data for Draft EPA Method 1621 and statistically-derived QC acceptance criteria that reflect method performance. The results of the analyses in Phases 3 and 4 of this study will be evaluated using common statistical procedures (References 8.2, 8.7 - 8.8) such as:

- *t*-tests, to determine if the mean result for an analyte differs between "treatments" (e.g., spiking levels or background levels of potential interferences)
- F-tests, to determine if the variance (standard deviation squared) for an analyte differs between "treatments"
- Analysis of variance (ANOVA), to determine how the differences between the components and the treatments affect overall variability

All study results will be subjected to statistical evaluations, and suspected outliers will be examined in detail by the GDIT Study Manager and the laboratory before they are excluded from use in developing method performance summaries. GDIT technical staff will perform the statistical evaluations that will be used for the establishment of method QC acceptance criteria; however, GDIT will not perform the evaluation for the Youden pair graphical method. ASTM may choose to use the data provided by EPA for that evaluation.

EPA and GDIT will use the results from the replicate samples to develop QC acceptance criteria for initial precision and recovery (IPR) studies, ongoing precision and recovery (OPR) samples, MS/MSD recovery limits, relative percent difference (RPD) limits, etc. A general description of the derivation of those QC acceptance criteria is provided in Appendix A and is based on EPA's protocol for evaluation of new methods (Reference 8.9).

EPA and GDIT will develop tables of method performance data, including precision and accuracy, as a function of analyte concentration that will provide an indication of expected performance of the procedures under typical conditions. Such tables will be included in the revised procedure as further evidence of its overall capabilities or limitations.

Finally, EPA and GDIT will prepare a formal report on the results of the multi-laboratory validation study. EPA will submit that draft report to appropriate levels of management review within EPA and revise the report as needed. The study report will also be shared with the ASTM D19 workgroup. If the study is successful and EPA decides to move forward with rulemaking to approve EPA Method 1621 at 40 CFR Part 136 for use in nationwide compliance monitoring, the study report and records from the study will be placed into the rulemaking docket (Ref 8.10).

Section 8 References

8.1 Draft Method 1621 Screening Method for the Determination of Adsorbable Organic Fluorine (AOF) in Aqueous Matrices by Combustion Ion Chromatography (CIC), April 2022, USEPA, Office of Water, Engineering and Analysis Division. https://www.epa.gov/system/files/documents/2022-04/draft-method-1621-for-screening-aof-in-aqueous-matrices-by-cic_0.pdf

8.2 ASTM 2013. ASTM D2777-21, Standard Practice for Determination of Precision and Bias of Applicable Test Methods of Committee D19 on Water, ASTM International, West Conshohocken, Pennsylvania.

8.3 US Environmental Protection Agency. 2002. *Guidance for Developing Quality Systems for Environmental Programs*. U.S. Environmental Protection Agency. Office of Environmental Information. EPA QA/G-1, EPA/240/R-02/008.

8.4 US Environmental Protection Agency. 1979. *Handbook for Analytical Quality Control in Water and Wastewater Laboratories.* U.S. Environmental Protection Agency. Office of Research and Development. EPA-600/4-79-019.

8.5 *Federal Register*. Vol. 82, No. 165. August 28, 2017. Definition and Procedure for the Determination of the Method Detection Limit. pp. 40939 – 40941, Federal Register Online via the Government Printing Office.

8.6 US Environmental Protection Agency. 1995. *Good Automated Laboratory Practices*. U.S. Environmental Protection Agency. Office of Information Resource Management. EPA-2185.

8.7 SAS Institute Inc. 1994. SAS/STAT User's Guide, Volume 2, GLM-VARCOMP. Version 6, 4th Edition, June 1994.

8.8 Berry, D. A.; Lindgren, B. W. 1990. Statistics: Theory and Methods. pp. 286-290, 600-618. Brooks/Cole Publishing Company. Pacific Grove, California.

8.9 US Environmental Protection Agency. 2016. *Protocol for Review and Validation of New Methods for Regulated Organic and Inorganic Analytes in Wastewater Under EPA's Alternate Test Procedure Program*. U.S. Environmental Protection Agency. Office of Water, Engineering and Analysis Division. EPA 821-B-16-001.

8.10 *Federal Register* Volume 82, Number 165. August 28, 2017. Clean Water Act Methods Update Rule for the Analysis of Effluent; Final Rule. pp. 40836 - 40941, Federal Register Online via the Government Printing Office.

Appendix A

Procedures for Derivation of QC Acceptance Criteria from the Validation Study Results

The information below has been excerpted from Reference 8.9, and all citations to section numbers and references apply to the original document, not this study plan. The calculations used for this study will be adjusted for the actual number of laboratories involved.

Quality Control Acceptance Criteria Development for New Methods

Method Detection Limits and Minimum Levels

Each laboratory participating in the validation study must perform an MDL study as described in Section 3.1.1. The organization responsible for developing the new method must establish an MDL for the method, using a pooled MDL from at least six laboratories. A pooled MDL is calculated from *m* individual laboratory MDLs by computing the square root of the mean of the squares of the individual MDLs and multiplying the result by a ratio of *t*-values to adjust for the increased degrees of freedom.

Note: The MDL values used in this calculation are those determined in each of the six or more laboratories. If one laboratory reports an MDL_s (from spiked samples), that value is used in conjunction with the MDL values from the other laboratories, including any values reported as MDL_b (from blanks).

$$MDL_{pooled} = \sqrt{\frac{d_1 \left(\frac{MDL_{Lab\ 1}}{t_{(0.99,d_1)}}\right)^2 + d_2 \left(\frac{MDL_{Lab\ 2}}{t_{(0.99,d_2)}}\right)^2 + \cdots d_i \left(\frac{MDL_{Lab\ m}}{t_{(0.99,d_m)}}\right)^2}{d_1 + d_2 + \cdots d_m}} \times t_{(0.99,[d_1 + d_2 + \cdots d_m]}$$

where:
m = The number of laboratories, and
d_i = The number of degrees of freedom used by Lab i to derive the MDL.

In the case of 9 laboratories with 7 replicates per laboratory, the equation simplifies to:

$$MDL_{pooled} = \sqrt{\frac{MDL_{Lab\ 1}^2 + MDL_{Lab\ 2}^2 + \cdots MDL_{Lab\ 9}^2}{9}} \times \frac{2.41}{3.14}$$

The organization responsible for developing the method must also use this MDL to develop an ML. Procedures for determining the ML are given in Section 3.1.1.

Calibration Linearity

The instrument or analytical system is then calibrated with six standards specified in the method to calculate an initial RSD for the response factor.

The RSD and the RSD limit for the CF, RF, or RR is determined as follows:

1. Calculate the mean and standard deviation of the CFs, RFs, or RRs for each laboratory.

$$Mean\ Factor = \overline{Factor} = \frac{\sum_{i=1}^{n} Factor_i}{n}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} (Factor_i - \overline{Factor})^2}{n-1}}$$

where:
Factor = The "Factor" terms are replaced by the CF, RF or RR terms, based on the quantitation approach described in the method in question, and
n = The number of calibration points used in each laboratory.

2. Calculate the relative standard deviation of the CFs, RFs, or RRs of each laboratory and analyte as:

$$RSD_i = 100 \times \frac{s_i}{\overline{Factor_i}}$$

where s_i and $\overline{Factor_i}$ are the standard deviation and mean of the CFs, RFs, or RRs for laboratory *i*.

3. Calculate the pooled RSD of the CFs, RFs, or RRs for each analyte from *all* laboratories. The pooled RSD is calculated as the square root of the mean of the squares of the sample RSDs from each individual laboratory. For example, for nine laboratories, the pooled RSD is calculated as:

$$RSD_{pooled} = \sqrt{\frac{RSD_1^2 + RSD_2^2 + RSD_3^2}{3}}$$

4. Calculate RSD_{max} as the smaller of 35% and:

$$RSD_{max} = k(RSD_{pooled})$$

where:
k = The square root of the 95th percentile of an F distribution with *n* - 1 degrees of freedom in the numerator and *m*(*n* - 1) degrees of freedom in the denominator,
m = The number of laboratories, and
n = The number of calibration points.

For nine laboratories using a five-point calibration (*m* = 9, *n* = 5), the value of *k* is 1.6. The maximum allowable specification for RSD_{max} is 35%.

Calibration Verification

As noted in Section 2.2, acceptance limits for calibration verifications can be determined in three different ways, each of which is described below.

The calibration verification criterion may be specified as a maximum relative distance between the mean CF, RF, or RR obtained by a future laboratory's initial calibration ($\overline{Factor}$) and the CF, RF or RR obtained from its calibration verification standard ($Factor_{VER}$). The maximum allowable deviation is based on the pooled relative standard deviation (RSD_{pooled}) calculated in Section 3.2.2.

1. Determine k_{VER} by multiplying the 97.5th percentile of a Student's *t* distribution with (m[n-1]) degrees of freedom times the square root of (1+1/*n*), where there are *n* points in the calibration and *m* laboratories:

$$k_{VER} = t\sqrt{\left(1 + \frac{1}{n}\right)}$$

For a five-point calibration, the Student's *t* value is 2.0, resulting in combined multipliers of 2.4 for a three-point calibration, and 2.2 for a five-point calibration.

2. The calibration verification criterion for the new method would then be stated as the maximum percent difference as follows:

$$Percent\ Difference = 100 \times \left(\frac{Factor_{VER} - \overline{Factor}}{\overline{Factor}}\right) \leq k_{VER}\ RSD_{pooled}$$

where:
Factor = The "Factor" terms are replaced by the CF, RF or RR terms, based on the quantitation approach described in the method in question

For example, if the calibration verification criterion, calculated as k_{VER} RSD_{pooled}, equals 17%, then the difference between the $\overline{Factor}$ from the initial calibration and the $Factor_{VER}$ from the calibration verification sample must be less than or equal to 17% of the $\overline{Factor}$.

When using either the concentration or the recovery approach, the calculations are very similar to those used for the "factor" limits shown above.

3. Calculate the upper and lower QC acceptance criteria for the known concentration of the analyte in the calibration verification standard, using the lower and upper percentages calculated in Step 2 above:

$$Lower\ limit = (Lower\ Percentage\ in\ Step\ 2) \times (Known\ Concentration\ in\ Standard)$$
$$Upper\ limit = (Upper\ Percentage\ in\ Step\ 2) \times (Known\ Concentration\ in\ Standard)$$

Alternatively, calibration verification criteria may be specified as the range of acceptable recoveries calculated for the analytes in the calibration verification standard, using the lower and upper percentages calculated in Step 2 above to create a window around 100% recovery.

Initial and Ongoing Precision and Recovery

For the IPR and OPR tests, QC acceptance criteria are calculated using the mean percent recovery and the standard deviation of recovery from the IPR tests of four aliquots of the reference matrix and the OPR test of one aliquot of the reference matrix (for a total of five samples) in nine laboratories. The QC acceptance criteria are developed using the following steps:

1. Calculate the mean percent recovery ($\overline{X}$) for each analyte, based on all data points from all laboratories, the between-laboratory standard deviation (s_b) of the mean results for each of the six or more laboratories (standard deviation of the nine laboratory means $\overline{X_1}$ + $\overline{X_2}$ + … $\overline{X_9}$), and the pooled within-laboratory standard deviation (s_w). The value of s_w is calculated as the square root of the mean of all within-laboratory variances. For example, for nine laboratories:

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

where:
$\overline{X}_j$ = The mean percent recovery for the *jth* laboratory
m = The number of laboratories, and
$\bar{X}$ = The overall mean of the percent recoveries from all laboratories

$$s_w = \sqrt{\frac{s_1^2 + s_2^2 + \cdots s_9^2}{9}}$$

Note: GDIT will provide direction to the participating laboratories to ensure they are spiking IPR and OPR samples at the same concentration.

2. QC acceptance criteria for IPR recovery - Calculate the combined standard deviation for interlaboratory variability and estimation of the mean (s_c) as:

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

where:
m = the number of laboratories, and
n = the number of data points per laboratory.

For 9 laboratories and 5 data points per laboratory, the calculation becomes:

$$s_c = \sqrt{\left(\frac{10}{9}\right) s_b^2 + \left(\frac{1}{20}\right) s_w^2}$$

3. Calculate the QC acceptance criteria for recovery in the IPR test by constructing a ± 2.3 s_c window around the mean percent recovery $\bar{X}$, where 2.3 is the 97.5th percentile Student's t value for 10 degrees of freedom (an estimated degrees of freedom based on the variance ratios observed with EPA Method 1625):

$$Lower\ limit\ (\%) = \bar{X} - 2.3 s_c$$
$$Upper\ limit\ (\%) = \bar{X} + 2.3 s_c$$

If more than 9 laboratories are used, the degrees of freedom for t will increase, but a complete calculation is beyond the scope of this document. An approximation of degrees of freedom equal to the number of laboratories will serve for most situations.

4. QC acceptance criterion for IPR precision - The maximum acceptable RSD for the four IPR aliquots is approximated by a 95% upper confidence limit around the observed RSD of the results from all of the laboratories. The RSD_{IPR} (computed as s_w divided by $\bar{X}$) is multiplied by the square root of a 95th percentile F value with 3 degrees of freedom in the numerator and m (n - 1) degrees of freedom in the denominator, where m = the number of laboratories, and n is the number of data points per laboratory. For example, the resulting multiplier on the RSD for nine laboratories and five data points per laboratory will then be 1.7, and the QC acceptance criterion for precision in the IPR test is calculated as follows:

$$\text{Maximum } RSD_{IPR} = (1.7) \text{ x } RSD_{IPR}$$

5. QC acceptance criteria for OPR recovery - Calculate the combined standard deviation for interlaboratory variability and estimation of the mean (s_c) as:

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

where:
m = the number of laboratories, and
n = the number of data points per laboratory.

For 9 laboratories and 5 data points per laboratory,

$$s_c = \sqrt{\left(\frac{10}{9}\right) s_b^2 + \left(\frac{4}{5}\right) s_w^2}$$

6. Calculate the QC acceptance criteria for recovery in the OPR test by constructing a ± 2.1 s_c window around the mean percent recovery $\bar{X}$, where 2.1 is the 97.5th percentile Student's *t* value for 19 degrees of freedom (an estimated degrees of freedom based on the variance ratios observed with EPA Method 1625):

$$Lower\ limit\ (\%) = \bar{X} - 2.1 s_c$$
$$Upper\ limit\ (\%) = \bar{X} + 2.1 s_c$$

If more than 9 laboratories are used, the degrees of freedom for *t* will increase, but a complete calculation is beyond the scope of this document. An approximation of degrees of freedom equal to the number of laboratories will serve for most situations.

Matrix Spike and Matrix Spike Duplicate

Results of the MS/MSD analyses performed in the validation study are used to develop the MS/MSD QC acceptance criteria. Calculate the MS/MSD performance criteria as follows:

1. Calculate the mean and sample standard deviation of the recoveries of each MS/MSD pair, and then compute the overall mean recovery $\bar{X}$, the between-laboratory standard deviation (s_b) of the mean results for each of the nine laboratories, and the pooled within-laboratory standard deviation (s_w) for each target analyte using the MS and MSD analyses.

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

where:
$\bar{X}_j$ = The mean percent recovery for the *jth* laboratory
m = The number of laboratories, and
$\bar{X}$ = The overall mean of the percent recoveries from all laboratories

In order to allow for interlaboratory variability, calculate the combined standard deviation (s_c) for interlaboratory variability and estimation of the mean as:

$$s_c = \sqrt{\left(1 + \frac{1}{m}\right) s_b^2 + \frac{1}{2} s_w^2}$$

where:
m = the number of laboratories.

For nine labs, this becomes:

$$s_c = \sqrt{\left(\frac{10}{9}\right) s_b^2 + \left(\frac{1}{2}\right) s_w^2}$$

2. QC acceptance criteria for MS/MSD recovery - Calculate the QC acceptance criteria for recovery in the MS/MSD test by constructing a ± 2.2s_c window around the mean percent recovery ($\overline{X}$) using the combined standard deviation. This factor comes from a *t* value for an estimated 11 degrees of freedom (based on this experimental design and variance ratios observed in Method 1625):

   $$Lower\ limit\ (\%) = \bar{X} - 2.2 s_c$$
   $$Upper\ limit\ (\%) = \bar{X} + 2.2 s_c$$

   If more than 9 laboratories are used, the degrees of freedom for *t* will increase, but a complete calculation is beyond the scope of this document. An approximation of degrees of freedom equal to the number of laboratories plus 2 will serve for most situations.

   Note: For highly variable methods, it is possible that the lower limit for recovery for both the IPR and OPR analyses will be a negative number. In these instances, the data should either be log-transformed and the recovery window recalculated, or the lower limit established as "detected," as was done with some of the methods in 40 CFR Part 136, Appendix A.

3. QC acceptance criteria for MS/MSD relative percent difference (RPD) - To evaluate a 95% upper confidence limit for precision, the RSD (computed using the pooled within-laboratory standard deviation s_w of the MS/MSD samples, divided by $\overline{X}$, is multiplied by the square root of the 95th percentile *F* value with 1 degrees of freedom in the numerator and m degrees of freedom in the denominator multiplied by the square root of 2 (i. e., $\sqrt{2}$), where *m* is the number laboratories. The resulting multiplier on the RSD for 3 laboratories will then be 3.2. The QC acceptance criterion for precision in the MS/MSD test (RPD_{max}) is calculated as follows:

   $$\mathrm{RPD_{max}} = 3.2\ \mathrm{RSD}$$

Blanks

Establish the QC acceptance criteria for blanks. The historical requirement has been that the concentration of an analyte in a blank must be below the ML or below one-third (1/3) the regulatory compliance level, whichever is higher. However, other limits (including those below the ML) may be used for a specific method. In instances where the level of the blank is close to the regulatory compliance level or the level at which measurements are to be made, it may be necessary to require multiple blank measurements and establish the QC acceptance criteria based on the mean of the blank measurements plus two standard deviations of the blank measurements.

Appendix B

Draft EPA Method 1621

Given the size of the method document, it will be provided as a separate file. The method can also be found at Draft Method 1621 Screening Method for the Determination of Adsorbable Organic Fluorine(AOF) in Aqueous Matrices by Combustion Ion Chromatography (CIC)

Note: Draft EPA Method 1621 in its current form does not represent a final product and the method has been released for public comment. EPA and GDIT anticipate revising the method as a result of the multi-laboratory validation study. Therefore, EPA and GDIT ask that readers and reviewers of this study plan not to distribute the plan without first consulting S. Bekah Burket (EPA/OW). If needed, a version of this study plan can be provided by GDIT. Laboratory participants in the study may be asked to sign a non-disclosure agreement prior to the start of the study.