**UCMR5 PFAS data for OPAL README**

03/05/2025

Contact Lloyd Hill for any questions (hill.lloyd@epa.gov)

**About**

This project folder contains the data and scripts to format EPA’s UCMR-5 PFAS monitoring data for use in the OPAL database. UCMR-5 samples for 29 different PFAS compounds (and lithium) in public drinking water systems. It is an ongoing project with expected completion in 2026. The most recent data release was in October 2024 which includes PFAS data for 7,214 public water systems (PWS). This release is estimated to represent 55% of the total data expected to be collected. More information on EPA’s UCMR program can be found here: <https://www.epa.gov/dwucmr/occurrence-data-unregulated-contaminant-monitoring-rule#5-data-finder>.

This project also adds location data (lat/lon coordinates) for the tested PWS to use in spatial analysis.

**Guide To Folder Contents**

* **data/**
  + **PWSID\_Centroids.csv** = Latitude and longitude of public water system (PWS) service area boundary centroids. Calculated from the EPA’s community water system boundaries shapefile: [https://services.arcgis.com/cJ9YHowT8TU7DUyn/arcgis/rest/services/Water\_System\_Boundaries/FeatureServer]. Also includes size of service area (square meters) and estimate of population served.
  + **UCMRunmatchedPWSIDSwZCTAS.csv** = This file contains PWS location data without shapefile or UCMR5-provided ZIP code information. Information on these PWSs from EPA’s envirofacts SDWIS search tool [https://enviro.epa.gov/envirofacts/sdwis/search]
  + **ZCTA\_2020\_Centroids.csv** = geographic centroids of all US ZIP code tabulated area (ZCTA) boundaries. ZCTA shapefiles representing the 2020 census were downloaded from the census’s TigerLine database [https://www.census.gov/cgi-bin/geo/shapefiles/index.php]
  + **zip\_to\_zcta\_2018.csv** = crosswalk file matching ZIP codes to ZCTAs. Downloaded from HUD [https://www.huduser.gov/portal/datasets/usps\_crosswalk.html]).
  + **ucmr5\_oct2024** = folder of most recent release of UCMR5 data (10/2024). Downloaded from [https://www.epa.gov/dwucmr/occurrence-data-unregulated-contaminant-monitoring-rule#5-data-finder].
* **outputs/**
  + **UCMR5\_PFASSamplingData\_03052025.csv** = final dataset of UCMR5 results formatted for OPAL database. Created in “ucmr5\_data\_processing.R” script.
    - UNIQUE\_ID = unique identifier generated for each UCMR5 sample (PWSID\_CollectionDate\_SampleID)
    - PWSID = public water system identifier
    - LAT = PWSID centroid latitutde (decimal degrees)
    - LON = PWSID centroid longitude (decimal degrees)
    - DATE = sample collection date (YYYY-MM-DD)
    - PFAS\_Concentration\_ngL = concentration of PFAS1 in ng/L. Zeros represent samples below the MRL (ie, non-detects), whereas NA’s represent PFAS compound not tested.
    - PFAS\_MRL\_ngL = Minimum Reporting Limit (MRL) of in ng/L of PFAS compound
    - Full list of PFAS compounds sampled can be found in table at end of document.
  + **UCMR5allPWSIDcoordinates.csv** = dataset of all UCMR-tested PWSIDs and their coordinates. Created in “ucmr5\_pwsid\_coordinates.R” script.
* **scripts/**
  + **ucmr5\_data\_processing.R** = this script imports UCMR5 PFAS sampling results and processes the data to match OPAL formatting. This script also adds location data (lat/lon) for all sampled PWSs.
  + **ucmr5\_pwsid\_coordinates.R** = this script compiles PWS location data from multiple data sources to create a data frame of all UCMR5-tested PWS and their centroid coordinates (lat/lon).

**UCMR5 tested PFAS**

|  |  |  |
| --- | --- | --- |
| **Label in Database** | **UCMR5 Abbreviation (Full PFAS Name)** | **CASRN** |
| X11ClPF3OUdS | 11Cl-PF3OUdS (11-chloroeicosafluoro-3-oxaundecane-1-sulfonic acid) | 763051-92-9 |
| X42FTS | 4:2 FTS (1H, 1H, 2H, 2H-perfluorohexane sulfonic acid) | 757124-72-4 |
| X62FTS | 6:2 FTS (1H, 1H, 2H, 2H-perfluorooctane sulfonic acid) | 27619-97-2 |
| X82FTS | 8:2 FTS (1H, 1H, 2H, 2H-perfluorodecane sulfonic acid) | 39108-34-4 |
| X9ClPF3ONS | 9Cl-PF3ONS (9-chlorohexadecafluoro-3-oxanone-1-sulfonic acid) | 756426-58-1 |
| ADONA | ADONA (4,8-dioxa-3H-perfluorononanoic acid) | 919005-14-4 |
| HFPODA | HFPO-DA (hexafluoropropylene oxide dimer acid) | 13252-13-6 |
| NEtFOSAA | NEtFOSAA (n-ethyl perfluorooctanesulfonamidoacetic acid) | 2991-50-6 |
| NFDHA | NFDHA (nonafluoro-3,6-dioxaheptanoic acid) | 151772-58-6 |
| NMeFOSAA | NMeFOSAA (n-methyl perfluorooctanesulfonamidoacetic acid) | 2355-31-9 |
| PFBA | PFBA (perfluorobutanoic acid) | 375-22-4 |
| PFBS | PFBS (perfluorobutanesulfonic acid) | 375-73-5 |
| PFDA | PFDA (perfluorodecanoic acid) | 335-76-2 |
| PFDoA | PFDoA (perfluorododecanoic acid) | 307-55-1 |
| PFEESA | PFEESA (perfluoro (2-ethoxyethane) sulfonic acid) | 113507-82-7 |
| PFHpA | PFHpA (perfluoroheptanoic acid) | 375-85-9 |
| PFHpS | PFHpS (perfluoroheptanesulfonic acid) | 375-92-8 |
| PFHxA | PFHxA (perfluorohexanoic acid) | 307-24-4 |
| PFHxS | PFHxS (perfluorohexanesulfonic acid) | 355-46-4 |
| PFMBA | PFMBA (perfluoro-4-methoxybutanoic acid) | 863090-89-5 |
| PFMPA | PFMPA (perfluoro-3-methoxypropanoic acid) | 377-73-1 |
| PFNA | PFNA (perfluorononanoic acid) | 375-95-1 |
| PFOA | PFOA (perfluorooctanoic acid) | 335-67-1 |
| PFOS | PFOS (perfluorooctanesulfonic acid) | 1763-23-1 |
| PFPeA | PFPeA (perfluoropentanoic acid) | 2706-90-3 |
| PFPeS | PFPeS (perfluoropentanesulfonic acid) | 2706-91-4 |
| PFTA | PFTA (perfluorotetradecanoic acid) | 376-06-7 |
| PFTrDA | PFTrDA (perfluorotridecanoic acid) | 72629-94-8 |
| PFUnA | PFUnA (perfluoroundecanoic acid) | 2058-94-8 |