User Guide to Downloading and Using SYR3 Data from EPA's Website

To support the national contaminant occurrence and exposure assessments performed under the Six-Year Review process, EPA analyzes compliance monitoring data from public water systems (PWSs) for regulated drinking water contaminants. This analysis allows EPA to characterize the frequency of occurrence, the levels found, and the geographic distribution of contaminants to help the Agency determine if there may be a meaningful opportunity to improve public health protection. EPA conducted a voluntary data request from the states and primacy agencies to obtain the compliance monitoring data necessary to analyze national contaminant occurrence in support of the third Six-Year Review (SYR3). This data request was conducted through the Information Collection Request (ICR) process. EPA requested that states and primacy agencies submit their SDWA compliance monitoring data collected between January 2006 and December 2011. For more information on the process undertaken to request the voluntary submission of compliance monitoring data by the states, see the third Six-Year Review ICR renewal (75 FR 6023, USEPA, 2010).

Through extensive data management efforts, quality assurance evaluations, and communications and consultations with state data management staff, EPA established a single contaminant occurrence dataset that consists of compliance monitoring data from 54 out of 67 states/primacy agencies (46 states plus Washington, D.C. and the tribal data). This dataset is referred to as the National Compliance Monitoring ICR Dataset for the third Six-Year Review (or "SYR3 ICR Dataset"). The 54 states/primacy agencies that provided data for the SYR3 ICR Dataset comprise 95 percent of all PWSs and 92 percent of the total population served by PWSs nationally, and are geographically representative of PWSs nationwide. The SYR3 ICR Dataset was used to estimate a variety of occurrence measures to characterize the national occurrence of regulated contaminants in public water systems to support the Six-Year Review process.

The SYR3 ICR Dataset is the largest, most comprehensive set of drinking water compliance monitoring data ever compiled and analyzed by EPA to inform decision making. EPA conducted a quality control evaluation of these data submitted by states and other primacy agencies, and assembled these data into a database. The database is more than twice the size of the one collected to support of the Second Six-Year Review (SYR2) with more than 47 million records from approximately 167,000 public water systems, serving approximately 290 million people nationally. The dataset includes the results of all compliance monitoring data (all sample analytical detections and non-detections) from January 2006 to December 2011 for regulated chemical phase contaminants, radionuclides, disinfectants and disinfection byproducts (D/DBPs), DBP precursors, microbial contaminants, disinfectant residuals and treatment information. Note that only the data that passed the QA/QC process are posted online.

Additional reference material is available to assist with the assessment of the SYR3 data.
- EPA's Six-Year Review website
- The Data Management and Quality Assurance/Quality Control Process for the Third Six-Year Review Information Collection Rule Dataset (USEPA, 2016a)

The data are posted online in several zip files. Each zip file includes text files for multiple contaminants/parameters. The number of records and contaminants/parameters included in each file vary. The remainder of this document is organized as follows:

- Section 1 describes the data being posted for phase chemicals, radionuclides and disinfection byproducts.
- Section 2 describes the data being posted for disinfection byproduct precursors.

- Section 3 describes the data being posted for microbial contaminants and associated disinfectant residuals.
- Section 4 describes data being posted for additional parameters.
- Section 5 describes the treatment data being posted.
- Section 6 describes the data quality considerations of the SYR3 ICR data.
- Section 7 describes supplemental data sources being posted.

User Guide to Downloading and Using SYR3 Data from EPA's Website

Section 1: Phase Chemicals, Radionuclides and Disinfection Byproducts

Exhibit 1 contains a list of the data elements, column names and a brief description of the data for each data element included in each of the SYR3 ICR text files for the individual phase chemicals, radionuclides and disinfection byproducts.

Exhibit 1: Six-Year 3 Data Field Names and Definitions

Data Element	Column Name	Description
Contaminant Identification Code	Analyte ID	4-digit Safe Drinking Water Information System (SDWIS) contaminant identification number for which the sample is being analyzed.
Contaminant Name	Analyte Name	Common name of contaminant for which the sample is being analyzed.
State Code	State Code	2- digit state code. Note that the state code "IM" refers to non-community water system data from the State of Illinois.
Public Water System Identification Number (PWSID)	PWSID	The code used to identify each PWS. The code begins with the standard 2-character postal state abbreviation or region code; the remaining 7 numbers are unique to each PWS in the state.
System Name	System Name	Name of the PWS.
Federal Public Water System Type Code	System Type	A code to identify whether a system is: • Community Water System (C); • Non-Transient Non-Community Water System (NTNC); or • Transient Non-Community Water System (NC).
Retail Population-served	Retail Population Served	Retail population served by a system.
Adjusted Total Population-served	Adjusted Total Population Served	Total population served by a system, adjusted to reduce double-counting of population served by purchasing water systems.
Source Water Type	Source Water Type	Type of water at the source. Source water type can be: • Ground water (GW); • Surface water (SW); • Purchased Surface Water (SWP); • Purchased Ground Water (GWP); • Ground Water Under Direct Influence of Surface Water (GU); or • Purchased Ground Water Under Direct Influence of Surface Water (GUP).
Facility Identification Code	Water Facility ID	A unique identifier for each water system facility.
Water Facility Type	Water Facility Type	Type of water system facility: • CC = Consecutive Connection; • CH = Common Headers; • CW = Clear Well; • DS = Distribution System; • IG = Infiltration Gallery; • IN = Intake; • OT = Other; • PC = Pressure Control; • PF = Pumping Facility; • RS = Reservoir; • SI = Surface Impoundment; • SP = Spring; • SS = Sampling Station; • ST = Storage; • TM = Transmission Main (Manifold); • TP = Treatment Plant; • WH = Well Head; • WL = Well; or

Data Element	Column Name	Description
		• XX = unknown.
Sampling Point Identification Code	Sampling Point ID	A unique identifier for each sampling point location.
Sampling Point Type	Sampling Point Type	Location type of a sampling point: • DS = Distribution System; • EP = Entry point; • FC = First Customer; • FN = Finished Water Source; • LD = Lowest Disinfectant Residual; • MD = Midpoint in the Distribution System; • MR = Point of Maximum Residence; • PC = Process Control; • RW = Raw Water Source; • SR = Source Water Point; • UP = Unit Process; or • WS = Water System Facility Point.
Source Type Code	Source Type Code	Type of water source, based on whether treatment has taken place. Source type can be: • Finished (FN); • Raw (RW); or • Unknown (null or X).
Sample Type Code	Sample Type Code	Type of sample: • CO = Confirmation; • MR = Maximum Residence Time; • RP = Repeat; or • RT = Routine.
Laboratory Assigned Identification Number	Laboratory Assigned ID	Unique lab identification, used to link up the total coliform positive (TC+) and *E. coli* / fecal coliform samples.
Six Year ID	Six Year ID	Unique identifier for each analytical result.
Sample Identification Number	Sample ID	Identifier assigned by state or the laboratory that uniquely identifies a sample.
Sample Collection Date	Sample Collection Date	Date the sample was collected, including month, day, and year.
Detection Limit Value	Detection Limit Value	Limit below which the specific lab indicated they could not reliably measure results for a contaminant with the methods and procedures used by the lab.
Detection Limit Unit	Detection Limit Unit	Units of the detection limit value.
Detection Limit Code	Detection Limit Code	Indicates the type of Detection Limit reported in the Detection Limit Value column (e.g., the Minimum Reporting Level, Laboratory Reporting Level, etc.)
Sample Analytical Result - Sign	Detect	The sign indicates whether the sample analytical result was: • (0) "less than" means the contaminant was not detected or was detected at a level "less than" the MRL. • (1) "equal to" means the contaminant was detected at a level "equal to" the value reported in "Sample Analytical Result - Value."
Sample Analytical Result - Value	Value	For detections, this field is equal to the actual numeric (decimal) value of the analysis for the chemical result; for non-detections, this field is blank.
Sample Analytical Result - Unit of Measure	Unit	Unit of measurement for the analytical results reported (usually expressed in either µg/L or mg/L for chemicals; or pCi/L for radionuclides).

User Guide to Downloading and Using SYR3 Data from EPA's Website

Data Element	Column Name	Description
Presence Indicator Code	Presence Indicator Code	Indication of whether results of an analysis were positive or negative for TC, EC and FC. • P = Presence • A = Absence.
Residual Field Free Chlorine	Residual Field Free Chlorine mg/L	Amount of free chlorine residual (in mg/L) found in the water after disinfectant has been applied. These concentrations were measured in the field at the same time and location as coliform samples (TC-EC-FC samples).
Residual Field Total Chlorine	Residual Field Total Chlorine mg/L	Amount of total chlorine residual (in mg/L) found in the water after disinfectant has been applied. These concentrations were measured in the field at the same time and location as coliform samples (TC-EC-FC samples).

Summary of SYR3 Phase Chemicals, Radionuclides and Disinfection Byproduct Data

Exhibit 2 provides a count of states, total number of sample records and systems for each phase chemical, radionuclide and disinfection byproduct whose data is posted online. The user may want to compare their counts of records downloaded for each contaminant of interest to this table to ensure that all of the records were correctly downloaded and imported. Note that these record counts reflect the data after the QA/QC process.

Exhibit 2: Six-Year Review 3 Data Summary for Contaminants/Parameters

Contaminant	Analyte ID	Number of States with Data	Total Number of Sample Records	Total Number of Systems	Zip Filename
Phase Chemicals					
1,1,1-Trichloroethane	2981	50	374,181	55,735	SYR3_PhaseChem_1
1,1,2-Trichloroethane	2985	50	371,877	55,733	SYR3_PhaseChem_1
1,1-Dichloroethylene	2977	50	379,522	55,728	SYR3_PhaseChem_1
1,2,4-Trichlorobenzene	2378	50	369,032	55,725	SYR3_PhaseChem_1
1,2-Dibromo-3-chloropropane (DBCP)	2931	50	188,597	37,226	SYR3_PhaseChem_1
2,3,7,8-TCDD (Dioxin)	2063	30	20,244	3,216	SYR3_PhaseChem_1
2,4,5-Trichlorophenoxypropionic Acid (Silvex)	2110	50	126,887	36,897	SYR3_PhaseChem_1
2,4-Dichlorophenoxyacetic acid (2,4-D)	2105	50	131,047	37,690	SYR3_PhaseChem_1
Alachlor	2051	50	153,083	42,955	SYR3_PhaseChem_1
Antimony	1074	49	164,961	50,532	SYR3_PhaseChem_1
Arsenic	1005	50	297,354	54,845	SYR3_PhaseChem_1
Asbestos	1094	39	12,084	5,785	SYR3_PhaseChem_1
Atrazine	2050	50	162,134	44,310	SYR3_PhaseChem_1
Barium	1010	49	165,387	50,711	SYR3_PhaseChem_2
Benzo(a)pyrene	2306	50	131,437	34,341	SYR3_PhaseChem_2
Beryllium	1075	49	164,392	50,195	SYR3_PhaseChem_2
Cadmium	1015	49	165,247	50,583	SYR3_PhaseChem_2
Carbofuran	2046	50	122,110	34,614	SYR3_PhaseChem_2
Chlordane	2959	49	128,870	35,685	SYR3_PhaseChem_2
Chromium (Total)	1020	49	167,251	50,597	SYR3_PhaseChem_2
cis-1,2-Dichloroethylene	2380	50	376,300	55,734	SYR3_PhaseChem_2

User Guide to Downloading and Using SYR3 Data from EPA's Website

Contaminant	Analyte ID	Number of States with Data	Total Number of Sample Records	Total Number of Systems	Zip Filename
Cyanide	1024	49	119,659	36,907	SYR3_PhaseChem_2
Dalapon	2031	49	146,702	36,005	SYR3_PhaseChem_2
Di(2-ethylhexyl)adipate (DEHA)	2035	50	133,169	34,628	SYR3_PhaseChem_2
Di(2-ethylhexyl)phthalate (DEHP)	2039	49	133,523	33,923	SYR3_PhaseChem_2
Dinoseb	2041	50	126,014	36,701	SYR3_PhaseChem_2
Diquat	2032	46	69,829	17,906	SYR3_PhaseChem_2
Endothall	2033	45	61,972	15,538	SYR3_PhaseChem_3
Endrin	2005	50	136,623	38,453	SYR3_PhaseChem_3
Ethylbenzene	2992	50	372,709	55,754	SYR3_PhaseChem_3
Ethylene Dibromide (EDB)	2946	49	184,784	37,499	SYR3_PhaseChem_3
Fluoride	1025	49	256,237	47,227	SYR3_PhaseChem_3
Glyphosate	2034	45	70,016	18,502	SYR3_PhaseChem_3
Heptachlor	2065	50	137,286	38,691	SYR3_PhaseChem_3
Heptachlor Epoxide	2067	50	137,081	38,625	SYR3_PhaseChem_3
Hexachlorobenzene	2274	50	137,816	38,498	SYR3_PhaseChem_3
Hexachlorocyclopentadiene	2042	50	140,004	38,743	SYR3_PhaseChem_3
Lindane (gamma-Hexachlorocyclohexane)	2010	50	139,076	39,260	SYR3_PhaseChem_3
Mercury (Inorganic)	1035	49	164,558	50,552	SYR3_PhaseChem_3
Methoxychlor	2015	50	139,744	39,187	SYR3_PhaseChem_3
Monochlorobenzene	2989	50	371,311	55,676	SYR3_PhaseChem_3
Nitrate (as N)	1040	49	1,157,522	132,176	SYR3_PhaseChem_3
Nitrite (as N)	1041	49	445,544	85,742	SYR3_PhaseChem_3
o-Dichlorobenzene	2968	50	370,929	55,732	SYR3_PhaseChem_4
Oxamyl (Vydate)	2036	50	121,508	34,518	SYR3_PhaseChem_4
p-Dichlorobenzene	2969	50	371,276	55,739	SYR3_PhaseChem_4
Pentachlorophenol	2326	50	140,486	40,322	SYR3_PhaseChem_4
Picloram	2040	50	128,401	37,445	SYR3_PhaseChem_4
Polychlorinated biphenyls (PCBs)	2383	44	86,405	21,571	SYR3_PhaseChem_4
Selenium	1045	49	165,672	50,568	SYR3_PhaseChem_4
Simazine	2037	50	156,862	43,240	SYR3_PhaseChem_4
Styrene	2996	50	370,368	55,731	SYR3_PhaseChem_4
Thallium	1085	49	164,156	50,522	SYR3_PhaseChem_4
Toluene	2991	50	373,021	55,748	SYR3_PhaseChem_4
Toxaphene	2020	49	127,187	37,043	SYR3_PhaseChem_4
trans-1,2-Dichloroethylene	2979	50	371,580	55,633	SYR3_PhaseChem_4
Xylenes (Total)	2955	50	323,477	51,074	SYR3_PhaseChem_4
Radionuclides					
Alpha Particles	4000	47	60,803	13,309	SYR3_Rads
Beta Particles	4100	41	43,278	11,531	SYR3_Rads
Combined Radium-226 & -228	4010	42	73,018	15,805	SYR3_Rads

User Guide to Downloading and Using SYR3 Data from EPA's Website

Contaminant	Analyte ID	Number of States with Data	Total Number of Sample Records	Total Number of Systems	Zip Filename
Uranium	4006	49	86,208	12,155	SYR3_Rads
Disinfection Byproducts					
Total Trihalomethanes	2950	46	532,002	36,691	SYR3_THM
Bromoform	2942	42	433,636	34,788	SYR3_THM
Chloroform	2941	42	434,624	34,839	SYR3_THM
Bromodichloromethane	2943	42	433,663	34,815	SYR3_THM
Dibromochloromethane	2944	42	433,141	34,735	SYR3_THM
Total Haloacetic Acids	2456	45	475,592	33,518	SYR3_HAA
Monochloroacetic acid	2450	36	283,260	25,202	SYR3_HAA
Dichloroacetic acid	2451	36	282,778	25,221	SYR3_HAA
Trichloroacetic acid	2452	36	282,732	25,213	SYR3_HAA
Monobromoacetic acid	2453	36	282,799	25,196	SYR3_HAA
Dibromoacetic acid	2454	36	282,986	25,210	SYR3_HAA
Bromate	1011	29	8,884	222	SYR3_Bromate_Chlorite
Chlorite	1009	28	25,989	220	SYR3_Bromate_Chlorite

User Guide to Downloading and Using SYR3 Data from EPA's Website

Section 2: Disinfection Byproduct Precursors

Data for three disinfection byproduct precursors are being posted online: total organic carbon (TOC), alkalinity and pH. In addition to the "full" datasets for TOC and alkalinity, a "paired" TOC dataset was created that included, for each treatment plant, the average monthly concentrations of TOC and alkalinity in source (raw) water paired with the corresponding average finished water concentration of TOC. The "paired" TOC dataset was used to evaluate the percent removal of TOC using the SYR3 data; see Chapter 7 and Appendix C in USEPA (2016d) for more details on the "paired" TOC dataset.

Exhibit 3 contains the list of data elements, column names, and a brief description of the data for each data element included in the "paired" TOC dataset. For a list of data elements included in the "full" TOC, alkalinity and pH datasets, refer to Exhibit 1.

Exhibit 3: SYR3 "Paired" TOC Dataset Field Names and Definitions

Data Element	Column Name	Description
Public Water System Identification Number (PWSID)	PWSID	The code used to identify each PWS. The code begins with the standard 2-character postal state abbreviation or region code; the remaining 7 numbers are unique to each PWS in the state.
Sample Collection Date (Month)	Month	Month (1 through 12).
Sample Collection Date (Year)	Year	Year (2006 through 2011).
Retail Population-served	Retail Population Served	Retail population served by the water system.
Federal Public Water System Type Code	System Type	Water system type according to federal requirements. C = Community water system NTNC = Non-transient non-community water system
Source Water Type	Source Water Type	Primary water source for the water system. GU = Ground water Under Direct Influence of Surface Water GW = Ground Water GWP = Purchased Ground Water SW = Surface Water SWP = Purchased Surface Water
Facility Identification Code	Water Facility ID	Unique identifier for each water system facility.
State Facility Identification Code	State Facility ID	Identifier for each water system facility that is unique within a particular state.
State Assigned Identification Code	State Assigned ID Code	A state-assigned value which identifies the water system facility.
Raw water TOC average concentration	Avg Of Raw TOC (mg/L)	Monthly average (in mg/L) total organic carbon (TOC) concentration in raw water.
Raw water alkalinity average concentration	Avg Of Raw Alkalinity (mg/L)	Monthly average (in mg/L) alkalinity concentration in raw water.
Finished water TOC average concentration	Avg Of Finished TOC (mg/L)	Monthly average (in mg/L) total organic carbon (TOC) concentration in finished water.

User Guide to Downloading and Using SYR3 Data from EPA's Website

Summary of SYR3 Disinfection Byproduct Precursor Data

Exhibit 4 provides a count of states, total number of sample records and systems for TOC, alkalinity and pH.

Exhibit 4: Six-Year Review 3 Data Summary for TOC, Alkalinity and pH

Contaminant	Analyte ID	Number of States with Data	Total Number of Sample Records	Total Number of Systems	Zip Filename
Disinfection Byproduct Precursors - Full Datasets					
Total Organic Carbon	2920	32	232,567	2,836	SYR3_Precursors
Alkalinity	1927	38	201,682	15,059	SYR3_Precursors
pH	1925	40	208,203	25,509	SYR3_Precursors
Disinfection Byproduct Precursors - Reduced Dataset					
Paired TOC-alkalinity dataset[1]	N/A	22	65,771	1,208	SYR3_PairedTOC-Alkalinity

[1] The "paired" TOC-alkalinity dataset includes average monthly concentrations of TOC and alkalinity in source (raw) water paired with the corresponding average finished water concentrations of TOC.

User Guide to Downloading and Using SYR3 Data from EPA's Website

Section 3: Microbials and Associated Disinfectant Residuals

Summary of SYR3 Microbial and Residual Data

Data for three microbial contaminants (total coliforms, *E. coli*, and fecal coliform) and associated disinfectant residual data are being posted online. A "full" dataset includes all data for total coliforms (TC), *E. coli* (EC), and fecal coliform (FC) and associated disinfectant residual data (when available) that have passed the initial QA process. A "reduced" dataset includes a subset of the data for disinfecting systems with disinfectant residual. These data were used to support the analyses in USEPA (2016c). Only the data with paired chlorine residual concentrations (free and/or total chlorine) were included in the analysis; thus, these TC-EC-FC data represent only a subset of all total coliform results submitted via the SYR3 ICR. See Appendix A in USEPA (2016c) for details on the QA/QC documentation for both the full and the reduced microbial datasets.

For a list of data elements included in the full TC, EC, and FC datasets, refer to Exhibit 1. For a list of data elements included in the Reduced Dataset for Analysis of Disinfecting Systems with Disinfectant Residuals, refer to Exhibit 5.

Exhibit 5: SYR3 Reduced Dataset for Analysis of Disinfecting Systems with Disinfectant Residuals - Field Names and Definitions

Data Element	Column Name	Description
Contaminant Identification Code	Analyte ID	4-digit Safe Drinking Water Information System (SDWIS) contaminant identification number for which the sample is being analyzed.
Contaminant Name	Analyte Name	Common name of contaminant for which the sample is being analyzed.
State Code	State Code	2- digit state code. Note that the state code "IM" refers to non-community water system data from the State of Illinois.
Public Water System Identification Number (PWSID)	PWSID	The code used to identify each PWS. The code begins with the standard 2-character postal state abbreviation or region code; the remaining 7 numbers are unique to each PWS in the state.
System Name	System Name	Name of the PWS.
Federal Public Water System Type Code	System Type	A code to identify whether a system is: • Community Water System (C); • Non-Transient Non-Community Water System (NTNC); or • Transient Non-Community Water System (NC).
Retail Population-served	Retail Population Served	Retail population served by a system.
Source Water Type	Source Water Type	Type of water at the source. Source water type can be: • Ground water (GW); • Surface water (SW); • Purchased Surface Water (SWP); • Purchased Ground Water (GWP); • Ground Water Under Direct Influence of Surface Water (GU); or • Purchased Ground Water Under Direct Influence of Surface Water (GUP).
Facility Identification Code	Water Facility ID	A unique identifier for each water system facility.
Water Facility Type	Water Facility Type	Type of water system facility: DS = Distribution System.
Sampling Point Identification Code	Sampling Point ID	A unique identifier for each sampling point location.
Sampling Point Type	Sampling Point Type	Location type of a sampling point: • DS = Distribution System; • EP = Entry point;

Data Element	Column Name	Description
		• FC = First Customer; • FN = Finished Water Source; • MD = Midpoint in the Distribution System; • MR = Point of Maximum Residence; • RW = Raw Water Source; • SR = Source Water Point; or • WS = Water System Facility Point.
Source Type Code	Source Type Code	Type of water source, based on whether treatment has taken place. Source type can be: • Finished (FN); • Raw (RW); or • Unknown (null or X).
Sample Type Code	Sample Type Code	Type of sample: • RP = Repeat; or • RT = Routine.
Six Year ID	Six Year ID	Unique identifier for each analytical result.
Sample Collection Date	Sample Collection Date	Date the sample was collected, including month, day, and year.
Presence Indicator Code	Presence Indicator Code	Indication of whether results of an analysis were positive or negative for TC, EC and FC. • P = Presence • A = Absence.
Residual Field Free Chlorine	Residual Field Free Chlorine mg/L	Amount of free chlorine residual (in mg/L) found in the water after disinfectant has been applied. These concentrations were measured in the field at the same time and location as coliform samples (TC-EC-FC samples).
Residual Field Total Chlorine	Residual Field Total Chlorine mg/L	Amount of total chlorine residual (in mg/L) found in the water after disinfectant has been applied. These concentrations were measured in the field at the same time and location as coliform samples (TC-EC-FC samples).

Exhibit 6 provides a count of states, total number of sample records and systems for total coliform, *E. coli*, fecal coliform, and their associated free and total chlorine residual concentrations for both the full and reduced datasets.

Exhibit 6: Six-Year Review 3 Data Summary for Microbials and Associated Disinfectant Residuals

Contaminant	Analyte ID	Number of States with Data	Total Number of Sample Records	Total Number of Systems	Zip Filename
Microbials and Residuals – Full Datasets					
Total coliform	3100	46	9,766,686	113,548	SYR3_TC-DR-06-08; SYR3_TC-DR-09-11
E. coli	3014	44	1,804,329	55,509	SYR3_EC-FC-DR
Fecal coliform	3013	39	264,090	17,821	SYR3_EC-FC-DR
Microbials and Residuals - Reduced Dataset					
Total coliform	3100	41	4,750,432	36,753	SYR3_Microbes_DR

Contaminant	Analyte ID	Number of States with Data	Total Number of Sample Records	Total Number of Systems	Zip Filename
E. coli	3014	35	889,570	18,896	SYR3_Microbes_DR
Fecal coliform	3013	25	64,304	2,986	SYR3_Microbes_DR
Field free chlorine residual[1]	N/A	--	4,007,235	33,054	SYR3_Microbes_DR
Field total chlorine residual[1]	N/A	--	2,521,771	17,757	SYR3_Microbes_DR

[1] Measured in the field at the same time and location as coliform samples were collected.

Summary of Reduced Dataset for Analysis of Undisinfected Ground Water Systems

Data for total coliforms, *E. coli*, and fecal coliform paired with system disinfection status are also posted online. To simplify statistical modeling of the TC, EC, and FC data for that analysis, the data for each system and month were reduced to a small number of summary counts: (a) the total number of routine samples assayed, (b) the number of routine samples testing positive for TC, (c) the total number of TC positive routine samples tested for EC and (d) the number of routine samples testing positive for EC. Rather than include a record for each sample assayed, the reduced dataset includes, for each water system and month, counts of the routine and repeat samples for TC, EC and FC. (See Exhibit 7.) In the final "reduced" dataset, there are data for a total of 80,692 water systems from 39 states/entities. (The zip file containing these data is "SYR3_Microbes_GW.") See Appendix D in USEPA (2016c) for details on the steps used to produce this reduced dataset.

A subset of these data were used to represent "undisinfected" ground water systems. In this analysis, "undisinfected" ground water systems referred to those that do not practice disinfection or have very low disinfectant residuals (i.e., less than 0.1 mg/L). These data were used to support additional analyses in USEPA (2016c). See Appendix F in USEPA (2016c) for details on the analysis of undisinfected ground water systems.

Exhibit 7: SYR3 Reduced Dataset for Analysis of Undisinfected Ground Water Systems - Field Names and Definitions

Data Element	Column Name	Description
Public Water System Identification Number (PWSID)	PWSID	Public water system identification number (PWSID).
Sample Collection Date (Month)	Month	Month (1 through 12).
Sample Collection Date (Year)	Year	Year (2006 through 2011).
Retail Population-served	Retail Population Served	Retail population served by the water system.
Federal Public Water System Type Code	System Type	Water system type according to federal requirements. C = Community water system NTNC = Non-transient non-community water system
Source Water Type	Source Water Type (GW-SW)	Primary water source for the water system. GW = Ground Water (also includes Purchased GW)

Data Element	Column Name	Description
		SW = Surface Water (also includes Purchased SW; Ground water Under Direct Influence of SW; and Purchased Ground Water Under Direct Influence of SW)
Disinfection Status	Disinfecting?	An indication if the system disinfects its water (Y = Yes; blank = No). All systems with a source water type = "SW" were assumed to be disinfecting. Note: An explanation of the determination of the ground water systems' disinfection status is included on pages 2 and 3 of this document.
Count Routine TC samples	TC Samples (routine)	The count of routine total coliform (TC) samples.
Count Routine TC+ samples	TC+ Samples (routine)	The count of routine TC positive samples.
Count Routine EC samples	EC Samples (routine)	The count of routine *E. coli* (EC) samples.
Count Routine EC+ samples	EC+ Samples (routine)	The count of routine EC positive samples.
Count Routine FC samples	FC Samples (routine)	The count of routine fecal coliform (FC) samples.
Count Routine FC+ samples	FC+ Samples (routine)	The count of routine FC positive samples.
Count Repeat TC samples	TC Samples (repeat)	The count of repeat TC samples.
Count Repeat TC+ samples	TC+ Samples (repeat)	The count of repeat TC positive samples.
Count Repeat EC samples	EC Samples (repeat)	The count of repeat EC samples.
Count Repeat EC+ samples	EC+ Samples (repeat)	The count of repeat EC positive samples.
Count Repeat FC samples	FC Samples (repeat)	The count of repeat FC samples.
Count Repeat FC+ samples	FC+ Samples (repeat)	The count of repeat FC positive samples.

User Guide to Downloading and Using SYR3 Data from EPA's Website

Section 4: Additional Parameters

Data for 11 additional parameters have been provided; however, these parameters did not undergo the same quality assurance evaluations as the parameters that were analyzed as part of the SYR3 process. For more information on the quality assurance evaluations performed for these parameters, see USEPA (2016a). Exhibit 8 provides a count of states, total number of sample records and systems for the additional parameters whose data are being posted online. For a list of data elements included in the data posted online for these additional parameters, refer to Exhibit 1.

Exhibit 8: Six-Year Review 3 Data Summary for Additional Parameters

Parameter	Analyte ID	Number of States with Data	Total Number of Sample Records	Total Number of Systems	Zip Filename
Additional Parameters[1]					
Heterotrophic bacteria	3001	18	48,908	797	SYR3_AdditionalAnalytes
Enterococci	3002	2	9	3	SYR3_AdditionalAnalytes
Giardia lamblia	3008	5	426	42	SYR3_AdditionalAnalytes
Chlorine[2]	0999	11	1,505,286	3,673	SYR3_AdditionalAnalytes
Chloramine[2]	1006	5	58,012	474	SYR3_AdditionalAnalytes
Chlorine dioxide	1008	10	7,181	22	SYR3_AdditionalAnalytes
Residual chlorine[2]	1012	3	70,582	1,081	SYR3_AdditionalAnalytes
Free residual chlorine data[2]	1013	1	5,852	741	SYR3_AdditionalAnalytes
SUVA	2923	2	2,447	34	SYR3_AdditionalAnalytes
UV-254	2922	2	2,010	31	SYR3_AdditionalAnalytes
DOC	2919	4	16,669	163	SYR3_AdditionalAnalytes

[1] Coliphage was requested in the SYR3 ICR, however, no coliphage records passed the quality assurance evaluation.
[2] Reported independently of the coliform sample results.

Section 5: Treatment Data

Exhibits 9 and 10 provide a comprehensive summary of the data elements included in the treatment information within the SYR3 ICR database. EPA has posted these data online; however, it is important to note that the treatment information did not undergo the same quality assurance evaluations as the analytes that were analyzed as part of the SYR3 process.

Exhibit 9 identifies the data elements used in the treatment information tables and a description of each data element. However, a majority of these data elements are not populated. Exhibit 10 represents the database relationships between tables in the SYR3 ICR treatment database. This diagram shows how the treatment tables relate to one another. Bolded field names are primary keys, or unique fields, designated to identify all table records. Primary keys contain a unique number for each row of data. Italicized field names are foreign keys that serve as the link (connection) between two or more related tables. Relationships between key fields in different tables are illustrated by the lines connecting the tables.

Exhibit 9: Treatment Data Dictionary (Filename: SYR3_Treatment)

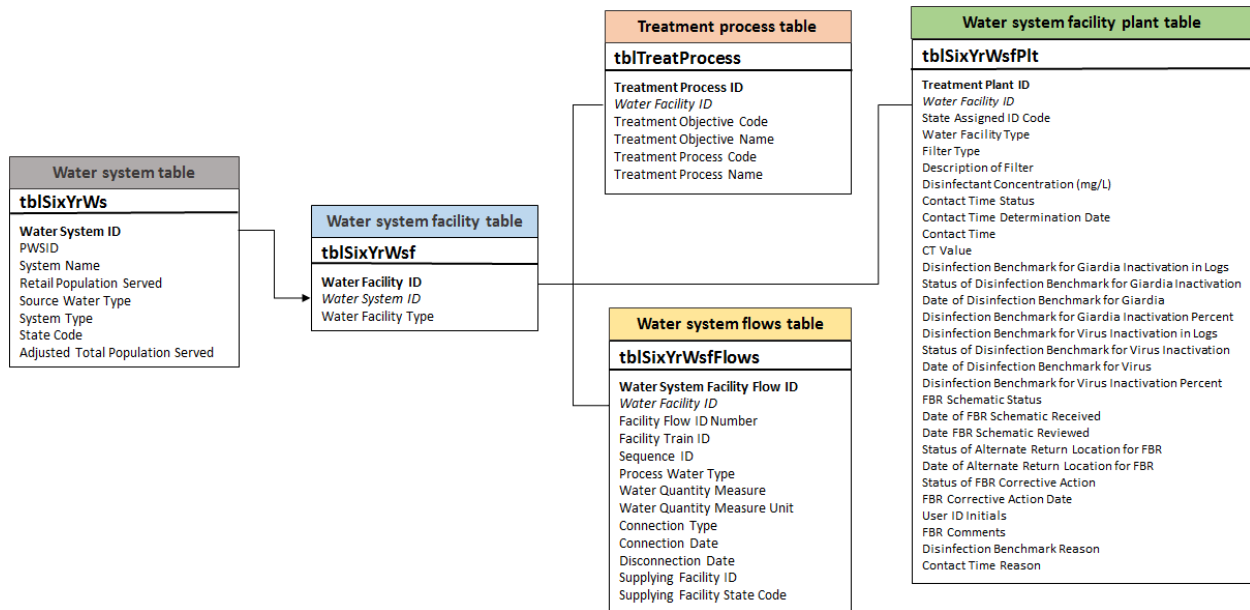| Data Element	Description
Water system facility plant table (tblSixYrWsfPlt)	
Treatment Plant ID	Unique identifier for each treatment plant water system facility record.
Water Facility ID	Identifier that relates each record to the unique record in the tblSixYrWsf table.
State Assigned ID Code	A state-assigned value which identifies the treatment plant water system facility.
Water Facility Type	The value extracted from SDWIS/State will be "TP" (treatment plant). The values from non SDWIS states include "TM" (transmission manifold) and "ST" (storage).
Filter Type	Unfiltered (UF), Conventional Filtration (CF), Direct Filtration (DF), Diatomaceous Earth (DE), Other (OT), and other permitted values that the System Administrator may add.
Description of Filter	A description of the filter.
Disinfectant Concentration (mg/L)	Disinfectant Concentration in mg/L.
Contact Time Status	Contact Time Status. Permitted values are: RQD – Required; NRQD - Not Required; REQT – Requested; RECV – Received; URVW - Under Review; RVWD – Reviewed; APVD – Approved; DTMD – Determined; DENY – Denied; RESB – Resubmitted.
Contact Time Determination Date	Date the Contact Time was determined
Contact Time	Contact Time in minutes--the number of minutes the water was in contact with the disinfectant in order to be properly disinfected. The range of values is 0001 to 2400.
CT Value	CT value in mg x min/liter.
Disinfection Benchmark for Giardia Inactivation in Logs	The disinfection profile benchmark for Giardia inactivation in Logs.
Status of Disinfection Benchmark for Giardia Inactivation	The status of the disinfection profile benchmark for Giardia inactivation. See CONTACT_TIME_STAT for permitted values and description.
Date of Disinfection Benchmark for Giardia	The date the disinfection virus benchmark was determined.
Disinfection Benchmark for Giardia Inactivation Percent	The disinfection profile benchmark for Giardia inactivation percent.
Disinfection Benchmark for Virus Inactivation in Logs	The disinfection profile benchmark for virus inactivation in Logs.
Status of Disinfection Benchmark for Virus Inactivation	The status of the disinfection profile benchmark for Virus inactivation. See CONTACT_TIME_STAT for permitted values and description.
Date of Disinfection Benchmark for Virus	The date the disinfection virus benchmark was determined.

Data Element	Description
Disinfection Benchmark for Virus Inactivation Percent	The disinfection profile benchmark for virus inactivation percent.
FBR Schematic Status	Under the Filter Backwash Rule, a water system is required to submit a schematic of this treatment plant to the primacy agency for review to demonstrate the percentage of filter backwash that is returned to the treatment plant influent. See CONTACT_TIME_STAT for permitted values and description.
Date FBR Schematic Received	Date primacy agency received treatment plant schematic to demonstrate the percentage of filter backwash that is returned to the treatment plant influent.
Date FBR Schematic Reviewed	Date primacy agency completes review of treatment plant schematic and determines the percentage of filter backwash that is returned to the treatment plant influent.
Status of Alternate Return Location for FBR	The status of a request from the water system to request an alternate location for return of the filter backwash.
Date of Alternate Return Location for FBR	The date that the water system requested an alternate location for return of the filter backwash.
Status of FBR Corrective Action	The status of corrective action by the water system as required by the primacy agency after review of the schematic of the filter backwash flow in the treatment plant.
FBR Corrective Action Date	The date that the water system achieved the corrective action required for the filter backwash.
User ID Initials	The User ID of the person who created this record.
FBR Comments	A memo field into which a user may enter comments about the Filter Backwash Recycling Rule.
Disinfection Benchmark Reason	Text description associated with the Disinfection Benchmark Reason.
Contact Time Reason	Text description associated with the Contact Time.
Treatment process table (tblTreatProcess)	
Treatment Process ID	Unique identifier for each treatment record.
Water Facility ID	Identifier that relates each record to the unique record in the tblSixYrWsf table.
Treatment Objective Code	A coded value that categorizes the treatment objective.
Treatment Objective Name	The name of the treatment objective.
Treatment Process Code	A coded value that categorizes the treatment process.
Treatment Process Name	The name of the treatment process.
Water system flows table (tblSixYrWsfFlows)	
Water System Facility Flow ID	Unique identifier for each water system facility flow record.
Water Facility ID	Identifier that relates each record to the unique record in the tblSixYrWsf table.
Facility Flow ID Number	Identifier for each water system facility flow entry that is unique when combined with SixYrWsf_ID.
Facility Train ID	This attribute identifies the water system facilities that are part of the same flow.
Sequence ID	This attribute identifies the order of the water system facilities in a specific flow.
Process Water Type	A system administrator controlled code of the type of water flowing between the facilities.
Water Quantity Measure	A value that represents the number of gallons of water purchased.
Water Quantity Measure Units	A coded value which specifies the unit of measurement for the quantity of water purchased.
Connection Type	Categorizes the type of connection between the water system facilities.
Connection Date	The date of the connection of the water system facility to another water system facility.
Disconnection Date	The date of the disconnection of the water system facility from another water system facility.
Supplying Facility ID	Identifier for each supplying water system facility that is unique when combined with TINWSF0ST_CODE.
Supplying Facility State Code	State in which the supplying facility is located using the states' two letter abbreviation.

Website Exhibit 10: Treatment Data Diagram

Treatment process table

tblTreatProcess

- **Treatment Process ID**
- *Water Facility ID*
- Treatment Objective Code
- Treatment Objective Name
- Treatment Process Code
- Treatment Process Name

Water system facility plant table

tblSixYrWsfPlt

- **Treatment Plant ID**
- *Water Facility ID*
- State Assigned ID Code
- Water Facility Type
- Filter Type
- Description of Filter
- Disinfectant Concentration (mg/L)
- Contact Time Status
- Contact Time Determination Date
- Contact Time
- CT Value
- Disinfection Benchmark for Giardia Inactivation in Logs
- Status of Disinfection Benchmark for Giardia Inactivation
- Date of Disinfection Benchmark for Giardia
- Disinfection Benchmark for Giardia Inactivation Percent
- Disinfection Benchmark for Virus Inactivation in Logs
- Status of Disinfection Benchmark for Virus Inactivation
- Date of Disinfection Benchmark for Virus
- Disinfection Benchmark for Virus Inactivation Percent
- FBR Schematic Status
- Date of FBR Schematic Received
- Date FBR Schematic Reviewed
- Status of Alternate Return Location for FBR
- Date of Alternate Return Location for FBR
- Status of FBR Corrective Action
- FBR Corrective Action Date
- User ID Initials
- FBR Comments
- Disinfection Benchmark Reason
- Contact Time Reason

Water system table

tblSixYrWs

- **Water System ID**
- PWSID
- System Name
- Retail Population Served
- Source Water Type
- System Type
- State Code
- Adjusted Total Population Served

Water system facility table

tblSixYrWsf

- **Water Facility ID**
- *Water System ID*
- Water Facility Type

Water system flows table

tblSixYrWsfFlows

- **Water System Facility Flow ID**
- *Water Facility ID*
- Facility Flow ID Number
- Facility Train ID
- Sequence ID
- Process Water Type
- Water Quantity Measure
- Water Quantity Measure Unit
- Connection Type
- Connection Date
- Disconnection Date
- Supplying Facility ID
- Supplying Facility State Code

Section 6: SYR3 Data Considerations

The SYR3 ICR data is of reasonable quality and is representative and appropriate for use to support national, scientifically-defensible findings. Data has undergone appropriate quality assurance evaluation and enough states provided compliance monitoring to be representative for national-scale analyses. EPA used the data in analytical activities informing decisions for Six Year 3. The data include sufficient information for users to be able to reproduce the SYR3 analyses.

There are a few limitations of the final SYR3 ICR dataset that should also be acknowledged. There may be different levels of completeness for different contaminants within the dataset. In some cases, the number of records per state ranged from less than one hundred records up to more than 1 million records for a given contaminant. States might not have submitted data for certain contaminants if they have monitoring waivers for the contaminant. States may grant waivers to PWSs to reduce monitoring frequencies, and it is possible that no samples were collected by systems during the SYR3 period of review. Other states may have submitted data for these contaminants under the ICR; however, the data were not in a format compatible with the SYR3 ICR dataset. Furthermore, there were four states and some other tribes/territories whose data are missing entirely from the analysis.

A thorough QA/QC process was undertaken to evaluate these SYR3 ICR data used for analyses. However, it is possible that data entry errors may still exist in the final SYR3 ICR Dataset. The QA/QC review focused only on the data elements essential for analysis.

For a complete discussion of the SYR3 ICR dataset, including a description of the quality assurance/quality control review, refer to USEPA (2016a) and USEPA (2016b). For more detailed information on the microbial contaminants' occurrence analysis, refer to USEPA (2016c). For more detailed information on the occurrence analysis of contaminants regulated under the Stage 1 and Stage 2 Disinfectant and Disinfection Byproducts Rules, refer to USEPA (2016d).

Instructions on Importing SYR3 Datasets to Excel

These text files are tab delimited and have no text qualifier. Field names are included in the first row of each file. A basic understanding of Microsoft Excel is necessary to effectively use these instructions. Using Microsoft Excel 2013 or a newer version is recommended due to the size of the dataset(s). Note, however, that the complete SYR3 ICR Dataset is too large to be imported into Excel. The data are available for download for each parameter and should be imported into a data management system that supports large datasets for analysis.

Part One: Downloading and Importing Data (Note that instructions may vary depending on the version and software used to import data.)

1. Begin by reviewing the SYR3 ICR Dataset Summary (Exhibit 2) and in particular note the table of Data Field Names and Definitions (Exhibit 1).

2. Access the SYR3 ICR data by going to the Six-Year Review homepage. Click on the link for "Six-Year Review 3."

3. Click on the desired zip file and select **Save As** to save the file to your computer.

4. Navigate to the location on your computer where you saved the zip file and unzip or extract the zip file contents by clicking **Open with** and using Win Zip or Microsoft Compression.

5. Open a blank workbook in Microsoft Excel.

6. In the workbook, select **Data** among the tabs at the top of the page.

7. On the far left, top of the screen, go to the **Get External Data** section and select **From Text**.

8. You will be prompted to select a text file. Locate the text files you unzipped or extracted in Step 4, and click **Import** on the text file that of interest.

9. The Text Import Wizard – Step 1 of 3 will appear. The default settings will be displayed and should have **Delimited** selected as the **Original data type**. Select the checkmark box next to **My data has headers.** Click **Next>**.

10. The Text Import Wizard – Step 2 of 3 will appear. The default settings will be displayed and should have **Tab** selected as the **Delimiter** while **Treat consecutive delimiters as one** should be **unselected**. Select **Text qualifier** as **{none}** from the dropdown menu. Click **Next>.**

11. The Text Import Wizard – Step 3 of 3 will appear. The default settings will be displayed and will specify each column data format a**s General**. Click **Finish**. See #18 for further details about formatting.

12. The **Import Data** prompt will appear. Click **OK**. This import may take several minutes.

13. Save the Excel spreadsheet file.

Part Two: Filtering and Formatting Data in Excel

14. To efficiently search, have cell A1 selected, choose **Data** among the tabs on the top of the page and click on the **Filter**. Each header title for each column now will have a small dropdown arrow displayed.

15. Filtering the data:
    a. If you want to look for a specific public water system, click the dropdown arrow for "PWSID" or "System Name." Within the search field, type the name and select from the displayed list.
    b. If you want to search for a different public water system, click the dropdown arrow and "Clear Filter from PWSID" or "Clear Filter from System Name."
    c. If you want to filter the data by contaminant, select "Analyte Name."

16. Multiple filters can be applied for example, allowing you to look for an individual water system's data for a specific contaminant of interest.

17. De-select **Filter** in the top menu bar and the entire database will again be displayed.

18. Note, all column formats are imported as the default **General** formatting. Column formats must be individually, manually changed in excel after the download is complete to aid in data analysis. Use

the **Home** screen in excel, highlight the column and select the format from the drop down menu. Suggested formats are:

a. **Text** for: Analyte Name, State Code, PWSID, System Name, System Type, Source Water Type, Water Facility Type, Sampling Point Type, Source Type Code, Sample Type Code, Laboratory Assigned ID, Sample Collection Date, Detection Limit Unit, Detection Limit Code, Value Unit, Presence Indicator Code.

b. **Number** for: Analyte ID, Retail Population Served, Adjusted Total Population Served, Water Facility ID, Sampling Point ID, Six Year ID, Sample ID, Detection Limit Value, Detect, Value, Residual Field Free Chlorine mg/L, Residual Field Total Chlorine mg/L.

Section 7: Supplemental Data Sources

Several supplemental data sources were used to support the national contaminant occurrence and exposure assessments performed under the Six-Year Review process. These supplemental data sources are described below.

Disinfection Byproducts (DBP) Information Collection Rule (ICR) (Filename: DBPICR_Aux1)

The DBP ICR "Aux 1" database houses monitoring data from large public water systems (PWSs serving a population greater than or equal to 100,000) from the 18-month period of July 1997 to December 1998. A total of 296 water systems reported data; included in the database are monitoring results for microbials and DBPs, plant treatment, source water characteristics and disinfectant type information. This database was previously used in the development of the Stage 2 D/DBPR. Refer to McGuire et al (2002) for additional information.

For the SYR3 review, this database was used for several purposes, including the following: to investigate changes in disinfection practices; to evaluate changes in DBP precursor occurrence and removal; and to evaluate chlorate occurrence and co-occurrence of chlorate and chlorite. Refer to USEPA (2016d) and USEPA (2016f) for additional information.

Within the "Aux 1" version of the database, there are 31 relational tables within the database, plus several other tables providing additional information such as descriptions of each table, data element, attribute, etc.

The DBP ICR (Aux 1) database is posted online in Microsoft Access. The data documentation file is posted alongside the data. This documentation explains to the user all of the various data elements and tables included in the database.

EPA ICR Treatment Study Database (TSD) (Filename: ICR_TSD)

The ICR TSD was constructed to manage the treatment study data submitted by the systems required to conduct DBP precursor removal studies under the 1996 ICR. Results from 99 treatment studies (63 granular activated carbon (GAC) and 36 membrane studies), are reported in this database. This database was previously used in the development of the Stage 2 D/DBPR. Refer to McGuire et al. (2002) for additional information.

For the SYR3 review, this database was further used to evaluate the reduction of brominated DBP formation by GAC. Refer to USEPA (2016d) for additional information.

The TSD is posted online in Microsoft Access. There is a data documentation file (entitled "TS Database User's Guide") posted alongside the data to provide an explanation to the user all of the various data elements and tables included in the database.

TSD files posted online:

1. ***TSDatabase.accdb*** (the TSD Access database file) – 28 MB
2. ***TSDB_Documents***: Includes pdf documents that users access from the database's "Documentation" section
    a. *BenchPilotManual*: ICR Manual for Bench- and Pilot-Scale Treatment Studies

      b. *DataSprdShtMnl*: ICR TS Data collection spreadsheets User's Guide
      c. *GAC Base Analysis Doc*: Base Analysis Document: GAC Studies
      d. *Membrane Base Analysis Doc*: Base Analysis Document: Membrane Studies
      e. *TS Database User's Guide*: Treatment Study Database User's Guide
3. **TSLIB_DB_DCS**: Excel Data Collection Spreadsheets for all samples
4. **TSLIB_DB_Graph**: PDF Graphical Summary Files for all samples
5. **TSLIB_SumRpt**: PDF Summary Reports for all samples

Structure of TSD Files for Posted Online:

1. Download and save **ICR_TSD** to your local hard drive in "C:\".
2. Extract the files from **ICR_TSD** and rename the destination folder as "C:\icr". See screenshot below for an example of the structure and location of files once the data have been extracted and saved locally.
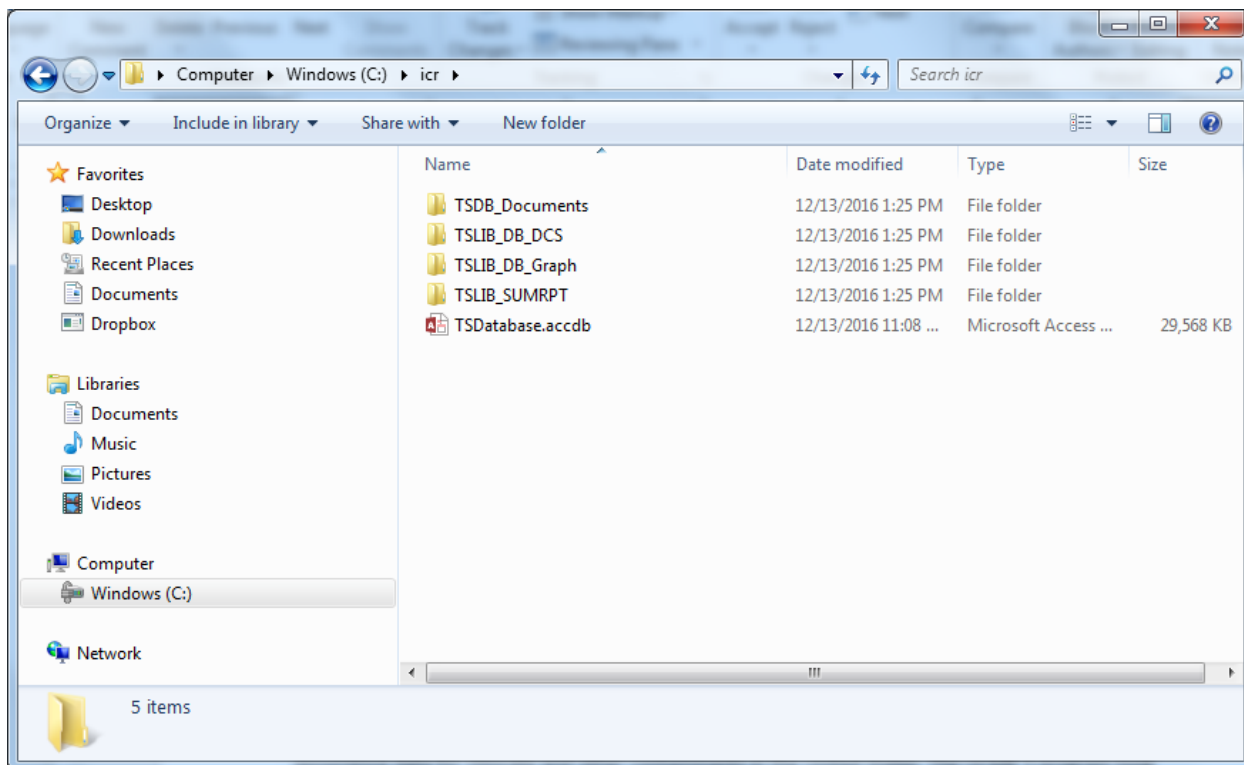
Second Unregulated Contaminant Monitoring Rule (UCMR 2) Data

Data are available for nitrosamine occurrence in finished drinking water in public water systems (PWSs) from the nationally representative monitoring completed under the Second Unregulated Contaminant Monitoring Rule (UCMR 2). UCMR 2 monitoring included monitoring for all six nitrosamines discussed in the SYR3 nitrosamine support document (USEPA, 2016e): N-nitrosodi-n butylamine (NDBA), N nitrosodiethylamine (NDEA), N nitrosodimethylamine (NDMA), N nitrosodi-n propylamine (NDPA), N nitrosomethylethylamine (NMEA) and N nitrosopyrrolidine (NPYR).

UCMR 2 monitoring, conducted between January 2008 and December 2010, provided data about nitrosamine occurrence; these data are available from the agency's website (https://www.epa.gov/dwucmr/occurrence-data-unregulated-contaminant-monitoring-rule#2).

Third Unregulated Contaminant Monitoring Rule (UCMR 3) – July 2016 version (Filename: UCMR3_July2016)

The data available for chlorate occurrence in finished drinking water in PWSs are from the nationally representative monitoring completed under the third round of the Unregulated Contaminant Monitoring Rule (UCMR 3). The UCMR 3 monitoring provides nationally representative contaminant occurrence data for chlorate and other contaminants in the United States. The UCMR 3 program took place from 2012 to 2015.

The UCMR 3 occurrence analyses presented in SYR3 chlorate support document (USEPA, 2016f) are based on data collected through May 2016 and released in July 2016 (USEPA, 2016g). EPA expects a relatively small amount of data reporting to continue after July 2016. The UCMR 3 dataset will not be considered "final" until early 2017. EPA does not anticipate that there will be any substantial difference between findings based on the July 2016 dataset and findings based on the final dataset.

Safe Drinking Water Information System (SDWIS) Information

The Safe Drinking Water Information System (SDWIS) contains information about public water systems and their violations of EPA's drinking water regulations, as reported to EPA by the states. Several versions of SDWIS datasets were used to support the national contaminant occurrence and exposure assessments performed for SYR3. This section provides the applicable SDWIS dataset file names on EPA's occurrence data webpage, and describes how these data were used for SYR3.

Note that the varying activity issues in the SDWIS datasets described below could cause confusion about the understanding of the data being presented. For example, there are active and inactive systems, non-public systems, systems that have merged with other systems and potential future systems included in the SDWIS datasets. The inactive, non-public and potential future systems were not used in the occurrence analyses but are included in the data posted online. There are also systems that have been inactive for many years.

SDWIS 2011 Pivot Tables (Filename: SDWIS2011_Pivot)

SDWIS inventory data were used to assess representativeness of SYR3 ICR data on both state and national levels. This is discussed further in chapter 6 of the D/DBPR support document (USEPA, 2016d). Note: the data within this file represents data ending in FY 2013. The file does contain information from 2010 to 2013; however, only the 2011 data were used for this analysis.

SDWIS Violation Data (Filename: SDWISViolations_2006-2011)

SDWIS violation data were used to assess violation rates and representativeness of populations. EPA conducted this assessment for the IOCs, SOCs, VOCs, and radionuclides.

2011 SDWIS/FED Freeze (Filename: SDWIS2011_Freeze)

A SDWIS/FED freeze from December 2011 was used to populate missing inventory information (e.g., source water type or population served) for some of the non-SDWIS states. This version of SDWIS was also used to evaluate the completeness of the data submitted for SYR3.

Note that Safe Drinking Water Information System (SDWIS) Quarterly Freeze is a copy of the data contained in SDWIS as of a specific year and quarter and includes all information available in the system at that time.

2010 SDWIS/FED Freeze (Filename: SDWIS2010_Freeze)

A SDWIS/FED freeze from December 2010 was used to identify the system type and for the national extrapolation of small system occurrence data for chlorate. Refer to the SYR3 chlorate support document (USEPA, 2016f) for additional information.

SDWIS Buyers-Sellers (Filename: SDWISBuyers_Sellers)

A list of buyer-wholesaler relationships from a fourth quarter 2010 SDWIS/FED freeze was used to adjust the population values of the wholesale systems to include the population of the systems that they sell water to (the purchased water systems). Refer to "The Analysis of Regulated Contaminant Occurrence Data from Public Water Systems in Support of the Third Six-Year Review of National Primary Drinking Water Regulations: Chemical Phase Rules and Radionuclides Rules" (USEPA, 2016b) for additional information.

2005 SDWIS Freeze (Filename: SDWIS2005_Freeze)

A 2005 SDWIS freeze was used in the occurrence analyses of nitrosamines to categorize PWSs by their source water type and by the size of the population served. Refer to the SYR3 nitrosamine support document (USEPA, 2016e) for additional information.

LT2 Round 1 Monitoring Data

In support of its LT2 analyses, EPA used data from the Data Collection and Tracking System (DCTS) pull from April 2012, which contained 44,944 records representing all system sizes. EPA posted the original and "cleaned-up" datasets on the EPA website at: https://www.epa.gov/dwsixyearreview/long-term-2-enhanced-surface-water-treatment-lt2-rule-round-1-source-water. Refer to the LT2 support document (USEPA, 2016h) for additional information.

User Guide to Downloading and Using SYR3 Data from EPA's Website

References

McGuire, M.J., J.L. McLain, and A. Obolensky (eds.). 2002. *Information Collection Rule Data Analysis*. Denver, CO: American Water Works Research Foundation Research Foundation and American Water Works Association, 600 p. Available online at http://www.waterrf.org/PublicReportLibrary/90947.pdf.

United States Environmental Protection Agency (USEPA). 2010. Agency Information Collection Activities; Submission to OMB for Review and Approval; Contaminant Occurrence Data in Support of EPA's Third Six Year Review of National Primary Drinking Water Regulations (Renewal). Notice: February 5, 2010, Volume 75, Number 24, Page 6023-6024.

USEPA. 2016a. The Data Management and Quality Assurance/Quality Control Process for the Third Six-Year Review Information Collection Rule Dataset. EPA-810-R-16-015. December 2016.

USEPA. 2016b. The Analysis of Regulated Contaminant Occurrence Data from Public Water Systems in Support of the Third Six-Year Review of National Primary Drinking Water Regulations: Chemical Phase Rules and Radionuclides Rules. EPA 810-R-16-014. December 2016.

USEPA. 2016c. Six- Year Review 3 Technical Support Document for Microbial Contaminant Regulations. EPA-810-R16-010. December 2016.

USEPA. 2016d. Six- Year Review 3 Technical Support Document for Disinfectants/Disinfection Byproducts Rules. EPA-810-R-16-012. December 2016.

USEPA. 2016e. Six- Year Review 3 Technical Support Document for Nitrosamines. EPA-810-R-16-009. December 2016.

USEPA. 2016f. Six- Year Review 3 Technical Support Document for Chlorate. EPA-810-R-16-013. December 2016.

USEPA. 2016g. Third Unregulated Contaminant Monitoring Rule Dataset. July, 2016 version.

USEPA. 2016h. Six- Year Review 3 Technical Support Document for Long- Term 2 Enhanced Surface Water Treatment Rule. EPA-810-R-16-011.