U.S. EPA Office of Compliance

Water Quality Indicators (WQI) Project Background and Technical Specifications

September 2024

Please contact Rusty Wasem (wasem.russell@epa.gov) with questions, comments, or feedback.

Contents

Previous Versions and Description of Changes

June 2017: First Published Version

September 2017:

Add most-recent-48-month statistics

Extend search for speciation(units2) to the MethodSpecificationName column in the U.S. Water Quality Portal (WQP) "Biological" Result Retrieval

November 2019

Previous versions of the WQI utilized only data where the nutrient ecoregion was known, after adding GUI options to use a Harmful Algal Bloom (HAB) Benchmark, user generated criteria, and Top 50-500, knowing the nutrient ecoregion or being sampled in a nutrient ecoregion was no longer a prerequisite for data to be used in the WQI.

Previous versions of the WQI utilized only data where the WQP's parameter countrycode="US." It was later discovered many contributors to the WQP do not populate the countrycode.

Spatial filters will be used to remove data from monitoring locations that are not located in the U.S. or Territories of the U.S.

Ecoregion ID is still derived when applicable (e.g. monitoring locations in the CONUS).

Processing Summary and Data Usability section added.

Added Appendix A: Alternate Data Format.

August 2021

Adds Pathogen data for e. coli, enterococcus, and fecal coliform. Pathogen analysis includes the following waterbody types:

- Rivers/Streams
- Lakes/Reservoirs
- Salt Water/Marine

The Beach Action Value of 70 cfu/100ml was applied to analysis of e. coli and enterococcus at all waterbody types.

The Red Book Value 1976 of 200 cfu/100ml was applied to analysis of fecal coliform at all waterbody types.

Generated outputs (original format, and v22 / v24 format) are identical to that of the nitrogen and phosphorus analysis.

September 2024

Clarifies that the WQI's data processing treats MPN/CFU/CCE/100ml as being equivalent.

Clarifies that salt and marine water body types are considered for pathogen pollutants.

Background:

The Water Quality Indicators Project was originally developed as a component of U.S. Environmental Protection Agency's (EPA) "Clean Water Action Plan." The Action Plan sought to better align inspections with areas of poor water quality – particularly focusing on enforcement at facilities that have violations and contribute to water impairment. After reviewing information about watersheds that have been designated as having poor water quality, it became clear that a large number of watersheds have not been assessed. EPA was concerned that shifting resources into impaired watersheds might miss important areas that have not been designated as impaired. To address this data gap, EPA designed the WQI to identify potential hotspots using water station monitoring data, that when compared to water quality criteria, could yield a secondary data source showing possible problem areas. The WQI set out to use ambient water quality data in the EPA database STOrage and RETrieval of Water Quality Data (STORET) (currently known as the Water Quality Exchange or WQX) and U.S. Geological Survey (USGS) database National Water Information System (NWIS) to flag "hotspots" which could then be evaluated by interested EPA, state, and tribal staff. The development of the Water Quality Indicators (WQI) project coincided with an initiative by the Office of Enforcement and Compliance Assurance (OECA) to begin using data analytics to integrate and assess large data sets for problem identification. Under the WQI project, EPA used data analytic methods to bring together and compare these data sources. After compiling a data set and establishing a data refresh process, EPA developed mapping and data visualization tools that will help users display and assess the data. To develop the methods and techniques needed to produce the WQI, EPA convened a workgroup that included representatives from EPA Regions 2, 4 and 10, OAR, ORD, OEI, OW, OECA, OCSPP, and USGS, all whose contributions are gratefully acknowledged.

Scope:

Initially, it was decided that hotspot identification would consist of the comparison of water quality measurements with ambient water quality criteria based on national standards as indexed within defined ecoregions. EPA recognizes that states and tribes have developed water quality criteria that may differ from national standards; however, given resources, it was beyond the reach of the initial Indicators project to attempt to identify and mobilize all the appropriate state criteria. Instead, EPA designed the user interface so that users can input their own concern levels, while making default settings correspond to the CWA 304(a) recommended criteria for nutrients. Data collected from before 2009 would not be used, and the initial pollutants included in the project were limited to the nutrients nitrogen and phosphorus – a choice which would support analysis of many sectors.

Later, when the WQI added pathogen pollutants, the 2012 Recreational Water Quality Criteria for E. Coli and Enterococci, and the 1976 "Red Book" Recreational and Shellfish Harvesting Criteria for Fecal Coliforms were used.

Approach and Data Use:

The U.S. Water Quality Portal (WQP) ambient water quality monitoring data were used. The period of interest was initially set to eight years (yearbegin=2009, yearend=2016) – with the intention of eventually increasing it to the most recent ten years - and relevant statistics (percentiles, counts) for each month and over all the months [yearend-yearbegin + 1) *12 months] in the period -- would be determined for all monitoring locations nationally by pollutant for which there were data. These results were then associated with a graphical user interface which would visually identify nationally locations where the water quality measurements fell into one of several contiguous percentage ranges, extending from below the criteria to above, as desired by the user. This interface was made available to government users for exploration during a trial period held from April 2016 to December 2016 via ECHO and EPA's Office of Enforcement and Compliance Assurance (OECA) collected comments from participating users. WQI Version 1.0 was released for government use in February 2017. In 2021, EPA added the pathogen pollutants Enterococcus, E. Coli and Fecal Coliform. EPA is working to expand the project to encompass additional pollutants and expects to engage with interested government stakeholders on this expansion. The current WQI data file can be downloaded.

Vetting the data was a major undertaking that is still considered a work in progress -- involving identifying protocols for

recognizing detection limits (for non-detects); for rejecting poor quality measurement data; for repairing poor quality or ambiguous reporting (e.g. multiple meanings of "total"); for recognizing duplicate samples; and for establishing a protocol for combining the particular population of measurements of those fractions of nitrogen and phosphorus found in the data into total nitrogen and total phosphorus, the measures used in the 304(a) criteria.

WQX and NWIS measurement submissions have an extensive set of metadata questions, which, when answered properly, assist greatly in determining which measurements to use for a given analysis and which not to use. When crucial metadata are incorrect or incomplete, measurements which would otherwise be usable may be deemed unusable, as was the case with about half the nitrogen and phosphorus measurements in the Version 1.0 of the WQI released in February 2017 (see current and historical File Processing Summaries).

Some of the more crucial fields (WQP field names) are: OrganizationIdentifier, MonitoringLocationIdentifier, ActivityIdentifier, ActivityStartDate, CharacteristicName, ResultSampleFractionText, ResultDetectionConditionText, ResultMeasureValue, ResultMeasure/MeasureUnitCode(units1 and units2), DetectionQuantitationLimitTypeName, DetectionQuantitationLimitTypeMeasure/MeasureValue, DetectionQuantitationLimitTypeMeasure/MeasureUnitCode(units1 and units2). Units1 is the concentration, units2 is the reference species (e.g., in "2 mg/l as N", units1="mg/l" and units2="N"). Agencies with incomplete or inaccurate crucial metadata can improve their agency representation rate in the WQI by adding and/or correcting the metadata. The WQX data in the WQP are updated every week. The WQI plans to refresh the data in April and October.

Virus samples are usually measured from at least 1 liter of water and for the purposes of processing pathogen data, the WQI treats colony forming unit (CFU), calibrator cell equivalent (CCE), and most probable number (MPN) as equivalent.

Source Data and Water Quality Criteria

WQP is a cooperative service sponsored by USGS and EPA that integrates publicly available water quality data from USGS' NWIS and EPA's WQX.

The EPA water quality data originate from the WQX, the EPA's repository of water quality monitoring data are collected by water resource management groups across the country. Organizations, including states, Tribes, watershed groups, other federal agencies, volunteer groups, and universities, submit data to the WQX to make their data publicly accessible.

The USGS water quality data originate from the NWIS Database. The database contains current and historical water data from more than 1.5 million sites across the nation. It is used by state and local governments, public and private utilities, private citizens, and other federal agencies involved with managing water resources. All publicly available discrete water quality data are currently available through the WQP.

The WQP data are served up in two files: a Site file and a Result file. The element MonitoringStationIdentifier is the key that links the two files. There are multiple result records (measurements) for one Site record.

The first version of the WQI only included nitrogen and phosphorus. Ideally the state water quality criteria for nitrogen and phosphorus should be used, but when the WQI project began few states had total nitrogen and total phosphorus criteria and when criteria existed, they were described idiosyncratically by each state sometimes in unstructured text. Consequently, any effort to extract and normalize state criteria was tabled, and instead, the CWA 304(a) national criteria are used, which for nutrients including nitrogen and phosphorus defer to Aggregate Nutrient Ecoregional (ANE) criteria The ANE criteria were published by EPA in the early 2000s and "represent conditions of surface waters that have minimal impacts caused by human activities, starting points to identify more precise numeric levels for nutrient parameters needed to protect aquatic life, recreational, or other uses on site-specific or subregion-specific conditions" There are 14 ANEs, covering the conterminous 48 states only. The ANE consist of aggregates of "old" Level III ecoregions (not to be confused with later Level III ecoregions).

The N and P criteria applicable to waterbodies fall into one of two categories -- river/stream or lake/reservoir.

The second major release of the WQI added monitoring data for the pathogen pollutants e. coli, enterococcus, and fecal coliforms. The Beach Action Value of 70 cfu/100ml was applied to analysis of e. coli and enterococcus at all waterbody types. The Red Book Value 1976 of 200 cfu/100ml was applied to analysis of fecal coliform at all waterbody types.

Pathogen criteria was applied to waterbodies categories:

- Rivers/Streams
- Lakes/Reservoirs
- Salt Water/Marine

Previous versions of the WQI utilized only data where the nutrient ecoregion was known, after adding GUI options to use a Harmful Algal Bloom (HAB) Benchmark, user generated criteria, and Top 50-500, knowing the nutrient ecoregion or being sampled in a nutrient ecoregion was no longer a prerequisite for data to be used in the WQI.

Previous versions of the WQI utilized only data where the WQP's parameter countrycode="US." It was discovered than many contributors that provide sampling data to the WQP do not populate the countrycode. Spatial filters are used to remove data from monitoring locations that are not located in the U.S. or Territories of the U.S.

Output

For initial implementation a CSV format record was decided upon, consisting of these elements taken directly from the WQP Site file:

MonitoringLocationIdentifier
OrganizationFormalName(OrganizationIdentifier)
MonitoringLocationName
LatitudeMeasure
LongitudeMeasure
CEightDigitCode
StateCode
CountyCode
MonitoringLocationTypeName,

These elements derived by WQI Site file protocols:
WaterBodyType ("R"[iver/Stream] or "L"[ake/Reservoir])
Ecoregion Identifier (I, II, III…XIV,noeco)

These elements derived by Result file processing:
Up to ten groups of 12-values per group. If the first value in a group is:

N: The pollutant is nitrogen and the interval for the summary statistics is the WQI entire period of record, which has a length in months of (yearend-yearbegin+1)*12 (from Jan yearbegin to Dec yearend).

N48: The pollutant is nitrogen and the interval for the summary statistics is the 48-month period ending on the date of the Portals retrieval. (The entire period of record is assumed to be greater than 48 months.)

P: The pollutant is phosphorus and the interval for the summary statistics is the WQI entire period of record, which has a length in months of (yearend-yearbegin+1)*12 (from Jan yearbegin to Dec yearend).

P48: The pollutant is phosphorus and the interval for the summary statistics is the 48-month period ending on the date of the WQP retrieval. (The entire period of record is assumed to be greater than 48 months.)

The pathogens follow the same format above as Nitrogen and Phosphorus:

- E. Coli
- E. Coli 48
- Enterococcus
- Enterococcus48
- Fecal Coliform
- Fecal Coliform48

Each time one of the above markers (N, N48, P, P48, E. Coli, E. Coli48, Enterococcus, Enterococcus48, Fecal Coliform, Fecal Coliform48) occurs, it will be followed by the following eleven fields:

Number of samples
Number of samples rejected because of incomplete metadata
Samples maximum
Samples median
Samples time-weighted median
Samples 70th percentile
Samples time-weighted 70th percentile
Samples 90th percentile
Samples time-weighted 90th percentile
Number of sample values greater than the default 304(a) criteria
Maximum of monthly medians

When a month (yymm) appears instead of a marker as the first value in a group, the month will be followed by the first ten of these eleven fields and the period of record is the month. ("Maximum of monthly medians" does not apply to monthly data)

Only the months with data appear in the record. The sequence of fields in the record is N, N48, P, P48, months in ascending yymm order.

Here is an example record with 44 used and 0 rejected N samples, and 55 used and 0 rejected P samples (the white space added for clarity is not in the record):

21ARIZ_WQX-SRTON053.87,ARIZONA DEPARTMENT OF ENVIRONMENTAL QUALITY(21ARIZ_WQX),TONTO CREEK - BELOW BEAR FLATS,34.2833333,-111.0708333,15060105,04,007,River/Stream,R,II,

N,44,0,1.000,0.370,0.370,0.433,0.430,0.575,0.575,43,0.520,N48,22,0,0.910,0.360,0.350,0.410,0.410,0.460,0.460,21,0.412,1305,5,0,0.700,0.520,0.520,0.564,0.575,0.650,0.700,5,1306,6,0,1.000,0.375,0.430,0.502,0.575,0.788,1.000,6,1307,5,0,0.540,0.400,0.400,0.456,0.470,0.512,0.540,5,1308,6,0,0.550,0.317,0.325,0.347,0.370,0.460,0.550,6,1309,6,0,0.910,0.412,0.430,0.445,0.460,0.685,0.910,6,1407,5,0,0.785,0.410,0.410,0.450,0.460,0.655,0.785,5,1408,6,0,0.370,0.265,0.270,0.285,0.300,0.335,0.370,5,1409,5,0,0.420,0.290,0.290,0.314,0.320,0.380,0.420,5,

P,55,0,4100.000,50.000,50.000,85.800,92.000,153.000,155.000,55,2125.000,P48,22,0,230.000,91.000,90.000,130.000,130.000,154.500,155.000,22,127.250,0903,1,0,17.500,17.500,17.500,17.500,17.500,17.500,17.500,1,0904,2,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,2,0908,1,0,78.000,78.000,78.000,78.000,78.000,78.000,78.000,1,0909,1,0,77.000,77.000,77.000,77.000,77.000,77.000,77.000,1,1007,2,0,4100.000,2125.000,4100.000,2915.000,4100.000,3705.000,4100.000,2,1008,1,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,1,1009,1,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,25.000,1,1106,1,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,1,1107,1,0,15.000,15.000,15.000,15.000,15.000,15.000,15.000,1,1109,2,0,50.000,50.000,50.000,50.000,50.000,50.000,50.000,2,1305,4,0,36.000,32.500,34.000,34.200,34.000,35.400,36.000,4,1306,5,0,35.000,32.500,32.500,32.900,33.000,34.200,35.000,5,1307,5,0,500.000,49.000,49.000,361.800,440.000,476.000,500.000,5,1308,6,0,87.000,59.000,68.000,73.750,79.500,83.250,87.000,6,1309,6,0,230.000,127.250,150.000,152.500,155.000,192.500,230.000,6,1407,5,0,150.000,50.000,50.000,130.000,150.000,150.000,150.

000,5,1408,6,0,190.000,88.000,120.000,125.000,130.000,160.000,190.000,6,1409,5,0,130.000,90.000,90.000,91.600,92.000,114.800,130.000,5

Here is a sample record for Enterococcus:

1083919_WQX-26.2W,Bronx River Alliance (Volunteer)(1083919_WQX),Piermont- Pirelli Park,41.0434200000,-73.9137000000,02030101,36,087,River/Stream,R,XIV,
Enterococcus,3,0,75.000,63.000,63.000,67.800,75.000,72.600,75.000,1,75.000,Enterococcus48,3,0,75.000,63.000,63.000,67.800,75.000,72.600,75.000,1,75.000,1707,1,0,63.000,63.000,63.000,63.000,63.000,63.000,63.000,0,1708,1,0,75.000,75.000,75.000,75.000,75.000,75.000,75.000,1,1710,1,0,0.110,0.110,0.110,0.110,0.110,0.110,0.110,0

Here is a sample record for E. Coli:

11NPSWRD_WQX-GUCO_HFSP,National Park Service Water Resources Division(11NPSWRD_WQX),Hoskins Farm Spring,36.1291860000,-79.8525990000,03030002,37,081,Spring,R,IX,

E. Coli,10,3,461.100,18.450,25.900,173.540,203.500,248.970,461.100,4,488.400,E. Coli48,1,0,461.100,461.100,461.100,461.100,461.100,461.100,461.100,1,461.100,1110,2,0,11.000,10.950,11.000,10.970,11.000,10.990,11.000,0,1201,1,0,160.700,160.700,160.700,160.700,160.700,160.700,160.700,1,1204,2,0,6.300,4.150,6.300,5.010,6.300,5.870,6.300,0,1207,1,0,2.000,2.000,2.000,2.000,2.000,2.000,2.000,0,1310,1,0,225.400,225.400,225.400,225.400,225.400,225.400,225.400,1,1401,1,0,203.500,203.500,203.500,203.500,203.500,203.500,203.500,1,1404,0,2,,,,,,,,,1407,0,1,,,,,,,,1611,1,0,25.900,25.900,25.900,25.900,25.900,25.900,25.900,0,1708,1,0,461.100,461.100,461.100,461.100,461.100,461.100,461.100,1

Site File Processing

As noted in the previous section ***Output***, eight original elements are needed from the Site file and one additional elements, waterbody type must be derived. Ecoregion ID is still derived when applicable (e.g. nutrient monitoring locations in the CONUS).

Waterbody Type

For waterbody type, the Site file is read and an inventory of MonitoringLocationTypeName (the Site file element name for waterbody type) was made. There were about 120 different types. The two columns of this inventory (count, waterbody type) were put into the first two columns of a spreadsheet, and two additional columns were added on the right, one for whether or not the named type could legitimately be compared to the criteria, and, if so, the second column a place to indicate whether the river(R), lake(L), or salt(S) criterion should be used (see Figure T).

Count (From 2015)	MonitoringLocationTypeName	Use?	Criterion type: R=rivers/streams, L=lakes/reservoirs S=salt water/marine
120	(no type present)	No	
1,473	Aggregate groundwater use	No	
1,248	Aggregate surface-water-use	No	
5,196	Atmosphere	No	
2	BEACH Program Site-Channelized stream	Yes	R
1,551	BEACH Program Site-Estuary	Yes	S

821	BEACH Program Site-Great Lake	Yes	L
444	BEACH Program Site-Lake	Yes	L
2	BEACH Program Site-Land	No	
6	BEACH Program Site-Land runoff	No	
4,195	BEACH Program Site-Ocean	Yes	S
241	BEACH Program Site-River/Stream	Yes	R
1	BEACH Program Site-Waste sewer	No	
797	Borehole	No	
3,317	Canal Drainage	Yes	R
550	Canal Irrigation	Yes	R
4,624	Canal Transport	Yes	R
540	Cave	No	
7,525	CERCLA Superfund Site	No	
561	Channelized Stream	Yes	R
17	Combined Sewer	No	
105	Constructed Wetland	Yes	R
68,660	Estuary	Yes	S
792	Facility Industrial	No	
1,759	Facility Municipal Sewage (POTW)	No	
1,867	Facility Other	No	
1,082	Facility Privately Owned Non-industrial	No	
266	Facility Public Water Supply (PWS)	No	
9	Facility: Cistern	No	
146	Facility: Combined sewer	No	
4,485	Facility: Diversion	No	
182	Facility: Field, Pasture, Orchard, or Nursery	No	
114	Facility: Golf course	No	
217	Facility: Laboratory or sample-preparation area	No	
25	Facility: Landfill	No	
1,377	Facility: Outfall	No	
172	Facility: Pavement	No	
49	Facility: Septic system	No	
245	Facility: Storm sewer	No	
93	Facility: Waste injection well	No	
169	Facility: Wastewater land application	No	
105	Facility: Wastewater sewer	No	
254	Facility: Water-distribution system	No	
2,719	Facility: Water-use establishment	No	
3	Floodwater	Yes	R
1	Floodwater non-Urban	Yes	R
2,458	Floodwater Urban	Yes	R
11	Gallery	No	
14	Glacier	No	
844	Great Lake	Yes	L
78,406	Lake	Yes	L
24,752	Lake, Reservoir, Impoundment	Yes	L

5,681	Land	No	
632	Land Flood Plain	No	
800	Land Runoff	No	
275	Land: Excavation	No	
487	Land: Outcrop	No	
1	Land: Playa	No	
41	Land: Shore	No	
456	Land: Sinkhole	No	
985	Land: Soil hole	No	
1	Land: Volcanic vent	No	
347	Landfill	No	
60	Leachate-Lysimeter	No	
3	Local Air Monitoring Station	No	
293	Mine/Mine Discharge	No	
304	Mine/Mine Discharge Adit (Mine Entrance)	No	
228	Mine/Mine Discharge Tailings Pile	No	
148	Mine/Mine Discharge Waste Rock Pile	No	
17,746	Ocean	Yes	S
466	Ocean: Coastal	Yes	S
5,065	Other-Ground Water	No	
2,126	Other-Surface Water	Yes	R
591	Pipe, Unspecified Source	No	
54	Pond-Anchialine	Yes	L
86	Pond-Stormwater	Yes	L
3	Pond-Wastewater	Yes	L
11,833	Reservoir	Yes	L
308,728	River/Stream	Yes	R
41	River/Stream Ephemeral	Yes	R
222	River/Stream Intermittent	Yes	R
1,728	River/Stream Perennial	Yes	R
1,098	Riverine Impoundment	Yes	R
210	Seep	No	
5	Spigot / Faucet	No	
36,684	Spring	Yes	R
5	State/Local Air Monitoring Station	No	
1,032	Storm Sewer	No	
134,125	Stream	Yes	R
2,634	Stream: Canal	Yes	R
2,055	Stream: Ditch	Yes	R
169	Stream: Tidal stream	Yes	R
100	Subsurface	No	
20	Subsurface: Cave	No	
794	Subsurface: Groundwater drain	No	
1,586	Subsurface: Tunnel, shaft, or mine	No	
816	Subsurface: Unsaturated zone	No	
292	Survey Monument	No	

268	Test Pit	No	
61	Waste Pit	No	
31	Waste Sewer	No	
1,427,116	Well	No	
392	Well: Collector or Ranney type well	No	
45	Well: Extensometer well	No	
911	Well: Hyporheic-zone well	No	
643	Well: Interconnected wells	No	
1,195	Well: Multiple wells	No	
26,589	Well: Test hole not completed as a well	No	
675	Wetland	Yes	R
683	Wetland Estuarine-Emergent	Yes	S
30	Wetland Estuarine-Forested	Yes	S
18	Wetland Estuarine-Scrub-Shrub	Yes	S
306	Wetland Lacustrine-Emergent	Yes	R
2,621	Wetland Palustrine-Emergent	Yes	R
57	Wetland Palustrine-Forested	Yes	R
7	Wetland Palustrine-Moss-Lichen	Yes	R
31	Wetland Palustrine-Shrub-Scrub	Yes	R
152	Wetland Riverine-Emergent	Yes	R
644	Wetland Undifferentiated	Yes	R

Figure T. MonitoringLocationType Names Found in WQP Site Data and Decision on Use

The WQI workgroup manually filled in the two rightmost columns (whether the waterbody type should be used and, if "yes", whether the river/stream or lake/reservoir criteria should be applied). Some types were obviously inapplicable because they are not ambient surface water locations, but some – such as estuaries – were excluded for nutrient records over uncertainty of criteria applicability.

Salt and marine water body types were included when pathogen records were added as pathogen criteria exists for these water body types.

Obtaining ANE for Each Nutrient Monitoring Location

Once the WQP Site data have been retrieved, to obtain ecoregion a point-in-polygon search of the appropriate (ANE) shapefiles is performed, and monitoring locations from the survivors of the waterbody-type screen that also had latitude and longitude placing them outside any nutrient ecoregion are assigned to "ecoregion noeco".

The 14 nutrient ecoregions cover only the conterminous U.S., so locations in Alaska, Hawaii, Guam, etc. are "noeco". Offshore monitoring locations with lat/long outside the ecoregions' polygon boundaries are also included among the "noecos". (There has been some interest in considering these locations as if they were inside the ecoregion they are geographically nearest to, but this was not done.)

A modified Site file was created containing:

From the original Site file:

MonitoringLocationIdentifier
MonitoringLocationName

LatitudeMeasure
LongitudeMeasure
CEightDigitCode
StateCode
CountyCode
MonitoringLocationTypeName,

And from the additions made above:

R, L, or S criterion
ANE identifier or “noeco”

This file is called the “tagged sites” file.

Result File Processing

The WQP Result data are retrieved for media=water, starting date, characteristic type and characteristic names desired. This query over selects to prevent inadvertently missing characteristic names wanted but not in a characteristic group.

The WQP Result file has one record for each measurement.

Locations where there are no nutrient measurements after the start date of the period involved are dropped.

A sample is defined as all records with the same value for MonitoringLocationIdentifier, ActivityIdentifier and ActivityStartDate. OrganizationIdentifier was not used as all the MonitoringLocationIdentifiers were unique.

Multiple determinations of the same pollutant on the same day are regarded as quality control exercises and are averaged to produce only one measurement. Multiple determinations are defined as records with the same values for CharacteristicName and ResultSampleFractionText, in addition to being in the same sample.

Depth was not included in the test for duplicate determinations, with the result that determinations differing only by depth in the same sample will be averaged.

For each record

Screen
ActivityMediaName: “Water”

ActivityStartDate: yearbegin-01-01 to yearend-12-31

CharacteristicName After removing “as xxx” or “xxx”, if present as rightmost word(s), one of these values not preceded by the // symbol:

```
                "Ammonia",
                "Ammonia and ammonium",
                "Ammonia-nitrogen",
                "Ammonium",
//              "Ammonium-Nitrogen",
                "Inorganic nitrogen (ammonia, nitrate and nitrite)",
                "Inorganic nitrogen (nitrate and nitrite)",
                "Kjeldahl nitrogen",
                "Nitrate",
//              "Nitrate-Nitrite",
//              "Nitrate-Nitrogen",
```

"Nitrite",
"Nitrogen",
"Nitrogen Kjeldahl",
// "Nitrogen, ammonia (NH3)",
"Nitrogen, mixed forms (NH3), (NH4), organic, (NO2) and (NO3)",
// "Nutrient-nitrogen",
"Organic Nitrogen",
"Organic nitrogen"

For nitrogen, the // names do not appear in the SRS substance search and were not used, although they show up in the data (about 65,000 samples). It was not clear how these measurements would figure in the total N derivation.

// "Inorganic phosphorus",
// "Organic phosphorus",
// "Ortho-Phosphate-Phosphorus",
// "Orthophosphate",
// "Phosphate",
"Phosphate-phosphorus",
"Phosphorus",
// "Polyphosphate",
// "Soluble Reactive Phosphorus (SRP)"

For phosphorus, the // names cannot be used in combination to derive total P.

CharacteristicName specifies the pollutant. The Characteristic names shown above were determined to be relevant for identifying all the N and P data needed.

Vetting the measurement

Abbreviations
AE ResultDetectionConditionText
AH ResultMeasureValue
AI ResultMeasure/MeasureUnitCode (units1 and units2)
BG DetectionQuantitationLimitTypeName
BH DetectionQuantitationLimitTypeMeasure/MeasureValue
BI DetectionQuantitationLimitTypeMeasure/MeasureUnitCode(units1 and units2)
BR MethodSpecificationName (units2)

Units1 is the concentration, units2 is the reference species (e.g., in “2 mg/l as N”, units1=”mg/l” and units2=”N”).

Check ResultDetectionConditionText(AE)

ResultDetectionConditionText(AE)	Action
*Present >QL	Automatic Hotspot flag
Present Above Quantification Limit	Automatic Hotspot flag
*Present	If AH is null then don’t use measurement
Detected Not Quantified	If AH is null then don’t use measurement
*Not Reported	Don’t use measurement
Not Reported	Don’t use measurement
*<	Use detection limit protocols
*Non-detect	Use detection limit protocols

*Not Detected	Use detection limit protocols
*Present <QL	Use detection limit protocols
<	Use detection limit protocols
ND	Use detection limit protocols
Not Detected	Use detection limit protocols
Present Below Quantification Limit	Use detection limit protocols
Any other text or no text	Not a detection limit

There are two, possibly different, detection limit protocols, one for combining subspecies ("subspecies detection limit protocol") and the other for determining summary statistics ("statistics detection limits protocol").

Definition of a valid detection limit measurement:
If AE indicates measurement is a detection limit, AH is null and BH and BI (units1 and units2) are present, use BH/BI as the detection limit. If BI units2 is null, seek units2 at the end of Characteristic Name. If units2 is still not determined, look for a units2 in the table in S**ubspecies Combination Protocol – Nitrogen** (infra). If units2 remains undetermined, look for units2 in BR. If either units1 or units2 remain(s) undetermined, or if BG is "Drinking Water Maximum" or "Water Quality Standard or Criteria", drop the measurement.

BG DetectionQuantitationLimitTypeName
Drinking Water Maximum
Estimated Detection Level
Historical Lower Reporting Limit
Instrument Detection Level
Laboratory Reporting Level
Long Term Method Detection Level
Lower Quantitation Limit
Lower Reporting Limit
Method Detection Level
Method Detection Limit (MDL)
Practical Quantitation Limit
Sample-specific min detect conc
Upper Quantitation Limit
Upper Reporting Limit
Water Quality Standard or Criteria

Definition of a valid quantification measurement:
If AE is null, AH is present, and AI (units1) [preferred] or BI (units1) are present, and units2 is present in (preferred order) AI, Characteristic Name or BR, the record is used (and is not a detection limit). Otherwise, drop the measurement.

Convert nitrogen concentration to mg/l as N and phosphorus concentration to ug/l as P, if not already in these units.

Continue until all N and P species in the sample have been processed.

Determination of Total N and/or Total P

If Total N as N is present, it is used.

	Characteristic	Fraction	Speciation	Note	Comment
Total N	Nitrogen, mixed forms (NH3), (NH4), organic, (NO2) and (NO3)	Total	?	Total N, No Combo Needed	Assumed as N
	Nitrogen	Total	NS		
	Nitrogen	Total	N		
	Nitrogen	Total	?		Assumed as N

Otherwise
Derive values for total N from N subspecies, if possible, as follows:

Subspecies Combination Protocol – Nitrogen

The subspecies in the following table found in the Result file have been identified as combinable:

Row	Characteristic Name	Fraction	Species	Remarks	
1	Kjeldahl Nitrogen	Total	N		
1	Nitrogen Kjeldahl	Total	N		
2	Ammonia-nitrogen	Total	N		
3	Ammonia-nitrogen	Total	?	Assume as N	
4	Ammonia and ammonium	Total	N		
5	Ammonia	Total	N		
6	Nitrate	Total	N		
7	Nitrogen	Dissolved	?	Assume as N	
8	Nitrogen, mixed forms (NH3), (NH4), organic, (NO2) and (NO3)	Dissolved	?	Assume as N	
9	Kjeldahl Nitrogen	Dissolved	N		
9	Nitrogen Kjeldahl	Dissolved	N		
10, 11	Ammonia	Dissolved	N		
12	Ammonium	Filterable	NH4	Convert to as N	
13	Ammonium	Filterable	N		
14	Inorganic Nitrogen (Nitrate & nitrite)	Dissolved	N		
15	Ammonia and ammonium	Dissolved	N		
16	Nitrate	Filterable	NO3	Convert to as N	
17	Nitrate	Dissolved	N		
18	Nitrite	Dissolved	N		
19	Organic Nitrogen	Dissolved	?	Assume as N	
20	Nitrogen	Suspended	?	Assume as N	
		Filterable=Dissolved ?=Not Specified			

The combinations yielding total N are these:

7 + 20

8 + 20
9 + 14 + 20
4 + 6 + 18 + 19 + 20
1 + 6 + 18
2/3/5 + 6 + 18 + 19 + 20
4 + 6 + 18 + 19 + 20
10/11 + 14 + 12/13 + 19 + 20
15 + 16/17 + 18 + 19 + 20
1 + 14
1 + 6/16/17 + 18

If Phosphorus or Phosphate-phosphorus is present, use this value as total P.

	Characteristic	Fraction	Speciation	Note	Comment
Complete	Phosphorus	Total	as P	Total P. No conversion needed.	
	Phosphorus	Total	NS		
	Phosphorus	Total ?	NS		
	Phosphate-phosphorus	Total	as P		
	Phosphate-phosphorus	Total	NS		As P assumed

A review of the phosphorus subspecies available in the WQP data did not identify any combination of them which would produce total P.

	Phosphorus	Dissolved	as P	Dissolved/Soluble P fraction only. Needs Suspended/Particulate P	
		NS	NS		
		NS	as P		
	Phosphate-phosphorus	Dissolved	NS	Dissolved/Soluble P fraction only. Needs Suspended/Particulate P	
		Dissolved	as P		
		NS	NS		
		NS	as P		
	Phosphate	Dissolved	as P	Dissolved/Soluble P fraction only. Needs Suspended/Particulate P	
			NS		Assumed reported as PO4, may require conversion factor of 0.33 to P
	Orthophosphate	Total	as P	Total Inorganic P. Needs Organic Fraction	
		Dissolved	NS	Dissolved/ Soluble Inorganic P only. Less	May need conversion if

			than dissolved P fraction	reported as PO4
	Dissolved	as P		
	NS	as P	Not Specified. Inorganic P. Could be Total or Dissolved orthophosphate but not Total P.	
	NS	as PO4		Need conversion of value with 0.33 to P
	NS	NS		

Figure P. Combination of available P subspecies are insufficient to derive Total P.

There may be additional, less obvious, combinations of fraction, subspecies and units in the WQP data which were overlooked in the initial effort to identify them.

Non-Detect Handling

When Combining Species to Arrive at Total N

Deriving total N from N subspecies requires adding together a number of subspecies values, some of which may be detection limits.

When Deriving Summary Statistics

A summary statistic (e.g. maximum, percentile) requires a number of sample values be reduced to a single number. Some of these values may be detection limits.

Measurements below the detection limit (DL) of the measurement method, often referred to as "non-detects", are typically reported as "<DL". But how to combine or summarize non-detects and detections is a choice made anew by every investigator. It seems to be widely recognized that substitution methods (removing the "<" and using the DL, zero, ½ DL or a random value between zero and the DL) are inadequate but are still in use.

Studies cited above determined that simple substitution methods performed poorly in comparison to other procedures. Substitution of zero produced estimates of mean and median which were biased low, while substituting the reporting limit resulted in estimates above the true value. Results for the standard deviation and IQR, and for substituting one-half the reporting limit, were also far less desirable than alternative methods. With the advent of powerful desktop computers to perform more complex calculations there appears to be no reason to use simple substitution methods. As the choice of value to be substituted is essentially arbitrary without some knowledge of instrument readings below the reporting limit, and as large differences may occur in the resulting estimates, simple substitution methods are not defensible (Helsel and Hirst, USGS, 2002).

Antweiler and Taylor have this to say:
The main classes of statistical treatment of below-detection limit (left-censored) environmental data for the determination of basic statistics that have been used in the literature are substitution methods, maximum likelihood, regression on order statistics (ROS), and nonparametric techniques. These treatments, along with using all instrument-generated data (even those below detection), were evaluated by examining data sets in which the true values of the censored data were known. It was found that for data sets with less than 70%censored data, the best technique overall for determination of summary statistics was the nonparametric Kaplan–Meier technique. ROS and the two substitution methods of assigning one-half the detection limit value to censored data or assigning a random number between zero and the detection limit to censored data were adequate alternatives. The use of these two substitution methods, however, requires a thorough understanding of how the laboratory censored the data. The technique of employing all instrument-generated data--including numbers below the detection limit--was found to be less adequate than the above techniques. At high degrees of censoring (greater than 70% censored data), no technique provided good estimates of

summary statistics. Maximum likelihood techniques were found to be far inferior to all other treatments except substituting zero or the detection limit value to censored data.

These and other studies argue persuasively that Kaplan-Meir or ROS should be used for non-detects; however, some time would be required to study the implementation and implications of these methods, and, in the interest of arriving at an initial version of the WQI product ASAP, the time-honored but non-defensible substitution of ½ the censored value(BH) as the actual value when computing the statistics was used. Perhaps the next version of the WQI file will do better. A consequence of this choice is that the DL is removed as soon as the measurement is read and so there is no need for methods for combining non-detects in this scenario because there are no non-detects to combine.

Statistics, Addition of Site file data and Output

At this point all the (usable) values for N or P for date range yearbegin to yearend for a monitoring location are available, statistics for the location (median, max, etc.) are computed, the "tagged sites" file data for the location are prepended to form the entire output record (see **V** supra), and the record is written.

The statistics chosen are conventional (maximum, various percentiles including median, but not interquartile range), and exploratory (the time-weighted percentiles). Time weighting applied a weight of w=(yearend-yearbegin+1)*12 to Dec yearend, w-1 to Nov yearend,…,1 to Jan yearbegin. The problem was how to avoid giving as much visibility to older data as to newer, unless, of course, equal visibility was desired. The weighting method is still open for discussion.

The presence of only a small number of samples can produce artificially large differences in percentiles.

Processing Summary and Data Usability

A processing summary for the most recent WQI refresh is available via the WQI Data Usability Improvement Project Dashboard. Via this dashboard, you can identify metadata deficiencies that limit the WQI's ability to use the data.

Original Format

MonitoringLocationIdentifier
OrganizationFormalName(OrganizationIdentifier)
MonitoringLocationName
LatitudeMeasure
LongitudeMeasure
CEightDigitCode
StateCode
CountyCode
MonitoringLocationTypeName,

These elements derived by WQI Site file protocols:
WaterBodyType ("R"[iver/Stream] or "L"[ake/Reservoir])
Ecoregion Identifier (I, II, III…XIV,noeco)

These elements derived by Result file processing:
Up to 10 groups of 12-values per group. If the first value in a group is:

N: The pollutant is nitrogen and the interval for the summary statistics is the WQI entire period of record, which has a length in months of (yearend-yearbegin+1)*12 (from Jan yearbegin to Dec yearend).

N48: The pollutant is nitrogen and the interval for the summary statistics is the 48-month period ending on the date of the WQP retrieval. (The entire period of record is assumed to be greater than 48 months.)

P: The pollutant is phosphorus and the interval for the summary statistics is the WQI entire period of record, which has a length in months of (yearend-yearbegin+1)*12 (from Jan yearbegin to Dec yearend).

P48: The pollutant is phosphorus and the interval for the summary statistics is the 48-month period ending on the date of the WQP retrieval. (The entire period of record is assumed to be greater than 48 months.)

Each time one of the above markers (N, N48, P, P48) occurs, it will be followed by the following eleven fields:

Number of samples
Number of samples rejected because of incomplete metadata
Samples maximum
Samples median
Samples time-weighted median
Samples 70th percentile
Samples time-weighted 70th percentile
Samples 90th percentile
Samples time-weighted 90th percentile
Number of sample values greater than the default 304(a) criteria
Maximum of monthly medians

When a month (yymm) appears instead of a marker as the first value in a group, the month will be followed by the first ten of these eleven fields and the period of record is the month. ("Maximum of monthly medians" does not apply to monthly data)

Only the months with data appear in the record. The sequence of fields in the record is N, N48, P, P48, months in ascending yymm order.

Here is an example record with 44 used and 0 rejected N samples, and 55 used and 0 rejected P samples (the white space added for clarity is not in the record):

21ARIZ_WQX-SRTON053.87,ARIZONA DEPARTMENT OF ENVIRONMENTAL QUALITY(21ARIZ_WQX),TONTO CREEK - BELOW BEAR FLATS,34.2833333,-111.0708333,15060105,04,007,River/Stream,R,II,

N,44,0,1.000,0.370,0.370,0.433,0.430,0.575,0.575,43,0.520,N48,22,0,0.910,0.360,0.350,0.410,0.410,0.460,0.460,21,0.412,1305,5,0,0.700,0.520,0.520,0.564,0.575,0.650,0.700,5,1306,6,0,1.000,0.375,0.430,0.502,0.575,0.788,1.000,6,1307,5,0,0.540,0.400,0.400,0.456,0.470,0.512,0.540,5,1308,6,0,0.550,0.317,0.325,0.347,0.370,0.460,0.550,6,1309,6,0,0.910,0.412,0.430,0.445,0.460,0.685,0.910,6,1407,5,0,0.785,0.410,0.410,0.450,0.460,0.655,0.785,5,1408,6,0,0.370,0.265,0.270,0.285,0.300,0.335,0.370,5,1409,5,0,0.420,0.290,0.290,0.314,0.320,0.380,0.420,5,

P,55,0,4100.000,50.000,50.000,85.800,92.000,153.000,155.000,55,2125.000,P48,22,0,230.000,91.000,90.000,130.000,130.000,154.500,155.000,22,127.250,0903,1,0,17.500,17.500,17.500,17.500,17.500,17.500,17.500,1,0904,2,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,2,0908,1,0,78.000,78.000,78.000,78.000,78.000,78.000,78.000,1,0909,1,0,77.000,77.000,77.000,77.000,77.000,77.000,77.000,1,1007,2,0,4100.000,2125.000,4100.000,2915.000,4100.000,3705.000,4100.000,2,1008,1,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,1,1009,1,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,25.000,1,1106,1,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,1,1107,1,0,15.000,15.000,15.000,15.000,15.000,15.000,15.000,1,1109,2,0,50.000,50.000,50.000,50.000,50.000,50.000,50.000,2,1305,4,0,36.000,32.500,34.000,34.200,34.000,35.400,36.000,4,1306,5,0,35.000,32.500,32.500,32.900,33.000,34.200,35.000,5,1307,5,0,500.000,49.000,49.000,361.800,440.000,476.000,500.000,5,1308,6,0,87.000,59.000,68.000,73.750,79.500,83.250,87.000,6,1309,6,0,230.000,127.250,150.000,152.500,155.000,192.500,230.000,6,1407,5,0,150.000,50.000,50.000,130.000,150.000,150.000,150.000,5,1408,6,0,190.000,88.000,120.000,125.000,130.000,160.000,190.000,6,1409,5,0,130.000,90.000,90.000,91.600,92.000,114.800,130.000,5

Rearranging the repeating portion of the record above for improved readability:

N,44,0,1.000,0.370,0.370,0.433,0.430,0.575,0.575,43,0.520,
N48,22,0,0.910,0.360,0.350,0.410,0.410,0.460,0.460,21,0.412,
1305,5,0,0.700,0.520,0.520,0.564,0.575,0.650,0.700,5,
1306,6,0,1.000,0.375,0.430,0.502,0.575,0.788,1.000,6,
1307,5,0,0.540,0.400,0.400,0.456,0.470,0.512,0.540,5,
1308,6,0,0.550,0.317,0.325,0.347,0.370,0.460,0.550,6,
1309,6,0,0.910,0.412,0.430,0.445,0.460,0.685,0.910,6,
1407,5,0,0.785,0.410,0.410,0.450,0.460,0.655,0.785,5,
1408,6,0,0.370,0.265,0.270,0.285,0.300,0.335,0.370,5,
1409,5,0,0.420,0.290,0.290,0.314,0.320,0.380,0.420,5,

P,55,0,4100.000,50.000,50.000,85.800,92.000,153.000,155.000,55,2125.000,
P48,22,0,230.000,91.000,90.000,130.000,130.000,154.500,155.000,22,127.250,
0903,1,0,17.500,17.500,17.500,17.500,17.500,17.500,17.500,1,
0904,2,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,2,
0908,1,0,78.000,78.000,78.000,78.000,78.000,78.000,78.000,1,
0909,1,0,77.000,77.000,77.000,77.000,77.000,77.000,77.000,1,
1007,2,0,4100.000,2125.000,4100.000,2915.000,4100.000,3705.000,4100.000,2,
1008,1,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,1,
1009,1,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,1,
1106,1,0,25.000,25.000,25.000,25.000,25.000,25.000,25.000,1,
1107,1,0,15.000,15.000,15.000,15.000,15.000,15.000,15.000,1,
1109,2,0,50.000,50.000,50.000,50.000,50.000,50.000,50.000,2,
1305,4,0,36.000,32.500,34.000,34.200,34.000,35.400,36.000,4,
1306,5,0,35.000,32.500,32.500,32.900,33.000,34.200,35.000,5,
1307,5,0,500.000,49.000,49.000,361.800,440.000,476.000,500.000,5,
1308,6,0,87.000,59.000,68.000,73.750,79.500,83.250,87.000,6,
1309,6,0,230.000,127.250,150.000,152.500,155.000,192.500,230.000,6,
1407,5,0,150.000,50.000,50.000,130.000,150.000,150.000,150.000,5,
1408,6,0,190.000,88.000,120.000,125.000,130.000,160.000,190.000,6,
1409,5,0,130.000,90.000,90.000,91.600,92.000,114.800,130.000,5

Doing a similar exercise with the Pathogen results for E. Coli:

Original Record:

11NPSWRD_WQX-GUCO_HFSP,National Park Service Water Resources Division(11NPSWRD_WQX),Hoskins Farm Spring,36.1291860000,-79.8525990000,03030002,37,081,Spring,R,IX,E. Coli,10,3,461.100,18.450,25.900,173.540,203.500,248.970,461.100,4,488.400,E. Coli48,1,0,461.100,461.100,461.100,461.100,461.100,461.100,461.100,1,461.100,1110,2,0,11.000,10.950,11.000,10.970,11.000,10.990,11.000,0,1201,1,0,160.700,160.700,160.700,160.700,160.700,160.700,160.700,1,1204,2,0,6.300,4.150,6.300,5.010,6.300,5.870,6.300,0,1207,1,0,2.000,2.000,2.000,2.000,2.000,2.000,2.000,0,1310,1,0,225.400,225.400,225.400,225.400,225.400,225.400,225.400,1,1401,1,0,203.500,203.500,203.500,203.500,203.500,203.500,203.500,1,1404,0,2,,,,,,,,1407,0,1,,,,,,,,1611,1,0,25.900,25.900,25.900,25.900,25.900,25.900,25.900,0,1708,1,0,461.100,461.100,461.100,461.100,461.100,461.100,461.100,1

Separated For Readability:

11NPSWRD_WQX-GUCO_HFSP,National Park Service Water Resources Division(11NPSWRD_WQX),Hoskins Farm Spring,36.1291860000,-79.8525990000,03030002,37,081,Spring,R,IX,

E. Coli,10,3,461.100,18.450,25.900,173.540,203.500,248.970,461.100,4,488.400,
E. Coli48,1,0,461.100,461.100,461.100,461.100,461.100,461.100,461.100,1,461.100,
1110,2,0,11.000,10.950,11.000,10.970,11.000,10.990,11.000,0,
1201,1,0,160.700,160.700,160.700,160.700,160.700,160.700,160.700,1,
1204,2,0,6.300,4.150,6.300,5.010,6.300,5.870,6.300,0,
1207,1,0,2.000,2.000,2.000,2.000,2.000,2.000,2.000,0,
1310,1,0,225.400,225.400,225.400,225.400,225.400,225.400,225.400,1,
1401,1,0,203.500,203.500,203.500,203.500,203.500,203.500,203.500,1,
1404,0,2,,,,,,,,
1407,0,1,,,,,,,,
1611,1,0,25.900,25.900,25.900,25.900,25.900,25.900,25.900,0,
1708,1,0,461.100,461.100,461.100,461.100,461.100,461.100,461.100,1

Appendix A: Alternate Data Format

While the original format works well with the WQI interface, it is not practical for 3rd party downloading and data analysis. An alternate format was created which separates the data into two tables.

RFEOutput2_v24_MonLocInfo (or **Table v24**)
RFEOutput2_v22_RRecs (or **Table v22**)

Table v24 contains all the non-repeating Monitoring Location Information:

MonitoringLocationIdentifier (key)
OrganizationFormalName
MonitoringLocationName
LatitudeMeasure
LongitudeMeasure
CEightDigitCode
StateCode
CountyCode
MonitoringLocationTypeName
WaterBodyType
EcoregionIdentifier

Table v22 contains repeating records of monitoring data keyed by concatenating **MonitoringLocationIdentifier Pollutant YYYYMM**:

MonitoringLocationIdentifier (key)
Pollutant (key)
48mo flag indicating if record falls into the 48-month period
YYYYMM (key) a zero value represents a summary record
Samples
Rejected
Maximum
Median
MedianW
70%
70%W
90%
90%W
Crits-304a
MaxMedian only populated in a summary record (when YYYYMM is 0)), otherwise is null

Sample:

Table v24

MonitoringLocationIdentifier	OrganizationFormalName	MonitoringLocationName	LatitudeMeasure	LongitudeMeasure	CEightDigitCode	StateCode	CountyCode	MonitoringLocatio	WaterBody	EcoregionIdentifier
1119USBR_WQX-BLA101	ureau of Reclamation(1	Payette River at Montor	43.9323	-116.3356	17050122	16	45	River/Stream	R	III
1119USBR_WQX-CBP017	Bureau of Reclamation(	PE16.4 WW AT HENDRIC	46.6727778	-119.15	17020016	53	21	Canal Irrigation	R	III
1119USBR_WQX-CBP029	Bureau of Reclamation(	PE16.4 WW AT COLUMB	46.5055556	-119.2591667	17020016	53	21	Canal Irrigation	R	III
1119USBR_WQX-CBP052	Bureau of Reclamation(	ESQUATZEL DIV CHNL A	46.3586111	-119.2566667	17020016	53	21	Canal Irrigation	R	III

Table v22

MonitoringLocationIdentifier	Pollutant	48mo	YYYYMM	Samples	Rejected	Maximum	Median	MedianW	70%	70%W	90%	90%W	Crits-304a	MaxMedian
1119USBR_WQX-BLA101	N		0	13	1	1.9	0.22	0.22	0.239	0.24	0.328	0.34	1	1.9
1119USBR_WQX-BLA101	N	N48	0	13	1	1.9	0.22	0.22	0.239	0.24	0.328	0.34	1	1.9
1119USBR_WQX-BLA101	N	N48	201611	1	1	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0	
1119USBR_WQX-BLA101	N	N48	201709	1	0	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0	
1119USBR_WQX-BLA101	N	N48	201801	3	0	0.18	0.17	0.17	0.174	0.18	0.178	0.18	0	
1119USBR_WQX-BLA101	N	N48	201803	1	0	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0	
1119USBR_WQX-BLA101	N	N48	201805	1	0	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0	
1119USBR_WQX-BLA101	N	N48	201806	1	0	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1	
1119USBR_WQX-BLA101	N	N48	201807	1	0	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0	
1119USBR_WQX-BLA101	N	N48	201808	1	0	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0	
1119USBR_WQX-BLA101	N	N48	201809	1	0	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0	
1119USBR_WQX-BLA101	N	N48	201810	1	0	0.217	0.217	0.217	0.217	0.217	0.217	0.217	0	
1119USBR_WQX-BLA101	N	N48	201811	1	0	0.239	0.239	0.239	0.239	0.239	0.239	0.239	0	
1119USBR_WQX-BLA101	P		0	13	0	500	17	17	19	19	33	34	3	500
1119USBR_WQX-BLA101	P	P48	0	13	0	500	17	17	19	19	33	34	3	500
1119USBR_WQX-BLA101	P	P48	201611	1	0	19	19	19	19	19	19	19	0	
1119USBR_WQX-BLA101	P	P48	201709	1	0	19	19	19	19	19	19	19	0	
1119USBR_WQX-BLA101	P	P48	201801	3	0	17	16	16	16.4	17	16.8	17	0	
1119USBR_WQX-BLA101	P	P48	201803	1	0	29	29	29	29	29	29	29	1	
1119USBR_WQX-BLA101	P	P48	201805	1	0	34	34	34	34	34	34	34	1	
1119USBR_WQX-BLA101	P	P48	201806	1	0	17	17	17	17	17	17	17	0	
1119USBR_WQX-BLA101	P	P48	201807	1	0	14	14	14	14	14	14	14	0	
1119USBR_WQX-BLA101	P	P48	201808	1	0	16	16	16	16	16	16	16	0	
1119USBR_WQX-BLA101	P	P48	201809	1	0	15	15	15	15	15	15	15	0	
1119USBR_WQX-BLA101	P	P48	201810	1	0	500	500	500	500	500	500	500	1	
1119USBR_WQX-BLA101	P	P48	201811	1	0	11	11	11	11	11	11	11	0	

Here is another image of the Table v22 with Enterococcus

MonitoringLocationIdentifier	Pollutant	48mo	YYYYMM	Samples	Rejected	Maximum	Median	MedianW	70%	70%W	90%	90%W	Crits-304a	MaxMedian
1083919_WQX-26.1W	Enterococcus		0	2	0	209	204	209	206	209	208	209	2	209
1083919_WQX-26.1W	Enterococcus	Enterococcus48	0	2	0	209	204	209	206	209	208	209	2	209
1083919_WQX-26.1W	Enterococcus	Enterococcus48	201707	1	0	199	199	199	199	199	199	199	1	
1083919_WQX-26.1W	Enterococcus	Enterococcus48	201710	1	0	209	209	209	209	209	209	209	1	
1083919_WQX-26.2W	Enterococcus		0	3	0	75	63	63	67.8	75	72.6	75	1	75
1083919_WQX-26.2W	Enterococcus	Enterococcus48	0	3	0	75	63	63	67.8	75	72.6	75	1	75
1083919_WQX-26.2W	Enterococcus	Enterococcus48	201707	1	0	63	63	63	63	63	63	63	0	
1083919_WQX-26.2W	Enterococcus	Enterococcus48	201708	1	0	75	75	75	75	75	75	75	1	
1083919_WQX-26.2W	Enterococcus	Enterococcus48	201710	1	0	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0	
1083919_WQX-28E	Enterococcus		0	2	0	2489	1244.523	0.045	1742.314	2489	2240.104	2489	1	2489
1083919_WQX-28E	Enterococcus	Enterococcus48	0	2	0	2489	1244.523	0.045	1742.314	2489	2240.104	2489	1	2489
1083919_WQX-28E	Enterococcus	Enterococcus48	201706	1	0	2489	2489	2489	2489	2489	2489	2489	1	
1083919_WQX-28E	Enterococcus	Enterococcus48	201710	1	0	0.045	0.045	0.045	0.045	0.045	0.045	0.045	0	
1083919_WQX-BR-GS-1	Enterococcus		0	1	0	4884	4884	4884	4884	4884	4884	4884	1	4884
1083919_WQX-BR-GS-1	Enterococcus	Enterococcus48	0	1	0	4884	4884	4884	4884	4884	4884	4884	1	4884
1083919_WQX-BR-GS-1	Enterococcus	Enterococcus48	201709	1	0	4884	4884	4884	4884	4884	4884	4884	1	

Notes:

Summary record for each Monitoring ID, Pollutant, 48 month combination is noted by the YYYYMM = 0.

MaxMedian only occurs in Summary record. Currently each monitoring location ID may have up to four summary records.