Water Quality Portal Data Discovery Tool - Adventures in Data Discovery!

2016 National Monitoring Conference
May 4, 2016

Presenter: Amy Wesley-Snider (RTI)

Co-authors: Danette Boezio, Peter Ilieve, Sam Sifleet, Marwa Salem (RTI) Dwane Young (EPA)
Presentation Outline

- Background
- WQP Data Discovery tool
- Future Adventures!
  - WQP Data Analysis Tool
- Conclusions
Background: Two years ago the adventure began…


2014 National Monitoring Conference
April 28 – May 2, 2014

Presenter: Amy Wesley-Snider (RTI)

Co-authors: Danette Boezio, Jay Rineer, Breda Munoz, Peter Ilieve, Michele Eddy, Phillip Jones (RTI)

RTI International is a trade name of Research Triangle Institute

3040 Cornwallis Road  Research Triangle Park, North Carolina, USA 27709

www.rti.org
WQP Data Discovery Tool

- Desktop Application
  - Opens in browser (IE, Chrome, Firefox)
- Query, Summarize, QC, Display data from the WQP
  - Assists users in selecting data for analysis
- R – Open source statistical programming language.
WQP Data Discovery Tool – Query Data

Select filters to build your data query

Click here to expand and close a panel.

Location

Sampling Parameters

Site Parameters

Click here to run the query.

Retrieve Data

WQP Web Service Query URL

http://www.waterqualitydata.us/Result/search?siteid=mimeType=tsv&sorted=no
WQP Data Discovery Tool – Query Data

Select filters to build your data query

Location

<table>
<thead>
<tr>
<th>Place</th>
<th>Point Location</th>
<th>Bounding Box</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose a State</td>
<td>Latitude (decimal degrees)</td>
<td>North</td>
</tr>
<tr>
<td></td>
<td>Longitude (decimal degrees)</td>
<td>South</td>
</tr>
<tr>
<td>Choose a County</td>
<td>Distance from point (miles)</td>
<td>East</td>
</tr>
<tr>
<td>HUC</td>
<td></td>
<td>West</td>
</tr>
</tbody>
</table>

Enter a huc code (ex: 03020201)
WQP Data Discovery Tool – Query Data

Select filters to build your data query

Location

Sampling Parameters

Date Range
(MM-DD-YYYY)
From
To

Sample Media
Characteristic Group
Characteristic
Select filters to build your data query

Location

Sampling Parameters

Site Parameters

Site Type

Organization

Site ID
**WQP Web Service Query URL.** At the base of the 'Query Data' screen the web service url is displayed. As you apply and modify query parameters on this screen the query url will automatically update.

WQP Web Service Query URL

http://www.waterqualitydata.us/Result/search?siteid=&mimeType=tsv&sorted=no
WQP Data Discovery Tool – Query Data

Your query returns 472 records from 7 stations.

You may import your data

Click here to Import Data.

Data Import Complete

Click close and proceed to the Check Data Tab.
WQP Data Discovery Tool – Check Data

Data Summary
The total imported data set contains 472 records from 7 stations representing 11 parameters with, 118 non-detects. 3 records without units. 0 records without methods. 1 duplicate records.

Date: 2016-03-01 15:12:55

Web Query:
http://www.waterqualitydata.us/Result/search?organization=YAN_WQP&siteId=8&uc=8&mimeType=tsv&sorted=mo

Please select a method to deal with Non-Detections
- Ignore Non-Detections - remove from data set
- Set Non-Detections equal to zero
- Set Non-Detections equal to the Limit of Detection
- Set Non-Detections equal to the 1/2 times the Limit of Detection

Available Data Sets
All Data:
This table displays all of the raw data records imported from the Water Quality Portal

Non Detects:
These are the records with values for the 'Result Detection Condition Text' field equal to 'Not Detected' or 'Present below Quantitation Limit'.

W/O Units:
These data records have no data entered in either the 'Result Measure - Measure Unit Code' or the 'Quantitation Limit Measure - Measure Unit Code' fields.
All Imported Data Records

This table displays all of the raw data records imported from the Water Quality Portal.

<table>
<thead>
<tr>
<th>Station</th>
<th>Name</th>
<th>ActivityStartDate</th>
<th>Characteristic</th>
<th>Result</th>
<th>ActivityMediaSubdivisionName</th>
<th>OrganizationFormalName</th>
<th>ActivityTypeCode</th>
</tr>
</thead>
<tbody>
<tr>
<td>YAN_WQX-BCR06</td>
<td>Beaver Creek</td>
<td>2013-05-13</td>
<td>Conductivity</td>
<td>488.20</td>
<td>Yavapai-Apache Nation</td>
<td>Sample-Routine</td>
<td></td>
</tr>
<tr>
<td>YAN_WQX-BCR06</td>
<td>Beaver Creek</td>
<td>2013-08-27</td>
<td>Conductivity</td>
<td>561.00</td>
<td>Yavapai-Apache Nation</td>
<td>Sample-Routine</td>
<td></td>
</tr>
<tr>
<td>YAN_WQX-BCR06</td>
<td>Beaver Creek</td>
<td>2013-10-29</td>
<td>Conductivity</td>
<td>459.10</td>
<td>Yavapai-Apache Nation</td>
<td>Sample-Routine</td>
<td></td>
</tr>
</tbody>
</table>
WQP Data Discovery Tool – Check Data

Click the button below to run a summary of the data

Click here to summarize data

This table shows summary statistics of all unique combinations of station, media, characteristic, unit, and sample fraction.
Data Filters. This panel provides users with several options for refining their data set. These filters drive the data available for display in the Interactive Map, table, station summary and Parameter/Unit Summary tabs.
WQP Data Discovery Tool – View Data

Map Displays Filtered Data

Please select a station on the map

Station currently selected:
Beaver Creek

Latitude 34.64019 Longitude -111.7819
Station ID YAN_WQX-BCR06, Station Name Beaver Creek, 7 sample(s), 43 results

* To populate the tabs ‘Station Summary’ and ‘Parameter/Unit Summary’, please select a station using the panel located on the map.
<table>
<thead>
<tr>
<th>Station</th>
<th>Name</th>
<th>Organization</th>
<th>Characteristic</th>
<th>Result</th>
<th>Activity/StartDate</th>
</tr>
</thead>
<tbody>
<tr>
<td>YAN_WQX-BCR06</td>
<td>Beaver Creek</td>
<td>YAN_WQX</td>
<td>pH</td>
<td>8.2</td>
<td>2013-02-26</td>
</tr>
<tr>
<td>YAN_WQX-BCR06</td>
<td>Beaver Creek</td>
<td>YAN_WQX</td>
<td>Dissolved oxygen (DO)</td>
<td>10.11</td>
<td>2013-02-26</td>
</tr>
<tr>
<td>YAN_WQX-BCR06</td>
<td>Beaver Creek</td>
<td>YAN_WQX</td>
<td>Temperature, water</td>
<td>9.9</td>
<td>2013-02-26</td>
</tr>
<tr>
<td>YAN_WQX-BCR06</td>
<td>Beaver Creek</td>
<td>YAN_WQX</td>
<td>Escherichia coli</td>
<td>13.4</td>
<td>2013-02-26</td>
</tr>
<tr>
<td>YAN_WQX-BCR06</td>
<td>Beaver Creek</td>
<td>YAN_WQX</td>
<td>Temperature, water</td>
<td>18.8</td>
<td>2013-05-13</td>
</tr>
</tbody>
</table>
WQP Data Discovery Tool – View Data

Summary for station Verde River Downstream
There are 35 records representing 10 characteristics at this station.

Sampling Frequency
Select parameters to view date range and frequency

- pH
- Suspended Sediment Concentration (SSC)
- Dissolved oxygen (DO)
- Temperature, water
- Escherichia coli
- Kjeldahl nitrogen
- Phosphorus
- Nitrogen
- Conductivity
- Inorganic nitrogen (nitrate and nitrite)

To populate the tabs ‘Station Summary’ and ‘Parameter/Unit Summary’, please select a station using the panel located on the map.
WQP Data Discovery Tool – View Data

Please select up to 3 parameter/unit combinations to plot

First Parameter/Unit Choice
- Temperature, water (deg C)

Second Parameter/Unit Choice
- Dissolved oxygen (DO) (mg/l)

Third Parameter/Unit Choice
- Nitrogen (mg/l)

* To populate the tabs 'Station Summary' and 'Parameter/Unit Summary', please select a station using the panel located on the map.

Click here to Print or download the graphic
WQP Data Discovery Tool Help

This tool imports data from the Water Quality Portal WQP using Web Services.

The WQP Data Discovery tool is a desktop application that provides an easy to use interface allowing users to query, summarize, QC and display data from the WQP. The tool uses open source R, a statistical programming language and several add-on packages, to visualize the data selected from the portal and to assist users in selecting data for analysis.

Click on an expandable panel below to view tool tips and troubleshooting guidance.

- Query Data
- Check Data
- View Data
- Troubleshooting
WQP Data Discovery Tool – Available Now!

Future Adventures!
WQP Data Analysis Tool

- Currently under development
- Works with the output from the WQP Data Discovery Tool
- Allow users to compare a water quality parameter value against an established criteria value.
Welcome. Please upload data.

You've successfully imported 40356 records downloaded from the WQP on 4/14/2016 11:17 using the url http://www.waterqualitydata.us/Result/search?countycode=US%3A37%3A183&siteid=&huc=&startDateLo=01-01-2010&startDateHi=04-14-2016&mimeType=tsv&sorted=no. This is the Filtered data with Non-Detections removed from data set.

<table>
<thead>
<tr>
<th>Station</th>
<th>Organization</th>
<th>Organization/FormatName</th>
<th>ActivityIdentifier</th>
<th>ActivityTypeCode</th>
<th>ActivityMediaName</th>
<th>ActivityMediaSubdivisionName</th>
<th>ActivityStartDate</th>
</tr>
</thead>
<tbody>
<tr>
<td>21NC003WQ-CPF126A4</td>
<td>21NC003WQ</td>
<td>NCDENR-DWQ WQX</td>
<td>21NC003WQ-NCLAKESMON326614051-13</td>
<td>Sample-Routine</td>
<td>Water</td>
<td>Surface Water</td>
<td>5/22/2013</td>
</tr>
<tr>
<td>21NC003WQ-CPF126A4</td>
<td>21NC003WQ</td>
<td>NCDENR-DWQ WQX</td>
<td>21NC003WQ-NCLAKESMON5207114051-13</td>
<td>Field Msr/Obs</td>
<td>Water</td>
<td>Surface Water</td>
<td>5/22/2013</td>
</tr>
<tr>
<td>21NC003WQ-CPF126A4</td>
<td>21NC003WQ</td>
<td>NCDENR-DWQ WQX</td>
<td>21NC003WQ-NCLAKESMON5207614051-13</td>
<td>Field Msr/Obs</td>
<td>Water</td>
<td>Surface Water</td>
<td>5/22/2013</td>
</tr>
</tbody>
</table>
WQP Data Analysis Tool

WQP STORET Data Analysis Tool  Data  Criteria  Station  View Data

Click here to generate a blank criteria file:

Generate

(Alternatively)

Please load criteria file

Choose CSV File

Browse...
WQP Data Analysis Tool

Download Data Analysis
Conclusions

✔ The data discovery tool streamlines the process of compiling and performing QA on data set from the WQP.

✔ Utilized Open Source software for development:
  ✔ Eliminates the costs associated with licensing fees.
  ✔ Allows for easier code sharing for customization by jurisdictions.

✔ Reinforces interest in the WQP/STORET Data Warehouse and bolsters the user community’s interest in providing data.
Download Info and Contacts


- For more information contact:
  - Dwane Young (EPA), 202-566-1214, Young.dwane@epa.gov
  - Amy Wesley-Snider (RTI), 919-316-3711, alwesley@rti.org