

Greetings STORET Community

- We will hold the monthly STORET User call on Thursday December 18, 2014 at 12 noon Eastern using:
 - Call number 866-299-3188, access 202-566-0399#
 - Please remember send an email to storet@epa.gov to indicate that you are on the call.
 - Webinar link: <https://epa.connectsolutions.com/storetusercall/>
- Happy holidays from the STORET Team.
- Agenda
 1. New STORET Data Discovery and Assessment Tools

turning knowledge into practice

Exploration of STORET Data Discovery and Analytical Tools

*STORET User Call
December 18, 2014*

**Charles Kovatch, USEPA
Amy Wesley-Snider, RTI**

RTI International is a trade name of Research Triangle Institute

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



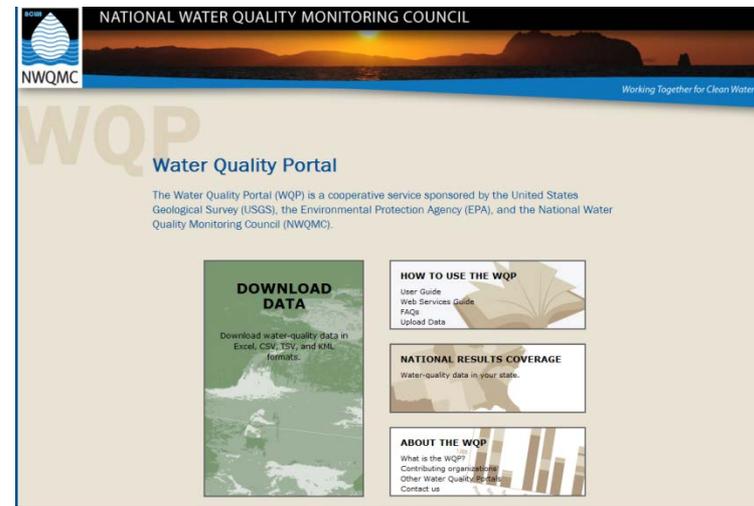
Presentation Outline

- Overview
 - What is driving this effort?
 - Data Discovery
 - Data Analytical Tools
 - Discussion and Feedback
 - Next Steps



What is Driving this Effort?

- Water data submission to STORET
- Use of the Water Quality Portal
- Activities associated with the use of R tools
- Planning to increase Portal outreach and develop analytical tools



Data Discovery

Data Discovery Tools

- STORET
 - Warehouse http://ofmpub.epa.gov/storpubl/dw_pages.querycriteria
 - Web Sercvices http://www.epa.gov/storet/web_services.html
 - User Tools <http://www.epa.gov/storet/otherapps.html>
- Water Quality Portal
 - Query <http://www.waterqualitydata.us>
 - Tools built with R (<https://github.com/USGS-R>)



Assessment Tools

Automated Assessment Tools

1. STORET Trends Analysis
2. Clean Water Act (CWA) Reporting
3. National Level Screening
4. Exploratory Data Analysis (EDA)
5. Biological Risk
6. Additional Functions



1. STORET Trends Analysis Tool

- Goal: Provide users with the ability to statistically identify trends in STORET Data.
- Pilot was run on a static copy of the STORET database.
 - Conducted a Seasonal Kendall Trend analysis for 22 characteristic types at all STORET stations.
 - Utilized open source “R” for this analysis.



1. STORET Trends Analysis Tool cont.

STORET Characteristics used for the Pilot Study

The screenshot displays the STORET Trends Analysis Tool interface. On the left, a list of characteristics is shown under the heading "Trends for Rivers/Streams". The "pH" characteristic is selected. On the right, a map shows several monitoring points marked with green and pink arrows. A pop-up window for point UTAHDWQ-4946900 provides details: "FOUNTAIN GREEN FW W", "Begin Date: Jul 1985", "End Date: Jan 2006", and links for "Result Plot" and "R.Outout".

STORET Trends

Trends for Rivers/Streams

- Absolute Positioning: Bottom left
- Arsenic
- Cadmium
- Copper
- Dissolved oxygen (DO)
- Fecal Coliform
- Lead
- Mercury
- Nitrogen, Kjeldahl
- Nitrogen, Nitrate (NO3) as NO3
- Nitrogen, Nitrite (NO2) + Nitrate (NO3) as N
- Nitrogen, ammonia (NH3) + ammonium (NH4)
- Nitrogen, ammonia (NH3) as NH3
- Nitrogen, ammonia as N
- pH
- Phosphorus
- Phosphorus as P
- Phosphorus, orthophosphate as P
- Solids, Total Suspended (TSS)
- Temperature, water
- Turbidity
- Zinc

RTI INTERNATIONAL *turning knowledge into practice*

UTAHDWQ-4995330
UTAHDWQ-4946900
UTAHDWQ-4932960
UTAHDWQ-4932730
UTAHDWQ-4932220
UTAHDWQ-4932370
UTAHDWQ-4932390
UTAHDWQ-4932450
UTAHDWQ-4955390
UTAHDWQ-4955300
UTAHDWQ-4931410

UTAHDWQ-4946900
FOUNTAIN GREEN FW W
Begin Date: Jul 1985
End Date: Jan 2006
[Result Plot](#)
[R.Outout](#)
Directions: [To here](#) - [From here](#)

Image State of Utah
Image © 2009 DigitalGlobe
Google
lat: 39.267567° lon: -111.016437° elev: 6439 ft Eye alt: 108.97 mi

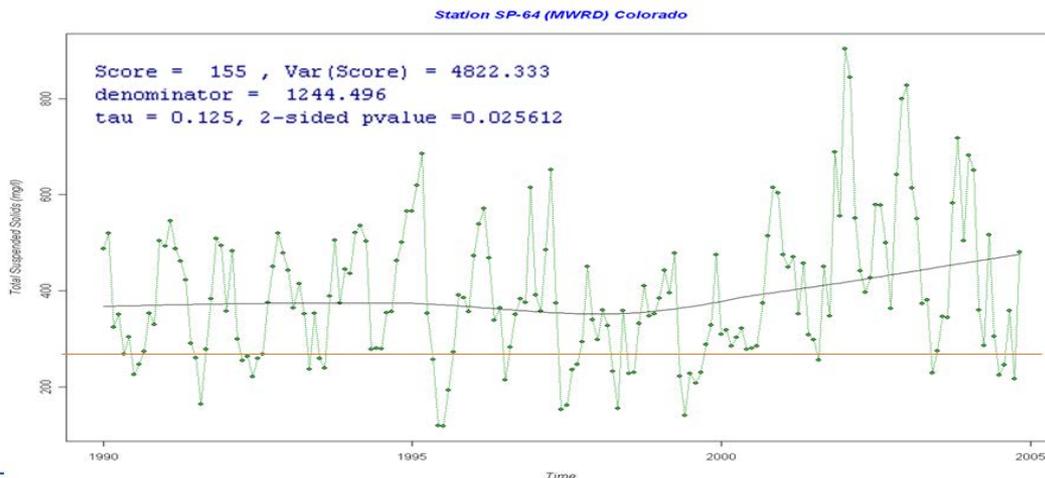
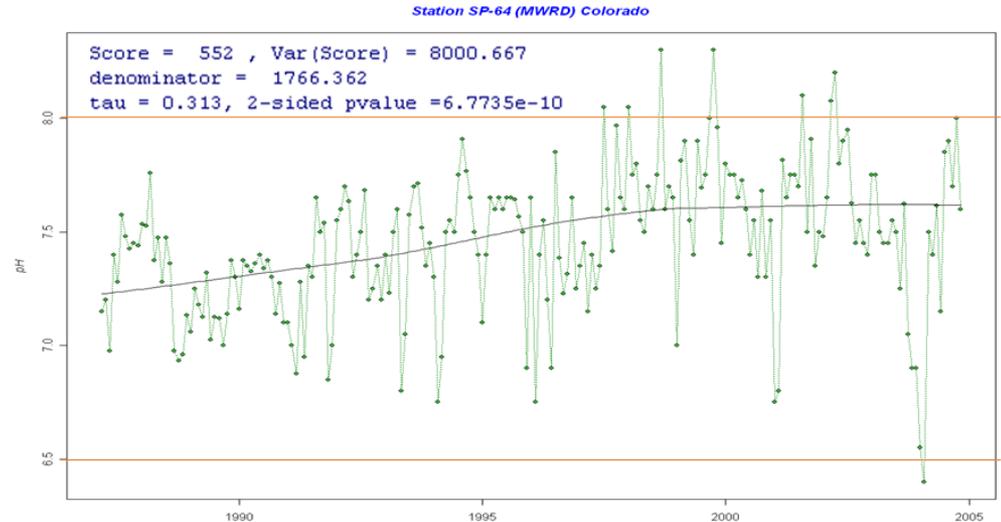
1. STORET Trends Analysis Tool cont.

- Stations that did not have an adequate period of records to support a trends analysis were removed.
 - Created a script to filter for stations having a time series with at least four measurements collected periodically for at least five years.
- A minimum of 20 measurements was determined to provide for the investigation of serial correlation (Yue and Wang, 2004)
- Produced trend plots and scores for each parameter.



1. STORET Trends Analysis Tool cont.

Trend Plot for pH produced by the STORET Trends Analysis Tool

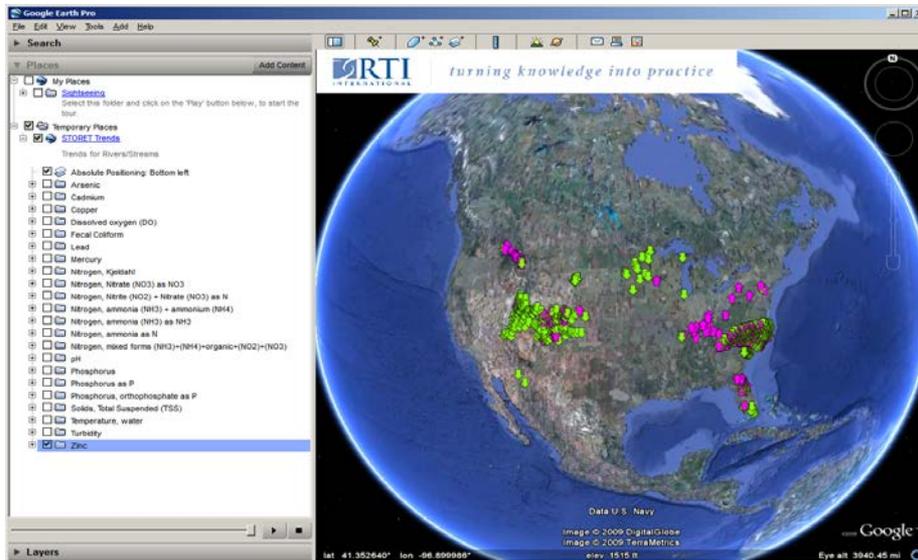


Trend Plot for TSS produced by the STORET Trends Analysis Tool

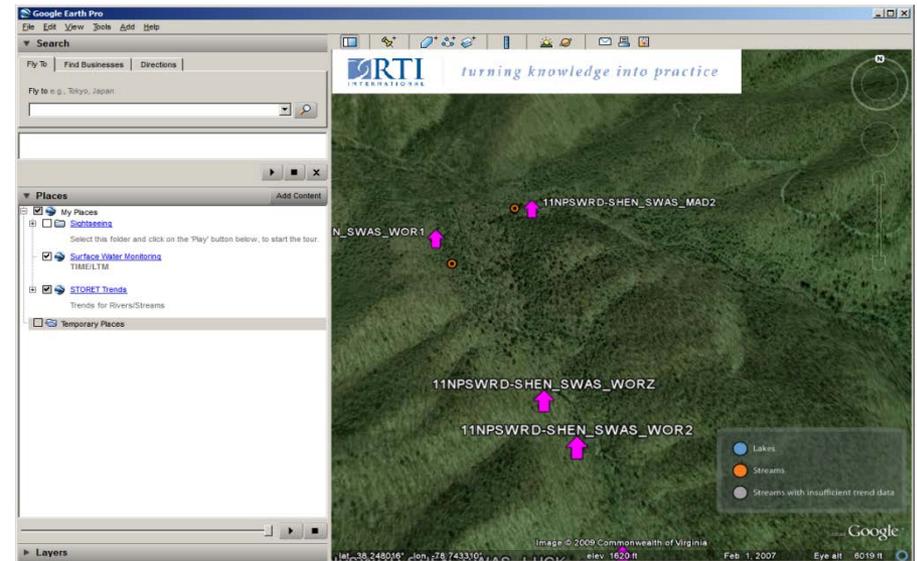


1. STORET Trends Analysis Tool cont.

- Tool produced a Google KML map interface for examining national characteristic trends
- Users can use interface to visualize the trend nationally or drill in on a specific station of interest.



STORET Trends Analysis Tool, National View



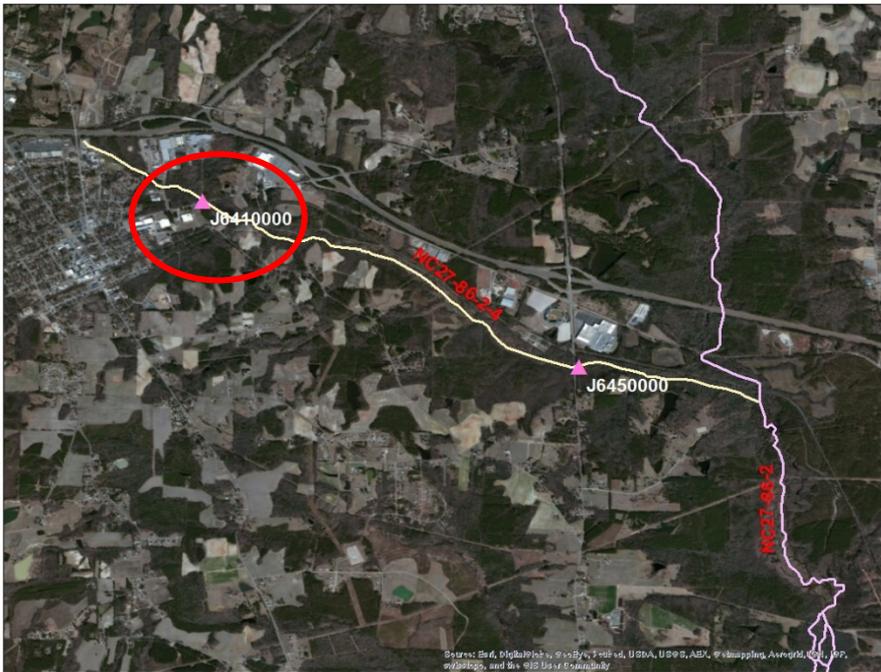
STORET Trends Analysis Tool, Drill Down View

2. CWA Reporting Tools

- Compare STORET/NWQ Portal data to a value (e.g., a WQS) for a parameter of interest.
- Tailored to state-specific needs: waterbody-specific water quality criteria and standards.
- Screening exceedances could be flagged, and samples with results of concern could be highlighted to steer a user to particular areas that may have water quality problems.

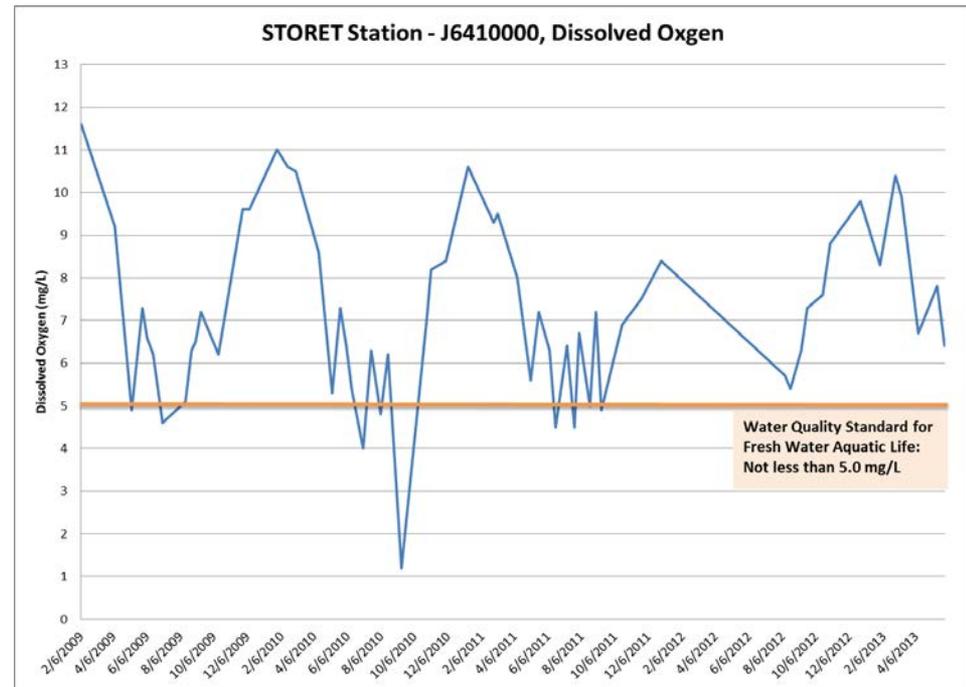


2. CWA Reporting Tools Cont.



Little Creek (West Side): NC27-86-2-4

Total # of Results	Total # of Exceedences	Min	Max	Average
60	8	1.2	11.6	7.185



3. National Level Screening

- Under CWA 304(a) EPA has established National Criteria to provide guidance for states/tribes to use in adopting Water Quality Standards.
- Create a tool to assist in screening against the national recommended water quality criteria.
- Example: Identify all observations of nitrogen/phosphorus that exceed the ecoregional nutrient criteria.



3. National Level Screening Cont.

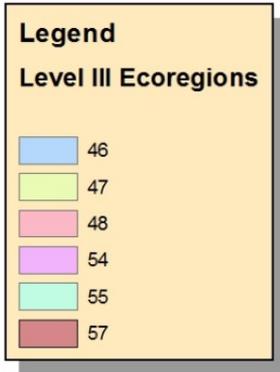
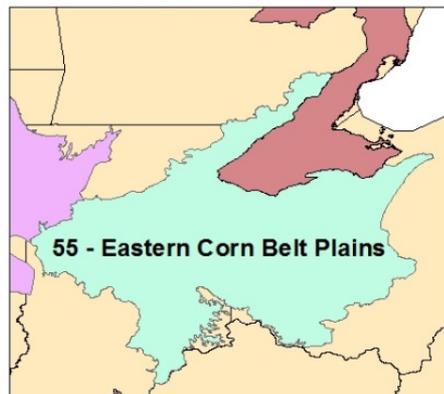
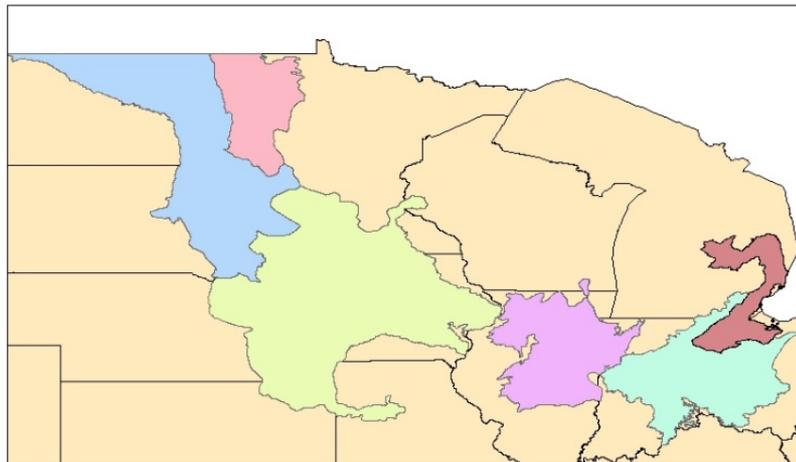
- EPA Developed Ecoregional Criteria for Total Phosphorus, Total Nitrogen, Chlorophyll *a* and Water Clarity.
- For this example Ecoregion VI, #55 *Eastern Corn Belt Plains* was used.

Table 3e. Reference conditions for level III ecoregion 55.

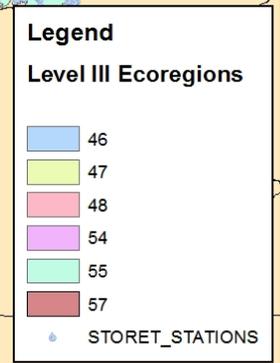
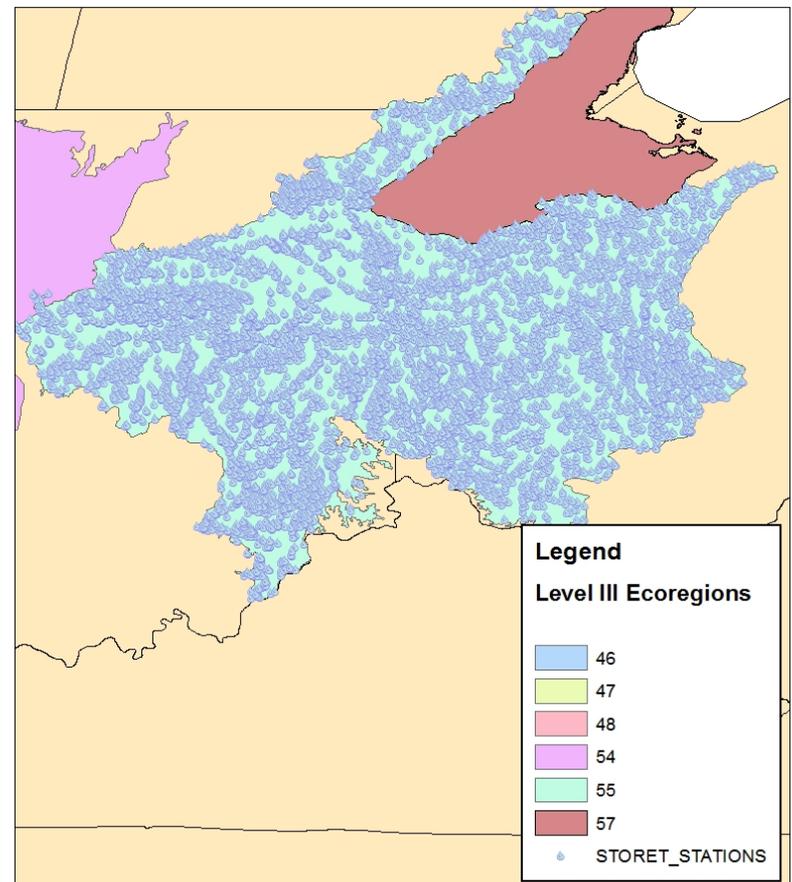
Parameter	No. of Streams N **	Reported values		25 th Percentiles based on all seasons data for the Decade	Reference Streams **
		Min	Max	P25-all seasons [†]	P75 - all seasons
TKN (mg/L)	198	0.05	3.5	0.4	
NO ₂ + NO ₃ (mg/L)	219	0.025	8.13	1.60	
TN (mg/L) - calculated	NA	0.075	11.63	2	
TN (mg/L) - reported	2 z	3.63	3.78	3.63	
TP (µg/L)	225	10	1820	62.5	
Turbidity (NTU)	1 z	10.4	10.4	10.4 zz	
Turbidity (FTU)	12	3.3	50.65	9.21	
Turbidity (JCU)	1 z	28	28	28 zz	
Chlorophyll <i>a</i> (µg/L) -F	—	—	—	—	
Chlorophyll <i>a</i> (µg/L) -S	8	4.32	19.24	6.62	
Chlorophyll <i>a</i> (µg/L) -T	8	6.67	22.72	7.99	
Periphyton Chl <i>a</i> (mg/m ²)	—	—	—	—	—

3. National Level Screening Cont.

Aggregate Nutrient Ecoregion 6



Level III Ecoregion - 55: Eastern Corn Belt Plains River and Stream Monitoring Stations



3. National Level Screening Cont.

- For this ecoregion there are over 12,000 stations
- Screening Tool would allow user to:
 - Define parameter,
 - Define date range,
 - Select statistics for data (e.g. number of samples that exceed the reference conditions for parameter)
- Screening tool would return results in user defined format (.xls, .csv, etc..)

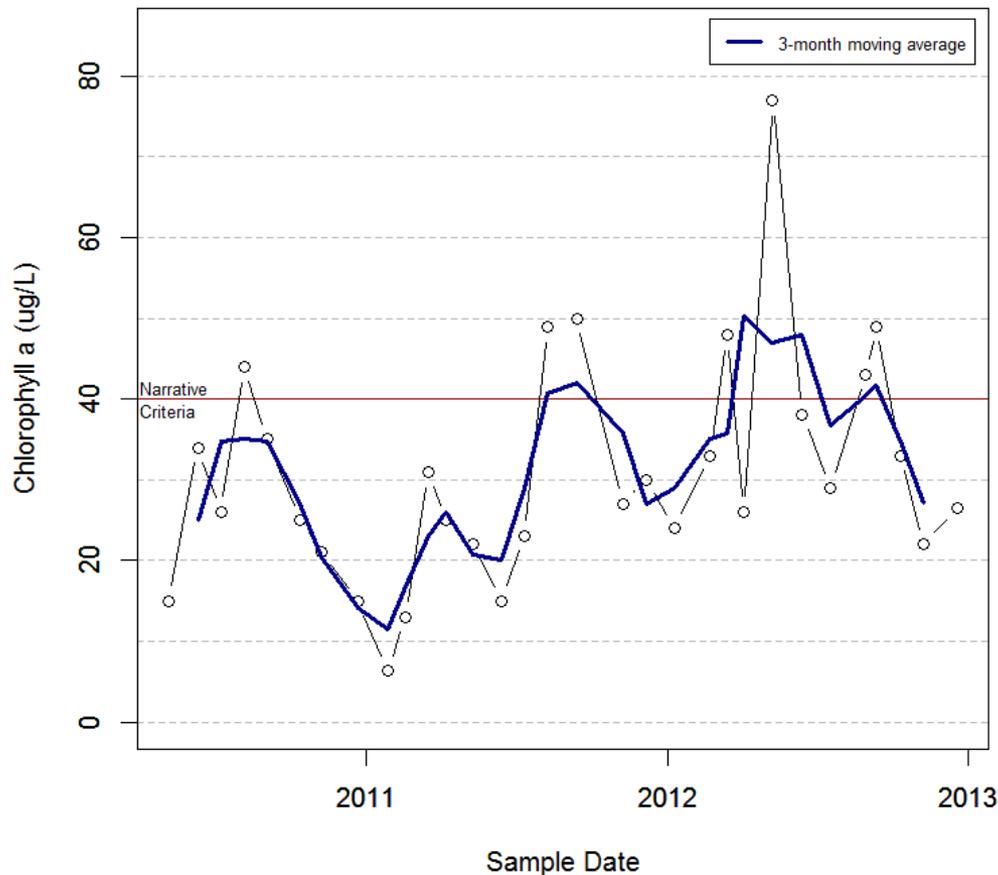


4. Exploratory Data Analysis

- Exploratory data analysis (EDA)
 - Help identify degrading waters that may need additional protection.
 - Highlight improving waters showing progress towards restoration.
- EDA tools can be used to:
 - Examine parameter trends over time.
 - Identify relationships between parameters and other variables.
 - Estimate the probability of threshold exceedances.
 - Characterize central tendency (mean, median) and variability of parameter values.
 - Compare outputs described above by: station, watershed, county, eco-region, sample month, season, etc.
- Tool outputs can be graphical and/or tabular

4. Exploratory Data Analysis-Time Series Data

Chlorophyll a
Station ID NEU0171B

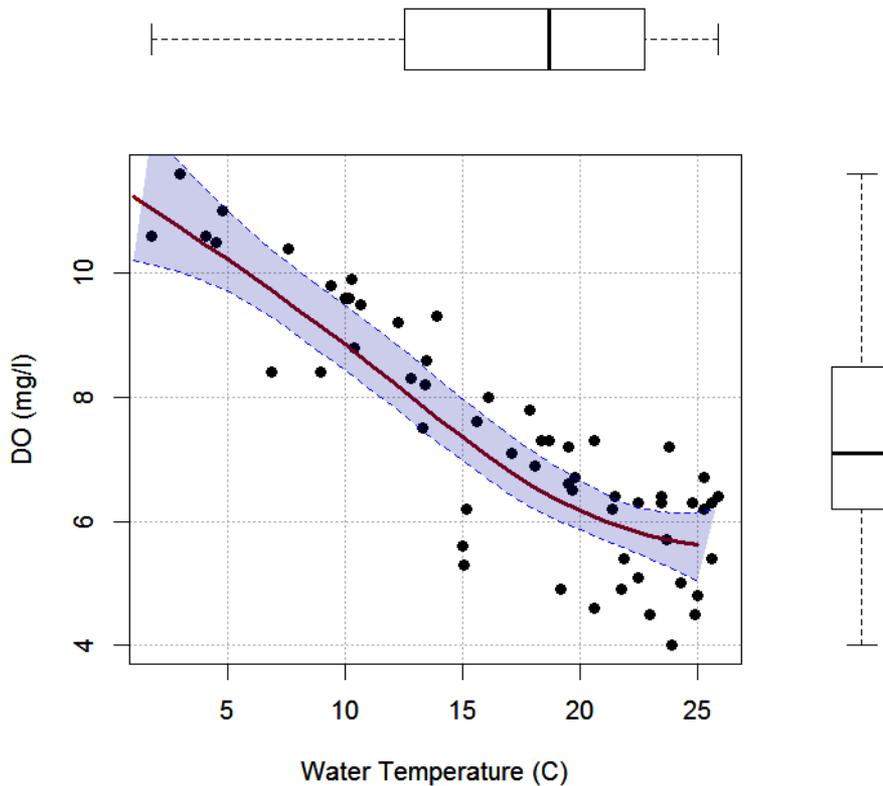


Time series data:

- Identify potential trends over time relative to WQ criteria or other thresholds.
- Add moving average to help clarify trends obscured by noisy (high variability) data.

4. Exploratory Data Analysis- Scatterplots/Trend lines

Relationship between DO and Water Temperature
Station ID J6410000

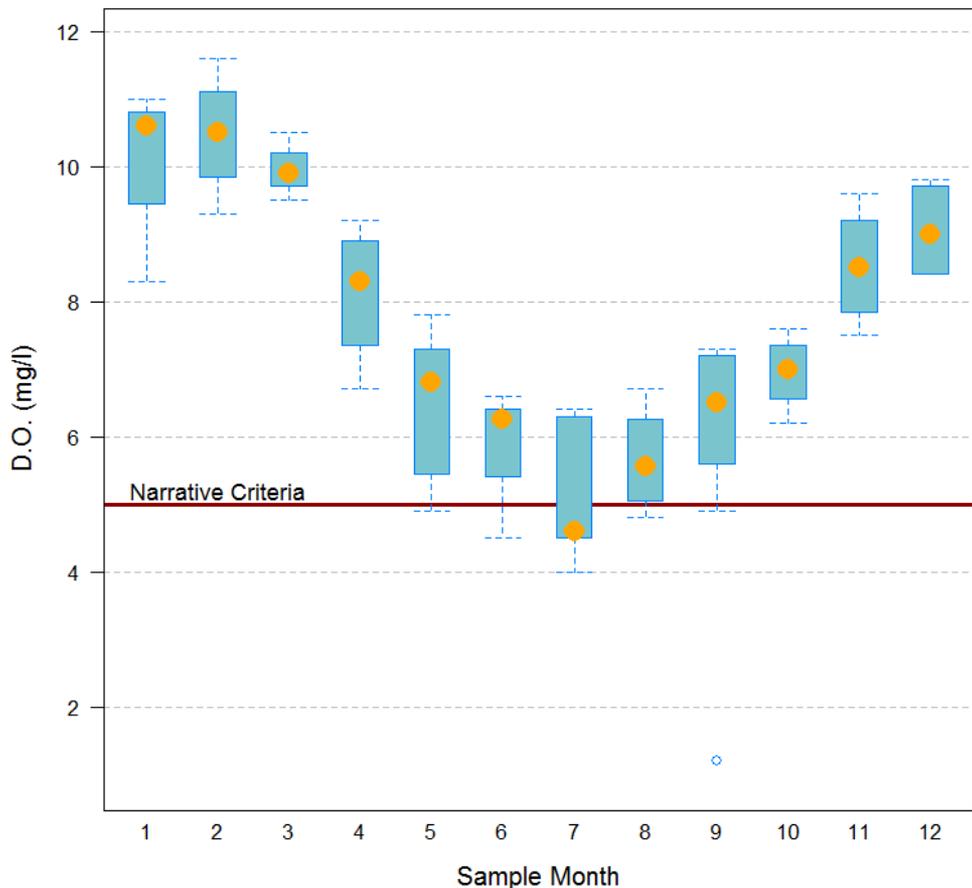


Scatterplots with trend lines:

- Identify the direction, shape, and strength of relationships between variables
- Estimate and display prediction intervals.

4. Exploratory Data Analysis - Boxplots

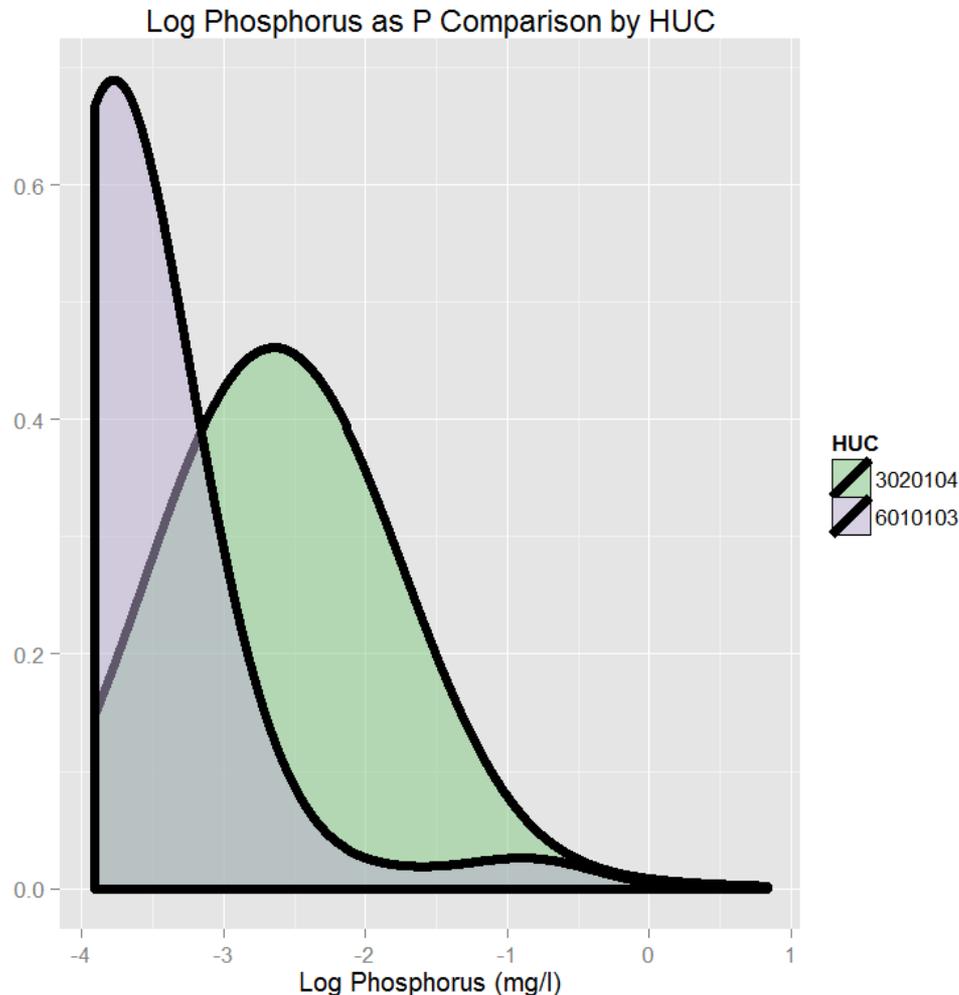
Seasonal Trends in Dissolved Oxygen
Station ID J6410000



Boxplots:

- Show median values, inner quartile range (25th to 75th percentile), outer quartiles and outliers.
- Identify potential spatial or temporal trends.

4. Exploratory Data Analysis– Density Estimates



Density plots:

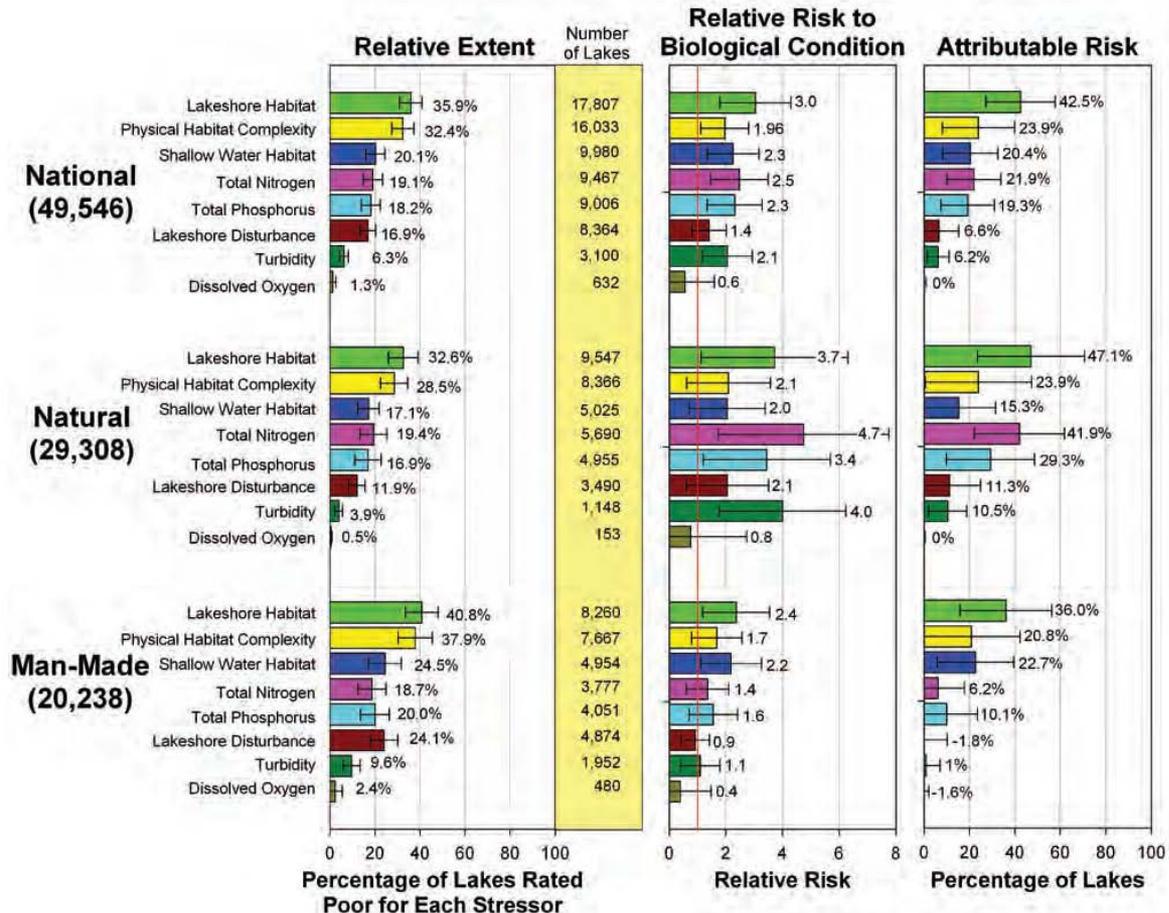
- Display central tendency (mean or median) and spread (variability) of parameter.
- Compare different spatial or temporal units (i.e., is one HUC more likely to experience an exceedance?).
- Identify skewed distributions (i.e., higher probability of extreme values).

5. Biological Risk

- Develop a tool to calculate relative and Attributable Risk for the NARS data set.
 - Relative risk: examines the relationship between a stressor and the response in a measure of biological condition.
 - Attributable risk: builds on relative risk by looking at the likelihood that poor biological condition will exist when stressor levels are rated poor, and at the extent to which those poor conditions exist in the assessed region.



5. Biological Risk



Relative extent, relative risk and attributable risk has been used to assess the condition of the nation's lakes (US EPA, 2009)



6. Additional Functions

- **Upstream/downstream summarization:** Select a pour point and summarize data upstream. Or once a station is selected, select any other stations with same parameter downstream for X km. Results could be plotted on the same graph(s) for comparison.
- **NHDPlus Catchments:** Report out the NHDPlus catchments associated with monitoring stations.
- **Ecoflows Calculation:** STORET contains biological and habitat data that could be useful for site-specific or cross-sectional ecological flow analysis. (Note: RTI has demonstrated the potential for such cross-sectional analyses using data from North Carolina)



Summary

- ✓ A range of discovery and analytical tools could be developed to access the wealth of data within the STORET and the Water Quality Portal.
- ✓ Make it easier to compile relevant information for state users and decision makers.
- ✓ Increase the value of water monitoring data.
- ✓ Utilize Open Source software for development:
 - ✓ Reduces or eliminates software costs.
 - ✓ Allows for easier code sharing and customization



Discussion

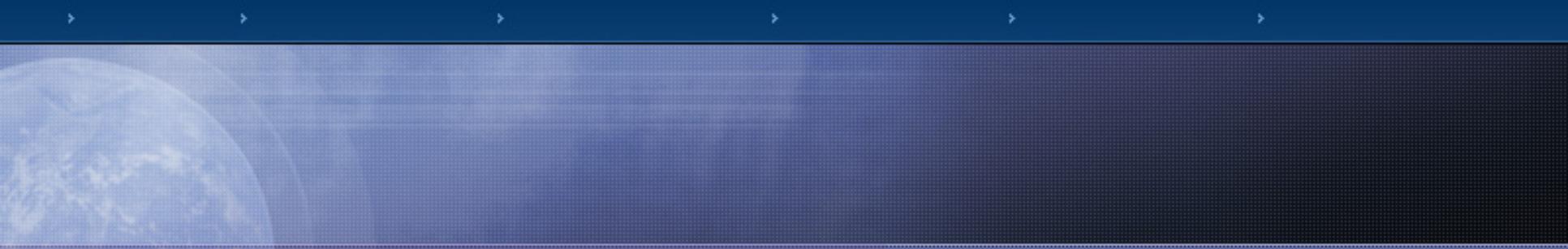
- What data discovery tools would you find the most helpful?
- What data analysis tools would you find the most helpful?



Next Steps

- Compile list of requirements gathering ideas
- Identify the first tool(s) to develop
- Launch tool(s) in 2015





Happy
Holidays

