

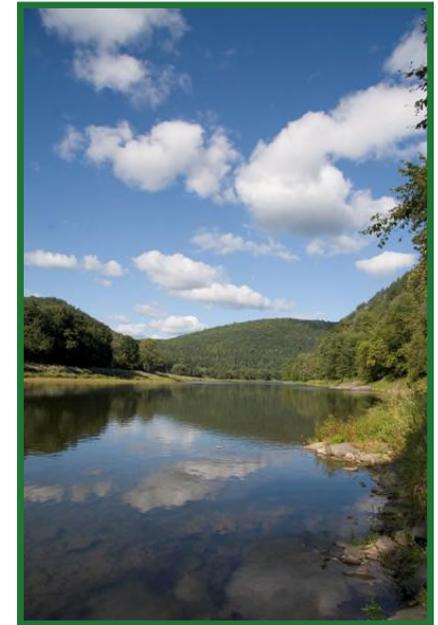
# Delaware River Basin Commission

## Enhancing Data Interpretation in the Delaware River Basin with R

**Robert Limbeck**  
Senior Aquatic Biologist

**John Yagecic, P.E.**  
Manager, Water Quality Assessment

10th National Monitoring Conference  
Tampa, FL  
May 2016



**Delaware River Basin Commission**

DELAWARE • NEW JERSEY  
PENNSYLVANIA • NEW YORK  
UNITED STATES OF AMERICA

# This Presentation

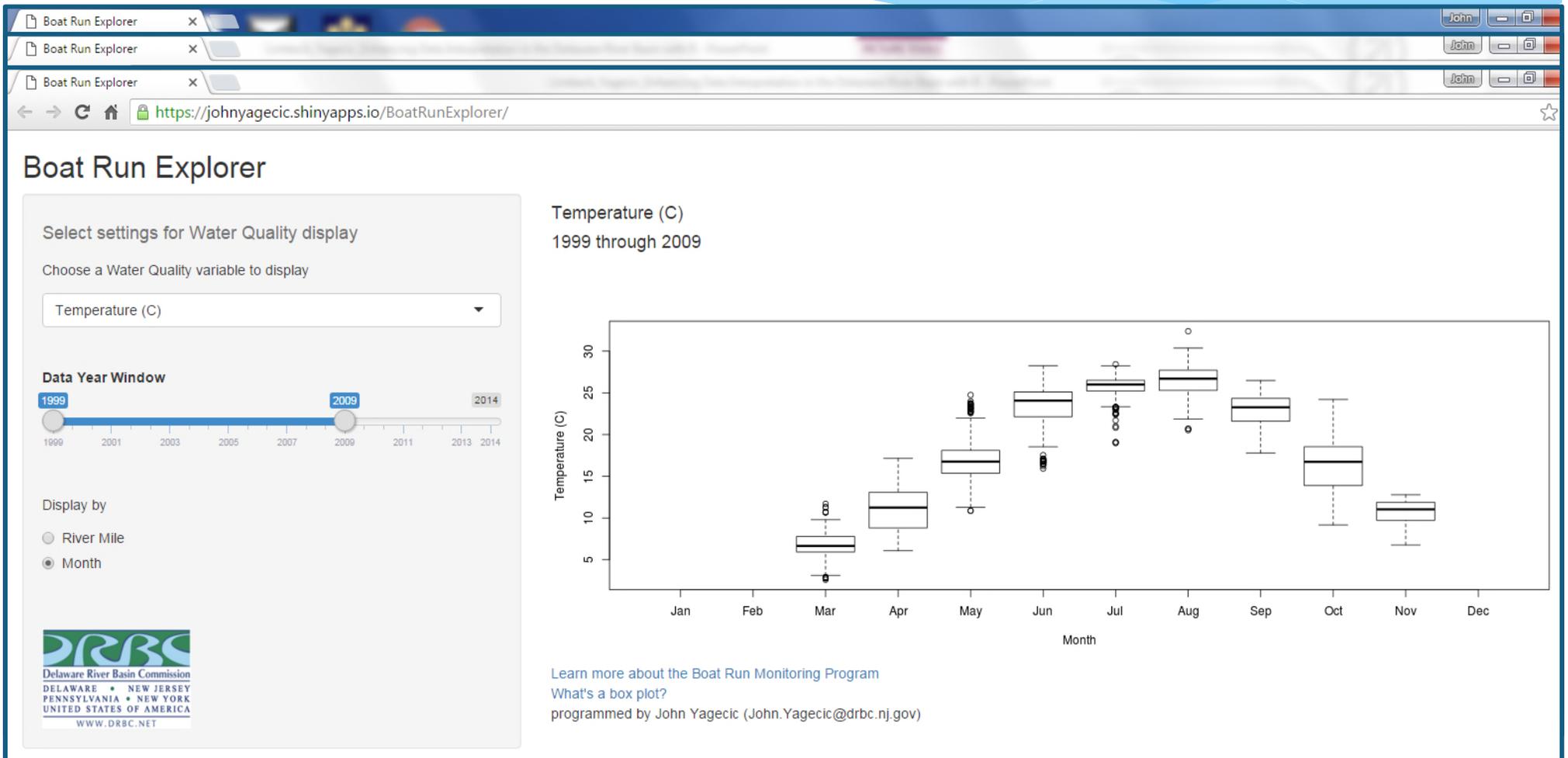
## \* Three Sample Projects by Delaware River Basin Commission

1. Boat Run Explorer
2. Near real-time flow & water quality dashboards
3. Optimization of complex response time for Dissolved Oxygen

# Project 1: Boat Run Explorer

- \* **Goal: to foster better public engagement with long-running monitoring program;**
  - \* Colleagues, stakeholders, students, public;
  - \* What to show?
  - \* Paper report? Would anyone read it?
  - \* Able to be updated?

# Boat Run Explorer



# Boat Run Explorer Details

## Direct Link

- \* <https://johnyagecic.shinyapps.io/BoatRunExplorer/>
  - \* Shiny App hosted by the shiny apps io server;
  - \* User selected time period;
  - \* By River Mile or by month;
  - \* **1000**'s of combinations most of which we haven't seen;
  - \* Links to program page and "what's a box plot" Wikipedia page;
  - \* Explore data spatial and temporal structure.
- \* User selected parameter:
    - \* Temperature;
    - \* Salinity;
    - \* Nitrate + Nitrite;
    - \* DO;
    - \* DO Sat;
    - \* Secchi depth;
    - \* Turbidity;
    - \* Others.

# Project 2: Daily Dashboards

- \* **Goal: near-real-time assessment of water quality and flow conditions;**
  - \* Take advantage of real-time data;
  - \* Better sense of system response & dynamics;
  - \* Improved ability to:
    - \* notify others
    - \* respond to problematic conditions;
    - \* remedy problematic conditions (if possible) .
  - \* <http://drbc.net/Sky/flows.htm>
  - \* <http://drbc.net/Sky/waterq.htm>

# Automated Process for Generating Daily Dashboards

Web Site

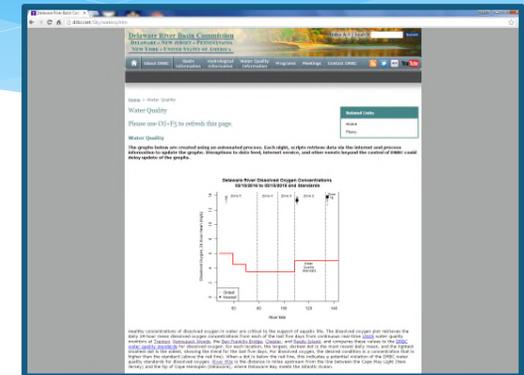
USGS NWIS



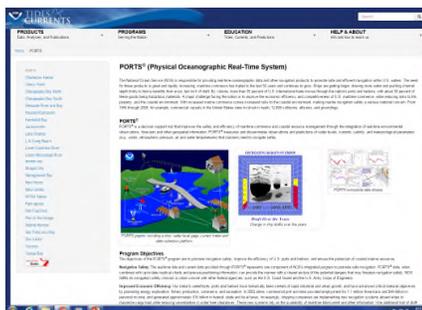
Query

Data

New Plots



NOAA PORTS

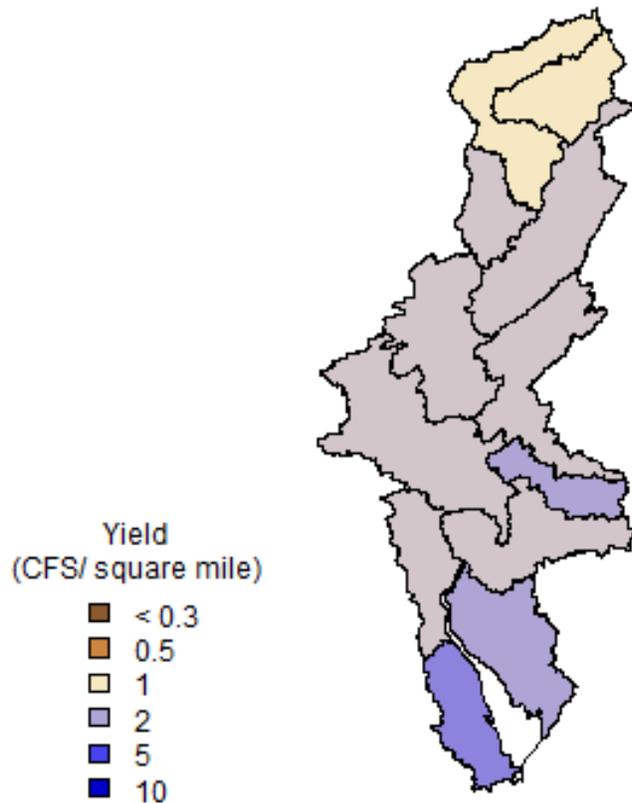


Data Processing & Plotting

- \* Fully Automated
- \* Overnight, every night
- \* R scripts executed in batch mode
- \* Called by Windows Task Scheduler

# Daily Dashboard Features: Animated choropleth

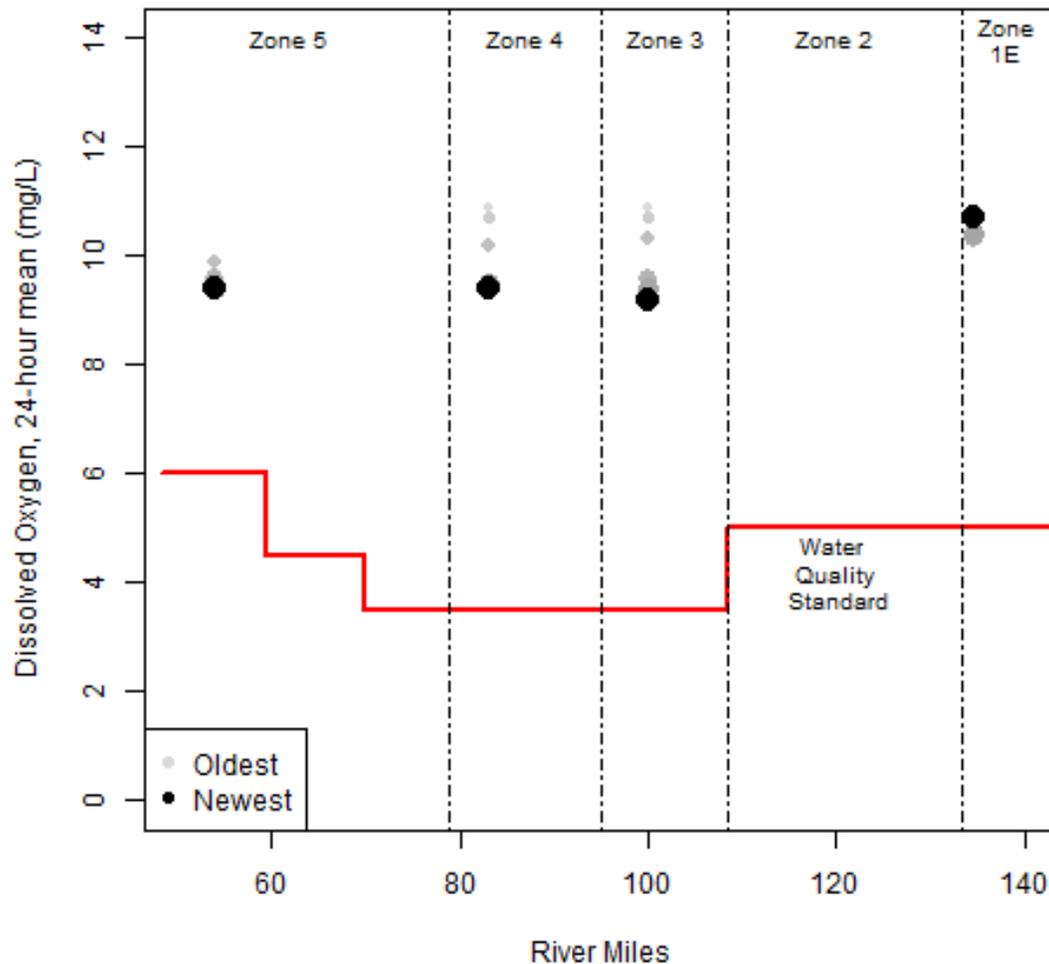
Delaware Basin daily water Yield  
(CFS/square mile) on 01/28/2016



- \* Retrieve daily mean Q from 140+ USGS gages in basin for previous 20 days;
- \* Divide daily Q by gage drainage area to get yield;
- \* Mean of all yields within a HUC8 for each day;
- \* Animate over 20-day time series.

# Daily Dashboard Features: DO Assessment and Trending

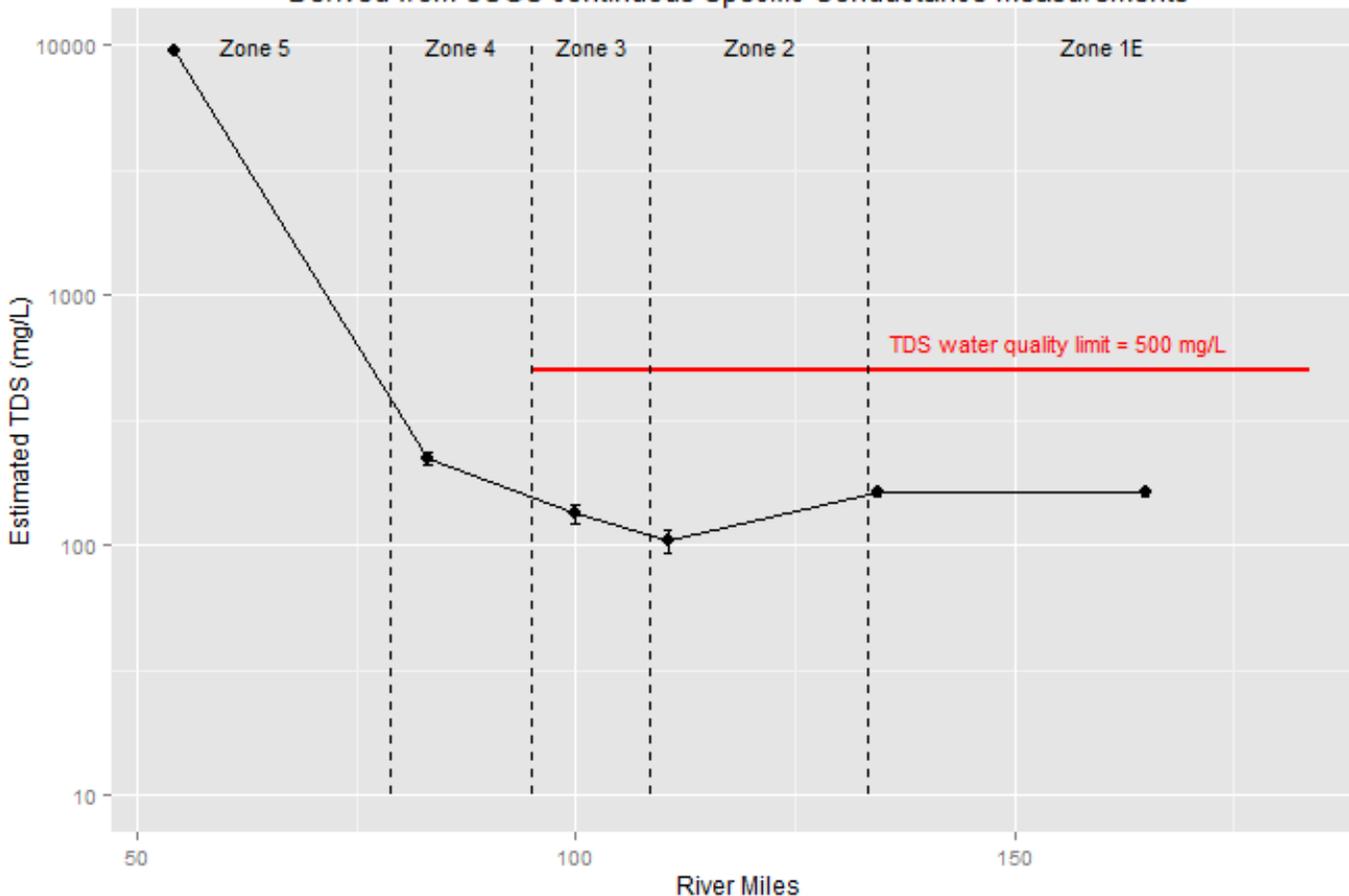
Delaware River Dissolved Oxygen Concentrations  
04/18/2015 to 04/23/2015 and Standards



- \* Retrieve daily mean DO from USGS sites last 5 days;
- \* Plot compared to DO standards (red line);
- \* Newest DO – big dark dot;
- \* Older DO – smaller lighter dot.

# Daily Dashboard Features: Total Dissolved Solids near real-time

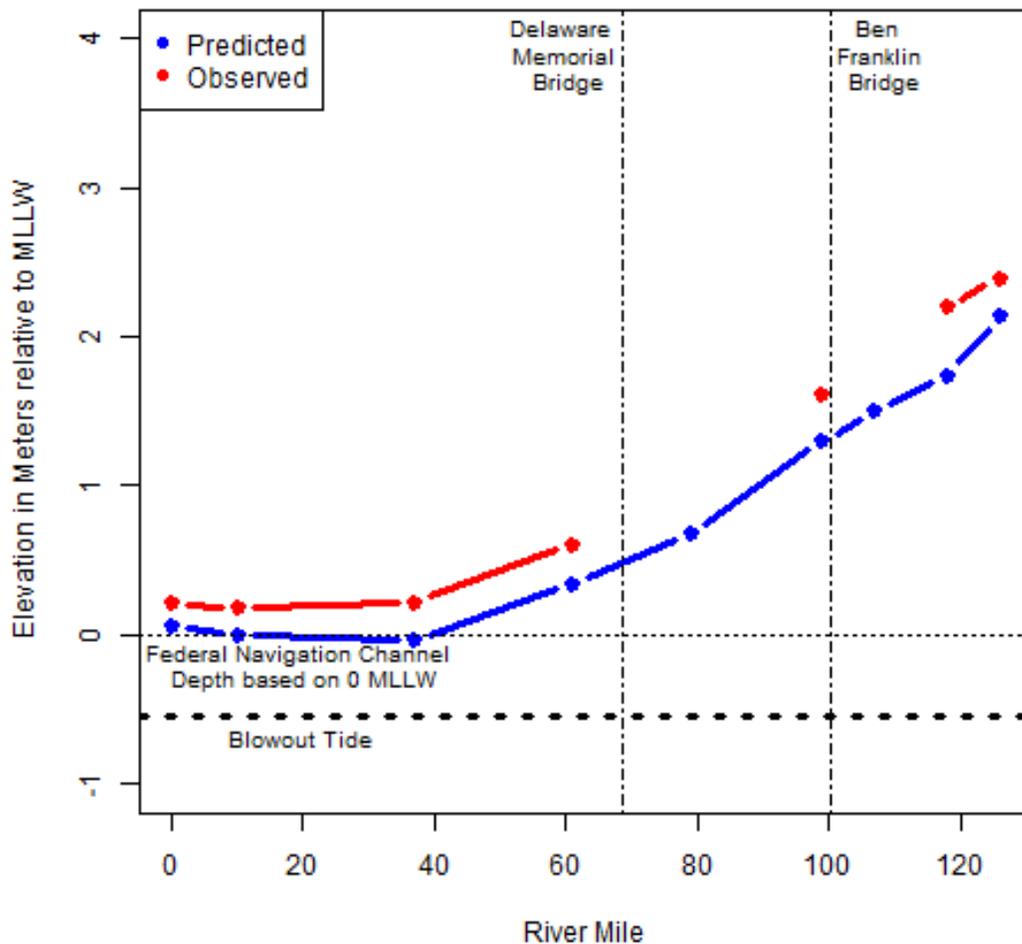
Mainstem Delaware River estimated maximum Total Dissolved Solids on 08/09/2015  
Derived from USGS continuous Specific Conductance measurements



- \* Real time specific conductance data;
- \* Location specific relationships to compute TDS (black dot) from SC;
- \* Compare near-real-time computed TDS to standards (red line).

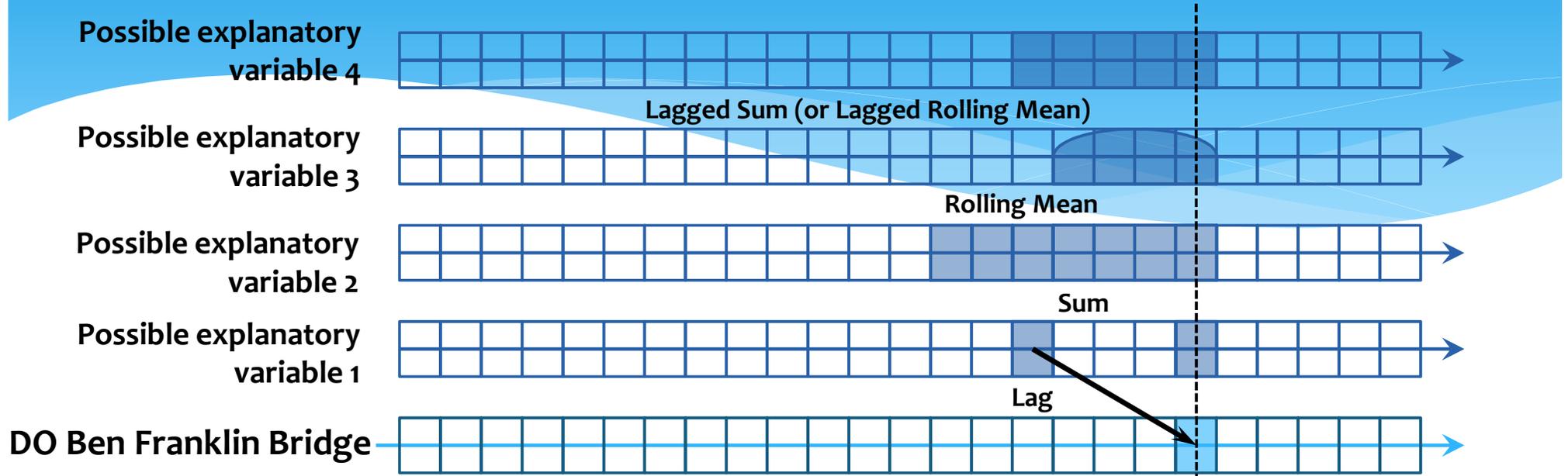
# Daily Dashboard Features: Animated Tidal Water Surface

Delaware Estuary Water Surface Elevation, 02/05/2016 00:00  
Data retrieved 02/09/2016



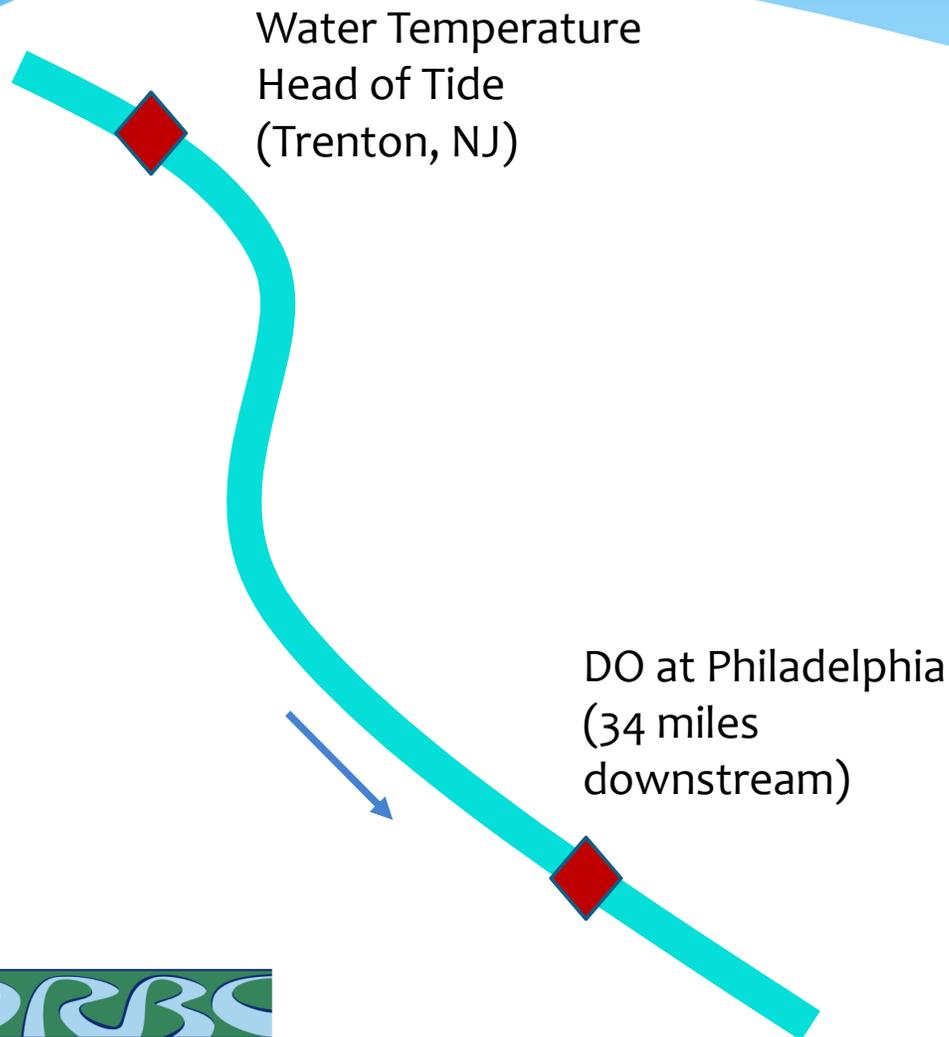
- \* Retrieve NOAA PORTS water surface time series data;
- \* observed last 5 days; predicted to end of current day;
- \* Longitudinal plot for each time step;
- \* Knit together to animation.

# Temporal Complexity Problem



	Explanatory Variables	Averaging Periods	Summing Periods	Lagging Periods	time Subtotal
Lag	35			30	1,050
Sum	17		30		510
Rolling Mean	35	30			1,050
Sum + Lag	17		30	30	15,300
Rolling Mean + Lag	35	30		30	31,500
		Possible Additional Explanatory Variables			49,410

# Project 3: Determine complex lag time between stimulus and response



- \* **Goal: to better understand relationships between stimulus and response variables;**
- \* Some relationship water temperature at Trenton and DO at Philadelphia but...
- \* Travel time...
- \* Residence time in estuary much higher...
- \* Lag time?
- \* Rolling mean of multiple days?
- \* Some combination of rolling mean and lag time?

# Power of a scripting language for examining data

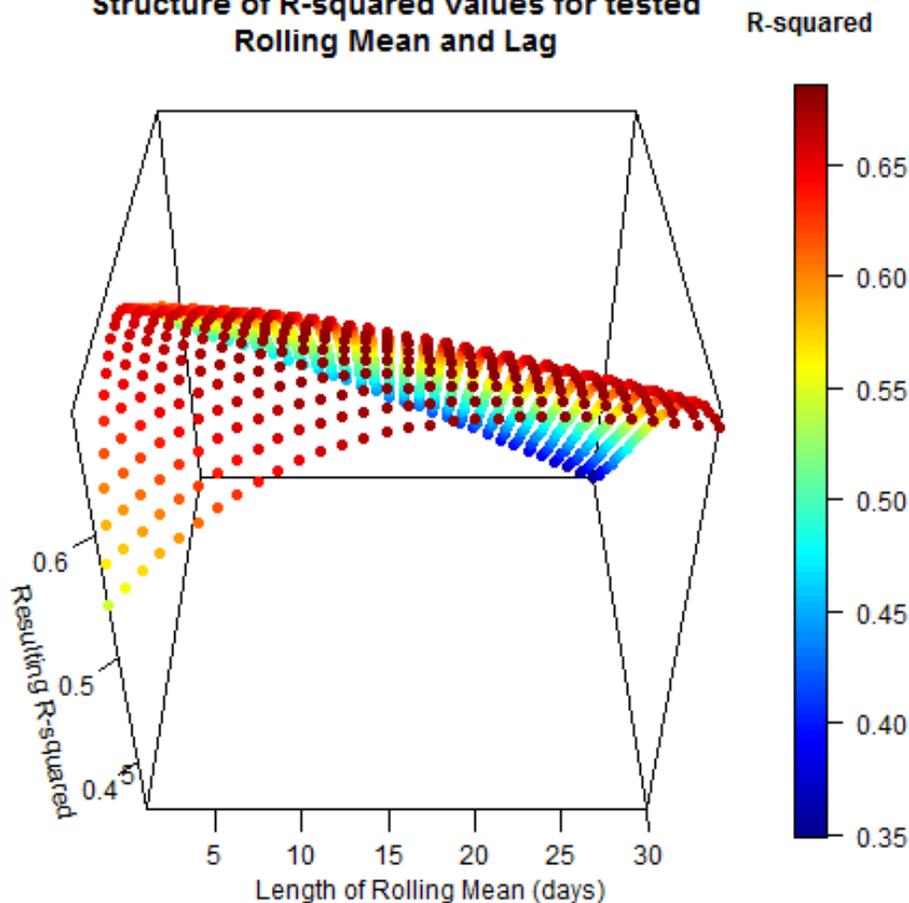
- \* 10 years of daily data (USGS);
- \* Can cycle through every combination of lag time, rolling mean (stimulus), lag + rolling mean;
- \* 49,000 possible combinations
- \* Separate linear regression model for each of 49,000
- \* Compare all resulting  $R^2$  values

```
1 for (i in 1:length(term1)){  
2   for (j in 1:length(term2)){  
3     for (k in 1:length(term3)){  
4       testmod<-lm(y~term1[i]+  
5         term2[j]+  
6         term3[k])  
7       summary(testmod)$adj.r.squared  
8     }  
9   }  
10 }  
11 }  
12 }  
13 }
```

\* pseudo-code only, not actual script

# Complex Lag Time Result

Structure of R-squared values for tested Rolling Mean and Lag



- \* Resulting  $R^2$  values (vertical axis) have structure - not random;
- \* One combination is better than all other combinations;
  - \* 16-day rolling mean temperature at Trenton lagged by 4 days;
- \* Can understand something about system that would be difficult to determine any other way.

# Thank You!

## Delaware River Basin Commission

[www.drbc.net](http://www.drbc.net)

[John.Yagecic@drbc.nj.gov](mailto:John.Yagecic@drbc.nj.gov)

[Robert.Limbeck@drbc.nj.gov](mailto:Robert.Limbeck@drbc.nj.gov)

609-883-9500

\* *P.S. – Always looking for collaborators to do something interesting.*