New Tools for Water Quality Data Access and Trend Analysis

An overview of the USGS R Packages: dataRetrieval and EGRET

Robert M. Hirsch & Jessica Thompson
USGS
1 May 2014
From Ralph Keeling

The only way to figure out what is happening to our planet is to measure it, and this means tracking changes decade after decade and poring over the records.

Keeling, 2008, Recording Earth’s vital signs, Science, p1771-1772
The challenge

How do we come to understand what is happening to water quality in large watersheds.

Is it getting better or worse?

Can we develop ideas of causative factors and changes in processes?

Can these be used to guide management choices?
Potomac River at Chain Bridge, Washington DC
Box plot of sample values by Water Year
Nitrate as N

Data through September 2002

The thrill of victory
Potomac River at Chain Bridge, Washington DC
Boxplot of Discharge on Sampling Date by Water Year
Data through September 2003

Potomac River at Chain Bridge, Washington DC
Box plot of sample values by Water Year
Nitrate as N

The agony of defeat
Potomac River at Chain Bridge, Washington DC
Boxplot of Discharge on Sampling Date by Water Year
Motivations for the method

- Describe the evolving behavior of the watershed. No mathematical straight-jacket!!
- Estimate both concentration & flux (averages as well as trends).
- Estimate the actual history but also a flow-normalized history.
- Resolve a serious bias in flux estimates.
- Be quantitative but also exploratory.
Data requirements

- Low intra-day variability (not flashy)
- Requires a complete daily discharge record
- Intended for >200 samples, but has been used for some purposes with as few as 60 samples
- Water quality samples cover most of the discharge range
- For trend studies: 20+ years
- For average flux computations: 5 – 10 years.
Susquehanna, 70,000 km² watershed

Choptank, 293 km² watershed
“Data without models are chaos, but models without data are fantasy”

Nesbit, Dlugokencky and Bousquet, Science, 31 January 2014, pp. 493-495
Use the data and a simple, highly-flexible smoothing model to decompose the data into 4 components.

1) Discharge related component
2) Seasonal component
3) Time trend
4) Random component

**Weighted Regressions on Time, Discharge and Season (WRTDS)**
Locally Weighted Regression

For any location in time - discharge space (t and Q) we assume that concentration (c) follows this model

\[ \ln(c) = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot \ln(Q) + \beta_3 \cdot \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon \]

But the coefficients should be smoothly changing as we move through the space.

Use weighted regression at many points in that space. The weight on each sample is determined by its “relevance” to that particular point in the space.
WRTDS view of the evolving behavior of nitrate

Choptank River near Greensboro, MD  Nitrate plus Nitrite, Filtered, as N
Estimated Concentration Surface in Color

How is this surface created?
Every dot is a data point from 1993 to 2012
Let’s say we want to use the data to estimate the expected value of concentration for January 1, 2003 at Q=500 cfs

Choptank River near Greensboro, MD  Nitrate plus Nitrite, filtered, as N
Locations of all available data
How do we set the weights for the regression? These are the same points we just saw, but the radius of the dot is proportional to weight assigned to that point for purposes of estimating concentration for January 1, 2003 at Q=500 cfs.

The weight depends on distance in: time, log discharge, and season from January 1, 2003 at Q = 500 cfs.
What if we wanted to make an estimate for January 1, 2003 but for $Q = 50$ cfs
Redo the weights for distance from that point

Choptank River near Greensboro, MD
Nitrate plus Nitrite, filtered, as N
Locations of all available data
To organize the work, let's make estimates for a fine mesh of points in this space. We will do it at 14 Q values and 177 time values, for a grid of 2,478 points.
Here are two, more detailed looks at this surface.
Now, for every one of 10,227 days in the record from 1985 through 2012:

We can use the date and the observed discharge to compute the expected value of concentration.

From that value we can compute the expected value of flux.
Then we can sum these estimates by year to compute estimates of annual mean concentration & annual mean flux.
Can we filter out this flow-driven variation to see the underlying change?
The “flow normalized concentration” on any given day is: $c=f(Q,T)$ integrated over the probability distribution of $Q$ for that day of the year.

Flow normalized flux is just $c \times Q$ integrated over discharge.

Sum those over the year to get annual flow-normalized mean concentration and flux.
Choptank River near Greensboro, MD  Nitrate
Water Year
Mean Concentration (dots) & Flow Normalized Concentration (line)

Concentration trend +54% 1985-2012

Choptank River near Greensboro, MD  Nitrate
Water Year
Flux Estimates (dots) & Flow Normalized Flux (line)

Flux trend +36% 1985-2012
Choptank River near Greensboro, MD  Nitrate
Water Year
Mean Concentration (dots) & Flow Normalized Concentration (line)

Concentration trend +3.2% 2007-2012

Flux trend -1.4% 2007-2012
Look at changes in just the last few years.

This is a graphic of differences 2007 to 2012

Hypothesis, cover crops are helping at higher flows particularly in the winter. Low flows are still responding to legacy of nitrate enriched groundwater.
Why all this complexity?

Different products for different purposes

• Concentration versus flux
• Actual history versus flow-normalized history
For understanding impact on the estuary ecosystem

We want the flux history
For understanding progress in the watershed, we want the flow-normalized flux history.
For understanding the changes in the rivers, we want the concentration history.
Susquehanna, 70,000 km$^2$ watershed
Sediment plume from Susquehanna watershed
Plume extends over 150 km down the Chesapeake Bay
Carrying:
Sediment, Phosphorus, Nitrogen

From the watershed and from storage in Conowingo Reservoir
Conowingo Dam during Tropical Storm Lee, September 2011, Reservoir is rapidly filling, Trap efficiency in decline
Susquehanna River at Conowingo, MD  Total Phosphorus
Estimated Concentration change from 1995 to 2011
Change from 1995-2012

0.4%/yr

3.1%/yr
Maumee River – 16,000 km²
Tributary to Lake Erie
Cyanobacter – Lake Erie
Dissolved Reactive Phosphorus, Maumee River, at Waterville, OH

Mid 1970’s

Late 1980’s

Early 2010’s
Dissolved Reactive Phosphorus, Maumee River, at Waterville, OH

**Concentration**

- Concentration in mg/L
- 1980-2010
- +44% 1989-2011

**Flux**

- Flux in millions of kg/year
- 1980-2010
- +80% 1989-2011
dataRetrieval & EGRET (Exploration and Graphics for RivEr Trends)

- **dataRetrieval**: ingests the data from multiple sources and puts them in a form suitable for analysis by EGRET or generic use of R
- **EGRET** includes the WRTDS computations: for trend analysis, flux computation, and exploration
- Also includes analysis of streamflow trends
- Produces a wide range of graphical and tabular outputs
dataRetrieval

• Can acquire data from web services (USGS or Storet) as well as from user-supplied files

• Includes capability for sample data, daily discharge, other daily values, sensor data, and meta data about site and parameter

• Structures the data to be conveniently used by the EGRET software
Getting Started

- Need to install R (freely downloaded from http://cran.us.r-project.org/)
- Next install the EGRET & dataRetrieval packages & the other R packages they depend on:
- In R, run the following commands:
  - `install.packages(c("zoo", "survival", "stringr", "fields", "spam", "XML", "Rcurl", "plyr", "reshape2"))`
  - `install.packages("dataRetrieval", repos="http://usgs-r.github.com", type="both")`
  - `install.packages("EGRET", repos="http://usgs-r.github.com", type="both")`
Then each time packages are used, they need to be loaded, using the commands

```r
library(dataRetrieval)
library(EGRET)
```

Once this is done you will have access to help and to the package vignettes.

To get help with a function (such as the function `getSampleData`) just type `?getSampleData`
How can we enter data

- For the water quality sample data
  - From USGS web services
  - From EPA-Storet web services
  - From a user supplied file
- For the daily discharge data
  - From USGS web services
  - From a user supplied file
- For the meta-data
  - From USGS web services
  - From user entries
```r
library(dataRetrieval)
library(EGRET)
siteID <- "01491000"
parameter_cd <- "00631"
startDate <- "1979-09-01"
endDate <- "2011-09-30"
Sample <- getSampleData(siteID, parameter_cd, startDate, endDate)
summary(Sample)
```
> Daily <- getDVData(siteID,"00060",startDate,endDate)

There are 11718 data points, and 11718 days.

> summary(Daily)

> length(Daily$Q)
[1] 11718


Censored values

All concentration data are treated as intervals.

• Let’s say reported concentration is 1 mg/L
• We code this as: ConcLow = 1.0 and ConcHigh = 1.0
• The interval for this data point is then 1.0 to 1.0

• For a value reported as “less than 1.0 mg/L”
• We code this as: ConcLow = NA and ConcHigh = 1.0
• The interval for this data point is then 0.0 to 1.0

All of the “weighted regressions” in WRTDS are really “survival regression” (the function survreg in R) which is design for data reported as an interval.
Censored values and compound analytes

Sometimes an analyte of interest is the sum of two or more measured analytes. Here is a real example for Total Nitrogen in the Susquehanna River, Maryland, April 27, 1988.

• The rule is: Compute Total N as Ammonia plus organic N, unfiltered + Nitrate plus nitrite, filtered

The two analyte values were reported as $<0.2$ and $0.9$ mg/L respectively. Therefore, this data point has ConcLow = 0.9 and ConcHigh = 1.1.

• The conventional left-censored approach calls this (0,1.1)
• WRTDS calls this (0.9 to 1.1)
For Non-USGS data from the Water Quality Portal

```r
> siteID <- "IL_EPA_WQX-BPK-07"
> characteristicName <- "Inorganic nitrogen (nitrate and nitrite)"
> startDate <- "2005-01-01"
> endDate <- "2012-01-01"
> Sample <- getSTORETSampleData(siteID,characteristicName,startDate,endDate)
> summary(Sample)
```

```r
> siteID <- "IL_EPA_WQX-BPK-07"
> characteristicName <- "Inorganic nitrogen (nitrate and nitrite)"
> startDate <- "2005-01-01"
> endDate <- "2012-01-01"
> Sample <- suppressWarnings(getSTORETSampleData(siteID,characteristicName,startDate,endDate))
> summary(Sample)
```

- **Date**
  - Min.: 2005-01-24
  - 1st Qu.: 2009-02-08
  - Median: 2010-01-07
  - Mean: 2009-05-21
  - 3rd Qu.: 2011-03-03
  - Max.: 2011-11-28
  - NA's: 8

- **ConcLow**
  - Min.: 0.041
  - 1st Qu.: 3.618
  - Median: 5.205
  - Mean: 4.834
  - 3rd Qu.: 6.560
  - Max.: 11.400

- **ConcHigh**
  - Min.: 0.0180
  - 1st Qu.: 0.1905
  - Median: 4.5950
  - Mean: 3.8710
  - 3rd Qu.: 6.225
  - Max.: 11.4000

- **Uncen**
  - Min.: 0.0
  - 1st Qu.: 1.0
  - Median: 1.0
  - Mean: 0.8
  - 3rd Qu.: 1.0
  - Max.: 1.0

- **ConcAve**
  - Min.: 0.0090
  - 1st Qu.: 0.1905
  - Median: 4.5950
  - Mean: 3.8692
  - 3rd Qu.: 6.2250
  - Max.: 11.4000

- **Julian**
  - Min.: 56636
  - 1st Qu.: 58112
  - Median: 58446
  - Mean: 58215
  - 3rd Qu.: 58866
  - Max.: 59135

- **Month**
  - Min.: 1.000
  - 1st Qu.: 4.000
  - Median: 6.500
  - Mean: 6.425
  - 3rd Qu.: 9.000
  - Max.: 12.000

- **Day**
  - Min.: 1.0
  - 1st Qu.: 96.5
  - Median: 184.0
  - Mean: 179.5
  - 3rd Qu.: 256.2
  - Max.: 349.0

- **DecYear**
  - Min.: 2005
  - 1st Qu.: 2009
  - Median: 2010
  - Mean: 2009
  - 3rd Qu.: 2011
  - Max.: 2012

- **MonthSeq**
  - Min.: 1861
  - 1st Qu.: 1910
  - Median: 1921
  - Mean: 1913
  - 3rd Qu.: 1934
  - Max.: 1943

- **SinDY**
  - Min.: -0.997643
  - 1st Qu.: -0.742051
  - Median: -0.008572
  - Mean: -0.011740
  - 3rd Qu.: 0.739359
  - Max.: 0.998674

- **CosDY**
  - Min.: -0.99908
  - 1st Qu.: -0.69469
  - Median: -0.16188
  - Mean: -0.07631
  - 3rd Qu.: 0.62395
  - Max.: 0.98673

```r
> length(Sample$Date)
[1] 40
> ```
Find a nearby USGS streamgage that can provide appropriate discharge data to use with these water quality data.
Retrieve daily discharge data

```r
> siteID <- "03336645"
> Daily <- getDVData(siteID,"00060",startDate,endDate)
```

There are 2557 data points, and 2557 days.

```r
> summary(Daily)

    Date         Q     Julian     Month     Day DecYear MonthSeq
Min. :2005-01-01 Min. : 0.1727 Min. :56613 Min. : 1.000 Min. : 1.0 Min. :2005 Min. :1861
1st Qu.:2006-10-02 1st Qu.: 1.4725 1st Qu.:57252 1st Qu.: 4.000 1st Qu.: 93.0 1st Qu.:2007 1st Qu.:1882
Mean :2008-07-02 Mean :12.2350 Mean :57891 Mean : 6.522 Mean :183.7 Mean :2009 Mean :1903
3rd Qu.:2010-04-02 3rd Qu.:12.4877 3rd Qu.:58530 3rd Qu.:10.000 3rd Qu.:275.0 3rd Qu.:2010 3rd Qu.:1924

    Qualifier   i   LogQ   Q7   Q30
Length:2557 Min. : 1 Min. :-1.7560 Min. : 0.1905 Min. : 0.2219
Class :character 1st Qu.: 640 1st Qu.: 0.3869 1st Qu.: 1.6768 1st Qu.: 2.3293
Mode :character Median :1279 Median : 1.5416 Median : 5.2548 Median : 7.6092
Mean :1279 Mean : 1.4538 Mean :12.2107 Mean :11.9463
3rd Qu.:1918 3rd Qu.: 2.5247 3rd Qu.:14.1463 3rd Qu.:19.9398
Max. :2557 Max. : 5.7771 Max. :151.2524 Max. :55.4217
NA's :6 NA's :29
```

```r
> length(Daily$Q)
[1] 2557
```

>
Data can then be used for EGRET or other analysis. This plot uses the `boxConcMonth()` function.
Unit values retrieval (not used by EGRET)

- Raccoon River at Van Metre, IA
- Nitrate sensor data
- March – Sept 2013

```r
> siteID <- "05484500"
> parameterCD <- "99133"
> UnitVals <- retrieveUnitNWISData
  (siteID, parameterCD, "2013-03-01",
   "2013-09-30")
> Summary(UnitVals)
```
<table>
<thead>
<tr>
<th>agency</th>
<th>site</th>
<th>dateTime</th>
<th>tz_cd</th>
<th>X18_99133_00011</th>
<th>X18_99133_00011_cd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Length: 17720

Min. : 2013-03-20 00:00:00 CST: 17720

Min. : 1.260 A: 17720

1st Qu.: 2013-05-08 16:11:15 1st Qu.: 2.540


3rd Qu.: 2013-08-15 15:33:45 3rd Qu.: 16.900

Max. : 2013-09-30 23:45:00 Max. : 20.700

> plot(UnitVals$dateTime, UnitVals$X18_99133_00011, type="l", ylim=c(0,22), yaxs="i")
Back to our main EGRET example, metadata is next

• The `getMetaData` function is used to retrieve, input, and save metadata about the current analysis

• It defines what data you have and where they came from

• Inputs are used for labels on figures & tables

• Inputs are used to name your saved workspaces, using abbreviations
INFO<-getMetaData(siteID,parameter_cd)

Your site for streamflow data is 01491000.

Your site name is CHOPTANK RIVER NEAR GREENSBORO, MD, but you can modify this to a short name in a style you prefer.

This name will be used to label graphs and tables.

If you want the program to use the name given above, just do a carriage return, otherwise enter the preferred short name (no quotes):

<cr>
The latitude and longitude of the site are: 38.99719, -75.78581 (degrees north and west).

The drainage area at this site is 113 square miles which is being stored as 292.6687 square kilometers.

It is helpful to set up a station abbreviation when doing multi-site studies, enter a unique id (three or four characters should work).

It is case sensitive. Even if you don't feel you need an abbreviation for your site you need to enter something (no quotes):

Chop
Your water quality data are for parameter number 00631 which has the name:' Nitrate plus nitrite, water, filtered, milligrams per liter as nitrogen '.

Typically you will want a shorter name to be used in graphs and tables. The suggested short name is:' Nitrate-nitrite '.

If you would like to change the short name, enter it here, otherwise just hit enter (no quotes):

Nitrate, filtered, as N
The units for the water quality data are: mg/l as N.

It is helpful to set up a constituent abbreviation when doing multi-constituent studies, enter a unique id (three or four characters should work something like tn or tp or NO3).

It is case sensitive. Even if you don't feel you need an abbreviation you need to enter something (no quotes):

no3
If you are using supplied data, you still must run the `getMetaData` command:

```r
> INFO <- getMetaData()
```

The program will then prompt you to enter metadata about your site and study.

All metadata is voluntary except the following required fields:

- A site name
- A parameter name
- A site abbreviation
- A parameter abbreviation
Two more commands before we can start our analysis of the data

> Sample<-mergeReport()

Discharge Record is 11718 days long, which is 32 years
First day of the discharge record is 1979-09-01 and last day is 2011-09-30
The water quality record has 654 samples
The first sample is from 1979-09-25 and the last sample is from 2011-09-29
Discharge: Minimum, mean and maximum 0.00991 4.08 246
Concentration: Minimum, mean and maximum 0.05 1.1 2.4
Percentage of the sample values that are censored is 0.15 %
Let’s look at the data before we proceed, the function is:

```r
> multiPlotDataOverview(qUnit=1)
```
We’ve gone to all this effort, let’s save our work

> savePath<-"/Users/rhirsch/Desktop/"
> saveResults(savePath)

Save it over and over as you proceed and add results
We now have 3 data frames

- Sample (654 rows, 14 columns)
- Daily (11,718 rows, 12 columns)
- INFO (1 row, 42 columns)
> modelEstimation()

- Runs the model in cross-validation mode
- Estimates the “surface” for concentration as a function of time and discharge
- Uses the surface to compute daily values of
  - Concentration
  - Flux
  - Flow-normalized concentration
  - Flow-normalized flux
- Adds those to the Daily data frame

User has choices about some of the parameters of the WRTDS model
Now what is in Daily?
It is a data frame that has dimension (11718, 19)

```
> summary(Daily)

  Date             Q   Julian   Month    Day  DecYear MonthSeq
Min. :1979-09-01  Min. :9.911e-03 Min. :47359 Min. :1.000 Min. :1.0 Min. :1980 Min. :1557
1st Qu.:1987-09-08 1st Qu.:9.345e-01 1st Qu.:50288 1st Qu.:4.000 1st Qu.:92.0 1st Qu.:1988 1st Qu.:1653
Mean   :1995-09-15  Mean   :4.082e+00 Mean   :53218 Mean   :6.529 Mean   :183.3 Mean   :1996 Mean   :1749
3rd Qu.:2003-09-22 3rd Qu.:14.616e+00 3rd Qu.:56147 3rd Qu.:10.000 3rd Qu.:274.0 3rd Qu.:2004 3rd Qu.:1845

  Qualifier  i  LogQ    Q7  Q30  Leap
Length:11718  Min. : 1  Min. :-4.61412 Min. :0.01808 Min. :0.09606 Min. :1.0
Class:character  1st Qu.: 2930 1st Qu.:0.06779 1st Qu.:0.99109 1st Qu.:1.17609 1st Qu.:92.0
Mode:character  Median : 5860 Median :0.87835 Median :2.55661 Median :2.87133 Median :184.0
Mean : 5860 Mean :0.76602 Mean :4.08154 Mean :4.08059 Mean :183.3
3rd Qu.: 8789 3rd Qu.:1.52945 3rd Qu.:4.91095 3rd Qu.:5.68036 3rd Qu.:274.0
Max. :11718 Max. :5.50678 Max. :84.00395 Max. :365.0

  yHat  SE  ConcDay  FluxDay  FNConc  FNFlux
Min. : -1.534999 Min. : 0.1347 Min. : 0.230 Min. : 1.633 Min. : 0.8208 Min. : 77.3
1st Qu.-0.006091 1st Qu.: 0.2186 1st Qu.: 1.033 1st Qu.: 98.053 1st Qu.:1.0524 1st Qu.:168.3
Median : 0.130494 Median : 0.2504 Median : 1.195 Median :248.578 Median :1.2042 Median :317.3
Mean : 0.117362 Mean : 0.2678 Mean : 1.194 Mean :365.030 Mean :1.1975 Mean :361.5
3rd Qu.: 0.258278 3rd Qu.: 1.353 3rd Qu.:482.075 3rd Qu.:1132.86 3rd Qu.:1536.0
Max. : 0.659426 Max. : 1.962 Max. :5705.826 Max. : 1.7017 Max. : 928.3
```
“Period of Analysis” concept in EGRET.

- Could be water year
- Could be calendar year
- Could be April-May-June
- Could be Dec-Jan-Feb-Mar
- Could be only May…

paStart = calendar month that starts Period
paLong = length of Period, in months
Period of analysis set up

Say we want calendar year
INFO <- setPA(paStart = 1, paLong=12)

Say we want April, May, June
INFO <- setPA(paStart = 4, paLong = 3)

Default is water year
Units in EGRET

Everything stored as:

$\text{m}^3/\text{s}, \ \text{kg/day}, \ \text{or mg/L}$

But each graphic or table has a wide choice of units (English and SI) that the user can select

Now lets see some trend results
plotConcHist(1980, 2012)

Choptank River near Greensboro, MD  Nitrate, filtered, as N
Water Year
Mean Concentration (dots) & Flow Normalized Concentration (line)

Concentration in mg/L
0.0  0.5  1.0  1.5  2.0
> plotFluxHist(1980,2012,fluxUnit=8)

Choptank River near Greensboro, MD  Nitrate, filtered, as N
Water Year
Flux Estimates (dots) & Flow Normalized Flux (line)
Choptank River near Greensboro, MD
Nitrate, filtered, as N
Water Year

<table>
<thead>
<tr>
<th>Year</th>
<th>Discharge (cfs)</th>
<th>Conc (mg/L)</th>
<th>FN_Conc</th>
<th>Flux (tons/yr)</th>
<th>FN_Flux</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>150.2</td>
<td>0.932</td>
<td>0.992</td>
<td>126.0</td>
<td>116</td>
</tr>
<tr>
<td>1981</td>
<td>78.3</td>
<td>1.025</td>
<td>0.988</td>
<td>73.7</td>
<td>119</td>
</tr>
<tr>
<td>1982</td>
<td>107.6</td>
<td>1.025</td>
<td>0.982</td>
<td>107.9</td>
<td>120</td>
</tr>
<tr>
<td>1983</td>
<td>176.1</td>
<td>0.995</td>
<td>0.983</td>
<td>143.9</td>
<td>122</td>
</tr>
<tr>
<td>1984</td>
<td>201.9</td>
<td>0.981</td>
<td>0.994</td>
<td>173.3</td>
<td>124</td>
</tr>
<tr>
<td>1985</td>
<td>53.6</td>
<td>1.056</td>
<td>1.012</td>
<td>53.8</td>
<td>127</td>
</tr>
<tr>
<td>1986</td>
<td>92.8</td>
<td>1.060</td>
<td>1.036</td>
<td>99.0</td>
<td>131</td>
</tr>
<tr>
<td>1987</td>
<td>119.1</td>
<td>1.079</td>
<td>1.062</td>
<td>125.2</td>
<td>134</td>
</tr>
<tr>
<td>1988</td>
<td>66.0</td>
<td>1.120</td>
<td>1.086</td>
<td>72.6</td>
<td>137</td>
</tr>
<tr>
<td>1989</td>
<td>198.2</td>
<td>1.057</td>
<td>1.108</td>
<td>180.4</td>
<td>140</td>
</tr>
<tr>
<td>1990</td>
<td>141.5</td>
<td>1.118</td>
<td>1.126</td>
<td>148.7</td>
<td>142</td>
</tr>
<tr>
<td>1991</td>
<td>97.0</td>
<td>1.174</td>
<td>1.144</td>
<td>108.1</td>
<td>144</td>
</tr>
<tr>
<td>1992</td>
<td>77.2</td>
<td>1.204</td>
<td>1.158</td>
<td>89.2</td>
<td>145</td>
</tr>
<tr>
<td>2007</td>
<td>151.2</td>
<td>1.408</td>
<td>1.382</td>
<td>176.0</td>
<td>164</td>
</tr>
<tr>
<td>2008</td>
<td>90.5</td>
<td>1.476</td>
<td>1.401</td>
<td>111.1</td>
<td>165</td>
</tr>
<tr>
<td>2009</td>
<td>130.0</td>
<td>1.410</td>
<td>1.420</td>
<td>147.4</td>
<td>165</td>
</tr>
<tr>
<td>2010</td>
<td>254.0</td>
<td>1.324</td>
<td>1.438</td>
<td>247.1</td>
<td>165</td>
</tr>
<tr>
<td>2011</td>
<td>185.2</td>
<td>1.441</td>
<td>1.458</td>
<td>175.0</td>
<td>164</td>
</tr>
</tbody>
</table>
How to see trends as numbers


tableChange(fluxUnit=5,yearPoints)
Choptank River near Greensboro, MD
Nitrate, filtered, as N
Water Year

### Concentration trends

<table>
<thead>
<tr>
<th>time span</th>
<th>change mg/L</th>
<th>slope mg/L/yr</th>
<th>change %</th>
<th>slope %/yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980 to 1995</td>
<td>0.21</td>
<td>0.014</td>
<td>21</td>
<td>1.4</td>
</tr>
<tr>
<td>1980 to 2011</td>
<td>0.47</td>
<td>0.015</td>
<td>47</td>
<td>1.5</td>
</tr>
<tr>
<td>1995 to 2011</td>
<td>0.26</td>
<td>0.016</td>
<td>22</td>
<td>1.4</td>
</tr>
</tbody>
</table>

### Flux Trends

<table>
<thead>
<tr>
<th>time span</th>
<th>change tons/yr</th>
<th>slope tons/yr/yr</th>
<th>change %</th>
<th>slope %/yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980 to 1995</td>
<td>31</td>
<td>2.1</td>
<td>27</td>
<td>1.8</td>
</tr>
<tr>
<td>1980 to 2011</td>
<td>48</td>
<td>1.6</td>
<td>42</td>
<td>1.3</td>
</tr>
<tr>
<td>1995 to 2011</td>
<td>17</td>
<td>1.1</td>
<td>12</td>
<td>0.73</td>
</tr>
</tbody>
</table>
I’m going to switch data sets to Nitrate for the Raccoon River at Des Moines Iowa
EGRET produces a diagnostic plot to help spot serious problems with the model. 

`fluxBiasMulti(fluxUnit=4)`
This same type of plot can be used to look at other models, here the LOADEST7 fluxBiasMulti(fluxUnit=4)
Diagnostics and potential problems with estimating mean flux, see:


Bottom line, look at the fit before you use a statistical model!!!
How difficult is it to make those contour plots?

>plotContours(yearStart=1970, yearEnd=2005, qBottom=2, qTop=200, qUnit=2, contourLevels=seq(0,300,50))
> plotConcTime(qUnit=1,qUpper=50,paLong=8,paStart=6,concMax=2.5)

Choptank River near Greensboro, MD, Nitrate, filtered, as N
Season Consisting of Jun Jul Aug Sep Oct Nov Dec Jan
For Discharge < 50 Cubic Feet per Second

> plotConcTime(qUnit=1,qLower=200,paLong=5,paStart=12,concMax=2.5)

Choptank River near Greensboro, MD, Nitrate, filtered, as N
Season Consisting of Dec Jan Feb Mar Apr
For Discharge > 200 Cubic Feet per Second
When all is said and done: The only way to figure out what is happening to our planet is to measure it, and this means tracking changes decade after decade and poring over the records.
dataDelivery and EGRET

- Information and software available at: https://github.com/USGS-R/EGRET/wiki