Outline

- Context

- General Assessment Approach

  - Organization, Monitoring, Assessment Inputs/Outputs

- The Way DHEC Used to Do Assessment

- The New Way – Overview of R Statistical Software

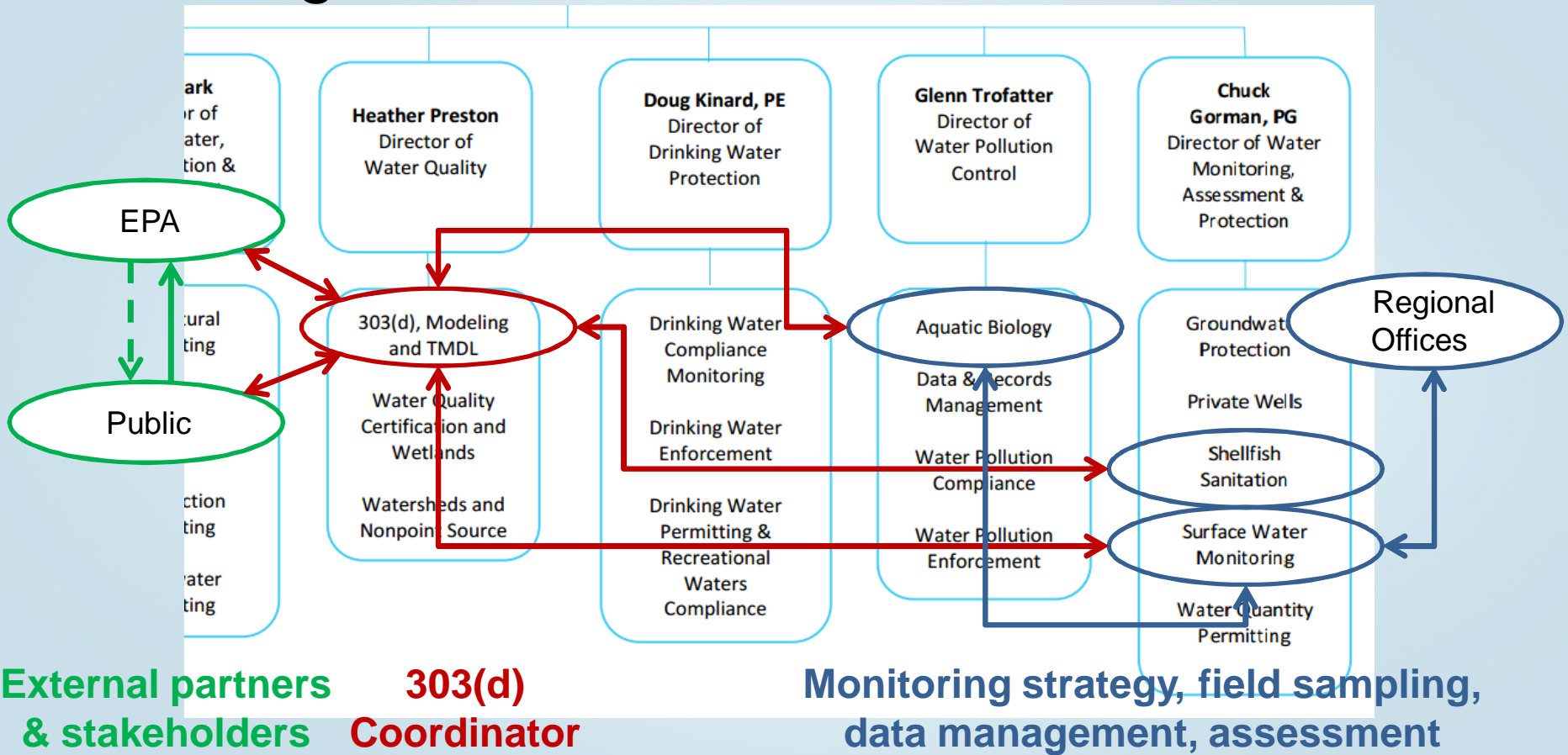- Resources for Additional Information

- Questions

Context

- Presenting work of DHEC's Surface Water Monitoring Section
  - Develop water quality monitoring strategy
  - Direct field sampling
  - Manage data
  - Perform the assessment for 303(d) & Integrated Report
- EPA's assessment staff have noted DHEC's R-based assessment tools and thought a presentation of DHEC's approach would be useful along side the Data Discovery Tool
- Presentation Goal: Overview of assessment approach and why R is useful for DHEC, not details or how it's done

South Carolina Department of Health and Environmental Control
Promoting and Protecting the Health of the Public and the Environment

D H E C
PROMOTE PROTECT PROSPER

Water Quality Monitoring

- Strategy updated annually
- Defines stations, parameters, sampling frequency, etc.
- Base Sites – 250 base sites sampled every other month
- Special Request Sites – 70 NPS 319 project sites in 2015, additional sites for TMDLs and WLA/NPDES permitting
- Statistical Survey Sites – 90 new randomly selected sites each year sampled once per month for 12 months
  - 303(d) List and 305(b) Assessment

Additional Monitoring

- Aquatic Biology Section
  - Fish tissue Hg and PCBs
  - Benthic macroinvertebrates
- Shellfish Sanitation Section
  - Fecal coliform sampling in shellfish harvesting areas
- Beach Monitoring
  - Enterococci at ocean beaches
- Outside Data
  - Requires approved QAPP

South Carolina Department of Health and Environmental Control
Promoting and Protecting the Health of the Public and the Environment

DHEC
PROMOTE PROTECT PROSPER

Monitoring Sites

2000 Sites Assessed
In 2016

Station Types
Macroinvertebrates
Beaches
Water Quality
Fish Tissue
Random
Shellfish Areas
Special Study
External QAPP

*Includes ~500
carryovers

Table Of Contents

Layers
☑ Draft_DHEC_303_16
MULTI_TYPE

- BEACH
- FISH
- FISH_SPECIAL
- MACRO
- MACROINVERTEBRATES
- MACRO_PERI
- MACRO_PERI_FISH_SPECIAL
- MACRO_PERI_RAND
- MACRO_PERI_SPECIAL
- MACRO_RAND
- MACRO_SPECIAL
- QAPP
- RANDOM
- SHELLFISH
- SPECIAL
- WQMS
- WQMS_FISH
- WQMS_FISH_SPECIAL
- WQMS_MACRO_PERI
- WQMS_MACRO_PERI_FISH
- WQMS_MACRO_PERI_SPECIAL
- WQMS_PERI_FISH
- WQMS_SPECIAL
☑ sc_bnd
☑ World Physical Map
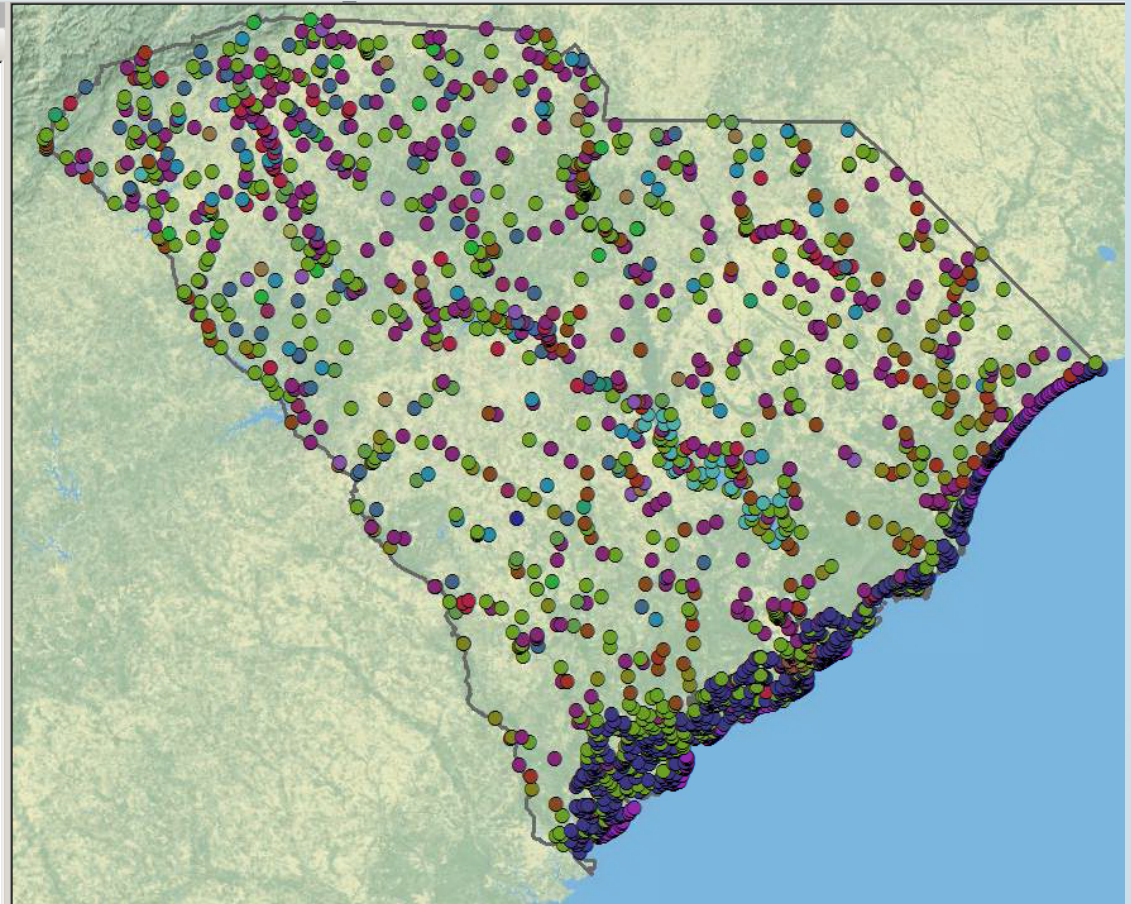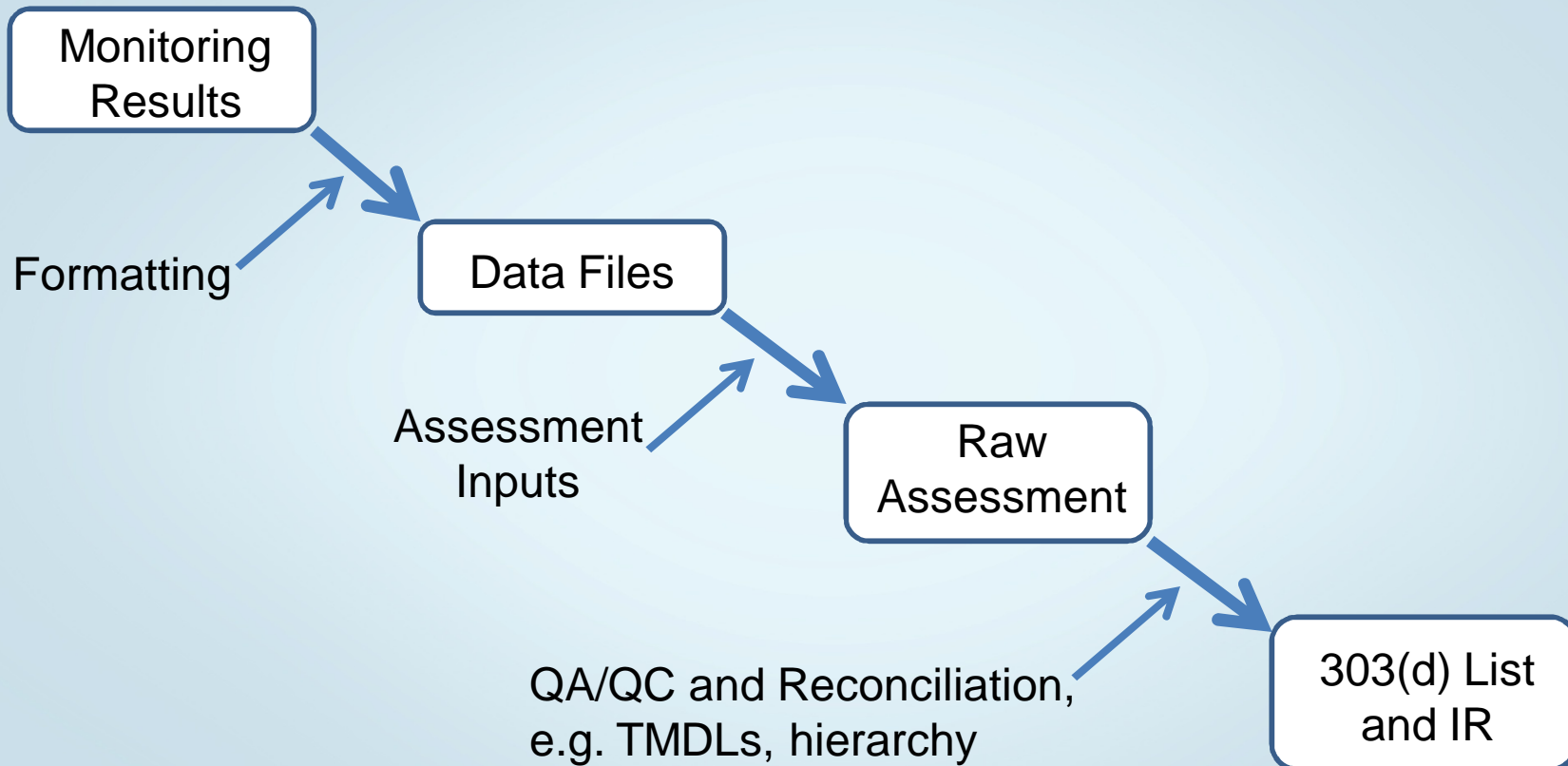
Classifications and Standards

- Water Classifications – Freshwater, Trout Water, Saltwater Class A, Saltwater Class B, Shellfish Harvesting Water

- Water Quality Standards

  - magnitude, frequency & duration components

  - acute and chronic criteria

  - Site-specific standards

  - pH, temperature & salinity dependence (ammonia toxicity)

- Waterbody Type – streams, lakes < 40 acres, lakes > 40 acres, estuaries, ecoregion

Assessment Methodology

- (F)ull , (P)artial , (N)ot supporting

- Bacteria and Conventional Parameters

  - F = ≤ 10% Exceeds

  - P = >10% and ≤ 25%

  - N = >25%

- Toxics

  - F = no more then 1 exceedance

  - P = 2 or more exceedance ≤10%

  - N = 2 or more and > 10%

- Partial and Not Supporting are on the 303(d) List

Data Processing

- Data Blending
  - Proximity – due to random generation of statistical survey site locations, data may be pooled for adjacent sites
  - Standards change during assessment window
  - Different sampling frequencies

- Trend Analysis
  - Seasonal Kendall's Tau Test
  - 15 years
  - Minimum 30 data points

Assessment Task

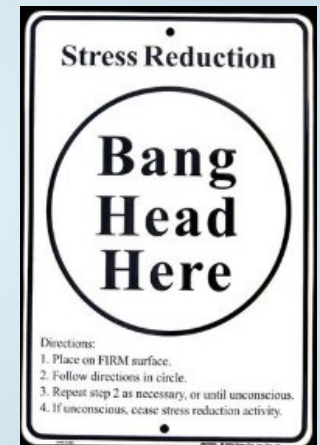- Compile monitoring results and format data files
- Record-by-record, determine appropriate criteria depending on water class, water quality standard, waterbody type, ancillary chemistry
- Station-by-station and parameter-by-parameter, determine exceedances, use support, and listing status
- Perform QA/QC
- Reconcile sites with TMDLs – if yes, category 4a; if no category 5 (303d List)
- Point is, it's a challenge

The Way DHEC Used to Do Assessment, or so I'm told

- Initially, some combination of legacy STORET tools and brute force FoxPro and Excel, **said to take months**

- Later, Agency programmers developed an assessment in C programming language and then another in SAS, apparently **didn't work very well**

- Then, legend has it, WQHydro trend analysis software was **lost in the fire** which gutted part of the monitoring annex

- Submittals were timely, but **assessments were hard**, and reliving those dark days with assessment staff can leave one with a terrible sense of **woe**

Stress Reduction

Bang Head Here

Directions:
1. Place on FIRM surface.
2. Follow directions in circle.
3. Repeat step 2 as necessary, or until unconscious.
4. If unconscious, cease stress reduction activity.

South Carolina Department of Health and Environmental Control
Promoting and Protecting the Health of the Public and the Environment

DHEC
PROMOTE PROTECT PROSPER

The Way DHEC Does Assessment Now

- Assessment staff use R statistical software to perform most assessment operations

- Most data management is done outside of the R environment (text files, Excel, Access)

- Modules and Functions in R access data, perform calculations, write results, create tables that become the final assessment

- Can be used to create data plots and KML files for mapping assessment results

What is R ?

- R is a language and environment for statistical computing and graphics

- Free open-source software

- Includes a core set of statistical packages

- Additional packages developed by the user community

- Highly adaptable to specific situations

- Provided and supported by:

  R Foundation for Statistical Computing

  (https://www.r-project.org/)

What R Looks Like

```
1   trenddata <- DATASET %>%
2     filter(TREND == "Y" & Year > LastDataYear - 15) %>%
3     select(c(STAT, PARM, DateTime, ResultTrend)) %>%
4     mutate(Month = format(DateTime, "%m")) %>%
5     mutate(Year = format(DateTime, "%Y")) %>%
6     left_join(MonthCrossWalk)
7
8   trendbimonth <- trenddata %>%
9     select(STAT,PARM,Year,MonthGroup,ResultTrend) %>%
10    group_by(STAT,PARM,Year,MonthGroup) %>%
11    summarise(ResultTrend = mean(ResultTrend))
12
13  trendbimonthcount <- trendbimonth %>%
14    select(STAT,PARM,ResultTrend) %>%
15    group_by(STAT,PARM) %>%
16    summarize(Count = n())
17
18  trendbimonth %<>% left_join(trendbimonthcount) %>%
19    filter(Count >= 30)
20
21  statvector <- unique(trendbimonth[, c("STAT", "PARM")])
22  yearvector <- sort(unique(trendbimonth$Year))
23  MonthGroupvector <- sort(unique(trendbimonth$MonthGroup))
24  yearMonthGroup <- merge(yearvector, MonthGroupvector)
25  names(yearMonthGroup) <- c("Year", "MonthGroup")
26  yearMonthGroup$Year <- as.character(yearMonthGroup$Year)
27  yearMonthGroup$MonthGroup <- as.character(yearMonthGroup$MonthGroup)
28
29  trendbimonthts <- merge(statvector,yearvector)
30
```

South Carolina Department of Health and Environmental Control
Promoting and Protecting the Health of the Public and the Environment

DHEC

PROMOTE PROTECT PROSPER

Why R Works For DHEC

- We didn't already have an efficient approach and weren't already invested in something else

- Talented individual on staff overcame the R learning curve and now does code, customization, etc.

- Flexible – R-based tools are readily adapted to SC specifics in house

- Free – no license issues

- Powerful – large user base and growing contributing group adding new functions

- Fast – staff time doing assessment reduced from months to days-weeks (depending on changes between cycles)

Standards Assessment

- Select parameters to assess

Microsoft Excel - Criterion.xls

File Edit View Insert Format Tools Data Window Help

Arial 10

B2 FW

	A	B	C	D	E	F	G	H	I	J	K	L
1	PARM	CLASS	WBODY	ECOREGION	LT	GT	TOXIC	NUTRIENTS	CHRONIC	ACUTE	DL	Blank
14	00076	TPGT				10	N	N			0.5	1
15	00076	TPGT-SP				10	N	N			0.5	1
16	00076	TPT				10	N	N			0.5	1
17	00300	FW-SP			4		N	N				1
18	00300	FW			5		N	N				2
19	00300	SA			5		N	N				2
20	00300	SA-SP			4		N	N				1
21	00300	SB			4		N	N				2
22	00300	SB-SP			5		N	N				1
23	00300	SFH			5		N	N				2
24	00300	TN			6		N	N				12
25	00300	TPGT			6		N	N				1
26	00300	TPGT-SP			5		N	N				2
27	00300	TPT			6		N	N				12
28	00400	FW			6	8.5	N	N				1
29	00400	FW-SP			5	8.5	N	N				2
30	00400	SA			6.5	8.5	N	N				1
31	00400	SA-SP			6.5	8.5	N	N				3
32	00400	SB			6.5	8.5	N	N				13
33	00400	SB-SP			6.5	8.5	N	N				13
34	00400	SFH			6.5	8.5	N	N				13
35	00400	TN			6	8	N	N				1
36	00400	TPGT			6	8.5	N	N				1
37	00400	TPGT-SP			6	8.5	N	N				3
38	00400	TPT			6	8.5	N	N				1
39	00600		L	BLUE RIDGE		0.35	N	Y				3
40	00600		L	PIEDMONT		1.5	N	Y				1
41	00600		L	SOUTHEASTERN PLAINS		1.5	N	Y				1
42	00600		L	MIDDLE ATLANTIC COASTAL PLAIN		1.5	N	Y				2
43	00610	FW					Y	N	0.179	1.32	0.05	2
44	00610	FW-SP					Y	N	0.179	1.32	0.05	2
45	00610	SA					Y	N	0.05	0.35	0.05	2
46	00610	SA-SP					Y	N	0.05	0.35	0.05	2
47	00610	SB					Y	N	0.05	0.35	0.05	4

Sheet1 / Sheet2 / Sheet3 /

Attributes Table

Microsoft Excel - ExportAttributes.csv

File   Edit   View   Insert   Format   Tools   Data   Window   Help

Arial   10

A1   STAT

	A	B	C	D	E	F	G	H
1	STAT	WBODY	CLASS1	CLASS2	PREVCLASS	ECOREGION	LONGITUDE	LATITUDE
2	MD-069	E	SB	SFH		SOUTHERN COASTAL PLAIN	-79.84221069	32.7728285
3	MD-209	E	ORW		SFH	SOUTHERN COASTAL PLAIN	-80.16433819	32.62230612
4	PD-040	S	FW-SP			SOUTHEASTERN PLAINS	-80.3341803	33.8744111
5	PD-141	S	FW			SOUTHEASTERN PLAINS	-79.85619422	34.30323939
6	RL-01011	L	FW			SOUTHEASTERN PLAINS	-80.1981181	33.5126767
7	RL-05398	L	FW			SOUTHEASTERN PLAINS	-79.6731037	34.6347623
8	RS-08233	S	FW			PIEDMONT	-80.46398683	34.63503148
9	S-287	S	FW			PIEDMONT	-81.18636514	34.05384144
10	ST-032	L	FW			MIDDLE ATLANTIC COASTAL PLAIN	-80.0111928	32.93241553
11								

Example Parameter - Ammonia

- Calculates a criterion for each individual result based on necessary ancillary field measurements
  - pH & temp for freshwater
  - pH, temp & salinity for saltwater
- Missing field values
  - Other measurements in the same month are averaged
  - Data from preceding and following month are averaged
  - If no data meeting above criteria, NH3 value not assessed
- Calculates the average of the standards exceedances

Ammonia Results Table

Assessment Output

Microsoft Excel - EPA_ASSESMENT.xls

	A	B	C	D	E	F	G	H	I	J	K	L
1	STATION	TIER	CLASS	WBODY	BIO_SUP	BIO_SOURCE	DO_N	DO_NX	DO_%X	DO_AVGX	DO_SUP	DO_Trend
2	MD-069		SB	E			45	0	0	0	F	NS
3	MD-069		SFH	E			45	3	6.667	4.827	F	NS
4	MD-209		SFH	E			45	16	35.556	4.233	N	NS
5	PD-040		FW-SP	S			12	8	66.667	1.683	N	D
6	PD-141		FW	S			11	1	9.091	4.75	F	NS
7	RL-01011		FW	L			47	0	0	0	F	
8	RL-05398		FW	L			12	0	0	0	F	
9	RS-08233		FW	S	P		11	1	9.091	3.82	F	
10	S-287		FW	S	P		12	0	0	0	F	D
11	ST-032		FW	L			45	1	2.222	4.71	F	NS
12	WAC-028	Tier I										
13	WAC-029	Tier I										
14	WAC-032	Tier I										
15	X-001				P							
16												

Supports Output

Microsoft Excel - EPA_SUPPORT.xls

	STATION	CLASS	WBODY	TIER	OVERALLUSE	RECUSE	ALUSE	AL_F_SUP	AL_P_SUP	AL_N_SUP	BIO	DO	PH	TURB	TN	NH3	TP	CD	CR	CU	PB	NI	ZN	HG	CHLA	FECAL	BEACH
2	MD-069	SB	E		F	F	F	TURB,DO,PH,NH3,CD,CR,CU,PB,NI,ZN,HG				F	F	F			F		F	F	F	F	F	F		F	
3	MD-069	SFH	E		F	F	F	TURB,DO,PH,NH3,CD,CR,CU,PB,NI,ZN,HG				F	F	F			F		F	F	F	F	F	F		F	
4	MD-209	SFH	E		N	F	N	TURB,PH,NH3,CD,CR,CU,PB,NI,ZN,HG		DO		N	F	F			F		F	F	F	F	F	F		F	
5	PD-040	FW-SP	S		N	N	N	TURB,PH,CD,CR,CU,PB,NI,ZN,HG		DO,NH3		N	F	F		N			F	F	F	F	F	F		N	
6	PD-141	FW	S		N	N	N	TURB,DO,PH,CD,CR,CU,PB,NI,ZN,HG		NH3		F	F	F		N			F	F	F	F	F	F		N	
7	RL-01011	FW	L		N	F	N	TURB,DO,TN,NH3,CD,CR,CU,PB,NI,ZN	PH	TP		F	P	F	F	F	N	F	F	F	F	F	F			F	
8	RL-05398	FW	L		N	F	N	DO,PH,TN,NH3,CD,CR,CU,PB,NI,ZN,HG		TURB,TP,CHLA		F	F	N	F	F	N	F	F	F	F	F	F	F	N	F	
9	RS-08233	FW	S		P	P	P	TURB,DO,PH,NH3,CD,CR,CU,PB,NI,ZN,HG	BIO		P	F	F	F		F			F	F	F	F	F	F		P	
10	S-287	FW	S		N	N	N	DO,PH	BIO	TURB	P	F	F	N												N	
11	ST-032	FW	L		N	F	N	TURB,DO,PH,TN,NH3,CD,CR,CU,PB,NI,ZN,HG		TP,CHLA		F	F	F	F	F	N	F	F	F	F	F	F	F	N	F	
12	WAC-028			Tier I	N	N																					N
13	WAC-029			Tier I	P	P																					P
14	WAC-032			Tier I	P	P																					P
15	X-001				P		P		BIO		P																	
16																											
17																											

Example KML File Created Using R
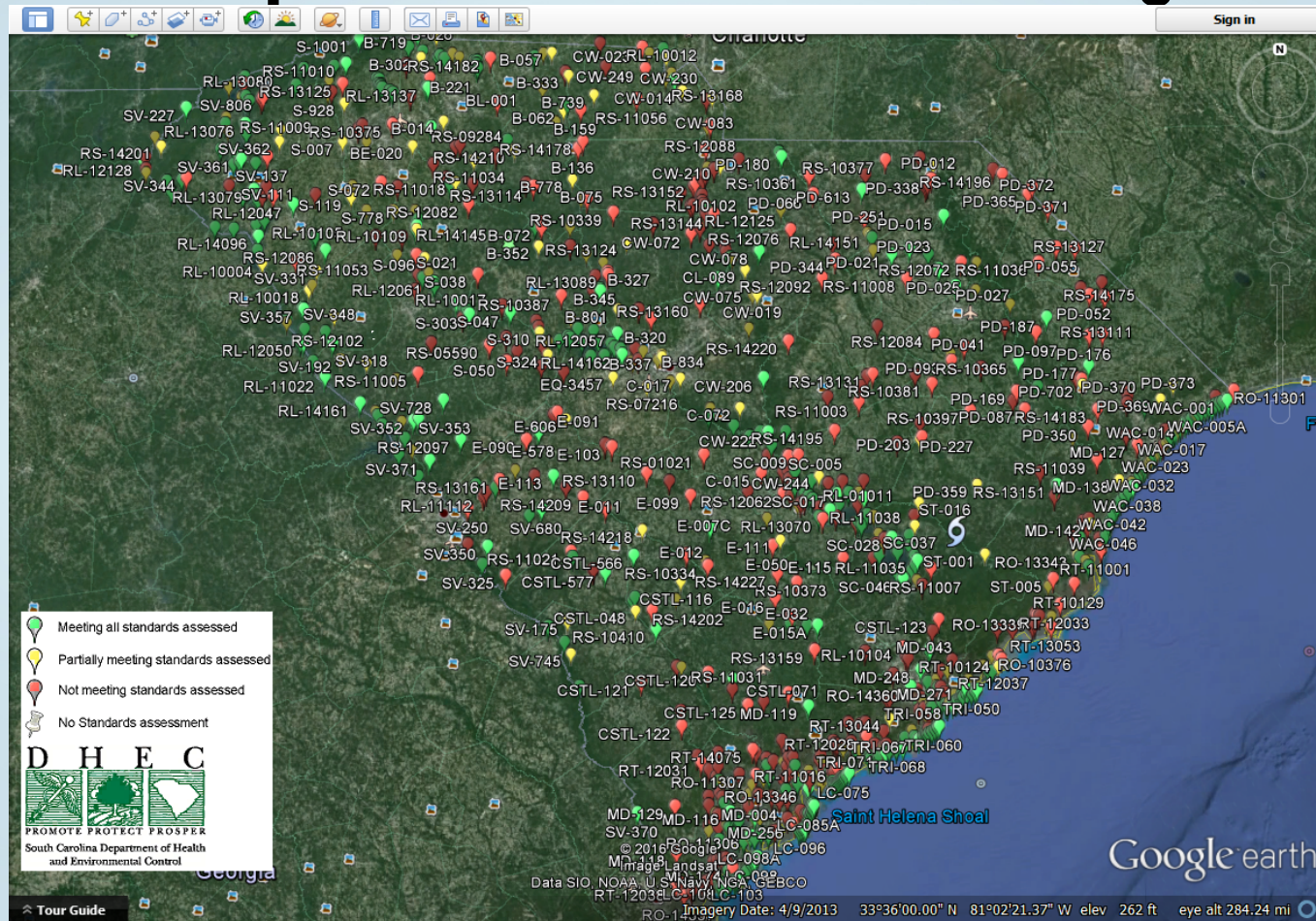
Example Data Plot and Listing History
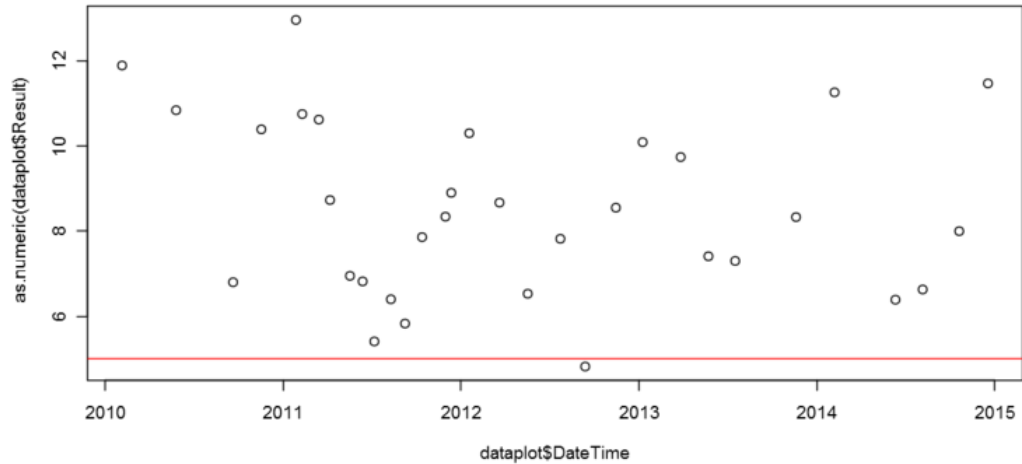
Helps with QA/QC

Counting on R to Help with New ATTAINS
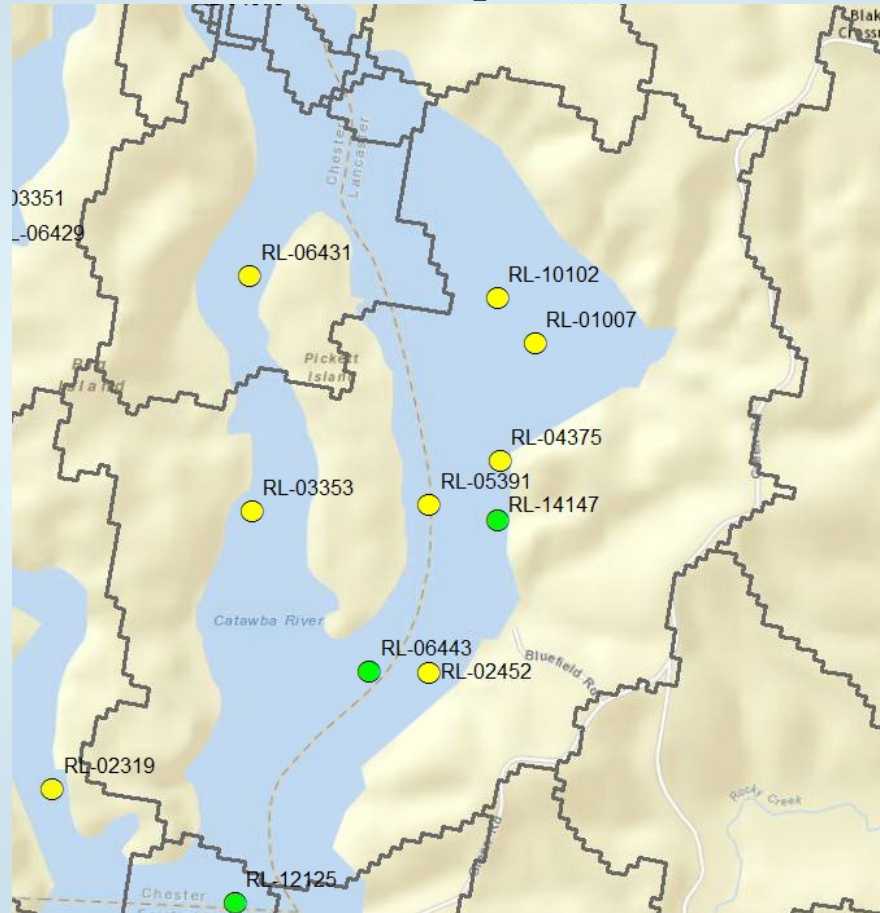
Resources

- R Foundation for Statistical Computing

  https://www.r-project.org/

- R project mailing list and archive

  https://www.r-project.org/mail.html

- Quick-R

  http://www.statmethods.net/

- R Tutorials

  http://ww2.coastal.edu/kingw/statistics/R-tutorials/

Contacts

Using R for assessment:
Bryan Rabon
Surface Water Monitoring Section
SCDHEC – Bureau of Water
803-898-4402
raboneb@dhec.sc.gov

General information:
Wade Cantrell, Manager
303(d), Modeling & TMDL Section
SCDHEC – Bureau of Water
803-898-3548
cantrewm@dhec.sc.gov