Statistical Analysis Phase 1

July, 2009

John W. Green, Ph.D, Ph.D
DuPont Applied Statistics

Data Considered

- Water Samples
  - 42,336 measurements on 5581 samples
- Sediment Samples
  - 10,800 measurements on 1788 samples
- Floodplain
  - 11,280 measurements on 627 samples
- Riverbank
  - 184 samples or predictions
- Loading
  - 40,032 measurements on 2224 samples
- Flow (Discharge)
  - Daily averages at three locations since 1926

Data Considered

- Storm Data
  - 10,062 measurements on 559 samples
    - Included in Loading dataset
- Samples have
  - GPS coordinates
    - Translated to NAD83 coordinates and RRM
  - Date collected
  - Many also have time of sample
  - Other identifying info

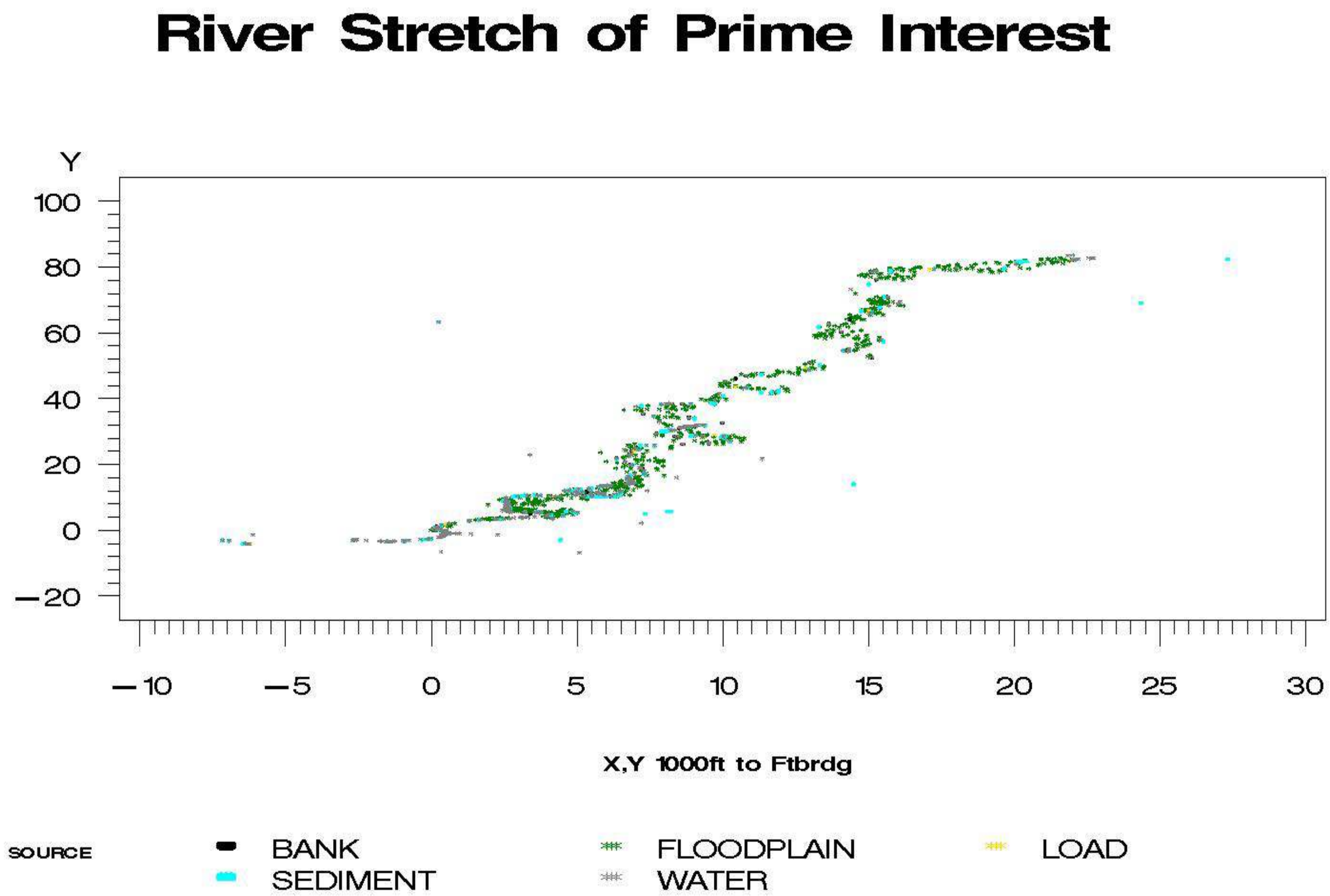
Data to be Included Later

- Biota
  - Approximately 35,000 measurements on 5000 samples
- Storm
  - 70 samples have not yet been identified as to time collected relative to storm beginning
    - Will be done by comparing flow at collection time against daily flow values

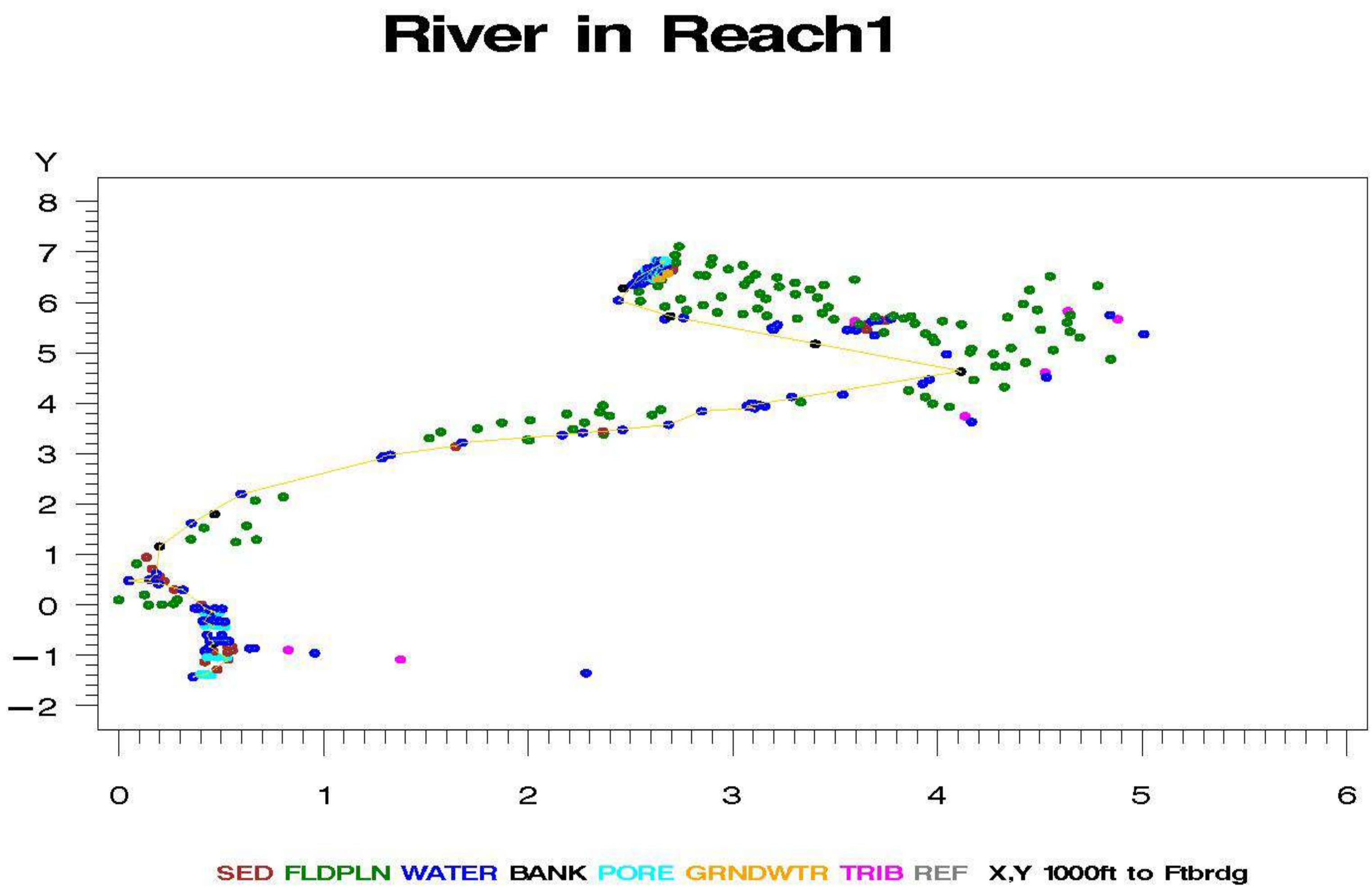
Treatment of Data

- Where replicates were collected, the values across reps of each measurement (chemical) were examined for consistency.
- A few unbelievable extremes (hi or low) were discarded
- Replicate means (or geometric means) were used for analysis
- Distributions were fit to each response

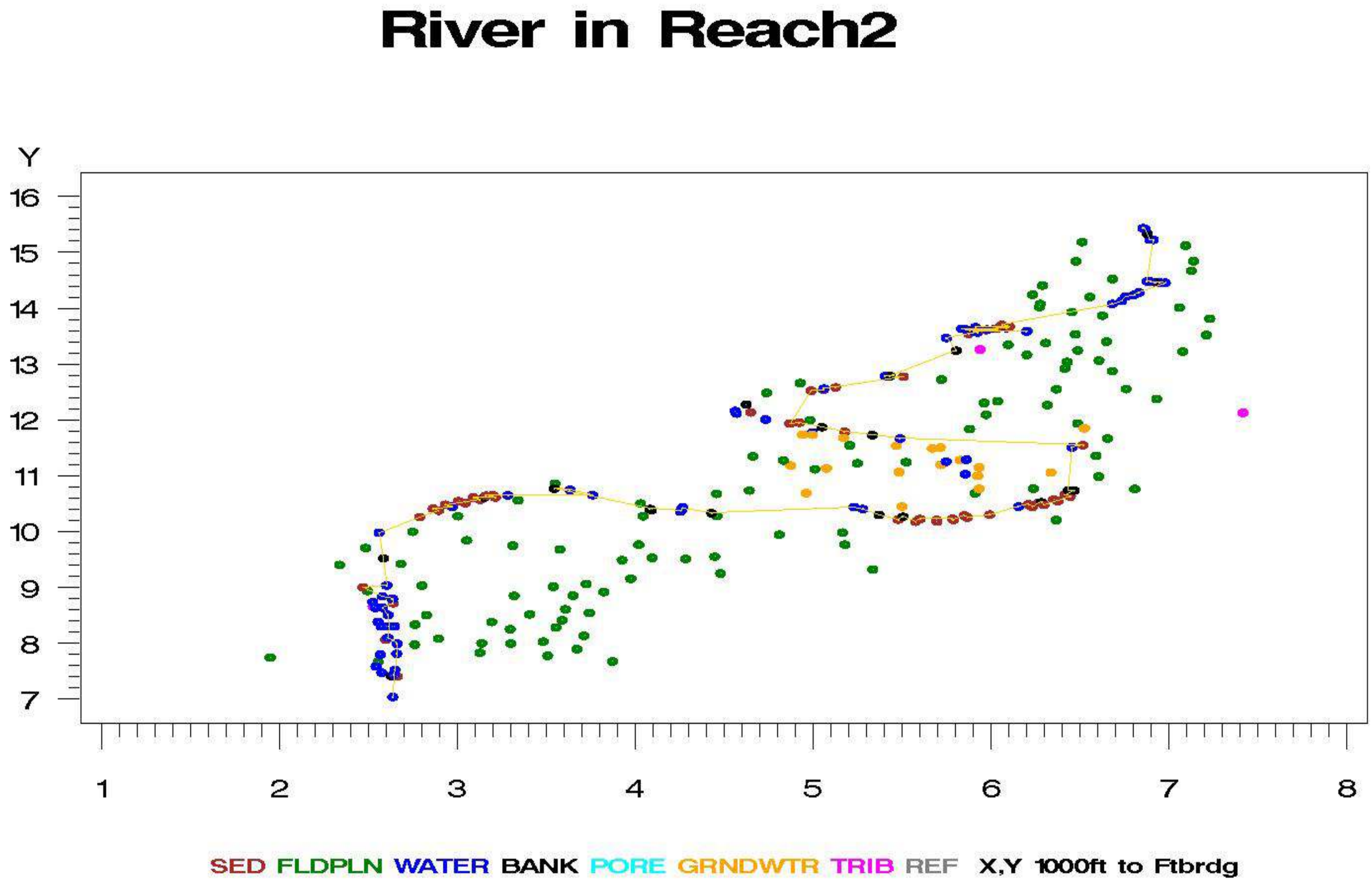
Sampling Points



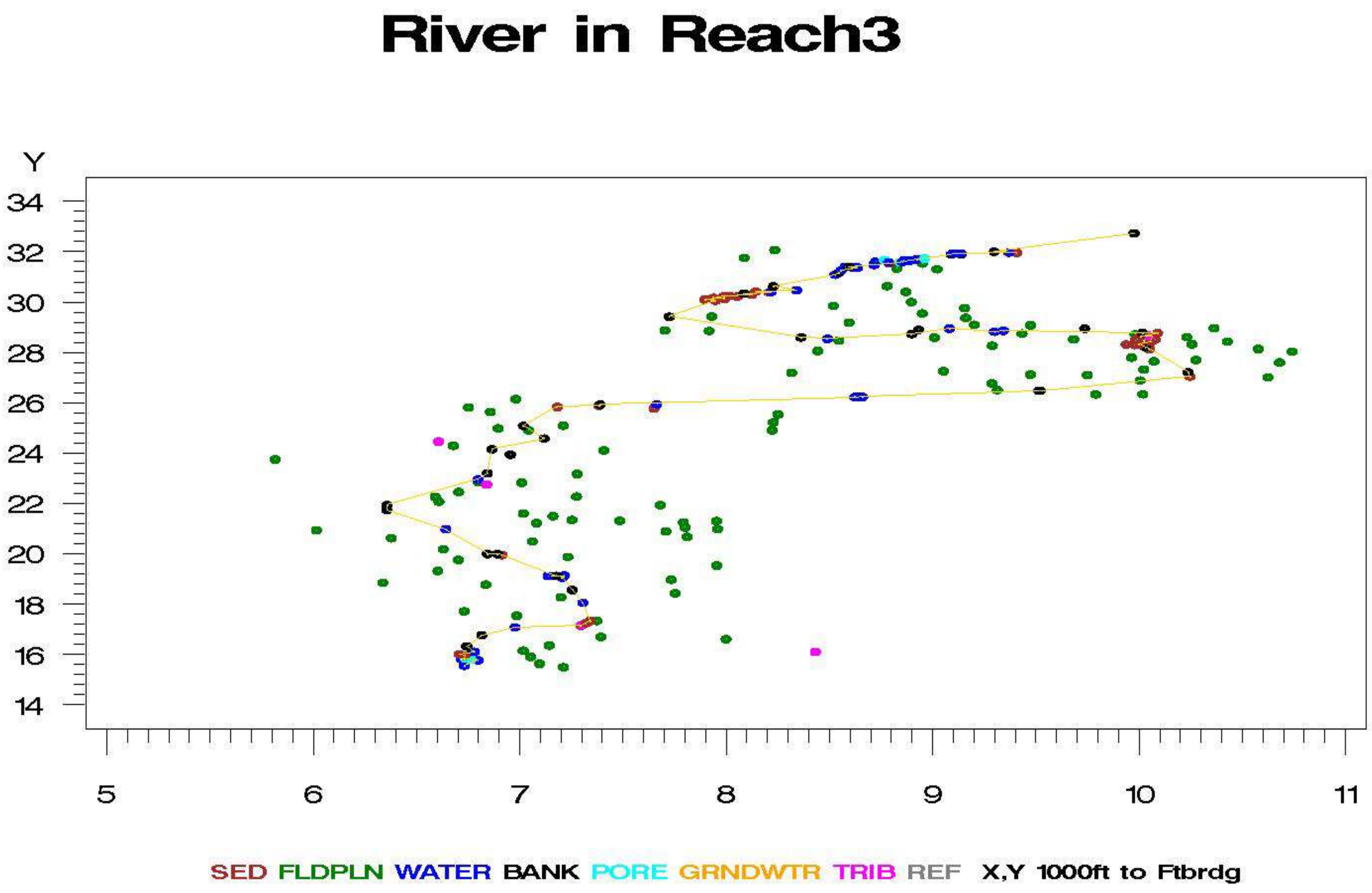
Sampling Points Reach 1



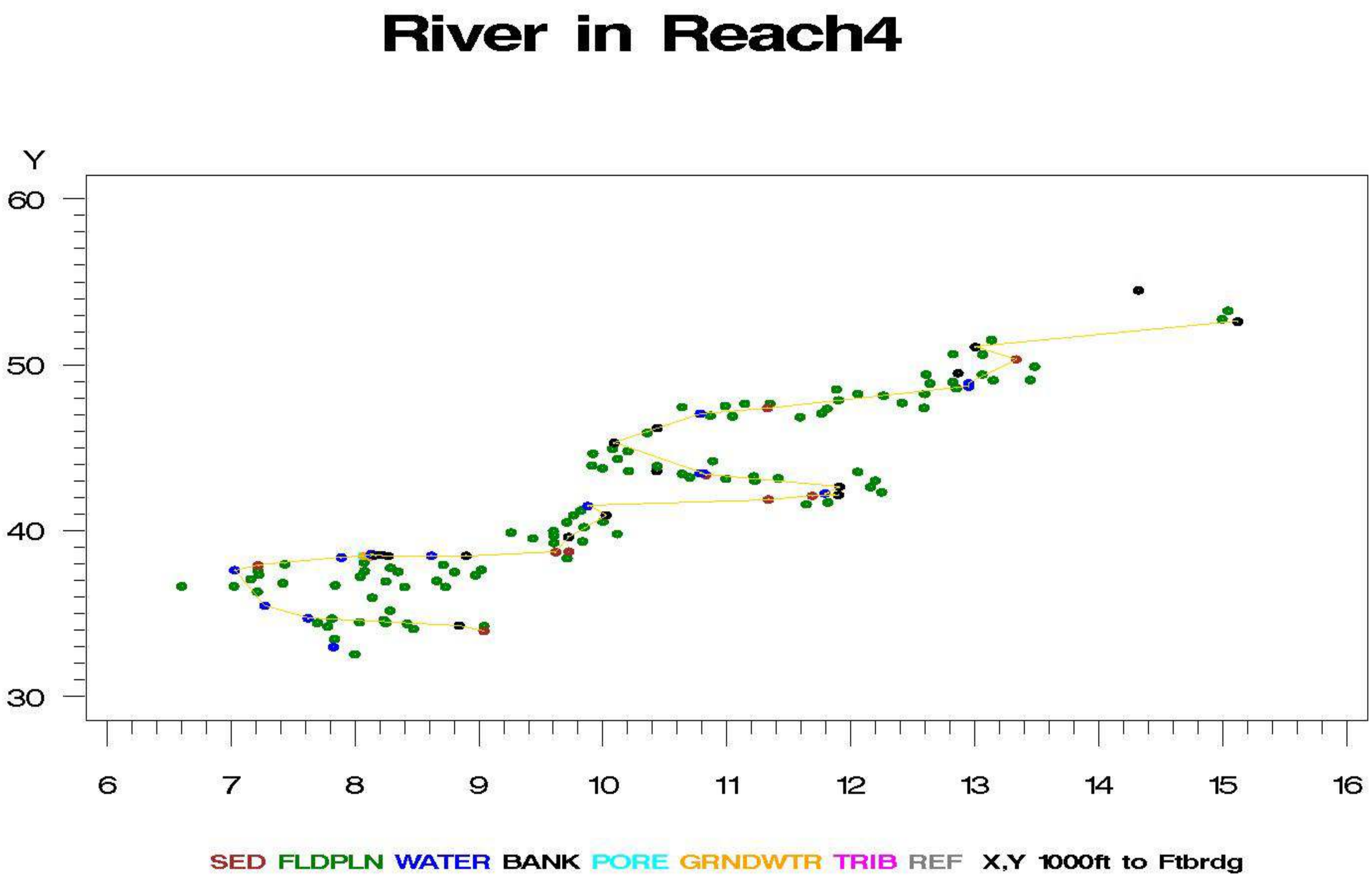
Sampling Points Reach 2



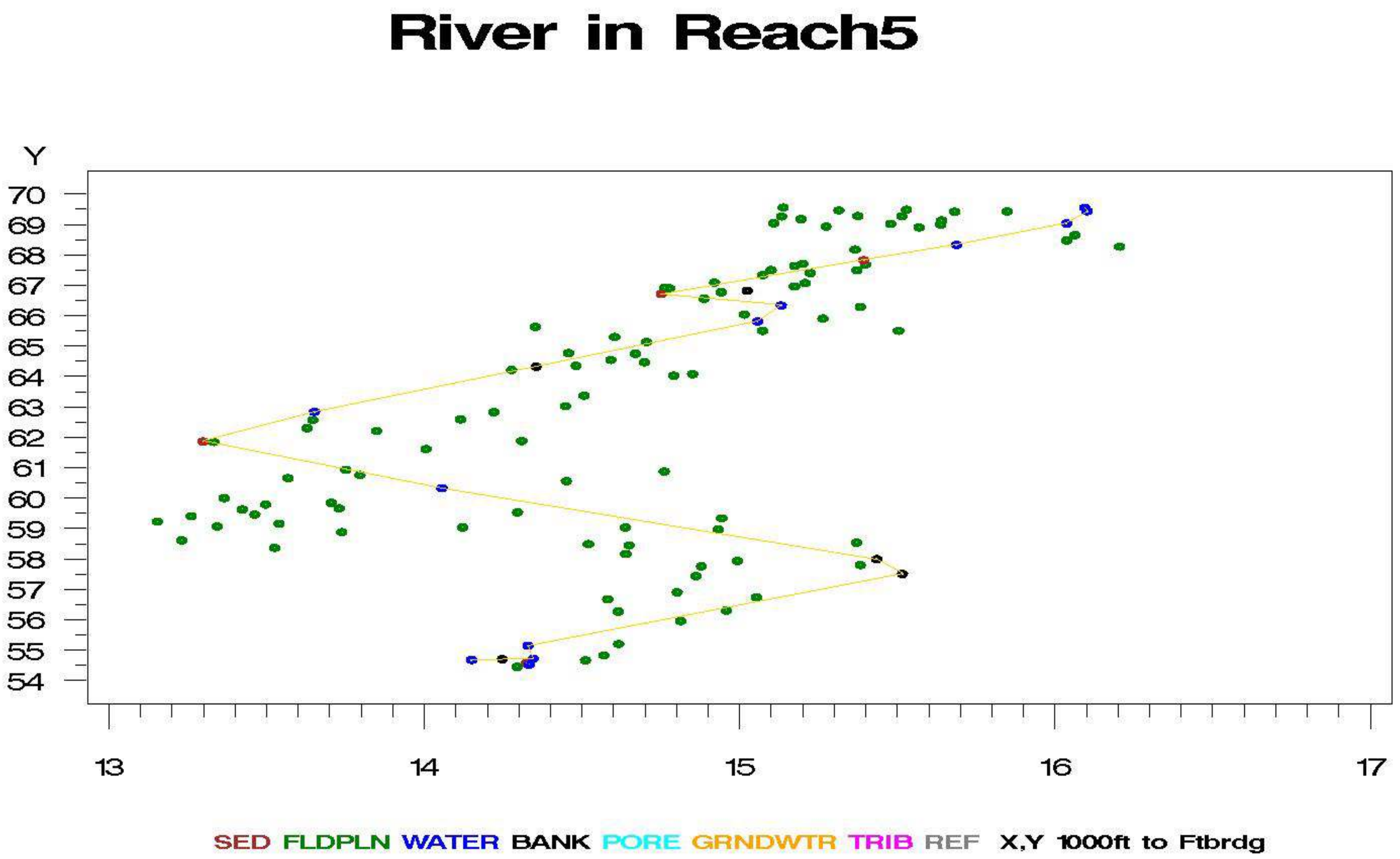
Sampling Points Reach 3



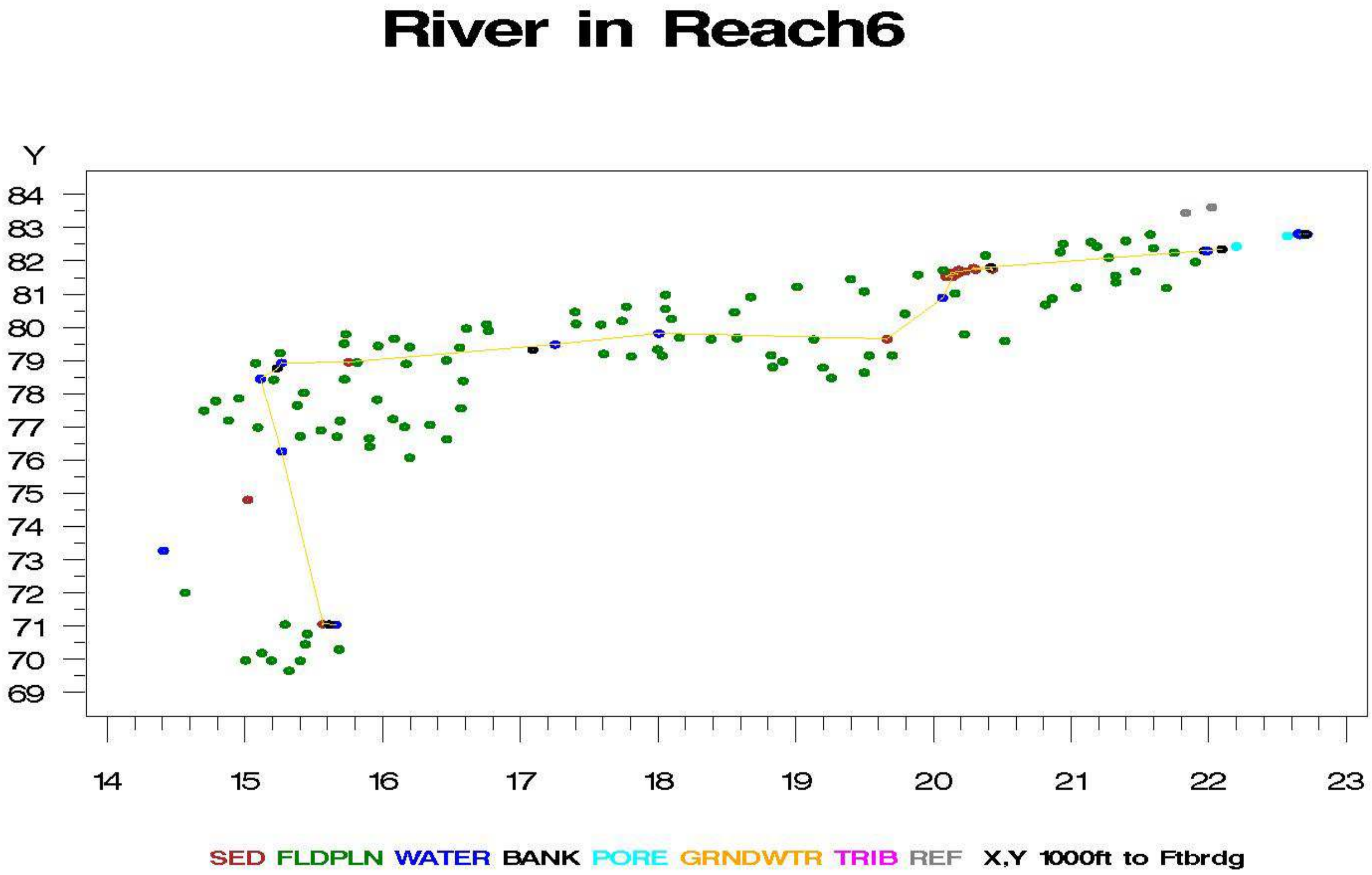
Sampling Points Reach 4



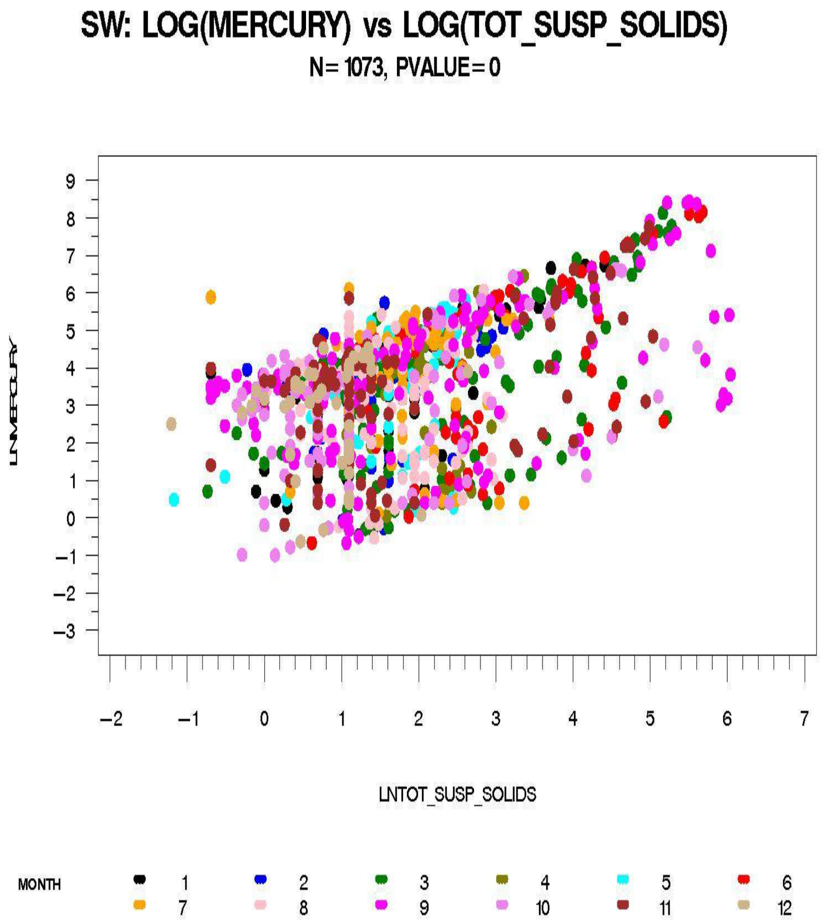
Sampling Points Reach 5

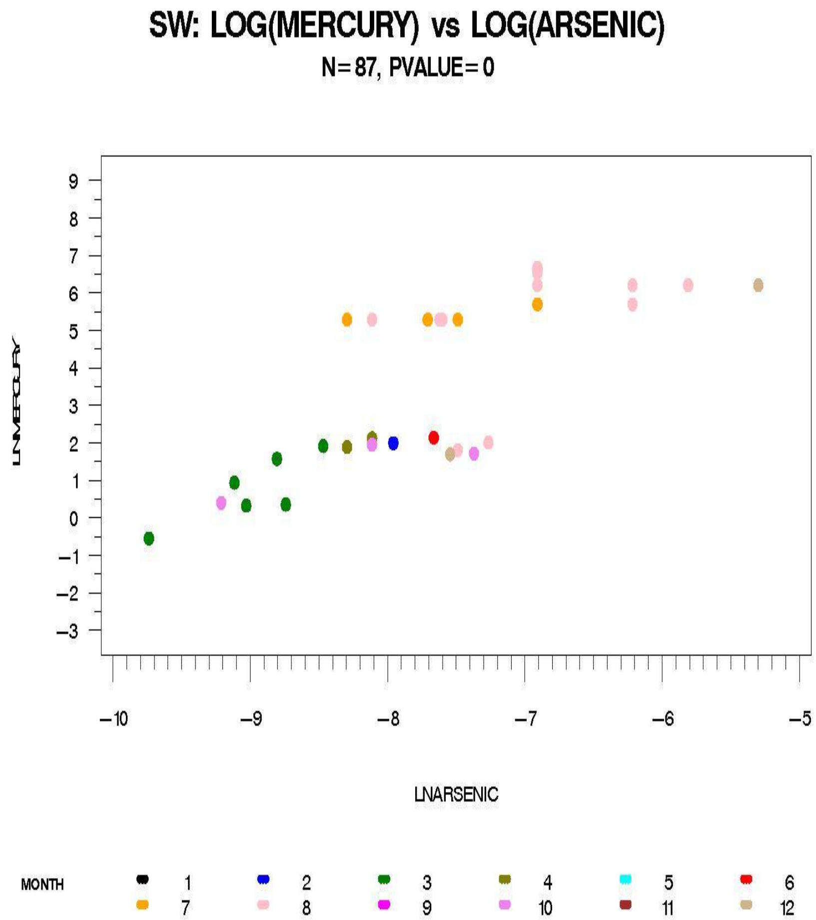


Sampling Points Reach 6



SW THg vs. Others

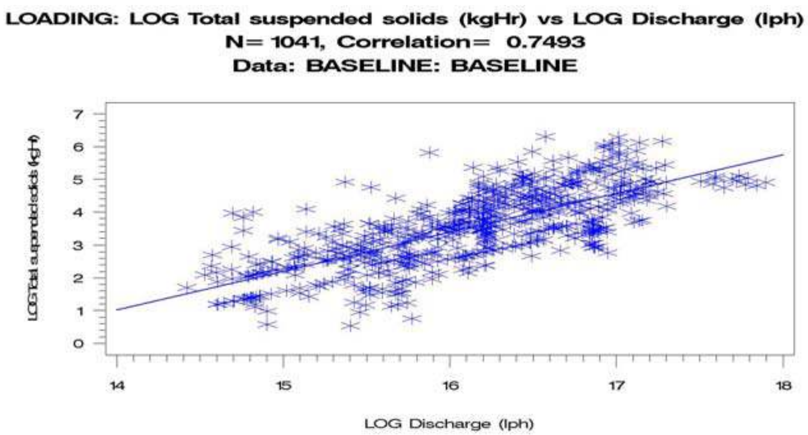


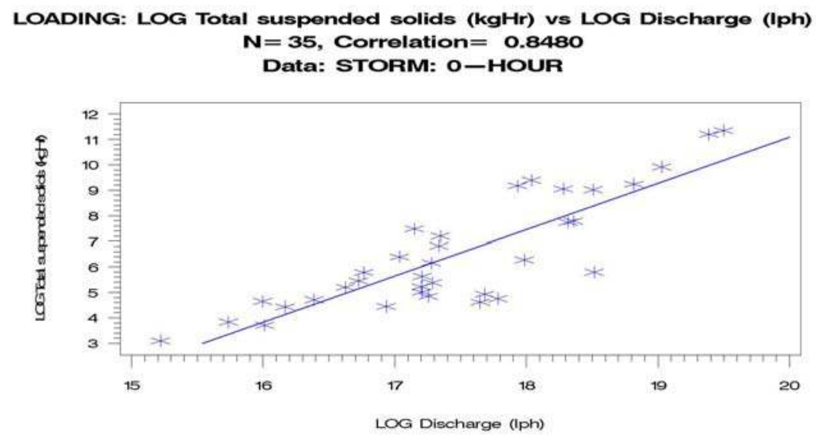


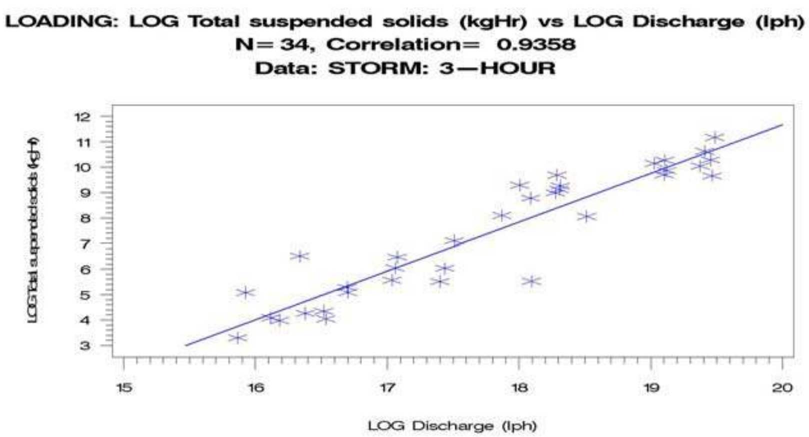
Total Suspended Solids

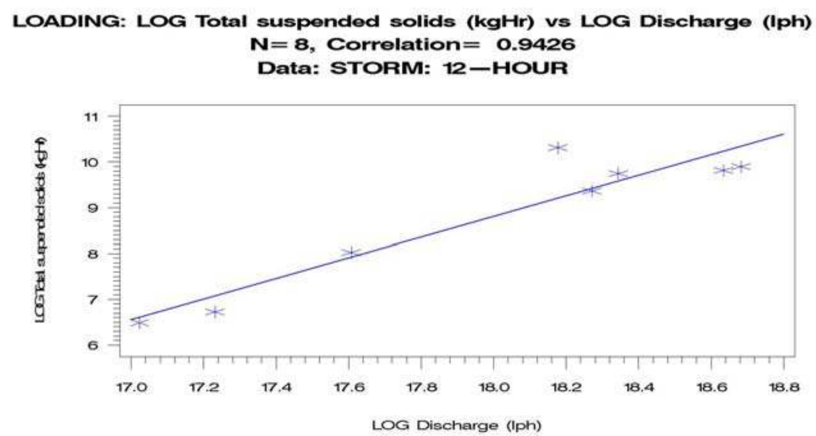
- Previous plot indicates a positive correlation of THg and TSS
- Expect a correlation of TSS and flow rate, especially storm related
- Next plot explores this idea

TSS vs. Storm Flow Rate

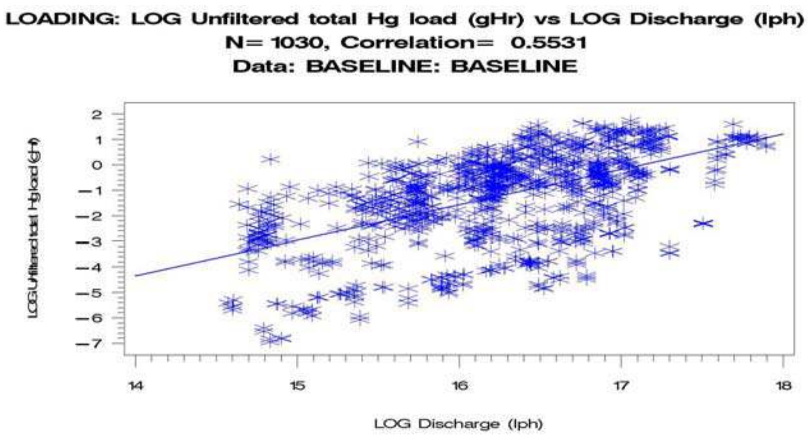


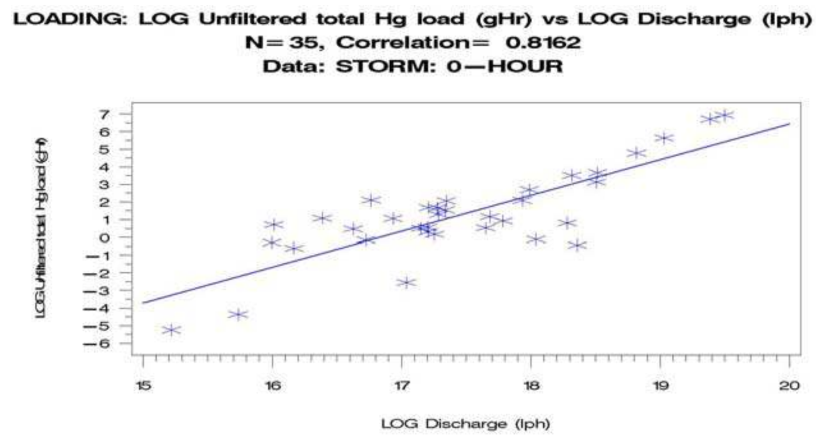


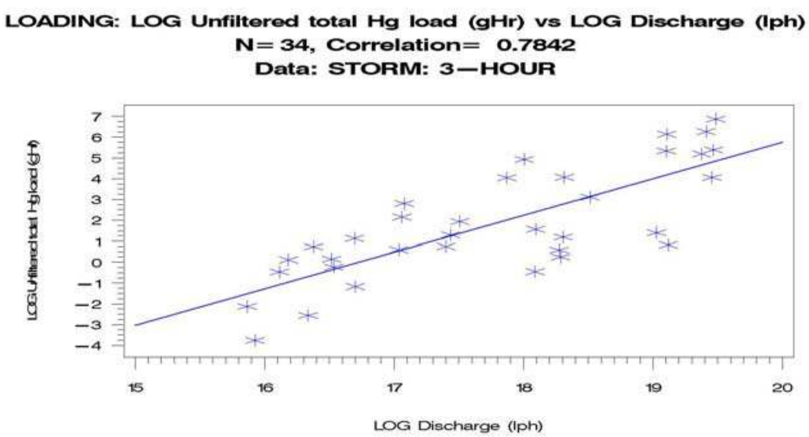


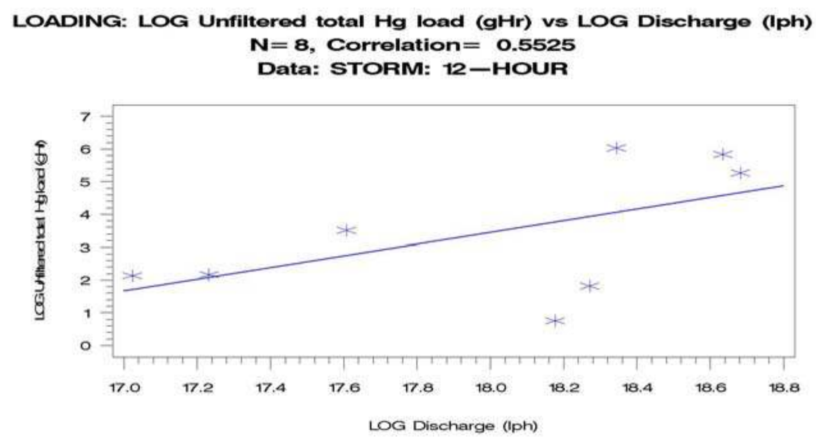


THg vs. Storm Flow rate

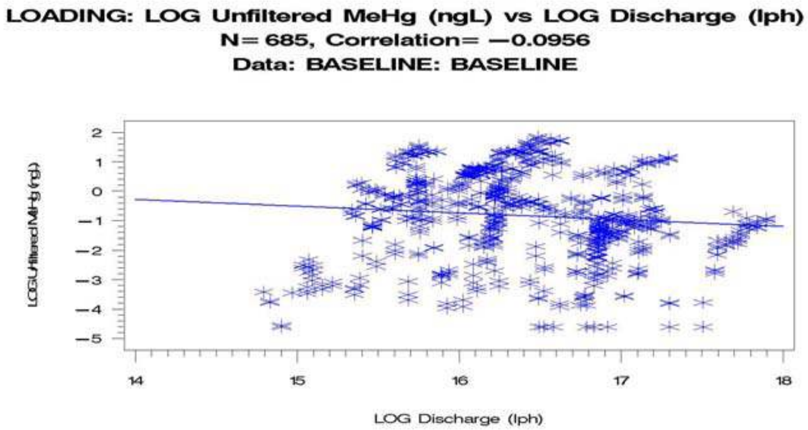


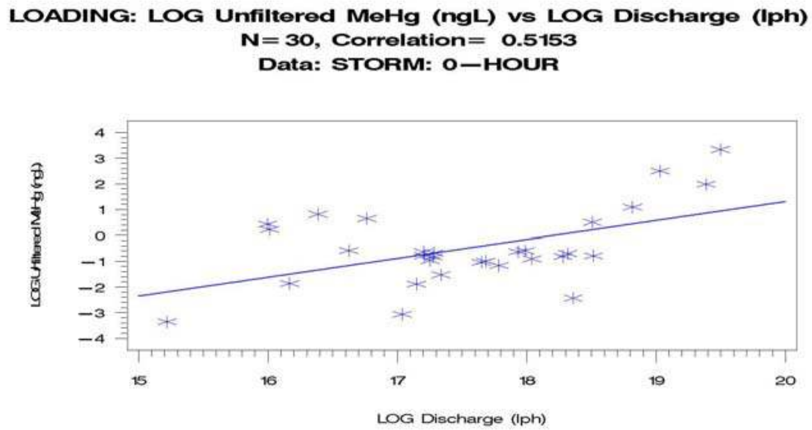


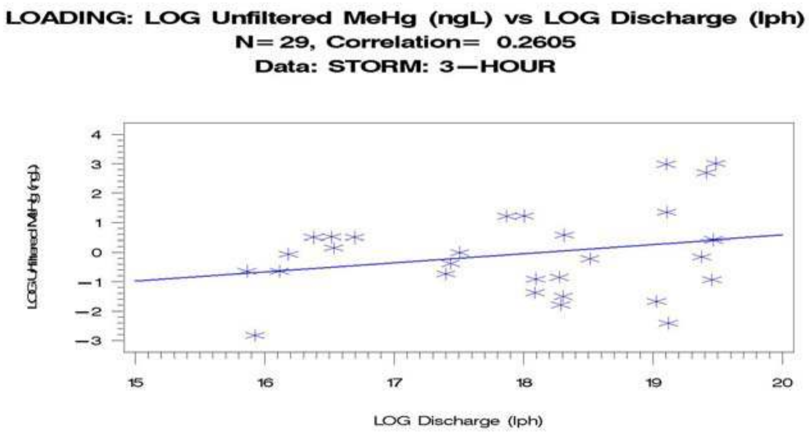


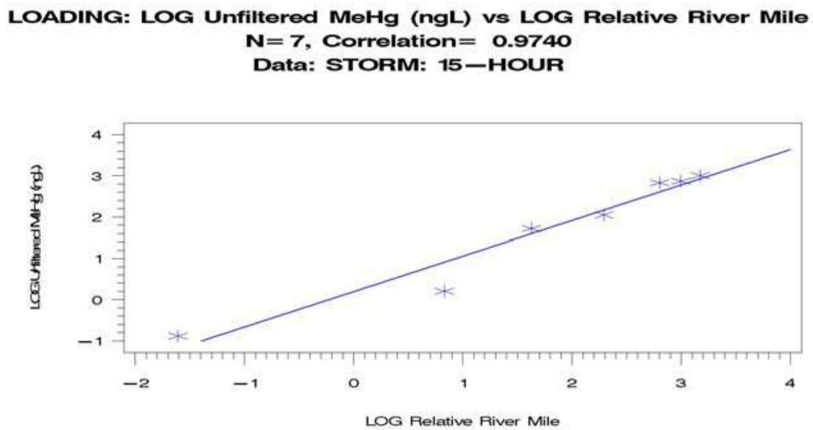


MeHg vs. FLOWRATE





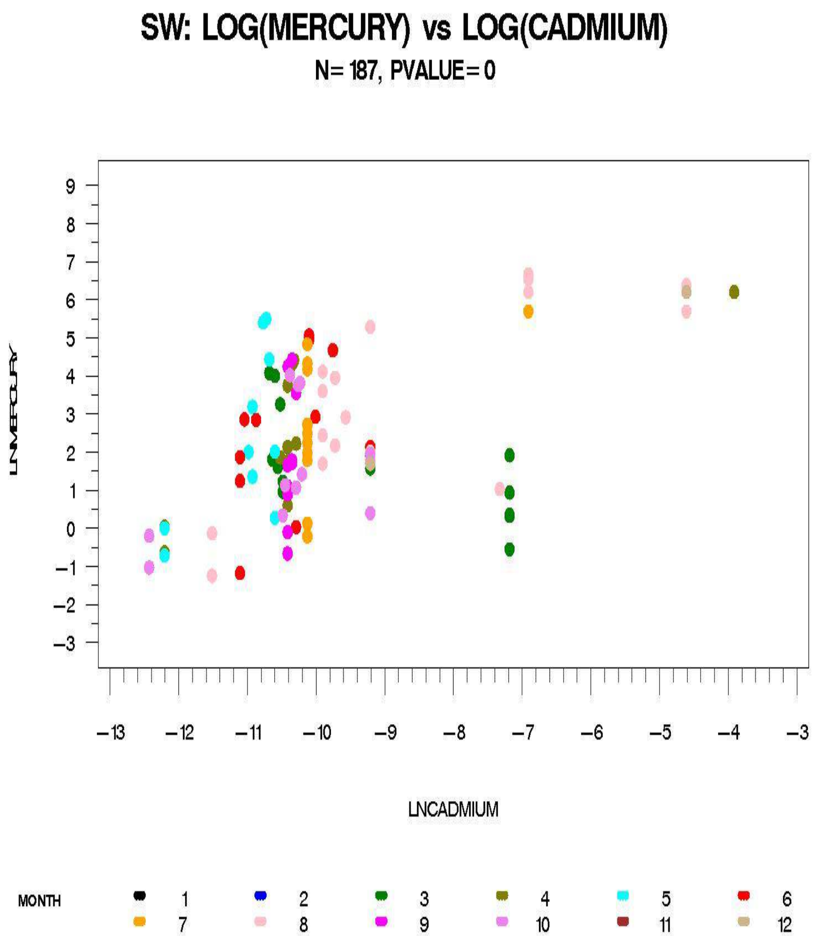


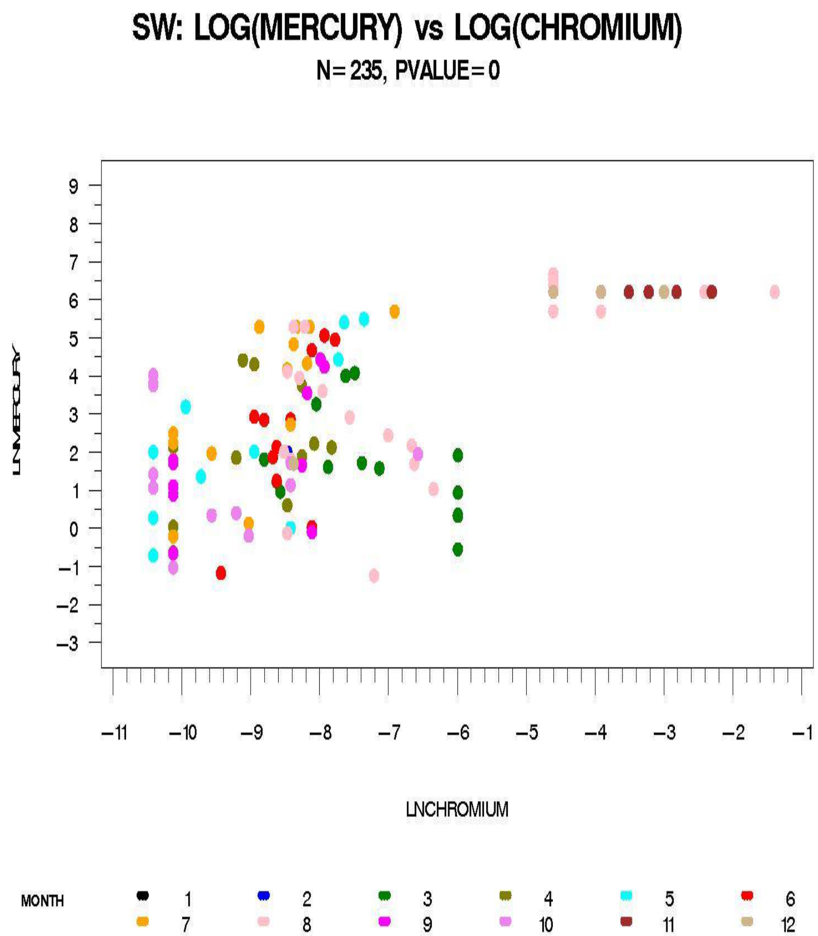


THg, MeHg vs. Flow rate

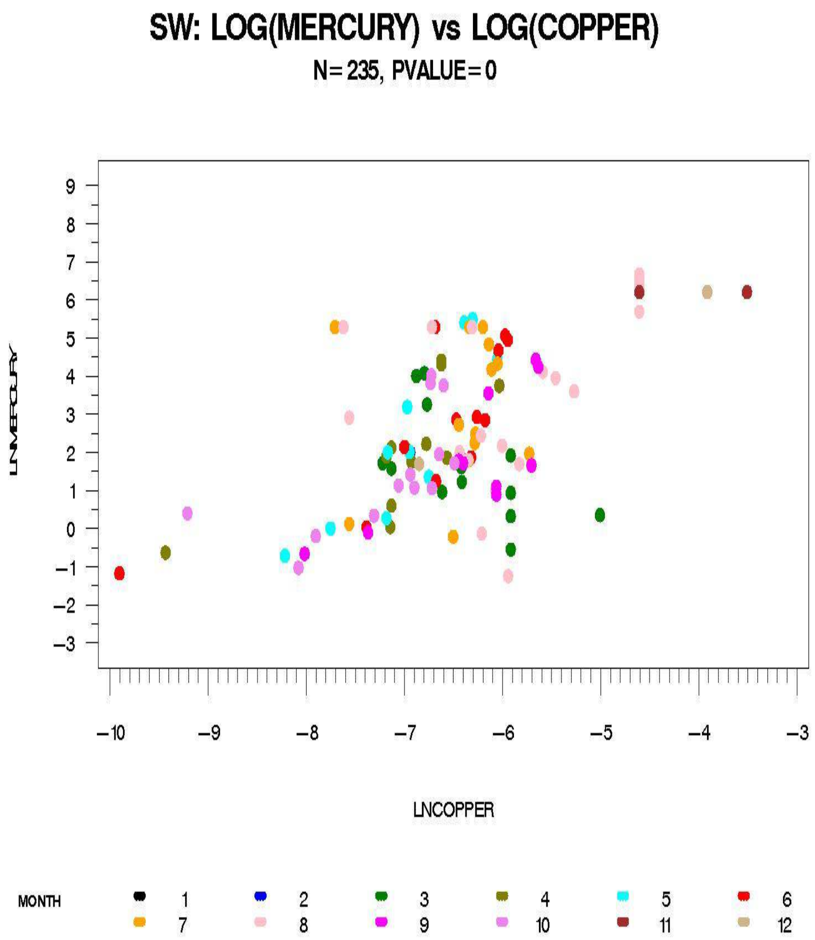
- Baseline negative correlation consistent with dilution of MeHg with higher flow volume.
- Storm plots suggests increased flow of MeHg into river during storms, but plots “noisy”
- THg positively correlated with flow rate at baseline and during storm day

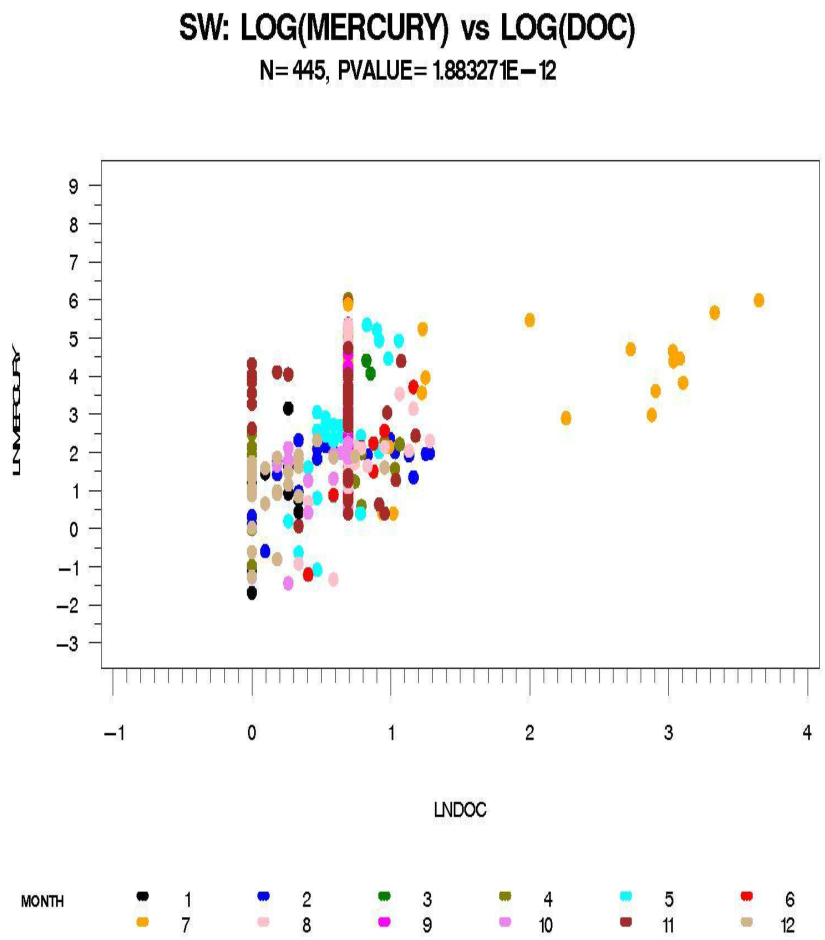
SW THg vs. Others



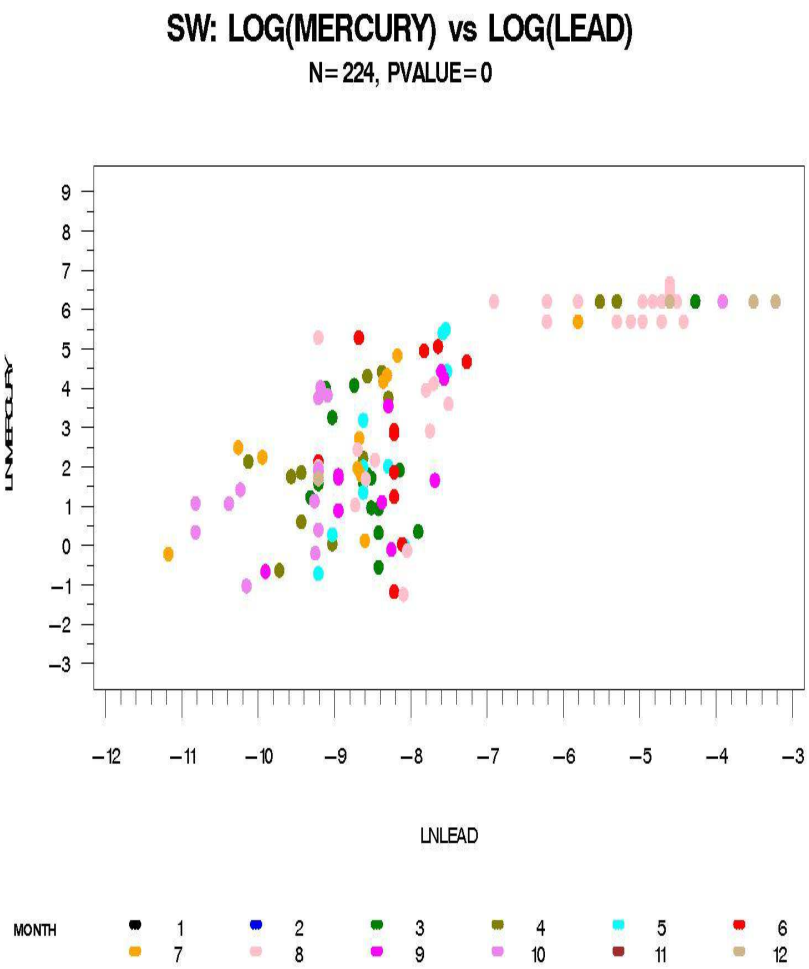


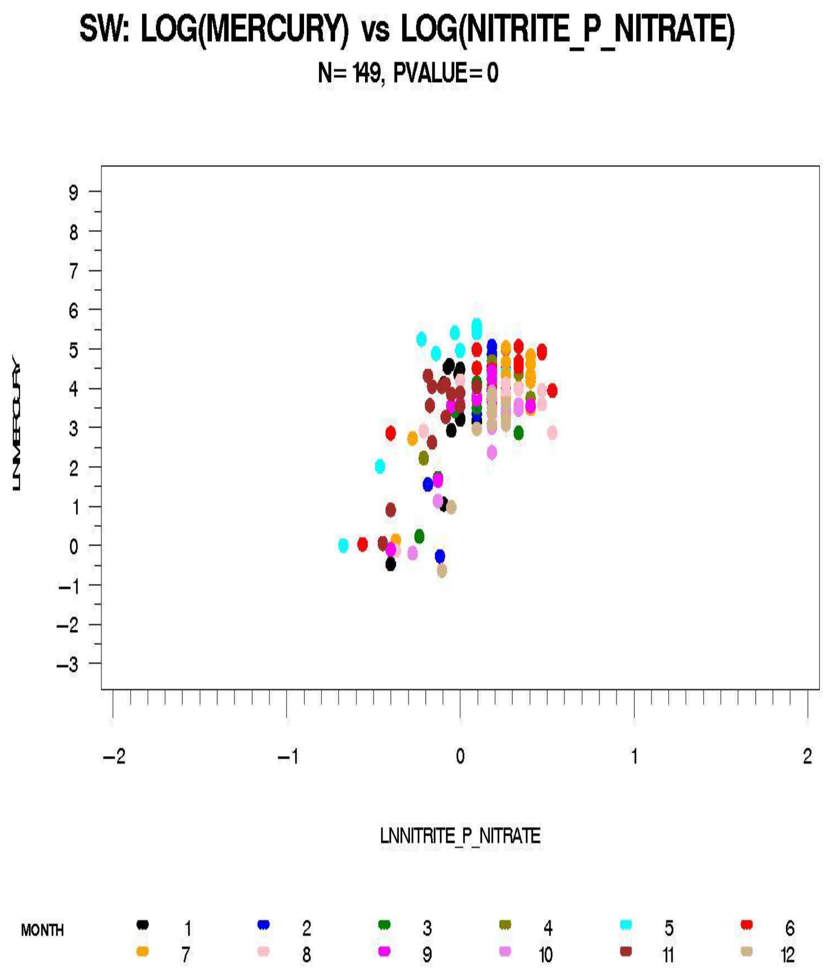
SW THg vs. Others



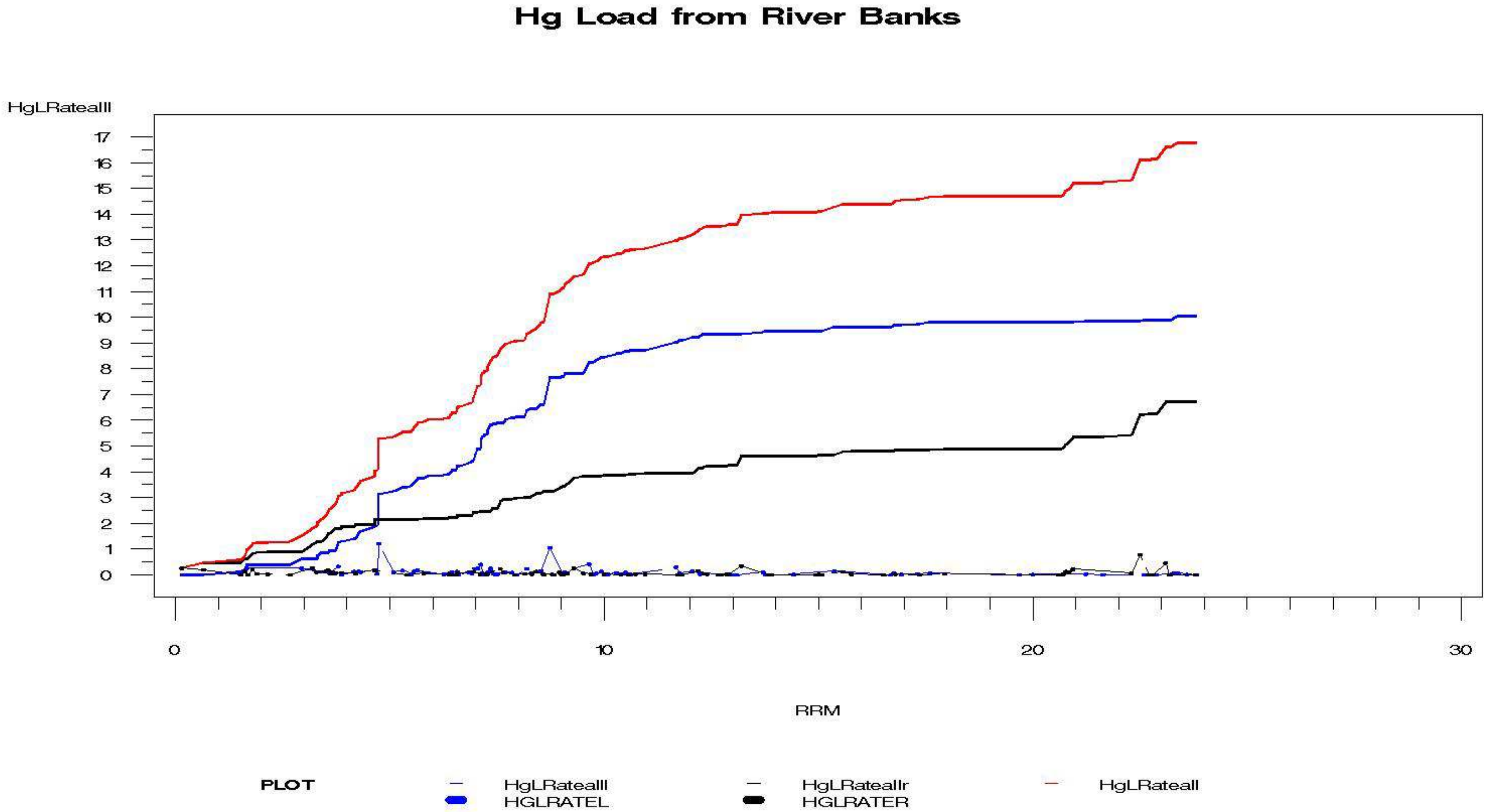


SW THg vs. Others





Annual THg Load From Bank



Mercury Entering South River Surface Water from Eroding Banks (ng/L)

Year	meanHg	mdHg	minHg	maxHg
2001	142.224	229.361	10.1260	432.658
2002	210.388	333.982	28.7134	543.913
2003	36.723	55.180	1.5230	218.816
2004	64.615	85.559	4.6774	260.780
2005	73.860	101.260	3.1106	292.876
2006	65.558	115.375	4.7005	307.048
2007	106.381	186.637	13.0390	328.223
2008	139.395	203.604	24.5638	432.658
2009	105.790	137.949	16.2709	209.197

Annualized rate is approximately 17 kg/Yr. Table based on Mean, median, minimum and maximum measured flow volumes.

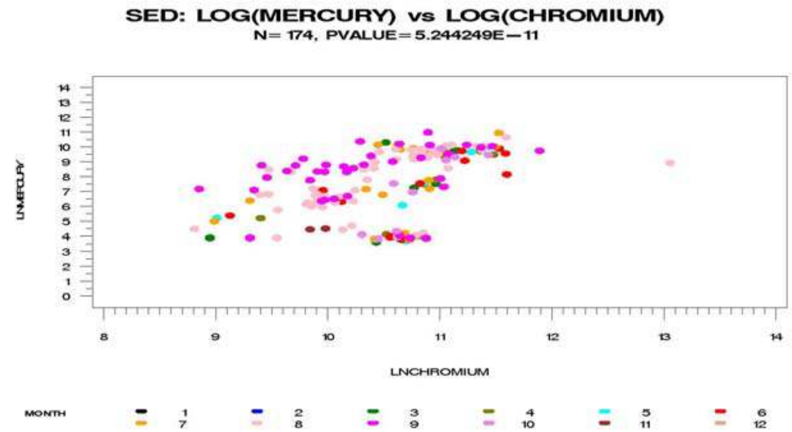
Table: Mercury Observed in South River (ng/L)

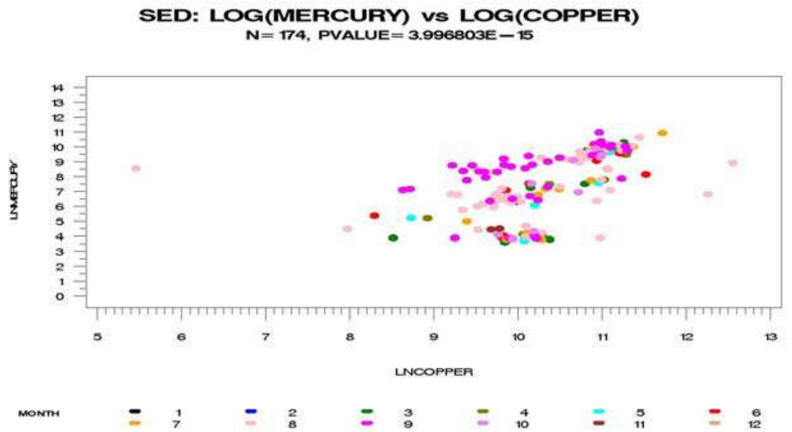
YEAR	mdHg	minHg	maxHg
2001	1.50000	1.500	38
2002	7.85000	1.500	449
2003	3.04000	1.500	415
2004	3.78460	0.690	2887
2005	5.43000	0.150	739
2006	7.92000	0.280	4656
2007	6.43000	0.100	3430
2008	8.73737	0.265	642

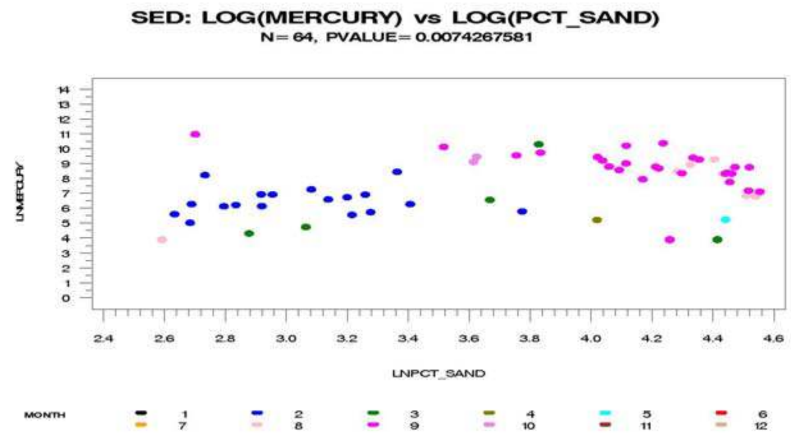
Table based on median, minimum, and maximum flow volumes and sample THg values. This works out to between 0.18 and 2 kg/Yr. Values from riverbank thus exceed measured SW THg levels, possibly due to deposition of THg from banks into sediment and floodplain.

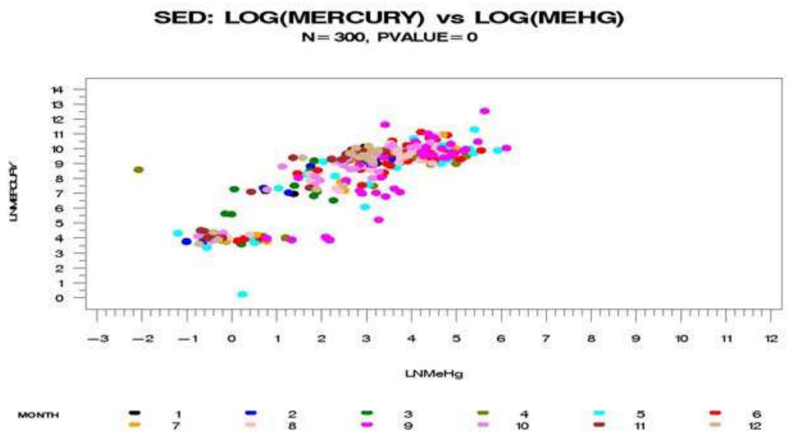
This, in turn suggests possibility of relationships between sediment chemicals and SW THg and MeHg. Next plots compare *sediment* THg to other chemicals

SED THg vs. Others

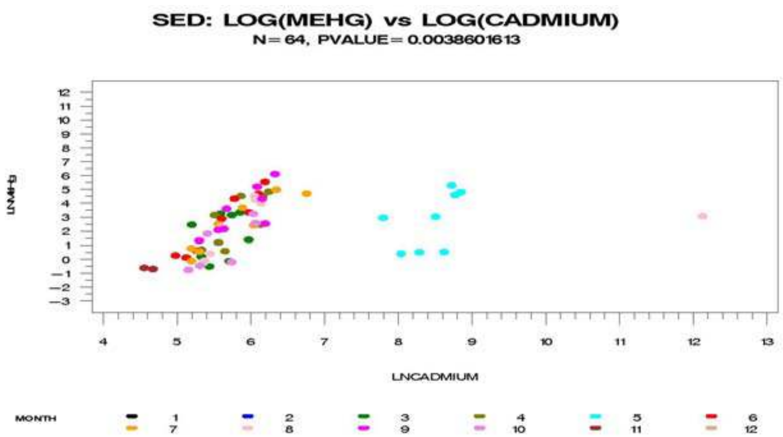


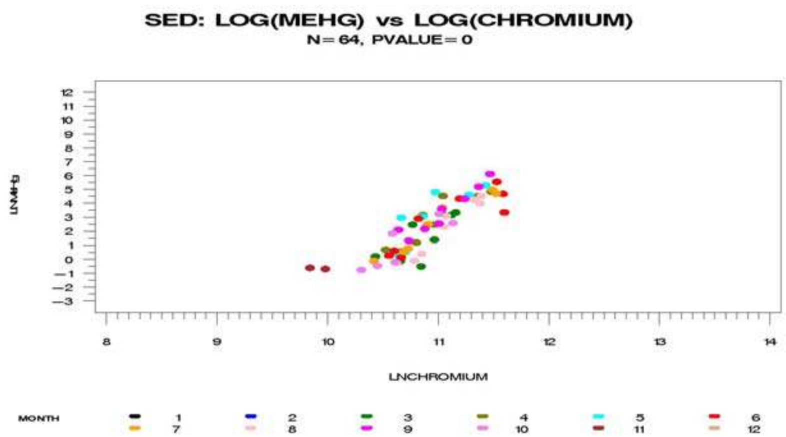


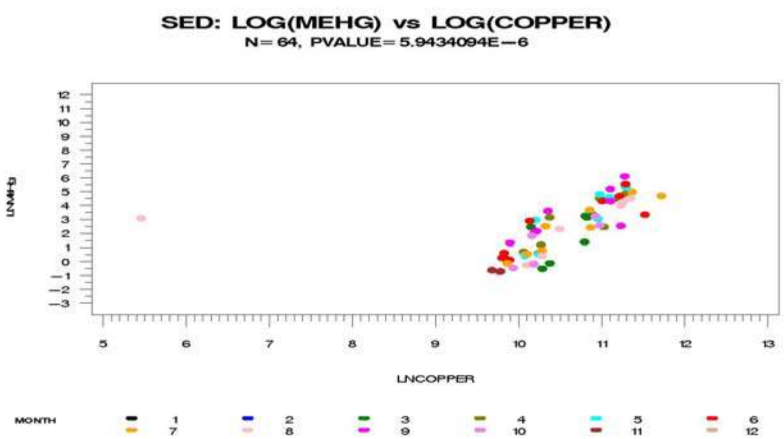


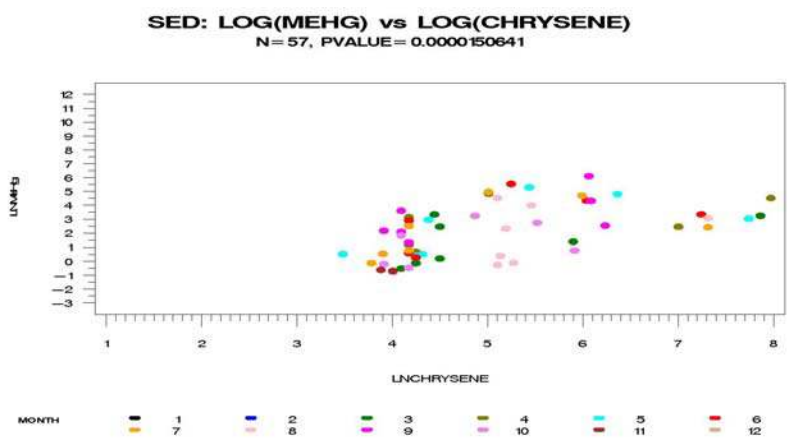


SED MeHg vs. Others









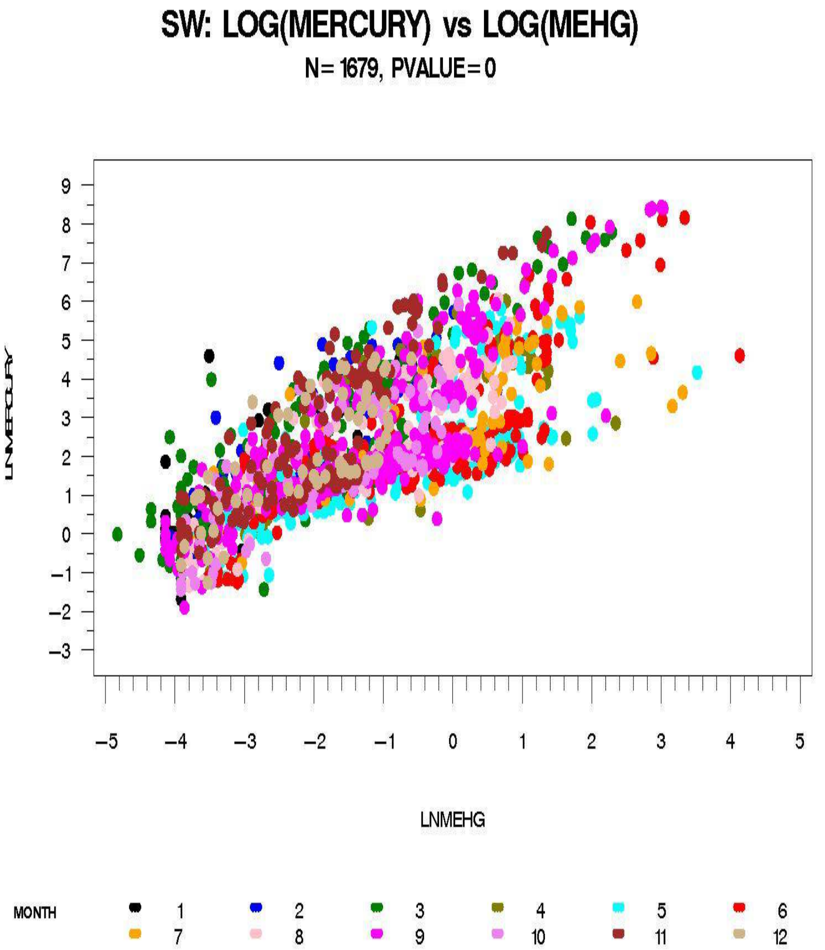
MeHg as Percent of THg

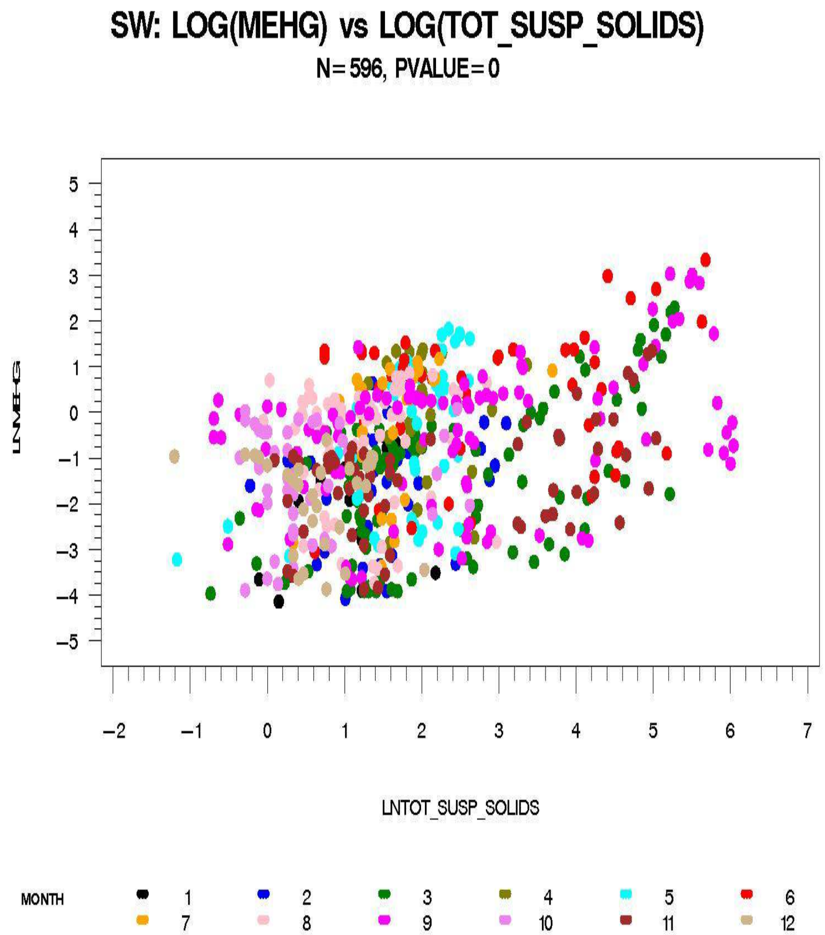
Year	Stream	%(MeHg/THg)	Q1	Q3	n	THg	MeHg
2004	SR	2.99	1.17	6.34	38	3.79	0.4975
2005	SR	5.54	2.09	11.37	303	5.43	0.2650
2006	SR	3.07	1.08	6.81	754	7.92	0.3925
2007	SR	2.81	0.93	7.03	417	6.43	0.3157
2008	SR	3.46	1.25	6.15	152	8.74	0.3477

Thus, MeHg is only a small fraction of THg in the South River. Similar calculations hold for sediment, floodplain, Tribs, reference streams.

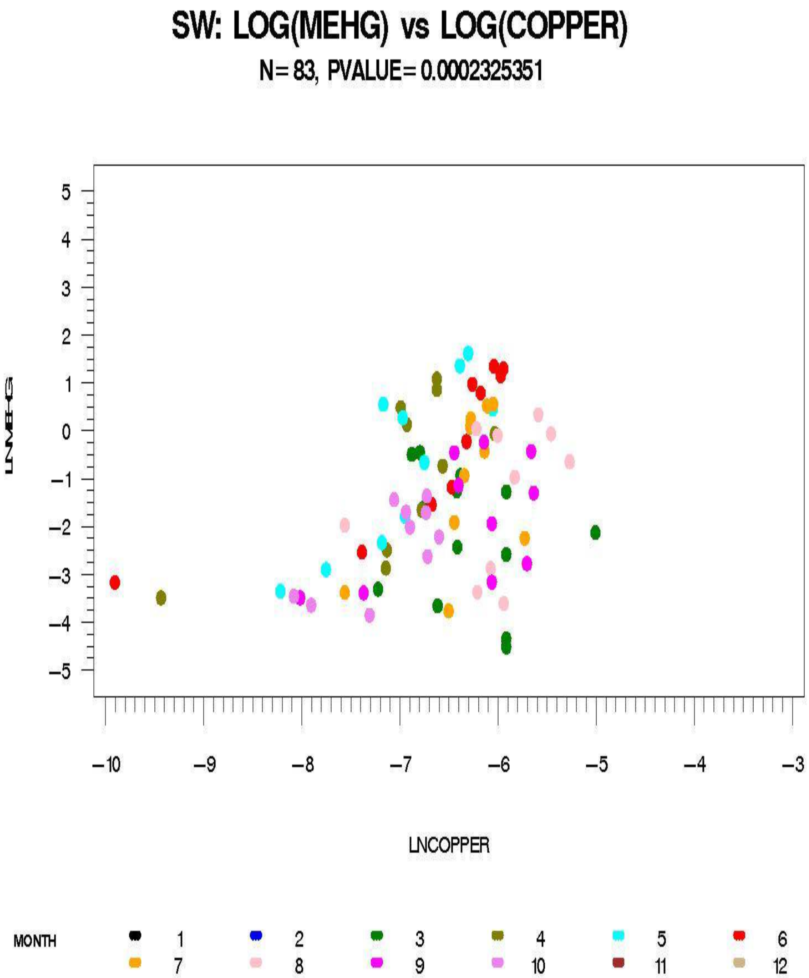
High correlation of THg and MeHg suggests similar models may apply

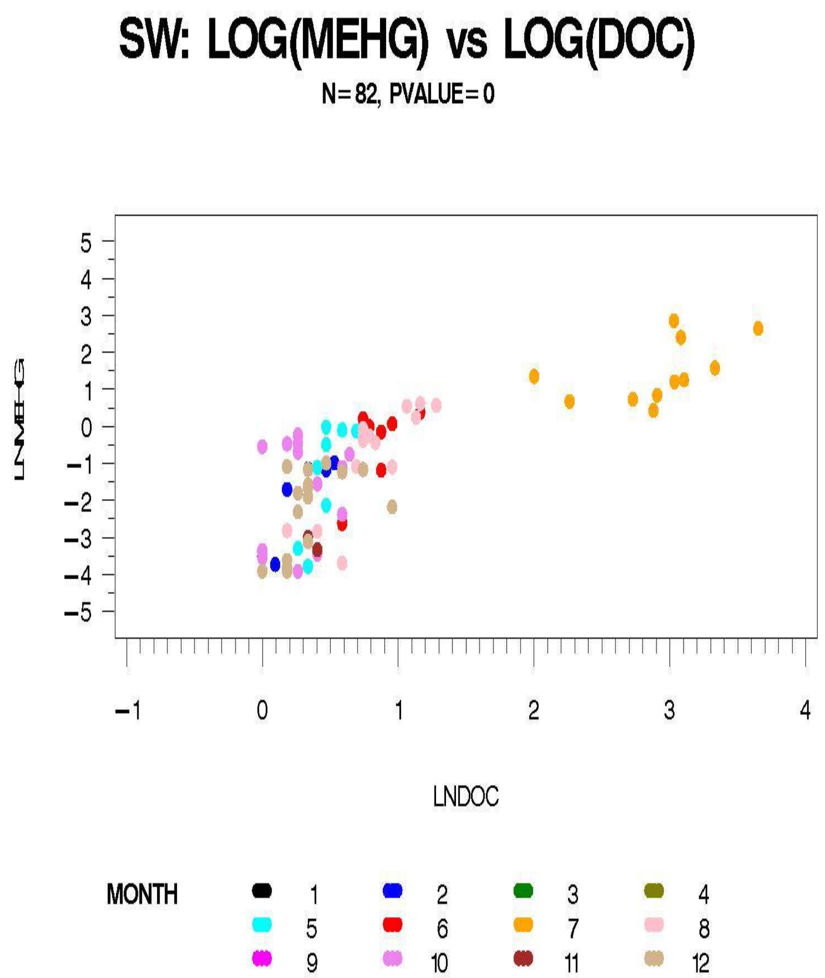
SW MeHg vs. Others





SW MeHg vs. Others

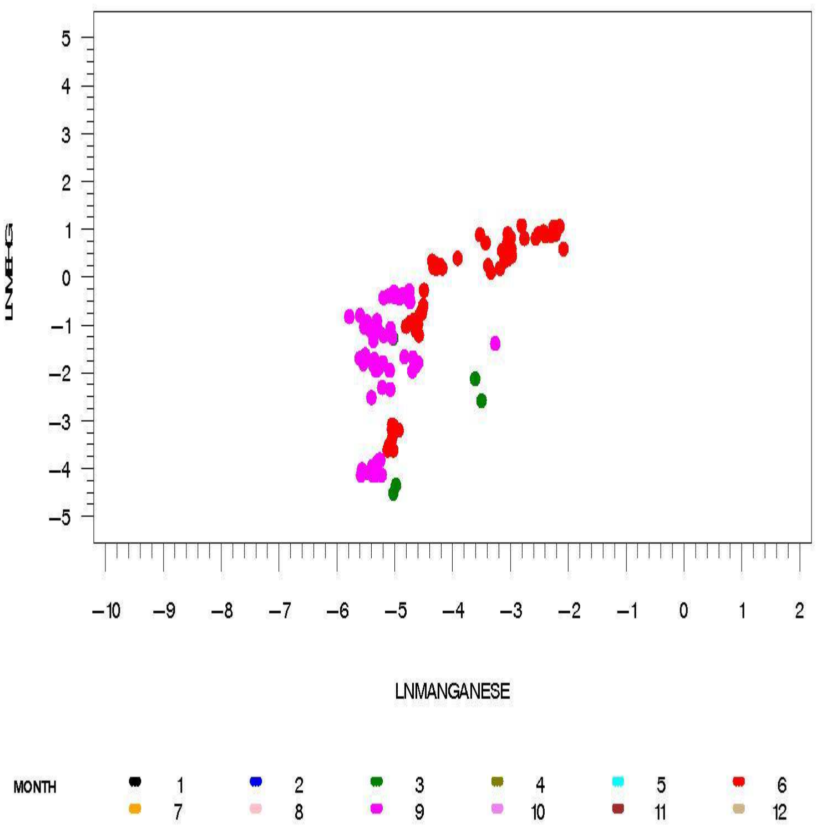




SW MeHg vs. Others

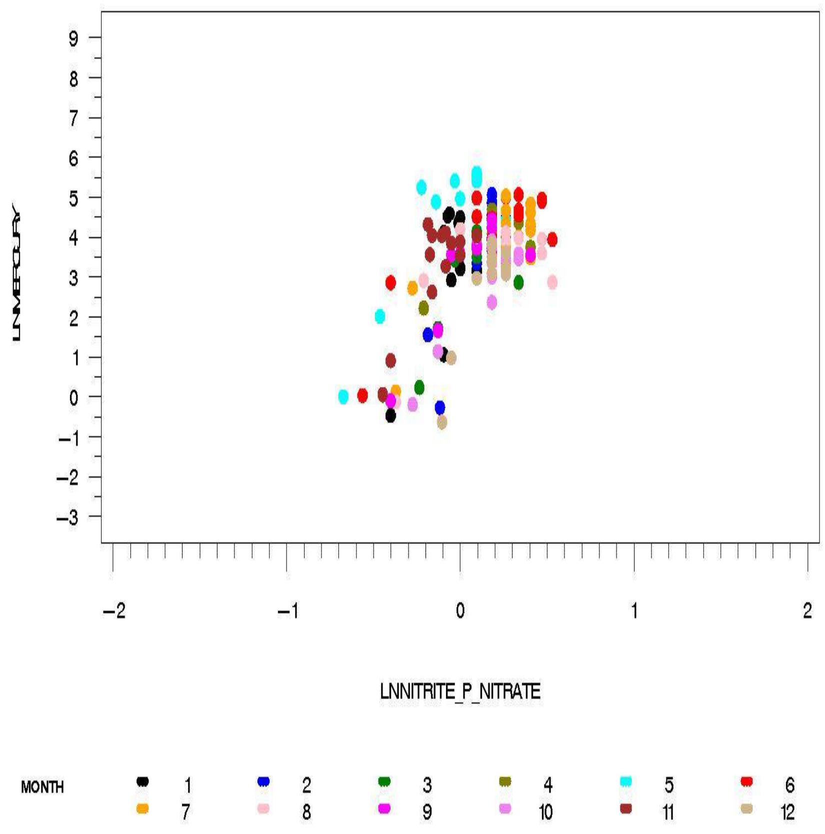
SW: LOG(MEHG) vs LOG(MANGANESE)

N= 115, PVALUE= 0

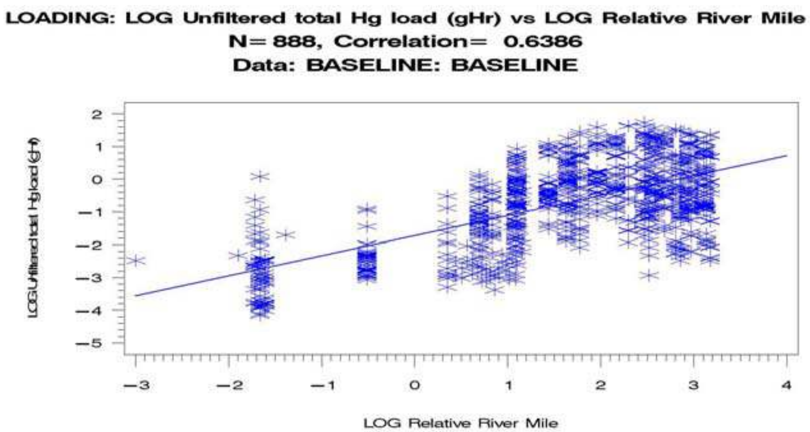


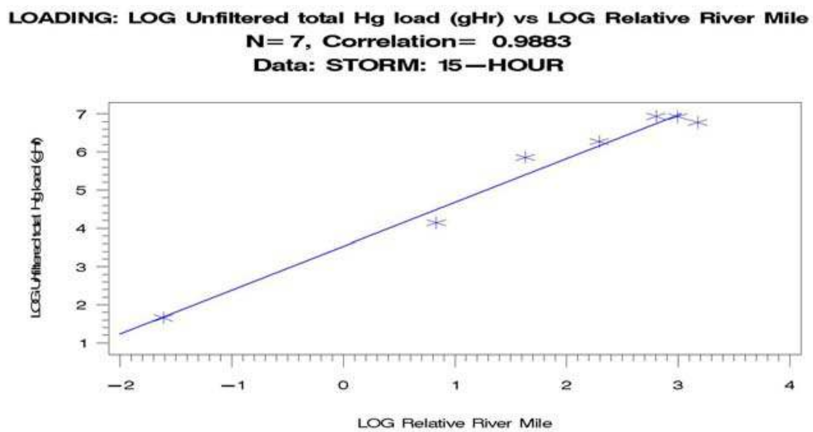
SW: LOG(MERCURY) vs LOG(NITRITE_P_NITRATE)

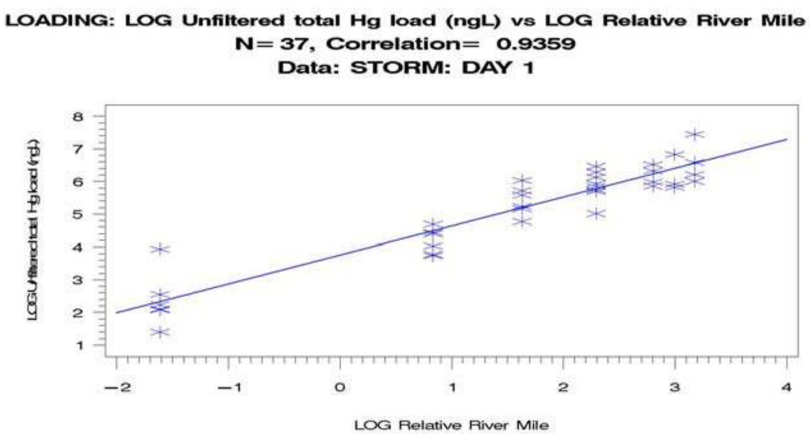
N= 149, PVALUE= 0



THg vs. RRM



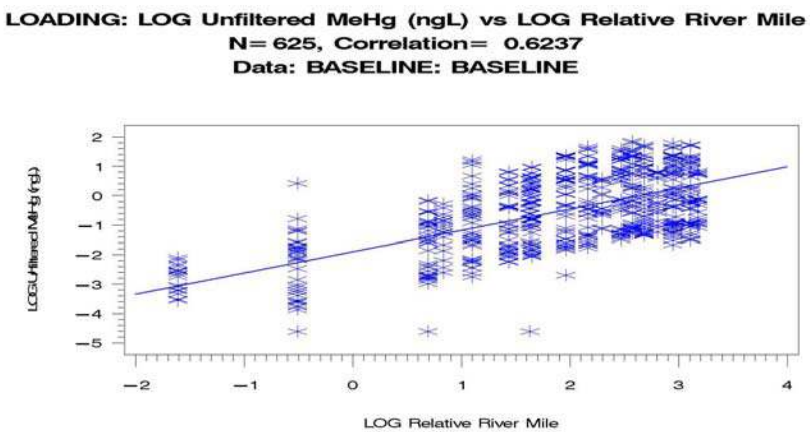


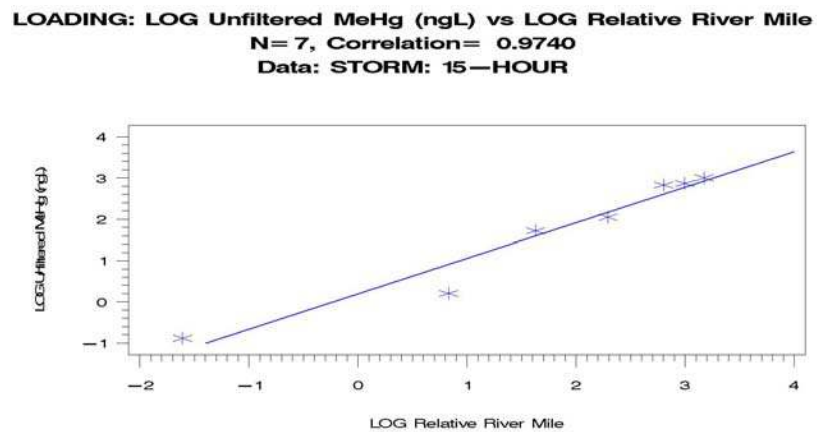


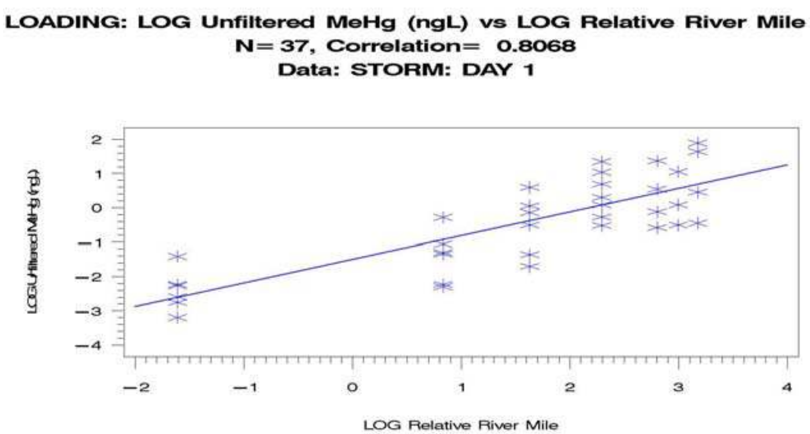
In baseline plot one can discern the "hump" between RRMs 4.5 and 12 (Ln(RRM)=1.5 to 2.5).

Storm plots may suggest greater flow of THg downriver during storms.

MeHg vs. RRM







No “hump” observed in the baseline plot between RRMs 4.5 and 12 (Ln(RRM)=1.5 to 2.5).

Storm plots may suggest greater flow of MeHg downriver during storms.

Principal Component Analysis

- PCA can identify relationships among a large number of chemicals that simple correlation analysis misses.
- PCA was done separately on
  - All SW chemicals and water quality criteria
  - All sediment chemicals
- This potentially provides useful associations to explain THg or MeHg
  - In fact, little useful info found in this way

SW PCA, Excl Ref Samples

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8
ALKALINITY	75 *	-30 *	15	-18	-20	7	5	-3
AMMONIA	14	23	29 *	17	28 *	-42 *	33 *	-15
BOD5DAY	-10	9	39 *	46 *	1	-10	19	-2
CALCIUM	39 *	-10	-65 *	39 *	10	-38 *	-18	5
CARBON_ORG	23	20	13	29 *	33 *	36 *	-23	16
CHLORIDE	48 *	-29 *	42 *	19	-41 *	-14	5	-7
COD	17	16	6	54 *	20	57 *	-13	4
DOC	33 *	-8	-38 *	26	-11	-13	7	12
F_COLIFORM	8	11	-12	-6	8	19	19	-9
HARDNESS	77 *	-12	-18	-19	5	7	1	-4
MAGNESIUM	39 *	-9	-65 *	38 *	10	-38 *	-18	4
MEHG	*12*	*15*	*-29 **	*-2*	*-19*	*8*	*46 **	*52 **
MERCURY	*8*	*11*	*-1*	*-9*	*-10*	*1*	*26*	*76 **
NITRATE	50 *	-7	12	-17	44 *	9	-21	18
NITRITE	11	27	7	7	45 *	-11	40 *	-7
NITRITE_P_NITRATE	40 *	-20	30 *	-10	25	3	-35 *	18
NITROGEN_KJELD	36 *	54 *	17	3	29 *	-32 *	20	-11
OPHOSPHATE	62 *	4	16	-13	47 *	8	8	5
ORTHOPHOSPHATE	46 *	24	-38 *	-3	-26	12	15	-29 *
PHOSPHORUS	51 *	61 *	17	-9	2	-11	-15	-5
POTASSIUM	72 *	-36 *	20	4	-8	-4	4	6
RESIDUE_VOLNONFIL	18	78 *	6	-7	-36 *	-9	-32 *	-1
SILICA	47 *	13	-39 *	-38 *	-4	36 *	22	-24
SODIUM	46 *	-32 *	41 *	41 *	-34 *	-7	-5	1
SPECIFIC_CONDUC	48 *	-12	-18	-13	-4	21	-1	-11
SULFATE	37 *	-9	8	1	-15	-2	13	-4
TOTFIX_SOLIDS_NONFIL	-6	30 *	-15	65 *	-10	46 *	9	-9
TOT_SUSP_SOLIDS	23	74 *	10	-14	-40 *	-9	-26	17
TURBIDITY	20	3	28	14	-27	12	30 *	-7

THg appears uncorrelated with other chemicals and water quality criteria in PCA.
MeHg mildly associated with low magnesium and calcium and high sodium

SED PCA, Excl Ref Samples

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8
ACENAPHTHENE	80 *	16	28	35 *	-5	11	4	25
ACENAPHTHYLENE	80 *	16	28	35 *	-5	11	4	25
ACETOPHENONE	70 *	28	22	14	-23	-7	22	-12
ALUMINUM	-6	-41 *	25	56 *	39 *	-8	-5	8
ANILINE	52 *	29	30	25	-30	2	25	-19
ANTHRACENE	93 *	8	14	17	-2	9	-1	18
ANTIMONY	-7	4	55 *	-27	13	-15	12	32 *
ARSENIC	-7	-1	60 *	-54 *	9	7	0	-13
BENZO_A	93 *	-6	-14	-17	6	1	-8	11
BENZO_APYRENE	95 *	-3	-15	-20	5	-5	-4	6
BENZO_B	95 *	-4	-16	-20	4	-4	-4	4
BENZO_GHI	96 *	-1	-14	-19	4	-9	-1	4
BENZO_K	96 *	-1	-13	-17	4	-7	-2	8
BERYLLIUM	1	-1	-68 *	48 *	-32	-3	21	18
CADMIUM	37 *	-1	-4	-7	-3	-28	26	-15
CHROMIUM	17	-41 *	17	16	-17	72 *	-8	-13
CHRYSENE	95 *	-4	-15	-19	5	-3	-5	6
COPPER	47 *	-21	-2	-1	-17	68 *	-19	-20
DIBENZ_AHA	87 *	14	-7	-18	-8	-22	14	-14
DRY_FRACTION	-10	-46 *	-32 *	-42 *	-12	34 *	8	30
FLUORANTHENE	95 *	-2	-16	-22	3	-10	0	1
FLUORENE	81 *	16	27	34 *	-5	10	5	24
INDENO_PYRENE	96 *	0	-13	-17	5	-11	-1	7
IRON	46 *	-64 *	18	24	21	0	21	-13
LEAD	66 *	23	2	6	13	20	-31	-21
LOI	-7	34 *	23	17	18	-48 *	1	0
MANGANESE	-34 *	-66 *	-10	10	47 *	-15	-4	24
MERCURY	***-5***	***39 ****	***-8***	***3***	***0***	***9***	***-36 ****	***5***
MeHg	***-4***	***33 ****	***-8***	***-1***	***2***	***-1***	***-28***	***10***
NAPHTHALENE	80 *	16	28	35 *	-5	11	4	25
NICKEL	-8	-72 *	-5	19	49 *	-25	8	7
Org_Cmplx_Mercury	-21	-33 *	-33 *	-13	25	31	-35 *	33 *
PCT_CLAY	-12	76 *	-16	2	9	-5	-40 *	8
PCT_H20	-3	39 *	-18	-2	60 *	39 *	52 *	-5
PCT_SAND	-53 *	-16	27	-7	-35 *	20	33 *	24
PCT_SILT	77 *	-44 *	-6	6	18	-7	4	-15
PHENANTHRENE	96 *	-2	-15	-20	4	-8	-1	4
PYRENE	95 *	-2	-16	-21	3	-8	-2	1
SELENIUM	54 *	-17	29	16	17	11	-16	-22
SILVER	-4	0	47 *	-38 *	-4	14	14	40 *
THALLIUM	1	-2	-72 *	51 *	-32 *	-2	18	10
TOTAL_SOLIDS	-6	58 *	-21	-2	58 *	35 *	38 *	-3
TOT_ORG_CARBON	-3	28	21	44 *	33 *	-4	-43 *	6
TVS	-9	74 *	-23	-2	51 *	28	20	0
ZINC	68 *	-45 *	1	19	16	13	-1	-28

THg and MeHg are associated w/ low heavy metals, high total volatile solids, total solids , percent clay

Method of Analysis

- Principal components of other measured compounds are viewed as possible, though unlikely, explanatory variables
  - Based on PCA
- Principal components of river bound biota are also viewed as possible explanatory variables, based on previous analyses
  - This will be explored

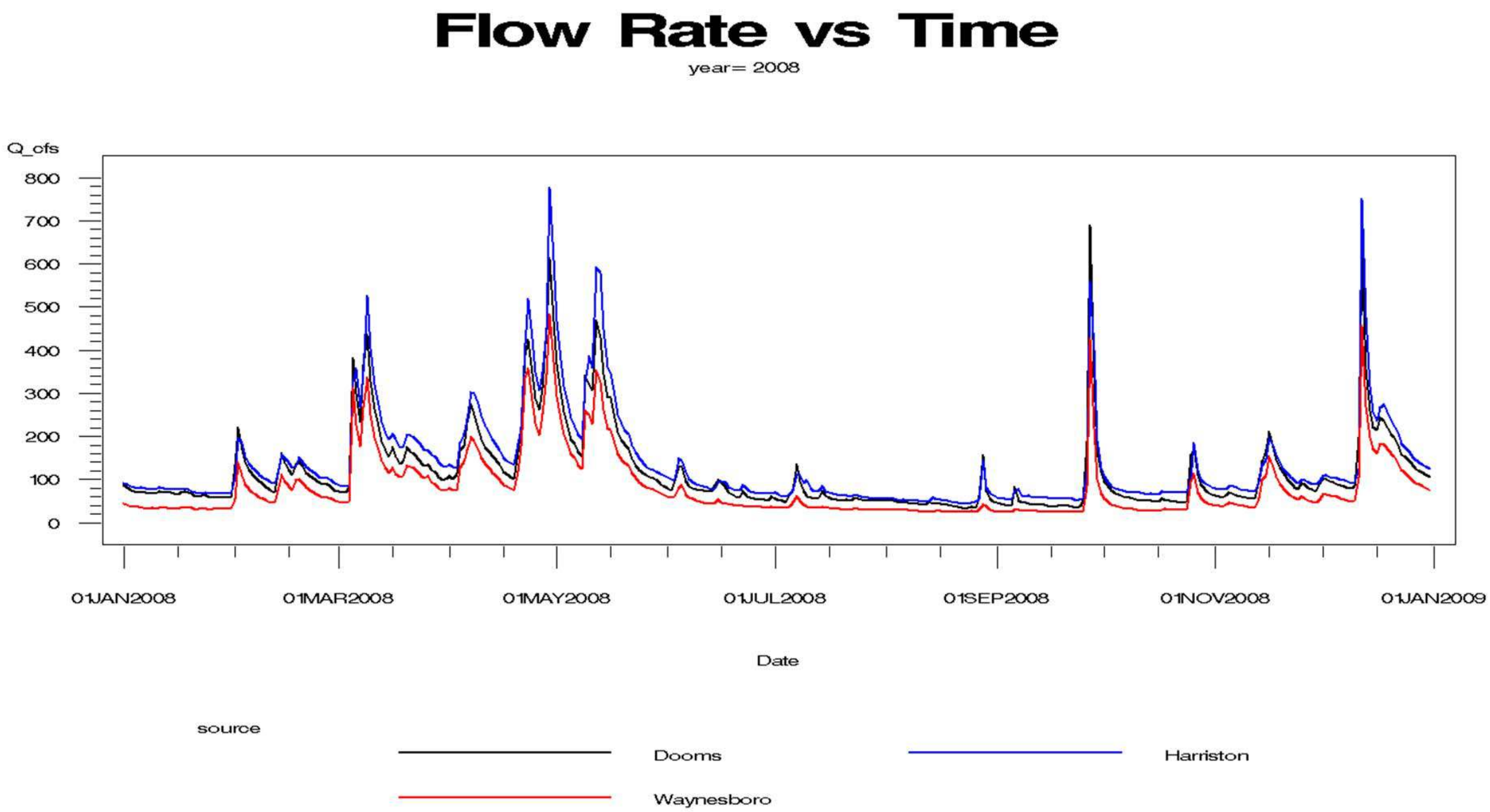
Method of Analysis: SW

- Regression model fit to South River SW THg with
  - Discharge/flow rate
    - Serves as surrogate for annual and seasonal variation and storms
    - 1-10-day "lookbacks" capture lingering storm effects
  - Month
    - Capture seasonal variation
  - Amount of other Hg, MeHg, and chemicals in sediment
    - 60-day, 1-3-RRM "lookbacks" capture relevant upstream sediment
    - Sediment chemicals of possible relevance from correlations and PCA
  - Loads from banks
    - 1-5 RRM and all upstream "lookbacks" by left/right and total capture relevant upstream loading from riverbank
  - Loads from tribs and floodplain
    - 1-3 RRM "lookbacks"
  - Flow rate at sample time and 1-10 day "lookbacks"
    - a surrogate for storm and seasonal effects
  - Land use at river edge

Method of Analysis: SW

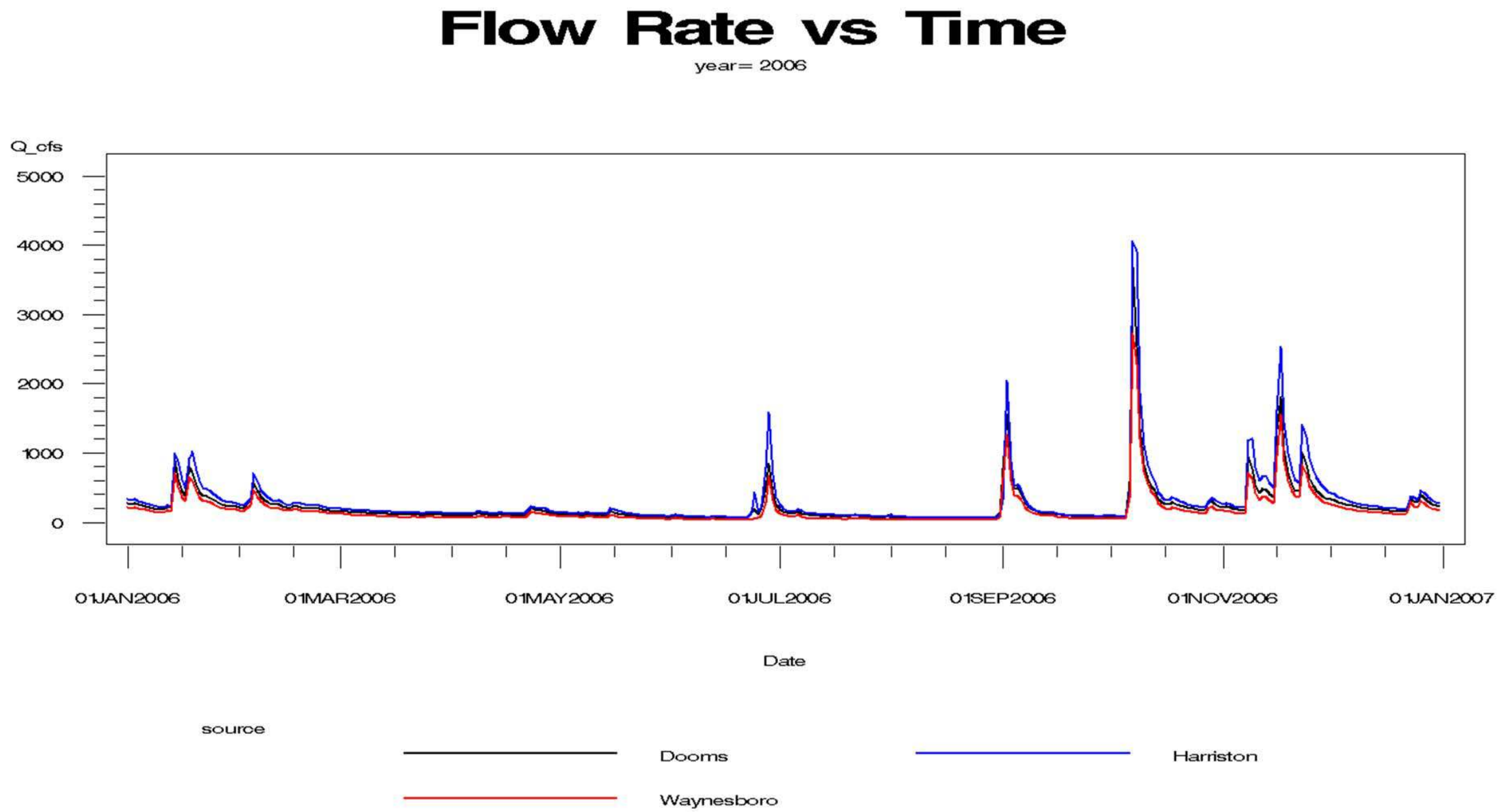
- Stepwise methods to select most important explanatory variables
- Residual analysis, regression diagnostics to refine model and identify important observations
- Bank loading values are annual means
  - Most of load from banks expected to enter river during high flow, i.e., storms
  - Interactions of load and flow in model

Flow Rates-2008



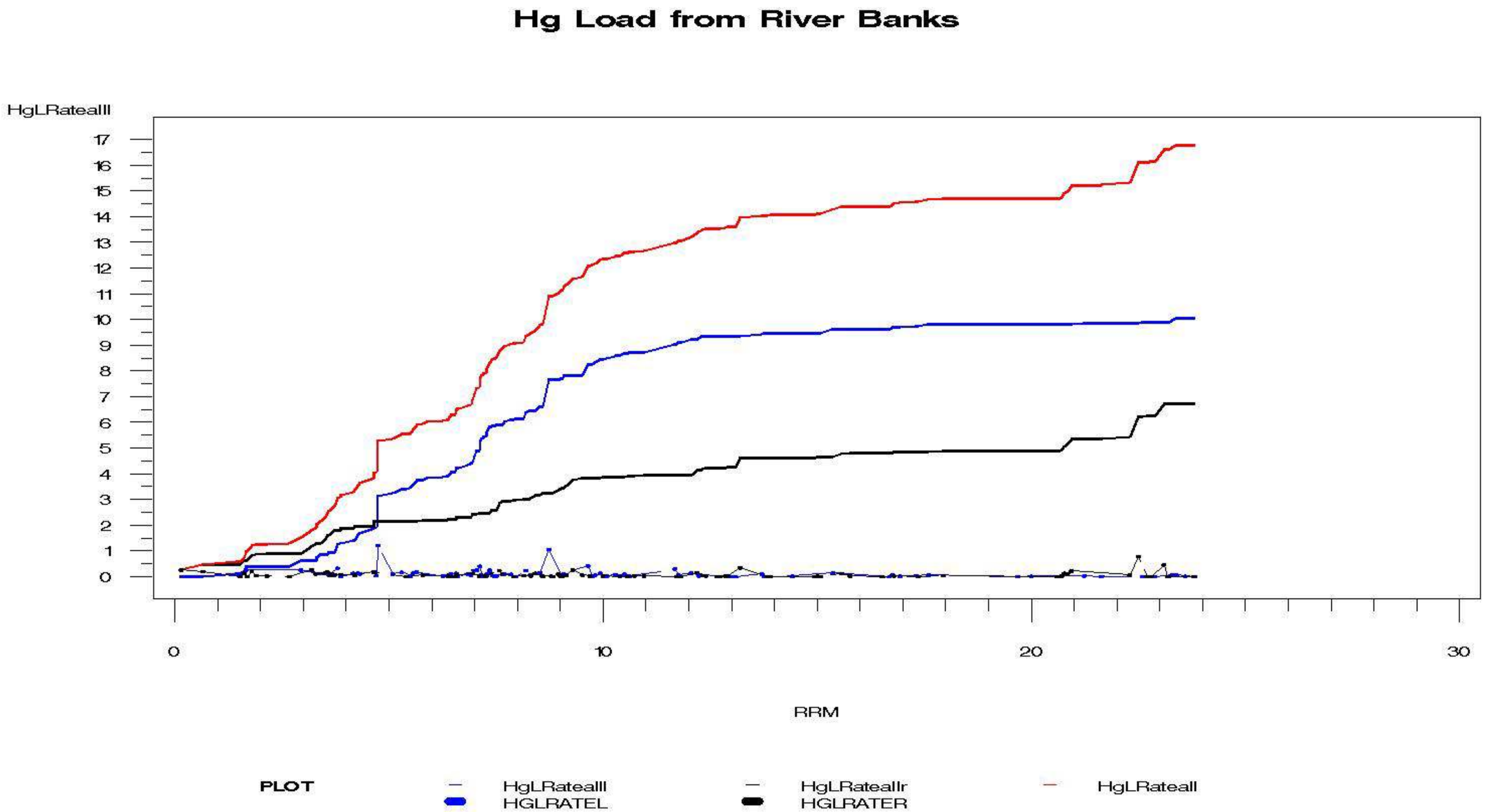
Storms and low and high flow rate periods easy to detect

Flow Rates-2006



Storms and low and high flow rate periods easy to detect. Note also the generally higher flow rates compared with 2008

River Bank THg Loading



The left bank contributes more THg than the right bank. Note hotspots near RRM 6, 9, and 23.

Model Results SW THg

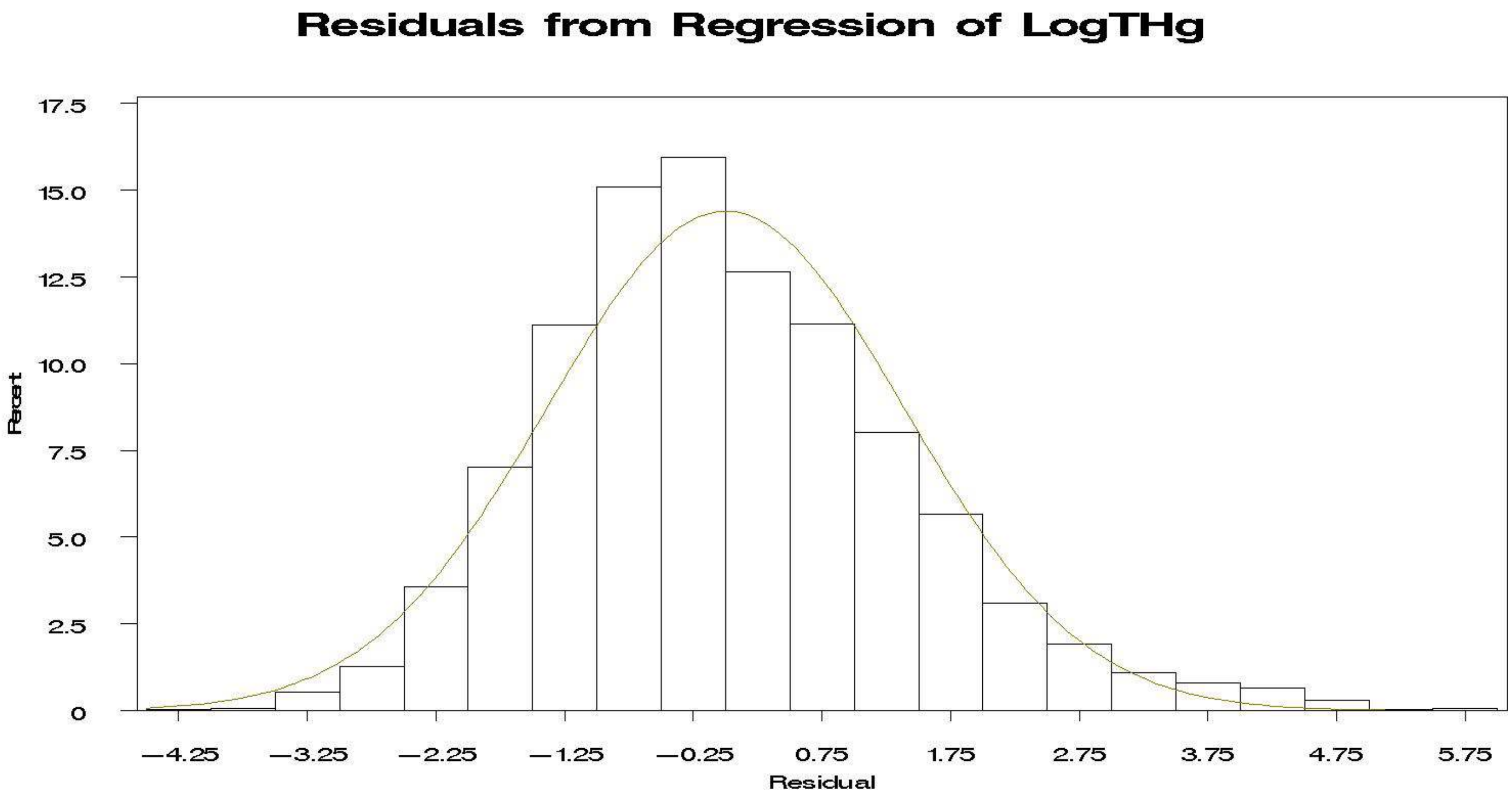
Variable	ParmEst	t Value	Pr > |t|	R-Square=0.38
Intercept	0.81116	13.66	<.0001	
MON5	0.23483	2.47	0.0136	
MON6	0.31795	3.38	0.0007	
MON7	0.84913	7.38	<.0001	
MON8	0.24790	2.38	0.0172	
MON9	0.39450	5.04	<.0001	
MON12	-0.37037	-2.23	0.0255	
MEHG_1	-0.01469	-1.80	0.0715	Sediment
MEHG_2	0.05526	5.64	<.0001	Sediment
CHROMIUM_3	-0.00000625	-4.95	<.0001	Sediment
MANGANESE_1	-3.7872E-7	-1.83	0.0671	Sediment
MANGANESE_2	3.090427E-7	2.51	0.0121	Sediment
PCT_SAND_1	0.01313	3.88	0.0001	Sediment
PCT_SAND_3	-0.01002	-3.19	0.0014	Sediment
HgLRateL1	-1.20371	-4.62	<.0001	Bankload
HgLRateR2	-2.72519	-11.20	<.0001	Bankload
HgLRateR3	2.01481	4.76	<.0001	Bankload
HgLRateR4	-0.71473	-2.16	0.0308	Bankload
HgLRate1	1.80014	6.79	<.0001	Bankload
HgLRateallr	0.48278	18.56	<.0001	Bankload

Model Results SW THg, cont

Variable	ParmEst	t Value	Pr > |t|	
QHG12	0.00063825	3.69	0.0002	Flow*Bankload
QHGR14	-0.00088644	-2.26	0.0237	Flow*Bankload
QHGALLR1	0.00008986	2.56	0.0106	Flow*Bankload
QHGR45	-0.00114	-3.30	0.0010	Flow*Bankload
QHGALLL4	0.00023877	5.79	<.0001	Flow*Bankload
QHGR54	0.00074518	2.11	0.0354	Flow*Bankload
QHGALLL5	-0.00012008	-4.03	<.0001	Flow*Bankload
QHGR101	0.00252	7.24	<.0001	Flow*Bankload
QHGALLR10	-0.00036138	-4.99	<.0001	Flow*Bankload
QHGR4	-0.00064384	-3.74	0.0002	Flow*Bankload

Note: Tribs do not appear.
Land use on shore does not appear.

Residual Analysis THg



Model Results SW THg, cont

- **Flow rate and riverbank load are the dominant factors associated with THg levels in surface water**
  - **Suggests storm/flood action and bank erosion**
- **Increased sediment THg and MeHg 1-3 RRM upstream indicative of elevated SW THg**
- **Near-by heavy metals not associated with high levels of SW THg**
  - **but heavy metals ~3RRM upstream indicative of elevated SW THg**
- **Omit sediment THg, MeHg = little effect on model**
- **Tribs, groundwater have no discernable effect**

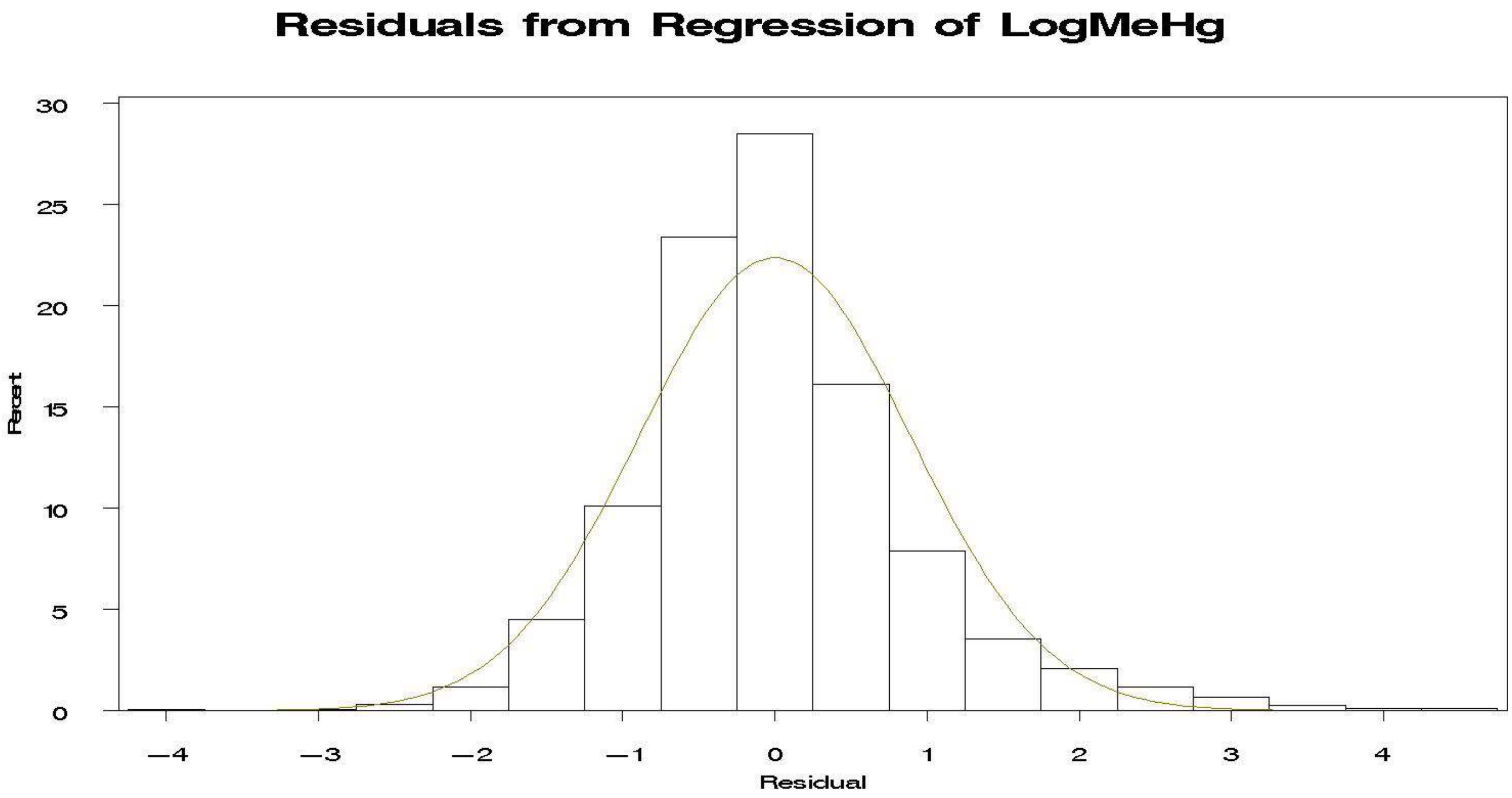
Model Results SW MeHg

Variable	ParmEst	t Value	Pr > |t|	R-Square=0.67
Intercept	-4.35518	-37.78	<.0001	
MON2	0.61782	4.11	<.0001	
MON3	0.66668	5.49	<.0001	
MON4	1.99039	12.87	<.0001	
MON5	2.22459	18.47	<.0001	
MON6	2.28985	18.86	<.0001	
MON7	2.23185	15.38	<.0001	
MON8	1.61692	12.33	<.0001	
MON9	1.72509	14.77	<.0001	
MON10	0.92074	6.48	<.0001	
MON11	0.66316	5.15	<.0001	
MON12	0.28619	1.81	0.0709	
PCT_CLAY_1	-0.01474	-2.95	0.0032	
fQcfsLag4	0.00056610	4.20	<.0001	
HgLRateL1	0.40513	5.76	<.0001	
HgLRateR3	1.18067	5.30	<.0001	
HgLRateR4	-0.58127	-2.84	0.0045	
HgLRatealll	0.24679	34.62	<.0001	
QHGR42	-0.00074079	-2.34	0.0196	
QHGR51	-0.00099123	-2.97	0.0030	
QHGR2	-0.00039687	-2.10	0.0356	
TRIBMEHG_3	0.05441	12.03	<.0001	

Model Results SW MeHg

- Riverbank load, season, flow rate are important factors to MeHg levels in surface water
  - Suggests storm/flood action and bank erosion
  - Right bank appears more important than left
- Decreased %clay indicative of elevated SW MeHg
- No association with heavy metals
- Upstream tribs or groundwater contribute to SW MeHg

Residual Analysis MeHg



Conclusions

- Riverbanks appear to be a major source of THg, MeHg to the river
- Model consistent with deposition of Hg in sediment
- Follow-up work will evaluate
  - 1) the reduction in mercury levels in surface water if input from specific sections of riverbank were eliminated
  - 2) the improvement, if any, of additional terms to the model
    - e.g., biota (incl. aquatic vegetation), environmental factors