

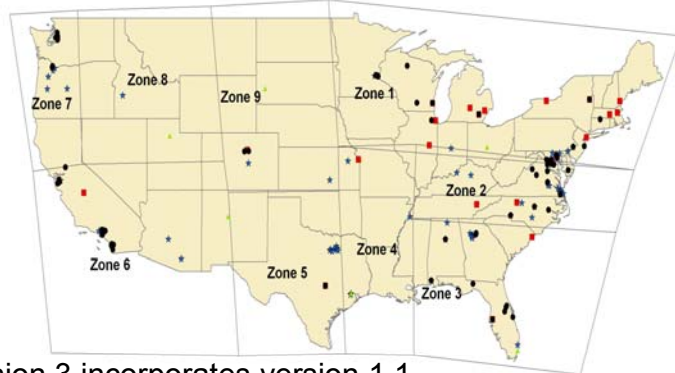
Stormwater Characteristics and Monitoring

Robert Pitt, Cudworth Professor of Urban Water Systems
The University of Alabama, Tuscaloosa, AL 35487

Stormwater NPDES Data Collection and Evaluation Project

- The University of Alabama and the Center for Watershed Protection were awarded an EPA 104(b)3 grant in 2001 to collect, review, and analyze selected Phase 1 NPDES stormwater permit data.
- We received an extension of the project in 2005 to expand the database to include under-represented areas. We recently completed 3.1 of the database (version 2 was not posted as it was an interim version that had not undergone complete QA/QC reviews).
- The National Stormwater Quality Database (NSQD) is available on the Internet.

Communities Included in NSQD version 3



Version 3 incorporates version 1.1 data, plus additional MS4 data, along with selected data from the International BMP Database, the USGS, and NURP.

Database Representation

- BMP
- NURP
- ▲ USGS
- ★ NSQD

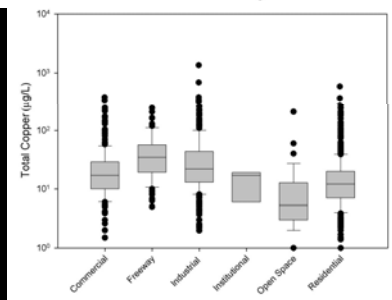
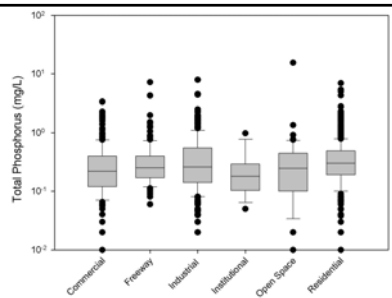
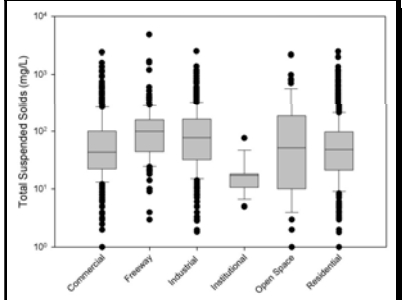
Number of Events and Geographical Coverage in NSQD ver. 3

RAIN ZONE	TOTAL EVENTS	PERCENTAGE
Zone 1- Great Lakes and Northeast	1,271	15
Zone 2- Mid Atlantic	3,984	46
Zone 3- Southeast	744	9
Zone 4- Lower Mississippi Valley	301	4
Zone 5- Texas	799	9
Zone 6- Southwest	417	5
Zone 7- Northwest	865	10
Zone 8- Rocky Mountains	24	0.3
Zone 9- Midwest	197	2
TOTAL	8,602	100

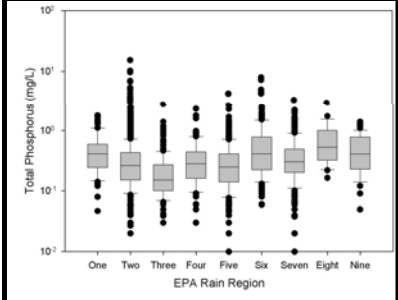
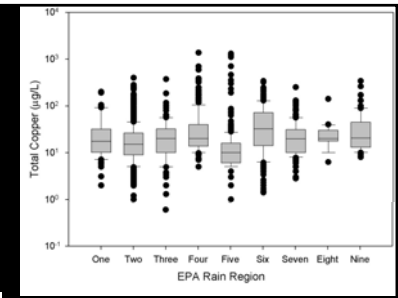
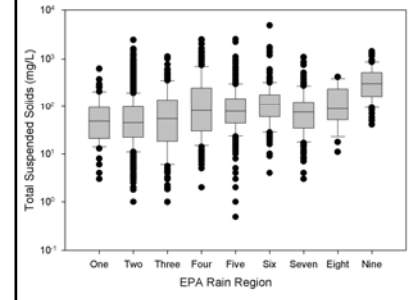
Number of Events and Land Use Coverage in NSQD ver. 3		
LAND USE	TOTAL EVENTS	PERCENTAGE
Residential	2,933	34
Commercial	1,080	13
Institutional	55	1
Industrial	893	10
Freeway	734	9
Open Space	125	2
Mixed Land Uses	2,782	31
TOTAL	8,602	100

Total Suspended Solids by Land Use and Geographical Area (mg/L)							
		1	2	3	5	7	All
Commercial	Mean	135	86	60	67	81	118
	Count	237	454	50	40	42	916
	COV	1.2	1.8	2.0	1.6	1.1	1.7
Industrial	Mean	177	78	96	244	182	171
	Count	100	304	82	43	24	719
	COV	1.4	1.0	1.3	1.6	1.2	1.7
Residential	Mean	140	85	107	109	100	123
	Count	332	1,388	122	107	170	2,386
	COV	1.2	1.7	1.6	1.0	0.9	2.0
ALL	Mean	155	97	95	138	126	137
	Count	1,132	3,466	420	488	443	6,780
	COV	1.6	1.7	1.5	1.5	1.7	2.2

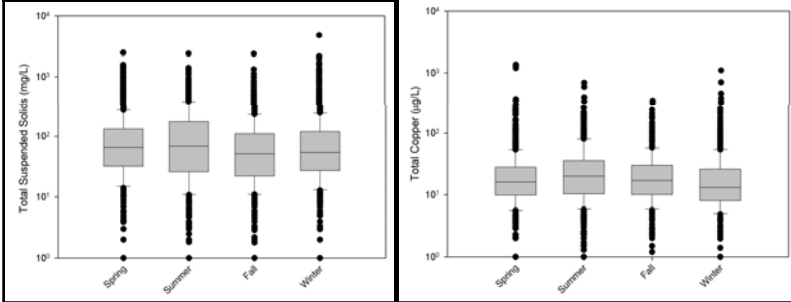
These grouped box-whisker plots sort all of the data by land use. Kruskal-Wallis analyses indicate that all constituents have at least one significantly different category from the others. Heavy metal differences are most obvious.



Residential area concentrations grouped by EPA rain zones. Zones 1-4 are east half of country, zones 5-9 are western half of country. Zones 3 and 7 are the wettest zones.



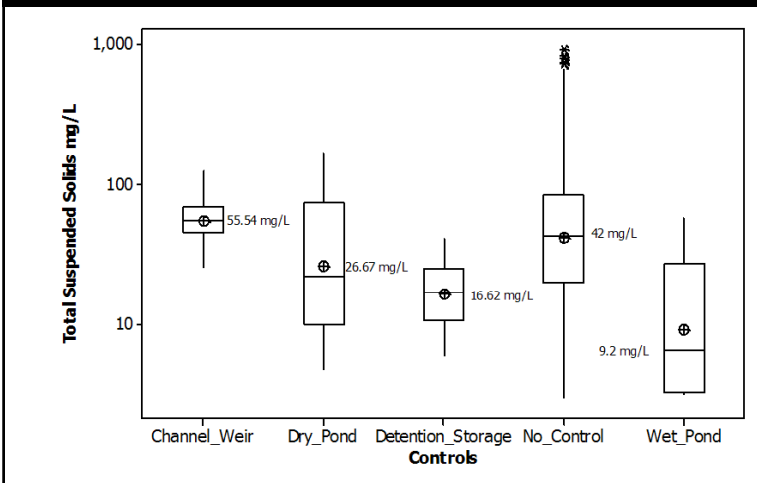
These grouped box-whisker plots sort residential data by sampling season. The most obvious difference is shown for fecal coliforms (a similar conclusion was found during NURP, EPA 1983). (These plots are only for residential data)



Main Factors and Interactions Affecting Outfall Concentrations

	Land Use (LU)	Season (SN)	Rain Zone (RN)	LU*SN	LU*RN	SN*RN	LU*RN*SN
TSS	<0.001	0.737	<0.001	0.017	<0.001	0.184	<0.001
BOD	<0.001	0.155	<0.001	0.001	<0.001	0.001	0.221
COD	<0.001	0.134	<0.001	0.034	<0.001	0.014	0.009
TP	<0.001	0.687	<0.001	0.055	<0.001	<0.001	<0.001
NO2+NO3	<0.001	0.108	<0.001	0.052	<0.001	0.034	0.057
Cu	<0.001	0.112	<0.001	0.623	<0.001	0.038	0.141
Pb	<0.001	0.765	<0.001	0.420	<0.001	0.285	0.012
Zn	<0.001	0.910	<0.001	0.936	<0.001	0.014	<0.001

Comparison of Stormwater Control Practices (Residential Land Uses EPA Rain Zone 2)



Why Monitor as Part of MS4 Permits?

- “Characterization” monitoring may not be necessary unless in under-represented areas or land uses.
- Monitoring at small scales (having homogeneous characteristics) more useful than for large multi-land use locations.
- More efficient to require monitoring to learn about processes (sources, transport, control, and effects) and for program assessment/validation.
- A coordinated monitoring program for an area would be much more efficient than a standardized “one-size-fits-all” approach.

Recommendations for Improved Future Regulatory Monitoring Activities

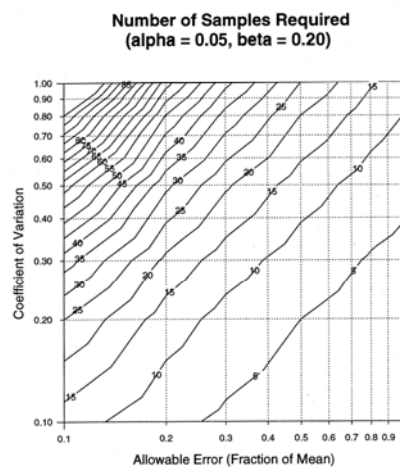
- Better site descriptions (drainage area delineation, effective percentage impervious area, transient and adjacent activities that may affect water quality) are always needed.
- Adequate on-site rain gauges and flow monitoring critical.
- Monitor for the complete event duration (not just “first flush,” or only for 3 hours)
- Statistical analyses indicated differences between automatic and manual sampling. Automatic flow-weighted composite sampling may be preferred in most cases, supplemented with bed load and floatables sampling.

Bedload samplers installed at WI DNR/USGS monitoring location. About 5% of annual sediment was in bedload fraction that was not captured by automatic samplers.



Experimental Design Number of Samples Needed

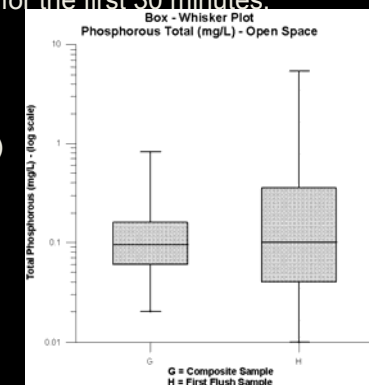
The number of samples needed to characterize stormwater conditions for a specific site is dependent on the COV and allowable error. For most constituents and conditions, about 20 to 30 samples may be sufficient for most objectives. Most Phase 1 sites only have about 10 events, but each stratification category usually has much more.



Comparison of First-Flush and Composite Samples

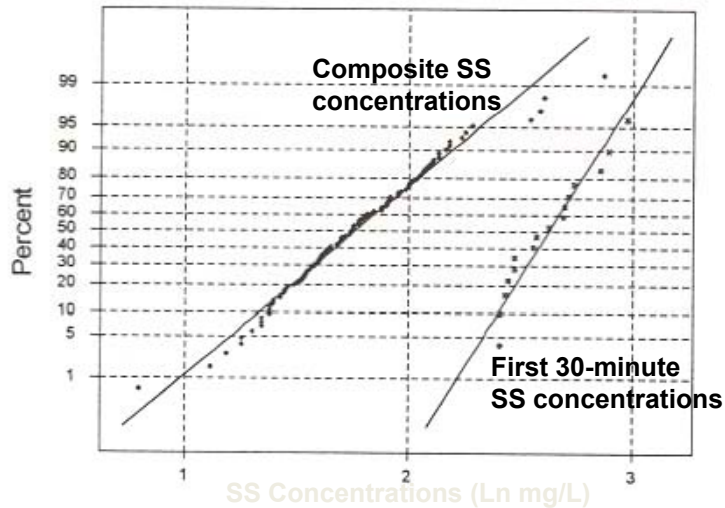
More than 400 paired samples were available for comparison.
The first-flush samples are for the first 30 minutes.

The Fligner-Policello (symmetrical about the medians) and the Mann-Whitney (symmetrical and same variance) non-parametric comparison tests were used to compare the paired first-flush concentrations with the whole storm composite concentrations. The Anderson-Darling test was used to test for normality.



Common for concentrations to be similar, but first-flush variance larger

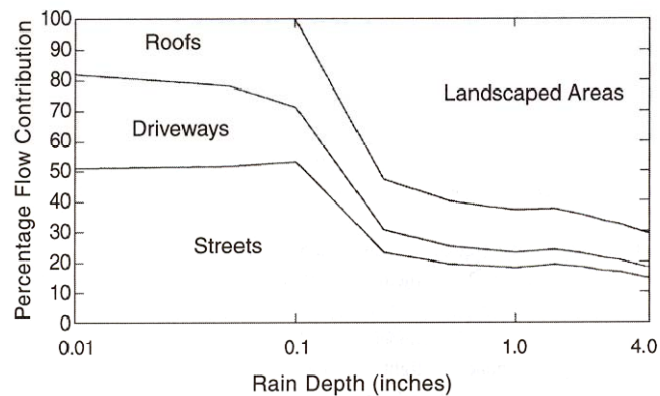
Example for commercial area suspended solids showing a significant first-flush effect:



First-Flush Observations

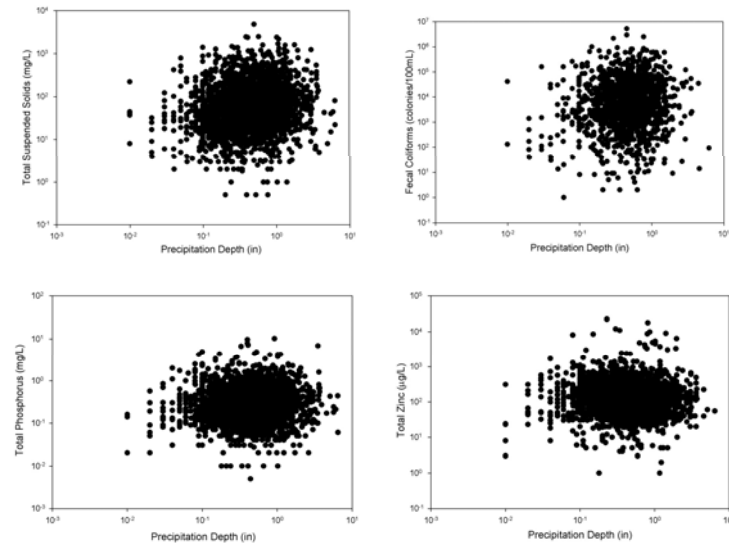
- COD, BOD₅, TDS, TKN, and Zn all had significant first-flushes for all land uses (except for open space).
- The ratio of the first-flush to composite concentrations ranged from 1.3 to 1.7 for these constituents.
- Turbidity, pH, fecal coliforms, fecal strep., total N, dissolved P, and orthophosphate did not have significant first-flushes for most of the separate land uses.
- No open space, and only a few institutional data sets had significant first-flushes.

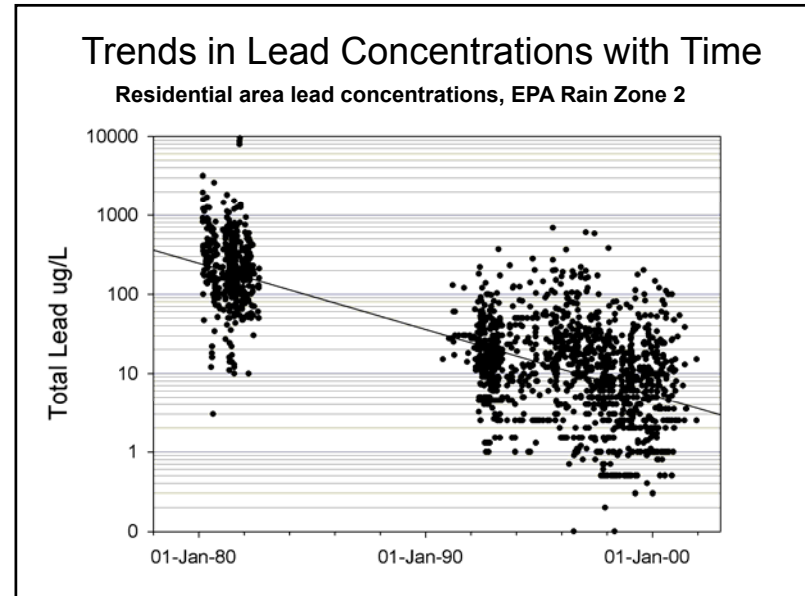
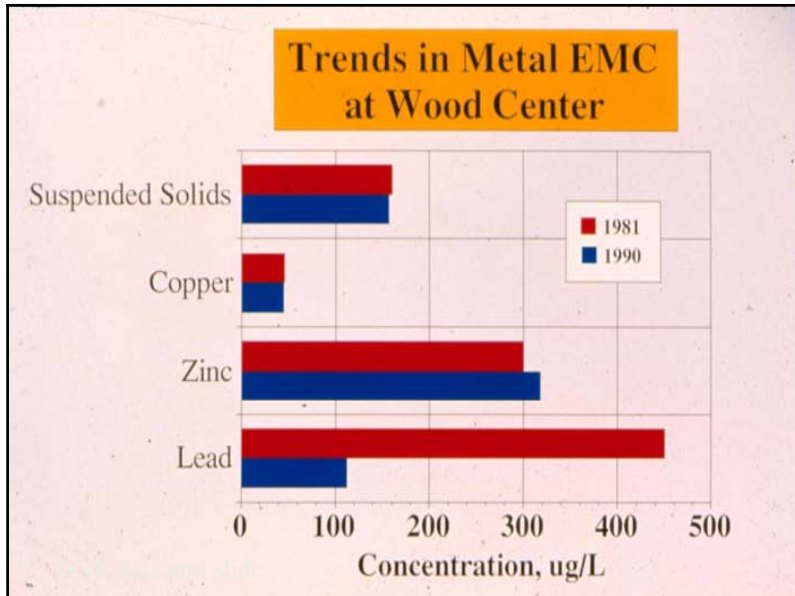
Flow Sources for Different Rain Depths



Runoff originates from different areas as the rain depth changes; "first flush" doesn't represent all flows. Routing of source hydrographs during complex rains results in mixing of first flushes from individual source areas and first flushes not commonly seen at outfalls, unless they drain areas have large impervious area fractions.

Plots of concentrations vs. rain depth typically show random patterns.





NSQD Conclusions

- Much concern expressed about use of Phase 1 MS4 data due to various experimental designs, different sampling and analytical procedures, etc.
- However, the large amount of data, the documentation available (although some hard to locate), and the wide range of conditions included in the monitoring programs, allow a great deal of information to be extracted and summarized.

NSQD Conclusions (cont.)

- The database can be used to evaluate the performance of stormwater controls, type of conveyance, sampling procedures, etc.
- Phase 1 MS4 data shows significant patterns for different land uses and geographical locations for most constituents.
- More data needed in under-represented areas for more complete evaluations.

Download the NSQD and supporting information at:
<http://unix.eng.ua.edu/~rpitt/Research/ms4/mainms4.shtml>

Basic Monitoring Strategy

- Scale-up of monitoring from pilot to full-scale control devices
- Need flexibility of small units and control to test many variables under large variety of conditions
- Need to verify with full-scale units to check performance under real-world conditions

Different Pilot-Scale Treatment Setups



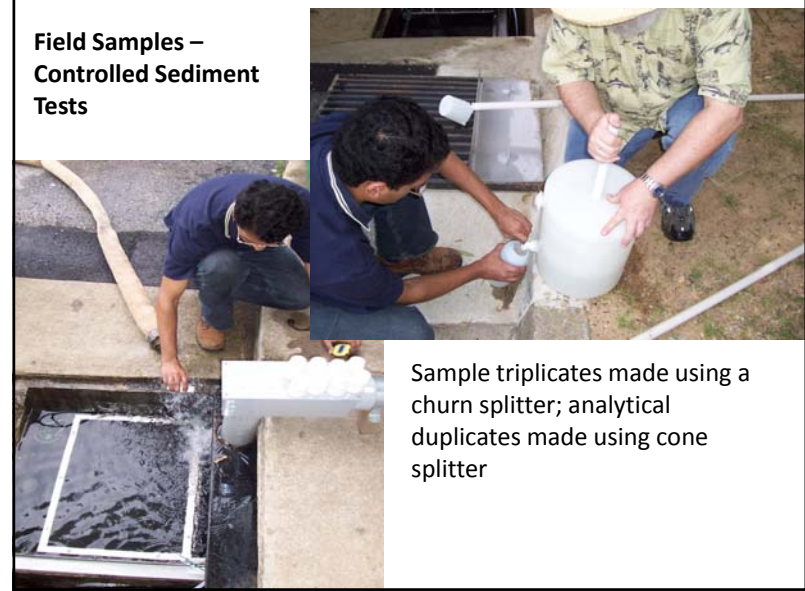
Milwaukee, WI, Ruby Garage Public Works Maintenance Yard MCTT Tests (0.25 acre site)



Minocqua, WI, MCTT Tests (2.5 acre site)

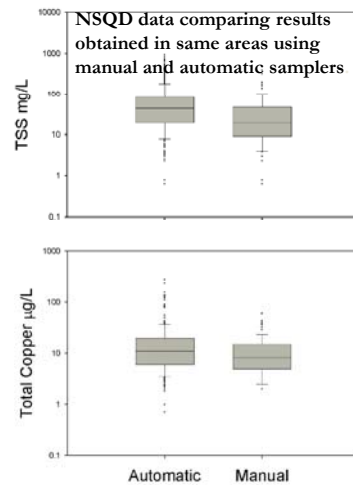


Field Samples – Controlled Sediment Tests

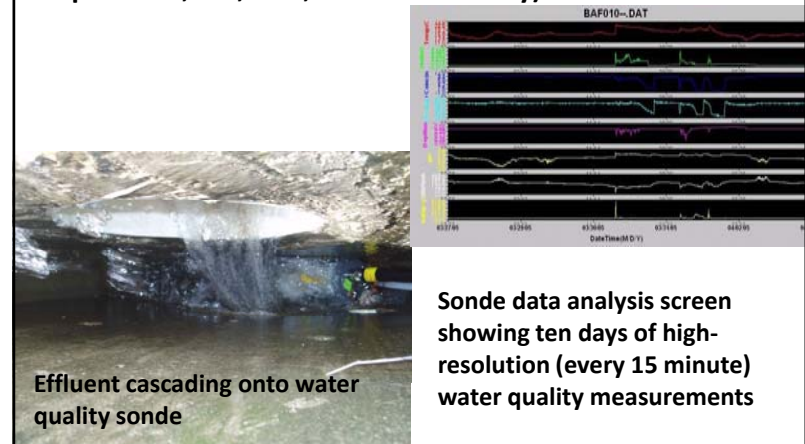


Sample triplicates made using a churn splitter; analytical duplicates made using cone splitter

May have small biases with automatic vs. manual sampling, but automatic sampling allows unattended operation under a variety of conditions and captures complete event. Manual sampling can better represent complete range of particulate matter in sample.



The use of continuous water quality sondes can supplement other sampling programs by providing high resolution data for a variety of constituents (turbidity, temperature, DO, ORP, and conductivity).



Effluent cascading onto water quality sonde

Sonde data analysis screen showing ten days of high-resolution (every 15 minute) water quality measurements

Issues Concerning Stormwater that May Need to be Addressed

- Rainfall patterns must be considered for area being studied, and accurate flow measurements are necessary as performance is commonly related to hydraulic conditions. Most flow instruments must be calibrated at the site.
- The variability of stormwater quality must be considered when designing a sampling program.
- Incorrectly reported data can have a very large effect on many statistical analyses
- Variability of stormwater quality does not always vary as anticipated (“first-flush” relatively rare, unless mostly paved areas and small drainage areas; little relationship with rain depth of event)
- Sources of flows and pollutants vary with land use and development characteristics

Probability distribution of rains (by count) and runoff (by depth).

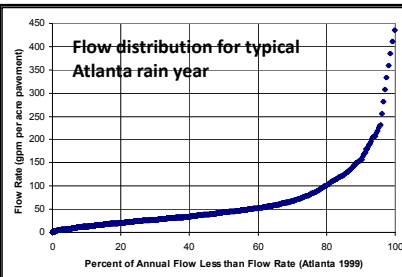
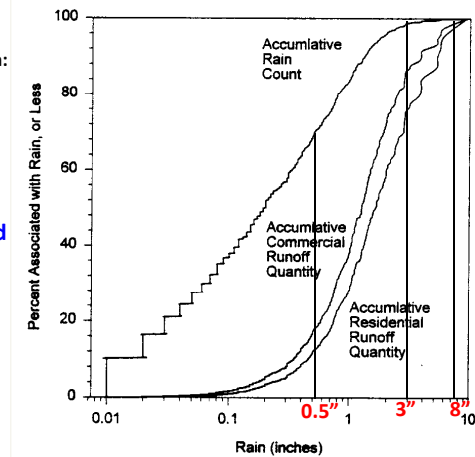
Central Alabama Rain Condition:
 <0.5" (<10 mm): 65% of rains (10% of runoff)

0.5 to 3" (10 to 75 mm): 30% of rains (75% of runoff) **We therefore need to focus on these rains!**

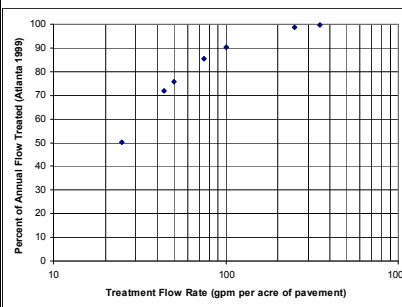
3 to 8" (75 to 400 mm): 4% of rains (13% of runoff)

>8" (>400 mm): <0.1% of rains (2% of runoff)

Birmingham, AL Rain & Runoff Distributions ('81-'89)

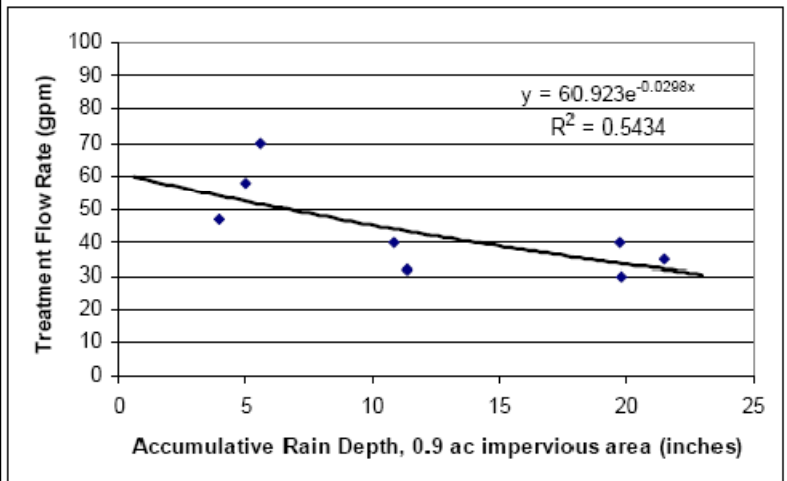


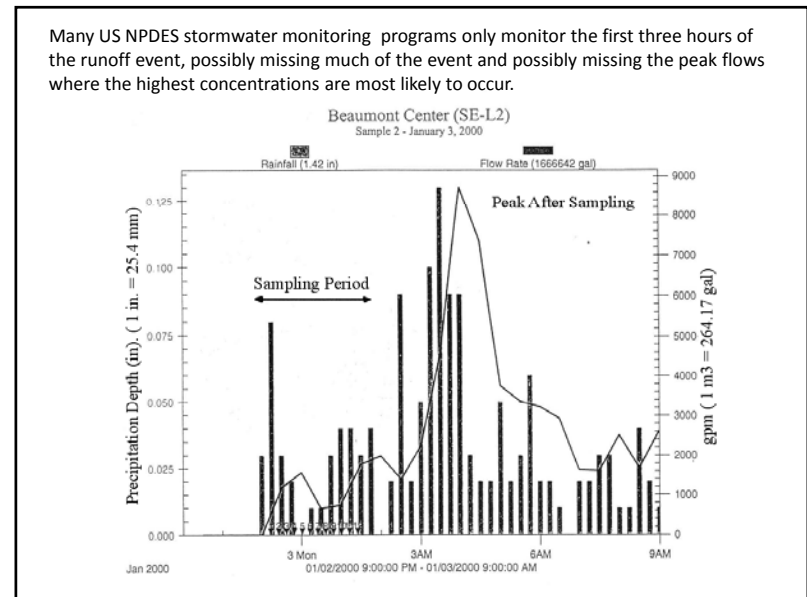
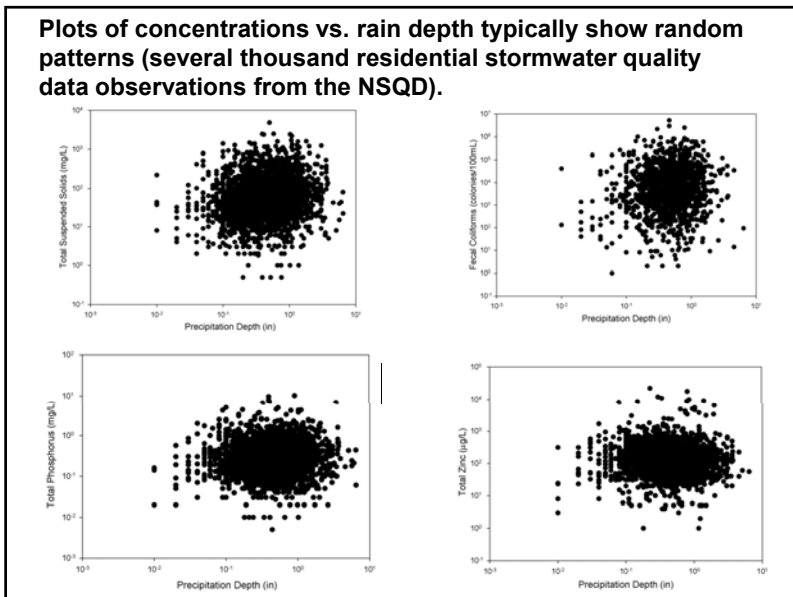
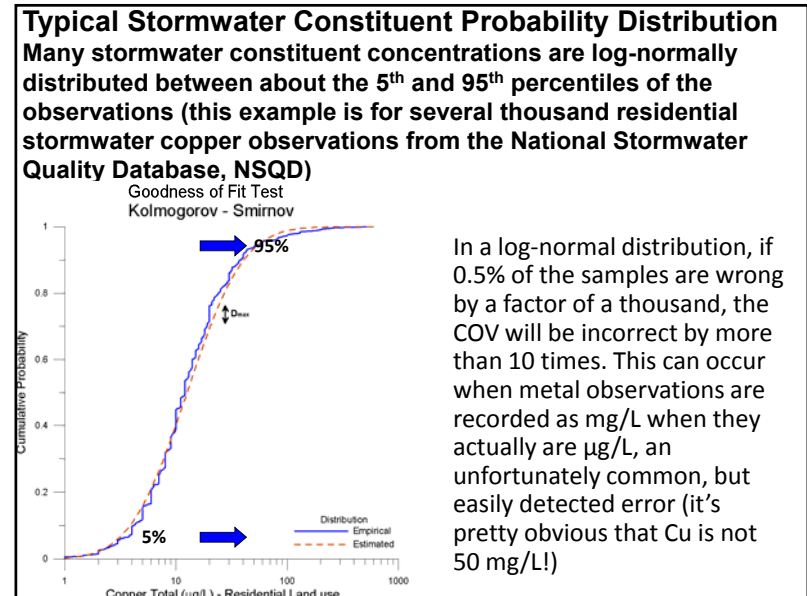
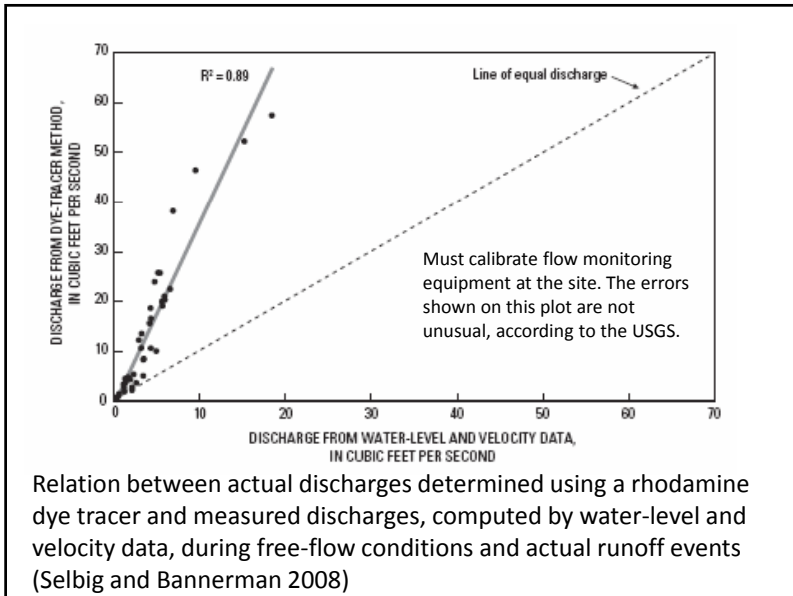
Continuous Simulation can be used to Determine Needed Treatment Flow Rates:
 - 90% of the annual flow for SE US conditions is about 170 gpm/acre pavement (max about 450).

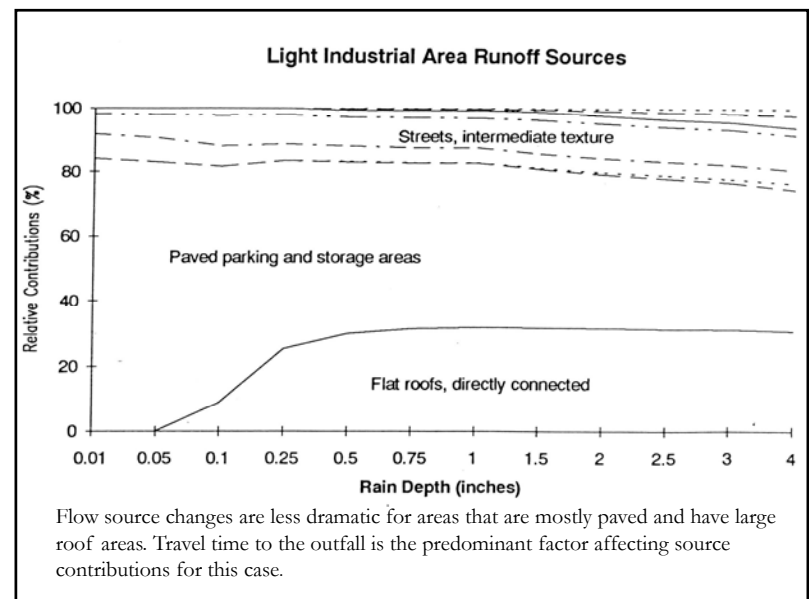
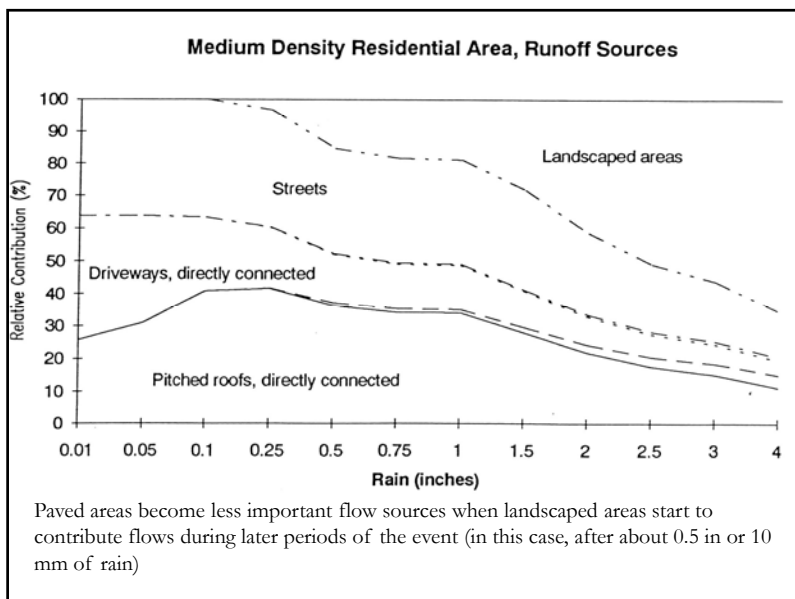
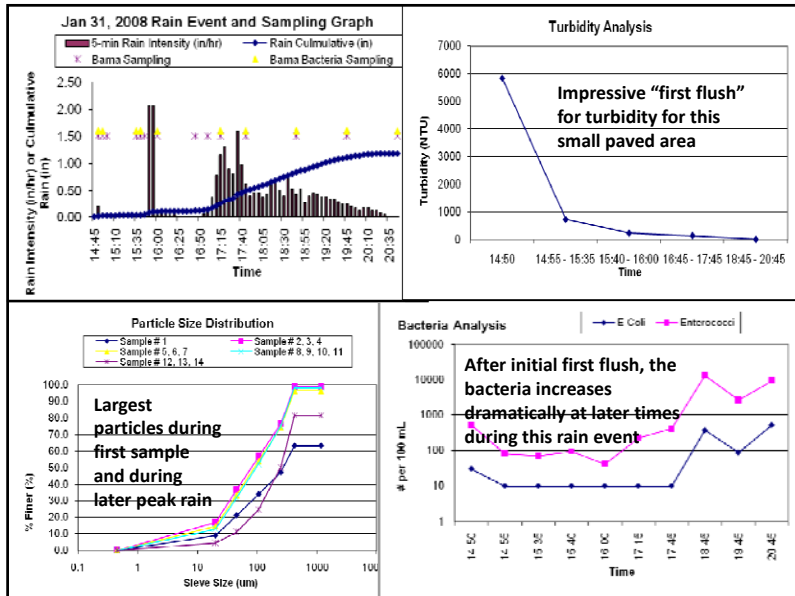


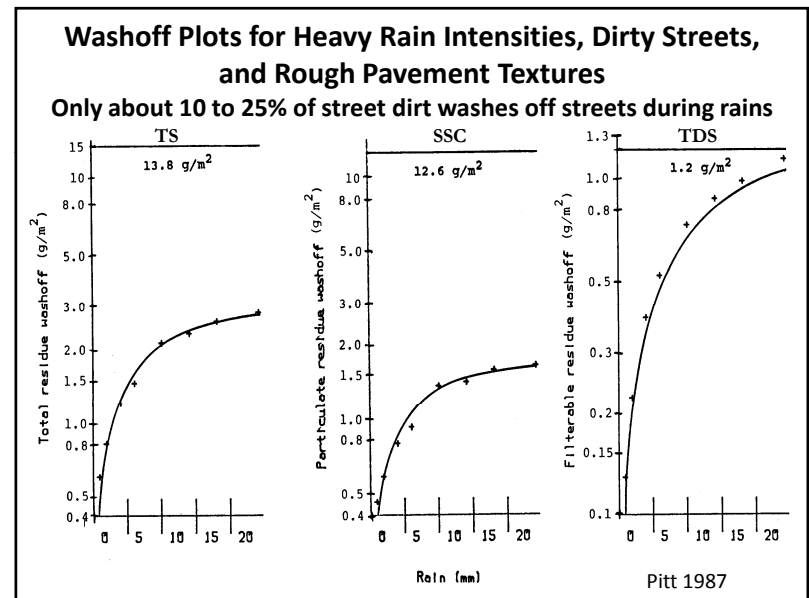
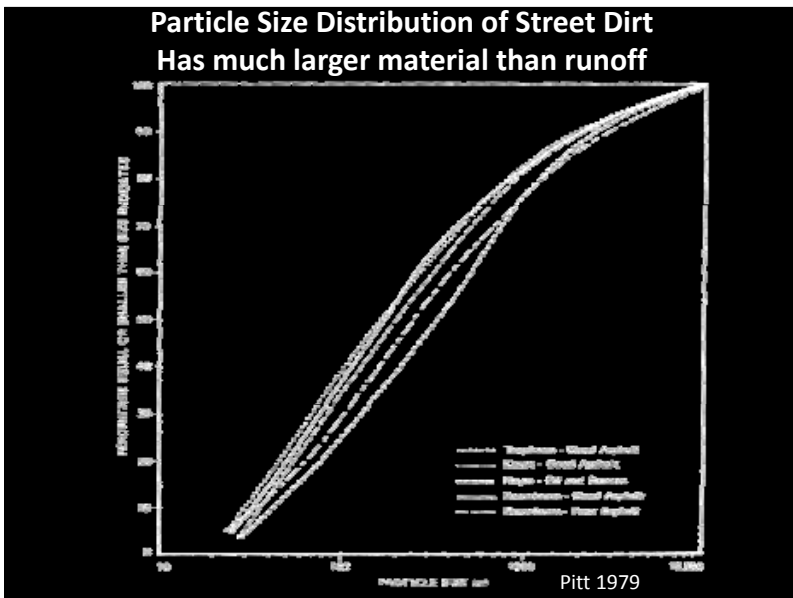
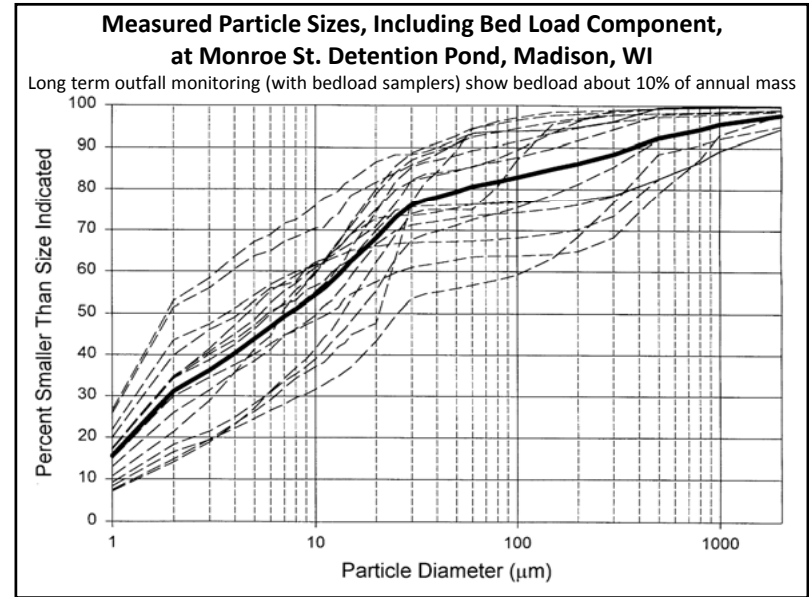
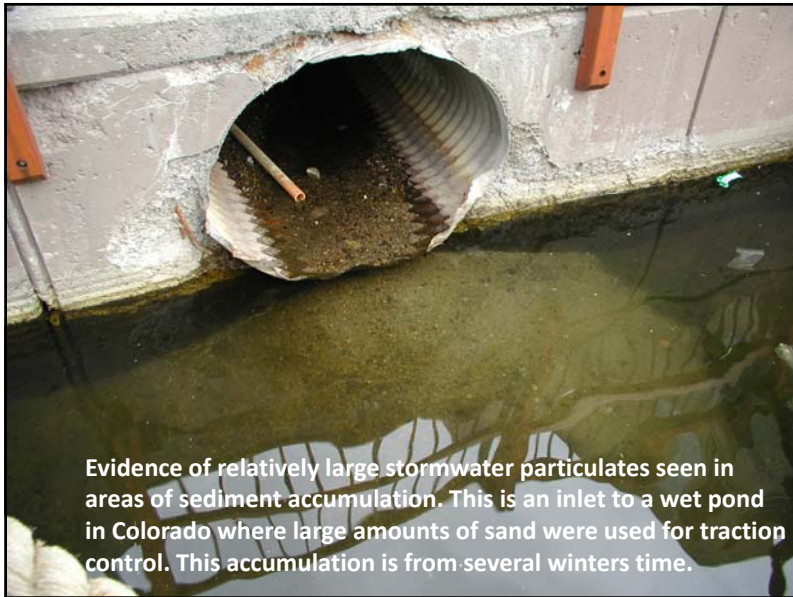
- treatment of 90% of annual runoff volume would require treatment rate of about 100 gpm/acre of pavement. More than three times the treatment flow rate needed for NW US.

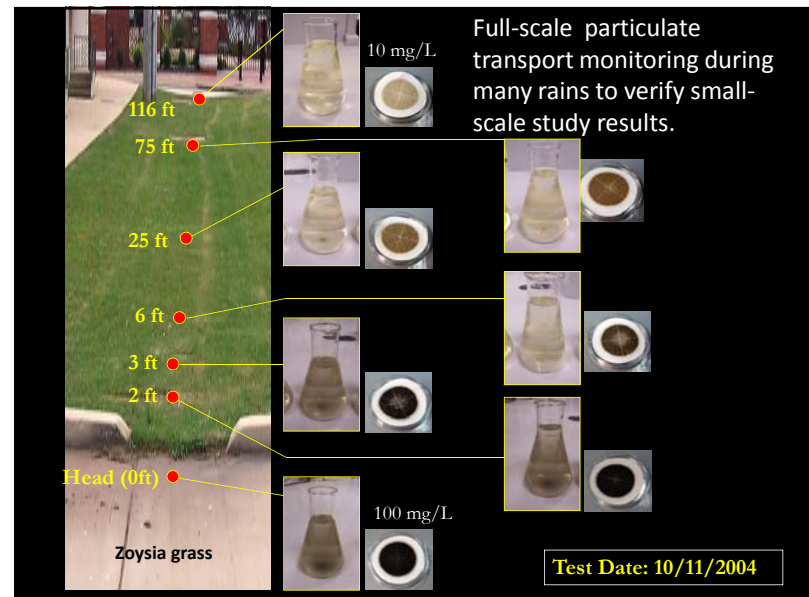
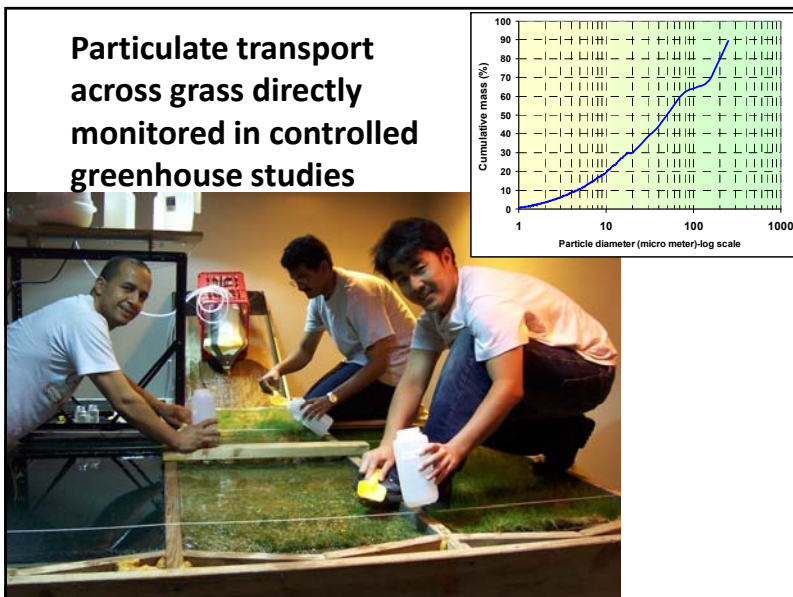
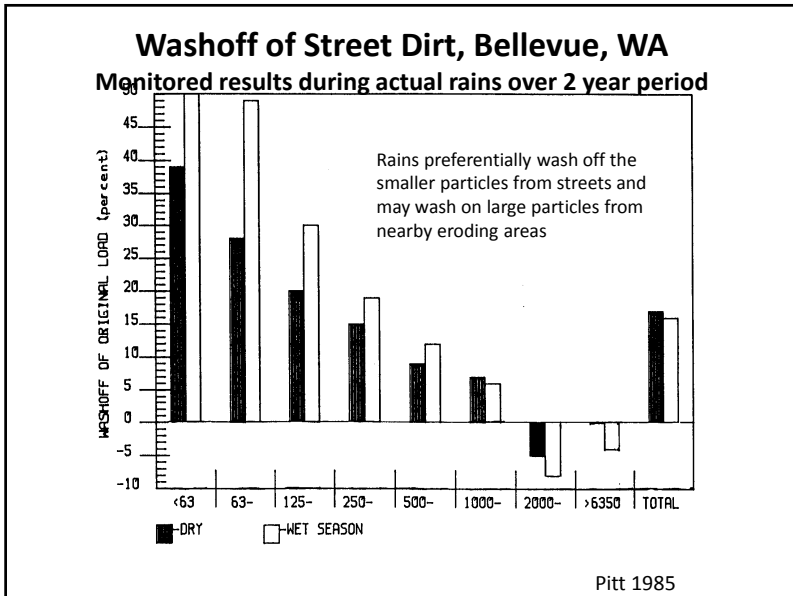
Treatment Flow Rate Changes during 10 Month Monitoring Period of Prototype UpFlo™ Filter





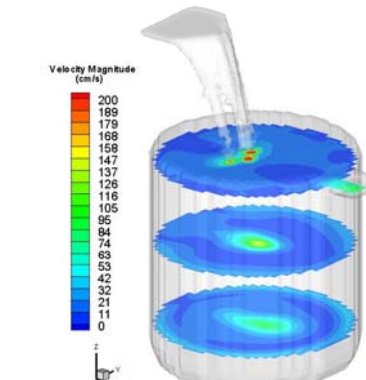






Physical and 3D-CFD Modeling

Scour tests of previously deposited sediment in sumps



CFD modeling being verified by full-scale 3D flow field measurements

Velocity and shear stress for different slopes and depths (2 ft pipe)

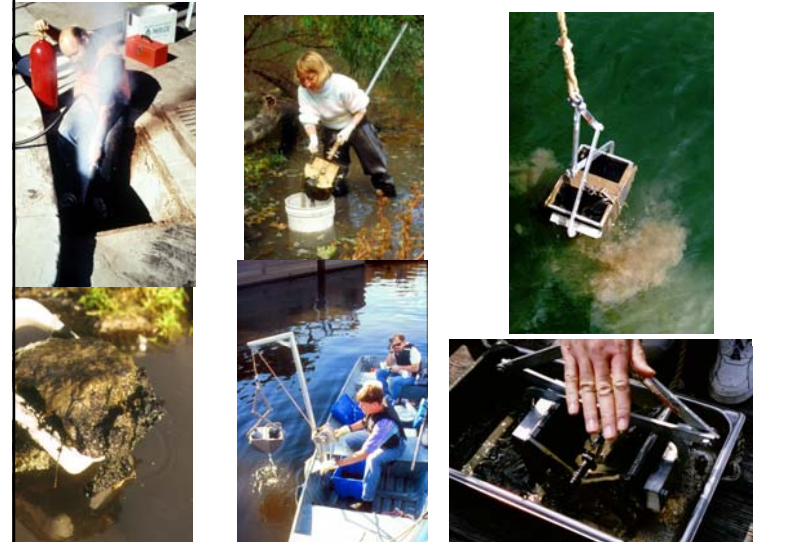
Depth/ Diameter ratio	Velocity (ft/sec) 0.1% slope	Shear stress (lb/ft ²) 0.1% slope	Velocity (ft/sec) 2% slope	Shear stress (lb/ft ²) 2% slope
	0.1	0.91	0.0081	4.1
0.5	2.3	0.031	10	0.62
1.0	2.3	0.031	10	0.62

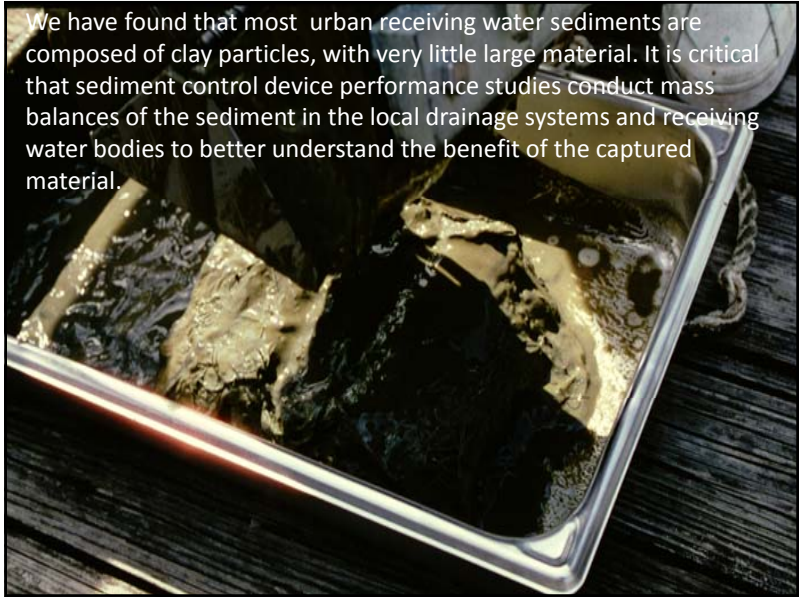
Pipes having small slopes allow particles >100 μm to settle and form permanent deposits, while pipes with large slopes will likely have moving beds of larger material.

Coarse Floatable Control and Monitoring Also Important in Many Areas



We have monitored sediment transport in storm drainage systems and accumulated sediment in urban receiving waters to quantify the fate and transport of urban stormwater particulates.





Stormwater Sampling

- Important to monitor sources, transport, and fate of stormwater pollutants.
- Need to program automatic samplers to collect samples under a wide-range of flow conditions.

Many types of runoff monitoring have been used to understand their transport and fate, from small source areas to outfalls.



Many stormwater monitoring configurations used over the years





It is difficult to program an automatic sampler to collect flow-weighted samples over a wide range of flow conditions.

use time-compositing instead of flow-weighted sampling and then manually composite the sample using the available flow data

use a large sample base in order to accommodate a wide range of runoff events

use two samplers located at the same location, one optimized for small events, the other optimized for larger events

Special Sampling and Handling Needs – solids processing

- A wide range of sample characteristics need to be considered in a sampling program
- Automatic samplers are not effective in collecting large particles; recovery of particles >250 μm is usually <50%, while they can be close to 100% effective for particles <100 μm .
- In most cases, the actual errors in annual mass discharges are <10%. However, complete mass balances need to be done as part of control practice monitoring to quantify the errors and to identify the large particle fraction.
- Particle size information is one of the most important stormwater characteristics affecting treatability, transport, and fate.
- Cone splitters need to be used to divide samples for analyses and SSC (suspended sediment concentration) should be used instead of TSS for the most repeatable results.
- Discrete particle size pollutant analyses on different particle sizes can also be important for treatability and fate analyses.

Particle Size Distributions for Stormwater

Suitable sampling and measurement methods must be able to handle a wide range of particle sizes

Generally, larger median particle sizes at source areas and inlets and smaller median particle sizes at outfalls to receiving waters. Stream sediments will accumulate the largest particulates discharged from outfalls.

USGS and WI DNR Monitoring Facility for Hydrodynamic Separator Tests, Madison, WI

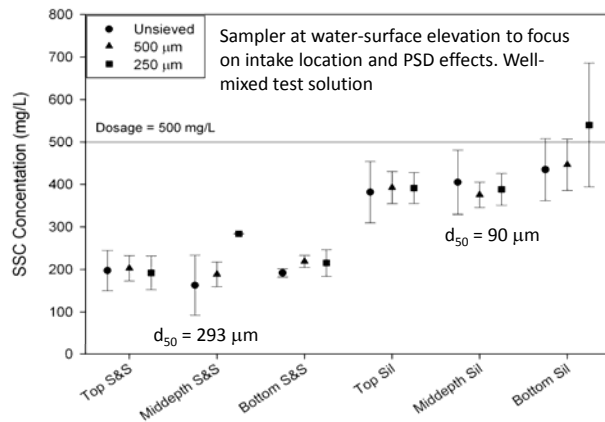


Results of Verification Monitoring of Hydrodynamic Separator (Madison, WI)

Sampled solids load in	1623 kg
Sampled solids load out	1218 kg
Trapped (by difference)	405 kg (25% removal)
Actual trapped total sediment by measuring captured material	536 kg (33% actual removal)
Total solids not captured by automatic samplers	131 kg out of 1623+131 kg missed (8%)

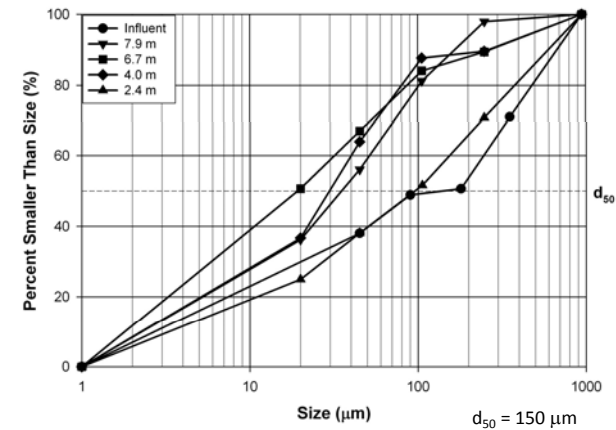
USGS data

Effect of Intake Location and Solids PSD



- Well-mixed water column required to not see biases in intake location. Smaller particle sizes less subjected to bias because they stay suspended.
- Sand-sized particles much more problematic; sands > 250 μm not highly recovered.

Sampler Height Effects



In this example, sampler heights >2.5 m resulted in fewer larger particles in sampler.

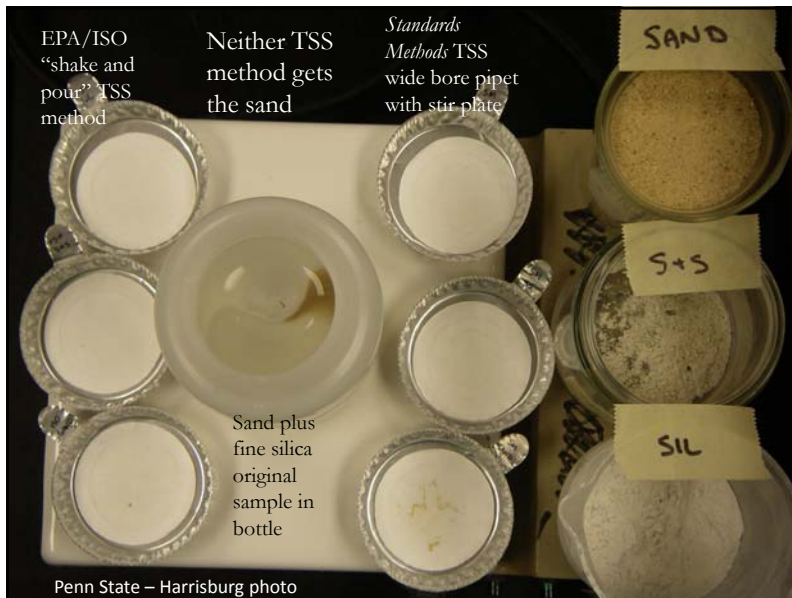
Stirred then settled sample, showing settleable solids (collected with automatic sampler during Madison, WI, high-efficiency street cleaner tests)



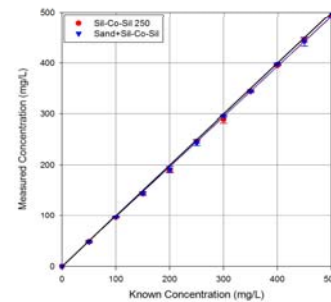
Comparison of three TSS/SSC analytical methods

	EPA TSS (160.2) ISO 11923	Standard Methods TSS (2540D)	USGS SSC (ASTM D3977-97(B))
Filter Nominal Pore Size	Not specified	< 2.0 μm	1.5 μm
Sample Mixing	Shake vigorously	Stir plate	Decant supernatant & flush bottle with DI
Aliquot Size	Not specified (normal 100 mL)	Not specified (normal 100 mL)	Entire sample
Method of Aliquot Collection	Pour aliquot into graduated cylinder	Pipet: mid-depth in bottle & mid-way between wall and vortex	Pour from original bottle

Tested differences between methods using samples from 50 – 500 mg/L particulate matter having two different particle size distributions (PSDs), d_{50} of 90 and 260 μm .



Suspended Sediment Concentrations Compared to Known Laboratory Additions

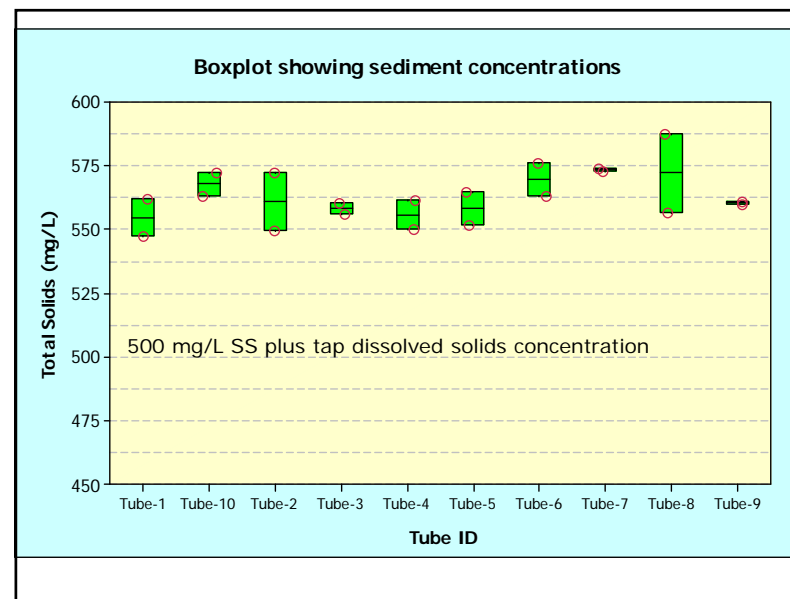
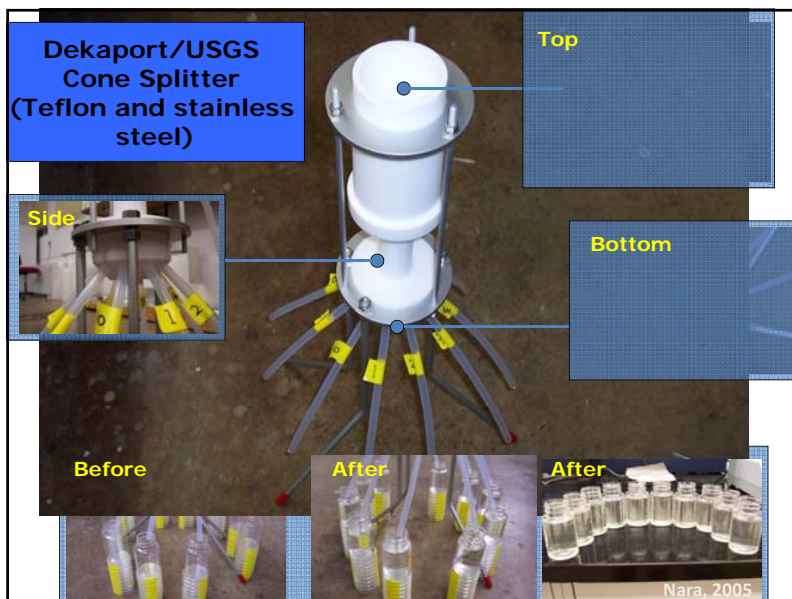


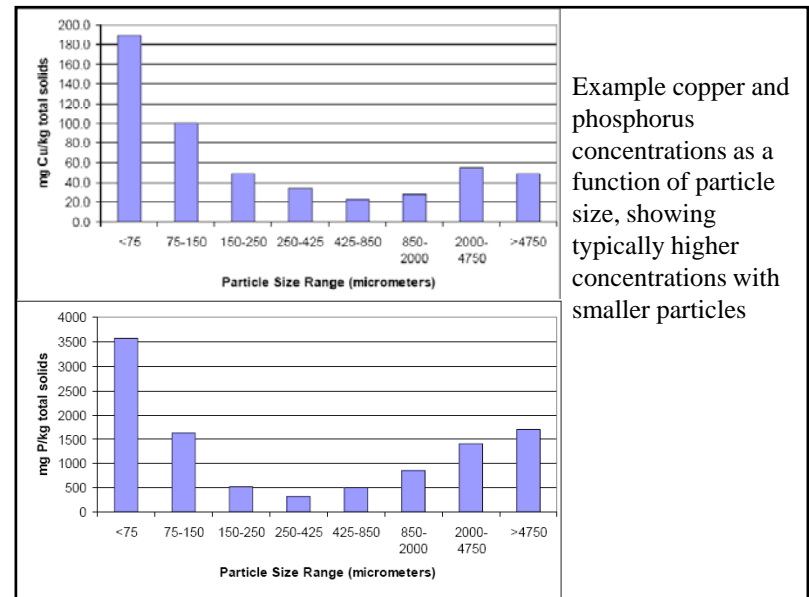
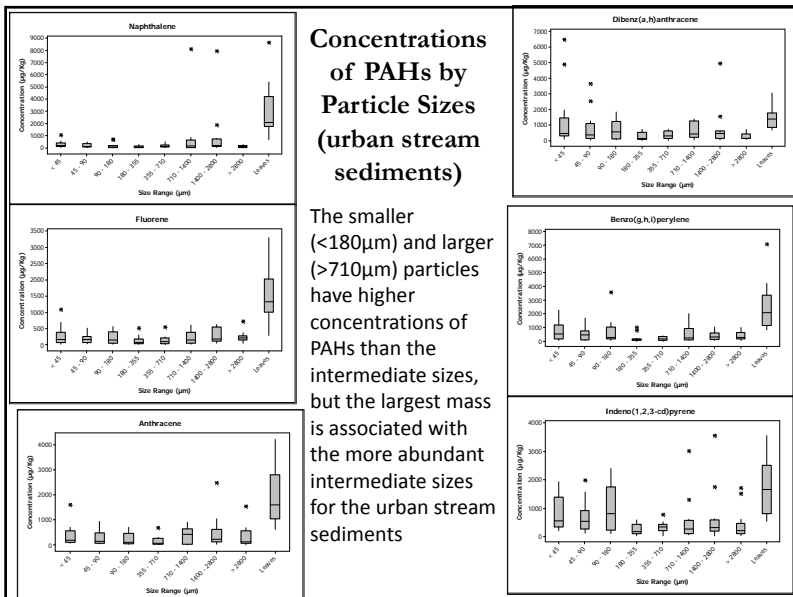
- SSC methodology closely matches known concentrations, regardless of sample concentrations or PSD.

Sample Preparation before Particle Size Association Tests

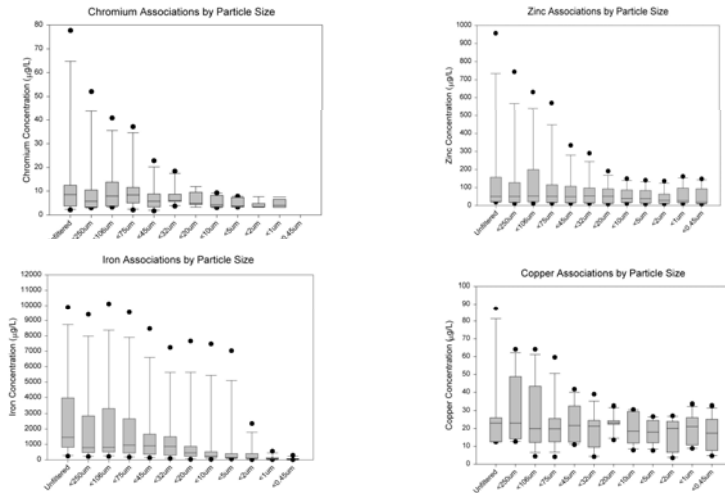
- These tests are used to obtain concentration and particle samples associated with different particle sizes.
- Samples are first split using a cone splitter, and the individual samples are individually separated using a variety of filters and sieves.
- The filtered portion for each separated subsample is then individually analyzed and the associations are determined by difference. Sediment samples can also be analyzed by examining the filters, or by removing some of the captured debris from the sieves.

Large sample volume (about 5 L) separated into subsamples using cone splitter. The sample is first poured through a 1,200 μm screen to remove leaves and grass clippings, and coarse sediment that would clog the splitter. This captured material is also analyzed.



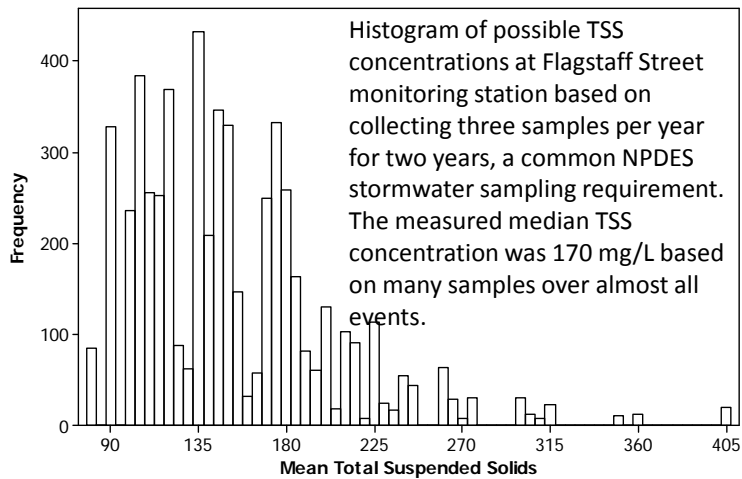


Tuscaloosa, AL, Stormwater Outfall Samples
Residual stormwater concentrations after removal of particles larger than size indicated



Experimental Design – monitoring to consider variability and objectives

- The large variability of stormwater quality requires a major sampling effort to obtain useful data
- Experimental design equations can be used to estimate the number of samples needed to meet the data quality objectives (power analysis)



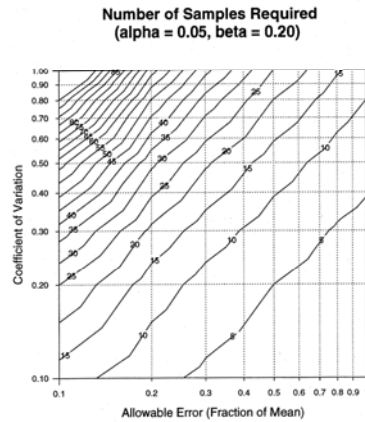
Errors in decision making are usually divided into confidence, or type 1 (alpha) and power, or type 2 (beta) errors:

(alpha) (type 1 error) - a false positive, or assuming something is true when it is actually false. An example would be concluding that a tested water was adversely contaminated, when it actually was clean. The most common value of alpha is 0.05 (accepting a 5% risk of having a type 1 error). Confidence is $1-\alpha$, or the confidence of not having a false positive.

(beta) (type 2 error) - a false negative, or assuming something is false when it is actually true. An example would be concluding that a tested water was clean when it actually was contaminated. If this was an effluent, it would therefore be an illegal discharge with the possible imposition of severe penalties from the regulatory agency. In most statistical tests, beta is usually ignored (if ignored, beta is 0.5). If it is considered, a typical value is 0.2, implying accepting a 20% risk of having a type 2 error. Power is $1-\beta$, or the certainty of not having a false negative.

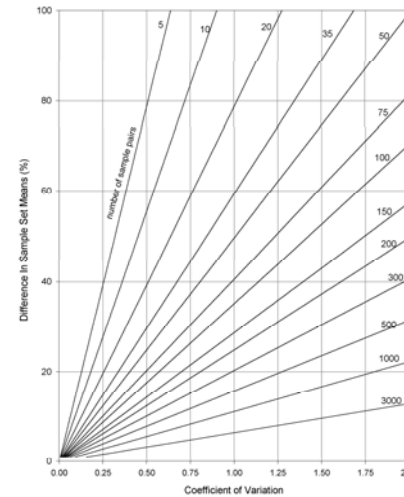
Experimental Design - Number of Samples Needed

The number of samples needed to characterize stormwater conditions for a specific site is dependent on the COV and allowable error. For most constituents and conditions, about 20 to 30 samples may be sufficient for most objectives. Most NPDES Phase 1 sites only have about 10 events, but each stratification category (land use for region of the US) usually has much more.



Burton and Pitt 2002

Number of Sample Pairs Needed
(Power=80% Confidence=95%)



Burton and Pitt 2002

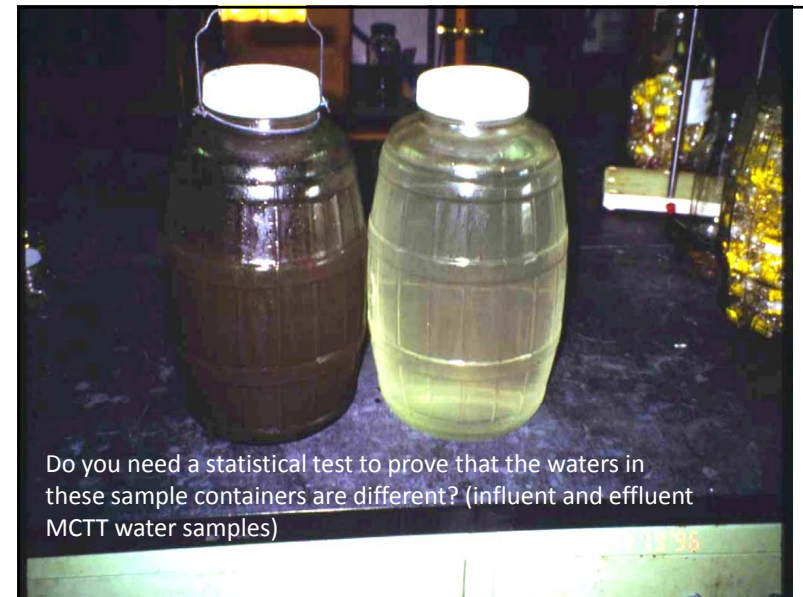
Experimental Design - Number of Samples Needed can be Large

Much information will be needed to confirm performance of stormwater controls for most constituents.

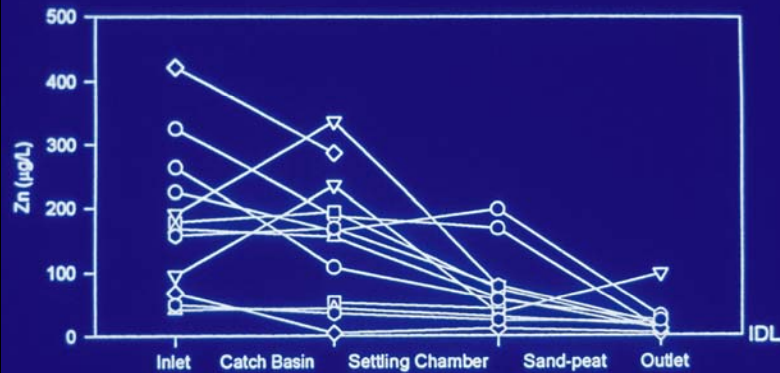
Obviously, easier to confirm removals when the differences between influent and effluent are greatest. Data sets having few samples cannot detect small and moderate differences. A power analysis before the monitoring program needs to be conducted to determine the level of control that can be detected with significance and to ensure that value meets the data quality objectives for the project.

Basic Data Analyses

- The most common goals for stormwater monitoring programs are comparisons (influent vs. effluent), characterization (for different conditions), and model building (relating effluent to influent conditions).
- Simple exploratory data analysis plots are very helpful (scatter plots, line graphs, histograms).
- Probability analyses are very important to compare the data sets directly and to help select the best and correct statistical tests
- ANOVA and residual analyses must be conducted with regression analyses to verify that the test assumptions have been met.

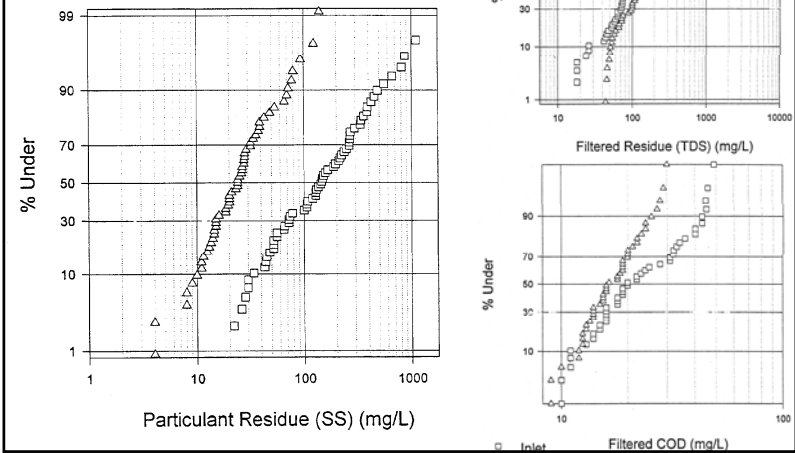


MCTT Pilot-Scale Test Results



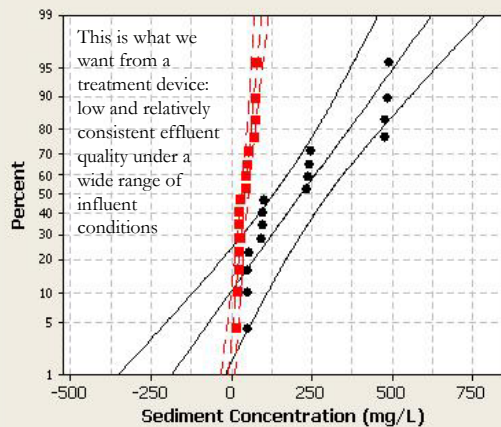
Simple line plots like this can help understand the role of different treatment processes under different conditions

Probability Plots of data from the Monroe St., Madison, WI, wet detention pond monitoring project (7 years duration), USGS and WI DNR



TSS Probability Plot of Influent and Effluent

Normal - 95% CI



This is what we want from a treatment device: low and relatively consistent effluent quality under a wide range of influent conditions

Variable	Mean	StDev	N	AD	P
TSS Influent (mg/L)	218.4	174.0	16	1.236	<0.005
TSS Effluent (mg/L)	42.72	22.38	16	0.932	0.013

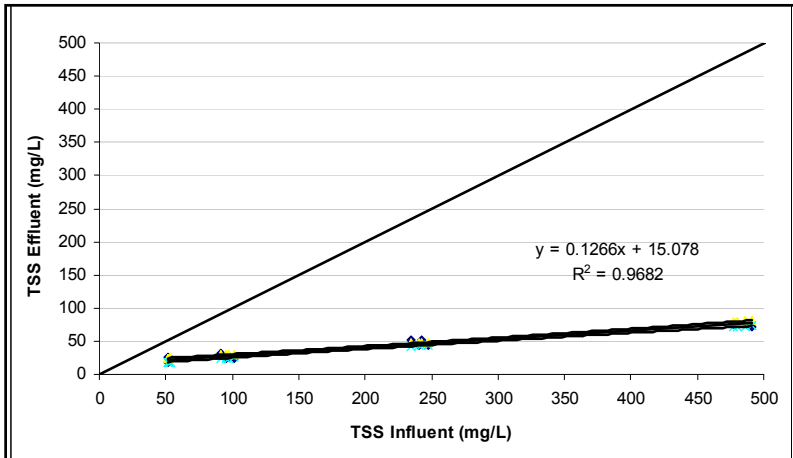
Influent and effluent particulate solids probability plots for current UpFlo™ filter testing. Neither normally distributed ($p < 0.05$) and certainly not the same variances. Therefore requires a suitable non-parametric statistical test to determine if the two sets are from the same population. Certainly doesn't look it, but need the numbers!

Mann-Whitney-Wilcoxon Test

n1	16
n2	16
Observed T	376
Expected T	264
Std. Dev. T	26.533
Test Statistic	4.221
p-Value (lower tail)	0.99999
p-Value (upper tail)	1.22E-05
p-Value (two tail) (Reject H0, if p-Value < 0.05)	2.43E-05
Significant Diff?	Yes

H0: Influent and Effluent Concentration is Same

Ha: Influent and Effluent Concentration is Differ



Scatter plot with calculated regression line and 95% confidence intervals (very narrow CIs because of good fit). Equation needs to be verified with ANOVA and residual analyses.

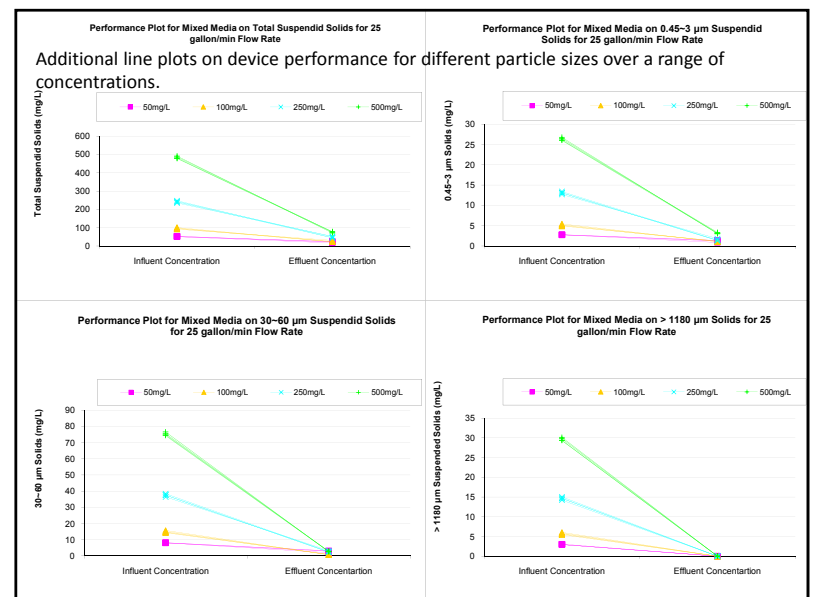
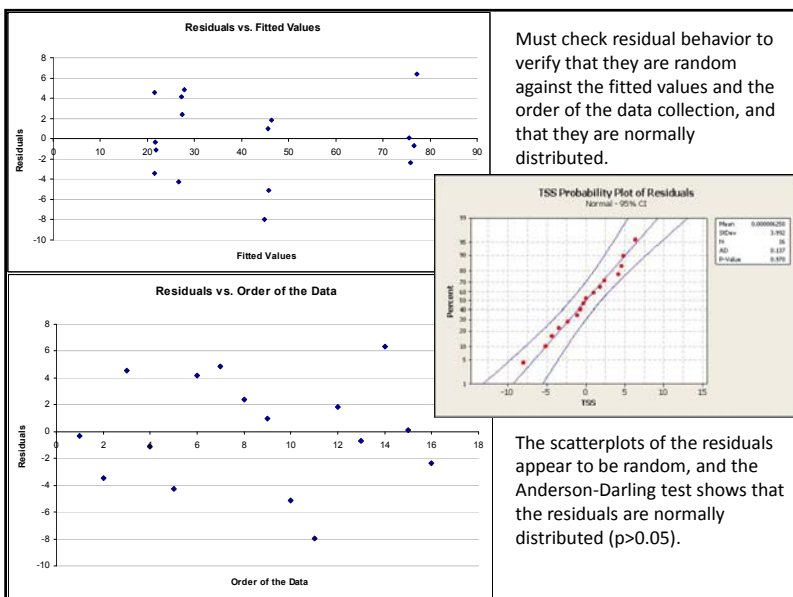
ANOVA to verify that the equation coefficients are significant (if not, remove the offending coefficient and re-analyze) and if the total equation is significant. In this case, both coefficients and the equation are highly significant, with each $p < 0.001$

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.984
R Square	0.968
Adjusted R Square	0.966
Standard Error	4.132
Observations	16

ANOVA						
	df	SS	MS	F	Significance F	
Regression	1	7275.225	7275.225	426.021	0.000	
Residual	14	239.080	17.077			
Total	15	7514.305				

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 90.0%	Upper 90.0%
Intercept	15.078	1.691	8.914	0.000	11.450	18.706	12.099	18.057
X Variable 1	0.127	0.006	20.640	0.000	0.113	0.140	0.116	0.137



Conclusions

There have been many stormwater monitoring strategies used over the years and we have learned a great deal about stormwater characteristics. It is possible to select a suitable approach based on the monitoring objectives, and to understand the limitations of the method.

It is important to examine as many elements of the urban area stormwater pollutant mass balance as possible during monitoring activities to appreciate the component being investigated.

Special sampling and handling is needed to obtain the best particulate solids information.

The study objectives may require a large sampling effort to obtain statistically valid results.

Basic data analyses are easy to perform, but care must be taken to ensure that the methods used are appropriate.