BMP Monitoring: Methods and Evaluations

Betty Rushton, Ph.D.
Southwest Florida Water Management District

Stormwater management for water quality control is a relatively recent technology. Initially urban stormwater was considered a water quantity problem solved by rapidly draining runoff into sewers, ditches or directly into lakes and rivers. By the mid-1970s studies showed over half the pollutant loads entering Florida waters came from non-point sources (stormwater) caused in part by paving, ditching and draining – techniques that increased storm runoff volume and pollutant loads. Regulations were written and by 1984 all new developments in Florida were required to install systems for the management and storage of surface waters. More recently local governments have initiated projects to retrofit existing urban areas. In response to these challenges, engineers have developed a wide variety of methods and devices for treating stormwater. In our studies we have identified over 70 methods and new ones are being marketed all the time. These methods demonstrate various levels of success and data are needed to quantify their pollution removal capabilities. Accurate data are also needed to meet the needs of watershed models, to obtain permits through the National Pollutant Discharge Elimination System (NPDES) program, and to help establish Total Maximum Daily Loads (TMDLs). Increasingly, uniform and reliable measurements are needed to quantify the efforts being made to protect our water resources.

This paper is not a comprehensive description for collecting detailed data for testing BMPs. That is much too large a subject, but it will focus on some of the steps you should take to develop a comprehensive study. First, it is important to read the quality assurance document prepared by your laboratory and the instruction manual that came with the equipment you have purchased. I also strongly recommend that you read the detailed manual developed by members of the Urban Water Resource Research Council of the American Society of Civil Engineers (ASCE) and the Environmental Protection Agency EPA (ASCE/EPA 2002). If you collect data using these guidelines it can become part of a valuable international data base comparing stormwater techniques. The manual and several other useful documents are available on the internet at:

www.bmpdatabase.org/docs.html/.

The purpose of this paper is to share some of the experiences encountered during our thirteen years of monitoring stormwater Best Management Practices (BMPs). It will focus on pitfalls encountered in conducting monitoring studies and try to provide some useful tips and techniques, especially those that are not a part of most manuals. Water hydrology will be discussed first, followed by water quality considerations and finally methods to evaluate data. Most of the ideas are shown in figures and tables with only brief discussions in the text, so that if you only look at the figures you can get a pretty good idea about the concepts being presented.

Measuring Hydrology

Calculating flow volume - Accurate measurements of flow volume are essential to nearly all aspects of water pollution control. For most stormwater flow applications, measurements are made in open channels either with structures (weirs, flumes, etc.) or within pipes by velocity meters. A useful document for determining the best type of primary measuring device for your application has been published by ISCO (Grant and Dawson 1995). For structures, years of testing by engineers have developed relationships (usually nonlinear) with water level measurement (head) and flow rates. These standardized formulas are available in most engineering hydrology textbooks such as Brater *et al.* (1982). In contrast, velocity meters determine flow by measuring velocity across a cross-sectional area at a specific point. Several of our site set-ups are shown in Figures 1a-1c.

Whether you use weir structures, flumes, pipe flow or velocity meters, they all require an accurate measurement of water level by a separate measuring device usually purchased from a vendor. Some typical instruments to measure water depth are listed in Table 1.

Table 1. Equipment for measuring depth of flow (ASCE/EPA 2002)

Method	Major Requirements
Float Gauge	Open channel flow with stilling well
Bubbler Tube	Open channel flow with no velocities greater than 5 ft/sec
Pressure Transducer	Better measurements if transducer stays submerged
Ultrasonic Sensor	Open channel flow with no wind, loud noises, turbulence, foam steam, or floating oil & grease
Ultrasonic uplook	Sediments or obstructions are likely to cause errors in measurements
Pressure Probe	Open channel flow with no organic solvents or inorganic acids
Acoustic Path	Large channels with base flow

Keep in mind that all of these sensors are only sending an electrical impulse and it is up to you to make these measurements accurate. Therefore, these high tech devices need frequent comparisons to an independent measuring device such as a rigid permanently installed staff gauge. *Actually your staff gauge and your field journal are the two most* *important pieces of equipment you have at your study site.* All of your other measurements should be calibrated against the staff gauge including the control elevation of your structure (level where water first goes over or through the structure). The level reading from your data logger or flow meter should be checked against the staff gauge reading during each site visit. It doesn't even really matter if the staff gauge is accurate in relation to the NGVD for the site since all of your measurements are relative to the control elevation of your structure and not the actual land elevation in relation to sea level.

<u>Accurate water level measurements</u> – To check the accuracy of your water level measurements and to assure that your level sensor has not drifted over time, it is a good idea to periodically check the level measurement with the staff gauge readings by developing a regression equation. An example is shown in Figure 2.

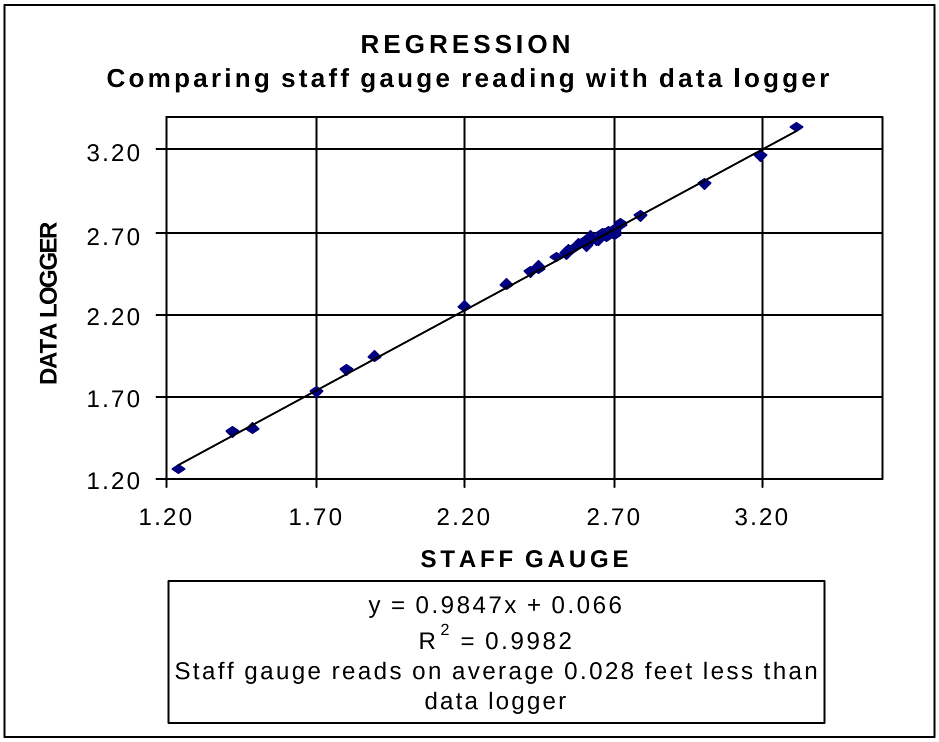

Figure 2. A regression equation is calculated between the staff gauge and the data logger readings.

The regression graph indicates that the staff gauge readings are less than the recorded level in the data logger. It now is the responsibility of the researcher to determine if this small one-third inch difference in level will skew his results. An example of how a small difference in level can result in a large difference in volume is demonstrated in Figure 3. It also shows how differences in sensors can also affect measurements. In this case the top of the hydrograph was flattened when using a float gauge water level sensor. This small one-half inch difference at the top of the hydrograph reduced the volume of the calculated flow from 62,000 cu ft. to 48,000 cu ft.

A good reference for collecting and reporting rainfall and flow data has been published by the United States Geological survey (Church et al. 1999). Another useful reference has been published by the Environmental Protection Agency (2001).

<u>What if a standard measuring device doesn't exist</u>? - You can still take measurements even if no standard weir exists, but standard weirs and flumes are still best. You can also check the accuracy of sub-standard primary measuring devices as well as investigating other processes taking place in the pond using some of the following techniques.

- A portable velocity meter and a section of channel upstream or downstream of the pond you are attempting to evaluate can be used to develop a stage discharge curve (Figure 4) (Rushton and Hastings 2001).
- A bucket and stop watch can measure low flows at different water levels to develop a relationship between level and flow (Figure 5)(Rushton *et al*. 1997).
- A dilution method can measure flow by adding a tracer solution. Radioactive and fluorescent dye tracers are most often used and flow rate is calculated by theoretical formulas. A fluorometer or other appropriate instrument is required for determining sample concentrations and complete vertical and lateral mixing is assumed. (Grant and Dawson 1995 and others).
- The Manning formula can be used to estimate flows in a channel or pipe. It requires knowledge of the channel cross section, liquid depth, slope of the water surface. And a roughness factor (Grant and Dawson 1995 and others).
- A variable gate method can be used to measure fluctuating flows in small pipes (Grant and Dawson 1995).
- A complete water budget of the pond can be calculated if one of the flows in the system has to be estimated. This uses the following relationship of all the inflows and outflows:

  Precipitation + inflow – outflow – ET +/- seepage = change in storage

  Some of the data for this type of estimate are presented in Table 2 (Rushton 2001).
- A micro scale evaluation of water levels in a pond or lake can often shed additional insight into processes taking place (Figure 6). If conditions are right, it is also a good method for estimating evapotranspiration and net seepage by comparing levels between night when very little evapotranspiration takes place and during the day. The water level measurements need to have been taken to four decimal places to accurately detect the small differences in water levels during a 24-hour period (Carr and Rushton 1995).
- If at all possible have a backup for equipment that can work off of a battery and provide a hard copy in case of electrical failure. We have had good luck with ISCO flow meters and some of these have been in operation at different sites for ten years. The new ones can give water level, flow and rainfall and many study sites use only this one instrument.
- Rainfall or models can be used to predict flow (Church et al. 1999).
- A number of passive samplers that can be used to collect flow-weighted water quality samples are described by USGS with useful diagrams (Bent et al 2001).

What can you do with the data ? – The type of monitoring equipment in use today allows for a tremendous amount of data to be automatically collected. The timing, frequency and duration of flow measurements are critical factors in monitoring accurate flows in small urban drains because of the rapid response to stormwater runoff and the wide ranges of flow over short periods of time. Church *et al* (1999) present a good discussion of these variables, which may be useful to you in programming your data loggers. At our sites, we measure hydrology data every minute and the logger averages it at fifteen-minute-intervals and can also record maximum and minimums during the interval. Even with this amount of summarization, we still have enormous, often unwieldy, files. Some researchers collect data every five minutes, but only collect it during rain events. Depending on the goal of the project, hydrology data can be either summarized by storm event or the entire record can be shown. Since we usually present all the hydrology data in figures in an appendix, we have found it is useful to develop a template to use for downloading the data. Using this method all figures follow a standardized format and are ready to go into a report without much more editing (Figure 7).

Often data, especially from velocity sensors, need to be smoothed. One method for smoothing data is using the running average technique available in most spreadsheet programs (Figure 8). Here we used a two-hour running average.

Figures developed from a spreadsheet program are also helpful for making corrections caused by equipment malfunction. Two examples are shown in Figure 9. This is also the place to make corrections to water level elevations that may have been detected such as those suggested by Figure 2.

Collecting Water Quality Data

Selecting the constituents to test – The constituents most often measured in stormwater studies include: ammonia, nitrite, nitrate, total nitrogen (organic nitrogen calculated), ortho, soluble or reactive phosphorus, total phosphorus, iron, copper, cadmium, zinc, total suspended solids (TSS), biological oxygen demand (BOD), chemical oxygen demand (COD), total organic carbon (TOC) and rarely polycyclic aromatic hydrocarbons (PAHs) and bacteria. Some of the problems associated with the various constituents include:

- Nitrogen recycles rapidly and collecting accurate measurements includes keeping temperature at 4 degrees C.
- BOD is a holdover from constituents measured for sewage treatment plants and COD is a better metric for evaluating stormwater.
- BOD, COD and TOC samples cannot be collected with automatic samplers if the procedures in Standard Methods used by most laboratories are followed.
- Total metals (usually measured) is not considered a good test by many people since only the soluble portion is available for use by organisms.
- The state standard for total metals requires that hardness also be measured since it is used in the calculation for the standard. The softer the water the more toxic the metal.

- It costs about five to six times more per sample to test for PAHs in water and concentrations are low enough that they are rarely detected. In our studies we only test for PAHs in the sediments and standards are being develop to evaluate these.
- The method for determining TSS was originally designed for the analysis of wastewater samples and is shown to be fundamentally unreliable for the analysis of natural-water samples (Gray et al 2000, Bent et al. 2001). Instead it is recommended that suspended-sediment concentrations (SSC) be used. Other researchers have also found problems with using TSS for analyzing pollutant loads (James1999). This is important because often TSS is the only constituent measured since many people believe it is indicative of other pollutants.
- Bacteria and viruses are rarely sampled in stormwater studies, but a number of microbial pathogens can be present in stormwater and have been implicated in water borne disease outbreaks in both humans and fish (Kurz 1998). The fecal coliform test used for wastewater is not comprehensive enough for stormwater studies and new metrics are being developed.
- Trash, litter and other debris are not measured in most stormwater studies and yet this is one of the most polluting aspects of storm runoff. These pollutants are removed by street sweeping, drop box inserts, swirl concentrators and baffle boxes and a standardized method for measurement needs to be developed. These pollutants are often discharged unmeasured from stormwater ponds and other treatment alternatives since they don't fit in our sample bottles.

Quality assurance testing - In addition to the quality assurance tests run by the laboratory, you also need to take samples to check your field equipment and sampling techniques. Methods for doing this include taking water quality samples using deionized water through the equipment. Our samples indicated that the clear plastic tubing should be placed in PVC pipe to shield it from the sun, since sunlight encourages algae growth in the tubing. Testing also indicated that changing the tubing about every six months is often enough under normal conditions. Table 3 and Table 4 show some of the results from testing automatic refrigerated samplers used for taking flow-weighted water quality samples.

Why flow-weighted samples are best – Flow-weighted samples take more individual aliquots with higher flows than with lower flows. In Figure 10, An example of three hydrographs, shown on the vertical axis, are compared to each individual aliquot of sample, horizontal lines. The data were copied from an ISCO strip chart and demonstrate how more samples are taken during high flows and fewer samples during low flow. The chart also shows the difficulty in estimating the right amount of sample to take with each aliquot since the bottle filled up before the end of the large storm.

Figure 11 shows some of the variation in water quality that occurs during different stages of the hydrograph and demonstrates the value of taking composite samples to collect representative samples for storm events. It also shows why flow-weighted samples give better estimates than timed samples. Although timed-samples are better than grab samples as will be discussed later. Figure 11 also compares two types of basins, one measured runoff directly from an asphalt parking lot while the other measured runoff after it had traveled through a small vegetated swale. Although concentrations were often higher in the basin with a swale, pollutant loads that included the volume of water discharged was less in the basin with a swale because of the opportunity to infiltrate into the soils.

How long can samples be left in collection bottles? – It is not always convenient to be at the site immediately after a storm event. Also, as often happens at an outflow station, it takes several days before the pond stops discharging. For example, our District rules have a five-day draw down period for wet-detention ponds. To test whether leaving samples in the refrigerator for three days without preservatives affected sample results, statistical tests were run. Duplicate samples were collected within 24-hours after a rain event, fixed with preservatives and taken to the lab to test our sampling technique. No significant differences were detected in the duplicate samples (Mann-Whitney alpha=0.80 and above for most constituents). Although there were no significant differences (alpha=0.10) for samples stored in the refrigerator or even in styrofoam coolers without ice, the probabilities were much greater for errors indicating that we might not be detecting a difference when there actually was a difference. The results are shown in Table 5 and Figures 12 and 13. Although samples held three days in the refrigerator may not be significantly different, they showed a greater probability than the duplicate samples, and samples not refrigerated showed an even greater probability that they were different. The results indicate it is best to collect and preserve samples right away. Although it is possible to pre-acidify sample bottles, we discovered the following problems: the preservatives evaporated, it required washing a lot of bottles with each storm event, and it was much more difficult to estimate the aliquot that should be collected to sample both small and large storm events.

Table 5. Probability that there is not a significant difference between samples held for three days without preservatives either in refrigerators (n=7) or in a dark cooler with no ice (n=10).

	Electric Refrigeration		No Ice 3days		Duplicates
Constituent	Sign Test	Mann Whitney	Sign Test	Mann Whitney	Mann Whitney
Ammonia (mg/l)	1.000	0.70	0.754	0.68	0.85
Nitrate (mg/l)	0.453	0.70	0.343	0.57	0.67
Total Nitrogen (mg/l)	1.000	0.60	0.343	0.35	0.85
Ortho-Phosphorus (mg/l)	0.453	0.61	0.109	0.47	0.88
Total Phosphorus (mg/l)	1.000	0.89	0.109	0.67	1.00
Turbidity (NTU)	0.219	0.65	na	na	na
Total Iron (ug/l)	0.375	0.94	0.754	1.00	1.00
Total Copper (ug/l)	0.219	0.70	1.000	0.96	0.82

Can grab samples be compared to composite samples? We were interested in finding out if grab samples taken immediately after storm events were comparable to flow weighted composite samples taken for the same storm. Since the pond should be well mixed by the time water is discharged, we thought the samples might not be significantly different at the outflow. Grab samples were compared to the results of composite samples for 21 rain events at a Florida Aquarium stormwater pond (Figure 14). Almost all samples measured significantly higher concentrations at the outflow with the grab samples (alpha=0.10). The exceptions were ammonia and total nitrogen, which showed no significant differences.

In contrast, grab samples at the inflow of the wet-detention pond measured significantly lower concentrations compared to composite samples for the six storms sampled (Figure 15). Once again the exceptions were ammonia and nitrogen, but this time ortho-phosphorus was also not significantly different (alpha=0.10). Nitrates at both the inflow and outflow measured consistently higher concentrations for the composite sample than for the grab sample, indicating the rapid recycling of nitrate into organic nitrogen or ammonia.

These results can be at least partially explained by the “first flush” effect described in Figure 11. Once storm water displaces the treated water in the pond, the early samples at the outflow represent cleaner water and only later does the untreated stormwater travel to the outflow in time to be a major part of the grab sample. In contrast, the grab sample taken at the inflow represents runoff from later in the storm. These results emphasize the importance of hydraulic residence time (the average amount of time stormwater remains in the pond before discharge downstream) for removing pollutants between storm events.

Fluctuations in Field Parameters - Field conditions such as pH, dissolved oxygen, temperature and specific conductivity affect the chemical and biological cycling and transformations that take place in a pond. For example, phosphorus and metals are released back out of the sediments and into the water column at low dissolved oxygen concentrations. The effects of daily fluctuations and the effect of rainfall on these parameters are shown in Figure 16. The diurnal cycle that reflects the daily progress of sunlight provides insight into processes taking place in the pond. Eutrophic ponds with a lot of phytoplankton in the water column have greater diurnal fluctuations than oligotrophic ponds. Conditions in summer usually create more fluctuations than in winter because more biological activity is taking place. Rainfall moderates the fluctuations at both the inflow and outflow and causes a reduction in concentrations. This is partially explained by the dilution effect of rain and is most obvious at the outflow station. This particular pond receives about 25 percent of its hydrologic input from rainfall on the pond.

Servicing the Research Site

If the research project is at all complicated, and most of them are, it is good idea to have the person(s) responsible for collecting data make an outline and schedule for the various tasks that must be completed. An excellent outline was developed for the two research sites at the Florida Aquarium by Rebecca Hastings where multiple samples (15-20) had to be collected for each storm event (Appendix A). The outline also made it possible for interns and student co-ops to service the site.

Evaluating Results

Calculating BMP efficiency - Best Management Practices (BMPs) with well defined inlets and outlets are usually evaluated using efficiency (% removal of pollutants from the inflow to the outflow). One problem is the variety of methods for calculating efficiency. A good evaluation of the various methods can be downloaded from the internet under the title "Development of BMP Performance Measures" (ASCE/EPA 2000):

www.bmpdatabase.org/docs.html

The methods most often used include:

- Efficiency ratio (ER)

$$\text{ER }(\%) = 1 - \frac{\text{average outlet EMC}}{\text{average inlet EMC}}$$

Event Mean Concentration's (EMCs) can be either collected as flow weighted composite samples in the field or calculated from discrete samples on a flow-weighted basis

- Summation of Loads (SOL)

$$\text{SOL }(\%) = 1 - \frac{\text{SOL outlet}}{\text{SOL inlet}}$$

Where: SOL = the sum of loads in cubic feet or cubic meters for all storms for a predefined period (usually a year)

SOL in = sum of loads at the inflow plus rain falling directly on the pond
SOL out = sum of loads at the outlet

Loads are calculated by multiplying the EMC by volume and converting to mass.

- Regression of Loads (ROL)

The regression of loads method developed by Martin and Smoot (1986) defines regression efficiency as the slope of a least squares linear regression of inlet loads and outlet loads of pollutants, with the intercept constrained to zero.

$$\text{ROL (\%)} = \text{slope} - \frac{\text{loads out}}{\text{loads in}}$$

Loads are calculated as in the SOL method

- Individual Storm Loads (ISL)

  Calculates an efficiency for each storm event based on the loads in and the loads out as calculated in SOL. The efficiency of the BMP for a single storm is given by:

  $$\text{ISL(\%)} = 1 - \frac{\text{load out}}{\text{load in}} \quad \text{Average efficiency} = \frac{\text{sum of individual ISL}}{\text{number of storm}}$$

The ASCE/EPA report (2000) used our data for the Tampa Office Pond to calculate the efficiencies of these various methods (Table 6). In the Tampa Office study we re-designed the same pond using three different criteria and then studied each design for eight months. In 1990 the pond had a 2.5-day hydraulic residence time (HRT), in 1994 the HRT was increased to 5 days and in 1995 the HRT was 14 days.

Table 6. Comparison of methods used for calculating removal efficiency (%) for Total Suspended Solids.

Year	ER	SOL	ROL	ISL
1990	0.59	0.71	0.79	0.29
1994	0.64	0.66	0.82	-0.02
1995	0.95	0.94	0.95	0.89

The data in Table 6 not only demonstrates differences between calculation methods, but shows the improvement in pond performance as hydraulic residence time was increased. Since more samples (42) were collected in 1995 compared to 22 storms sampled in 1990 and 27 storms in 1994, the percent removal by each method is much more consistent for 1995. This emphasizes the need for an adequate sample size for producing meaningful results.

Comparing data to State Standards - Often water coming into a stormwater pond is clean enough that further reductions in concentration are not possible and efficiency calculations are meaningless. For example, total suspended solids cannot usually be reduced below 10 mg/l and often, especially if storm runoff has been pre-treated by ditches or swales, it is already reduced to this level by the time it reaches the inflow. Another method for evaluating stormwater is to compare concentrations to State Water Quality Standards. State standards can be downloaded on the internets using the following address:

http://www.dep.state.fl.us/water/rules/62-302t.pdf

This method is also useful if you are unable to collect flow-weighted samples and the only alternative is to collect grab samples after storm events. We have found this method effective in our surveys for determining if stormwater systems permitted by the District are meeting the presumption of our rule that assumes systems built according to our rules meet State Standards. However, remember that grab samples taken after storm events at the outflow have significantly higher concentrations than flow-weighted samples (see Figure 14). We did find that grab samples taken in the surveys we conducted had more exceedances of standards than the composite samples from our detailed studies.

Acknowledgement

Special appreciation is due Rebecca Hastings for patiently collecting all of the samples for our quality assurance tests and for writing detailed instructions for site visits. Also thanks to David Carr for the study testing water quality before and after changing tubing in the automatic samplers. As always, I am grateful for the great work done by our laboratory staff who carefully analyzed our numerous water quality samples that cannot be collected according to any pre-determined schedule

Disclaimer

Although I have used ISCO equipment as examples in the paper, there are many other good products on the market and a researcher should consider many vendors and alternatives before making a selection. Also, be sure to consider the kind of support you will be receiving after the equipment is installed and paid for.

References

ASCE/EPA Determining Urban Stormwater Best Management Practice Removal Efficiencies. 2000. www.bmpdatabase.org/docs.html

ASCE/EP A guidance Manual for Best Management Practices (BMP) Monitoring. 2002. www.bmpdatabase.org/docs.html/.

Bent, G. C., J. r. Gray, K. P. Smithe and G. D. Glysson. 2001 A Synopsis of Technical Issues for Monitoring Sediment in Highway and Urban Runoff. United States Geological Survey, Open-File Report 00-497, Northborough, Massachusetts from USGS or from: www.stormwater-resources.com/Library/133PLUSGSTraceMonitoring.pdf

Brater, E. F., W. King, J. E. Lindell and C. Y. Wei. 1982. Handbook of Hydraulics. McGraw-Hill Companies Inc. New York, NY.

Carr, D. W. and B. T. Rushton. 1995. Integrating a Native Herbaceous Wetland into Stormwater Management. Southwest Florida Water Management District, 2379 Broad Street, Brooksville, FL 34604

Church, P. E., G. E. Granato, and D. W. Owens. 1999. Basic Requirements for Collecting, Documenting, and Reporting Precipitation and Stormwater-Flow Measurements. United States Geological Survey. Open-File Report 99-255. Northborough, Massachusetts. Available from USGS or from the web site below.
www.stormwter-resources.com/library/136PLUSGSMonitoring.pdf

Environmental Protection Agency (EPA). 2001. Techniques for Tracking, Evaluating, and Reporting the Implementation of Nonpoint Source Control Measures. III Urban. EPA Office of Water. EPA 841-B-00-007.

Grant, D. M. and B. D. Dawson. 1995. Isco Open Channel Flow Measurement Handbook. P.O. Box 82531, Lincoln, NE 68501.

Gray, J. R. , G. D. Glysson, L. M. Turcios, and G. E. Schwarz. 2000. Comparability of Suspended-Sediment Concentration and Total suspended Solids Data. United States Geological Survey. Water-Resources Investigations Report 00-4191.

James, R. B. 1999. Solids in Storm water Runoff.
www.stormwater-resources.com/library/070BmeasuringSolids.doc

Kurz, R. C. 1998. Removing Microbial Indicators from Stormwater Using Sand Filtration, Wet Detention, and Alum Treatment Best Management Practices. Southwest Florida Water Management District, 2379 Broad Street, Brooksville, FL 34604

Martin, E. H., and J. L. Smoot. 1986. Constituent-Load Changes in Urban Storm Water runoff routed through a Detention Pond-Wetland System in Central Florida. Water Resource Investigation Report 85-4310. United States Geological Survey.

Rushton, B. T. and R. Hastings. 2001. Florida aquarium Parking Lot: A Treatment Train apporach for Stormwater Management. Southwest Florida Water Management District, 2379 Broad Street, Brooksville, FL 34604.

Rushton, B. T., C. Miller, C. Hull and J. Cunningham. 1997. Three Design alternatives for Stormwater Detention Ponds. Southwest Florida Water Management District, 2379 Broad Street, Brooksville, FL 34604

Rushton, B. T. 2001. Treatment of Stormwater Runoff from an Agricultural Basin by a Wet-Detention Pond in Ruskin, Florida. Southwest Florida Water Management District, 2379 Broad Street, Brooksville, FL 34604