CONNECTING LAB TESTING TO FIELD TESTING
Can the Advantages of Each Method Be Combined?

John Moll, CEO CrystalStream Technologies

Chief Executive Officer, CrystalStream Technologies, 2424 Lance Court, Loganville, GA, 30052, USA; johnmoll@crystalstream.com

ABSTRACT

The problems with field testing are at best daunting, and at worst insurmountable. Can an accurate sample be taken from a pipe or other conveyance when it is under flow? "Accurate" means that a sample represents the entire water column top to bottom and side to side. If you can get such a sample, does it accurately depict all the flow that occurs during the slice of time that the sample represents? If it does, can you pair it with a sample taken below a BMP that accurately represents the "same" water after treatment? How many sample pairs do you need to accurately represent a five hour storm?

The problems with laboratory testing are just as serious as the problems with field testing, but in a different way. In a lab it is possible to accurately sample using calibration methods such as mass balance to establish proper "under flow" sampling points and techniques. It is easy to provide accurate flow measurements and provide continuous pair testing. Exact particle size definitions of the source material can be measured, as can the captured and "missed" or re-suspended material. Unfortunately, no natural storm has continuous, steady flows. Source areas have different particle size distributions at different times, and they differ from each other based on the size of the basin and its use. There are periods of time with no flow, where material can build up on a site, and periods of time within a manufactured treatment device with no flow, where numerous chemical and physical processes have effects on the material. The material itself is not a nice clean batch of sediment, but is a soup of organics, chemicals, natural and man- made sediments, grit, trash and litter. A laboratory test cannot duplicate these conditions.

In short, natural storms have real conditions and materials we need to capture and understand, but which cannot be accurately sampled. Laboratory "storms" are more easily sampled but cannot duplicate real flow conditions and real materials.

This paper will attempt to tie the two methods together by using an actual field study and certified laboratory performance curves for a specific manufactured device and will report the following:

1. Using a "Standard" rainfall hydrograph, what removal rate would be predicted for two very different storms by comparing the storm flows to performance curves?
2. Using measured field results for actual storms, how does the observed removal rate compare to the rate predicted by the performance curves when the field data logs are examined for flow rates?
3. What observations can be made from the data that will help models better predict real storm characteristics versus SCS and Rational Hydrograph generalizations.
4. Can lab performance curves be used to reasonably predict field performance?

DATA AND DOCUMENTS UTILIZED

The performance curves used in this paper were derived from laboratory testing on a CrystalStream Model 956 that was conducted at Alden Laboratories in 2006. The removal efficiency rates used herein were taken directly from their report[1]. Field testing data were taken from the U.S. Environmental Protection Agency's (USEPA) "Environmental Technology Verification" (ETV) report[2] which was based on testing conducted in Griffin, Georgia. The "Standard" hydrographs were developed using standard engineering methods. The rational hydrograph utilized was the "Dekalb Rational Hydrograph", which is in common use in the area where the field testing was done. The hydrograph was developed to have a peak flow of 1.78 cfs using a 5-minute time of concentration, which equals the "1.2 Inch First Flush" peak flow as calculated in the "Georgia Stormwater Management Manual"[3]. The "SCS" hydrograph was also developed from the Georgia manual, once again calculated to produce a peak flow of 1.78 cfs. These two methods of developing a "standardized" rainfall curve are the most widely used by hydrologists and differ greatly in their characteristics. The rational method spreads the storm over a time period that is 10 times the time of concentration for the basin, where the "SCS" method spreads the storm over a much longer period, up to 24 hours.

METHODOLOGY

The performance curves shown below (Figures 1 and 2) were developed by Alden Labs and came from 6 data points established by the rate of flow through the device. The material injected into the device as OK-110 with an average diameter (D_{50}) of about 90 microns. Alden carefully calibrated their influent and effluent testing ports to reflect the actual concentration of materials in the water column and then took samples of the device under flow to determine removal rates. This testing was labeled "Indirect" sampling. After the "Indirect" sampling was completed, Alden drained the device, and directly measured what the device had trapped. The amount trapped was compared to the amount injected, and the results were labeled "Direct" sampling. The similarity of the results from the two methods verified the calibration of the "Indirect" sampling method.

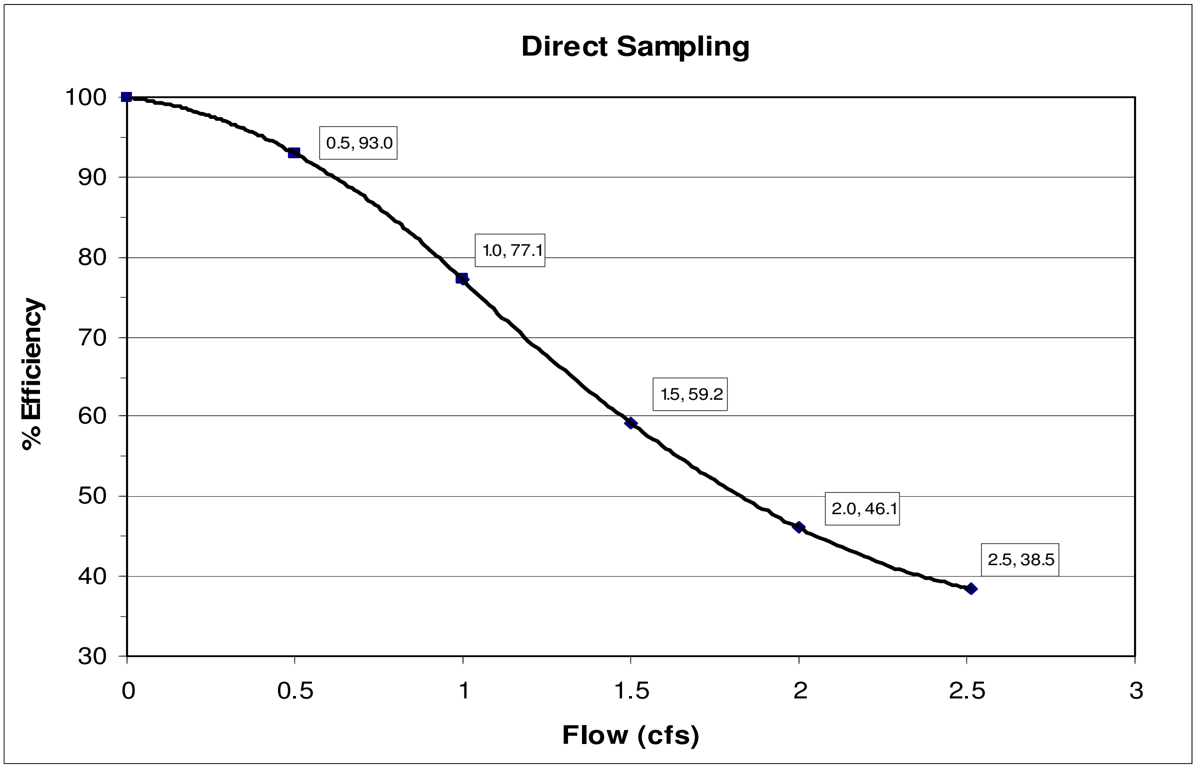

Figure 1

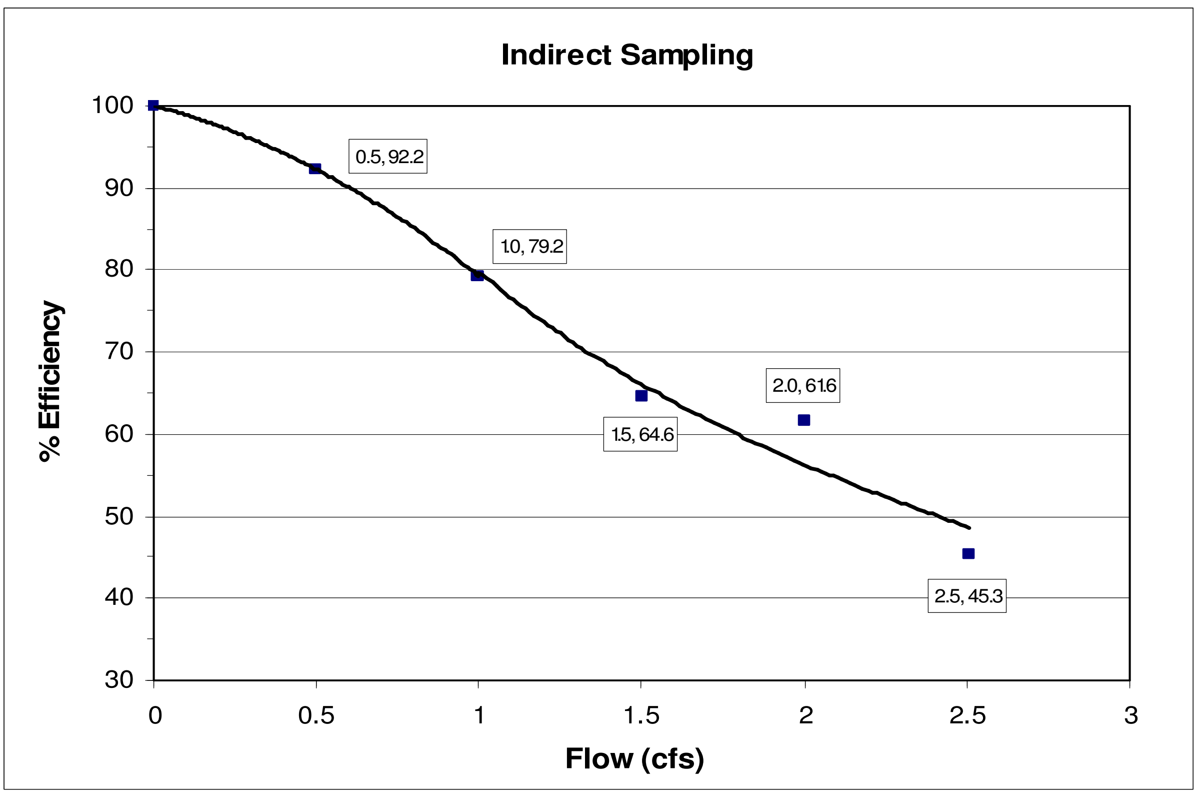

Figure 2

The 6th data point for each method was not included in the graphs above. At 5 cfs, the "Indirect" removal efficiency was 27.4%, and the "Direct" removal efficiency was 19.4%. These values were used to extend the performance curves to higher flows. Neither curve conformed to an easily derived low order equation, so straight line interpolation was used to develop interim values for flows from 0 to 5.0 cfs at 0.10 cfs intervals. A decision was made to use the indirect curve for this paper, because of the close correlation and the fact that the field sampling was indirect sampling.

Because of the similar geometry of the Model 956 device that was lab tested and the Model 1056 that was field tested, the performance curves were assumed to be similar. The hydraulic loading rate (HLR) of the field-tested device was lower than the lab tested device by about 10%, so the field device might be expected to have slightly better performance. The variation between the two devices was that the field-tested Model 1056 had an exit chamber 1 foot larger (3.5 feet versus 2.5 feet for the Model 956). The larger exit chamber provides a larger nutrient screening surface (coconut fiber filter), and less turbulence at high flows. The HLR is lower because the Model 1056 has 50 square feet of surface area, versus 45 square feet of surface area for the lab tested Model 956. Otherwise, the vaults are identical.

When "average" efficiencies were calculated, they were flow weighted. An efficiency of 70% on a flow of 2.0 cfs would count 20 times as much as an efficiency of 90% on a flow of 0.1 cfs. The average of those two efficiencies would be 70.95%, not 80%. Throughout this document, all average efficiencies are flow weighted. During the actual field events evaluated, there were periods of zero flow, which show a theoretical removal efficiency of 100%, but the flow weighted efficiency calculated was zero, because of the weighted calculation. Average intensities were calculated for the field study, but only actual rainfall intensities were utilized in our evaluation, as based on real time rainfall data recorded during the event.

RATIONAL HYDROGRAPH EVALUATION

The rational hydrograph and the removal efficiencies calculated during this idealized storm are shown in Figure 3. The average removal efficiency calculated from the performance curves is 82.5%. As seen in the chart, removal efficiencies remained above average for about the first 20 minutes and the last twenty minutes of this storm and dipped down to about 62% at the peak flow of 1.78 cfs. To chart these two values together, the peak flow of 1.78 cfs is shown at the 100 percent mark. Other interim flows can be calculated by reading the percentage on the "Efficiency" axis and multiplying the percentage as a decimal times 1.78 cfs.

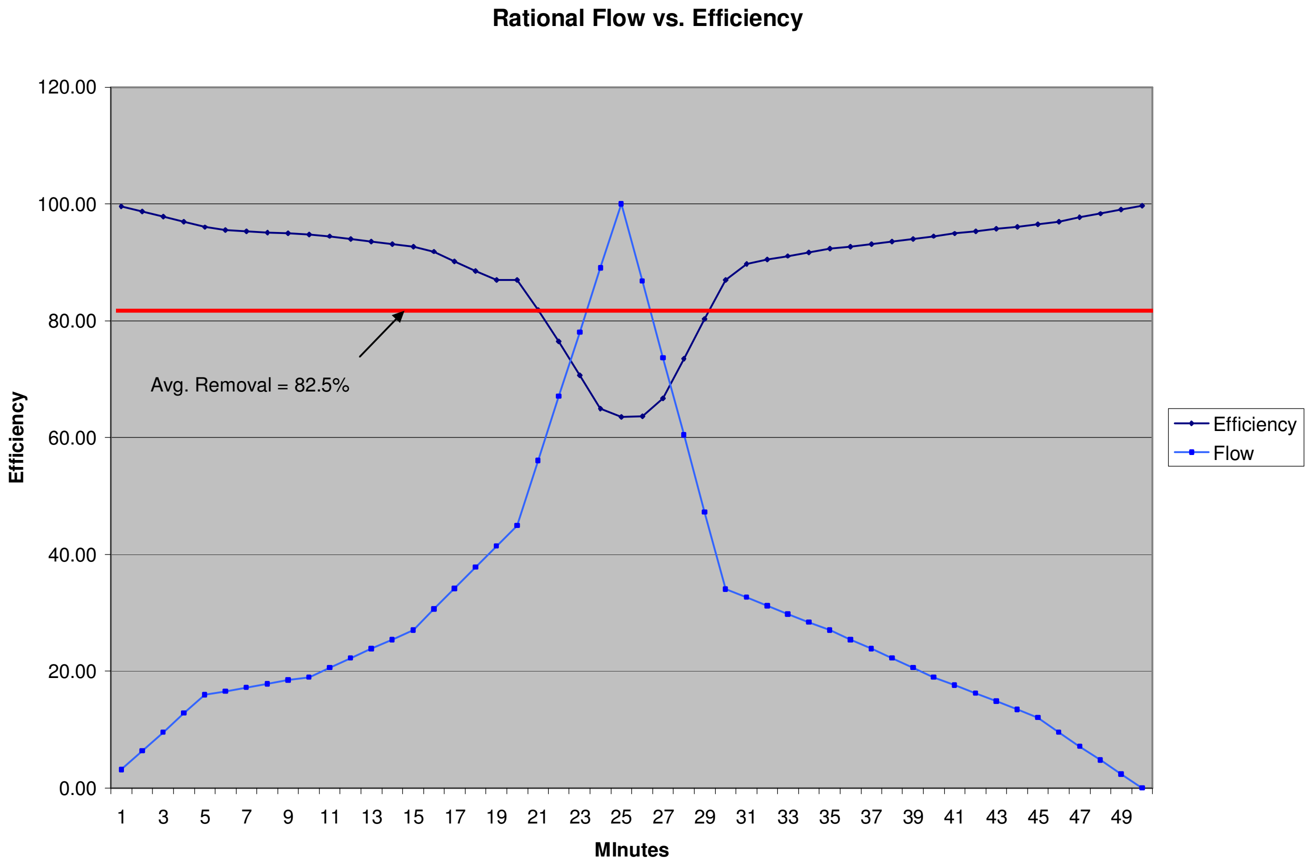

Figure 3

This type of storm represents one of the most aggressive ways to display an event with a peak of this size. These hydrographs are used typically to size pipes and detention facilities for the "worst case" scenario. It should be noted that, while this storm has a peak that equals the Georgia Stormwater Management Manual's peak flow calculation for a 1.2 inch first flush storm, the short duration of the storm produces a total flow volume of 1,597.17 cubic feet, versus the 4,356 cubic feet that would be expected from a raw volumetric calculation for 1.2 acre-inch storm.

SCS HYDROGRAPH EVALUATION

The "SCS" hydrograph and the removal efficiencies calculated during this idealized storm are shown in Figure 4. The average removal efficiency calculated from the performance curves is 90.4%. As seen in the chart, removal efficiencies remained above average for about the first 390 minutes and the last 450 minutes of this storm, and dipped down for about 20 minutes, to a low of about 62% at the peak flow of 1.78 cfs. The average removal rate for this storm is higher than the rational hydrograph due to the long duration of relatively low flows which tend to offset the peak flow removal rates, which are lower, but short in duration. This method of storm modeling would show enhanced performance but is probably not very typical of actual field conditions. To chart these two values together, the peak flow of 1.78 cfs is shown at the 100% mark. Other interim flows can be calculated by reading the percentage on the "Efficiency" axis and multiplying the percentage as a decimal times 1.78 cfs.

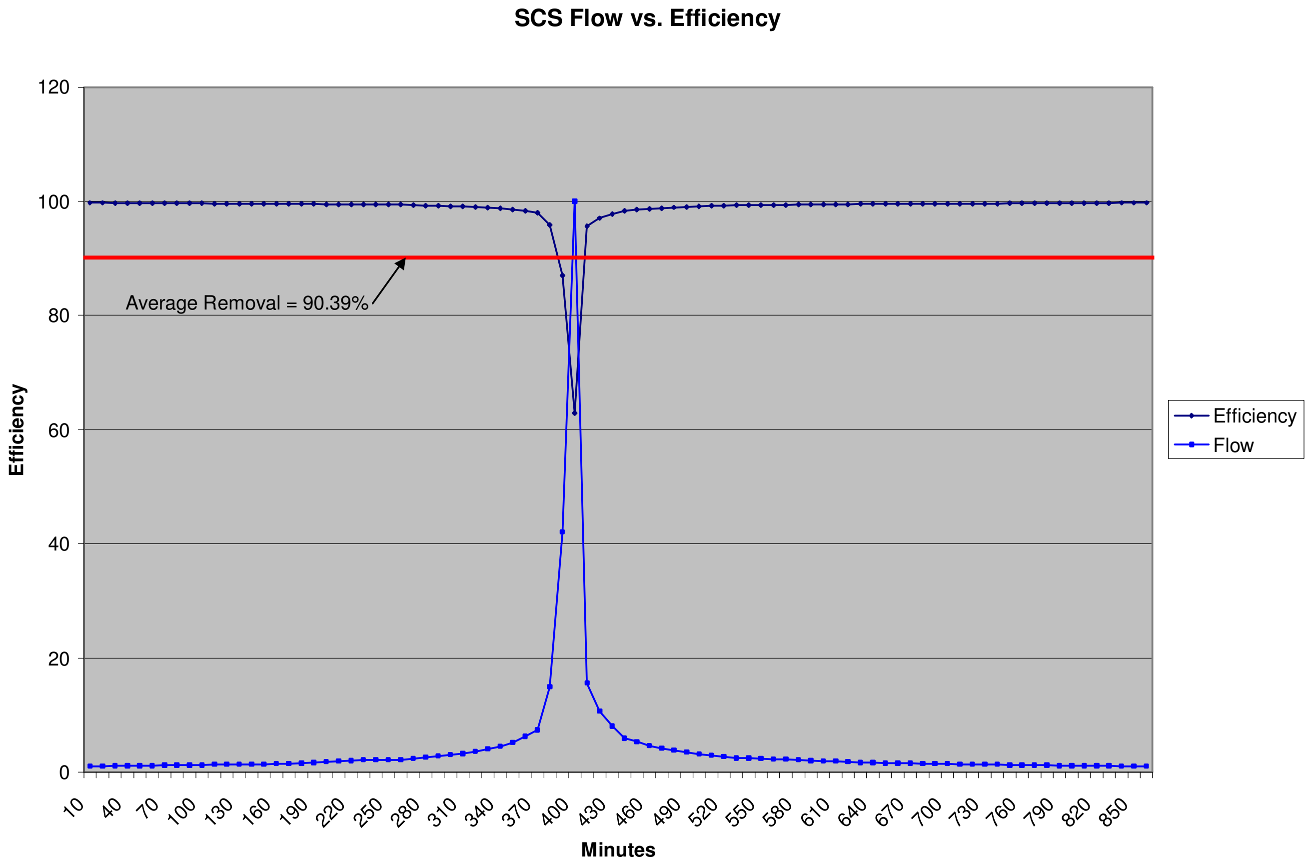

Figure 4

This type of storm represents a less aggressive way to display an event with a peak of this size. These types of hydrographs are also used to size pipes and detention facilities, but they are most often applied to larger areas with one type of terrain. They are based on a rainfall depth, rather than a certain rainfall rate (intensity), so they lend themselves for "first flush" rainfall depth calculations. It should be noted that, while this storm has a peak that equals the Georgia Stormwater Management Manual's peak flow calculation for a 1.2 inch first flush storm, the basic assumptions concerning initial abstraction and related SCS parameters are slightly different between the two methods. This storm as modeled produces a total flow volume of 3,897.23 cubic feet, versus the 4,356 cubic feet that would be expected from a raw volumetric calculation for 1.2 acre-inch storm.

ACTUAL FIELD DATA EVALUTAION USING ETV STORMS

The two removal efficiencies determined from evaluating the rational hydrograph and the SCS hydrograph (82.5% and 90.4%) can be compared to the 89% removal efficiency rate established by the ETV program's field testing. Looking at rainfall models and utilizing laboratory performance curves is useful in that it establishes a reasonably close correlation to the EPA's third-party testing of the CrystalStream water quality vaults, but the modeled storms bear little resemblance to actual rainfall events. Figure 5 is "Event 3" from the ETV test as reported by the monitoring instruments at the site.

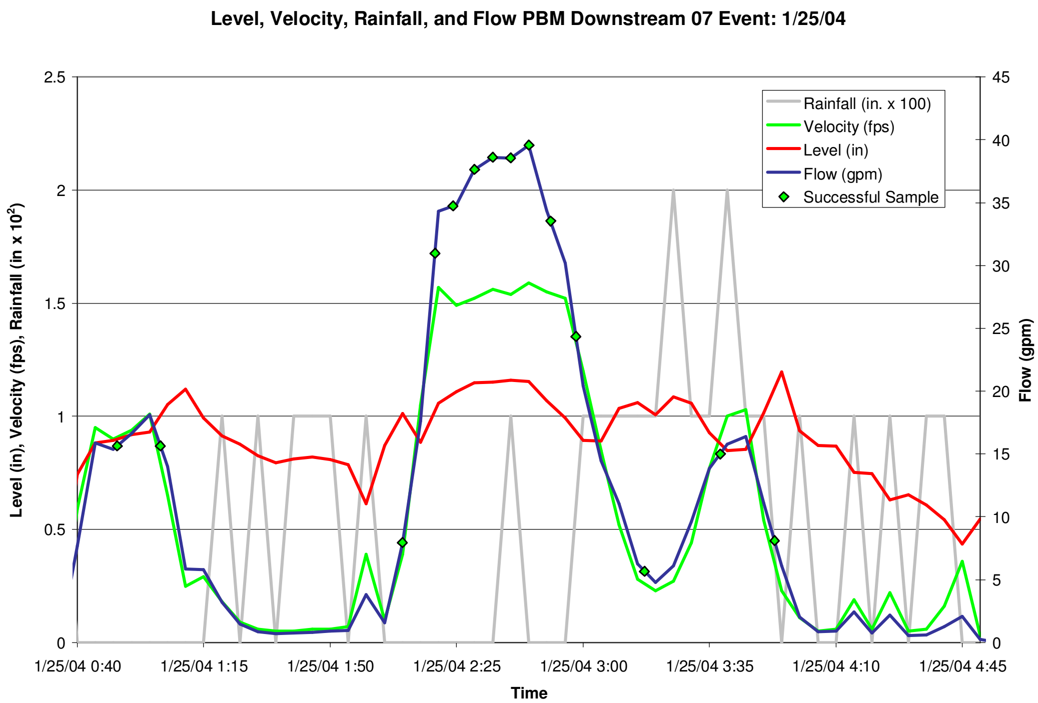

Figure 5

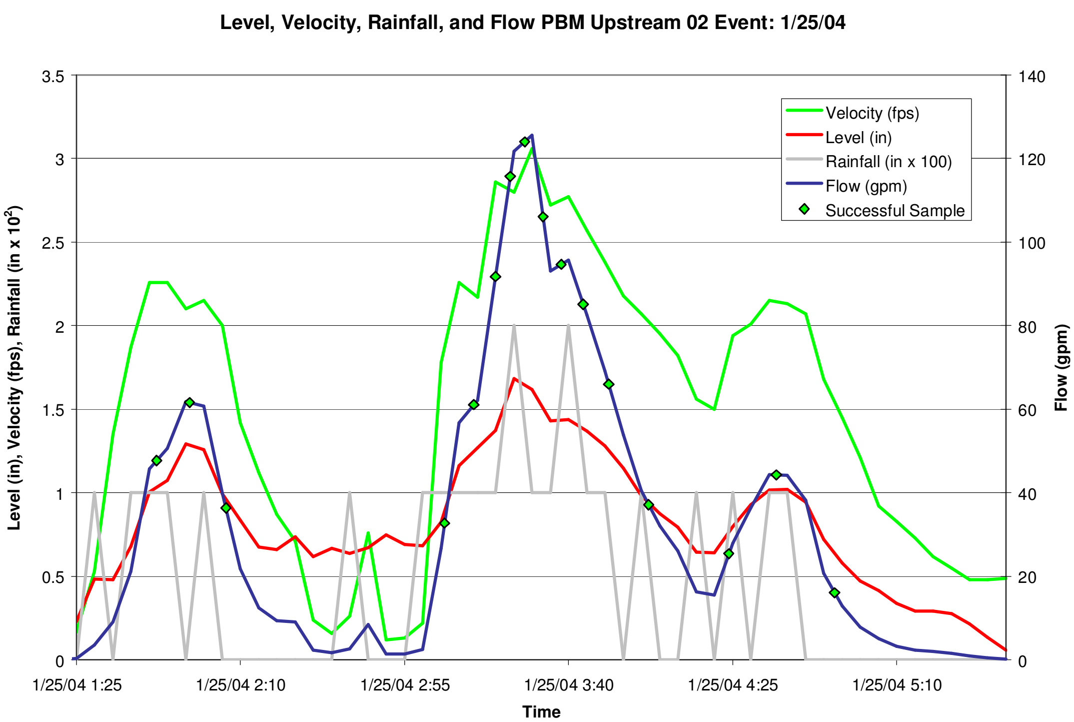

Figure 6

This is a real storm recorded in real time spanning 255 minutes. This data was taken at the downstream station just below the vault and differs significantly from the data taken at the upstream station shown in Figure 6. Overall, the trends are similar, but a comparison of these two "versions" of the same event shows how problematic that field testing can be. It is important to understand that these two sets of data were taken above and below a 10 foot by 5 foot vault with essentially no storage capacity. Almost as soon as water flows into a small vault, it begins to flow out, so that inflow and outflow should be close to equal. In addition, the vault has a standing pool, but it stores no water, so that the total water flowing into the vault and out of the vault should be identical. The nature of the measuring equipment, however, does not recognize temporary rises in water level due to trash screening, or a tilting fiber filter, and the meters tend to get fooled as to flow and volume. To understand this difference, there needs to be a study of how the data differs, and why. Once this is done, projecting performance curves over the field data will be more meaningful.

The field study was conducted on a 4.047 acre basin with 2.836 acres of lawn (a cemetery), 1.145 acres of roadway, and 0.066 acres of sidewalks. The net "C" factor for runoff was calculated to be 0.635. The runoff for each storm was calculated using this "C" factor, the acreage, and the rainfall depth recorded on the rain gauge. This calculated volume is shown on Figure 7, along with the volumes measured by the automated sampling equipment upstream and downstream of the vault. The rational volume in white tends to be larger than either recorded volume, and might suggest that the rational modeling was a little high. The more troubling aspect of this graph is the difference in the upstream and downstream runoff volumes plotted in red and blue. The meters should have determined that the same amount of water entered and exited this small vault, but in events 7, 8, 9, 12, 13, 14 and 15, the volumes were very different, with the upstream volumes often at one-half of the downstream volumes or less.

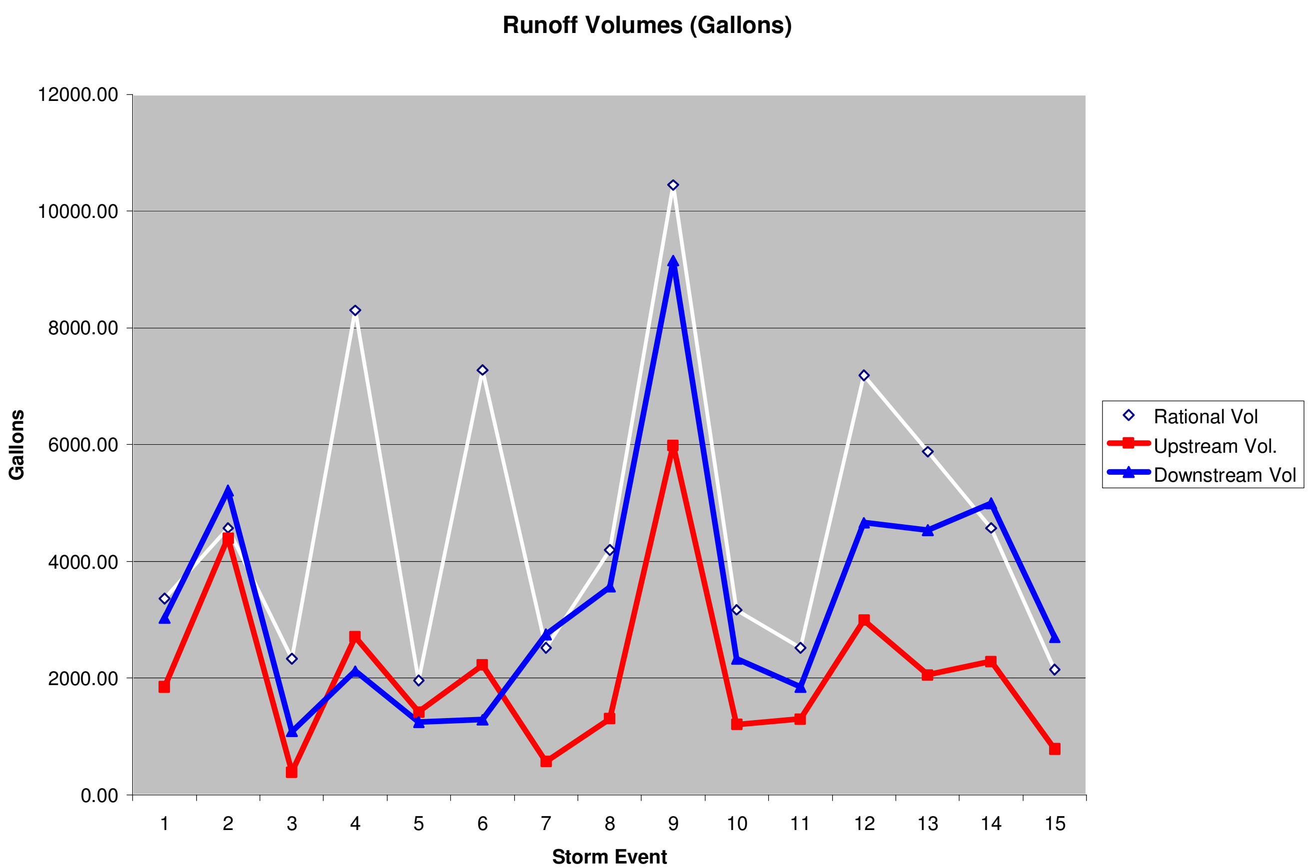

Figure 7

A similar picture is shown when charting the peak flow rates as calculated for the basin versus the recorded peak flows upstream and downstream of the vault by the automated samplers. These flow rates are shown in Figure 8.

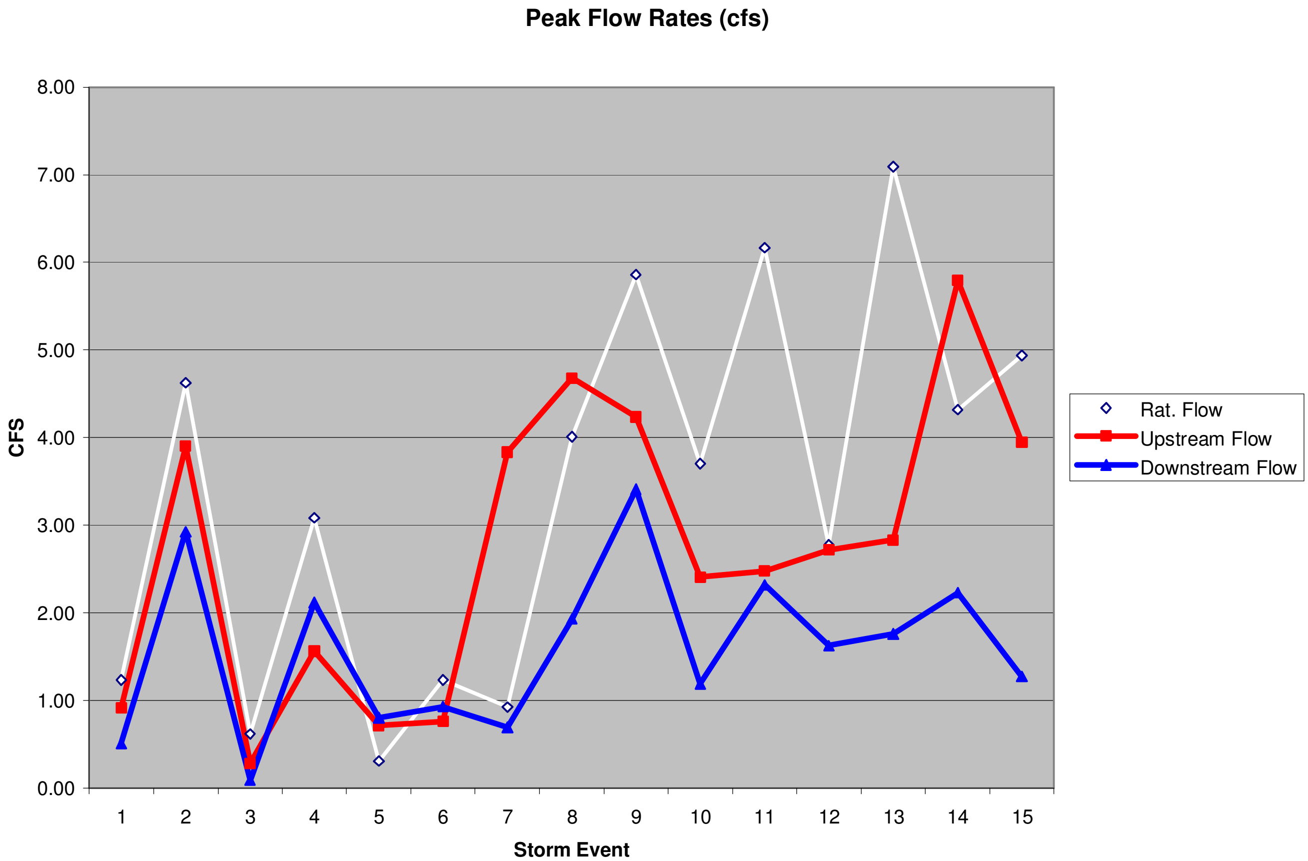

Figure 8

As with the volumes, the peak flow rates tend to be recorded as less than the rational model would suggest. Also, as with the volumes, there are very serious differences between the upstream flow rates and the downstream flow rates for some storms. In particular, events 7, 8, 10, 14 and 15 show large variations.

These variations are pointed out to demonstrate that while field conditions are obviously paramount in determining how a system might operate, data collection is problematic, even in the same basin, during the same storm, and on the same pipe in locations within a few feet of each other. Figure 9 is the detail of one short storm shown in five minute increments for the upstream and downstream data. Based on the rate of flow, the percentage of removal was calculated from the “indirect” lab performance curve, and is plotted. The flow rate in cfs is also plotted. This data is for Event 11, a storm where the upstream and downstream totals agreed pretty closely and showed good correlation. Even in this “good” event, the duration of the storm was shown as double on the downstream instruments. Figure 10 is the detail for Event 14, displayed the same way. The variability of the flows in Event 14 has a dramatic effect on the removal percentage. This variability has made it necessary to show each field event in light of both the upstream and the downstream data as recorded in the field.

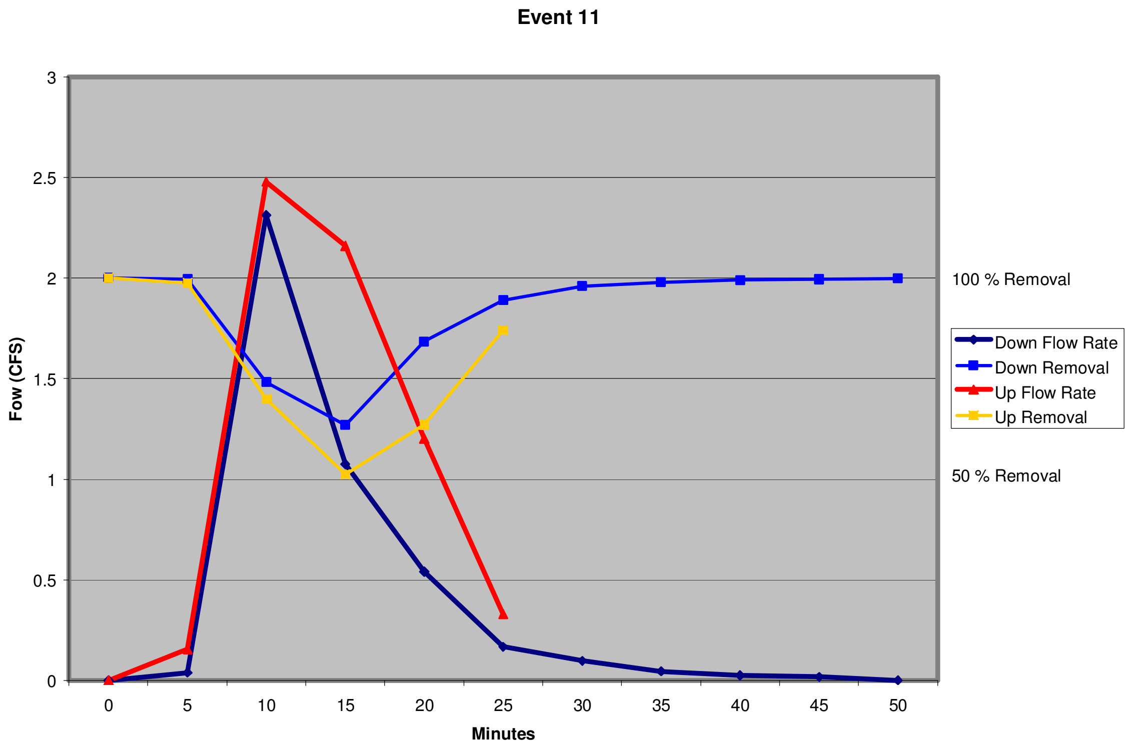

Figure 9

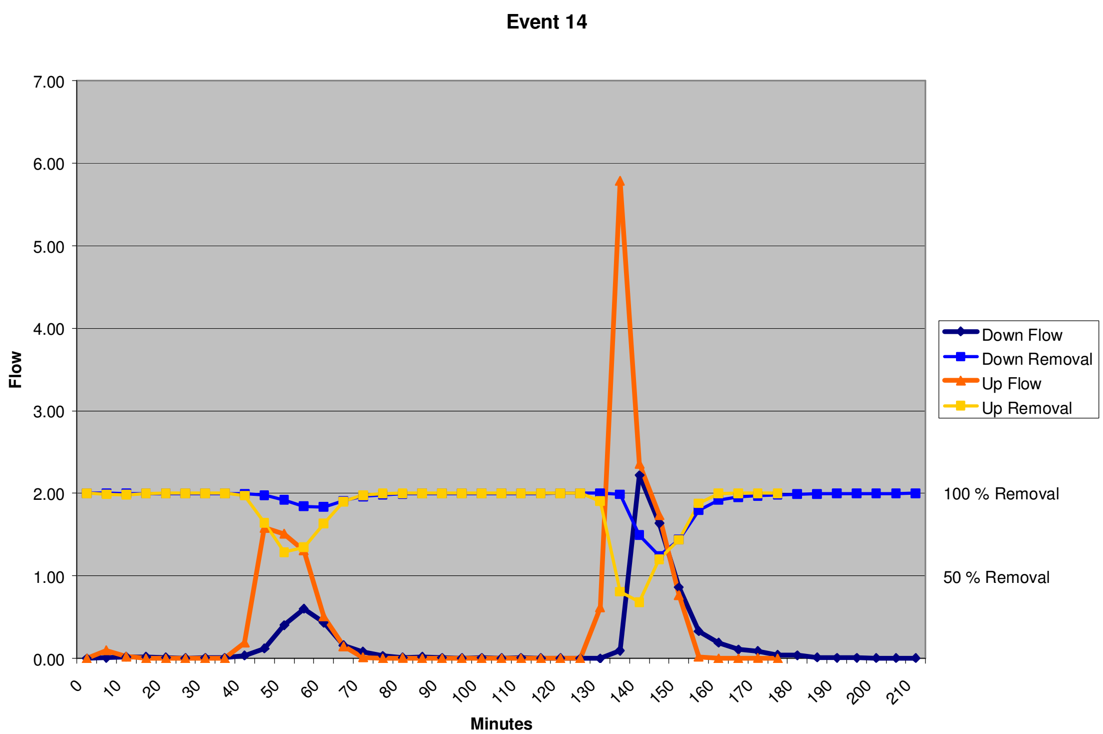

Figure 10

The differences in the upstream and downstream data are significant and can be eliminated by better testing equipment, but the differences are not so great as to render the field data useless. Figure 11 illustrates this as it displays the calculated upstream and downstream efficiencies based on flow, the ETV published overall efficiency, the ETV sand removal efficiency, and the ETV "Fines" removal efficiency. The ETV study deemed all particles over 62.5microns as sand, and the remainder as "Fines".

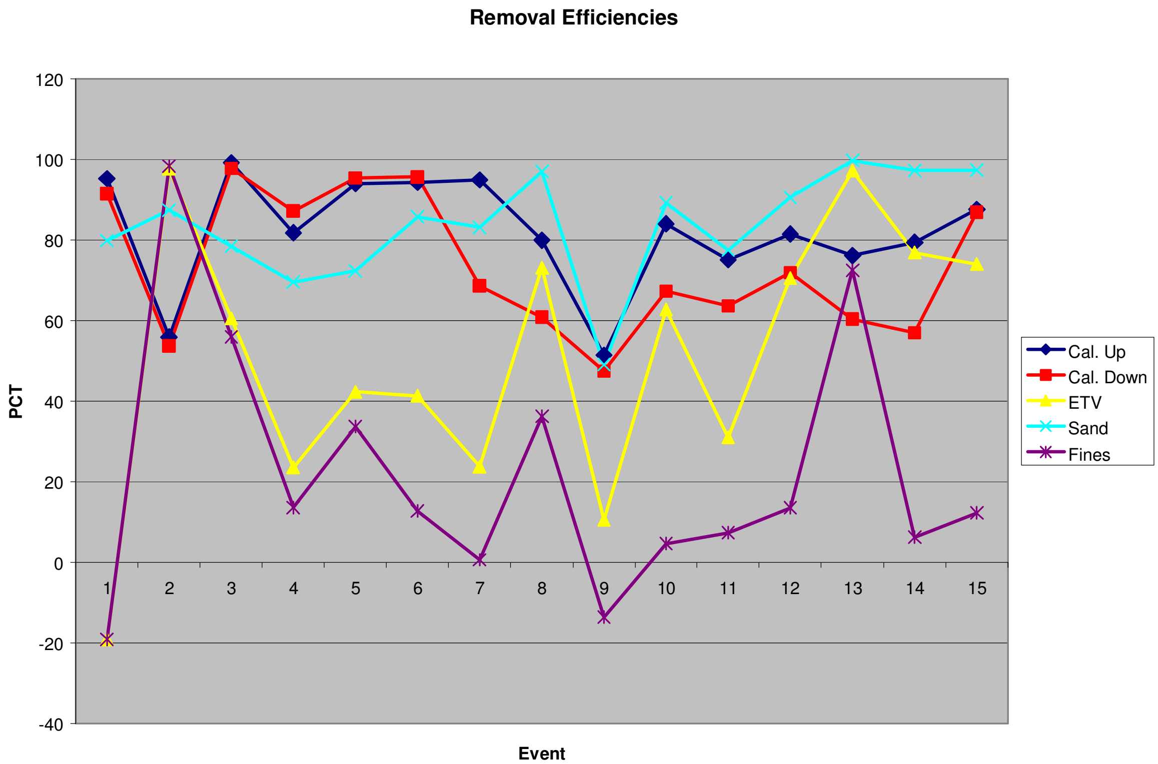

Figure 11

The Event 1 data, shown as negative, was off scale low at -41% for the ETV overall, and -151% for the ETV fines. Although vaults do not typically add particles, this was a relatively long, placid storm which could have washed materials out of the vault, if the antecedent event was especially weak. There was no reason to discard the data, so it was included in the ETV evaluation, and this study. One major trend evident on this storm-by-storm breakdown is the tracking of the OK-110 (sand) laboratory curve data, and the field sand data. Even with the variation between upstream and downstream calculated data, the correlation is very close. This indicates that laboratory performance curves can be useful for predicting removal rates when particle sizes are relatively well matched. While the full particle size definition of the field material was not measured, it can be inferred that the overall distribution was close to that of the OK-110 material used in the lab. The second major trend is that the ETV published efficiencies more closely track the "fines" removal efficiency, although there is also a definite influence from the sand fraction. This storm-by-storm tracking is indicative of the variation between events, where efficiencies can be vastly different.

The major goal of field and laboratory testing is to indicate how a vault will operate over all events and in all conditions. Long term efficiencies must be estimates based on the data available. As extensive and costly as the ETV study was in Griffin, GA, it is only for 15 storms in one basin. The full-scale laboratory tests were also expensive and lengthy, but they included only one homogeneous material tested at just a few speeds. Neither method reflects true field operations very well. With this understood, it is still necessary to draw some conclusions about removal efficiency and how the available data can be used to predict how a vault will work on a long-term basis in the field.

SUMMARY

It is important to remember that the efficiencies shown in Figure 11 are for individual storms. The ETV study published an overall efficiency of 89% for the CrystalStream vault studied in the field. This overall efficiency was determined by a flow weighted method, where the "sum of loads" was calculated for all the samples taken in all the storms. Two large storms accounted for a large portion of the overall weight of materials removed by the CrystalStream device, but this is predictable and normal under real field conditions. Figure 12 shows the average, flow weighted removal rates for the laboratory calculated performance, the ETV overall removal rate, ETV sand removal, and ETV "Fines" removal.

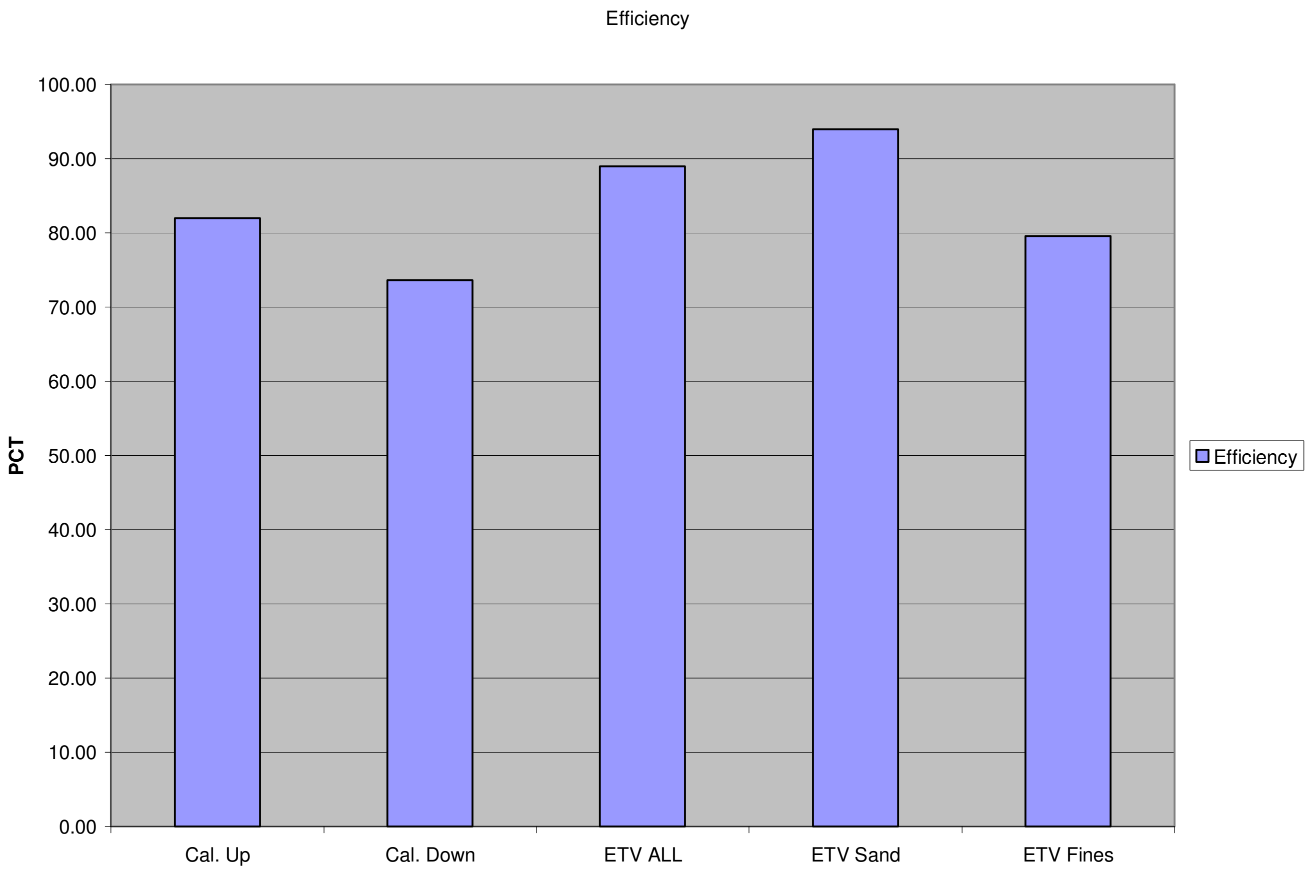

Figure 12

This overall picture shows that the laboratory performance curves generated come very close to predicting the overall field efficiency. This is despite some obvious problems with comparing laboratory data with field data.

The most difficult problem to overcome is the nature of the material used in the laboratory. Pure silica sand has a relative density of 2.65, where CrystalStream research has shown that the average density of material removed from our vaults is closer to 1.70. The major contributor to less density is trash and vegetative debris. The particle size definition of field materials is also unlikely to match laboratory materials. Some efforts have been made to add two or more materials together to give a more "realistic" mixture of particle sizes, but true field material varies greatly, depending on the basin characteristics. One "mix" cannot simulate field conditions.

There are two other layers of complexity that complicate matching laboratory curves to field performance. The first is that the particle size definition of the material is not static during a storm. The mixture changes as conditions change. Not only is there an effect to have smaller particles early in a storm, based on antecedent conditions, there is an effect from higher flows on the surface entraining larger materials, and higher flows in pipes and other conveyances being capable of transporting larger materials. As rainfall rates and flows constantly change, the material in the water column changes in other ways, as well. The concentration of materials will vary during a storm as well. Some researchers believe that there is an "irreducible" concentration of materials that simply cannot be removed from stormwater, which is a concept that might apply to very small flows on the "back side" of a large event, as well as at other times. Very low concentrations in small flows might even show "negative" removal rates in a vault where essentially clean water is flowing into a vessel that is loaded with material from earlier events.

An ideal laboratory performance curve would include varying the particle size definition and concentration as flows increase. A baseline could be established by testing two materials separately, such as a 20 to 40 micron silt based material and something like OK-100, both at a static concentration and at varying flows. Then a static ratio of the same materials, such as 30 percent of the lighter material, and 70% of the heavier material could be tested at a static concentration in the same way. The final step would be to test again, starting with a low concentration of a "fines" dominated particle size distribution, gradually increasing concentration and increasing the particle size at the same time. This initial laboratory effort would expose weaknesses in how variable testing can be done to better emulate field conditions. This type of work is needed, because field testing conditions are going to change. Testing equipment is not well suited for the chaos of "real world" stormwater. Even if the instruments work perfectly, and a sample represents the perfect cross-section of the water column, can the sample be said to truly represent that portion of the storm? With actual field fluctuations of rainfall and flow, can we ever be sure exactly when to sample? The fact is that we cannot do field testing very well.

After all the difficulties with the field data, and the fact that laboratory materials and conditions cannot emulate the field materials and conditions very well, we find that our comparisons of laboratory and field efficiencies correlate very well. Although some portion of the correlation must lie within the realm of compensating errors, it is obvious that the concept has proven to be viable. We have sampled in the field as best we know how and sampled in a controlled laboratory as accurately as possible, and there is a reasonable comparison. The underlying physics of sedimentation, entrainment and transportation makes the concept viable. Even though stormwater flows are chaotic, we know a lot about water flows and the behavior of particles in water. This indicates that we can develop better laboratory performance curves that can closely predict field performance.

REFERENCES

1. Verification Testing of the CST 956 Water Quality Vault, James T. Mailloux, Alden Laboratories Document 206-303_H1181C, October, 2006.
2. Environmental Technology Verification Report (Draft), Stormwater Source Area Treatment Device, Practical Best Management of Georgia Inc., CrystalStream ™ Water Quality Vault, Model 1056, Prepared by: NSF International, 2005
3. Georgia Stormwater Management Manual, Volume 2, Technical Handbook, First Edition, August 2001, State of Georgia Environmental Protection Division

Note: Figures 1 and 2 are from the Alden Laboratories report. Figures 5 and 6 are from ETV detail data.