Potential Data Analysis Methodology to Evaluate the Performance of Manufactured BMPs

Masoud Kayhanian[1*], Robert M. Roseen[2], James H. Lenhart[3], Greg Williams[4]

[1*]Department of Civil and Environmental Engineering, University of California, One Shields Avenue, Engineering III, Davis, CA 95616; PH (530) 752-8957; FAX (530) 752-7872; email: mdkayhanian@ucdavis.edu

[2]The University of New Hampshire Stormwater Center, Department of Civil Engineering, 35 Colovos Road, Durham, NH 03824; robert.roseen@unh.edu

[3]CONTECH Stormwater Solutions Inc., 11835 NE Glenn Widing Dr., Portland, OR 97220, LenhartJ@contech-cpi.com

[4]Monteco Research & Development Centre, 2596 Dunwin Dr., Mississauga, ON L5L 1J5 gwilliams@monteco.com

ABSTRACT

Evaluating the performance of manufactured BMPs on a consistent and scientifically sound approach is beneficial for both the service provider and the services recipient. To do this properly, it is important that these devices need to be tested under a standard set of protocols. The testing data must be collected, reported, and validated prior data analysis. The testing, data collection, data reporting and validation will be addressed under a separate ASCE/EWRI subcommittee. The focus of this paper is to address the data analysis and performance evaluation of manufactured BMPs.

To address this issue the existing statistical data analysis methods and performance evaluation that potentially could be used for manufactured BMPs have been examined in this paper. Special attention was devoted to the data distribution and the issue of normality since that will influence the selection of suitable data analysis approach. In general, it has been concluded that the stormwater data is log-normally distributed. The existing BMP performance evaluation has also been evaluated and the effluent probability plot has been recommended to determine the performance evolution of manufactured BMPs.

INTRODUCTION

Background. Since the passage of the United State Clean Water Act in 1972 and consequently through NPDES amendment in 1983, many local, state and industry are obligated to treat the stormwater runoff before discharging them to a receiving water. Initially with Phase I, the regulatory compliance was only applicable to communities with population higher than 100,000. Since as early as 2002, Phase II extended the same regulatory framework to communities with population of 10,000 or higher and to areas along the urban fringe that together have a residential population of at least 50,000 and an overall population density of at least 1,000 people per square mile. For this reason, demand for the use of compact treatment systems has increased and as a result numerous manufacturers responded and developed wide ranges of structured best management practices (BMPs) to remove different pollutants. These treatment systems are defined as manufactured treatment devices (MTDs). Some organizations such as Caltrans, who had the technical capability and financial resources tried to develop their own treatment systems that is more suitable on their right-of-way facilities (http://www.dot.ca.gov/hq/env/stormwater/special/newsetup/index.htm#new_technology). In recent years, however, the use of MTDs are rapidly increasing in order to meet escalating water quality regulatory requirements; particularly in re-development and new development areas where land space is limited and potential use of other BMPs are not practical. As the use of MTDs has been increased, the demand for some type of certification for their effectiveness in pollutant removal and meeting water quality standard has also been increased on national level. Several attempts have been made on state and regional level (TAPE, TARP, ETV). However, no consistent protocol has been established for national use.

In response to this national demand, a manufactured BMP certification committee was established under the EWRI's stormwater infrastructure committee of water, wastewater, and stormwater council (WWSC) and the wet weather flow technology committee of the urban water resources research council (UWRRC) (Guo et al., 2007). In addition, parallel efforts are being undertaken by ASTM and WERF. To avoid inconsistencies between the organizations, liaisons from the task committee were established for these other groups. To develop the certification guidelines, several subcommittees were established and each subcommittee was responsible for specific task to develop the final certification guidelines document. The subcommittees to tackle specific tasks include: data reporting, data analysis and performance evaluation, lab testing, field monitoring, scaling, and maintenance. Deliverables from each sub-committee will be a technical guidance manual that will be published by EWRI. The work of each sub-committee will be presented at the annual ASCE/EWRI Congress conference meeting and finally will be published as a peer reviewed ASCE journal paper.

This effort has been broadened to be consistent with the reporting requirements for the International Stormwater BMP database (BMPDB) developed by a broad coalition of partners including the Water Environment Research Foundation (WERF), ASCE Environmental and Water Resources Institute (EWRI), USEPA, Federal Highway Administration (FHWA) and the American Public Works Association (APWA). The purpose of the BMPDB is to provide scientifically sound information to improve the design, selection and performance of BMPs. Continued population of the database and assessment of its data will lead to a better understanding of factors influencing BMP performance and help to promote improvements in BMP design, selection and implementation.

Focus of the paper. The objective of this paper is to summarize the data analysis and performance evaluation task and to prepare a conference manuscript to be presented at the ASCE/EWRI Congress conference at Kansas City, 2009. The specific objectives of this manuscript are to present (1) the potential data analysis approaches, and (2) to evaluate potential methodologies for assessing manufactured BMP performances. From this preliminary work, there will be an opportunity to recommend consistent protocol to evaluate the performance of manufactured devices.

POTENTIAL DATA ANALYSIS APPROACHES

Data Quality and Reporting Prior Statistical Analysis. The central goal of a BMP verification program is first to collect scientifically defensible data in fulfillment of the verification objectives. Statistical methods may be used in the design of the lab, pilot, or field testing to determine data quality objectives (DQOs). DQOs are an important and useful step to ensure that the data produced by laboratory or field testing is sufficient, both in terms of quality and quantity, to establish meaningful confidence levels for the BMP performance evaluation.

The use of standardized procedures to collect BMP performance data and reporting them is essential for successful verification program – not just from one BMP to another, but also from project to project, region to region, year to year. If consistent methods are not applied, it may be difficult to perform data comparisons, and the value of the data may be diminished. For instance, flow can vary significantly throughout a runoff event, and rainfall intensity and runoff volume (Kayhanian et al., 2007). Flow-proportioned composite samples are therefore considered to be the most representative sampling regimen for runoff monitoring. The preferred set-up for water quality evaluation in BMPs is to employ automated monitoring equipment to collect an equal sample volume (aliquot) for every increment of a pre-set runoff flow volume. The composite sample ideally needs to cover the full event hydrograph and accounts for variation in flow and/or runoff quality throughout the runoff event. Grab sampling, or even time-proportioned compositing, by contrast, may generally produce data that are not representative of any given runoff event. Grab sampling provides no information on the contaminant mass load, which in many cases is the ultimate goal. Depending on the sample timing and the nature of the events monitored, this can lead to a substantial bias in the reported results. However, grab samples are typically needed for collecting solid samples (including sediment and litter) where automatic samplers may be impractical.

Regardless of method of sampling collection, the verification protocol should provide a series of steps to be taken both in the field and in the lab to ensure the overall quality of data. The product of hard work related to representative sample collection, analytical testing method, and proper QA/QC will still be a relatively useless pile of data reports if there isn't consistency in and an established protocol for data reporting. The data reporting protocols help ensure that data from all tested BMPs will be reported in consistent format – and that the data records will include sufficient information to permit their full use within a common database. For instance, if half of data for total suspended solids are reported as TSS and the remaining half as "Total Suspended Solids", extracting the data from a database may not include all dataset. Similarly, reporting inconsistent unit for TSS as PPM or mg/L may create similar problem on data extraction and data analysis. In addition, a well-designed database having a user-friendly interface feature can be developed to provide an efficient and uniform data storage and data retrieval as an important and essential component of a successful verification program.

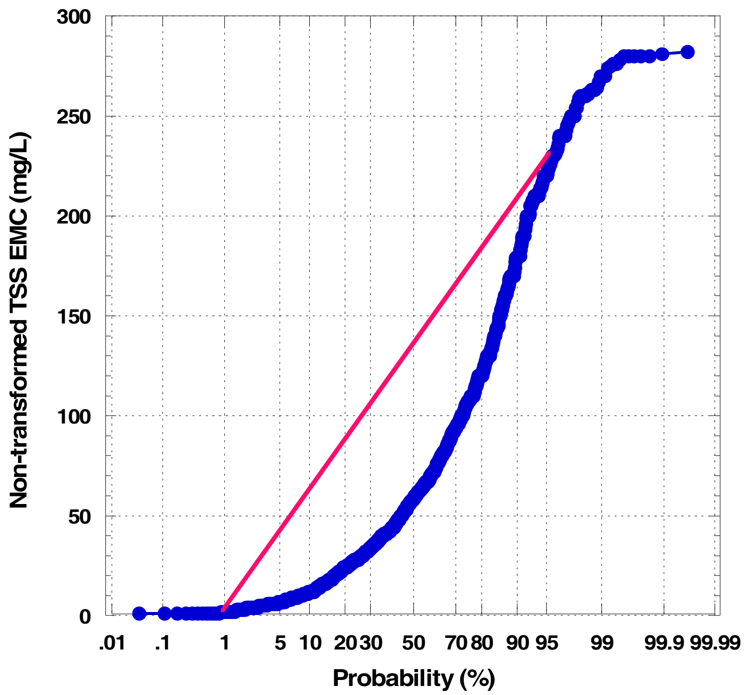
Performing detail statistical analysis and performance evaluation of manufactured BMPs will require additional data collection such as flow, physical characteristics of the manufactured device, monitoring site characteristics, and storm event characteristics. These data can be included as part of the same database. The principal data components potentially to be collected under each category are summarized in Table 1. Fields and content guidelines for each of these categories will be described in detail by the Data Reporting subcommittee.

Table 1. Summary of potential data to be collected as part of the manufactured BMP verification.

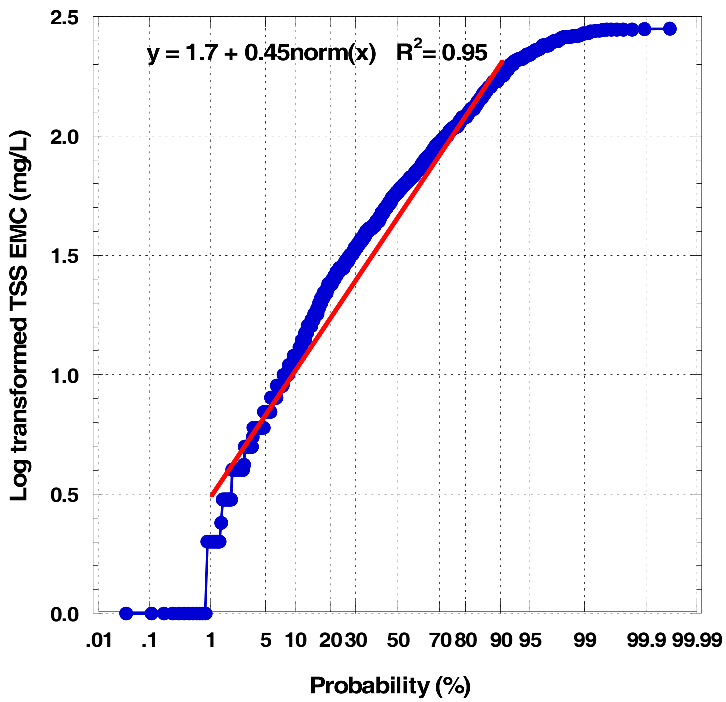
Data collection category	Data component collected under each category
WQ Parameter and solids characteristics	*Water Quality:* conventional, nutrients, aggregate organics, metals (total and dissolved) PAHs, etc…… *Solids:* trash, total suspended solids (TSS), total dissolved solids (TDS), suspended solid concentration (SSC), particle size distribution (PSD)
Storm event and hydraulic characteristics	Storm event starting and ending time, rainfall characteristics, inflow and outflow volume, antecedent dry period, etc……….
Device physical characteristics	Device name, principal pollutant removal mechanism, design characteristics, size, etc……………….
Monitoring site characteristics	Drainage area, imperviousness, number of lanes for highway, AADT, etc……….

Statistical Analysis. This section of the paper is organized to deal with some of the basic statistical analysis that can be performed to address specific question. General questions and statistical approach that are commonly used to address them are summarized in Table 2.
Detail discussion of each topic presented is beyond the scope of this paper and will be presented in final certification guidelines documentation. An overview of recommended statistical analyses for BMP performance monitoring can be found in the recently updated EPA manual (EPA, 2009). The selection of an appropriate statistical test based on the distribution of data (e.g., normality verification) is further discussed below.

Testing the Normal Distribution of Data. Various methods are available to test the normality of data distribution. The best method to verify normal distribution of stormwater runoff data is the probability plot of the original data and transformed data. Based on the past experience, the majority of the environmental data including stormwater runoff characterization and BMP performance data typically are not normally distributed. Log transformation is commonly the most suitable method to verify the normal distribution of data. A sample distribution of stormwater data for TSS prior to transformation and after log-transformation is shown in Figure 1. As shown, the data distribution before transformation resembles an exponential function. After log-transformation of the data, the majority of the data plots linearly (at least between 10 percentile and 90 percentile range) and hence assumption of log-normal distribution is appropriate. The assumption of log-normally distributed data is important when performing a comparative statistical analysis (Shumway et al, 2002; Kayhanian et al, 2002).



(a) non-transformed TSS data



(b) Log-transformed TSS data

Figure 1. Probability distribution of TSS EMC before and after log-transformation.

Table 2. Summary of Statistical Approach Addressing Specific Question in BMP Testing.

Statistical topic	Question to be addressed	Statistical approach
Sample size	How many samples are needed to achieve desired confidence? After a few years of field sampling or number laboratory testing, how would we decide whether we need more samples than initially planed?	USEPA statistical methods for practitioners: EPA QA/G-9S
Selection of an appropriate statistical test	How do we verify whether data are normally distributed? How do we very that the data variability of two or more groups is similar? This question is specifically useful when preparing the BMP probability plots.	Graphical data presentation using histogram, box plot, normal probability plots. Numerical evaluation through Shapiro-Wilk W test
Addressing non-detect in data set	How do we account for non-detect results?	Regression on order statistics (ROS)
Examining outlier in data set	How do we prepare graphical and numerical data summaries to identify potential outliers?	Histogram, box plot, standard and log-transformed probability plots, and scatter plot
Probability of meeting water quality requirements	How do we estimate how often the average BMP effluent concentration would meet a water quality standard or regulatory discharge limit? How do we estimate the BMP percentage removal of a pollutant with a specified confidence level?	This should be available directly from the effluent probability plot of laboratory or field testing data, assuming the field data sample is large enough to be representative.
BMP effectiveness for specific pollutant removal	In an influent-effluent approach or before-after approach, how do we decide whether a given BMP is effective in removing a pollutant?	Hypothesis testing through null hypothesis (H_0) or alternative hypothesis (H_A). Statistical test performed include student's t-test with equal variance and t-test with unequal variance. Both tests assume a particular probability distribution.
Comparing effectiveness of two or more BMPs	How do we compare the effectiveness of two or more pilot BMPs under similar test condition or under a given geographic location?	Paired t-test comparing the mean of effluent concentration for normal or log-normally distributed data. Comparing three or more BMPs use the analysis of variance (ANOVA). Specific analysis include: parametric ANOVA (equal variances), Welch ANOVA (unequal variance).
Developing a linear regression equation	How does BMP effectiveness vary as a function of factors such as storm event characteristics, BMP design variables, operational, maintenance practices?	Multiple linear regression using the following form: $y = \beta_0+\beta_1x_1+\beta_2x_2+\ldots\beta_m x_m+\varepsilon$ where, y=response variable, x=model variable, β = model parameters, and ε random error
Effectiveness of BMPs over time	How can we tell if the effectiveness of a BMP is changing over time within a year or multiple years?	Graphic representation using time-series plot or Mann-Kendall test.

The normality can further be tested using goodness of fit test applying the Shapiro-Wilk W test. For example the data presented in Table 3 can be tested for normality using graphical representations of data using standard probability, histograms, or box and whisker plots. The results of the Shapiro-Wilk W test for the previous data set are shown in Table 3 and illustrates that the original data did not fit the normal distribution but was log- normally distributed. The brief discussion presented here could justify the log-transformation of manufactured BMP data and may be used as a basis to perform other statistical analysis.

Table 3. Summary of Shapiro-Wilk W test to verify the normal distribution of a typical data.

Test #	BMP effluent	Shapiro-Wilk W test result			
1	10.9	*Parameter estimate prior transformation:*			
2	4.3	*Parameter*	*Estimate*	*Lower 95%*	*Upper 95%*
3	6.9	μ	11.84	8.75	14.94
4	0.7	σ	13.90	12.03	16.47
5	2.9				
6	79.6	*W test result:*			
7	9.6	*W = 0.642*	*Prob<W = <0.0001*		
8	4.0				
9	4.8	*Since the probability is less than 0.05, we can conclude that the*			
10	14.3	*assumption of normal distribution is not reasonable.*			
11	14.5				
12	4.4	*Parameter estimate after log- transformation:*			
13	12.2	*Parameter*	*Estimate*	*Lower 95%*	*Upper 95%*
14	13.5	μ	1.99	1.77	2.22
15	8.6	σ	1.02	0.88	1.21
16	8.4				
17	11.7	*W test result:*			
18	9.7	*W = 0.974*	*Prob<W = <0.105*		
19	17.8				
20	3.8	*Since the W is much higher that the W of the original data and*			
21	4.6	*probability is greater than 0.05, we can not reject the normal*			
22	11.6	*distribution of log-transformed data.*			
23	3.6				
24	1.5				
25	10.3				
26	16.0				

MANUFACTURED BMP PERFORMANCE EVALUATION

Review of the historical approaches. During the past ten years numerous methods have been applied to evaluate the performance of constructed and manufactured BMPs. These methodologies are summarized in Table 4. A few of the performance evaluations shown in Table 4 (e.g., influent and effluent concentration or load basis) are more commonly used while others are less common. The approach for using the performance evaluation in most cases is straight forward. Two other methods specifically the "lines of comparative performance" (Minton, 1999) and "performance expectation function" (Lenhart, 2007) have also been introduced and their application have not universally been adapted. However, among all performance evaluation the probability method may be the most comprehensive way to evaluate manufactured BMPs and it is gradually being recognized as method of choice as it has been recommended by the USEPA/ASCE (2002). The potential application of this technique is further discussed below.

Table 4 Summary of Methodologies used for BMP Performance Evaluation.

Method	BMP Performance Evaluation Method	Approach/Formula	Applicable certification program
1	Pollutant EMC reduction for single test or storm event	$\text{Efficiency} = 100 \times (1 - \frac{\text{Effluent EMC}}{\text{Influent EMC}})$	ETV, TARP, TAPE, EPA
2	Pollutant load reduction for single test or storm event	$\text{Efficiency} = 100 \times (1 - \frac{\text{Effluent load}}{\text{Influent load}})$	ETV, TARP, TAPE, EPA/ASCE
3	Pollutant load reduction for multiple tests of storm events	$\text{Efficiency} = 100 \times (1 - \frac{B}{A})$	ETV, TARP, TAPE, EPA/ASCE
4	Mean pollutant concentration removal efficiency	$\text{Efficiency} = 100 \times (1 - \frac{\text{Average effluent Conc.}}{\text{Average influent Conc.}})$ This technique does not require flow measurement and it might be useful for pollutants with grab sampling where no flow weighted data is available.	EPA/ASCE BPM performance monitoring
5	Regression of loads (ROL)	Compute effluent and influent load same as method 2; perform linear regression based on outflow and inflow load and determine the R^2 and slope of regression line. Under special circumstance, the regression lines may be forced through (0,0) as it may produce a more realistic lead removal.	EPA/ASCE BPM performance monitoring
6	Relative to achievable	$\text{Efficiency} = 100 \times (1 - \frac{\text{Attainable limit Conc.}}{\text{Average influent Conc.}})$	EPA/ASCE BPM performance monitoring
7	Relative to water quality standard	$\text{Efficiency} = 100 \times (\frac{\text{avg. infl. conc. - avg. effl. conc.}}{\text{avg. infl. conc. - WQ standard conc.}})$	EPA/ASCE BPM performance monitoring
8	Effluent probability method	This method provides comprehensive BMP performance evaluation by presenting the standard parallel probability plot of both influent and effluent EMCs for all storm events or performance tests.	EPA/ASCE BPM performance monitoring

Effluent load = effluent EMC × effluent volume
Influent load = influent EMC × influent volume

$$B = \left[\sum (\text{effluent EMC}_1 \times \text{flow } V_1) + (\text{effluent EMC}_2 \times \text{flow } V_2) + ...(\text{effluent EMC}_n \times \text{flow } V_n)\right]$$

$$A = \left[\sum (\text{influent EMC}_1 \times \text{flow } V_1) + (\text{influent EMC}_2 \times \text{flow } V_2) + ...(\text{influent EMC}_n \times \text{flow } V_n)\right]$$

Effluent Probability Method. Effluent probability method is probably the most comprehensive method that could be used to determine the performance evaluation of manufactured BMPs. Under this approach, a standard probability plot is prepared for both inflow and outflow pollutant EMC or load. Before, preparing the probability plot, however, non-parametric (or if applicable parametric) statistical tests should be conducted to indicate if the difference between influent and effluent are statistically significant. In addition, the log normality should be checked and the probability plot may be prepared based on log transformed data. If the log transformed data deviates significantly from normality, other transformations can be explored to fit the data distribution. Three example of log-normally distributed probability plots are shown in Figures 1 through 3. Figure 1 shows the probability plot for TSS data for a BMP that there is a clear difference between influent and effluent event mean concentration. Figure 2, shows the TSS data for another BMP that there is no distinct difference between influent and effluent EMC. Under this situation, no removal efficiency can be expected. Figure 3 shows the results for TSS data for a condition when a BMP performed fairly well when influent TSS concentration was high but achieved no removal when TSS concentration was low.

One problem associated with the probability method is the lack relationship with direct influent and effluent measurement. For instance, a BMP performance is much higher under correlated data shown in 2 vs. 1 (see top Figure). Ideally, the BMP performance can remain within a narrow range with high probability of meeting effluent requirement regardless of influent concentration or testing condition. For example, the sample probability plots could be used to determine if BMP could meet certain regulatory or standard effluent quality requirement. As an example, under ideal BMP performance (see top Figure 3), the probability plot shows that effluent TSS quality of 30 mg/L can be met at all time. Multiple probability plots can also be produced under a single plot to compare the performances of two or more BMPs.

Another issue that worth to discuss with respect to BMP performance is that the pollutant removal generally decreases as influent concentration decreases. This may not be a practical problem since the point where BMP performance starts to decrease is often close to the point where the influent concentration is low enough that is does not exceed the discharge requirement and hence a high level of treatment is not needed. Similarly, if majority of pollutants are associated with smaller particle size, some manufacturer devices may not be able to sufficiently meet the discharge load requirement at full capacity flow rate. However, it has been shown that this phenomenon is largely dependent on the dominant unit process used in certain MTDs and as such filtration systems do not exhibit much variability in removal (Roseen et al, 2006). Nevertheless, there is value in requiring BMPs to report performance even at low concentrations, recognizing that removal may not as important if discharge limits are being met. At the other extreme, discharge limits may exceeded at very high influent concentrations even though the BMP were capable of removing 80% of influent concentration. Therefore, testing the manufacture BMPs under wide ranges of influent flow and pollutant concentration is essential in order to determine the true performance characteristics.

SUMMARY AND CONCLUSIONS

Generally, the manufactured BMP performance data collected over a range of flow rates is not normally distributed. This can impact the choice of statistics used to describe performance. Log transformation found to be a good choice for normal distribution and hence the statistical data analysis should be performed accordingly. The choice of statistical method usually related to question to be addressed and must be performed on consistent basis throughout the verification program.

Consistent, data collection, data management, data retrieval, and data analysis is an important and essential integrated component of a successful manufactured BMP verification program. This may apply to both laboratory and field testing; although it is more realistic for laboratory generated data so lab data collection should be a requirement. For proper performance evaluation, the manufactured BMP may be tested at 25, 50, 100 and 125 percent of designed flow. BMP testing at or above 100 percent of designed flow rate is especially useful to assess scouring.

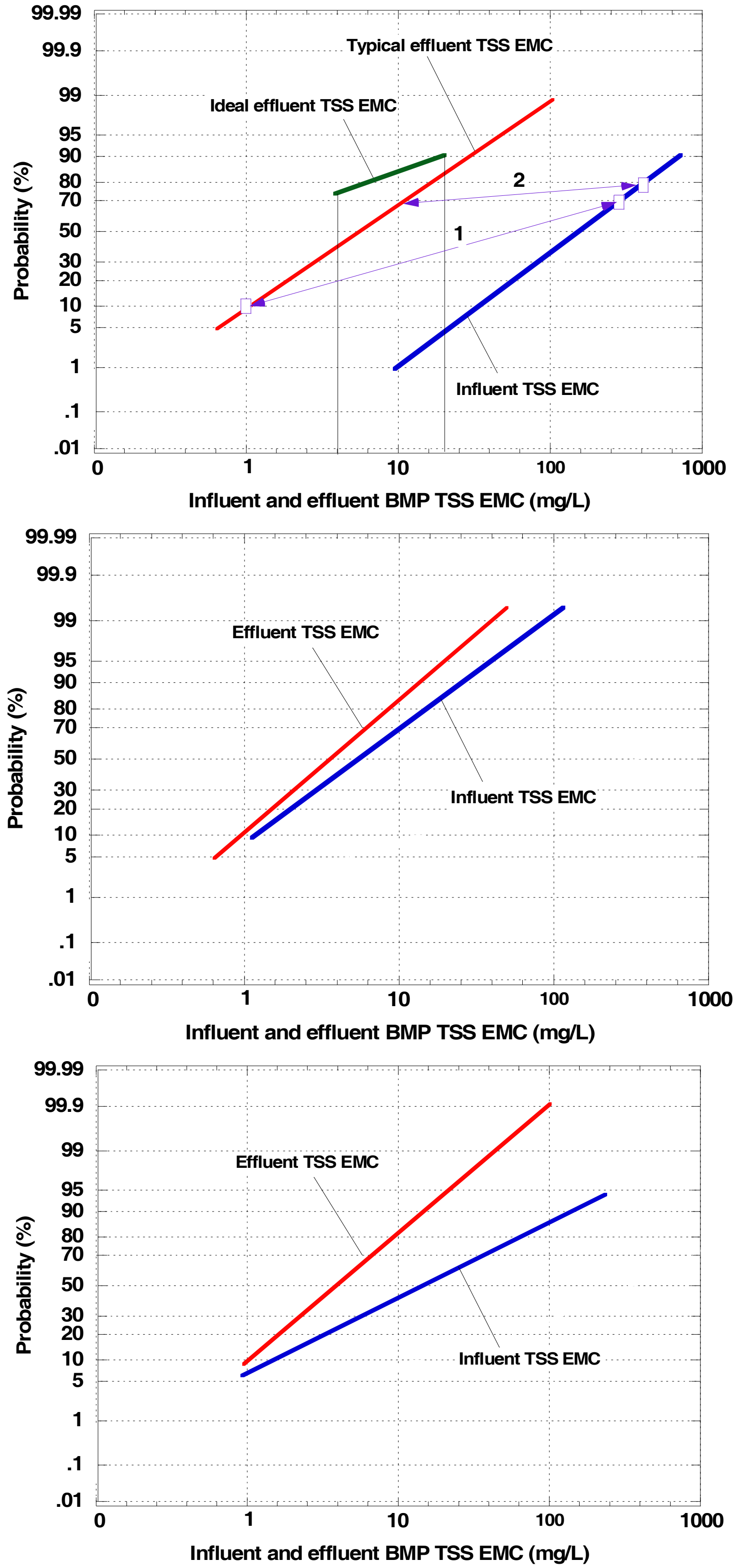


Figure 3. Probability plot for influent and effluent BMP TSS EMC under three different performance conditions.

ACKNOWLEDGEMENTS
The lead author greatly acknowledges the contribution of Robert Roseen, Jim Lenhart, and Greg Williams for their valuable comments and suggestions. This paper will be presented at the EWRI congress conference at Kansas City and the travel expense for the paper presentation will partially be supported by the ASCE/EWRI task force committee.

REFERENCES

American Society of Civil Engineers and U.S. Environmental Protection Agency (2002). *Urban Stormwater BMP Performance Monitoring, A Guidance Manual for meeting the National Stormwater BMP Database Requirements*, April, EPA-821-B-02-001.
http://www.bmpdatabase.org/docs/Urban/Stormwater/BMP/Performancemonitoring.pdf

ETV Verification Protocol (2002). *Stormwater Source Area Treatment Technologies*, Draft 4.1, March. http://www.epa.gov/etv/pubs/04_vp_stormwater.pdf

Guo, Q., G. England and E,C. Johnston (2007) Development of Certification Guidelines for Manufactured BMPs. Proceedings of EWRI conference, Hawaii, May 17-21.

Kayhanian, M., A. Singh, and S. Meyer (2002). " Impact of non-detects in water quality data on estimation of constituent mass loading." *Water Science and Technology,* 45(9), 219-225.

Kayhanian, M., A. Singh, C. Suverkropp, S. Borroum (2003). "The Impact of Annual Average Daily Traffic on Highway Runoff Pollutant Concentrations." *Journal of Environmental Engineering*, Vol. 129(11), 975-990.

Kayhanian, M., C. Suverkropp, A. Ruby, and K. Tsay (2007). "Characterization and Prediction of Highway Stormwater Pollutant Event Mean Concentrations." *Environmental Management*, 85(2), 279-295.

Lenhart, James (2007) "BMP Performance Expectation Functions – A Simple Method for Evaluating Stormwater Treatment BMP Performance Data." 9th Biennial Stormwater Research and Watershed Management Conference, Orlando, FL.

New Jersey Department of Environmental Protection (2006). *New Jersey Tier II Stormwater Requirements – Amendments to TARP Tier II Protocol*, Division of Science, Research and Technology, NJDEP, Trenton, New Jersey, January. http://www.state.nj.us/dep/dsr/bscit/NJStormwater_TierII.pdf

Minton, Gary (2005). Stormwater Treatment. Resource Planning Associates, Seattle, Washington, U.S.A.

Roseen, R. M., Ballestero, T. P., Houle, J. J., Avellaneda, P., Wildey, R., and Briggs, J. F. (2006). "Performance evaluations for a range of stormwater LID, conventional structural, and manufactured treatment strategies for parking lot runoff under varied mass loading conditions." *Transportation Research Record: Journal of the Transportation Research Board*(No. 1984), 135–147.

Shumway, R., R. Azari, and M. Kayhanian (2002). Statistical Approaches to Estimating Mean Water Quality Concentrations with Detection Limits. *Environmental Science and Technology*, Vol. 36, No. 15, 3345-3353.

Technology Acceptance Reciprocity Partnership (2003). *Protocol for Stormwater Best Management Practice Demonstrations*, Endorsed by California, Massachusetts, Maryland, New Jersey, Pennsylvania, and Virginia, Final Protocol August 2001, Updated July 2003.
http://www.state.nj.us/dep/dsr/bscit/StormwaterProtocol.pdf

USEPA, ASCE, Consultants, G., and District, U. D. a. F. C. (2009). "Urban Stormwater BMP Performance Monitoring." *EPA-821-B-02-001*, Washington DC.

Washington State Department of Ecology (2002). *Guidance for Evaluating Emerging Stormwater Treatment Technologies, Technology Assessment Protocol – Ecology (TAPE)*, October 2002 (Revised June 2004), Publication Number 02-10-037. http://www.ecy.wa.gov/biblio/0210037.html