Stormwater Control Measure Bypass Pollutant Concentrations Based On Storm Runoff Concentrations

SR-14-10; July 2014

Aaron Richter, EIT

City of Austin
Watershed Protection Department
Environmental Resource Management Division

Abstract

Stormwater runoff is recognized as a nonpoint source of pollution causing degradation to the receiving waters of the United States. Stormwater Control Measures (SCMs) are designed to capture stormwater runoff and treat the influent water to an acceptable level before it is released to receiving waters. The size of the SCM plays a large role in the load of pollution that reaches the receiving water because the size determines the amount of water that is captured for treatment and the threshold of stored volume at which point the runoff is routed around the SCM. The stormwater that is routed around the control is known as the bypass and the concentration of pollutants in the bypass can contribute greatly to degradation of the receiving waters. The concentration of the pollutants in the bypass must be determined to size SCMs appropriately to achieve target pollutant removal. Stormwater runoff was monitored in Austin, TX, for pollutant concentrations and the volume of runoff at different time intervals of runoff events. Piecewise regression was used to form a relationship between runoff measured in inches and the concentrations of chemical oxygen demand, Escherichia. coli, lead, total nitrogen, total phosphorus, total suspended solids, and zinc. For all pollutants, the piecewise regression showed a strong drop in concentration during the initial runoff from a storm followed by a slower decline in concentration. Thus, higher runoff capture volumes result in lower pollutant concentrations in the bypass which leads to SCMs sized to capture water based on the pollutant concentrations for bypass, rather than oversized to capture water based on the mean concentrations in runoff.

Introduction

Nonpoint source pollution from stormwater runoff is recognized as a cause for the degradation of receiving waters in the United States. In order to protect against the degradation of receiving waters within the Barton Springs Zone, which includes the Barton Springs segment of the Edwards Aquifer and the watersheds that contribute to recharge of the segment within the City of Austin jurisdiction, the Save Our Springs (SOS) Ordinance was adopted in Austin, TX, in 1992. This ordinance pertains to development and associated stormwater runoff within the Barton Springs Zone in the City of Austin jurisdiction. More specifically, the SOS Ordinance requires development within the Barton Springs Zone to achieve non-degradation (based on average annual loading) of the water quality in the Barton Springs Zone (City of Austin Land Development Code Chapter 25-8 Article 13). In order to comply with the

SOS Ordinance, the post-development average annual load for water quality pollutants listed in the ECM must be less than or equal to the average annual load pre-development for each pollutant.

The City of Austin's Environmental Criteria Manual (ECM) documents the steps necessary to show non-degradation of water quality under the SOS Ordinance (ECM Section 1.6.9). In 2013, City of Austin Watershed Protection Department (WPD) staff began the process of updating the ECM calculations and tables used to demonstrate compliance with the technical requirements of the SOS Ordinance. Prior to this time, the pollutant removal efficiency in percent (also known as efficiency ratio) was calculated for water quality pollutants listed in the ECM. The updated method outlined in the ECM considers only the average annual loads for water quality pollutants pre- and post-development instead of pollutant removal efficiency.

Part of the updated calculation in determining the total pollutant load from a developed site is the pollutant load in water that bypasses the stormwater control measures (SCMs), which are more commonly known as best management practices (BMPs). Determining the bypass pollutant load is simply a product of the pollutant concentration in the bypass and the volume of water that bypasses the SCM. In order to accurately calculate the pollutant bypass load, WPD staff wished to provide data-driven bypass concentrations for each pollutant to be considered in pre- and post-development average annual load comparisons.

In the initial stages of a storm, the concentration of pollutants in the stormwater runoff can be substantially higher than the concentration of pollutants later in the storm (Sansalone and Buchberger 1996, Larsen et al. 1998, Krebbs et al. 1999). This period of the storm which can contain a relatively large percentage of the total pollutants in stormwater runoff has been termed the 'first flush' (Gupta and Saul 1996, Larsen et al. 1998, Hager 2001). Rainfall and runoff data collected from 1984 to 1988 in the Austin area suggest that the pollutant loads removed in the first half inch of runoff can account for up to 50% of the total average storm load (City of Austin 1990).

A SCM captures the initial runoff during a storm and eventually reaches a volumetric capacity, at which time additional runoff bypasses the SCM. Thus, the SCM holds and treats the stormwater which contains the highest concentration of pollutants and bypasses the stormwater which should contain a lower pollutant concentration. The concentration of pollutants in the runoff that bypasses the SCM is thus dependent on the storage capacity of the SCM. Bypass pollutant concentrations for the duration of a runoff event are equal to or less than the runoff pollutant concentrations at the time when storage capacity of the SCM reaches its maximum and the bypass process begins. A conservative, simplifying assumption made by WPD staff is that bypass pollutant concentrations for the full bypass volume are constant and equivalent to the runoff pollutant concentrations at the time when the SCM reaches its maximum capacity or water quality volume.

WPD staff present equations which relate the concentrations of pollutants in the stormwater runoff to the amount of stormwater runoff measured in inches. A logical extension of this concept leads to an equation which relates the pollutant concentrations in the stormwater runoff, and thus the water that bypasses the SCM, to the storage capacity of a SCM measured in inches.

Methods

Stormwater runoff data used in this analysis spanned 592 storm events in Austin, TX, from 1984 to 2012. Water samples were collected from 56 different sites during the storm events. Some storm events occurred on only one site while other events occurred on multiple sites which lead to a total of 1,525 sampling events. Within each sampling event, runoff samples were collected at multiple time intervals to obtain pollutant concentrations and amount of runoff from a site in inches. The 56 sites ranged from 0-

97% impervious cover and spanned nine land uses which included commercial, commercial downtown, industrial, multi-family residential, single family residential, office, transportation, and mixed.

Pollutants collected from 1984 to 2012 include ammonia, chemical oxygen demand, copper, lead, nitrate/nitrite, organic carbon, total phosphorus, total Kjeldahl nitrogen, total suspended solids, and zinc. Total nitrogen was computed as the sum of nitrate/nitrite and total Kjeldahl nitrogen. Biochemical (biological) oxygen demand was initially collected in 1984 but was dropped from sampling in 2001. Cadmium and volatile suspended solids were added to the parameter list in 1993 and were collected through 2012. Fecal coliform and fecal streptococci were initially collected in 1984 but data collection for fecal streptococci stopped in 2001 and data collection for fecal coliform stopped in 2009. *Escherichia coli* (*E. coli*) replaced fecal coliform in 2009 but the last *E. coli* data was collected in 2011. Fecal coliform concentrations were transformed to *E. coli* concentrations through an equation developed to relate fecal coliform to *E. coli* in Austin stormwater (Richter 2013). Pollutants included in this analysis include chemical oxygen demand, *E. coli*, lead, total nitrogen, total phosphorus, total suspended solids, and zinc.

Regression is a common statistical method which employs the relationship between two (or more) quantitative variables so that the response variable can be predicted from the independent variable(s) (Kutner et al. 2005). Many applications of regression focus on estimating the change in the mean of the response variable according to the other variables. However, focusing on the mean of the response variable may lead to overestimates or underestimates of change in the response variable in heterogeneous distributions (Terrell et al. 1996, Cade et al. 1999). Mosteller and Tukey (1977) stated that it was possible to fit regressions to different portions of the response variable distribution and should be done to give a more complete description of the relationship between two or more variables.

"Regression quantiles" is a class of statistics developed in the 1970s by two econometricians that is an extension of the linear model for estimating change in all portions of the distribution of a response variable (Koenker and Bassett 1978). This class of statistics has become commonly known as quantile regression and is used to estimate a specified percentile of the distribution of the response variable conditional on a second variable (Cade et al. 1999, Koenker and Machado 1999). Put another way, a specified percentage of the response variable data points are less than or equal to the function developed using the quantile regression method. For example, using the 80^{th} percentile as the quantile will lead to a function for which 80% of the response variable data points are less than or equal to that function.

All runoff samples were aggregated into a single data set for each stormwater pollutant. Such a large combination of data can sometimes introduce heterogeneity into a data set which as noted above can lead to overestimates or underestimates of change in the response variable if only the mean of the response variable is analyzed (Terrell et al. 1996, Cade et al. 1999). This could be compounded by the combination of runoff samples from different land use types and impervious cover percentages. Thus, quantile regression was used to estimate changes of each pollutant throughout the distribution using the 5^{th} percentile to the 95^{th} percentile in 5% increments.

Prior to the quantile regression analysis, Box-Cox transformation analysis was performed for each pollutant to determine if there was a transformation of the pollutant that was appropriate for correcting the skewness of the error term distribution and nonlinearity of the data (Box and Cox 1964). In a Box-Cox transformation, λ is determined such that Y^{λ} is the most appropriate transformation of the response variable. For example, λ equal to 2 would indicate that the response variable should be squared. For all pollutants, the analysis determined λ to be close to 0 which would indicate that the most appropriate transformation would be to take the log of the response variable (Kutner et al. 2005). Thus, the log of each pollutant was used for regression analysis.

In addition, piecewise regression was performed on each pollutant with inches of runoff as the independent variable. Piecewise regression is used when different relationships are thought to exist between the response variable and independent variable at different ranges of the independent variable. The 'first flush' phenomena causes a large drop in concentration of a pollutant in the first few inches of a storm followed by a slow decrease in concentration of a pollutant throughout the remainder of the storm. Thus, piecewise regression was used to account for the 'first flush' phenomena. The value of the independent variable where the slope changes in a piecewise regression is known as the breakpoint. The value of the breakpoint was unknown and was estimated using the NLIN procedure in SAS9.2. The independent variables entered into the quantile regression analysis were the inches of runoff and the inches of runoff minus the breakpoint estimated from the NLIN procedure.

Results

The chemical oxygen demand (COD) breakpoint was 0.0053 inches and the R^2 value was 0.2339 for the general piecewise regression. Quantile regressions were performed on the COD data using the 0.0053 inches breakpoint for the 5^{th} to the 95^{th} percentile in 5% increments. Predictions for the intercepts and slopes of each quantile regression using COD as pollutant of interest can be calculated from model output in Appendix A. The COD concentration shared the same trend in all quantile regressions, where the concentration decreased rapidly until the breakpoint and then more slowly after the breakpoint.

The function used in the quantile regression was exponential so the slope of the line is related to the value of the pollutant. In a homogeneous distribution the slope of each function should be similar from percentile to percentile. However, if the variability in the data is not even over the distribution then the slope of the exponential functions would be different between percentiles. The higher variability at low values of runoff is another indicator for the 'first flush' phenomena. When pollutants have had a chance to build up between rain events, it is more likely that the concentration of that pollutant in stormwater will be very high at low levels of runoff until the pollutant has mostly been washed away. The pollutant concentration may not be very high at low values of runoff when the pollutant has not had time to build up between storms. The lower variability of each pollutant at high values of runoff would indicate that most of the pollutants have washed away before reaching high runoff values regardless of how much pollutant was available when the storm event began. In this case, if the variability in COD was higher at low values of runoff then the slope would become more negative but if the variability in COD was higher at high values of runoff then the slope would become more positive as the percentile increased.

For each 5% increase in percentile, the quantile regression intercept increased and the slope became more negative (Figure 1). The COD data has a higher variability in the lower values of runoff and the distribution is not homogeneous. This indicates that a normal least square regression which focuses on the mean would probably be a poor model for this data. A model which used the median or the mean of the data would underestimate the COD concentration trends in the higher COD data range.

It is important to characterize trends in the upper range of pollutant concentrations, but it is also important to recognize when abnormally high data points may be altering pollutant predictions. These abnormally high data points can be thought of as outliers to the data. Since it has been established that COD exhibits more variability in the lower ranges of runoff, the slope is expected to increase at a steady rate from percentile to percentile until the influence of outliers alters the slope more than previous percentile changes. For COD, the slope increased relatively evenly until the 90^{th} percentile was reached at which point the intercept and slope increased slightly more than was seen in previous percentile changes. The changes in intercept and slope were more drastic from the 90^{th} percentile to the 95^{th} percentile, which implies that the 95^{th} percentile was being affected by outliers.

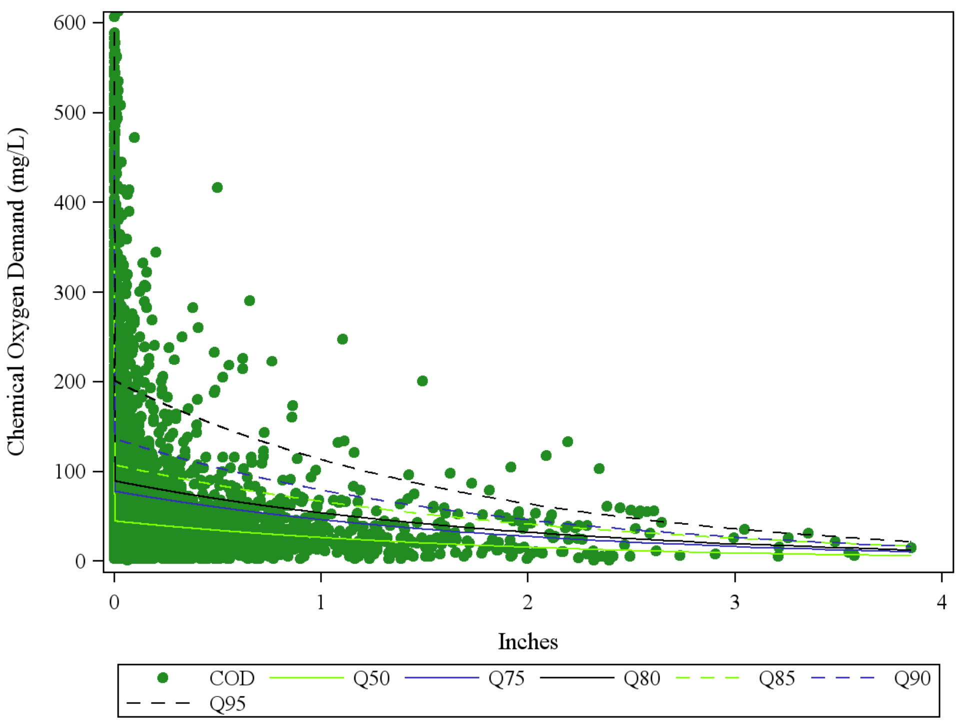

Figure 1: Chemical oxygen demand (mg/L) observations in all storm sampling events along with the quantile regression equations for the median (Q50), 75th (Q75), 80th (Q80), 85th (Q85), 90th (Q90), and 95th (Q95) percentile.

The *E. coli* breakpoint was 0.0091 inches and the R^2 value was 0.5336 for the general piecewise regression. Quantile regressions were performed on the *E. coli* data using the 0.0091 inches breakpoint for the 5th to the 95th percentile in 5% increments. Predictions for the intercepts and slopes of each quantile regression using *E. coli* as the pollutant of interest can be calculated from model output in Appendix B. Concentrations of *E. coli* decreased in every quantile regression after the inches breakpoint.

After the breakpoint, the quantile regression intercept increased and slope became more negative as the percentile of the quantile regression increased from the 40th percentile to the 95th percentile. The *E. coli* data also had a higher variability in the lower values of runoff and the distribution of *E. coli* was not homogeneous (Figure 2). As such, the slope was expected to become increasingly negative at a steady rate as the percentile increased for the quantile regression. No discernable pattern was detected in the change of slope from using the 5% incremental changes in percentile. This made it too difficult to determine where predictions of *E. coli* may be altered due to outlier data points and may be an indicator of the variability present in all ranges of *E. coli* data.

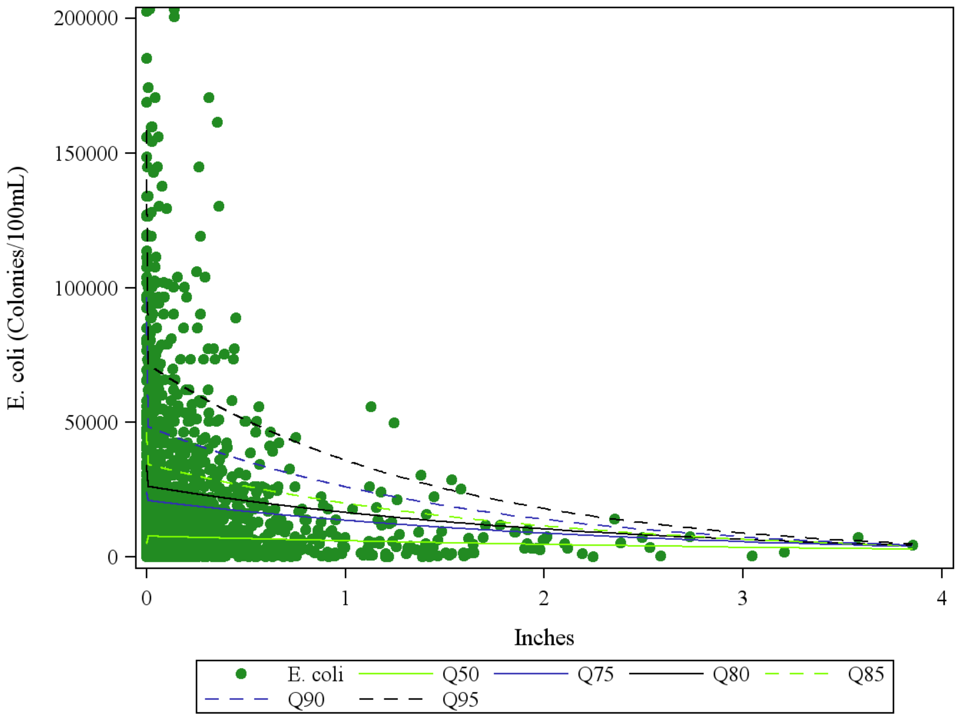

Figure 2: *E. coli* (colonies/100mL) observations in all storm sampling events along with the quantile regression equations for the median (Q50), 75th (Q75), 80th (Q80), 85th (Q85), 90th (Q90), and 95th (Q95) percentile.

The breakpoint for lead was 0.0140 inches and the R^2 value was 0.2124 for the general piecewise regression. Quantile regressions were performed on the lead data using the 0.0140 inches breakpoint for the 5th to the 95th percentile in 5% increments. Predictions for the intercepts and slopes of each quantile regression using lead as the pollutant of interest can be calculated from model output in Appendix C.

For lead, the slope of the regression line after the inches breakpoint was nearly zero until the 35th percentile was reached. The slope of the regression line after the inches breakpoint became more negative as the percentile modeled increased from the 35th percentile to the 95th percentile. Thus, lead data contained a higher variability in the lower value for inches of runoff (Figure 3). Disregarding the 5th to the 30th percentiles, the slope was expected to become increasingly negative at a steady rate as the percentile increased for quantile regression if a function was not being impacted by outlier data points. The slopes increased at a relatively steady rate for all percentiles. Only when comparing the slope change between the 90th and the 95th percentile functions did the change in slope appear to be slightly more than expected compared to other 5% percentile increments. It is difficult to truly determine if outlier data has impacted the 95th percentile function, but if the function has overestimated the slope change it appears to be only slightly overestimated due to outlier data.

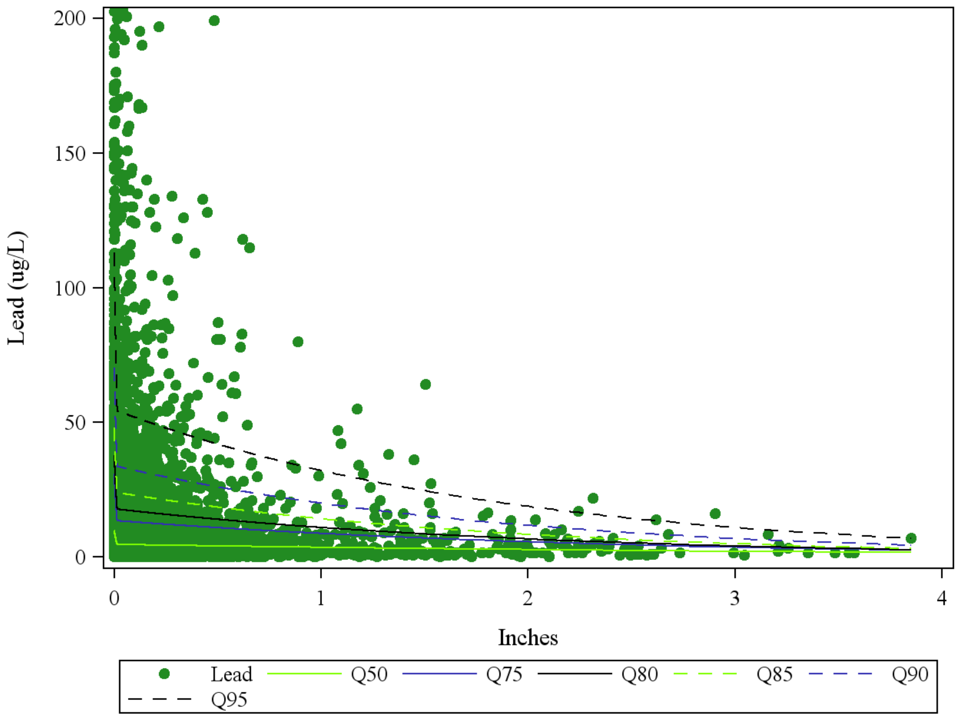

Figure 3: Lead (µg/L) observations in all storm sampling events along with the quantile regression equations for the median (Q50), 75th (Q75), 80th (Q80), 85th (Q85), 90th (Q90), and 95th (Q95) percentile.

The breakpoint for total nitrogen was 0.0145 inches and the R^2 value was 0.1905 for the general piecewise regression. Quantile regressions were performed on the nitrogen data using the 0.0145 inches breakpoint for the 5th to the 95th percentile in 5% increments. Predictions for the intercepts and slopes of each quantile regression using total nitrogen as the pollutant of interest can be calculated from model output in Appendix D.

The slope of regression lines after the 0.0145 inches breakpoint became more negative as the percentile of the regression function increased from the 5th percentile to the 95th percentile. Thus, the variability in total nitrogen data was higher for lower values of runoff. However, the slope changes between the percentile functions were not as large as slope changes for other pollutants, with the exception of total phosphorus. The lower change in slope is credited to the higher concentrations of nitrogen that exist in the mid- to high range of runoff instead of only occurring at low ranges of runoff (Figure 4).

For total nitrogen, the slope changes between regression functions increased negatively at a steady rate from the 10th percentile to the 90th percentile. Only when comparing the slope change between the 90th and the 95th percentile functions does the change in slope appear to be more than expected compared to other 5% percentile increments. The 95th percentile regression function does appear to be impacted by some outlier data points.

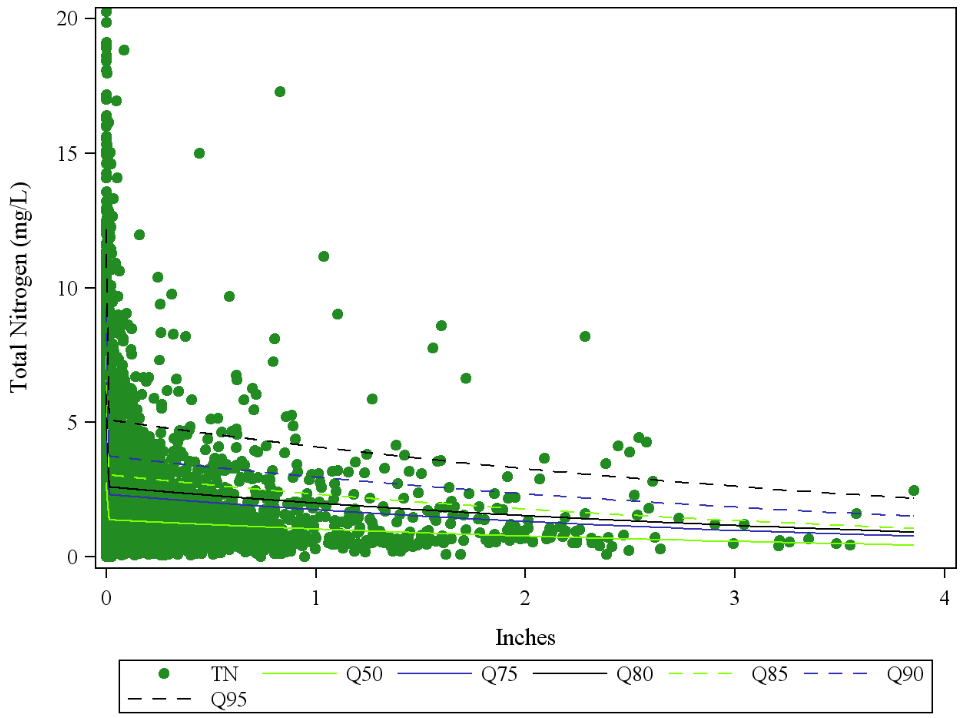

Figure 4: Total nitrogen (mg/L) observations in all storm sampling events along with the quantile regression equations for the median (Q50), 75th (Q75), 80th (Q80), 85th (Q85), 90th (Q90), and 95th (Q95) percentiles.

The breakpoint for total phosphorus was 0.00426 inches and the R^2 value was 0.1350 for the general piecewise regression. Quantile regressions were performed on the phosphorus data using the 0.00426 inches breakpoint for the 5th to the 95th percentile in 5% increments. Predictions for the intercepts and slopes of each quantile regression using total phosphorus as the pollutant of interest can be calculated from model output in Appendix E.

The slope of regression lines after the 0.00426 inches breakpoint became more negative as the percentile of the regression function increased from the 5th percentile to the 95th percentile. Thus, the variability in total phosphorus data was higher for lower values of runoff. Like total nitrogen, the slope changes between the percentile functions were not as large as slope changes for other pollutants. The lower change in slope is credited to the higher concentrations of phosphorus that exist in the mid- to high range of runoff instead of only occurring at low ranges of runoff (Figure 5).

The change in slope for the regression functions after the inches breakpoint was relatively constant between 5% increases in percentile with the exception of the increase from the 5th percentile to the 10th percentile. The function for predicted phosphorus using the 5th percentile was a constant with a zero slope so the change in slope from the 5th to the 10th percentile was large. No strange departure from the steady increase in slope change occurred at the higher percentiles. It did not appear that outlier data was impacting the regression functions for total phosphorus.

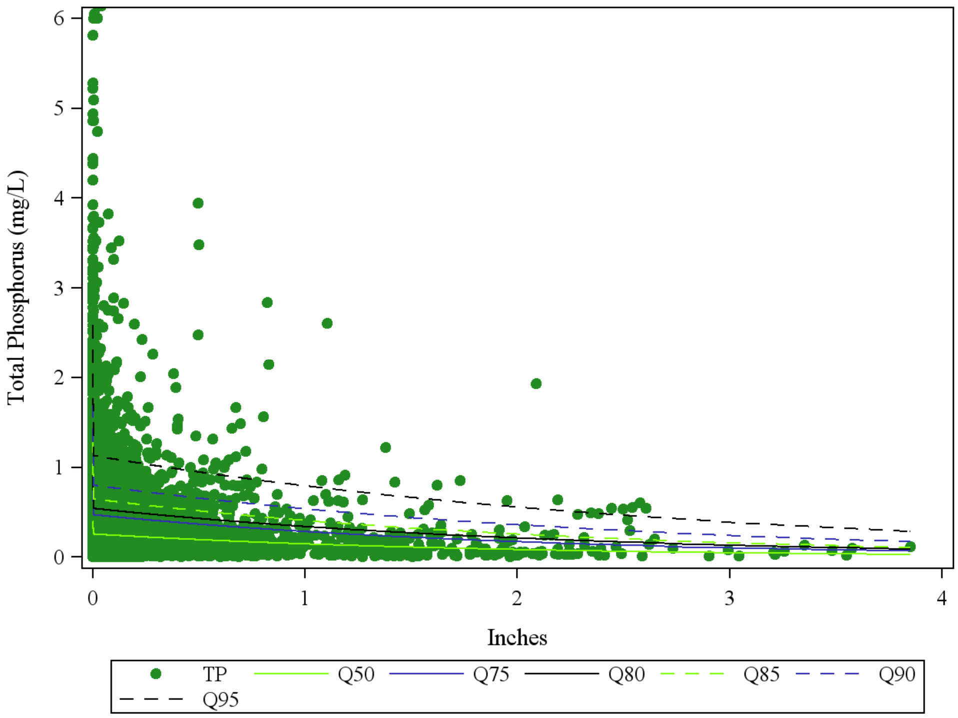

Figure 5: Total phosphorus (mg/L) observations in all storm sampling events along with the quantile regression equations for the median (Q50), 75th (Q75), 80th (Q80), 85th (Q85), 90th (Q90), and 95th (Q95) percentiles.

The breakpoint for total suspended solids was 0.00399 inches and the R^2 value was 0.1557 for the general piecewise regression. Quantile regressions were performed on the data using the 0.00399 inches breakpoint for the 5th to the 95th percentile in 5% increments. Predictions for the intercepts and slopes of each quantile regression using total suspended solids as the pollutant of interest can be calculated from model output in Appendix F.

The slope of regression lines after the 0.00399 inches breakpoint became more negative as the percentile of the regression function increased from the 20th percentile to the 95th percentile. There was no change of slope between the 5th percentile and the 20th percentile regression functions. Variability was higher in total suspended solids data at the lower values of runoff (Figure 6). There was no pattern in the change of slope between percentiles while comparing the percentiles at 5% increments. Impacts of any outlier total suspended solids data points on the upper percentile regression functions for total suspended solids was indeterminate.

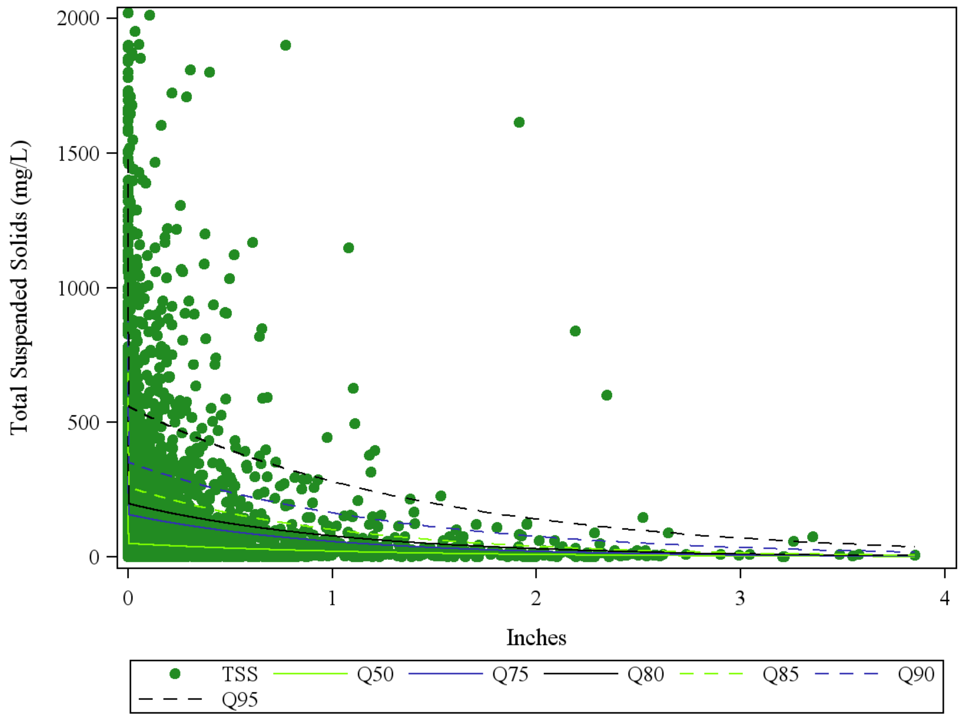

Figure 6: Total suspended solids (mg/L) in all storm sampling events along with the quantile regression equations for the median (Q50), 75th (Q75), 80th (Q80), 85th (Q85), 90th (Q90), and 95th (Q95) percentiles.

The breakpoint for zinc was 0.0147 inches and the R^2 value was 0.2391 for the general piecewise regression. Quantile regressions were performed on the data using the 0.0147 inches breakpoint for the 5th to the 95th percentile in 5% increments. Predictions for the intercepts and slopes of each quantile regression using zinc as the pollutant of interest can be calculated from model output in Appendix G.

The slope of regression lines after the 0.0147 inches breakpoint became more negative as the percentile of the regression function increased from the 20th percentile to the 95th percentile. Differences in slope had no pattern between the 5th percentile and the 20th percentile functions. Variability in zinc was higher for lower values of runoff (Figure 7). The change in slope for the regression functions after the inches breakpoint was relatively constant between 5% increases in percentile from the 50th percentile to the 80th percentile. Change in slope from the 80th to the 85th percentile, 85th to the 90th percentile, and 90th to the 95th percentile were much larger than was expected when comparing to changes of slope between more moderate percentiles like the 70th to the 75th percentiles. Thus, the 85th percentile, 90th percentile, and 95th percentile regression function may all be impacted by outlier data points.

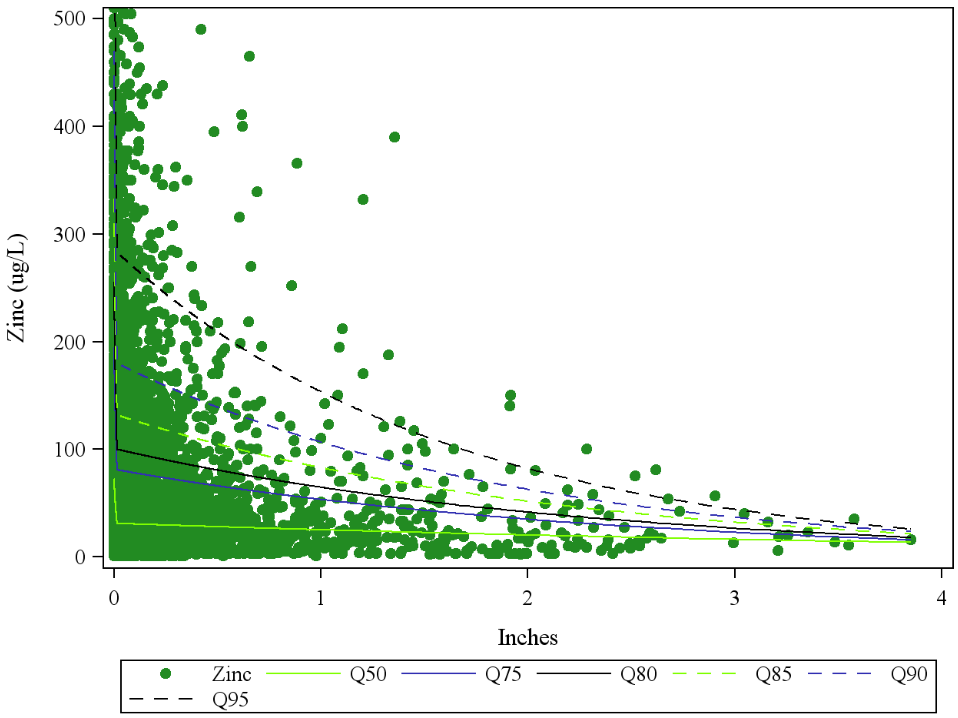

Figure 7: Zinc (µg/L) in all storm sampling events along with the quantile regression equations for the median (Q50), 75th (Q75), 80th (Q80), 85th (Q85), 90th (Q90), and 95th (Q95) percentiles.

Conclusions

The uneven distribution of variability throughout the data (increased at the lower ranges of runoff) confirmed that these data had heterogeneous distributions. A regression that focused on the mean of each pollutant would not appropriately estimate the change in the pollutants with respect to inches of runoff (Terrell et al. 1996, Cade et al. 1999). Regressions that focus on the mean are likely to have similar results to the median quantile regression and thus would underestimate the change in pollutants with respect to inches of runoff. The increased variability at low ranges of runoff also confirms that 'first flush' is relevant for each pollutant of concern assessed. This is less true for total nitrogen and total phosphorus when compared to the other pollutants of concern because the changes in slopes between the different regression percentiles were less for total nitrogen and total phosphorus. In fact, on certain land uses nutrients do not always adhere to the 'first flush' phenomenon (Glick 2009). However, because the land uses have been agglomerated for this analysis, the theory of 'first flush' can still applied to the nutrient pollutants assessed. It is thus appropriate to use these functions to predict a pollutant concentration at various rainfall runoff levels instead of using an event mean concentration for pollutants.

A goal of this analysis was to not only find the relationship between analyzed pollutants and the inches of runoff, but to recommend an equation relating the concentration of analyzed pollutants and inches of runoff that could be used to represent the amount of a pollutant running off a developed site location during a storm. This equation should be used to find the bypass concentration of stormwater runoff after a SCM captures a certain number of inches of runoff during a storm, thus helping to calculate an appropriately sized SCM to achieve regulatory load reduction requirements.

Developed sites can vary in design, geology, on-site vegetation, amount of impervious cover and other factors, which can all contribute to the concentration of pollutants in stormwater. Individual rainfall events can vary in rainfall intensity, antecedent rainfall conditions, and rainfall duration, which will also contribute to the concentration of pollutants in stormwater for any given event (Glick 2009). All of these conditions thus contribute to the variability inherent in the data set analyzed for this report. Developing a relationship between pollutant concentrations and inches of runoff that incorporated each of these conditions would be a complex and arduous task.

However, using an equation that focuses on high but not anomalous (outlier) concentration data points should alleviate the need to develop a relationship of concentration and runoff that incorporates site design, geology, impervious cover, antecendent rainfall, rainfall intensity, and other factors. This draws the focus to existing conditions which produce higher concentrations of pollutants in stormwater runoff while focusing on frequently occurring data which is likely to contribute to the average annual loading of stormwater pollutants. The higher percentiles in the quantile regression analysis would focus on the high concentration data points. Equations developed using the higher percentiles leads to SCMs sized appropriately to account for the variability in developed site runoff, without having to know the exact conditions on the site or in the rainfall event.

It is recommended that the equations for each pollutant matching the 80^{th} percentile regression be used as the appropriate relationship between pollutant concentration and runoff in inches of stormwater and thus the concentration for bypass flows of an SCM in the Austin area (Table 1). For each pollutant analyzed it appeared that the 80^{th} percentile regression equation was not impacted by any anomalous (outlier) data, while the 85^{th} percentile regression equation may have been impacted by anomalous data points for a few of the pollutants analyzed. The 80^{th} percentile regression equation should provide calculations that size SCMs for all site conditions that reach a level of treatment that maintain a non-degradation level of water quality. These calculations are not overly constrictive as the relationship is conservative but does not incorporate the extreme value concentrations and do not underestimate the change in the pollutants with respect to inches of runoff.

Table 1: The 80^{th} percentile regression equations for pollutants of concern based on inches of runoff. Only the piecewise equation after the breakpoint in inches is presented for each pollutant as this is the equation that is proposed for SCM sizing.

Pollutant	Breakpoint (inches)	Equation after breakpoint
Chemical Oxygen Demand (mg/L)	0.0053	$e\text{^}(4.4930 - 0.5104\ X\ (Inches))$
E. coli (colonies/100mL)	0.0091	$e\text{^}(10.1772 - 0.4647\ X\ (Inches))$
Lead (µg/L)	0.0140	$e\text{^}(2.8821 - 0.4893\ X\ (Inches))$
Total Nitrogen (mg/L)	0.0145	$e\text{^}(0.9574 - 0.2670\ X\ (Inches))$
Total Phosphorus (mg/L)	0.00426	$e\text{^}(-0.6134 - 0.4693\ X\ (Inches))$
Total Suspended Solids (mg/L)	0.00399	$e\text{^}(5.2895 - 0.9341\ X\ (Inches))$
Zinc (µg/L)	0.0147	$e\text{^}(4.6100 - 0.4424\ X\ (Inches))$

Appendix A: Intercept and slope predictions for chemical oxygen demand based on quantile regression models. Equations for a pollutant are given by Eq.1 if below the breakpoint in inches of runoff (0.0053) and Eq.2 if above or equal to the breakpoint in inches of runoff.

$$\text{Pollutant} = \exp(\text{intercept} + \beta1 \times \text{runoff}) \qquad \text{(Eq. 1)}$$

$$\text{Pollutant} = \exp(\text{intercept} + \beta1 \times \text{runoff} + \beta2 \times (\text{runoff} - \text{breakpoint}) \qquad \text{(Eq. 2)}$$

Quantile	Intercept	β1	β2
0.05	2.5664	-49.0131	48.4547
0.10	3.1781	-87.6950	87.2367
0.15	3.5264	-112.342	111.8437
0.20	3.7621	-131.057	130.5862
0.25	3.9042	-130.451	129.9844
0.30	4.0700	-135.667	135.2027
0.35	4.2471	-148.537	148.0567
0.40	4.4320	-159.657	159.1548
0.45	4.5539	-163.167	162.6616
0.50	4.7005	-170.704	170.1733
0.55	4.8437	-181.221	180.6881
0.60	4.9851	-187.809	187.2836
0.65	5.1648	-199.895	199.3804
0.70	5.2899	-200.337	199.7940
0.75	5.4424	-205.645	205.1221
0.80	5.6158	-212.362	211.8516
0.85	5.8621	-224.432	223.9545
0.90	6.1257	-228.771	228.2263
0.95	6.3790	-203.498	202.9239

Appendix B: Intercept and slope predictions for *E. coli* based on quantile regression models. Equations for a pollutant are given by Eq.1 if below the breakpoint in inches of runoff (0.0091) and Eq.2 if above or equal to the breakpoint in inches of runoff.

$$\text{Pollutant} = \exp(\text{intercept} + \beta 1 \times \text{runoff}) \quad \text{(Eq. 1)}$$

$$\text{Pollutant} = \exp(\text{intercept} + \beta 1 \times \text{runoff} + \beta 2 \times (\text{runoff} - \text{breakpoint}) \quad \text{(Eq. 2)}$$

Quantile	Intercept	β1	β2
0.05	5.0140	75.4472	-76.0149
0.10	5.6623	96.4184	-96.7861
0.15	6.1016	128.1468	-128.660
0.20	6.5437	119.1635	-119.593
0.25	6.9713	95.9314	-96.2416
0.30	7.3125	81.8089	-81.9752
0.35	7.6241	83.8825	-84.0948
0.40	7.8710	70.4692	-70.6051
0.45	8.1941	56.6645	-56.8565
0.50	8.4843	51.4894	-51.7427
0.55	8.8030	38.0262	-38.2672
0.60	9.1324	20.4885	-20.7731
0.65	9.3927	13.8670	-14.2080
0.70	9.7911	-8.1365	7.7905
0.75	10.0912	-15.6503	15.2165
0.80	10.4384	-29.1718	28.7071
0.85	10.7454	-33.1289	32.5798
0.90	11.4771	-75.9700	75.3459
0.95	11.9742	-87.7481	87.0553

Appendix C: Intercept and slope predictions for lead based on quantile regression models. Equations for a pollutant are given by Eq.1 if below the breakpoint in inches of runoff (0.0140) and Eq.2 if above or equal to the breakpoint in inches of runoff.

Pollutant = exp(intercept + β1 x runoff) (Eq. 1)

Pollutant = exp(intercept + β1 x runoff + β2 x (runoff – breakpoint) (Eq. 2)

Quantile	Intercept	β1	β2
0.05	0.0071	-171.324	171.8844
0.10	0.4055	-99.5859	99.7842
0.15	0.4164	-41.6822	41.7872
0.20	0.6591	-39.7510	39.7930
0.25	0.9457	-38.5891	38.5891
0.30	1.3836	-69.8655	69.8655
0.35	1.6270	-77.6932	77.6318
0.40	1.8600	-64.2528	64.0928
0.45	2.0836	-55.8761	55.6228
0.50	2.2828	-53.2402	52.9758
0.55	2.4857	-51.9536	51.6542
0.60	2.6741	-49.9408	49.6142
0.65	2.9167	-48.1848	47.7624
0.70	3.1718	-55.0316	54.6282
0.75	3.4340	-59.6670	59.2300
0.80	3.6322	-54.0678	53.5785
0.85	3.8730	-49.9684	49.4403
0.90	4.2534	-52.7676	52.2385
0.95	4.7276	-52.4311	51.8976

Appendix D: Intercept and slope predictions for total nitrogen based on quantile regression models. Equations for a pollutant are given by Eq.1 if below the breakpoint in inches of runoff (0.0145) and Eq.2 if above or equal to the breakpoint in inches of runoff.

$$\text{Pollutant} = \exp(\text{intercept} + \beta 1 \times \text{runoff}) \quad \text{(Eq. 1)}$$

$$\text{Pollutant} = \exp(\text{intercept} + \beta 1 \times \text{runoff} + \beta 2 \times (\text{runoff} - \text{breakpoint}) \quad \text{(Eq. 2)}$$

Quantile	Intercept	β1	β2
0.05	-0.9688	1.9770	-2.2613
0.10	-0.5006	-6.8395	6.5843
0.15	-0.1860	-15.9131	15.6796
0.20	0.0807	-24.4547	24.2073
0.25	0.2935	-31.6492	31.4147
0.30	0.4120	-32.3365	32.0771
0.35	0.5771	-36.7723	36.4992
0.40	0.7549	-42.2771	41.9982
0.45	0.8802	-44.3452	44.0453
0.50	1.0123	-47.8185	47.5257
0.55	1.1298	-49.0356	48.7251
0.60	1.2797	-52.5100	52.2177
0.65	1.3956	-54.2694	53.9727
0.70	1.5325	-56.2579	55.9464
0.75	1.6701	-57.3999	57.1153
0.80	1.8156	-59.4508	59.1838
0.85	2.0110	-61.9848	61.7132
0.90	2.2217	-62.2914	62.0550
0.95	2.4992	-60.2316	60.0098

Appendix E: Intercept and slope predictions for total phosphorus based on quantile regression models. Equations for a pollutant are given by Eq.1 if below the breakpoint in inches of runoff (0.00426) and Eq.2 if above or equal to the breakpoint in inches of runoff.

$$\text{Pollutant} = \exp(\text{intercept} + \beta 1 \times \text{runoff}) \qquad \text{(Eq. 1)}$$

$$\text{Pollutant} = \exp(\text{intercept} + \beta 1 \times \text{runoff} + \beta 2 \times (\text{runoff} - \text{breakpoint}) \qquad \text{(Eq. 2)}$$

Quantile	Intercept	β1	β2
0.05	-3.9120	0.0000	0.0000
0.10	-3.2189	79.3922	-79.9435
0.15	-2.6593	37.7567	-38.3440
0.20	-2.1201	-25.3691	24.8110
0.25	-1.7677	-63.6021	63.0442
0.30	-1.5606	-75.8182	75.2498
0.35	-1.3903	-82.0719	81.4956
0.40	-1.2385	-85.7663	85.1734
0.45	-1.0834	-95.0558	94.4920
0.50	-0.9163	-105.751	105.1869
0.55	-0.7443	-121.189	120.6471
0.60	-0.6351	-118.907	118.3785
0.65	-0.4780	-130.897	130.4305
0.70	-0.3023	-141.571	141.0810
0.75	-0.1776	-135.688	135.1795
0.80	0.0336	-152.341	151.8717
0.85	0.2700	-165.977	165.5107
0.90	0.5338	-178.080	177.6797
0.95	0.9524	-195.346	194.9909

Appendix F: Intercept and slope predictions for total suspended solids based on quantile regression models. Equations for a pollutant are given by Eq.1 if below the breakpoint in inches of runoff (0.00399) and Eq.2 if above or equal to the breakpoint in inches of runoff.

$$\text{Pollutant} = \exp(\text{intercept} + \beta1 \text{ x runoff}) \qquad \text{(Eq. 1)}$$

$$\text{Pollutant} = \exp(\text{intercept} + \beta1 \text{ x runoff } + \beta2 \text{ x (runoff} - \text{breakpoint)} \qquad \text{(Eq. 2)}$$

Quantile	Intercept	β1	β2
0.05	1.8941	-166.484	165.6315
0.10	2.5117	-196.172	195.5090
0.15	3.0969	-254.258	253.5241
0.20	3.5853	-291.388	290.5795
0.25	3.9101	-300.576	299.7646
0.30	4.2486	-313.585	312.7206
0.35	4.5554	-330.417	329.5918
0.40	4.8512	-350.939	350.1348
0.45	5.0557	-342.399	341.5977
0.50	5.2484	-336.199	335.3503
0.55	5.3982	-321.564	320.7310
0.60	5.5690	-302.882	301.9913
0.65	5.7654	-297.125	296.2261
0.70	5.9212	-279.806	278.8462
0.75	6.0765	-256.049	255.0378
0.80	6.3181	-258.741	257.8069
0.85	6.5280	-244.060	243.1202
0.90	6.8265	-242.449	241.6845
0.95	7.2984	-244.468	243.7781

Appendix G: Intercept and slope predictions for zinc based on quantile regression models. Equations for a pollutant are given by Eq.1 if below the breakpoint in inches of runoff (0.0147) and Eq.2 if above or equal to the breakpoint in inches of runoff.

$$\text{Pollutant} = \exp(\text{intercept} + \beta 1 \text{ x runoff}) \quad \text{(Eq. 1)}$$

$$\text{Pollutant} = \exp(\text{intercept} + \beta 1 \text{ x runoff } + \beta 2 \text{ x (runoff – breakpoint)} \quad \text{(Eq. 2)}$$

Quantile	Intercept	β1	β2
0.05	1.4268	-13.2610	13.1868
0.10	1.7933	0.0401	-0.3733
0.15	2.3634	-15.9429	15.6168
0.20	2.7935	-33.0240	32.9319
0.25	3.0991	-38.0203	37.8907
0.30	3.3676	-42.8169	42.6979
0.35	3.6482	-48.2604	48.1096
0.40	3.8901	-54.3085	54.1696
0.45	4.1109	-56.8202	56.6307
0.50	4.2793	-57.3969	57.1825
0.55	4.5028	-59.5125	59.2626
0.60	4.6913	-60.7108	60.4336
0.65	4.8512	-58.3895	58.0574
0.70	5.0556	-58.9194	58.5502
0.75	5.2730	-59.9410	59.5190
0.80	5.5439	-63.9761	63.5337
0.85	5.8073	-63.1234	62.6532
0.90	6.1534	-65.4251	64.8941
0.95	6.5988	-64.8203	64.1972

References

Box, GEP, and DR Cox. 1964. An Analysis of Transformations. Journal of the Royal Statistical Society Series B 26: 211-243.

Cade, BS, JW Terrell, and RL Schroeder. 1999. Estimating effects of limiting factors with regression quantiles. Oikos 91: 245-254.

City of Austin. 1990. The First Flush of Runoff and Its Effects On Control Structure Design. City of Austin, Watershed Protection Department, Environmental Resource Management, CM-90-01.

Glick, R. 2009. Stormwater Runoff Quality and Quantity from Small Watersheds in Austin, TX: Updated Through 2008. City of Austin, Watershed Protection Department, Environmental Resource Management, Water Quality Monitoring.

Gupta, K, and AJ Saul. 1996. Suspended Solids in Combined Sewer Flows. Water Science and Technology 33(9): 93-99.

Hager, M. 2001. Evaluating First Flush. Storm Water, the Journal for Surface Water Quality Professionals 2(6): 10-20.

Koenker, R, and GW Bassett. 1978. Regression Quantiles. Econometrica, 46: 33-50.

Koenker, R, and JAF Machado. 1999. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association 96: 458-468.

Krebs, P, P Holzer, JL Huisman, and W Rauch. 1999. First Flush of Dissolved Compounds. Water Science and Technology 39(9): 55-62.

Kutner, MH, CJ Nachtsheim, J Neter, and W Li. 2005. Applied Linear Statistical Models, 5th ed. McGraw-Hill Irwin, New York, NY.

Larsen, T, K Broch, and RM Andersen. 1998. First Flush Effects in an Urban Catchment Area in Aalborg. Water Science and Technology 37(1): 251-257.

Mosteller, F, and JW Tukey. 1977. Data analysis and regression. Addison-Wesley, New York, NY.

Richter, FA. 2013. Conversion of Fecal coliform to Escherichia coli Bacteria in Austin, TX. City of Austin, Watershed Protection Department, Environmental Resource Management, Water Resource Evaluation, DR-13-01.

Sansalone, JJ, and SG Buchberger. 1996. Characterization of Metals and Solids in Urban Highway Winter Snow and Spring Rainfall-Runoff. Transportation Research Record 1523, TRB, National Research Council, Washington, D.C. pp 147-159.

Terrell, JW, BS Cade, J Carpenter, and JM Thompson. 1996. Modeling stream fish habitat limitations from wedged-shaped patterns of variation in standing stock. Transactions of American Fisheries Society 125: 104-117.