

EXTERNAL MEMORANDUM

TO: Sadie Drescher, Chesapeake Bay Trust
FROM: Roxolana Kashuba, Melanie Edwards, and Sean Ryan, Exponent Inc.
DATE: March 6, 2020
PROJECT: 1901001.000/Restoration Research Award 16925
SUBJECT: Technical Memorandum #4; DRAFT BMP Evaluation Software and User Guide

The intent of this memorandum is to report and explain progress on the fourth phase of Exponent's FY2019 Restoration Research Program awarded contract. This phase of the project involved developing a preliminary version of BMP Evaluation Software, that allows users to interact with the results of the third phase of this project assessing the effect of sampling frequency and watershed characteristics on the uncertainty of pollutant load estimates.

The products delivered as part of this fourth project phase include:

- Best-fit predictive regression model results
 - Updated from Technical Memorandum #3
- BMP Evaluation Software implemented as a Shiny application¹ (files provided)
- User Guide for BMP Evaluation Tools (contained below)

1 Updated Predictive Regression Models

Technical Memorandum #3 reported results for regressions predicting the error in median load estimates for suspended sediment, total phosphorus, and nitrate from watershed characteristics for six different sampling methods. Upon re-coding these models in a form suitable for use in the BMP Evaluation Software, we updated² the best-fitting regression coefficients, shown in Tables 1–3 below, from the results presented in Technical Memorandum #3.

¹ Shiny (Chang, W., J. Cheng, J.J. Allaire, Y. Xie, and J. McPherson. 2018. Shiny: Web application framework for R. R package version 1.1.0. <https://CRAN.R-project.org/package=shiny>) is a package from RStudio that provides interactive web applications with R (R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>).

² These regressions were originally fit using generalized linear models that depend on starting values that can vary. Generalized linear models are ideal for handling complex error structures but unnecessary for the Gaussian error structure of the fitted models based on watershed characteristics. For repeatability and simplicity, we replaced the `glm()` function with the `lm()` function in the stepwise model-fitting procedure. Additionally, we corrected the log-transformation of the discharge predictor (which had been incorrectly log-transformed twice) which made minor changes to the fitted model for one sampling method each for suspended sediment and total phosphorus.

Table 1. Results of stepwise linear regression models for suspended sediment concentrations

Model Fit	Monthly		Weekly		Weekly+Storm		Seven-Hour		Flow-Paced		MDE MS4	
AIC	-12.19		28.76		125.78		111.72		98.28		77.91	
Adjusted R-squared	62.7%		69.3%		13.0%		35.7%		23.6%		3.5%	
Overall p-value	<0.0001		<0.0001		0.0094		<0.0001		0.0005		0.1168	
Residual standard error (df)	0.1972 (40)		0.3110 (39)		0.9661 (42)		0.8146 (41)		0.7067 (42)		0.5607 (42)	
Model Terms	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value
Intercept	0.08140	0.4594	-0.9875	0.0038	-2.265	0.0003	-0.7647	0.0054	1.148	0.1670	-1.115	0.0018
Log Watershed Size	-0.06871	0.0002	-0.1657	<0.0001	--	--	-0.2140	0.0002	--	--	--	--
Log Discharge	--	--	--	--	--	--	--	--	-0.1693	0.0005	--	--
Baseflow Index	--	--	2.253	0.0028	4.499	0.0094	--	--	--	--	1.537	0.1168
Flashiness Index	--	--	0.0203	0.0362	--	--	--	--	--	--	--	--
% Developed Low Intensity	-0.5668	0.0858	--	--	--	--	--	--	--	--	--	--
% Woody Wetlands	-14.30	<0.0001	-18.00	<0.0001	--	--	-16.98	0.0773	--	--	--	--

Table 2. Results of stepwise linear regression models for total phosphorus

Model Fit	Monthly		Weekly		Weekly+Storm		Seven-Hour		Flow-Paced		MDE MS4	
AIC	32.72		66.74		77.62		77.53		89.65		65.74	
Adjusted R-squared	27.0%		20.1%		30.6%		26.0%		5.6%		21.5%	
Overall p-value	0.0064		0.0208		0.0068		0.0028		0.1147		0.0284	
Residual standard error (df)	0.3914 (26)		0.7035 (26)		0.8363 (25)		0.8608 (27)		1.061 (27)		0.6813 (25)	
Model Terms	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value
Intercept	-2.190	<0.0001	-2.270	<0.0001	-4.208	<0.0001	-3.366	<0.0001	-3.450	<0.0001	-2.084	0.0002
Log Discharge	0.0431	0.1054	--	--	--	--	--	--	--	--	--	--
Baseflow Index	--	--	--	--	7.1545	0.0012	--	--	--	--	2.676	0.0780
Flashiness Index	--	--	--	--	0.0322	0.1610	--	--	--	--	--	--
% Developed High Intensity	--	--	--	--	--	--	--	--	--	--	-10.52	0.1413
% Developed Open Space	--	--	1.434	0.1742	--	--	4.099	0.0028	2.501	0.1150	--	--
% Woody Wetlands	-15.31	0.0020	-18.95	0.0212	-26.00	0.0142	--	--	--	--	-21.32	0.0130

Table 3. Results of stepwise linear regression models for nitrate

Model Fit	Monthly		Weekly		Weekly+Storm		Seven-Hour		Flow-Paced		MDE MS4	
AIC	47.38		64.80		71.91		84.52		65.47		42.75	
Adjusted R-squared	35.5%		32.2%		6.6%		0%		27.4%		30.8%	
Overall p-value	0.0048		0.0153		0.1257		<0.0001		0.02814		0.0097	
Residual standard error (df)	0.6108 (20)		0.8762 (19)		1.061 (21)		1.424 (22)		0.8890 (19)		0.5524 (20)	
Model Terms	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value
Intercept	-1.786	0.0003	0.4709	0.5445	-1.782	<0.0001	-2.463	<0.0001	-1.004	0.2228	-1.485	0.0004
Log Watershed Size	-0.1070	0.1039	-0.3309	0.0026	--	--	--	--	-0.1912	0.0908	-0.1285	0.0355
Baseflow Index	--	--	-1.517	0.2460	--	--	--	--	--	--	--	--
% Developed Medium Intensity	--	--	--	--	--	--	--	--	5.310	0.2024	--	--
% Developed High Intensity	--	--	--	--	--	--	--	--	--	--	11.42	0.1349
% Developed Open Space	2.516	0.0216	--	--	--	--	--	--	--	--	--	--
% Woody Wetlands	--	--	-14.80	0.1923	-18.24	0.1260	--	--	-26.83	0.0255	--	--

2 User Guide for BMP Evaluation Software

2.1 General Purpose of Software

The BMP Evaluation Software is designed to guide users to optimal and economical sampling strategies for detection of anticipated water quality improvements that result from best management practices (BMPs). The BMP Evaluation Software consists of two tools:

- (1) Educational Tool
- (2) Management Tool.

Both tools will dynamically update results based on changes to the user inputs.

2.1.1 Education Tool

Continuous measurement of pollutant concentrations in stormwater is both physically and financially impractical. Thus, it is important to understand how well a fixed number of water samples can estimate annual average concentrations in a watershed and whether the error in these estimates exceeds the expected pollutant load reduction that result from BMPs.

The Educational Tool allows users to select a watershed and adjust the number of samples collected within a year. This user-selected number of samples is then randomly drawn from 2017 and 2018 near-continuous data for the selected watershed monitoring site and used to estimate annual average pollutant loads. The error introduced from the reduced sampling is calculated by subtracting the sampled estimate from the “true” (full dataset) estimate. The user can adjust the sample size and observe the impact on the error in estimating annual average pollutant load.

2.1.2 Management Tool

The Management Tool compares six sampling methods with respect to median error in predicting annual pollutant load from watershed characteristics. The predicted annual load estimates, and thus the error estimates, incorporate uncertainty from measured data as well as sampling frequencies and the predictive relationships with key hydrologic and land-use properties of watersheds.

The user can input values for various watershed characteristics (Table 4, below) to understand how they may affect annual pollutant load estimates. The Management Tool shows the predicted estimate of median error in annual pollutant load (difference between predicted load and “true” load based on the complete measured data) for each sampling method (Table 5, below). The uncertainty in the estimates are characterized by 95% prediction limits. The error estimates and uncertainty ranges for each sampling method are compared to user-input percent expected annual reduction in pollutant load. Sampling methods with error ranges less than the expected load reduction will be sufficient to detect the anticipated reduction. However, sampling methods with error

ranges greater than the expected load reduction will not provide sufficient data to detect the anticipated load reduction because of variability in conditions, sampling frequency, and prediction error.

The Management Tool will highlight which sampling methods are able to detect the expected load reductions, the annual cost of those sampling methods, as well as the watershed characteristics predictive of sampling error for each sampling method. The Tool will update automatically as the user changes the input values for any of the watershed characteristics, expected load reductions, or cost per sample.

2.2 How to Install the BMP Evaluation Software

There are two ways to install the BMP Evaluation Software:

- (1) Run from within RStudio
 - using provided file: “BMP_APP_R.zip”, or
- (2) Install standalone desktop (Windows) application
 - using provided file: “BMP_APP_installer.zip”.

2.2.1 Run from within RStudio

Step 1. Install RStudio³ and R⁴.

Step 2. Download “BMP_APP_R.zip”, unzip file and extract all contents.

Step 3. Open the “BMP_APP_R” folder and double-click the file “BMP_Evaluation_shiny.RProject”.

Step 4. In RStudio, open the file “Run_BMP_App.R” from the “BMP_APP_R” folder.

Step 5. Install the “shiny” package by running lines 1–10 of the file “Run_BMP_App.R”. (This only needs to be done the first time you run the app.)

Step 6. Run line 10 or click the “Run App” button to launch the BMP Shiny App.

To view in new window, click the “Open in Browser” button.

Step 7. You can stop running the application by clicking the “Stop” button.

2.2.2 Install standalone desktop (Windows) application

This version does not require R, RStudio, or other programs to run.

³ RStudio Team (2019). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL <https://rstudio.com/>

⁴ This tool set was developed under R version 3.6.1 and may not function correctly under different versions of R.

- Step 1. Download “BMP_APP_installer.zip”, unzip file and extract all contents.
- Step 2. Open the “BMP_APP_installer” folder and double-click the file “setup_BMP_APP.exe”.
- Step 3. Read the license agreement, select the “I accept the agreement” radio button, and click Next if you accept.
- Step 4. Read the disclaimer information, and click Next if you accept.
- Step 5. Select the folder where you would like the BMP_APP to be installed, ensure you have the required disk space (~286 MB), and click Next.
- Step 6. Select the folder where you would like the BMP_APP Start Menu shortcut to be installed, and click Next.
- Step 7. Select whether you would like a BMP_APP desktop shortcut, and click Next.
- Step 8. Click Install.
- Step 9. Read the Information screen, and click Next.
- Step 10. If you would like to run the BMP_APP immediately after installation, keep the “Launch BMP_APP” box checked. If not, uncheck this box. Click Finish.

There are three ways to launch the BMP_APP after installation:

- (1) Double-click the BMP_APP desktop shortcut, if you chose to install it.
- (2) Double-click BMP_APP from the Start Menu location you selected during installation.
- (3) Double-click BMP_APP.bat (Windows Batch File) in BMP_APP folder.

To uninstall the BMP_APP, double-click the file “unins000.exe” in the BMP_APP folder.

2.3 How to use the BMP Evaluation Software

There are two tools contained in the BMP Evaluation Software:

- (1) Educational Tool
- (2) Management Tool

Select which Tool you would like use by clicking on either the Educational Tool tab or the Management Tool tab at the top of the screen. The user can select inputs for each tool separately using the drop-down menus in the left sidebar. The last tab, Software Information, contains the following details about the software version, developers, and disclaimers:

BMP Evaluation Software

Version: 1.0.0

Date: March 6, 2020

Authors: Josh Thompson (Watershed Restoration and Protection Program in Annapolis), Roxolana Kashuba (Exponent, Inc.), Melanie Edwards (Exponent, Inc.), and Sean F. Ryan (Exponent, Inc.)

Disclaimer: The user acknowledges that R and RStudio are open source platforms and that the user covenants that it will comply with any and all license requirements included with the R and RStudio packages, including but not limited to, those found at <https://www.r-project.org/Licenses/> and <https://rstudio.com/about/opensourcelicenses/>. Subject to the foregoing, permission is hereby granted, free of charge, to any user obtaining a copy of this software tool and associated documentation files (the "Software Tool"), to deal in the Software Tool without restriction, including without limitation the rights to use, copy, modify, merge, distribute, and/or sublicense copies of the Software Tool, and to permit persons to whom the Software Tool is furnished to do so, subject to the following conditions: THIS SOFTWARE TOOL IS PROVIDED "AS IS", WITHOUT REPRESENTATIONS AND WARRANTY OF ANY KIND, EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS BE REQUIRED TO PROVIDE ONGOING CODEBASE SUPPORT, UPKEEP, BUG FIXES OR HOSTING AFTER DELIVERY. IN NO EVENT SHALL THE AUTHORS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE TOOL OR THE USE OR OTHER DEALINGS IN THE SOFTWARE TOOL. THE RESPONSIBILITY FOR APPLICATION OF THE OUTPUT FROM THE SOFTWARE TOOL RESIDES WITH THE USER.

Built using: R 3.6.1 and R packages shiny, tidyverse, RColorBrewer, DT, shinydashboard, ggplot2, zoo, hydroTSM, stats, utils, ggpubr, shinycssloaders, and openxlsx.

Description: This program is designed to evaluate how sampling method (interval) affects the ability to detect the efficacy of Best Management Practices (BMPs) in reducing annual pollutant loads and identify the most economical method. Users can input watershed characteristics, expected annual load reductions resulting from BMPs, and estimated cost-per-sample. These inputs are used to calculate, from previously fitted models, the 95th percentile prediction intervals for the median estimated percent error in annual load estimates for each sampling method and total cost. The tool will then highlight which methods are able to detect expected pollutant reductions and rank them by cost.

Funding: This project was funded by the Pooled Monitoring Initiative led by the Chesapeake Bay Trust with the support from Maryland Department of Transportation State Highway Administration, the Maryland Department of Natural Resources, and the National Fish and Wildlife Foundation through the U.S. Environmental Protection Agency's Chesapeake Bay Program Office.

2.3.1 How to use the Educational Tool

Select the Educational Tool tab at the top of the screen. Open the Educational Tool menu on the left side of the screen. (The menu will not open automatically after you select the tab, and needs to be opened independently.)

2.3.1.1 Educational Tool Inputs

The Educational Tool lets the user select the values of two inputs:

- (1) *USGS High-Frequency Monitoring Site*: Select USGS high-frequency monitoring site from the drop-down menu.
- (2) *Sample size (per year)*: Enter number of water quality samples per year in the provided text field.

2.3.1.2 Educational Tool Outputs

The Educational Tool randomly sub-samples the user-selected number of samples per year from the high-frequency monitoring data for two years (2017 and 2018) for the user-selected monitoring site, calculates the error in simulated pollutant load estimates (relative to the complete data), and provides two types of outputs:

- (1) *Simulated sampling based on sample size (top set of plots)*: Plots for each pollutant (nitrate, total phosphorus, and suspended sediment concentration) showing the sub-samples of the user-selected sample size over time, compared to the discharge of the high-frequency dataset from which the sub-samples were randomly drawn.
- (2) *Effect of sample size on the error of load estimates (bottom set of plots)*: Boxplots showing the distribution of error in annual load estimates for the user-selected sample size for each pollutant and year simulated. Results are based on 100 simulations of the user-selected sample size.

Users can change the sample size and observe how the pollutant time series, and the error rate in estimating load for each pollutant, changes as a function of sample size. The intent is to allow the user to better visualize how changing annual sample size influences the representativeness of sampling events over time and the magnitude of error in annual load estimation.

2.3.2 How to use the Management Tool

Select the Management Tool tab at the top of the screen. Open the Management Tool menu on the left side of the screen. (The menu will not open automatically after you select the tab, and needs to be opened independently.)

2.3.2.1 Management Tool Inputs

The Management Tool lets the user enter the values of three categories of input variables in the provided text fields (scroll to the bottom of the page to access all values):

- (1) *Watershed Characteristics*: Enter ten different qualities of a watershed that may be related to annual load estimation error due to sampling method. Table 1 lists these ten watershed characteristics and the range of values for each characteristic in the dataset used to develop the regression models implemented in this Tool. If the user enters a value outside this range for any input variable, that cell will turn red and a message will pop up instructing the user to select a value within the fitted dataset range (for example, if the user enters “90” for “Percent impervious surface”, the message displayed over the red cell will read: “Value must be less than or equal to 50.”). The Tool will still calculate results; however, these results are highly uncertain as the inputs are beyond the calibration range of the model.
- (2) *Expected Annual Load Reduction*: Enter the percent annual load reduction expected from implemented BMPs or other watershed management activities, for suspended sediment, total phosphorus, and nitrate loads, to determine whether this load reduction is detectable via different sampling methods given error rates in load estimation simulated for each method.
- (3) *Monitoring Costs*: Enter analytical cost per sample to calculate and compare annual cost of each monitoring program.

Values for watershed characteristics default to median values from the complete dataset.

Values for expected annual load reduction default to 5%, as evaluated in Hypothesis 1.

Value for analytical cost per sample defaults to \$35.⁵

⁵ <https://www.abe.iastate.edu/soupir/files/2016/11/Water-Quality-Research-Laboratory-Method-Price-List.pdf>

Table 4. Descriptions and ranges of watershed characteristics

Watershed Characteristic	Description	Minimum Value	Maximum Value
Watershed Size (square miles)	Area of the watershed.	0.14	6252
Discharge (cubic meters per year)	Cumulative annual volume of water moving through watershed.	2.45E+05	1.02E+10
Flashiness Index (unitless ratio)	Standard hydrological metric for gauging the flashiness of a watershed: flow rate (m ³ /s) that is exceeded 5% of the time divided the flow rate that is exceeded 95% of the time.	2.2	53.9
Baseflow Index (%)	Modeled baseflow totals (contribution from groundwater and interflow, etc.) in a given year divided by the total flow. ⁶	0.093	0.67
% Woody Wetlands	Percent of watershed area where forest or shrubland vegetation accounts for greater than 20% of vegetative cover and the soil or substrate is periodically saturated with or covered with water.	0	5.3
% Impervious Surface	Percent of watershed area composed of materials that water does not permeate through (e.g., concrete, asphalt, etc.).	0.33	50
% Developed Open Space	Percent of watershed area with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20% of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.	2.1	35
% Developed Low Intensity	Percent of watershed area with a mixture of constructed materials and vegetation. Impervious surfaces account for 20–49% percent of total cover. These areas most commonly include single-family housing units.	0.4	37
% Developed Medium Intensity	Percent of watershed area with a mixture of constructed materials and vegetation. Impervious surfaces account for 50–79% of the total cover. These areas most commonly include single-family housing units.	0.022	62
% Developed High Intensity	Percent of watershed area with highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80–100% of the total cover.	0	8

⁶ Calculated using a digital recursive filter following standard hydrological methods; <https://www.bgs.ac.uk/research/groundwater/waterResources/thames/BFI.html>

2.3.2.2 Management Tool Outputs

For each pollutant and sampling method, the Management Tool calculates 95th percentile prediction intervals on median percent error in annual load estimates predicted from regression models with watershed characteristics (shown in Tables 1–3), and the annual cost of each method. The Tool compares the percent error in estimates of pollutant load to the expected percent reduction in pollutant load as a result of BMPs, and ranks sampling methods based on their ability to detect expected reductions and cost to do so. Table 5 describes the six sampling methods evaluated.

Every time the user sets values (or uses default values) for Watershed Characteristics, Expected Annual Load Reductions, or Monitoring Costs, the Tool automatically updates outputs. There are three types of outputs shown on three tabs for each pollutant:

- (1) *Plot tab*: Plots showing, for each sampling method (x-axis), the 95th percentile prediction intervals on median percent error in annual load estimates (y-axis), calculated from best-fit regression with watershed characteristics. These percent error intervals attributable to sampling method and regression prediction are compared to expected percent reduction in annual load attributable to BMP implementation (horizontal red dashed line) to determine whether that percent reduction in annual load is detectable by each sampling method.
- (2) *Parameters tab*: Chart showing which watershed characteristics (rows) were predictive of median percent error in annual load estimates for each sampling method (columns) in best-fit linear regression models (shown in Tables 1–3).
- (3) *Summary tab*: Table showing results of sampling method optimization evaluation sorted and ranked by sampling methods (rows) able to detect expected reduction in annual pollutant load for the lowest cost. In the middle column, sampling methods are designated as green, yellow, or red based on whether expected load reduction is always, sometimes, or never detectable (i.e., load estimation error is less than, overlapping with, or greater than expected load reduction). Annual cost of each sampling method is calculated based on the number of samples required for each method per year (shown in Table 5) multiplied by the analytical cost per sample entered by the user.

The user can change any of the inputs to evaluate the impact of different values on recommended sampling programs.

Regressions reported in Tables 1–3 are used to calculate the error range estimates shown in the Management Tool; however, the software is built such that the underlying coefficients can be updated easily should additional data or revised regression models be appropriate in the future.

2.4 Advanced Options

Since the BMP Evaluation Software was created using free, open-source R code, the underlying regression models, sampling intervals, user-selected inputs, and outputs can all be amended by editing the Software's source code.

Linear model objects (“.Rdata” files) are located in the BMP_APP\Input folder and currently represent regression fits reported in Tables 1–3, but can be replaced with new linear model objects representing alternate regression model fits for use in calculating prediction intervals reported in the Management Tool.

Number of annual samples for each sampling method is currently set to the values reported in Table 5 for the purposes of monitoring program cost calculation. Number of annual samples can be amended by editing the “Sampling_interval” cells in the file “Sampling_interval_by_method.csv” located in the BMP_APP\Fixed_params folder.

The BMP_APP\Rscripts folder contains the functions used to process user-selected values and generate outputs. The advanced user can explore and change these functions by editing the “.R” script files. However, great caution should be taken when editing these files as they may cause the application to no longer function correctly or at all.

Table 5. Descriptions of sampling methods assessed

Sampling Method	Number of Annual Samples Analyzed	Simulation Description
Monthly	12	Sample taken every month (7 am–6.30 pm, M–F)
Weekly	52	Sample taken every week (7 am–6.30 pm, M–F)
Weekly + Storm	72 ^a	Sample taken every week (7 am–6.30 pm, M–F) with a daily sample taken when flow >90 th percentile.
Seven-Hour	1,248	Sample taken every 7 hours.
Flow-Paced	52	Sample taken when flow exceeds cumulative threshold (threshold set to yield an average 80 sampling pumps per week and aggregated into weekly composite ⁷).
MDE MS4 Permittee Requirements	12	Samples taken every month with preference given to flow >90 th percentile storm events.

MDE – Maryland Department of the Environment; MS4 - NPDES Municipal Separate Storm Sewer System

^aThis is based on the average number of annual samples collected in the 1,000 simulations of Weekly + Storm sampling conducted for this analysis.

3 Discussion and Conclusions

The Management Tool shows high load estimation error rates for most Watershed Characteristic inputs, and no one sampling method consistently has the smallest estimation error. This is likely because the load estimation error calculated by the Management Tool is a function of both sampling uncertainty (e.g., the ability of 12 monthly samples per year to represent annual

⁷ <https://www.sciencedirect.com/science/article/abs/pii/S0925857418303537>

loading for that year), and regression prediction uncertainty (the ability of each set of best-fit watershed characteristics to accurately predict load estimation error).

Median load estimation error from sampling uncertainty alone was as high as 252% (for suspended sediment from Weekly + Storm sampling), while the highest median error due to sampling uncertainty was 86.4% for total phosphorus and 78.7% for nitrate. Table 6 shows the full ranges of median error due to sampling uncertainty for all pollutant and sampling method combinations. These error rates were the dependent variable in the multiple linear regressions relating load estimation error to watershed characteristics (reported in Tables 1–3).

Table 6. Percent median error ranges for each pollutant and sampling method due to sampling uncertainty alone

Sampling Method	Pollutant	Median Load Error Range Due to Sampling Uncertainty (%)
Monthly	Suspended Sediment ^a	26.7–91.6
	Total Phosphorus ^b	6.8–33.8
	Nitrate ^c	5.8–78.7
Weekly	Suspended Sediment	10.5–106
	Total Phosphorus	2.4–66.2
	Nitrate	1.8–74.2
Weekly + Storm	Suspended Sediment	3.8–252
	Total Phosphorus	1.3–86.4
	Nitrate	0.5–56.8
Seven-Hour	Suspended Sediment	1.5–79.7
	Total Phosphorus	1.6–54.4
	Nitrate	0.7–67.9
Flow-Paced	Suspended Sediment	2.60–92.2
	Total Phosphorus	0.6–43.2
	Nitrate	1.7–45.7
MDE MS4	Suspended Sediment	17.7–215
	Total Phosphorus	3.3–70.7
	Nitrate	5.6–61.3

MDE – Maryland Department of the Environment; MS4 - NPDES Municipal Separate Storm Sewer System

^a Simulated using 1,000 iterations each from 44 watershed/year combinations.

^b Simulated using 1,000 iterations each from 29 watershed/year combinations.

^c Simulated using 1,000 iterations each from 23 watershed/year combinations.

The second source of the large error predictive intervals estimated by the Management Tool is the poor predictive ability of watershed characteristics in explaining load estimation error due to sampling uncertainty. As shown in Tables 1–3, best-fit model adjusted r-squared values ranged

from 3.5 to 69.3% for suspended sediment, 5.6 to 30.6% for total phosphorus, and 0.0 to 35.5% for nitrate. This means the regression models did not explain a lot of the variability in absolute load error, resulting in large prediction intervals accounting for all that uncertainty.

Figure 1 shows an example of how poorly baseflow index explains load estimation error in suspended sediment from Weekly + Storm sampling. At a baseflow index value of 0.5, the best-fit linear regression model predicted an absolute percent error rate of 157% (Figure 1 is shown in log-scale, with a fit of -0.015). However, due to the large amount of variability in log(error) unexplained by the best-fit regression model, 95th percentile prediction intervals for load estimation error ranged from 20.3 to 1215% absolute error. These are the intervals reported by the Management Tool.

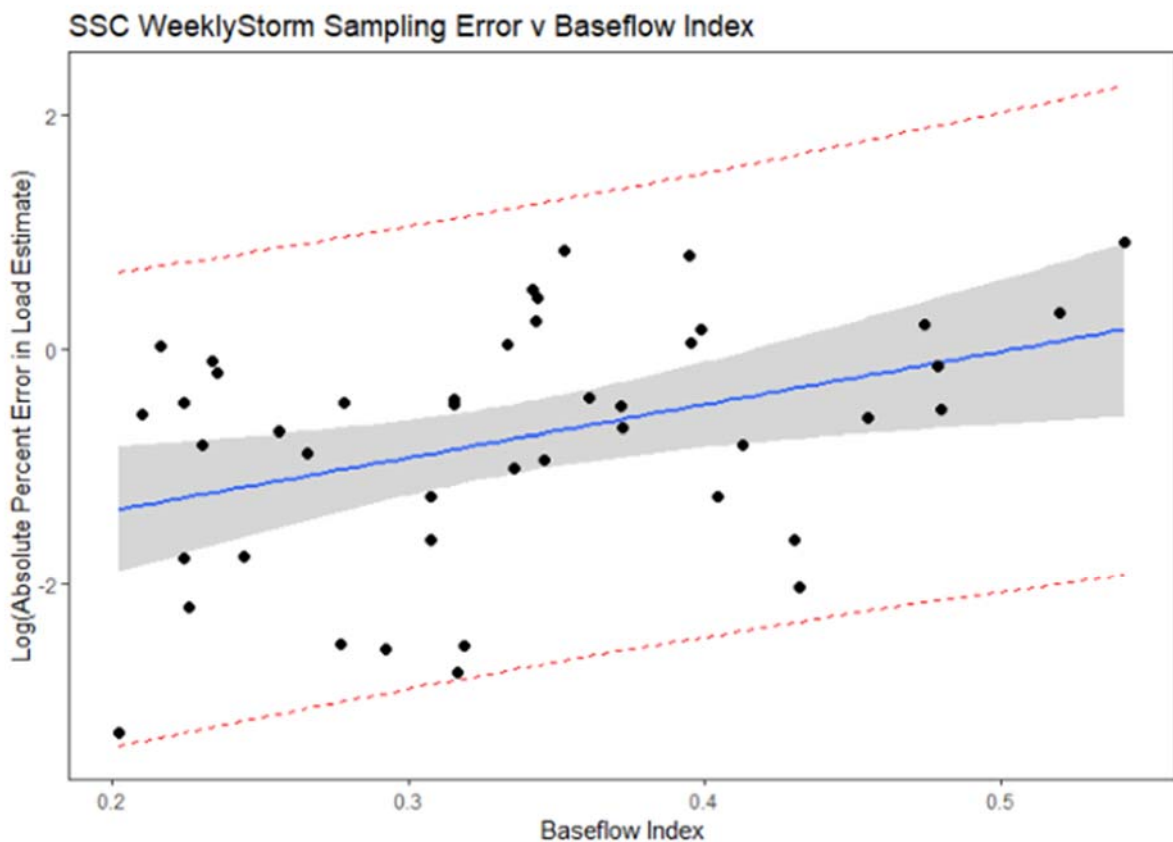


Figure 1. Relationship between Baseflow Index and log-transformed load estimation error for suspended sediment concentration (SSC) Weekly + Storm sampling.

Additionally, as shown in Tables 1–3 and in the Parameters tabs of the Management Tool, not every watershed characteristic is predictive of error rates for every pollutant/sampling method combination, as estimated through the best-fit regression models described in Technical Memo #3. As such, some error ranges shown in the Management Tool do not update upon changes to the values of watershed characteristics that are not predictive of error rates for those particular pollutant/sampling method combinations. For example, percent impervious surface was not a significant predictor of load estimation error in any regression model.

4 Approach for Next Deliverable

The next steps include review and exploration of this BMP Evaluation Software by Anne Arundel County and the Chesapeake Bay Trust and compilation of comments and suggested improvements and amendments. The Tools will then be revised as a final project product.

5 Summary

This Phase IV tool development created an interface for stakeholders to explore and engage with relationships between sampling frequency and precision and accuracy of sediment, total phosphorus, and nitrate load estimates for different watershed characteristics.

Percent load estimation errors can be very large due to both high sampling uncertainty and high regression prediction uncertainty when trying to explain error patterns from watershed characteristics. A possible simplified version of the BMP Evaluation Software may include removal or recategorization of the regression uncertainty error to more specifically focus on just sampling uncertainty error and its magnitude relative to expected or desired load reduction detection. Regression coefficients can be replaced, re-formatted, or removed to better achieve Management Tool goals.

Feedback from stakeholders regarding usefulness and functionality of the BMP Evaluation Software will be very helpful in deciding how to amend it to best meet stakeholder needs, elucidate the effect of sampling uncertainty in detection and quantification of BMP impacts, and inform monitoring design and evaluation going forward.