Exponent®

E X T E R N A L   M E M O R A N D U M

TO:	Sadie Drescher, Chesapeake Bay Trust
FROM:	Joshua Thompson, Exponent Inc.
DATE:	October 31, 2019
PROJECT:	1901001.000/Restoration Research Award 16925
SUBJECT:	Report #1; Data Deliverable

Exponent was awarded a contract for the FY2019 Restoration Research Program referenced above. The intent of this memorandum is to explain how, as part of this award, data were collected, aggregated, and processed and to explain the rationale for the inclusion and exclusion of data (or sites) and the assumptions made when estimating annual pollutant loads. This memorandum will also provide a "road map" to describe the statistical analyses that will be performed on the formatted and processed data, which will form the second deliverable. Applying these analyses will test the three hypotheses stated in Exponent's proposal (dated February 20, 2019) and clarified in Exponent's response to reviewer comments (dated March 27, 2019). Specifically, the hypotheses to be tested are:

1. An increase in best management practice (BMP)-treated area by 0.6% of a watershed area reduces pollutant export by 5%.
2. Stream water sampling at seven-hour frequencies using an automated sampler is sufficient to reduce maximum load estimate error rates to 15%.
3. The uncertainty of pollutant loads estimated at different sampling frequencies significantly differs with watershed size, land use, area of impervious surface, rainfall, and hydrology.

1 Methods - Data Collection

Hydrological, chemical, and geospatial data were aggregated from the sources documented in Table 1. Municipal Separate Storm Sewer System (MS4) permittee hydrological data were requested from each jurisdiction directly. Although most jurisdictions provided data, hydrological data from Charles County were not provided.[1] U.S. Geological Survey (USGS)

[1] Although Charles County shared chemistry data, their consultants were not able to provide the associated hydrological data.

hydrological and chemistry data were retrieved directly from their online data repository using the R package "dataRetrieval."[2,3]

Land cover data was retrieved directly from the Multi-Resolution Land Characteristics Consortium (MRLCC). To delineate each watershed, topographic data was retrieved from the National Hydrography Dataset (NHD) Plus Version 2. BMP information was obtained from the Maryland Department of Environment's (MDE) StormwaterPrint Geodatabase. Raw data used in these analyses were provided as part of the October Interim Progress Report, submitted to the Trust on October 11, 2019.

Table 1: Sources of compiled data for Restoration Research Award number 16925

Data	*Source*	*Contact(s)/Source*	*Date Contacted*	*Data Provided?*
Baltimore City MS4 hydrology data	Permittee and USGS	Robert McAulay; Robert.Mcaulay@baltimorecity.gov	08/27/2019	Yes; instantaneous stream discharge received for two watersheds.
Anne Arundel County MS4 hydrology data	Permittee	Chris Victoria; pwvict16@aacounty.org Janis Markusic; pwmark02@aacounty.org	08/27/2019	Yes; instantaneous stream discharge received for two watersheds.
Carroll County MS4 hydrology data	Permittee	Robert Flora-Nakoski; rfloranakoski@carrollcountymd.gov Byron Madigan; madigan@carrollcountymd.gov	08/27/2019	Yes; instantaneous stream discharge received for two watersheds.
Frederick County MS4 hydrology data	Permittee	Don Dorsey; DDorsey1@FrederickCountyMD.gov Shannon Moore; SMoore@FrederickCountyMD.gov	08/27/2019	Yes, although only received annual mean discharge for two sites.

2 R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

3 De Cicco, L.A., Hirsch, R.M., Lorenz, D., Watkins, W.D., 2018, dataRetrieval: R packages for discovering and retrieving water data available from Federal hydrologic web services, doi:10.5066/P9X4L3GE

Data	***Source***	***Contact(s)/Source***	***Date Contacted***	***Data Provided?***
Charles County MS4 hydrology data	Permittee	Karen Wiggen; WiggenK@charlescountymd.gov	08/27/2019	No. Exponent was referred to the consulting firm who collected the data, but despite several attempts data were not shared.
StormwaterPrint Geodatabase and MS4 chemical data	Permittee & Maryland Department of Environment	Jeff White; jeff.white@maryland.gov https://mdewin64.mde.state.md.us/SSDS/SWP/index.html	02/06/2019	Yes; water chemistry, event duration, event total discharge, BMP information.
Impervious surface and land cover data	MRLCC	https://www.mrlc.gov/	N/A	Yes
Topographical data	NHD Plus Version 2 (U.S. EPA and USGS)	https://www.epa.gov/waterdata/nhdplus-national-hydrography-dataset-plus	N/A	Yes
USGS hydrological and chemistry data	Retrieved from USGS using R package “dataRetrieval”	https://github.com/USGS-R/dataRetrieval	N/A	Yes

2 Methods – Data Processing

2.1 Hydrological data

Quality assurance was conducted after receiving hydrological data from the permittees and the USGS (sources described in Table 1).[4] Data were first plotted to identify periods of missing data and to identify obvious errors in the data (i.e., sensor drift, negative discharge values, or periods when the stream froze). Erroneous data points resulting from sensor malfunction or frozen stream water were not considered scientifically credible and were consequently removed from the dataset.[5]

---

4 It was assumed in most cases that some level of quality assurance was also conducted by the consulting firm or jurisdiction collecting the data. For the USGS data, as all data go through a thorough quality assurance, all data was also checked visually as an additional validation.

5 Only data points that were obviously the result of sensor errors were removed. Data were removed if values remained static for greater than 48-hours and were obviously not the result of low flows or if flows were

Following a review of data credibility, the completeness of each dataset was evaluated on a calendar year basis, as the percentage of actual observations out of the number total possible observations for each calendar year (e.g., for a stream monitored at 15-minutes intervals, the completeness would be calculated as the number of actual observations out of a possible 35,040). Calendar years with less than 90% data completeness were removed and were not considered for future analysis.

2.2 Chemistry data

As stated in Exponent's proposal, the Interstate Commission on the Potomac River Basin (ICPRB) was retained by MDE to aggregate the MS4 permittee chemistry data, and these data have undergone a partial quality assurance/quality control process.[6] After receiving the database, Exponent conducted additional quality assurance on the chemistry data for the MS4 sites stated in Table 1. Following U.S. Environmental Protection Agency guidance,[7] values below the analytical detection limit for a given parameter were replaced with half of their reported detection limit. This approach was selected as such values composed less than 15% of all observations. No supplemental quality assurance was applied to USGS data after an evaluation of their rigorous quality assurance process, which was considered adequate for the purposes of this analysis.

2.3 Load estimation

Following quality assurance, loads and flow-weighted mean concentrations (FWMC) were computed from reported concentration and discharge measurements. Traditional load estimation methods were applied to compute loads for each individual year.[8] Other methods, such as weighted regression techniques used by USGS, require long-term data collected at regular intervals, a condition that was not met by these data.[9] The Paris Commission load estimation

---

negative. If data contained outliers which were scientifically credible, then data were retained for further analysis to prevent "cherry-picking."

6 ICPRB. 2018. Analysis of Monitoring Data Collected under Maryland's Municipal Separate Storm Sewer System (MS4) Permits: Database Design and Preliminary Analysis of Water Chemistry. Interstate Commission on the Potomac River Basin. Rockville, MD. pp. 1–30.

7 U.S. EPA. 2000. Guidance for Data Quality Assessment: Practical Methods for Data Analysis, EPA QA/G-9, QA00.

8 Thompson, J., R. Cassidy, D.G. Doody, and R. Flynn. 2014. Assessing suspended sediment dynamics in relation to ecological thresholds and sampling strategies in two Irish headwater catchments. Science of the Total Environment, 468, pp.345–357.

9 Hirsch, R.M., and L.A. De Cicco. 2015. User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data (version 2.0, February 2015). pp. 1–93. In: U.S. Geological Survey Techniques and Methods Book 4.

algorithm was selected for consistently being the most accurate load estimation method.[10] This method estimates the load (mass of a pollutant exported over a calendar year), L_E, as:

$$L_E = \frac{K \sum_{i=1}^{n}(C_i Q_i)}{\sum_{i=1}^{n} Q_i} \overline{Q_r}$$

(Eq. 1)

Where K is a constant, which is the duration of the record in days, C_i is the instantaneous concentration measured within each grab sample (mass per volume), Q_i is the corresponding instantaneous discharge measurement (volume per unit time), and $\overline{Q_r}$ is mean discharge value over the period of observation (volume per unit time).

Permittee chemical data reported to MDE contains both baseflow and stormflow samples. Baseflow samples are simply instantaneous concentrations measured in a grab sample (i.e., direct measurements of C_i in equation 1), whereas stormflow samples are event mean concentrations (EMCs), which are flow-weighted mean concentrations of three grab samples, C_i, taken during a storm event. An event mean concentration (mass per volume), EMC, is calculated over the duration of the storm event as:

$$EMC = \frac{\sum_{i=1}^{n}(C_i Q_i)}{\sum_{i=1}^{n} Q_i}$$

(Eq. 2)

Where C_i is the instantaneous concentration measured in each grab sample during the storm event (mass per volume) and Q_i is the corresponding instantaneous discharge measurement during the storm event (volume per unit time).

In nearly all cases, raw data used to calculate EMCs (i.e., C_i and Q_i values in equation 2) were not provided by the MS4 permittees. To calculate an annual load using both baseflow and stormflow samples collected by the MS4 permittees, each EMC (mass per volume) was treated as equivalent to C_i in equation 1. To estimate the corresponding Q_i required in equation 1, the total discharge reported for each EMC (total volume) was divided by the duration of the monitored event (total time) to get the corresponding discharge value as a volume per unit time. In cases where the total discharge for a given storm was not reported to MDE, the total discharge for the duration of the storm was summed from the raw discharge data. This summed discharge (total volume) was then divided by the storm duration (total time) to get the discharge value as a volume per unit time. Data were then used within Equation 1 to estimate annual loads

10 Cassidy, R., and P. Jordan. 2011. Limitations of instantaneous water quality sampling in surface-water catchments: comparison with near-continuous phosphorus time-series data. Journal of Hydrology 405(1-2):182–193.

for the MS4 monitoring sites. Data collected by the USGS were all reported as instantaneous values (i.e., not as EMCs), and no adjustments were required for their use within equation 1.

Annual flow-weighted mean concentrations, $FWMC$, were then computed as:

$$FWMC = \frac{L_E}{\int_{t_{start}}^{t_{end}} Q_i(t)dt}$$

(Eq. 3)

Where L_E is the estimated annual pollutant load, and total annual discharge is calculated over the calendar year (from t_{end} to t_{start}).

2.4 High-frequency data

High-frequency in-situ measurements (defined as measurement collected at sub-hourly intervals) of nitrate and turbidity, and the corresponding discharge measurements were also retrieved from the USGS. For each monitoring site with high-frequency data, the associated water chemistry data, measured in grab samples, were also retrieved from the USGS.

Using both the high-frequency data and associated water chemistry data, for each monitoring site, turbidity was assessed as a surrogate for suspended sediment concentration (SSC) and total phosphorus by performing a regression of the in-situ turbidity measurement and the corresponding SSC or total phosphorus concentration. Lab measured SSC and total phosphorus concentrations were paired with in-situ turbidity measurements by selecting the turbidity measurement that was closest in time to the water chemistry measurements.[11] For each watershed, a robust regression model was applied to the paired data and model coefficients were retained for future use for the next deliverable. Robust regression was performed using the "robustbase" package in R.[12]

2.5 Geospatial data

Watersheds were delineated using a 25- by 25-m digital elevation model, with topographical data retrieved from the NHD Plus Version 2 dataset. The latitude and longitude of the monitoring locations, where stream discharge and chemistry were measured, were provided by MDE and the USGS and were used as the locations of the watershed outlets.

Once watersheds were delineated, land cover and impervious surface spatial area was summed for the years of data made publicly available by the MRLCC: 2001, 2006, 2011, and 2016. To

[11] Specifically, measurements that were closest in time within a 14-minute window. Many turbidity observations were collected at 15-minute intervals, and therefore, a 14-minute window limited the occurrence of multiple matching pairs.

[12] Maechler, M., P. Rousseeuw, C. Croux, V. Todorov, A. Ruckstuhl, M. Salibian-Barrera, T. Verbeke, M. Koller, E.L.T. Conceicao, and M.A. di Palma. 2019. robustbase: Basic Robust Statistics R package version 0.93-5. http://CRAN.R-project.org/package=robustbase

develop a timeseries, it was assumed that data for the year 2001 was representative of the years before 2001 and up to 2005, data for the year 2006 was representative of the years 2006–2010, data for the year 2011 was representative of the years 2011–2015, and data for the year 2016 was representative of the years 2016–2018. Because most land cover data do not extend further back than 2001, this was considered the preferred approach, rather than integrating older land cover data from other sources. Data from other sources do not have comparable land cover classes and use different methods to quantify land cover.

3 Approach for Next Deliverable

Many of the years that were presented as "complete" in a report by the ICPRB to MDE were not complete.[13] Consequently, per the quality assurance protocol described in Section 2.1, several years of data were removed due to incomplete (<90%) discharge records.

As such, the intended future analysis cannot be conducted as originally stated and must be amended. As stated in Exponent's proposal, data were intended to be analyzed using a linear effect model, by regressing the annual pollutant load over time, while also accounting for the changing BMP-treated area and impervious surface over time and the watershed where the data originate from. However, because of the incomplete record, and therefore unequal time intervals, regressing loads over time would produce erroneous results. Consequently, the temporal component will be removed from the analysis, and our methods will therefore be consistent with those of Reisinger et al. 2019.[14] Although our previous criticisms of this approach still apply, by incorporating this analysis within a linear mixed-effects model, the nested structure of pollutant loads and FWMC within watersheds will be addressed using fixed and random effects within one single fitted model analysis

Consequently, the future statistical analysis will be simplified, and using a linear mixed-effects model, pollutant load and/or pollutant FWMC will be regressed against BMP-treated area and impervious surface percentage, controlling for watershed location by including it as a random factor:

$$y_{ij} = \beta_1 x_{1ij} + \beta_2 x_{2ij} + b_{i1} z_{1ij} + \varepsilon_{ij}$$

(Eq. 4)

where y_{ij} is the value of the pollutant load or FWMC for a particular ij case, β_1 and β_2 are the fixed-effect coefficients for BMP-treated area and impervious surface percentage, x_{1ij} and x_{2ij} are the fixed-effect variables for BMP-treated area and impervious surface percentage for

[13] As stated in Exponent's proposal, the ICPRB was retained by MDE to aggregate these data, and all data in the database have undergone a partial quality assurance/quality control process.

[14] Reisinger, A.J., E. Woytowitz, E. Majcher, E.J. Rosi, K.T. Belt, J.M. Duncan, S.S. Kaushal, and P.M. Groffman. 2019. Changes in long-term water quality of Baltimore streams are associated with both gray and green infrastructure. Limnology and Oceanography, 64, pp. S60–S76.

observation j in group i, b_{i1} is the watershed random-effect coefficient, z_{1ij} is the watershed random-effect variable, and ε_{ij} is the error for case j in group i.

This linear mixed-effect model will be conducted in three stages. Initially, the analysis will be conducted on the MS4 data alone to understand whether the current data collection methods required by the MDE are sufficient to detect a statistically significant effect of BMP-treated area on annual pollutant loads and FWMCs. After this initial analysis, MS4 data will be combined with USGS data, with the analysis repeated to understand whether additional data from other Maryland watersheds with contrasting land cover improve the explanatory power of the model variables.

Finally, if significant relationships between water quality and BMP-treated area are observed, the raw data used to calculate loads and FWMC in Equations 1 and 3 will be standardized to 12 samples a year and then iteratively thinned at intervals from 90 to 10%. For each level of thinning, loads will be recalculated with a reduced dataset, and the linear mixed-effects analysis repeated to understand the minimum level of information required to detect an effect of BMP-treated area on water quality.

The road map for future analysis of the high-frequency data is unchanged from Exponent’s proposal to the Chesapeake Bay Trust. Coefficients derived from the robust regression model between turbidity and total phosphorus and turbidity and SSC will be used to transform the near-continuous turbidity data into near-continuous total phosphorus and SSC data. Data will then be thinned to assess the effect of sampling frequency on the uncertainty of load and FWMC estimates using the sampling methods defined in Table 2 and load estimation algorithm described in Equation 1. An example R script with example data is provided in the folder “Example Sampling Analysis” to demonstrate how this analysis will be conducted.

Table 2: Descriptions of sampling methods to be assessed

Sampling Method	No. Permutations	Description
Monthly	1,000	Sample taken every month (7 am–6.30 pm, M–F)
Weekly	1,000	Sample taken every week (7 am–6.30 pm, M–F)
Weekly + Storm	1,000	Sample taken every week (7 am–6.30 pm, M–F) with a daily sample taken when flow >10th percentile.
Seven-Hour	1,000	Sample taken every 7 hours.
Flow-Paced	600	Sample taken when flow exceeds cumulative threshold (threshold set to yield an average 80 pumps per week and aggregated into weekly composite).
MDE MS4 Permittee Requirements	1,000	Samples taken from 12 storms with monthly samples taken during episodes of extended low flow.

4 Summary

Although most stream chemistry data passed the quality assurance review, it should be noted that there were many instances where hydrological data were inadequate. Consequently, there were several years where data could not be used to estimate annual loads and FWMCs. As hydrological data are not required to be reported to MDE, it appears that MDE has a limited ability to assess whether EMCs reported by permittees are realistic or based on viable hydrological data. Exponent's review of permittee hydrological data revealed that in several cases, data contain significant gaps and suffer from sensor errors, either due to drift or due to battery failure.

Another limitation of the current reporting format is that EMCs are reported to MDE based on loads estimated for single storm events throughout the year. This format complicates the estimation of *annual* loads. Although assessing the water quality performance of BMPs on a storm-by-storm basis may provide useful insight, assessing the effect of BMPs on annual loads and annual flow-weighted mean concentrations over time is a more suitable approach for assessing the aggregated effect of BMP installation over time. BMP performance on a storm-by-storm basis is likely to be variable and, therefore, a less accurate assessment of BMP performance than using loads and flow-weighted mean concentrations on an annual basis.

After reviewing available data, Exponent has aggregated all relevant data, and data are now fit for analysis as outlined in the proposal and summarized in Section 3 of this memorandum. Data are provided as an attachment to this memorandum, which contains metadata that will assist the Chesapeake Bay Trust in understanding the data, data definitions, and, together with this memorandum, how the data were compiled.