



Regression modeling of combined sewer overflows to assess system performance

Matthew A. Bizer ^{a,*} and Christine J. Kirchoff ^{a,b}

^a Department of Civil & Environmental Engineering, University of Connecticut, 261 Glenbrook Rd, Unit 3037, Storrs, CT 06269-3037, USA

^b Current address: Department of Civil & Environmental Engineering and School of Engineering Design and Innovation, Penn State University, 213 Hammond, University Park, PA 16802, USA

*Corresponding author. E-mail: matthew.bizer@uconn.edu

 MAB, 0000-0002-1587-7740; CJK, 0000-0002-2686-6764

ABSTRACT

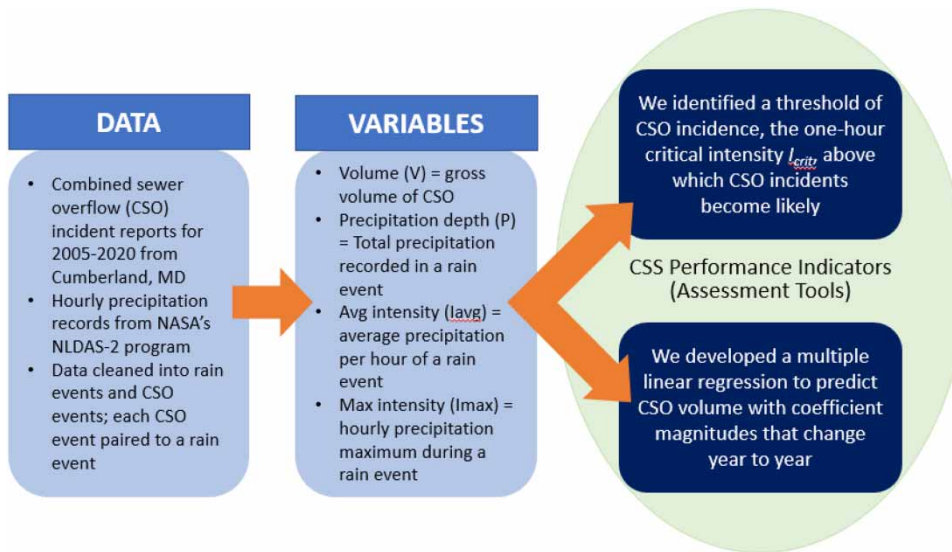
Combined sewer overflows (CSOs) occur when untreated raw sewage mixed with rainwater, runoff, or snowmelt is released during or after a storm in any community with a combined sewer system (CSS). Climate change makes CSOs worse in many locales; as the frequency and severity of wet weather events increases, so do the frequency and volume of CSO events. CSOs pose risks to humans and the environment, and as such, CSS communities are under regulatory pressure to reduce CSOs. Yet, CSS communities lack the tools needed, such as performance indicators, to assess CSS performance. Using the city of Cumberland, Maryland as a case study, we use public data on CSOs and precipitation over a span of 16 years to identify a new critical rainfall intensity threshold that triggers likely CSO incidence, and a multiple linear regression model to predict CSO volume using rainfall event characteristics. Together, this indicator and modeling approach can help CSS communities assess the performance of their CSS over time, especially to evaluate the effectiveness of efforts to reduce CSOs.

Key words: combined sewer overflow (CSO), combined sewer system (CSS), extreme rainfall events, performance indicators, stormwater

HIGHLIGHTS

- Rainfall events in two watersheds are sorted by single-hour maximum intensity (mm/h) to identify a CSO incidence threshold.
- The critical intensity threshold (I_{crit}), a novel indicator of CSS performance independent of rainfall measurements, is developed.
- A multiple linear regression model is developed that predicts CSO volumes from the depth and average intensity of the associated rainfall event.
- The two-pronged approach of I_{crit} and the multiple regression model serve as new performance indicators to assess CSS performance under changing precipitation, and relatively unchanging conditions of urbanization and population.

GRAPHICAL ABSTRACT



INTRODUCTION

Climate changes including more frequent and extreme wet weather events (Easterling *et al.* 2017) stress aging wastewater and stormwater infrastructure systems designed for the climate of the past (Rudberg *et al.* 2012; Schoen *et al.* 2015). These stresses are particularly acute in combined sewer systems (CSS), in which stormwater is collected in the same pipes as domestic and/or industrial wastewater. Thousands of CSOs occur each year in the United States (US Environmental Protection Agency 2022a), Europe (Environmental Agency 2020; Scottish Water 2020), and elsewhere around the world. Changing precipitation patterns may lead to more severe and more frequent incidents of combined sewer overflow (CSO), releases of untreated raw sewage mixed with stormwater from CSS (Astarai-Imani *et al.* 2012; Fortier & Mailhot 2015; Mahaut & Andrieu 2019; Botturi *et al.* 2021). Though climate change is forecast to reduce rainfall events and wet extremes in some areas (Christidis & Stott 2022), much of North America (Easterling *et al.* 2017) and northern Europe (Christidis & Stott 2022) are expected to face more rainfall and worsening wet extremes. In these regions, urbanizing areas with combined sewers are particularly vulnerable, as increasing urbanization, absent mitigation of additional flow volumes, also increases the frequency and volume of CSO discharges (Semadeni-Davies *et al.* 2008; Abdellatif *et al.* 2015). Indeed, emerging research on predicting CSO occurrence suggests CSO events will worsen in urban areas under climate change (Semadeni-Davies *et al.* 2008; Schroeder *et al.* 2011; Astarai-Imani *et al.* 2012; Sandoval *et al.* 2013; Fortier & Mailhot 2015; Yu *et al.* 2018).

When CSOs harm a community or surrounding environment, government regulators may respond by taking enforcement actions requiring CSS municipalities to reduce CSO discharges or otherwise act to improve water quality. In the United States, this typically takes the form of a consent decree from the US Environmental Protection Agency (USEPA) ordering the reduction or elimination of CSO discharge volumes (US Environmental Protection Agency 1994). Some U.S. municipalities end CSO discharges altogether by separating their sewer systems from their stormwater collection systems, while other municipalities, particularly those who serve historically marginalized communities, cannot afford to separate their systems. Existing methods for assessing CSS performance relative to planned CSO reductions rely upon complex hydraulic modeling approaches that are not easily transferable and yield results with wide uncertainty ranges (Andrés-Doménech *et al.* 2012; Bachmann-Machnik *et al.* 2021; Roseboro *et al.* 2021). Yet, simpler approaches to CSS assessment with lower uncertainties often have trade-offs, such as ignoring CSO magnitudes entirely and instead measuring only water quality (Madoux-Humery *et al.* 2013; Romei *et al.* 2021). To understand the effects of climate change on CSO discharges, rather than measuring water quality, it is simpler to assess the frequency and volume of CSO discharges. Using the city of Cumberland, Maryland as a case study, we build on prior research to develop a simple yet transferable statistical model to predict the occurrence and volume of CSO incidents and propose a new performance indicator of critical intensity (I_{crit}) based on CSO occurrence and rainfall intensity to assess CSS performance.

Schroeder *et al.* (2011) proposed a simple statistical approach to predict whether a CSO will occur from precipitation data. They identified a critical rainfall depth (N_{crit}), the total rainfall amount (in mm) of an antecedent rainfall event, above which a CSO was probable ($\geq 85\%$) to occur and below which a CSO was improbable ($\leq 17\%$) to occur, for each of four watersheds in Berlin, Germany, in 2007. Expanding on this work, Fortier & Mailhot (2015) identified N_{crit} and fit a sigmoid probability function for CSO incidence on the antecedent rainfall depth for outfalls in Quebec, Canada. Both Schroeder *et al.* (2011) and Fortier & Mailhot (2015) used antecedent rainfall depth (mm of precipitation) in their modeling procedures. Consistent with these, Sandoval *et al.* (2013) found that rainfall depth is an effective predictor of CSO volume and duration. Sandoval *et al.* (2013) also found that the maximum precipitation intensity of the antecedent rainfall event is the most effective predictor of CSO volume, duration, and flow rate. More recently, Yu *et al.* (2018) and Hung *et al.* (2020) suggested that the intensity (mm/h), whether recorded at intervals of 1 min (Yu *et al.* 2018) or calculated as an average (Hung *et al.* 2020), of an antecedent rainfall event is more effective at identifying probabilistic CSO thresholds in outfalls in urban watersheds, and Li *et al.* (2021) demonstrated the use of both depth and intensity of rainfall to effectively predict the volumes and peak rates of flows within CSS pipes. Underscoring all of this, Todeschini (2012) found that, as the climate changes, increasing rainfall intensities present the most significant challenges for urban stormwater management, as more intense average rainfall events will increase CSO discharge volumes regardless of other changes in annual precipitation patterns.

Here we demonstrate a simple and transferable statistical approach using precipitation to predict CSOs, which can be used to assess CSS performance over time. The case study for this research is Cumberland, Maryland, which operates a small wastewater treatment plant (average flow 8–13 million gallons (~36.4–49.2 million L) per day with a treatment capacity of 15 million gallons (~56.8 million L) per day) (US Environmental Protection Agency 2022b) served by a small CSS that includes a moderately environmentally disadvantaged, low-income urban area, and that sits mostly within the Wills Creek and Potomac Lower North watersheds. The USEPA requires Cumberland to reduce its annual CSO volumes to specific targets by 2023. While other Maryland CSS communities separated their stormwater and sewer systems, Cumberland has adopted an approach to reduce CSOs by rehabilitating the existing system, implementing asset management, and expanding the carrying and storage capacity of the CSS. Cumberland has documented all CSO incidents since 2005 and made that information publicly available, in accordance with 2005 Maryland state regulations on tracking sewage spills (Maryland Department of the Environment 2005). To develop our predictive model of CSO incidence and volume from precipitation variables, we analyze 16 years of Cumberland CSO reports, which include data on overflow incidence, overflow causes, discharged volumes, and times and locations of overflows, spanning January 2005 through December 2020, in both the Wills Creek and Potomac Lower North watersheds, in combination with historical precipitation data from Phase 2 of the North American Land Data Assimilation System (NLDAS-2) covering the same period. We assume no change in urbanization and impervious cover over the time period because urbanization and population are not expanding in Cumberland.

METHODS

Sewer overflow data collection and processing

The Maryland Department of the Environment (MDE) maintains the Maryland Reported Sewer Overflow Database, which documents and makes publicly available all reports of combined sewer overflows in Maryland since 2005. After a sewer overflow, publicly owned sewer systems or their operators in Maryland are required to accurately report, within 24 hours, the overflow's gross volume (total sewage spilled) and net volume (total sewage spilled minus any sewage cleaned up), as well as the location, date, reported cause (e.g. wet weather, pipe obstruction, mechanical failure), and receiving body of water for each incident (Maryland Department of the Environment 2005). CSO incident reports from the City of Cumberland were downloaded from the MDE website for the period of January 2005 to December 2020. Maps of the Cumberland sewer system were used to assign the city's 11 CSO outfalls to one of three watersheds, as appropriate. This process assigned six outfalls to the Potomac Lower North watershed, four to the Wills Creek watershed, and one to the Evitts Creek watershed (Figure 1). During the 16 years of the data set, the outfalls in the Wills Creek and Potomac Lower North watersheds overflowed dozens of times per year, while the outfall in Evitts Creek either did not overflow at all or experienced very few overflows per year. With the limited CSO data available, the outfall for Evitts Creek was not included in the analysis. All CSOs included in the analysis had a reported cause of wet weather. The final data set consists of 4,356 CSO discharges caused by precipitation, including 2,102 in the Wills Creek watershed and 2,254 in the Potomac Lower North watershed.

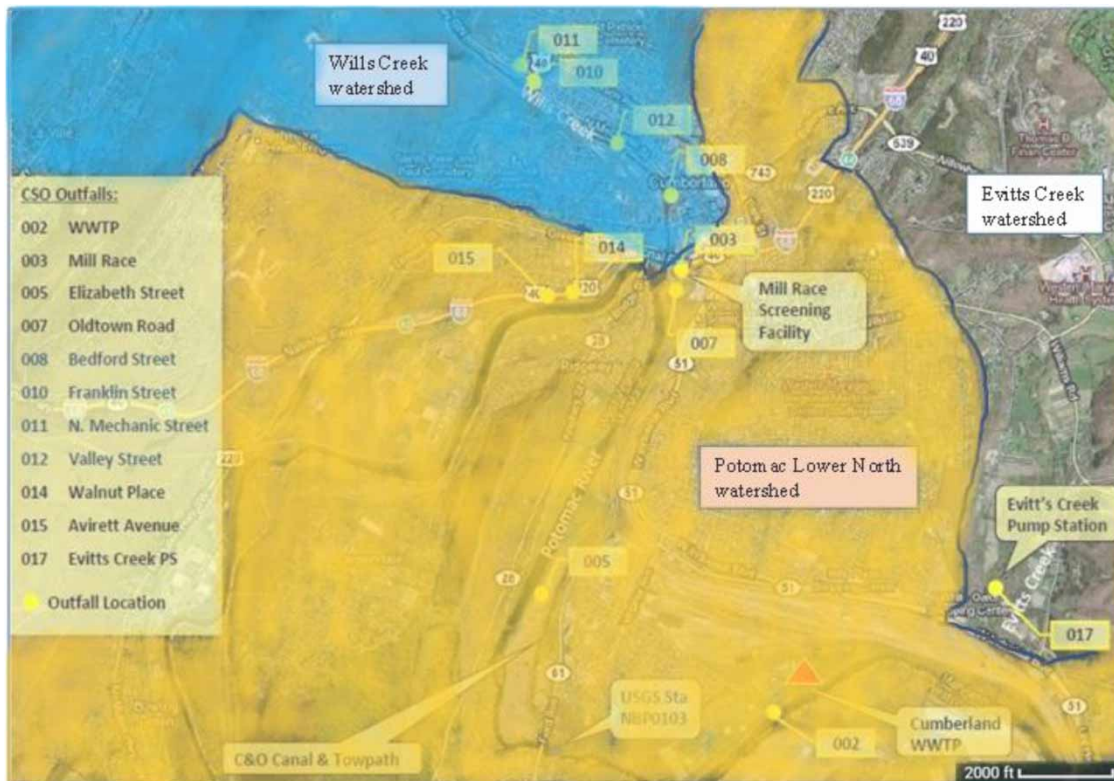


Figure 1 | Map of the Cumberland, MD area annotated with the city's 11 CSO outfalls and 3 watersheds (blue shading = Wills Creek, yellow shading = Potomac Lower North, unshaded = Evitts Creek). Wills Creek outfalls are labelled in blue in the legend, while the Potomac Lower North and Evitts Creek outfalls are labelled in black. This figure is adapted from the public document 'CSO Status in Maryland', a 2019 Maryland Department of the Environment (MDE) presentation (Cheng 2019), with approximate watershed boundaries (blue) from MDE's official Map of 8-Digit Watersheds (Maryland Department of the Environment undated).

Precipitation data collection and processing

North American Land Data Assimilation System Phase 2 (NLDAS-2) precipitation data at 1/8th-degree geospatial resolution and hourly temporal resolution were used for analysis. NLDAS hourly precipitation data are developed using atmospheric forcing and routing models from the execution of four Land Surface Models (LSMs) on a data set assimilated from the rainfall data of the North American Regional Reanalysis (NARR) and Climate Prediction Center (CPC), both of which are based on direct rainfall gauge observations from across the continent that have been disaggregated and adjusted for topography using the Parameter-elevation Regression on Independent Slopes Model (PRISM) (Xia *et al.* 2009). Due to the coarse geospatial resolution, only one set of hourly precipitation data was available for the Cumberland area. Discrete rainfall events were identified as rainfall periods separated by either 12 + consecutive hours of precipitation intensity below 0.2 mm/h (Fortier & Mailhot 2015), or 2 + consecutive hours of zero precipitation (Hung *et al.* 2020). The latter criterion was applied when 12 + consecutive hours of precipitation intensity below 0.2 mm/h failed to discretize rainfall events enough to pair individual rainfall events with daily CSO events.

We adopted the approach of Fortier & Mailhot (2015) to pair CSOs with rainfall events. Each CSO discharge was assigned a 'start time' equal to the discovery time in the CSO incident report, if such a discovery time was given, or 12.00 noon when no discovery time was given. Forty eight percent of CSOs included a discovery time, leaving 52% with a 12.00 noon start time assigned. For each CSO incident, we selected discrete rainfall events for pairing if they occurred between 48 hours before the 'start time' and 12 hours after the start time (considering the latter permissible to accommodate both uncertainty in CSO reports and the coarse geospatial resolution of the NLDAS-2 rainfall data). When a CSO incident had multiple discrete rainfall events that could be paired to it, one event was chosen based on greater total rainfall depth, greater rainfall intensity, and occurrence of the day before the CSO date, in that order of priority. This procedure and prioritization are adapted from the

methodology of pairing CSO and rainfall events used by Fortier & Mailhot (2015). Additionally, rainfall amounts from any other discrete rainfall events between this selected event and the CSO 'start time' were then added into the selected event to create a composite total picture of antecedent precipitation for each CSO incident. This process resulted in the tabulation of 5,405 rainfall events with and without CSO incidence, including 2,696 in the Wills Creek watershed and 2,709 in the Potomac Lower North watershed, with each rainfall event's maximum and average intensities, antecedent dry period, and total rainfall depth recorded from the NLDAS-2 data set. The associated CSOs were grouped by watershed and date, and then each group was summed into a single CSO event, to model CSOs as watershed-scale responses to wet weather. This resulted in consolidated lists of 715 paired CSO-rainfall events in the Wills Creek watershed and 599 paired CSO-rainfall events in the Potomac Lower North watershed, which were used for all subsequent analysis.

Predicting CSO incidence and evaluating CSO incidence thresholds

Using the CSO-rainfall event data, we modeled CSO incidence as a response variable to precipitation conditions (predictor variables) to identify and evaluate CSO incidence thresholds for the Potomac Lower North and Wills Creek watersheds. First, we identified a critical total rainfall depth N_{crit} (mm) which best discriminates CSO events from non-CSO events for each year of data in each watershed, following the method demonstrated by Schroeder *et al.* (2011), i.e. maximizing the average of the proportions of CSO-rainfall events above the threshold and of non-CSO rainfall events below the threshold. We performed logistic regressions using R (R Core Team 2021) of CSO incidence on each of rainfall depth (P , mm), maximum rainfall intensity (I_{max} , mm/h), average rainfall intensity (I_{avg} , mm/h), and antecedent consecutive dry period (CD , h), to identify the best performing predictors. Event maximum intensity was identified as an effective predictor of CSO incidence based on logistic regressions producing higher odds ratios without vast uncertainties, and as such, we identified a critical maximum intensity I_{crit} (mm/h) that best discriminates CSO events from non-CSO events for each watershed in each year, adapting the method of Schroeder *et al.* (2011) to use event maximum intensity instead of depth. The yearly sets of I_{crit} and N_{crit} values were comparatively evaluated for their ability to predict CSOs, and for correlations to year-to-year climate variables such as total annual rainfall, counts of all rainfall events, and counts of more severe rainfall events, derived from the NLDAS-2 data.

Modeling and predicting CSO volume

Using the CSO-rainfall event data, we performed ordinary least-squares (OLS) regression for gross CSO volume on associated event rainfall depth (P) for the CSO events in each watershed and year, based on Fortier & Mailhot (2015). Further analysis, performed in R, of bivariate correlations and regression diagnostics found that while the event rainfall depth is correlated with CSO volume, OLS regression produced highly skewed residuals (see Supplementary Material, Figure S1). To correct this, we transformed the data using natural logarithms, and then performed regressions on the transformed (lognormal) data. We used R programming to develop, evaluate, and compare multiple regression models of CSO volume ($\ln V$) on the natural logarithms of candidate predictors including event rainfall depth ($\ln P$), maximum rainfall intensity ($\ln I_{max}$), average rainfall intensity ($\ln I_{avg}$), and interactions between these predictors, using the full 16-year data sets for all CSO incidents in each of the Wills Creek and Potomac Lower North watersheds. We then applied the two most promising candidate models to the single-year CSO-rainfall event data sets to judge their historic consistency in correlation and diagnostics to identify the best reliable regression model. This model can be applied over each year (or other designated periods of time) to evaluate CSS performance.

RESULTS AND DISCUSSION

CSO incidence

CSO incidence is treated as a binary response variable to precipitation conditions, with total rainfall depth, maximum intensity, average intensity, and length of the dry period preceding rainfall as possible predictor variables. Welch two-sample t-tests were performed on the means of each variable in rainfall events with CSOs versus those without CSOs, and found statistically significant differences in these means ($p < 0.001$) in each watershed (see Supplementary Material, Figure S2).

N_{crit} values ranged between 3.1 and 6 mm in the Wills Creek watershed, and between 3.3 mm and 7.0 mm in the Potomac Lower North watershed, for each year of data from 2005 through 2020. In any given year, N_{crit} values correctly discriminate between 72 and 87% of rainfall events. These probabilities were lower than expected and generally lower than the values demonstrated in Schroeder *et al.* (2011).

Importantly, in year-to-year summary data with the annual sum of NLDAS-2 rainfall events representing annual rainfall totals, we found that N_{crit} values in each watershed positively correlate to the total annual rainfall recorded over Cumberland (Pearson's $r = 0.626$ ($p = 0.0095$) in the Wills Creek watershed; Pearson's $r = 0.535$ ($p = 0.0329$) in the Potomac Lower North watershed). Because N_{crit} is not independent of yearly precipitation changes, N_{crit} is not a good indicator of system performance over time.

To identify a better-performing indicator of CSO incidence, we first performed logistic regressions of CSO incidence in each watershed on each predictor over all 16 years of rainfall events (Table 1 and Figure 2), and then for events by year (Figure 3). Logistic regression is best interpreted via odds ratios. Odds here refers to the probability of occurrence divided by probability of non-occurrence of an outcome (in this case, CSO discharge). Odds ratio refers to the increase in odds of the outcome that accompanies a single-unit increase in the predictor variable. For example, in a logistic regression of CSO discharge occurrence on event precipitation depth (mm), the odds ratio equals the increase in odds of a CSO discharge accompanying a 1.0-mm increase in depth. Mathematically, the odds ratio is constant for all values of the predictor variable. Using the complete data set, average intensity was the predictor with the largest odds ratio, followed by maximum intensity, and both had significantly greater odds ratios than total depth. Pre-rainfall dry period carried an odds ratio of only 1.002–1.005 over the full dataset. Underscoring this result, logistic regression curves showed that despite the statistically significant difference between mean preceding dry period of CSO and non-CSO events, pre-rainfall dry period lacks explanatory power (Figure 2), so this variable was excluded from the yearly subset regressions.

On the yearly scale, odds ratios for logistic regressions of CSO incidence on average intensity ranged from 1.38 to 8.47, as seen in Figure 3. This was consistently the highest odds ratio in any given year, but with great variability. Also, at the annual scale, much higher odds ratios also had much higher uncertainties in the odds ratio. Having odds ratios with high uncertainties and variable magnitudes meant that average intensity would not be useful for assessing system performance. Yearly logistic regressions of CSO incidence on maximum intensity produced more consistent yearly odds ratios (ranging from 1.37 to 2.91), with moderate uncertainties. Odds ratios on maximum intensity were also statistically significantly higher than those for CSO incidence versus total depth in each watershed in each year. Therefore, we used maximum intensity to define a new probabilistic CSO threshold, the critical intensity I_{crit} , in each watershed and year.

Annual values for I_{crit} ranged between 1.42 mm/h and 2.13 mm/h for the Wills Creek watershed, and between 1.70 mm/h and 2.40 mm/h for the Potomac Lower North watershed (Figure 4(a)). In any given year, I_{crit} values correctly discriminate between 75 and 87% of rainfall events (Figure 4(b)). These probabilities are not a statistically significant improvement from the N_{crit} values: the average annual fraction of events successfully identified improves by only 1.24% in the Wills Creek watershed ($t = 0.7706$, $p = 0.447$), and by less than 0.1% in the Potomac Lower North watershed ($t = 0.0414$, $p = 0.967$).

However, unlike N_{crit} values, I_{crit} values are independent of annual total rainfall, as demonstrated by ordinary least-squares regressions (OLS) of I_{crit} and N_{crit} on total annual rainfall (Figure 5). The OLS regression lines in Figure 5 represent the statistically significant relationships of N_{crit} to total annual rainfall detailed previously, and statistically insignificant relationships of I_{crit} to the same (Pearson's $r = 0.352$, $p = 0.181$ in the Wills Creek watershed; Pearson's $r = 0.483$, $p = 0.058$ in the Potomac Lower North watershed). In addition, Figure 5 shows greater year-to-year scatter of N_{crit} values in years with differing rainfall, while I_{crit} remains more consistent. This is reflected in the magnitudes of the slopes of the OLS regression lines; for I_{crit} , these slopes are greater than an order of magnitude reduced in addition to lacking statistical significance. Taken all together, for the aim of identifying a climate-independent CSS performance indicator, I_{crit} is preferable to N_{crit} because it does not increase or decrease with different amounts of total annual rainfall.

Table 1 | Odds ratios resulting from logistic regression of CSO incidence on each predictor in each Cumberland watershed, for all rainfall events (2005–2020)

Predictor	Wills Creek		Potomac Lower North	
	N	Odds Ratio [95% CI]	N	Odds Ratio [95% CI]
Total rainfall depth (P , mm)	2,696	1.157 [1.140–1.174]	2,709	1.170 [1.152–1.188]
Max intensity (I_{max} , mm/h)	2,696	1.759 [1.667–1.856]	2,709	1.791 [1.696–1.892]
Avg intensity (I_{avg} , mm/h)	2,696	2.823 [2.508–3.198]	2,709	2.790 [2.472–3.148]
Preceding dry period (CD , h)	2,696	1.004 [1.002–1.005]	2,709	1.003 [1.002–1.005]

CI = confidence interval.

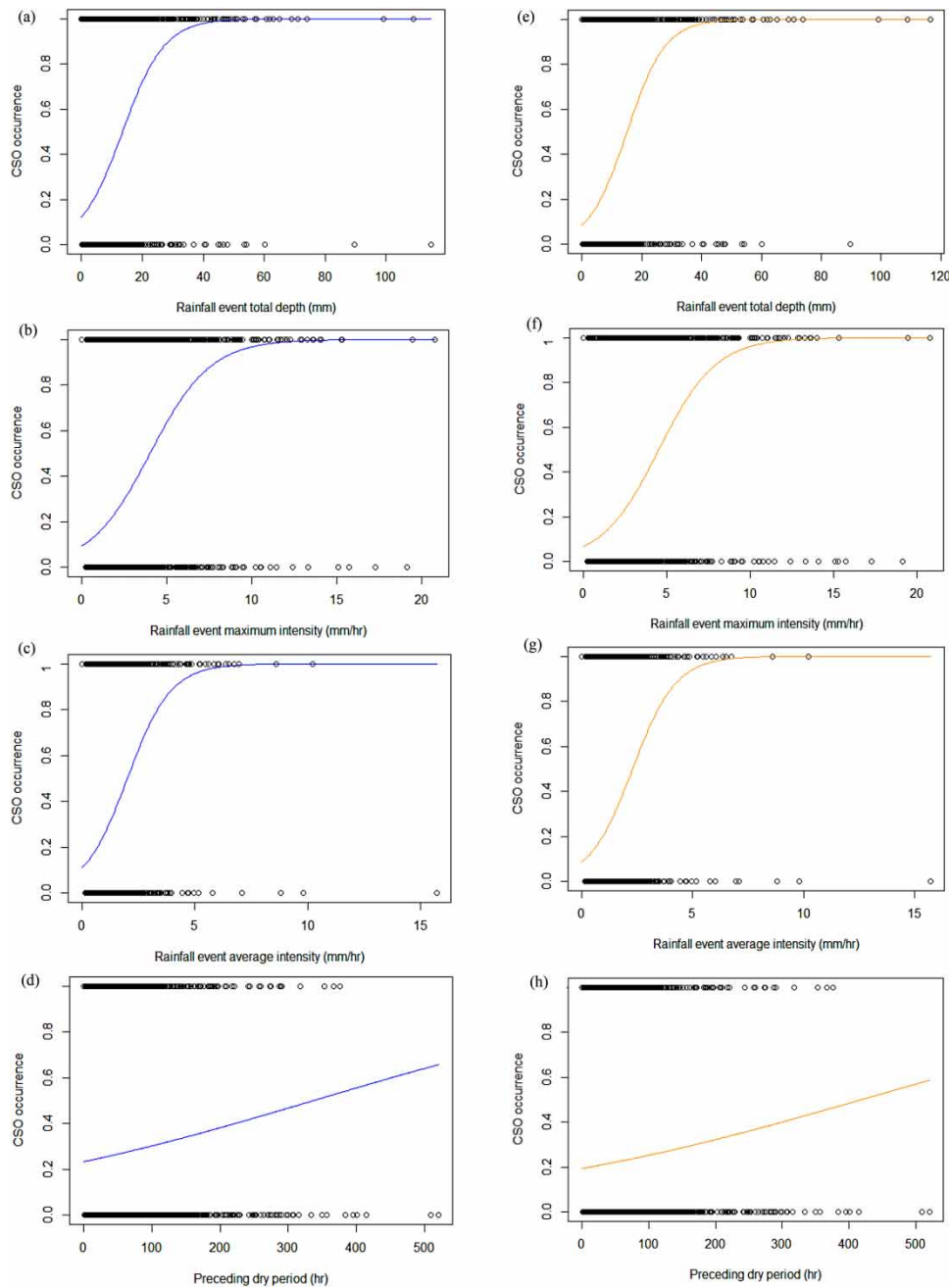


Figure 2 | Logistic regression curves for CSO incidence on each predictor in the Wills Creek (a–d) and Potomac Lower North (e–h) watersheds, for all rainfall events (2005–2020). The logistic regressions for CSO incidence on total precipitation depth (a,e), maximum precipitation intensity (b,f), and average rainfall intensity (c,g) all show traditional probabilistic S-curves, but the logistic regressions for CSO incidence on preceding dry period (d,h) indicate an ineffective logistic predictor. As a result, we discounted antecedent dry period as a predictor for CSO incidence.

Modeling CSO volume

CSO volume is treated as a response to the related rainfall event's precipitation depth (P , mm), average intensity (I_{avg} , mm/h), maximum intensity (I_{max} , mm/h), and antecedent consecutive dry period (CD , h). These variables have generally right-skewed distributions with high variance and spread, and moderate, manageable amounts of correlation between them. Correlation matrices run over the full data set in both watersheds found P to have the strongest correlation with CSO volume (V , gal,

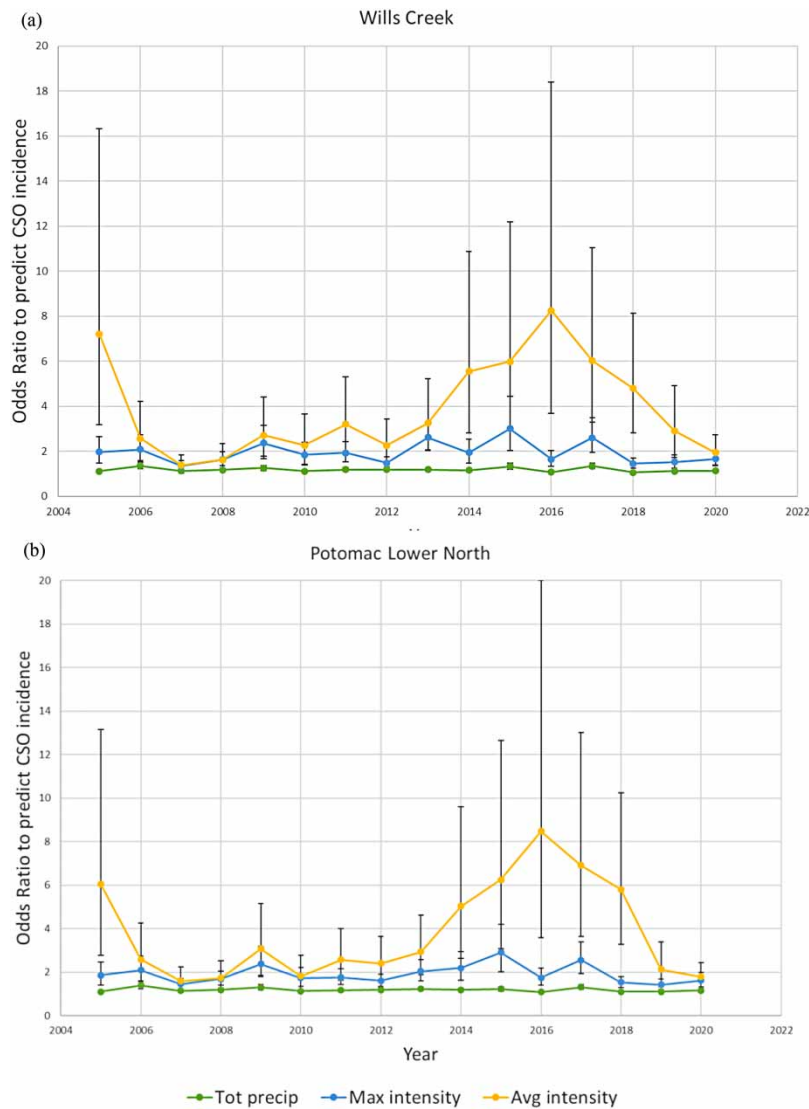


Figure 3 | Yearly odds ratios for logistic regressions of CSO incidence on each of total event precipitation depth (green), maximum event intensity (blue), and average event intensity (yellow) in each of the Wills Creek (a) and Potomac Lower North (b) watersheds.

where 1 gal = 3.7854 L) at Pearson's $r = 0.24$ in the Wills Creek watershed and Pearson's $r = 0.38$ in the Potomac Lower North watershed (Supplementary Material, Table S1). However, OLS regression of V on P produced invalid regression results; we observed that the residuals of these regressions were highly skewed, and there were many outliers (see Supplementary Material, Figure S1). These issues were resolved by treating all variables as lognormal distributions. In turn, the univariate distributions of $\ln V$, $\ln P$, $\ln I_{avg}$, and $\ln I_{max}$ are approximately normal. Bivariate linear regressions explained highly inconsistent amounts of variance in $\ln V$ that were not always statistically significant against a null hypothesis of no correlation (with $\alpha = 0.05$), and often with fit that still indicated nonlinearity. Thus, we deemed multiple linear regression to be a preferred approach.

Model building was undertaken in each watershed using the full 16-year data set ($N = 715$ in Wills Creek, $N = 599$ in Potomac Lower North). During model building, maximum intensity and antecedent dry period were both discarded from consideration for the regression. The maximum intensity $\ln I_{max}$ had multicollinearity with other predictors, a weaker correlation with $\ln V$ than other predictors, and added insignificant explanatory power to regressions that already included other predictors ($p > 0.05$ by general F-test). Meanwhile, antecedent dry period was discarded due to its lack of correlation with $\ln V$. This left $\ln P$ and $\ln I_{avg}$ as the available predictors. In addition, because precipitation depth and average intensity

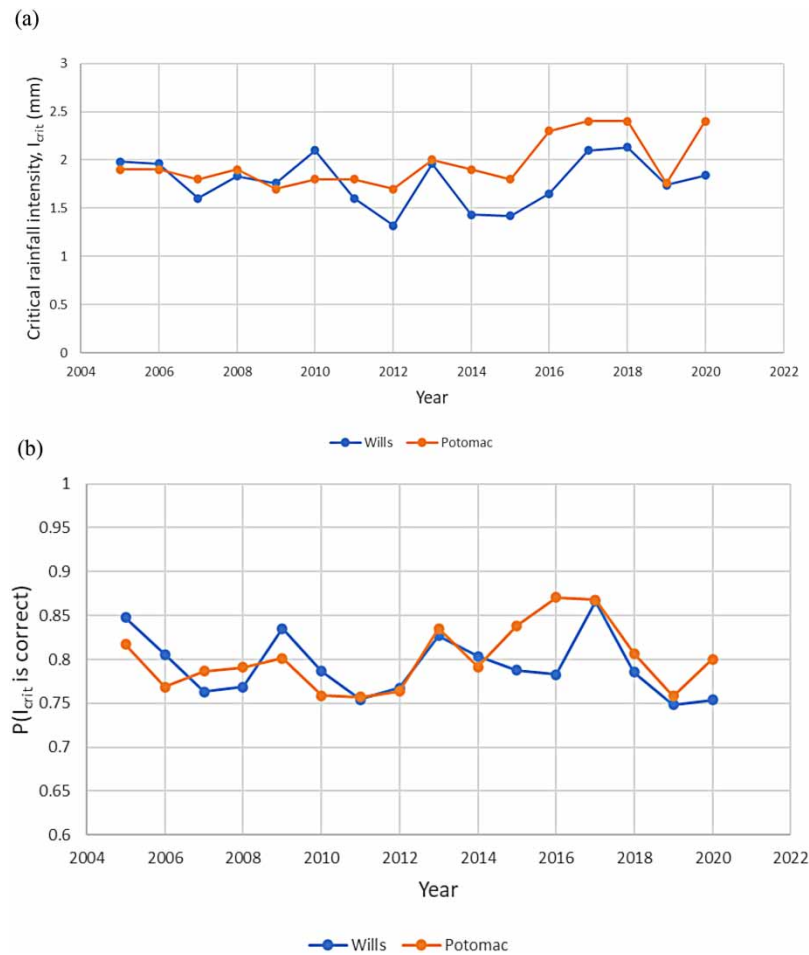


Figure 4 | (a) The determined annual values of critical maximum rainfall intensity (I_{crit} , mm/hr), above which a CSO was likely to occur, in each of the Wills Creek and Potomac Lower North watersheds in Cumberland, MD, in each year (2005–2020). Significant changes in these values between years are rare, and investigation of their causes would require more data regarding CSO mitigation interventions or other structural changes made to the CSS. (b) The probability of the critical intensity in each year correctly predicting whether a given incident will trigger a CSO; or, the fraction of rainfall events correctly discriminated into CSO or non-CSO events. This reflects the predictive ability of I_{crit} .

have moderate correlation with sufficient tolerances, we included an interaction term. We also considered a model developed from a stepwise procedure, which featured a quadratic relationship to depth instead of the interaction term. Outliers that had a starkly higher magnitude of residual error compared to the fitted value from the regression, or that carried high leverage values and oversized influence on regression coefficients, were removed. Such outliers were identified systematically, using DFBETAS analysis (i.e. compiling the change in regression parameters for each event's hypothetical removal from the data set) to measure influence on regression coefficients. Data points were considered outliers and removed if they had both disproportionately high influence on regression coefficients, and leverage values greater than the rule of thumb value of two times the number of predictors plus one, divided by the sample size. (Occasionally, an outlier with an extremely high leverage value or an extremely high influence in DFBETAS analysis would be removed without consideration for the other criterion.) The interpretability, diagnostic concerns (if any), and year-to-year consistency of yearly regression results, particularly in the values of R^2 and mean squared error (MSE), were then compared for both models. While the two models produced similar R^2 values, we selected the first model as it had more consistent uniform relationships of residuals to fitted values, normal distributions of residuals, and less variance in its regression coefficients when examined on an annual basis:

$$\ln(V) = \beta_0 + \beta_1 \ln(P) + \beta_2 \ln(I_{avg}) + \beta_3 \ln(P)\ln(I_{avg})$$

This model has an annual average coefficient of determination of $R^2 = 0.42$ (95%CI: 0.21–0.63) in the Wills Creek watershed, and $R^2 = 0.44$ (95%CI: 0.20–0.68) in the Potomac Lower North watershed, with yearly results displayed in Figure 6. The

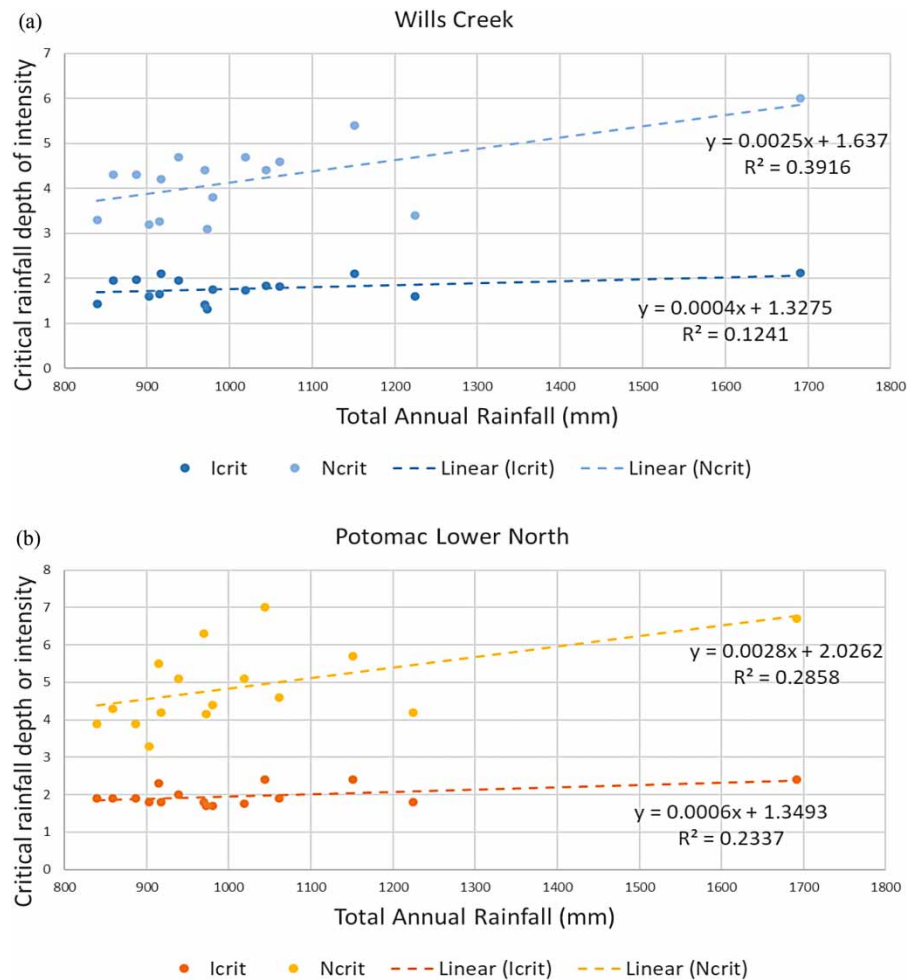


Figure 5 | Ordinary least-squares regression of N_{crit} (lighter) and I_{crit} (darker) values on total annual rainfall (sum of depth of all rainfall events) over the Wills Creek (a) and Potomac Lower North (b) watersheds. Annual values of I_{crit} are independent of changes in annual rainfall, whereas N_{crit} is sensitive to such changes.

model carries an annual average MSE of 2.75 (95%CI: 0.86–4.65) in the Wills Creek watershed, which is lower than the annual average MSE of 3.56 (95%CI: 0.79–6.33) in the Potomac Lower North watershed. There are a few possible reasons for this difference in error. It may be because the rainfall data from NLDAS-2 is calculated for a point much nearer to the CSO outfalls along Wills Creek than to those of the Potomac Lower North watershed, underscoring the importance of geospatially accurate hourly rainfall data in applying this model. The difference in error could also be due to differences in the number and/or geospatial distribution of outfalls. The Wills Creek watershed has four outfalls clustered near to one another along the eponymous creek, while the seven outfalls in the Potomac Lower North watershed are spread across a larger area.

This multiple regression model performs better than the simpler regressions of the natural logarithm of volume ($\ln V$) on the natural logarithm of either event precipitation depth ($\ln P$) or intensity ($\ln I_{avg}$) alone. Figure 6 presents how the yearly coefficients of determination compare to the bivariate regression of $\ln V$ vs $\ln P$, showing the bivariate correlation to be much less reliable. Although the bivariate correlation is sometimes higher than that of the multiple regression, the multiple regression's intensity terms prevent such drastic losses of correlation as can be sometimes observed in the bivariate model, and dramatically reduce the variance among R^2 values of the model. Thus, the multiple regression is the preferred model.

The consistency of this model from year to year makes it a reasonable choice to assess or identify changes in CSS performance, including to evaluate the impacts of climate change and interventions intended to build resilience to climate change. For example, in 2021, Cumberland installed a 50 million gallon (~227 million L) capacity storage volume to mitigate CSOs. The city plans to further increase capacity and mitigate CSOs by installing 78 in (1.98 m)-diameter overflow carrier

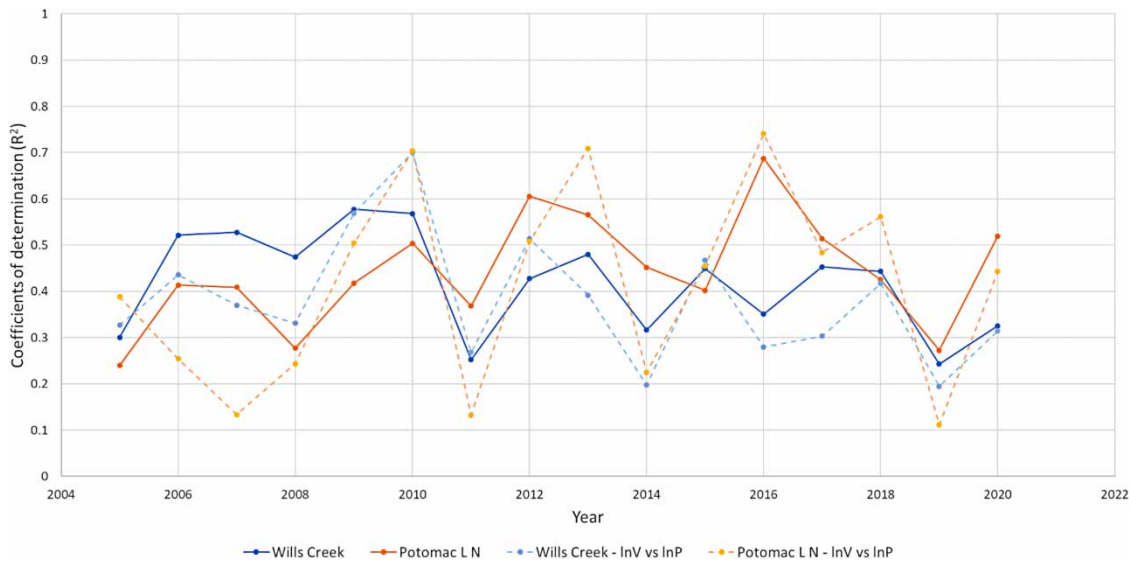


Figure 6 | R^2 coefficients of determination for the multiple regression model in the Wills Creek (blue) and Potomac Lower North (orange) watersheds in each year (2005–2020), and coefficients of determination for the simpler regression of $\ln V$ on $\ln P$ for each watershed (dashed) for comparison. The multiple regression model can dependably explain 30–60% of variance in CSO volumes. Data on interventions or other changes made in the system may help tease out reasons for variation in these results.

pipes in the mid-2020s. Changes in both the I_{crit} thresholds and the regression results measured over each year (or a longer time period) could be used to evaluate the success of these interventions in mitigating CSOs in Cumberland, or adapted for analysis in other communities with comparable data availability.

CONCLUSIONS

Using data on combined sewer overflows and hourly precipitation in Cumberland, Maryland over the years 2005–2020, we identified annual thresholds of 1-hour precipitation intensity above which CSO incidence is predicted (and below which it is not predicted) in each year. Then, we built a regression model to predict CSO volume from the precipitation depth and average intensity of the preceding rainfall event. These annual thresholds and the regression model can be used in tandem to assess the performance of a combined sewer system (CSS), such as Cumberland's, over time.

We developed this as a two-pronged model to predict CSO incidence and volume. Changes to the model outputs can be read over individual time periods (of 1 year or longer) as indicators of changing CSS performance over time. The model consists of:

1. Identifying a critical maximum rainfall intensity (I_{crit}) which most successfully discriminates CSO-associated rainfall events from non-CSO rainfall events over time
2. Performing multiple regression of the natural logarithm of CSO volume (gal) on the natural logarithms of the associated rainfall event's total precipitation depth (mm) and average intensity (mm/h) with an interaction term

We determined that I_{crit} is a more statistically valid performance indicator to evaluate CSS performance than the critical rainfall depth N_{crit} proposed in prior literature. Although the two proposed performance indicators appear about equally capable of predicting CSO occurrences (within the limitations of this study), I_{crit} is independent of fluctuations in wet weather from year to year. In addition, our multiple regression model predicts CSO volumes better than any bivariate correlation of CSO volume to an individual wet-weather variable, or other regression models considered.

Data limitations were encountered in this study, including the temporal and spatial resolution of rainfall data and metadata for CSOs. Rainfall data at 1-hour intervals were chosen because these data were the finest temporal scale data available over the entire 16-year period of analysis. While measurements at even finer time intervals could produce a multiple regression with greater ability to predict CSO volume, finer-scale temporal data were not available. Continuation flow data within each CSO outfall, connected impervious cover area per CSO outfall, amount of infiltration water, and storage volume per connected impervious area were also not available for the Cumberland CSS.

Cumberland is a small municipal system (treatment capacity 15 million gallons (~56.8 million L) per day) with a relatively small, non-expanding urban area. As such, this assessment protocol may be less directly transferable to very large or very small systems that may experience different ranges of CSO discharge volumes and/or qualitatively different engineering issues, or that may be expanding or changing the amounts of impervious cover. We recommend replication of the model-building process in such systems to validate or modify the model and potentially controlling for changes in impervious cover, rather than extrapolating the regression of CSO discharge volumes well beyond the range of CSO discharge volumes in Cumberland.

Future work should consider what magnitude of change in the model's performance indicators – I_{crit} values, regression coefficients, and regression statistics – can be observed in association with documented past interventions to one or more applicable CSS(s) of interest. This approach (I_{crit} plus multiple regression) could help evaluate the effect of those interventions on reducing CSO incidents and volumes in a CSS. Alternatively, these performance indicators could be applied using future precipitation projections to identify how future changes in precipitation in a CSS watershed may impact CSS performance, assuming urbanization and the CSS remains relatively unchanged. Because this methodology is not dependent on the physical characteristics of an individual CSS itself, the modeling procedure could be scaled up to assess broader CSS performance across a group of systems or municipalities of similar size, to determine if CSSs in general are keeping up with wet-weather changes under climate change, and identify locales in need of assistance with doing so.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Citlalli Rojas Huerta and Larissa Chraim for data cleaning and initial analysis; Dr Eric Loken for regression assistance and helpful feedback; Dinah Voyles Pulver for assistance with data access and informative input; and Robert Smith, PE, Director of Engineering at the City of Cumberland, Maryland, for informative input and helpful feedback. Kirchhoff was partially supported by the National Science Foundation under Grant CMMI-1944664. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

DATA AVAILABILITY STATEMENT

All relevant data are available from an online repository or repositories (Bizer & Kirchhoff, 2022).

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Abdellatif, M., Atherton, W., Alkhaddar, R. M. & Osman, Y. Z. 2015 Quantitative assessment of sewer overflow performance with climate change in northwest England. *Hydrological Science Journal* **60** (4), 636–650. doi:10.1080/02626667.2014.912755.
- Andrés-Doménech, I., Montanari, A. & Marco, J. B. 2012 Efficiency of storm detention tanks for urban drainage systems under climate variability. *Journal of Water Resources Planning and Management* **138** (1), 36–46. doi:10.1061/(ASCE)WR.1943-5452.0000144.
- Astaraie-Imani, M., Kapelan, Z., Fu, G. & Butler, D. 2012 Assessing the combined effects of urbanisation and climate change on the river water quality in an integrated urban wastewater system in the UK. *Journal of Environmental Management* **112**, 1–9. doi:10.1016/j.jenvman.2012.06.039.
- Bachmann-Machnik, A., Bruning, Y., Bakhshipour, A. E., Krauss, M. & Dittmer, U. 2021 Evaluation of combined sewer system operation strategies based on highly resolved online data. *Water (Switzerland)* **13** (6), 751–770. doi:10.3390/w13060751.
- Bizer, M. & Kirchhoff, C. 2022 *Regression Modeling of Combined Sewer Overflows to Assess System Performance: Cumberland, MD Case Study Data*. Available from: osf.io/w42cb (accessed 19 September 2022).
- Botturi, A., Ozbayram, E. G., Tondera, K., Gilbert, N. I., Rouault, P., Caradot, N., Gutierrez, O., Daneshgar, S., Frison, N., Akyol, Ç., Foglia, A., Eusebi, A. L. & Fatone, F. 2021 Combined sewer overflows: a critical review on best practice and innovative solutions to mitigate impacts on environment and human health. *Critical Reviews in Environmental Science and Technology* **51** (15), 1585–1618. doi:10.1080/10643389.2020.1757957.
- Cheng, Y.-D. 2019 *CSO Status in Maryland*. Available from: https://www.potomacdwsp.org/wp-content/uploads/2019/08/CSO_ICPRB.pdf (accessed 31 October 2021).
- Christidis, N. & Stott, P. 2022 Human influence on seasonal precipitation in Europe. *Journal of Climate* **35** (15), 5215–5231. doi:10.1175/JCLI-D-21-0637.1.

- Easterling, D. R., Kunkel, K. E., Arnold, J. R., Knutson, T., LeGrande, A. N., Leung, L. R., Vose, R. S., Waliser, D. E., Wehner, M. F., 2017 Precipitation change in the United States. In: *Climate Science Special Report: Fourth National Climate Assessment* (Wuebbles, D. J., Fahey, D. W., Hibbard, K. A., Dokken, D. J., Stewart, B. C. & Maycock, T. K., eds). U.S. Global Change Research Program, Washington, DC, pp. 207–230.
- Environmental Agency 2020 *Consented Discharges to Controlled Waters with Conditions*. Available from: <https://www.data.gov.uk/dataset/55b8eaa8-60df-48a8-929a-060891b7a109/consented-discharges-to-controlled-waters-with-conditions> (accessed 14 September 2022).
- Fortier, C. & Mailhot, A. 2015 Climate change impact on combined sewer overflows. *Journal of Water Resources Planning and Management* **141** (5), 04014073. doi:10.1061/(asce)wr.1943-5452.0000468.
- Hung, F., Harman, C. J., Hobbs, B. F. & Sivapalan, M. 2020 Assessment of climate, sizing, and location controls on green infrastructure efficacy: a timescale framework. *Water Resources Research* **56** (5). doi:10.1029/2019WR026141.
- Li, S., Kazemi, H. & Rockaway, T. D. 2021 Statistical modelling of hydrological performance in a suite of green infrastructure practices. *Water Science and Technology* **84** (12), 3663–3675. doi:10.2166/wst.2021.447.
- Madoux-Humery, A. S., Dorner, S., Sauvé, S., Aboufadel, K., Galarneau, M., Servais, P. & Prévost, M. 2013 Temporal variability of combined sewer overflow contaminants: evaluation of wastewater micropollutants as tracers of fecal contamination. *Water Research* **47** (13), 4370–4382. doi:10.1016/j.watres.2013.04.030.
- Mahaut, V. & Andrieu, H. 2019 Relative influence of urban-development strategies and water management on mixed (separated and combined) sewer overflows in the context of climate change and population growth: a case study in Nantes. *Sustainable Cities and Society* **44**, 171–182. doi:10.1016/j.scs.2018.09.012.
- Maryland Department of the Environment 2005 *Maryland Reported Sewer Overflow Database*. maryland.gov. Available from: <https://mde.maryland.gov/programs/water/compliance/pages/reportedseweroverflow.aspx> (accessed 30 April 2021).
- Maryland Department of the Environment no date *Map of Maryland's 8-Digit Watersheds*. Available from: <https://mde.maryland.gov/programs/water/tmdl/datacenter/pages/8digitwatershed.aspx> (accessed 30 June 2021).
- R Core Team 2021 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from: <https://www.R-project.org/> (accessed 29 August 2021).
- Romei, M., Lucertini, M., Renzoni, E. E., Baldrighi, E., Grilli, F., Manini, E., Marini, M. & Iagnemma, L. 2021 A detention reservoir reduced combined sewer overflows and bathing water contamination due to intense rainfall. *Water (Switzerland)* **13** (23), 3425–3434. doi:10.3390/w13233425.
- Roseboro, A., Torres, M. N., Zhu, Z. & Rabideau, A. J. 2021 The impacts of climate change and porous pavements on combined sewer overflows: a case study of the City of Buffalo, New York, USA. *Frontiers in Water* **3**, 725174. doi:10.3389/frwa.2021.725174.
- Rudberg, P. M., Wallgren, O. & Swartling, Å. G. 2012 Beyond generic adaptive capacity: exploring the adaptation space of the water supply and wastewater sector of the Stockholm region, Sweden. *Climatic Change* **114** (3–4), 707–721. doi:10.1007/s10584-012-0453-1.
- Sandoval, S., Torres, A., Pawlowsky-Reusing, E., Riechel, M. & Caradot, N. 2013 The evaluation of rainfall influence on combined sewer overflows characteristics: the Berlin case study. *Water Science and Technology* **68** (12), 2683–2690. doi:10.2166/wst.2013.524.
- Schoen, M., Hawkins, T., Xue, X., Ma, C., Garland, J. & Ashbolt, N. J. 2015 Technologic resilience assessment of coastal community water and wastewater service options. *Sustainability of Water Quality and Ecology* **6**, 75–87. doi:10.1016/j.swaqe.2015.05.001.
- Schroeder, K., Riechel, M., Matzinger, A., Rouault, P., Sonnenberg, H., Pawlowsky-Reusing, E. & Gnirss, R. 2011 Evaluation of effectiveness of combined sewer overflow control measures by operational data. *Water Science and Technology* **63** (2), 325–330. doi:10.2166/wst.2011.058.
- Scottish Water 2020 A resilient waste water network. In: *Shaping the Future of Your Water and Waste Water Services* (Archer, S. & Black, K., eds). Scottish Water, Dunfermline, Scotland, UK, pp. 17.
- Semadeni-Davies, A., Hernebring, C., Svensson, G. & Gustafsson, L.-G. 2008 The impacts of climate change and urbanisation on drainage in Helsingborg, Sweden: combined sewer system. *Journal of Hydrology* **350** (1–2), 100–113. doi:10.1016/j.jhydrol.2007.05.028.
- Todeschini, S. 2012 Trends in long daily rainfall series of Lombardia (northern Italy) affecting urban stormwater control. *International Journal of Climatology* **32** (6), 900–919. doi:10.1002/joc.2313.
- US Environmental Protection Agency 1994 USEPA CSO control policy. *Federal Register* **59** (75), 18688–18698. Available from: <https://www.epa.gov/sites/default/files/2015-10/documents/owm0111.pdf> (accessed 11 April 2022).
- US Environmental Protection Agency 2022a *Combined Sewer Overflows (CSOs)*. Available from: <https://www.epa.gov/npdes/combined-sewer-overflows-csos> (accessed 15 September 2022).
- US Environmental Protection Agency 2022b *Pollutant Loading Report (DMR) | ECHO | US EPA*. Available from: https://echo.epa.gov/trends/loading-tool/reports/dmr-pollutant-loading/?permit_id=MD0021598 (accessed 15 September 2022).
- Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., Mocko, D. & NCEP/EMC 2009 NLDAS primary forcing data L4 hourly 0.125 (0.125 degree V002). In: *NASA/GSFC/HSL* (Mocko, D., ed.). Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, Maryland, USA. doi:10.5067/6J5LHHOHZHN4. (accessed 14 February 2022).
- Yu, Y., Zhang, S., An, A. K. & Furumai, H. 2018 Simple method for calculating hydraulic behavior of combined sewer overflow from rainfall event data. *Journal of Water Resources Planning and Management* **144** (10), 04018061. doi:10.1061/(asce)wr.1943-5452.0000972.