JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

Vol. 50, No. 6 AMERICAN WATER RESOURCES ASSOCIATION December 2014

NITROGEN REMOVAL BY STORMWATER MANAGEMENT STRUCTURES: A DATA SYNTHESIS[1]

Benjamin J. Koch, Catherine M. Febria, Muriel Gevrey, Lisa A. Wainger, and Margaret A. Palmer[2]

ABSTRACT: A comprehensive synthesis of data from empirically based published studies and a widely used stormwater best management practice (BMP) database were used to assess the variability in nitrogen (N) removal performance of urban stormwater ponds, wetlands, and swales and to identify factors that may explain this variability. While the data suggest that BMPs were generally effective on average, removal efficiencies of ammonium (NH_4), nitrate (NO_3), and total nitrogen (TN) were highly variable ranging from negative (i.e., BMPs acting as sources of N) to 100%. For example, removal of NO_3 varied from (median ±1 SD) −15 ± 49% for dry ponds, 32 ± 120% for wet ponds, 58 ± 210% for wetlands, and 37 ± 29% for swales. Across the same BMP types, TN removal was 27 ± 24%, 40 ± 31%, 61 ± 30%, and 50 ± 29%. NH_4 removal was 9 ± 36%, 29 ± 72%, 31 ± 24%, and 45 ± 34%. BMP size, age, and location explained some of the variability. For example, small and shallow ponds and wetlands were more effective than larger, deeper ones in removing N. Despite well-known intra-annual variation in N fluxes, most measurements have been made over short time periods using concentrations, not flow-weighted N fluxes. Urban N export is increasing in some areas as large storms become more frequent. Thus, accounting for the full range of BMP performance under such conditions is crucial. A select number of long-term flux-based BMP studies that rigorously measure rainfall, hydrology, and site conditions could improve BMP implementation.

(KEY TERMS: best management practices; constructed wetland; detention pond; nitrogen removal; nutrients; performance; stormwater management; retention basin; urban areas; vegetated swale.)

Koch, Benjamin J., Catherine M. Febria, Muriel Gevrey, Lisa A. Wainger, and Margaret A. Palmer, 2014. Nitrogen Removal by Stormwater Management Structures: A Data Synthesis. *Journal of the American Water Resources Association* (JAWRA) 50(6): 1594-1607. DOI: 10.1111/jawr.12223

INTRODUCTION

Effective stormwater management is central to providing clean water and healthy rivers. Stormwater runoff severely degrades urban and downstream water bodies by producing flashy hydrographs, carrying sediment and contaminants into streams, altering channel morphology, and reducing freshwater biodiversity (Walsh *et al.*, 2005b). Watershed planning and management practices that mitigate the effects of increased impervious cover and pollutant inputs include engineered structures designed to slow and treat stormwater, as well as low impact development

[1]Paper No. JAWRA-13-0178-P of the *Journal of the American Water Resources Association* (JAWRA). Received August 16, 2013; accepted April 11, 2014. **Discussions are open until six months from print publication**.

[2]Formerly Postdoctoral Research Associate (Koch and Febria), Visiting Scientist (Gevrey), Research Professor (Wainger), Chesapeake Biological Laboratory, Center for Environmental Science, University of Maryland, Solomons, Maryland 20688; currently Postdoctoral Research Scholar (Koch), Center for Ecosystem Science and Society, Northern Arizona University, Box 5620, Flagstaff, Arizona 86011; currently Postdoctoral Research Fellow (Febria), School of Biological Sciences, University of Canterbury, Christchurch, New Zealand; and Professor (Palmer), Department of Entomology, University of Maryland, College Park, Maryland 20742 (E-Mail/Koch: ben.koch@nau.edu).

(LID) and green engineering practices that promote rapid infiltration (Davis, 2005; Dietz, 2007). Although LID and green engineering show great promise in ameliorating the negative consequences of stormwater in urbanizing watersheds (Dietz, 2007; Ahiablame *et al.*, 2012), engineered solutions such as detention ponds and treatment wetlands are necessary in many older cities and suburbs that already have substantial impervious footprints. Accordingly, stormwater best management practices (SW BMPs), such as ponds, wetlands, and vegetated biofilters and bioretention systems have been widely used throughout the United States (U.S.) to help control stormwater runoff volume in urban watersheds (Booth *et al.*, 2002; Hogan and Walbridge, 2007).

Greater sophistication in SW BMP design (Walsh *et al.*, 2005a) along with improved understanding of how urban design and infrastructure impact water quality (Hatt *et al.*, 2004) have advanced the ability of natural resource managers to reduce peak flows and the flux of pollutants to streams and rivers. Watershed models and optimization algorithms that simulate pollutant removal have been developed (e.g., Lee *et al.*, 2012) to help decide which types, how many, and where to build SW BMPs. Natural resource managers and water quality regulators also rely on summaries (e.g., Barrett, 2008; Simpson and Weammert, 2009) and databases (Winer, 2000; International Stormwater BMP Database, 2012) of the performance of different BMP types in the field. Such performance data may also be integrated into decision support tools (PLRM Development Team, 2009; USEPA, 2010a; Chesapeake Stormwater Network, 2012). Despite these resources, improvements in water quality are still less than desired in many regions throughout the U.S. (USEPA, 2013). Several possibilities may explain this continued degradation including improper SW BMP construction or siting, insufficient maintenance, and inaccurate or biased measurements and simulation of BMP performance.

A range of watershed, hydrologic, and site-specific factors (such as land cover, patterns of precipitation and discharge, and position in the catchment) may influence SW BMP performance, yet there is very little quantitative information on these factors (Strecker *et al.*, 2001; Barrett, 2008). Our goal was to synthesize the latest data on SW BMP effectiveness and determine how much of this variability could be explained by environmental factors. To ensure we obtained the highest quality information, we relied on BMP performance data from published empirical studies (i.e., we did not include data generated by models or estimates based on design specifications). Given the importance of elevated urban sources of nitrogen (N) in contributing to eutrophication of rivers and coastal waters (Castro *et al.*, 2003; Ator *et al.*, 2011; Howarth *et al.*, 2012), we focused on N removal. Although many SW BMPs (e.g., rain gardens, sand filters, and permeable pavements) warrant analysis, we concentrated on three classes of stormwater management structures that are among the most widely used SW BMPs in the U.S. (USEPA, 2004; Collins *et al.*, 2010): detention/retention ponds, constructed wetlands, and vegetated swales. For each of these three types, we compiled published, empirically measured N removal efficiencies, and a suite of corresponding environmental variables hypothesized to control N removal. In addition, we surveyed a major database of SW BMP performance to assess the availability of environmental variables that could explain variation in SW BMP performance.

We hypothesized that N removal would vary as a function of watershed features (e.g., size, land cover), BMP properties (e.g., type, age, size, and spatial configuration), and hydrologic factors (e.g., peak storm discharge, storm flow *vs.* base flow). The available data allowed us to test these hypotheses for short-term removal (over hours to weeks), but data were insufficient to assess annual or longer term performance. This is despite massive public investments in SW BMP infrastructure with the expectation of multidecadal water treatment benefits (Urbonas and Olson, 2011). Furthermore, despite the proliferation of modeling and decision support tools to help natural resource managers select and site BMPs, published empirical studies and a widely used SW BMP performance database rarely reported data on environmental variables hypothesized to influence performance. Given the pressing need to provide natural resource managers and water quality regulators with more complete information, we emphasize the importance of measuring SW BMP performance over time scales of years and under a range of ecological, hydrological, and landscape conditions.

METHODS

Literature Search

We conducted a literature search for published studies on detention/retention ponds, constructed wetlands, and vegetated swales (hereafter ponds, wetlands, and swales) in the Web of Science® database (2012; Thomson Reuters, New York) using the keyword terms listed in Table S1. We used Web of Science® because of the available databases; it had the most extensive coverage of environmental science and engineering journals. We systematically screened the abstracts (and when necessary, complete articles)

TABLE 1. Definitions of the Three Types of Stormwater Best Management Practices (BMPs) Surveyed in the Data Synthesis.

BMP Type		Definition
Ponds	Dry ponds (detention ponds)	A basin that temporarily impounds stormwater runoff and empties completely within a short time (usually <24 h)
	Dry extended detention ponds	A basin that temporarily impounds stormwater runoff and empties slowly (usually within 24-48 h)
	Wet ponds (retention ponds)	A basin that intercepts stormwater runoff and holds a permanent pool of water
Wetlands	Constructed wetlands	A basin or series of basins and channels that intercepts stormwater and contains wetland vegetation
	Wet swales	A shallow vegetated channel with low infiltration capacity that temporarily stores stormwater
Swales	Grass swales (dry swales)	A shallow vegetated channel that promotes infiltration of stormwater as it is conveyed down the watershed

of all publications (n = 701) returned in our initial search and selected only those containing data that matched our specific criteria. We focused exclusively on studies conducted in watersheds dominated by urban or suburban land uses, and only considered articles that contained: (1) empirically measured estimates of N removal and (2) estimates of influent and effluent N constituent loads or concentrations for SW BMPs. This selection process yielded 324 individual observations of N removal performance for SW BMPs ($n_{\text{dry pond}}$ = 34, $n_{\text{wet pond}}$ = 72, n_{wetland} = 195, n_{swale} = 23) in urban or suburban watersheds from 30 peer-reviewed studies (Table S2). "Ponds" included wet (retention) ponds, dry detention ponds, and dry extended detention ponds; "wetlands" included constructed wetlands both within and outside the stream channel, and wet swales; "swales" included dry swales and grass swales (Table 1; USEPA, 2004). These data represented 11 different N constituents (Figure 1); we restricted subsequent analyses to 246 data points for the three most commonly reported constituents of water quality concern: ammonium + ammonia (hereafter NH_4), nitrate or nitrite + nitrate (hereafter NO_3), and total nitrogen (TN).

Measuring SW BMP Performance

For each of those 246 observations, we used removal efficiency as our metric of SW BMP performance:

$$\text{Removal efficiency} = \frac{(\text{influent N} - \text{effluent N})}{\text{influent N}} \times 100 \tag{1}$$

Removal efficiencies are not perfect measures of SW BMP performance and can vary with influent concentration (Barrett, 2005). However, alternative methods of assessing individual SW BMP performance such as the effluent probability method (Strecker *et al.*, 2001; Chen *et al.*, 2009) are data intensive, requiring raw monitoring data for their calculation. Because the studies included in our synthesis did not report raw data, we relied on removal efficiencies to enable including as many studies as possible.

We based our analysis primarily on reported N removal efficiencies that were calculated using loads of N (in units of mass) entering and exiting SW BMPs over a specified storm event or base-flow interval. Authors of some studies calculated N removal efficiency from flow-weighted average N concentrations, where N concentrations measured over the course of a storm hydrograph were weighted by the proportion of total flow measured for each sampling interval. For the purpose of estimating N removal efficiency, those flow-weighted measures (units: mg N/l) are equivalent to using N loads (units: mg N). Finally, to provide adequate sample sizes for the environmental variables we tested, our analysis also included some cases of N removal efficiencies calculated from unweighted average N concentrations because no corresponding data on discharge levels or loads were reported. Thus, our NH_4/NO_3/TN dataset of removal efficiency consisted of 99 load-based or flow-weighted measurements and 147 estimates based on unweighted influent and effluent concentrations.

For each observation, we recorded a suite of environmental variables related to watershed attributes, SW BMP characteristics, and hydrologic conditions. We tested the relationships between N removal and the following environmental factors: watershed impervious cover, maximum event discharge, number of storm events, flow conditions (storm flow *vs*. base flow), quality of data (raw concentration *vs*. flow-weighted concentration), presence of permanent water, presence of vegetation, ratio of SW BMP area to contributing watershed area, SW BMP age, SW BMP depth, SW BMP volume, position in treatment train, and cumulative number of SW BMPs in treatment train. Treatment trains consist of serially connected SW BMPs along stormwater flow paths (USEPA, 2004).

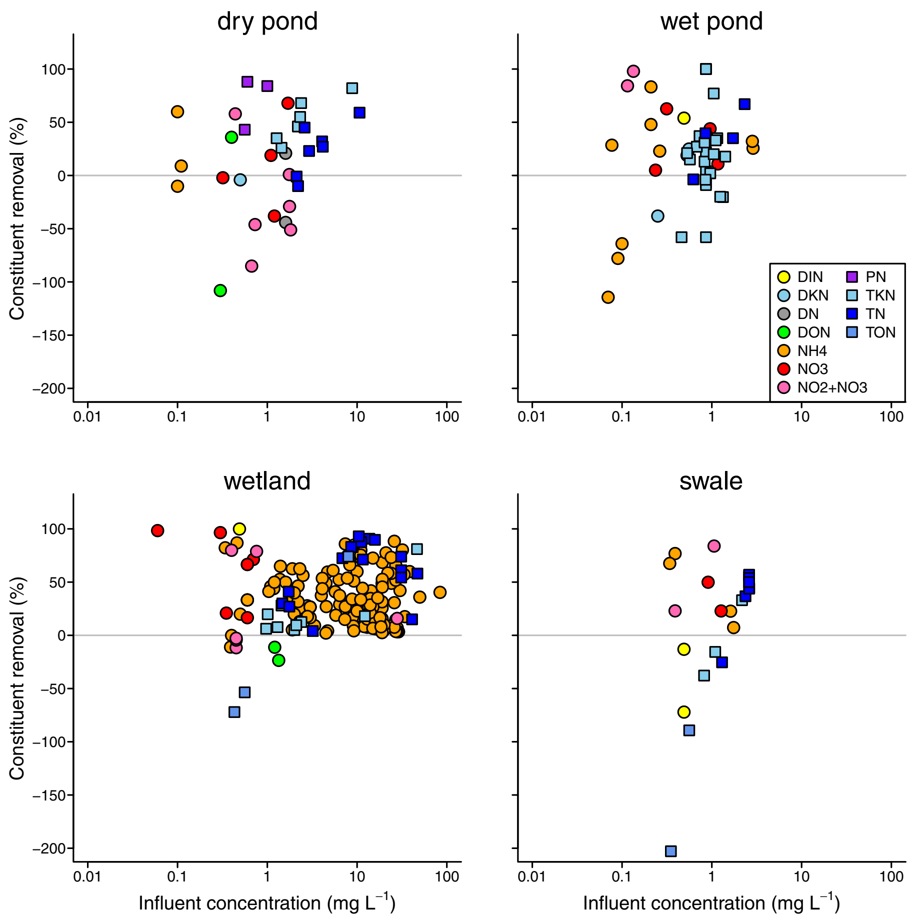

FIGURE 1. Nitrogen (N) Removal Efficiency and Concentration for Constituents in Water Flowing into Dry Ponds, Wet Ponds, Wetlands, and Swales. Data are from a comprehensive synthesis of empirical studies of stormwater best management practice (SW BMP) performance measured over time scales of hours to weeks. Over those time periods, stormwater management structures were generally effective (removal efficiency > 0) but highly variable in removing N across a wide concentration range. Solid gray lines denote no net effect of SW BMPs on N levels. DIN, dissolved inorganic nitrogen; DKN, dissolved Kjeldahl nitrogen; DN, dissolved nitrogen; DON, dissolved organic nitrogen; NH_4, ammonium; NO_3, nitrate; NO_2 + NO_3, nitrite + nitrate; PN, particulate nitrogen; TKN, total Kjeldahl nitrogen; TN, total nitrogen; TON, total organic nitrogen. One outlier (NO_3) does not appear on the wetland panel: 0.54 mg/l, −794%.

Data Analysis

We conducted our data analysis in two stages. In the first stage, we examined the overall performance and variability of SW BMPs. To assess the effect of including unweighted concentration-based estimates of N removal efficiency, we conducted parallel analyses on the subset of observations based solely on load or flow-weighted N data. Thus, for the complete dataset and the smaller load-based subset, we conducted the following analyses. We calculated the proportion of observations with N removal efficiency greater than zero, indicating some removal of N. We used permutation *F*-tests (Manly, 2007) to compare mean N removal efficiencies of dry ponds, wet ponds, wetlands, and swales. To assess variability in measured N removal efficiencies, we calculated sample standard deviations.

In the second stage of our data analysis, we tested how watershed- and SW BMP-level factors (Table 2) might control SW BMP performance. We focused on those observations from the complete dataset for which information on associated environmental variables was available. This analysis included both load-based and concentration-based measures of N removal efficiency, as there were insufficient data for an exclusively load-based analysis. Very few observations had data for most watershed- and SW BMP-level factors, making a multivariate analysis unfeasible. To make use of all available data, we analyzed the relationship between N removal efficiency and explanatory variables independently for each environmental factor and each N constituent (for $n \geq 20$). We used ordinary least-squares regression (for continuous variables) and permutation tests (for categorical variables), and restricted our analyses to ponds and wetlands, as these shared similar basin-type features. To compare the degree of variability between levels of categorical environmental variables,

TABLE 2. Percentage of Records Measuring N Removal from Stormwater Ponds or Wetlands That Reported Data for Relevant Environmental Variables. This study focused only on published, peer-reviewed articles. The total number of sites is 29 for this study and 92 for the International Stormwater BMP Database.

Variable	% of Sites with Data	
	This Study	**International Stormwater BMP Database (2012)**
Watershed impervious cover	28	72
Maximum event discharge	38	38
BMP age	59	71
BMP depth	62	17
BMP volume	52	21
BMP area: watershed area	59	20
Treatment train position	100	2
Vegetation presence	76	40
Flow-weighted or raw concentration	93	100
Flow conditions (base flow or storm flow)	100	86
Watershed area	79	96
Location (latitude, longitude)	34	100
Maintenance schedule	0	30

Note: BMP, best management practice.

we used permutation tests for equality of variances (Beersma and Buishand, 1999).

In addition to our survey of published literature, we analyzed data in the International Stormwater BMP Database (International Stormwater BMP Database, 2012) to assess the availability of environmental variables that could explain variation in SW BMP performance. We calculated the proportion of pond and wetland sites with N removal data that also reported data on the same environmental variables we examined in the literature-based data synthesis.

TABLE 3. Variability in Performance of SW BMPs as Represented by Sample Standard Deviations (SD) of Removal Efficiencies of Ammonium (NH_4), Nitrate or Nitrite + Nitrate (NO_3), and Total Nitrogen (TN) for Dry Ponds, Wet Ponds, Wetlands, and Swales. Sample sizes are in parentheses.

Constituent	Dry Pond	Wet Pond	Wetland	Swale
Including all data				
NH_4	36% (3)	72% (11)	24% (141)	34% (4)
NO_3	49% (10)	120% (19)	210%[1] (16)	29% (4)
TN	24% (7)	31% (5)	30% (19)	29% (7)
Including only load-based or flow-weighted data				
NH_4	36% (3)	68% (5)	28% (22)	6.6% (2)
NO_3	49% (10)	130% (16)	230%[2] (14)	29% (4)
TN	24% (7)	na (1)	20% (8)	29% (7)

[1]SD = 38% when a single extreme observation (−794%) is excluded.
[2]SD = 40% when a single extreme observation (−794%) is excluded.

RESULTS

For each constituent we examined, available data suggest that SW BMPs were effective at removing N across a wide range of influent concentrations (Figure 1). The majority (89% for the complete dataset, 77% for the load-based subset) of observations across the three common N constituents and BMP types showed a reduction in N, but performance of ponds, wetlands, and swales was highly variable. Results from the complete dataset and the load-based subset of data both revealed that wet ponds were 2-3 times more variable than other BMPs in their ability to remove NH_4 and NO_3 (Table 3, Figure 2). While wetlands also showed high variability in NO_3 removal, this finding was largely driven by a single observation. Data on site and watershed environmental variables hypothesized to influence SW BMP performance were sparsely reported in the International Stormwater BMP Database, and were only marginally more prevalent in the peer-reviewed, empirical studies (Table 2).

Removal efficiencies of NH_4, NO_3, and TN in the full dataset ranged from negative values (i.e., SW BMPs acting as sources of N) to 100% for stormwater ponds and wetlands (Figure 2A). Swales were represented by fewer data points (n = 15), and efficiencies ranged from <25% up to 85% for the dissolved constituents NH_4 and NO_3. N removal efficiencies for swales did not exceed 60% for TN. Removal efficiencies did not differ statistically among SW BMP types for NH_4 ($F_{3,155} = 1.50$, $p = 0.253$), NO_3 ($F_{3,45} = 0.150$, $p = 0.958$), or TN ($F_{3,34} = 2.22$, $p = 0.106$). Results from the load-based subset of the data were similar to those from the complete dataset, although diminished sample sizes reduced apparent variability in some cases (Figure 2B, Table 3).

Few studies provided information on critical environmental factors that could have influenced SW BMP performance (Table 2) which limited our ability

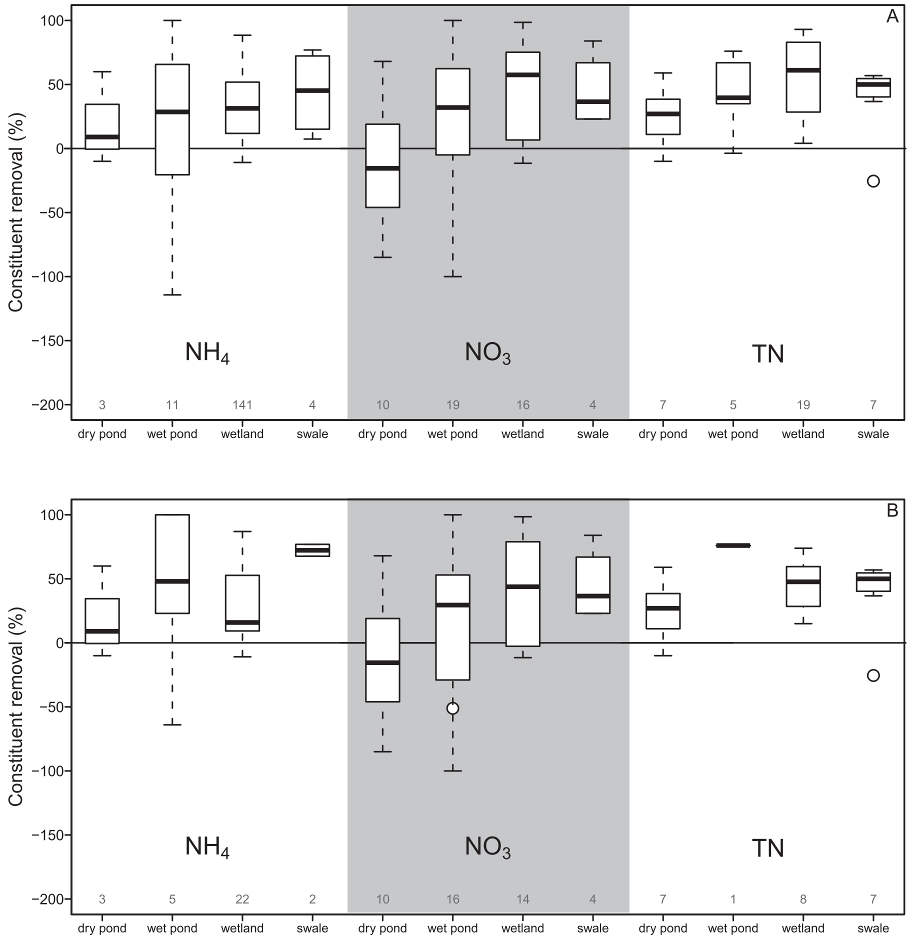

FIGURE 2. Nitrogen (N) Removal Efficiencies of Dry Ponds, Wet Ponds, Wetlands, and Swales Are Highly Variable for All N Constituents. Panel A shows the full dataset and panel B includes only those N removal efficiencies calculated exclusively from loads or flow-weighted measurements. For each panel, two outliers are not shown: −437 (NO_3, wet pond) and −794 (NO_3, wetland). Within each constituent group, stormwater best management practice types are not significantly different (permutation test, $p > 0.05$). Sample sizes are indicated below each box-and-whisker plot. NH_4, ammonium; NO_3, nitrate or nitrate + nitrite; TN, total nitrogen.

to explain variability in removal efficiencies to just some of the data (Table 4). For NH_4, removal efficiencies did not vary with maximum event discharge, flow conditions, or data quality (i.e., raw *vs*. flow-weighted concentrations), but the positioning of stormwater ponds and wetlands within treatment trains did affect performance. NH_4 removal efficiency increased along a treatment train (Figure 3) and was also less variable ($F_{79,74} = 1.86$, $p = 0.047$). In addition, NH_4 removal efficiency increased along a treatment train (Figure 3) and also increased with the cumulative number of SW BMPs in a treatment train (Table 4). Deeper (≥1 m) ponds and wetland basins did not remove NH_4 as effectively as shallower (<1 m) basins (Figure 4), and ponds and wetlands with large volumes did not perform as well as smaller SW BMPs (Table 4), although these effects were weak ($R^2 = 0.11$ and 0.03, respectively). NH_4 removal declined with SW BMP age and as the ratio of SW BMP area to contributing watershed area increased (Figure 5, Table 4), however, these variables explained only a small fraction of variance in SW BMP performance ($R^2 = 0.03$ and 0.13, respectively). TN removal efficiencies were higher for ponds and wetlands with permanent water and under base flow *vs*. storm flow conditions (Table 4). In addition, flow-weighted TN removal efficiencies were lower than those based on concentration data alone (Table 4), indicating that both concentration and flow data (i.e., N fluxes) may be necessary to accurately measure SW BMP performance. NH_4 removal efficiency decreased with the number of storm events composing that estimate of removal (Table 4), suggesting that longer term estimates of SW BMP performance that integrate many storms may be lower than short-term estimates based on single events.

Plots of N removal efficiency *vs*. influent concentration revealed that stormwater ponds, wetlands, and swales generally served as N sinks; however, in a few cases N removal efficiencies were negative, indicating

TABLE 4. Relationships between Selected Environmental Variables and Nitrogen (N) Removal Performance of Stormwater Ponds and Wetlands for Ammonium (NH_4), Nitrate or Nitrite + Nitrate (NO_3), and Total Nitrogen (TN). We used ordinary least-squares regression to test for relationships between N removal efficiency and continuous environmental variables: watershed impervious cover, maximum event discharge, number of storm events, ratio of SW BMP area to contributing watershed area, and SW BMP age, depth, and volume. We used permutation tests to detect relationships between N removal efficiency and categorical environmental variables: presence of permanent water (dry ponds *vs.* wet ponds and wetlands), treatment train position (1st-2nd position *vs.* 3rd-5th position), cumulative number of SW BMPs in a treatment train (1-2 *vs.* 3-5), presence of vegetation, data quality (raw *vs.* flow-weighted concentrations), and flow conditions (storm flow *vs.* base flow). Data include both load-based and concentration-based measures of N removal efficiency. Significant *p*-values are shown in bold.

Constituent	Variable	N	Model	R^2	p	β_0	β_1	D
NH_4	Watershed impervious cover (%)	9	Insufficient data	—	—	—	—	—
	Maximum event discharge (l/s)	130	NH_4 removal (%) = $\beta_0 + \beta_1 \times \log_{10}$[maximum event discharge]	0.007	0.332	38.7	−4.20	—
	BMP age (years)	142	NH_4 removal (%) = $\beta_0 + \beta_1 \times \log_{10}$[BMP age]	0.032	**0.033**	30.4	−9.67	—
	BMP depth (m)	139	NH_4 removal (%) = $\beta_0 + \beta_1 \times \log_{10}$[BMP depth]	0.111	**<0.0001**	13.8	−48.8	—
	BMP volume (m^3)	132	NH_4 removal (%) = $\beta_0 + \beta_1 \times \log_{10}$[BMP volume]	0.032	**0.039**	64.9	−9.04	—
	BMP area: watershed area	141	NH_4 removal (%) = $\beta_0 + \beta_1 \times$ [BMP area:watershed area]	0.127	**<0.0001**	34.7	−1,230	—
	Number of storm events	149	NH_4 removal (%) = $\beta_0 + \beta_1 \times$ [number of storm events]	0.030	**0.034**	35.5	−0.580	—
	Dry *vs.* wet	3 *vs.* 152	Permutation test for difference in means	—	0.448	—	—	−13.2
	Treatment train position (1-2 *vs.* 3-5)	80 *vs.* 75	Permutation test for difference in means	—	**0.001**	—	—	−15.1
	Cumulative BMPs (1-2 *vs.* 3-5)	162 *vs.* 29	Permutation test for difference in means	—	**<0.0001**	—	—	−47.6
	Vegetation (absent *vs.* present)	1 *vs.* 145	Insufficient data	—	—	—	—	—
	Flow-weighted *vs.* raw concentration	18 *vs.* 123	Permutation test for difference in means	—	0.926	—	—	−0.680
	Storm flow *vs.* base flow	11 *vs.* 137	Permutation test for difference in means	—	0.426	—	—	7.01
NO_3	Watershed impervious cover (%)	5	Insufficient data	—	—	—	—	—
	Maximum event discharge (l/s)	7	Insufficient data	—	—	—	—	—
	BMP age (years)	17	Insufficient data	—	—	—	—	—
	BMP depth (m)	19	Insufficient data	—	—	—	—	—
	BMP volume (m^3)	11	Insufficient data	—	—	—	—	—
	BMP area: watershed area	26	NO_3 removal (%) = $\beta_0 + \beta_1 \times$ [BMP area:watershed area]	0.037	0.345	−7.73	1,180	—
	Number of storm events	39	NH_4 removal (%) = $\beta_0 + \beta_1 \times$ [number of storm events]	0.061	0.130	−44.0	2.49	—
	Dry *vs.* wet	10 *vs.* 35	Permutation test for difference in means	—	0.928	—	—	−7.55
	Treatment train position (1-2 *vs.* 3-5)	43 *vs.* 0	Insufficient data	—	—	—	—	—
	Cumulative BMPs (1-2 *vs.* 3-5)	57 *vs.* 5	Permutation test for difference in means	—	0.428	—	—	−33.7
	Vegetation (absent *vs.* present)	6 *vs.* 26	Permutation test for difference in means	—	0.399	—	—	−45.9
	Flow-weighted *vs.* raw concentration	26 *vs.* 5	Permutation test for difference in means	—	0.231	—	—	−63.3
	Storm flow *vs.* base flow	24 *vs.* 13	Permutation test for difference in means	—	0.765	—	—	23.7
TN	Watershed impervious cover (%)	3	Insufficient data	—	—	—	—	—

(continued)

TABLE 4. Continued.

Constituent	Variable	N	Model	R^2	p	β_0	β_1	D
	Maximum event discharge (l/s)	13	Insufficient data	—	—	—	—	—
	BMP age (years)	18	Insufficient data	—	—	—	—	—
	BMP depth (m)	12	Insufficient data	—	—	—	—	—
	BMP volume (m^3)	10	Insufficient data	—	—	—	—	—
	BMP area: watershed area	5	Insufficient data	—	—	—	—	—
	Number of storm events	29	NH_4 removal (%) = $\beta_0 + \beta_1 \times$ [number of storm events]	0.002	0.834	47.5	0.134	—
	Dry *vs.* wet	7 *vs.* 24	Permutation test for difference in means	—	**0.025**	—	—	−28.6
	Treatment train position (1-2 *vs.* 3-5)	31 *vs.* 0	Insufficient data	—	—	—	—	—
	Cumulative BMPs (1-2 *vs.* 3-5)	34 *vs.* 4	Permutation test for difference in means	—	0.408	—	—	14.0
	Vegetation (absent *vs.* present)	7 *vs.* 19	Permutation test for difference in means	—	0.237	—	—	−16.7
	Flow-weighted *vs.* raw concentration	14 *vs.* 14	Permutation test for difference in means	—	**0.038**	—	—	−23.7
	Storm flow *vs.* base flow	10 *vs.* 16	Permutation test for difference in means	—	**0.050**	—	—	−22.8

Notes: BMP, best management practice.
N, sample size (single values); sample size of group 1 *vs.* sample size of group 2 (two values).
D, difference in group means (mean of group 1 − mean of group 2).

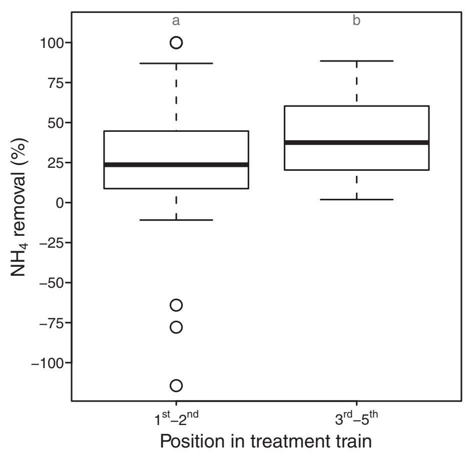

FIGURE 3. Ammonium (NH_4) Removal Efficiencies Increase for Stormwater Ponds and Wetlands Located Farther along a Treatment Train of Serially Linked Stormwater Best Management Practices (first or second position *vs.* third, fourth, or fifth position; difference in group means = −15.1, p = 0.001).

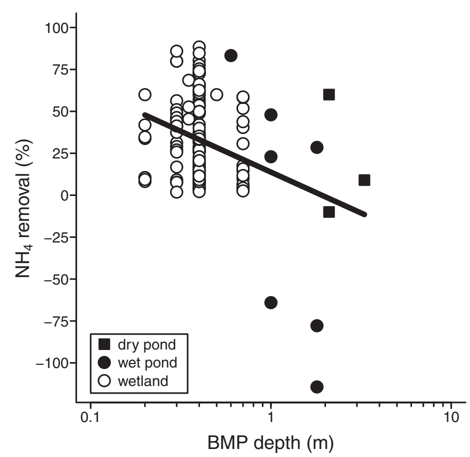

FIGURE 4. Ammonium (NH_4) Removal Efficiencies Decline with Stormwater Best Management Practice (BMP) Depth ($y = 13.8 - 48.8 \cdot \log_{10}[x]$; n = 139, R^2 = 0.111, p < 0.0001). Data are for ponds and wetlands.

cases where SW BMPs served as net sources of N to downstream waters (Figure 1). For all N species, such cases only occurred at ambient concentrations <2 mg/l, a level that is consistent with findings reported in many studies (Winer, 2000; Barrett, 2008; Chesapeake Stormwater Network, 2012). Nonetheless, despite the pattern of net N removal above this threshold, most estimates of N removal efficiency were made over time periods of hours to weeks. There was virtually no data on the long-term performance of SW BMPs.

DISCUSSION

Our analysis of empirical measurements of SW BMP performance revealed substantial variability in the ability of ponds, wetlands, and swales to remove N from surface waters. Although we identified several environmental factors that helped explain a small part of this variability, there was a remarkable lack of data on such factors in the peer-reviewed literature and in a

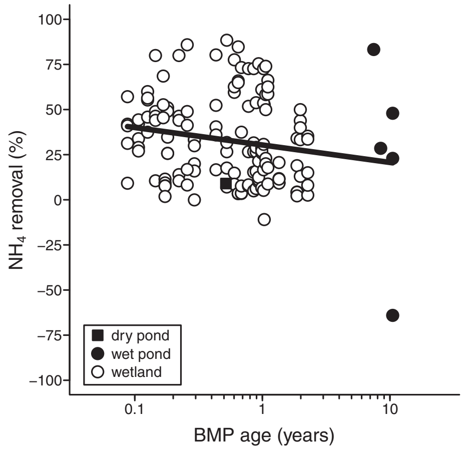

FIGURE 5. Ammonium (NH_4) Removal Efficiencies Decline with Stormwater Best Management Practice (BMP) Age ($y = 30.4 - 9.67 \cdot \log_{10}[x]$; $n = 142$, $R^2 = 0.032$, $p = 0.033$). Data are for ponds and wetlands.

major SW BMP performance database (Table 2). Environmental data were especially sparse for NO_3 and TN, two N constituents of particular concern to water quality regulators. Furthermore, available data on SW BMP performance were from assessments on time scales $\ll$1 year, meaning there is very little empirical information on the long-term effectiveness of using SW BMPs to control excess N. This is particularly concerning since many regions have high year-to-year variability in the total amount of precipitation and in the size of individual rain events—both of which influence nutrient flux to streams (Meyer and Likens, 1979; Grimm, 1987; Inamdar *et al.*, 2006). Reversing the trend of declining water quality in urban and suburban catchments will require accurate accounting of variability in SW BMP performance over a wide range of watershed and climate conditions. Such accounting can enable resource managers to know the lifetime performance of different types of SW BMPs relative to their costs, as well as when and where various types are most effective. We therefore emphasize the importance of investing in a limited number of well-designed long-term studies that fully report relevant environmental data to enable evaluation of performance under a range of real-world conditions and time spans. The results of such studies have great potential to help resource managers make decisions.

Environmental Factors Controlling SW BMP Performance

Removal of NH_4 by ponds and wetlands declined with increasing depth, volume, and relative area, suggesting that SW BMP size and shape are factors governing performance for this constituent. NH_4 is the most easily assimilated form of dissolved inorganic nitrogen, and larger, deeper ponds and wetland basins may have a reduced capacity for biofilms to take up dissolved N due to low light levels and relatively less benthic surface area and emergent vegetation in contact with water (Lee *et al.*, 2009). Very deep ponds may have anoxic conditions that inhibit coupled nitrification-denitrification and therefore minimize removal of NH_4 from the water. Deep ponds are frequently put in place to trap sediments carried in storm runoff (Bachand and Horne, 2000) and while NH_4 is known to sorb to fine sediment particles (Lee *et al.*, 2009), our results suggest that this is not a substantial sink for N. Larger stormwater ponds are effective at reducing peak storm flows to streams and as such, may not be capable of delivering both flood control and nutrient removal ecosystem services (Collins *et al.*, 2010). Stormwater BMPs are increasingly expected to perform a range of functions, including slowing the flow of water downstream, retaining sediments, and removing nutrients and other contaminants (e.g., N, phosphorus, heavy metals, fecal coliforms (USEPA, 2004), but meeting all these needs with one design may not be possible.

Our synthesis suggests that installing SW BMPs in series as part of treatment trains may be a sound approach for simultaneously maximizing N removal and peak flow and sediment reduction. N removal efficiency increased not only cumulatively but also for *individual* BMPs farther along a treatment train. Thus, treatment trains that begin with a large pond or wetland that is effective at slowing runoff and reducing particulate delivery may also minimize N export via a series of smaller cells that target dissolved nutrients (Wong *et al.*, 1999). Successively shallower cells may have high surface area-to-volume ratios, greater light penetration, and increasing water temperatures that are favorable for biotic uptake of N. Sediment forebays, which are depressions near the inlets of stormwater ponds and wetlands, can also increase retention of coarse sediments and particulate nitrogen and thereby minimize clogging of drainage infrastructure and ease maintenance requirements (USEPA, 2004).

Even given the short time intervals over which performance was measured, estimates of performance declined with SW BMP age (Figure 5). This pattern was noisy and may be due to lack of proper maintenance of SW BMPs. Others have reported that even limited maintenance (e.g., raking out detention ponds) can prevent declines in BMP performance (Erickson *et al.*, 2010). However, few stormwater ponds receive regular maintenance; for example, Klein (2012) reported that most of the estimated 32,000 SW BMPs in Maryland were maintained less than once every three years. Of the 30 studies we

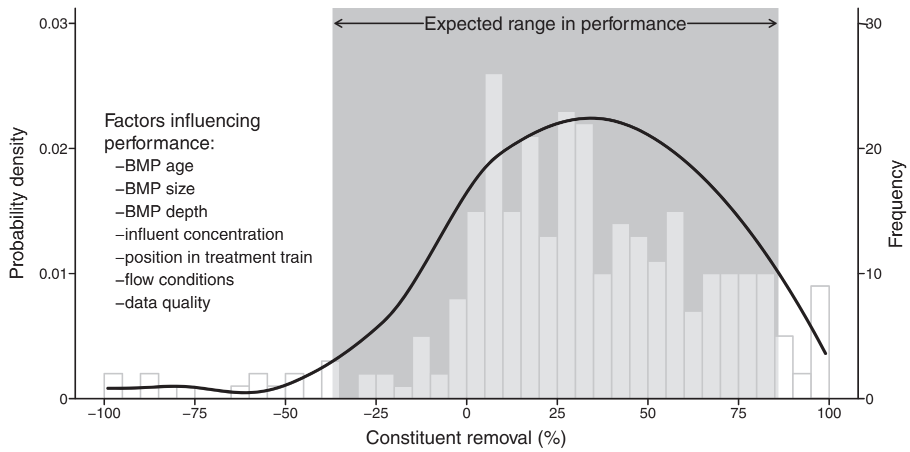

FIGURE 6. Accounting for the Full Distribution of Nitrogen Removal Efficiencies Rather Than Using a Single (median) Value Could Improve the Ability of Local Officials to Meet Nutrient Reduction Targets. Understanding the factors driving extreme values in performance can narrow the range of expected removal efficiencies for stormwater best management practices (BMPs) implemented in a specific watershed context. The histogram shows combined NH_4, NO_3, and TN removal efficiencies from the literature survey for all SW BMP types. The solid line is a scaled local polynomial regression (Loess) fit to the raw data. Observations less than −100 are not shown.

surveyed, none reported the level of maintenance performed for the BMPs, and maintenance data were reported in only 30% of cases in the International Stormwater BMP Database (Table 2). Collecting and reporting such maintenance data are crucial to evaluating why older management structures may have reduced N removal capacity.

Although the data from our literature survey showed no evidence that removal efficiency varied with loading rate (as estimated by maximum event discharge), this metric was poorly reported and others have found that hydraulic loading rate can be an important driver of SW BMP performance (Carleton *et al.*, 2001; Strecker *et al.*, 2001). Indeed, the capacity for many environmental factors to drive variation in measurements of SW BMP performance underscores the need to emphasize the full distributions characterizing performance rather than single, mean values (Figure 6).

Using Variability in Performance to Inform Decisions

Properly accounting for the full distribution of N removal efficiencies rather than using one value could impact the decisions of natural resource managers. For example, based on the magnitude of variability we observed from the field studies, a TN removal efficiency value for wet ponds that is just one standard deviation lower than the median equates to 1.4 times more N export from a moderately urbanized watershed with average N retention capacity. Assuming a 90-ha suburban watershed in the Chesapeake Bay region with average TN loading (7 kg/ha/yr) (USEPA, 2010b), moderate catchment N retention (73%) (Filoso and Palmer, 2011), and 50% of its area treated by stormwater ponds, this difference is equivalent to an additional 98 kg N exported annually, or 16% of the annual input of N to the watershed. Accounting for such variability could alter: (1) the expectations of what portion of nutrient load is controllable with SW BMPs (USEPA, 2010b), (2) how to calculate nutrient reductions for compliance with regulation of water quality trading (Maryland Department of Environment, 2011), and (3) whether to invest in expensive SW BMPs that have highly variable performance. For instance, we found evidence that ponds and wetlands of different depths and volumes fell on opposite sides of the distribution of N removal efficiencies (Figure 6) and that older ponds and wetlands underperformed compared to newer SW BMPs. Assuming that ponds and wetlands have no variation in their N removal capacities could lead to underestimates of N removal in areas with new, well-maintained, shallow cells, and overestimates in areas with old, poorly maintained, deep cells.

There are many regions where decisions on SW BMPs are based on single performance values. For example, management and restoration of the Lake Tahoe basin in California and Nevada rely on the Tahoe Integrated Information Management System which uses fixed levels of nutrient reduction by each type of SW BMP to model nutrient inputs to the basin (PLRM Development Team, 2009). Similarly, in the eastern U.S. entities such as the Chesapeake Bay Program, which involves six states and the District of Columbia, rely on models (e.g., Chesapeake Bay Watershed Phase 5.3 Model) (Chesapeake Bay Program, 2013; USEPA, 2010a) that use a single N removal efficiency value for each SW BMP type

TABLE 5. Comparison of Total Nitrogen Removal Efficiencies from Our Study and Other Published Works. Sample sizes are in parentheses if reported. Classifications of SW BMP types follow USEPA (2004).

BMP Type		This Study	International Stormwater BMP Database, 2012	Winer, 2000 (NPRPD[1])	Simpson and Weammert, 2009	Chesapeake Bay Program, 2013[2]
Ponds	Dry ponds	27% (5)	0% (32)	5% (2)	38% (4)	5%
	Dry extended detention ponds	18% (2)	—	31% (4)	31% (2)	20%
	Wet ponds[3]	40% (5)	27% (273)	33% (20)	21% (30)	20%
Wetlands	Wetlands[4]	61% (19)	2% (959)	30% (23)	19% (10)	20%
Swales	Swales[5]	50% (7)	7% (156)	84% (12)	—	10-45%[6]

Notes: BMP, best management practice. All removal efficiencies (except Chesapeake Bay Program) are medians.
[1]NPRPD: National Pollutant Removal Performance Database.
[2]Single values used in the Scenario Builder for the Chesapeake Bay Watershed Phase 5.3 Model. Values have been discounted to account for real-world performance reductions associated with poor design, installation, and maintenance of some SW BMPs.
[3]Wet ponds include "wet ponds" and "retention ponds."
[4]Wetlands include "constructed wetlands" and "wet swales."
[5]Swales include "grass channels", "grass swales", and "dry swales" without underdrains, and do not include "bioswales."
[6]The Chesapeake Bay Program provides different removal efficiencies for swales on poor and high-quality soils.

(Table 5). The Chesapeake Bay Program recognizes that those values will change according to runoff volume capacities (Chesapeake Stormwater Network, 2012; Chesapeake Bay Program, 2013). Nonetheless, for both the Chesapeake and Tahoe examples, incorporating a distribution of values into decisions or model runs could improve the ability to identify realistic N reduction strategies. Furthermore, relying on a distribution of removal efficiencies rather than on a single representative value would enable natural resource managers to assess the risk of not meeting a specified threshold level of performance for a given SW BMP. Such risk of failure is likely to be much more useful than an estimated single performance value and its associated uncertainty when deciding which and how many SW BMPs to invest in.

Resource managers and regulators will always be confronted with less information than they would like. Thus, regardless of the reasons for high uncertainty in performance of a SW BMP type (e.g., how and where it was implemented or the extent to which it was maintained), decision makers may prefer other SW BMPs with less uncertainty even if they have lower average performance. Over the long term, most nutrients are transported during few—relatively rare—high-flow events (Meyer and Likens, 1979) when nitrogen retention, even in highly engineered SW BMP structures, is low or absent (Palmer *et al.*, 2013). Compounding this problem, recurrence intervals of big storms are becoming more frequent (Groisman *et al.*, 2004, 2005), so we can expect such large, single-event deliveries to occur more frequently. Therefore, resource managers may want to select SW BMPs that continue to perform well at the highest flows to minimize the risk of relatively rare but large pulses of N that exceed total exports during all other flows. Of course resource managers may not always have such choices, as sometimes site conditions dictate the selection and placement of SW BMPs (USEPA, 2004). However, since it is unrealistic to assume that performance of a specific SW BMP will ever be known with a high level of certainty given the large number of variables that could influence performance, knowledge of relative differences in uncertainty should be used to manage risks (Peterman and Anderson, 1999). Given the lack of information on performance and the high degree of hydrologic uncertainty associated with climate change, it may make the most sense to focus on SW BMPs that operate effectively under a range of conditions and to avoid SW BMPs with widely varying performance.

Improving Stormwater Management

Rigorous and consistent monitoring of SW BMP effectiveness is crucial for informing sound management decisions. We found that some studies calculated nutrient loads by measuring concentrations and discharge at frequent intervals for multiple storm events, whereas others reported raw nutrient concentrations collected haphazardly throughout a single storm event. Many studies did not report crucial details of sample collection methods such as locations, dates, timing, discharge, or contributing watershed characteristics (Table 2). Such data could be analyzed in powerful ways. For example, sufficient geographically referenced location data would enable comparisons of SW BMP performance in different physiographic provinces or soil types. Databases such as the International Stormwater BMP Database (International Stormwater BMP Database, 2012) and

TABLE 6. Minimum Recommended Monitoring Variables for Studies of SW BMP Performance.

Data Category	Monitoring Variable
BMP	Hydraulic residence time
	Size (e.g., volume, area, length, depth)
	Location (geographically referenced coordinates)
	Date of construction
	Maintenance schedule
	Design features (e.g., type of soil, substrate, media)
Watershed	Contributing watershed area
	Impervious cover
	Number, type, and location of other watershed BMPs
Hydrology	Precipitation (site specific)
	Influent/effluent discharge
Pollutant	Influent/effluent fluxes of constituent(s)

Note: BMP, best management practice.

the National Pollutant Removal Performance Database (Winer, 2000) provide centralized catalogues of empirical SW BMP performance measurements (Table 5), however, much of the available performance data lacked information on the SW BMP design specifications, maintenance schedules, and the environmental, watershed, and hydrologic conditions necessary to identify likely drivers of variability (Table 2). Such site-specific factors may help explain the variability in average TN removal efficiencies observed across multiple databases and reviews (Table 5).

Natural resource managers and water quality regulators urgently need more complete information to reverse the continued trend of worsening urban water quality. Uncertainty in performance may always be high for some SW BMPs, however, decision makers could realize substantial benefits from even a few SW BMP monitoring efforts that are targeted at understanding how the performance of a given type of SW BMP varies with precipitation, position in the watershed, or other environmental factors. In addition, enhancing data quality by standardizing the methods used to assess SW BMP performance would strengthen the ability of managers to draw strong inferences from multiple independent studies. Based on our analysis, we recommend that long-term monitoring of SW BMPs includes a minimum set of variables on BMP size, location, watershed characteristics, precipitation, discharge, and nutrient flux (Table 6). Detailed data on local hydrologic conditions are especially needed. More comprehensive reporting of site conditions and environmental factors could greatly expand our understanding of the controls on SW BMP performance by enabling a thorough analysis of interactions among variables. Such information is vital given that SW BMPs are increasingly expected to provide multiple freshwater ecosystem services in the face of growing urban populations and climate change.

CONCLUSIONS

The health of freshwater ecosystems depends on effective implementation of SW BMPs. Our synthesis of available monitoring data revealed wide variability in SW BMP performance and found only limited evidence for factors such as BMP size, age, and position in the watershed explaining this variability. In light of our findings, we offer two broad recommendations for improving SW BMP implementation: (1) Properly accounting for the full distribution of SW BMP performance in setting nutrient reduction goals, and (2) Targeted long-term monitoring of SW BMPs that include standardized measurements of environmental factors and nutrient loads.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Keyword search terms used to identify papers.

Table S2. References of papers included in the data synthesis.

ACKNOWLEDGMENTS

We thank B. Bierwagen, K. Boomer, J. Carleton, N. Detenbeck, L. Dube, S. Filoso, D. Hogan, T. Johnson, S. Julius, J. Loperfido, G. Moglen, K. Prestegaard, S. Schwartz, R. Shedlock, D. Weller, and J. West for discussions and ideas that helped shape this research. Comments from five anonymous reviewers improved the manuscript. This project was funded by the USEPA Global Change Research Program (GS-10F-0502N to MAP) and received synthesis support from the National Socio-Environmental Synthesis Center (SESYNC) under funding from NSF (DBI-1052875). The views expressed in this article are those of the authors and do not necessarily reflect the policies of the USEPA.

LITERATURE CITED

Ahiablame, L.M., B.A. Engel, and I. Chaubey, 2012. Effectiveness of Low Impact Development Practices: Literature Review and Suggestions for Future Research. Water, Air, and Soil Pollution 223:4253-4273.

Ator, S.W., J.W. Brakebill, and J.D. Blomquist, 2011. Sources, Fate, and Transport of Nitrogen and Phosphorus in the Chesapeake Bay Watershed: An Empirical Model. U.S. Geological Survey Scientific Investigations Report 2011-5167, 27 pp.

Bachand, P.A.M. and A.J. Horne, 2000. Denitrification in Constructed Free-Water Surface Wetlands: I. Very High Nitrate Removal Rates in a Macrocosm Study. Ecological Engineering 14:9-15.

Barrett, M.E., 2005. Performance Comparison of Structural Stormwater Best Management Practices. Water Environment Research 77:78-86.

Barrett, M.E., 2008. Comparison of BMP Performance Using the International BMP Database. Journal of Irrigation and Drainage Engineering 134:556-561.

Beersma, J.J. and T.A. Buishand, 1999. A Simple Test for Equality of Variances in Monthly Climate Data. Journal of Climate 12:1770-1779.

Booth, D.B., D. Hartley, and R. Jackson, 2002. Forest Cover, Impervious-Surface Area, and the Mitigation of Stormwater Impacts. Journal of the American Water Resources Association 38:835-845.

Carleton, J.N., T.J. Grizzard, A.N. Godrej, and H.E. Post, 2001. Factors Affecting the Performance of Stormwater Treatment Wetlands. Water Research 35:1552-1562.

Castro, M.S., C.T. Driscoll, T.E. Jordan, W.G. Reay, and W.R. Boynton, 2003. Sources of Nitrogen to Estuaries in the United States. Estuaries 26:803-814.

Chen, C.-F., J.-Y. Lin, C.-H. Huang, W.-L. Chen, and N.-L. Chueh, 2009. Performance Evaluation of a Full-Scale Natural Treatment System for Nonpoint Source and Point Source Pollution Removal. Environmental Monitoring and Assessment 157:391-406.

Chesapeake Bay Program, 2013. Estimates of County-Level Nitrogen and Phosphorus Data for Use in Modeling Pollutant Reduction: Documentation for Scenario Builder Version 2.4. Revised January 2013.

Chesapeake Stormwater Network, 2012. Recommendations of the Expert Panel to Define Removal Rates for New State Stormwater Performance Standards, pp. 1-29.

Collins, K.A., T.J. Lawrence, E.K. Stander, R.J. Jontos, S.S. Kaushal, T.A. Newcomer, N.B. Grimm, and M.L. Cole Ekberg, 2010. Opportunities and Challenges for Managing Nitrogen in Urban Stormwater: A Review and Synthesis. Ecological Engineering 36:1507-1519.

Davis, A.P., 2005. Green Engineering Principles Promote Low-Impact Development. Environmental Science & Technology 39:338A-344A.

Dietz, M.E., 2007. Low Impact Development Practices: A Review of Current Research and Recommendations for Future Directions. Water, Air, and Soil Pollution 186:351-363.

Erickson, A.J., J.S. Gulliver, J.-H. Kang, P.T. Weiss, and C.B. Wilson, 2010. Maintenance for Stormwater Treatment Practices. Journal of Contemporary Water Research & Education 146: 75-82.

Filoso, S. and M.A. Palmer, 2011. Assessing Stream Restoration Effectiveness at Reducing Nitrogen Export to Downstream Waters. Ecological Applications 21:1989-2006.

Grimm, N.B., 1987. Nitrogen Dynamics during Succession in a Desert Stream. Ecology 68:1157-1170.

Groisman, P.Y., R.W. Knight, D.R. Easterling, T.R. Karl, G.C. Hegerl, and V.N. Razuvaev, 2005. Trends in Intense Precipitation in the Climate Record. Journal of Climate 18:1326-1350.

Groisman, P.Y., R.W. Knight, T.R. Karl, D.R. Easterling, B. Sun, and J.H. Lawrimore, 2004. Contemporary Changes of the Hydrological Cycle over the Contiguous United States: Trends Derived from In Situ Observations. Journal of Hydrometeorology 5:64-85.

Hatt, B.E., T.D. Fletcher, C.J. Walsh, and S.L. Taylor, 2004. The Influence of Urban Density and Drainage Infrastructure on the Concentrations and Loads of Pollutants in Small Streams. Environmental Management 34:112-124.

Hogan, D.M. and M.R. Walbridge, 2007. Best Management Practices for Nutrient and Sediment Retention in Urban Stormwater Runoff. Journal of Environmental Quality 36:386-395.

Howarth, R., D. Swaney, G. Billen, J. Garnier, B. Hong, C. Humborg, P. Johnes, C.-M. Mörth, and R. Marino, 2012. Nitrogen Fluxes from the Landscape Are Controlled by Net Anthropogenic Nitrogen Inputs and by Climate. Frontiers in Ecology and the Environment 10:37-43.

Inamdar, S.P., N. O'Leary, M.J. Mitchell, and J.T. Riley, 2006. The Impact of Storm Events on Solute Exports from a Glaciated Forested Watershed in Western New York, USA. Hydrological Processes 20:3423-3439.

International Stormwater BMP Database, 2012. International Stormwater BMP Database Version 01 04 12. www.bmpdatabase.org, *accessed* June 2012.

Klein, R., 2012. Making StormPrint Better: A Critical Step in Restoring the Chesapeake Bay through Expanded Citizen Involvement in Stormwater BMP Maintenance. Community & Environmental Defense Services, Owings Mills, Maryland.

Lee, C., T.D. Fletcher, and G. Sun, 2009. Nitrogen Removal in Constructed Wetland Systems. Engineering in Life Sciences 9:11-22.

Lee, J.G., A. Selvakumar, K. Alvi, J. Riverson, J.X. Zhen, L. Shoemaker, and F. Lai, 2012. A Watershed-Scale Design Optimization Model for Stormwater Best Management Practices. Environmental Modelling & Software 37:6-18.

Manly, B.F.J., 2007. Randomization, Bootstrap and Monte Carlo Methods in Biology. Chapman & Hall, Boca Raton, Florida.

Maryland Department of Environment, 2011. Accounting for Stormwater Wasteload Allocations and Impervious Acres Treated: Guidance for National Pollutant Discharge Elimination System Stormwater Permits. Maryland Department of the Environment, Baltimore, Maryland.

Meyer, J.L. and G.E. Likens, 1979. Transport and Transformation of Phosphorus in a Forest Stream Ecosystem. Ecology 60:1255-1269.

Palmer, M.A., S. Filoso, and R.M. Fanelli, 2013. From Ecosystems to Ecosystem Services: Stream Restoration as Ecological Engineering. Ecological Engineering, doi:10.1016/j.ecoleng.2013.07.059.

Peterman, R.M. and J.L. Anderson, 1999. Decision Analysis: A Method for Taking Uncertainties into Account in Risk-Based Decision Making. Human and Ecological Risk Assessment: An International Journal 5:231-244.

PLRM Development Team, 2009. Pollutant Load Reduction Model (PLRM) Model Development Document. Tahoe Integrated Information Management System, Santa Cruz, California.

Simpson, T. and S. Weammert, 2009. Developing Nitrogen, Phosphorus and Sediment Reduction Efficiencies for Tributary Strategy Practices: BMP Assessment, Final Report. University of Maryland, Mid-Atlantic Program, College Park, Maryland.

Strecker, E.W., M.M. Quigley, B.R. Urbonas, J.E. Jones, and J.K. Clary, 2001. Determining Urban Storm Water BMP Effectiveness. Journal of Water Resources Planning and Management 127:144-149.

Urbonas, B. and C.C. Olson, 2011. Assessment of Stormwater BMP Cost Effectiveness. Stormwater 12:10-21.

USEPA, 2004. The Use of Best Management Practices (BMPs) in Urban Watersheds. U.S Environmental Protection Agency, National Risk Management Research Laboratory, Edison, New Jersey.

USEPA, 2010a. Chesapeake Bay Phase 5.3 Community Watershed Model. EPA 903S10002 - CBP/TRS-303-10. U.S. Environmental Protection Agency, Chesapeake Bay Program Office, Annapolis, Maryland.

USEPA, 2010b. Chesapeake Bay Total Maximum Daily Load for Nitrogen Phosphorus and Sediment. U.S Environmental Protection Agency, Water Protection Division, Philadelphia, Pennsylvania.

USEPA, 2013. Water Quality Assessment and Total Maximum Daily Loads Information. www.epa.gov/waters/ir/, *accessed* April 2013.

Walsh, C.J., T.D. Fletcher, and A.R. Ladson, 2005a. Stream Restoration in Urban Catchments through Redesigning Stormwater Systems: Looking to the Catchment to Save the Stream. Journal of the North American Benthological Society 24:690-705.

Walsh, C.J., A.H. Roy, J.W. Feminella, P.D. Cottingham, P.M. Groffman, and R.P. Morgan, 2005b. The Urban Stream Syndrome: Current Knowledge and the Search for a Cure. Journal of the North American Benthological Society 24:706-723.

Winer, R., 2000. National Pollutant Removal Performance Database for Stormwater Treatment Practices (Second Edition). Center for Watershed Protection, Ellicott City, Maryland.

Wong, T.H.F., P.F. Breen, and N.L.G. Somes, 1999. Ponds *vs*. Wetlands - Performance Consideration in Stormwater Quality Management. Proceedings of the 1st South Pacific Conference on Comprehensive Stormwater and Aquatic Ecosystem Management, Auckland, New Zealand, pp. 223-231.