Review Article

The Environmental and Ecological Benefits of Green Infrastructure for Stormwater Runoff in Urban Areas

Miquel A. Gonzalez-Meler1*, Lisa A. Cotner1, Dean A. Massey2, Moira L. Zellner2,3 and Emily S. Minor1,3

1Ecology and Evolution Group, Department of Biological Sciences, University of Illinois at Chicago, USA
2Department of Urban Planning and Policy, University of Illinois at Chicago, USA
3Environmental Institute of Science and Policy, University of Illinois at Chicago, USA

Abstract

Water runoff from impervious surfaces threatens urban ecosystems, public health and property values. Traditional stormwater management systems are often overwhelmed after big storms, prompting the evaluation of alternative green infrastructure (GI) strategies to improve stormwater management. Here, we present a synthesis to determine the effectiveness of GI—detention basins, filtration devices, biofiltration, constructed wetlands, green roofs, and permeable pavement—in reducing runoff volumes and peak flows and in mitigating water pollutant loads by testing and using surrogates such as total suspended solids (TSS) and total nitrogen (TN) from storm runoff. In general, all infrastructures reduced stormwater quantity and/or improved runoff water quality at a local scale, and their performance was comparable to more traditional stormwater management approaches (i.e., detention basins). There was a general agreement between the peer-reviewed data and the best management practice (BMP) database for most GI effectiveness, particularly with respect to water quality. Our analysis shows, however, that the effectiveness of most GI was highly variable, possibly due to climate, influent concentration, or scale. Despite the variability in stormwater runoff performance, most GI can potentially provide valuable habitat for wildlife in urban settings. GI can be designed to promote additional ecosystem services in urban areas, such as habitat for flora or pollinators that can aid in urban gardens or C sequestration, among many others.

INTRODUCTION

Stormwater management is a challenge faced by urban and agricultural development at all spatial scales. As urban development increases the percent of impermeable cover within a watershed [1], stormwater volume, peak flow and concentration of non-point source pollutants increase [2]. In urban areas, traditional gutter and storm sewer systems are often inadequate for reducing the quantity of stormwater or decreasing pollutant loads [3]. In agricultural or rural areas, drainage systems quickly channel large volumes of water, sediment, and dissolved pollutants to waterways [4]. In both urban and rural settings, inadequate stormwater management can lead to flooding, erosion, and impaired aquatic habitats [5]. Best management practices (BMPs), such as detention and retention basins, are typically recommended for planning agencies to control discharge rates in developed and developing areas [6]. However, the effectiveness of such measures is very sensitive to context and, therefore, the identification of a “best” practice is proven to be difficult [7]. Traditional infrastructures fail to support other functions of a sustainable system, especially habitat and groundwater recharge [8]. The need for improved stormwater management has increased interest in the use of green infrastructure (GI). Local and state governments are increasingly promoting the use of GI mostly based on private and public reports such as those included in the US Environmental Protection Agency (EPA)/American Society of Civil Engineers (ASCE) International Stormwater Best Management Practices (BMP) Database (see Section 5)—or its perceived potential effectiveness. Here, we present a synthesis of peer-reviewed data on the effectiveness of GI to assess its viability for stormwater management, including policy and ecological implications.
At a regional scale, GI is broadly defined as a network of green spaces that provide natural ecosystem function and benefits to people through recreation, aesthetics and ecosystem services [9-11]. In the context of stormwater management, GI includes low impact development (LID) strategies implemented at the site level which aim to minimize the generation of runoff and associated pollution by using natural systems to collect, treat, and infiltrate rain where it falls [12]. For instance, runoff volume can be reduced through infiltration, evaporation, and evapotranspiration by plants [13]. Mechanisms for pollution removal include sedimentation, plant uptake [14], filtration [15], biofiltration [16], biodegradation and sorption [17]. GI such as swales or constructed wetlands are designed to achieve both runoff quantity and quality goals, while others are primarily designed to improve water quality (e.g., filters, green roofs, permeable pavement) or to reduce runoff volume and/or peak flow (e.g., rain barrels) [8,18]. Restored wetlands and vegetated swales will also contribute to reduce carbon emissions from cities [19], although these may be offset by CH₄ emissions. While individual assessments of GI performance have been conducted, a systematic comparison of their effectiveness has not been documented. Here, we consolidate and analyze effectiveness data for all documented categories of GI from peer-reviewed literature to help inform decisions on its adoption.

A few comparisons of multiple infrastructure have been limited to cost-benefit analyses [12], to description of a few case studies [8], and have rarely evaluated relative effectiveness [20]. We found sufficient published data to examine effectiveness of biofiltration [21-35], constructed wetlands [34,36-43], filtration systems [28,34,44-49], green roofs [50-55], and permeable pavement [34,56-68] and to compare them to traditional stormwater approaches like retention and detention basins [34,44,69-75]. We examined possible sources of variability in the effectiveness of GI, including design and scaling, maintenance, and geographical differences. We further compare our findings to the BMP Database, a prominent resource for practitioners, to identify commonalities, disparities, and shortcomings in both sources of information. We also discuss the policy and ecological implications (as GI may create habitat) of our findings.

INDICATORS OF EFFECTIVENESS

We selected runoff volume, peak flow, total suspended solids (TSS), and total nitrogen (TN) to evaluate the effectiveness of GI. Each of these relevant and widely used factors is representative of common stormwater management challenges in urban areas [5,76]. Reducing runoff volume (the amount of surface water resulting from a given storm event) and peak flow (maximum runoff volume per unit time) are fundamental goals for most green (and traditional) stormwater infrastructure, becoming key performance standards for implemented GI [77,78]. Changes in runoff volume and peak flow are responsible for many of the negative impacts of stormwater associated to urbanization (e.g., flooding, combined sewer overflows, erosion, low baseflow, or streambank entrenchment) [5, 12].

In addition to effects of GI on runoff quantity, GI has the potential to improve water quality. Two commonly used surrogate measures of water quality by agencies that regulate urban nonpoint source pollution are TSS (i.e., the amount of particulate matter suspended in water) and TN [79]. Suspended solids in stormwater can cause sedimentation in rivers and streams as well as transfer heavy metals and phosphorous over long distances [80-82]. The proportion of heavy metals in TSS and the size distribution of heavy metal particulates can vary between storms and across sites with different land uses, pH, antecedent dry period, and characteristics and TSS quantity [83, 84]. However, environmental protection agencies need to balance cost effectiveness with precision, resulting in the widespread use of TSS as a surrogate [85]. As a precursor to our analysis of GI effectiveness, we have confirmed that TSS reduction is a good measure of heavy metal reduction in water treated by green infrastructure.

Total nitrogen best describes the behavior of dissolved pollutants in general and nitrogen species in particular [82, 86]. Measuring TN in runoff waters is particularly important because excess nitrogen causes eutrophication and algal blooms, leading to reductions in dissolved oxygen and degradation of aquatic communities [87]. When used in combination with runoff volume and peak flow, TSS and TN represent a thorough way to evaluate and compare the ability of GI to mitigate threats to ecosystem and human health caused by stormwater.

Water quality can be measured in terms of concentration of pollutants (e.g., mg L⁻¹) or in terms of the mass or amount of a pollutant in water runoff, called load. Different scenarios can lead GI to reduce volume runoff, effectively decreasing pollutant loads without necessarily affecting TSS or TN effluent concentrations. If the ultimate goal is to improve water quality downstream, measuring total pollution loads in effluent runoff waters is as important as the effluent concentration because it also gives credit to techniques that reduce effluent volume. Unfortunately, effluent loads have not been traditionally reported in the literature (we found 39 and 29 GI sites reporting TSS and TN load reductions respectively, out of 214 sites with pollutant concentration data). Most studies report pollutant percent removal as the metric to evaluate GI effectiveness.

Pollutant percent removal has the benefit of standardizing pollution removal at the effluent normalized by influent quantity, allowing for direct comparisons between GIs with vastly different conditions [88]. However, pollutant percent removal may vary greatly depending on the size and frequency of storm events that can saturate water quality improvement capacity of GIs but still reduce pollutant loads by reducing runoff volume [88]. Effectiveness of GI is often reported as difference between effluent and influent values. However, when influent concentration of pollutants is already low, the resulting percent removal would be marginal providing little information on effectiveness of GIs under these conditions. Despite its shortcomings, percent concentration removal is the most widely reported effectiveness metric in the literature and it allows for comparison among many types of GI, particularly when effluent and influent data are not reported. Pollutant percent removal is also a scalable metric when used in combination with volume and peak flow reductions to estimate pollutant loads.

**Data selection criteria and data analyses**

Our review focused on five categories of GI (green roofs,
permeable pavement, constructed wetlands, bioinfiltration, and detention/retention basins) for which sufficient published data exist to evaluate reduction of peak flow, mitigation of runoff volume, performance efficiency of TSS and TN concentration, and pollutant load reduction (Table 1). We used keyword searches on ISI Web of Knowledge to compile an initial database of articles that were published/indexed prior to October 2010, returning 490 citations (Appendix I). About 236 of these citations contain data on GI effectiveness. From this subset, articles that had no replication in either space or over time, that did not specify the number of storm events monitored, or that focused exclusively on quantifying within-infrastructure characteristics such as the distribution of pollutants among sediment layers were also excluded. This resulted in 66 replicated studies with a total of 219 different sites (67% field sites, the remainder replicated lab studies) included in this review. Some articles contained data on more than one type of GI. To simplify terminology throughout the paper, we used the term “site” to indicate a single infrastructure configuration that was monitored over time, whether it was a set of all identical replicates in a particular laboratory study or a field site such as permeable pavement installations, bioinfiltration, constructed wetlands, and green roofs.

From each selected entry the following information was tallied for each GI type: the number of infrastructures (or infrastructure configurations) monitored over time, the number of storm events monitored, percent change in peak flow, runoff volume, TSS, TN, load data when available, and standard deviations (Table 2). We first confirmed that TSS removal was a reliable surrogate for heavy metal removal across green infrastructures. We found that reduction in TSS concentration (in mg L\(^{-1}\)) correlated with reduction in concentration of zinc, lead, copper, nickel, and manganese (in μg L\(^{-1}\)) for 34 GI sites including detention basins, bioinfiltration facilities, buffers, constructed wetlands, and filtration systems (\(R^2 = 0.90\), equation: y = 0.35x + 50.8). Analyses were based on the reduction between the influent and effluent concentration (in μg L\(^{-1}\)) of the two metals with the highest influent concentrations for each site. Total suspended solid reduction did not correlate well with reductions in soluble pollutants such as TN (\(R^2 = 0.004\)). It is noted that although TSS is a good general surrogate for heavy metals, TSS cannot be used when GI is implemented to reduce contamination of specific metal pollutants. In these cases, the specific pollutants of interest should be monitored.

Using TSS and TN as pollution indicators, we compared the relative effectiveness of GI at removing both particulate and dissolved fractions of contaminants, and summarized data

Table 1: Tally of Sites and Peer-Reviewed Publications for Each Green Infrastructure Category. The symbol * denotes instances with not enough data available for analysis of concentration reduction, peak flow reduction, or runoff volume reduction.

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>Total Sites</th>
<th>Total Articles</th>
<th>TN Sites</th>
<th>TSS Sites</th>
<th>Peak Flow Sites</th>
<th>Runoff Volume Sites</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bioinfiltration</td>
<td>84</td>
<td>20</td>
<td>43</td>
<td>76</td>
<td>7</td>
<td>6</td>
<td>Vegetated systems designed to facilitate the infiltration of stormwater and remove pollutants through infiltration media and / or vegetation uptake. Examples: bioretention areas, swales, infiltration basins</td>
</tr>
<tr>
<td>Detention/ Retention</td>
<td>15</td>
<td>12</td>
<td>6*</td>
<td>14</td>
<td>0*</td>
<td>0*</td>
<td>A traditional best management practice where water is detained in a manmade pond to reduce peak flow and allow pollutants to settle.</td>
</tr>
<tr>
<td>Permeable Pavement</td>
<td>23</td>
<td>14</td>
<td>8</td>
<td>11</td>
<td>5</td>
<td>12</td>
<td>Pavement which allows stormwater to infiltrate into underlying soil. Filters some pollutants.</td>
</tr>
<tr>
<td>Filtration</td>
<td>31</td>
<td>10</td>
<td>5*</td>
<td>31</td>
<td>0*</td>
<td>0*</td>
<td>A variety of devices which actively or passively filter pollutants out of stormwater. Many are proprietary designs. Often used in conjunction with other green infrastructure.</td>
</tr>
<tr>
<td>Green Roof</td>
<td>9</td>
<td>6</td>
<td>1*</td>
<td>0*</td>
<td>6</td>
<td>6</td>
<td>Roofs with a vegetated surface and substrate designed to reduce runoff through transpiration and evaporation and filter rainwater through media, vegetation, and geotextiles</td>
</tr>
<tr>
<td>Constructed Wetland</td>
<td>48</td>
<td>9</td>
<td>24</td>
<td>39</td>
<td>3*</td>
<td>0*</td>
<td>Manmade wetland intended to intercept runoff, reduce peak flows, decrease runoff volume and mitigate pollution</td>
</tr>
</tbody>
</table>

Table 2: Weighted Mean Percent Reductions and Estimated Load Reduction of Green Infrastructure. Storms is the number of storm events; Mean is the weighted mean percent; SD stands for weighted standard deviation; TSS load reduction is the estimated percent load reduction calculated from mean concentration and runoff volume reductions; TN load reduction is calculated from mean concentration, * represents instances with insufficient sample size to calculate weighted means.

<table>
<thead>
<tr>
<th>Infrastructure Type</th>
<th>Peak Flow Reduction</th>
<th>Runoff Volume Reduction</th>
<th>Total Suspended Solids Load Reduction</th>
<th>Total Suspended Solids Concentration Reduction</th>
<th>Total Nitrogen Concentration Reduction</th>
<th>Total Nitrogen Load Reduction</th>
<th>Estimated Load Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bioinfiltration</td>
<td>Storms 131 52 17 241 85 27 642 78 35 45 22 109 491 17 73 111 43 26 97 88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constructed Wetland</td>
<td>Detention * * * * * * 87 64 29 22 -8 152 * * * 20 21 93 * *</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filtration</td>
<td>* * * * * * * 323 59 36 61 94 3 * * 46 40 3 59 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green Roof</td>
<td>94 65 18 570 57 13 * * * * * * * * * * * * * * * -39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permeable Pavement</td>
<td>143 70 7 596 68 28 253 66 32 * * * 182 58 40 * * * 89 86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
on percent reduction of event mean concentration, or removal efficiency. In sites where removal efficiencies for either TSS or TN were not reported but runoff pollutant concentrations were reported for both treated and untreated water, we calculated the percentage difference between treated and untreated water for each storm event and used the average [71]. Average removal efficiencies (in percentages) for all sites and available pollutant concentration data (in mg L\(^{-1}\)) are presented in Appendix II.

To evaluate the reduction of TSS and TN loads by GI, we summarized effectiveness data by calculating the weighted mean and standard deviation percent load reduction for: bioinfiltration [21, 89-93], constructed wetlands [37], detention [94-96], and filtration [91, 97, 98]. Mean load reduction was calculated separately from mean concentration reduction because loads, measured in units of mass, are not directly comparable with concentrations (in mg L\(^{-1}\)). Because load reduction data were only directly available for a small subset of sites and GIs, we derived a second estimate of load reduction from the reported mean concentration reductions for TSS and TN and mean runoff volume reduction:

\[
L = 1 - ((1 - C) \times (1 - V))
\]

(1)

Where load reduction is represented by L, C is the weighted mean concentration reduction, and V is the weighted mean volume reduction.

Weighted averages of all parameters were used to minimize biases in the data introduced by uneven sample sizes between studies [99]. We calculated the weighted average reduction and weighted standard deviations for runoff volume and peak flow for each GI by the number of sites. Some studies considered between-storm-event variability, and thus reported averages and standard deviation or variance for TSS and TN removal efficiencies. We estimated between-storm-event variance for sites where these data were not reported [100], and used a weighted variance equation which includes both between-storm-event and between-site variation to calculate standard deviations [99, 100]. All these methods and analyses ensure the metrics used are readily scalable to consider issues of GI size, runoff volume and contaminant loads.

Sample sizes used in statistical analyses ranged from 9 (detention) to 73 sites (bioinfiltration) for TSS and 8 (permeable pavement) to 40 (bioinfiltration) sites for TN. There were insufficient data to calculate weighted average removal efficiency for either TSS or TN by buffers or green roofs or for TN by filtration. The few data reported for each GI did not allow us to establish significant differences in effectiveness across GI types. Nevertheless, we provide general trends, and establish general correlations between the average influent concentrations and pollutant percent removal based on the data.

**GREEN INFRASTRUCTURE PERFORMANCES**

**Water Quantity: Runoff Volume and Peak Flow**

A major attribute of stormwater infrastructures is retention and storage of water to avoid flooding [8]. Runoff quantity is also of particular importance because it is strongly related to pollution removal; reductions in runoff volume, even absent of any change in pollutant concentration, would result in lower total pollutant load entering stormwater systems and waterways.

Eighteen sites reported data on peak flow and 24 sites reported data on runoff volume reduction (Table 1). Permeable pavement, bioinfiltration, and green roofs reduced both peak flow and runoff volume from 52 to 85% (Figure 1). We did not have sufficient data from wetlands to calculate their effectiveness, with wide range reductions in peak flow (13-77%) and runoff volume (9-27%). Studies using detention or filtration systems met the research criteria for peak flow reduction. Filtration is designed exclusively for water quality improvement, not quantity, so that runoff volume or peak flow data are typically not reported for this GI type.

Reductions of at least 50% in peak flow and runoff by GIs will minimize impacts on downstream riparian ecosystems [5,101] while increasing groundwater recharge [102,103]. However, effluent flow rates, storm characteristics and GI design may cause variability in the performance of GI at reducing peak flow [31,67]. For instance, bioinfiltration basins reduced peak flow by 64% on average, but average peak flow reduction for storms with 2.54 cm or more of rain was only 25% [31]. Despite performance variability, GI have the potential to reduce peak flow at the neighborhood scale by 30 to 50% [102] when compared with grey infrastructure that contribute to high peak flows and provide no control of runoff volume [5]. This potential has been confirmed in field studies [50, 53].

**Water Quality: TSS and TN Concentration**

GI generally succeeds in reducing both TSS and TN event mean concentration, removing between 58 and 86% of TSS (Figure 1, TSS). Constructed wetlands were the least consistent in their performance (131% standard deviation compared to 28% in detention basins). This may be due to wetland’s sediment release to the effluent during large storm events [37]. Biologically-driven systems, such as bioinfiltration, perform better than hydrodynamic separators or filtration systems (Figure 1). Our findings are validated by a comparison of small-scale GI types under identical conditions that demonstrated at least 50% reduction in TSS [34]. Reduction of TSS levels by GIs would minimize associated negative impacts on aquatic ecosystems related to sediment deposition and heavy metals.

Concentration of TN of effluent waters was reduced by GIs but perhaps to a lesser extent than for TSS (Figure 1). None of the GIs consistently reduced the concentration of TN by more than 58 percent (Figure 1, TN), compared to the 58 - 80 % reduction of TSS (Figure 1, TSS). Weighted standard deviations of TN removal ranged from 10% (for filtration devices) to 73% (for bioinfiltration), indicating a high degree of variability in effectiveness between GI types. Some studies, particularly laboratory ones, had negative TN removal efficiencies. For example, vegetated bioinfiltration column mesocosms with some plant species such as Dianella revoluta show a net increase in nitrate in effluent waters [24] but in the field, bioinfiltration systems planted with Dianella spp. resulted in a statistically significant 19% reduction in TN [13]. In general, dissolved pollutants are generally more difficult to remove.
from stormwater than particulate pollution [36,104]. Although previous studies have found that lower influent concentrations of pollutants are related to lower removal efficiencies [94,105], the removal efficiency of pollutants by GIs was not related to influent concentration across all sites for both TSS (Figure 2) and TN (not shown). This further confirms the use of removal efficiency as a suitable measure of GI effectiveness for TSS and TN.

Load Reductions

Pollutant load, the total mass of pollution in runoff, can be reduced by decreasing pollutant concentration, by reducing runoff volume, or both [35]. Load reduction efficiency data were available for a subset of 29 sites for TN and 39 sites for TSS. Average TSS load reductions ranged from -7.6% for detention (n = 6) to 94% for filtration (n=5) (Table 2). Average TN load reductions ranged from 21% for detention (n=4) to 43% for bioinfiltration (n=7).

Reduction in effluent volume can indirectly improve water quality because water that stays on site will not carry pollution downstream [35] unless effluent pollutant concentration increases [90]. The reduced runoff volume, in combination with TSS and TN concentration reductions, means that GI practices could decrease the pollution load in effluent water. Concentration removal efficiency was generally high for TN and TSS (Table 2; Figure 1&2) and, combined with average volume reduction, was used to produce a rough estimate of expected average load reduction performance of GIs (Table 2). We estimated TSS load reductions at 97% for bioinfiltration and 89% for permeable pavement. Load reduction estimates for TN by bioinfiltration and permeable pavement were both approximately 86%. Results suggest that GI has great potential to reduce pollutant loads, whether through decreasing runoff volume, pollutant concentration, or both.

**Sources of Variation in Green Infrastructure Effectiveness**

Although GI generally reduced runoff volumes, peak flows, and pollutants, results indicate a high degree of variability in the effectiveness of a given infrastructure type, particularly for constructed wetlands (Figure 1, TN & TSS). Peak flow and runoff volume reductions may depend partially upon storm event and catchment characteristics [102,106]. The small number of studies also contributed to the variability in mean effectiveness. This study identifies sources of variability with regard to differences in design, scaling, geographical distribution, climate, and maintenance. Although data were not always sufficient to directly test whether these variables influenced the effectiveness of GI, they show emerging patterns in each potential source of variability.

**Design Variation and Effectiveness**

GI design should be site specific, considering local soil type, precipitation patterns, magnitude of storm events, the size and percent impervious cover in the catchment, and water quality problems and goals [52,106-108]. Water capacity and retention times are highly dependent upon infrastructure design and generally have a positive relationship with peak flow reduction and water quality improvement [33,108]. For example, the location of an infrastructure’s outflow can be designed to maximize both water residence time and pollution reduction [102] as well as to maximize habitat connectivity. Therefore, important aspects of design include scaling, functional components, such as vegetation, maintenance and climate.

**Scaling:** Infrastructure should be properly scaled for its drainage area by considering the proportion of impervious surface and historical rainfall to scale a given GI for a given size of storm [109]. Improperly sized GI may not meet performance expectations, especially if runoff saturates the infrastructure or has an insufficient water residence time. Scaling variables are site-dependent, and an ideal ratio of infrastructure size to drainage area has not been defined [42,110,111], except for constructed wetlands where a wetland-to-drainage ratio of at least 1:100 is recommended [37]. A comparison between sites that differed in wetland size to drainage area ratio found that properly scaled wetlands were no more effective at removing TSS or TN from effluent than those with a ratio below 0.01 (TSS: p = 0.94, TN: p = 0.09). Parameters such as water retention time and hydraulic loading rates (the amount of water flowing into a wetland each day) could be more important determinants of effectiveness than size as shown elsewhere [37]. This suggests that wetland area to
drainage area ratio need to be reevaluated as an adequate scaling unit. Wetlands which have a large storage volume in relation to their catchment area have been found to be efficient at removing nitrogen [41]. Therefore, a water volume ratio (volume of water discharged over volume of water retained by wetlands) may be more appropriate.

Functional Design Variation: Functional components of GI include infiltration substrate, vegetation components, underdrain bypasses or filtration media. Substrate choice in infiltration systems, such as the percentage of fine grains, impact the ability for a GI to retain and filter stormwater [90]. Hsieh [112] found that mulch or mulch mixed with sand and soil outperformed sand and soil substrates at removing nitrate, while Cho et al., [92] showed that increasing clay and silt content of bioinfiltration media can decrease infiltration rate and nitrogen removal. In bioinfiltration systems, plant species can influence GI performance because pollutants are differentially sequestered within tissues of plants, with some plant species being up to 20 times better at removing pollutants than other species [113]. Green roofs reduce peak flow and runoff volume by about 60% (Table 2). However, extensive green roofs, typically planted with sedum and other groundcover plants, are less effective at reducing runoff volume than intensive green roofs, which are planted with more substantial vegetation [114,115]. In addition, there is often variability in plant uptake rates of nitrogen over the growing season [31] so TN removal may not be as efficient by green roofs.

Due to site limitations, some GI cannot be designed to fully handle peak water volume. Underdrains, designed to transmit water to another green or grey stormwater infrastructure in situations where infiltration of all water is impractical, may negatively impact the effectiveness of other GIs such as rain garden bioinfiltration [27,116]. Similarly, infiltration basins are frequently designed to handle a 5 or 10-year storm; beyond that storm size, water bypasses the infrastructure, causing high levels of performance variability between storms [117]. Filter designs are diverse and utilize many possible filtration media, which may affect performance variability. For example, zeolite filters did not remove TSS from synthetic stormwater, however, Xsorb brand filters or AbTech Catch Basin Inserts removed almost all TSS under the same conditions [4,6,7].

Maintenance and Effectiveness

Maintenance issues commonly affecting GI performance are clogging [118] (when particulates reduce flow rate) and ineffective maintenance regimes [119] (e.g., inadequate catchbasin cleaning, street sweeping of permeable pavement). Clogging occurs when small particles fill voids in infiltration or filtration media and reduce water flow rates [118]. Sedimentation can completely impede infiltration, causing early saturation, release of water, and possibly pollutant leaching into groundwater [120]. Clogging can weaken functioning of GI as it was demonstrated for biofilters in Australia, where 40% of the GIs studied had limited hydraulic conductivity after 8 years of service [121]. Poorly maintained infiltration basins (>20 years old) tend to have both lower infiltration rates and higher pollutant concentrations in sediments compared with equivalent newer infiltration basins [122].

Within a permeable pavement installation, localized clogging may occur in heavy traffic areas and places where snow is piled in winter [123]. Researchers have also postulated that without regular maintenance to alleviate clogging, permeable pavement is unlikely to provide a water quality benefit [124]. However, Pratt et al. [67] reported that concrete block pavers in a nine-year old parking lot continued to have acceptable infiltration rates despite an absence of maintenance.

Another important maintenance concern relates to filters. Pollution removal can be maximized by customizing the maintenance regime for the surface area of the filter, the amount of impervious surface in the catchment area, and precipitation regimes. Filters with small surface area compared to the percent of impervious area in their catchments will clog rapidly and require frequent maintenance [28], highlighting the importance of proper scaling size of GI. Modeling techniques can reliably predict sediment trapping in some types of filters and may be useful in determining maintenance schedules and maximizing performance [125].

Climate

Climatic regions may affect GI effectiveness as temperature influences the biological and physical properties for pollution removal and infiltration [34,126]. Seasonal variability in runoff infiltration, due to lower hydraulic conductivity at lower temperatures, has been demonstrated in climates with frequent winter freezes [127]. Winter and summer TSS removal efficiency were similar for filtration, bioinfiltration, and retention infrastructure but winter TSS removal efficiency declined for stone swales and hydrodynamic separators [34]. Wet detention ponds show decreased removal efficiency for lead, zinc, and TSS in general during winter, but no change in cadmium and copper removal [74]. This implies that the established relationship between TSS and heavy metal pollution [81,128] may not be as strong in cold weather, at least for some metals. Therefore, practitioners should consider monitoring more pollutants in cold
weather, particularly in areas where heavy metal leaching is an identified issue.

Climatic regions can introduce biases in GI datasets, as most published studies are located in the mid-Atlantic region of the United States and fewer are from other climates. Our study reveals that most GI datasets originate from humid continental climate, characterized by wet, freezing winters. Although many biofiltration systems were pioneered in this region, there is a surprising paucity of data on wintertime performance of their GIs [126,129]. Future additions to the BMP database or published literature may shed some light on climate influence on GI performance.

**COMPARISON WITH THE INTERNATIONAL BMP DATABASE**

The International Best Management Practices (BMP) Database has collected data on stormwater runoff since 1996 and represents one of the most accessible and comprehensive information sources for practitioners about design and implementation of GI [7]. Most data in the BMP database have not been published in peer review journals and therefore did not meet the review criteria of our analyses. However, our study offers an opportunity to test the rigor and usefulness of the data rich BMP database, by comparing peer-reviewed findings with those in the BMP. The BMP database contains more than 180,000 water quality measurements and more than 11,500 runoff volume and peak flow measurements from at least 264 sites throughout the USA and two international sites (Canada and Sweden), located in 4 distinct precipitation patterns within North America (Mediterranean -California-, subtropical -Florida-, temperate –Virginia, Canada-, and semiarid –Texas-).

To quantitatively compare our results to results from the BMP database, we calculated 95% confidence intervals for effectiveness from the peer-reviewed dataset (Table 3) for detention (TSS), biofiltration (TN) and constructed wetlands (TSS and TN) and compared it with similar data from the BMP database reported by Geosyntec Consultants and Wright Water Engineers [130]. Our results from the peer-reviewed literature showed that all categories of GI, with the potential exception of constructed wetlands, significantly reduced both TSS and TN. In contrast, the BMP database showed that only detention ponds (TSS n=43; TN n=12) and biofilters (TSS n=56; TN n=46) significantly lowered concentrations of both TN and TSS (Table 3). In addition, media filters (a type of filtration n=33) were effective at reducing TSS. Channel-type constructed wetlands (n=3) showed a significant reduction in TN removal, with media filters (n=19) increasing TN levels. Calculated percent removal efficiencies from average influent and effluent concentrations of TSS and TN from the BMP database (Table 3) were comparable to the results reported in this study (It should be noted that our peer-reviewed effectiveness data and the BMP database data are not statistically independent because 22 BMP database sites, as of end of 2010, are also part of the peer-reviewed dataset) [130].

Together, the peer-reviewed studies that we evaluated and those in the BMP Database account for a very small proportion of all GI that is currently in use [7]. For instance, Liu and Wang [131] estimated that there are 150 detention and retention basins within a 178 km² watershed in Houston, TX, but only 80 basins from the entire USA are reported in the BMP database (and only 13 of these had sufficient TN data for this analysis [130]. This example illustrates the gap between the infrastructure in place and the infrastructure for which performance data are available and underscores the need for more widespread monitoring, publication, data standardization and data sharing. The 2009 Urban Stormwater BMP Performance Monitoring Manual [132] (hereafter referred to as the Monitoring Manual) lays out the requirements for data reporting in the BMP database, including location, watershed area, design storm, surface area and storage or treatment volume, and date put into service, but many of the sites do not include all these data. Data on catchment area of the stormwater infrastructure, the amount of impervious surface within the catchment area, maintenance type and frequency (both intended and actual), whether there is clogging, and climatic zone would further enhance the usefulness of the BMP database for cross-site comparisons. Standardized data reporting of these potential sources of variability would allow for robust analyses of the impact of each factor on infrastructure performance (see variability section above). We strongly encourage practitioners and researchers to submit their data to the BMP database in the required format laid out in the Monitoring Manual.

**POLICY IMPLICATIONS**

Given the variability in performance seen in this study, we recommend two types of measures accompanying any GI regulation. The first is the implementation of a systematic monitoring and reporting program requiring submission of standardized and complete data to the BMP database (or equivalent). Therefore, monitoring programs should include standardized protocols, quality controls and minimum data requirements to assess the effectiveness of a given infrastructure or cluster over time. GI frequently cost 5-30% less to construct and about 25% less to operate over its life cycle than traditional infrastructure and may be more cost effective than conventional practices [6,133, 134]. In addition, most GI types allow for more flexibility in adapting to changes in conditions and/or knowledge, whereas once gray infrastructure is built, it becomes more costly to reverse or modify it [9]. Initially, incentives may need to be provided to educate and recruit stakeholders for participation in monitoring.

Our analysis can provide guidance as to the selection of specific GI types for general pollutant removal, but site-specific reduction targets should override these recommendations. For example, if removal of specific heavy metals is the goal, TSS removal performance should not be used as an indicator to select a particular type of GI for a site. While rich site-specific information is currently available for individual sites, the range

---

Table 3: Percent reduction in TSS and TN concentrations by GI, derived from BMP database results. The symbol * denotes that reduction percentage falls within 95% CI from the peer-reviewed analyses from this study.

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>TSS Percent Reduction</th>
<th>TN Percent Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detention</td>
<td>58.9 *</td>
<td>-25.3</td>
</tr>
<tr>
<td>Wetland</td>
<td>52.9 *</td>
<td>45.8 *</td>
</tr>
<tr>
<td>Filtration</td>
<td>37.9</td>
<td>23.6</td>
</tr>
<tr>
<td>Biofiltration</td>
<td>54.1</td>
<td>17.0 *</td>
</tr>
</tbody>
</table>

of site types represented in these publications is very limited. Therefore extrapolation of available data (peer-reviewed or BMP) to non-represented sites should be done with caution. Any regulation of GI adoption for stormwater management will need to build in flexibility as more information is collected and the knowledge base is developed further to update the regulation.

ECOLOGICAL IMPLICATIONS

In urban areas, where space is limited, land cover often needs to serve multiple functions [135]. Fortunately, green infrastructure is an avenue to increase other ecosystem services in urban areas. With the right mix of plants, green roofs can provide significant energy and monetary savings, improve urban climates, and reduce greenhouse gas emissions [19,136-140]. Vegetated swales and constructed wetlands may act as long-term storage for carbon [141]. A number of green infrastructures can also be designed to provide wildlife habitat. Constructed wetlands, swales, and detention or retention areas may create suitable habitat for wetland species [142-145], especially if they are designed in a way to minimize pollutants. Similarly, green roofs can host a diversity of arthropod species [146-148] and may increase connectivity among other green spaces [149]. The USDA Natural Resource Conservation Service has developed specific management recommendations that meet water quality goals while providing beneficial habitat for target wildlife species (USDA-NRCS Fish & Wildlife Habitat Management Leaflets http://www.whmi.nrcs.usda.gov/technical/leaflet.htm; [6]). Creation of heterogeneous habitat in urban areas is a potential major benefit of GI, enhancing additional ecosystem services provided by green spaces.

SUMMARY

This study shows that the performance of GI is at the very least comparable to detention approaches to stormwater management (Figure 1), but with potential added benefits of water quality improvement, habitat creation, aesthetics and property values [150]. Green infrastructure decreases pollutant concentrations in effluent and reduces runoff volume and peak flow, but no infrastructure has clear performance superiority. Together, reduced concentrations and reduced volume result in lower total pollutant loads. Variability is high for all types of stormwater infrastructure and has diverse origins, including design and scaling variations, maintenance regimes, and climatic diversity. To maximize performance, GI should be designed and maintained with careful attention paid to local conditions and climate.

There is a great need for consistent reporting of GI performance. Most studies reported in the literature have not been included in the BMP database and, among sites included in the database, many essential pieces of data regarding scale and design specifications are missing. The consistent inclusion of additional data would allow analysis of sources of performance variability, including maintenance type and frequency, climate, and drainage area, to be site specific. Regulators and practitioners should take into account that GI provides benefits beyond those of traditional stormwater management, is more cost effective, and has the advantage of being more flexible than traditional practices. The ecological benefits of GI in urban areas (habitat, biodiversity, resilience, carbon sequestration, etc) need to be better documented in future studies.

ACKNOWLEDGMENTS

Funding for this project came primarily from the Illinois EPA and the American Recovery and Reinvestment Act of 2009. Authors also thank the National Science Foundation for additional support. We received valuable feedback from Hala Ahmed, Megan Elberts, Marty Jaffe, Brian Miller, Jay Rients, Hal Sprague, Steve Wise, Amy Belaire, and Douglas Johnston.

REFERENCES


54. Tapia Silva FO, Wehrmann A, Henze HJ. Model N. Ability of plant-based...


131. Liu H, Wang L. Mapping detention basins and deriving their spatial


Cite this article