Comparing Vegetation and Substrate Performances on Nutrient Removal and Biomass Establishment Using a Natural Swimming Pool Experiment

Margaret C. Hoffman[1], Robert Berghage[1], and Hong Wu[2]

Keywords. blue flag iris, *Iris versicolor*, lizard's tail, mass balance analysis, nitrogen, phosphorus, *Saururus cernuus*

Abstract. Natural swimming pools (NSPs) rely on the interaction of bog vegetation, bacteria, and substrate to maintain water quality. Nitrogen (N) and phosphorus (P) levels in NSPs are critical because of their involvement in eutrophication. We conducted a 15-week greenhouse study to address the significant literature gap regarding nutrient removal capabilities of substrates and vegetation in the low-nutrient environment of NSPs. We used mass balance analyses to compare the performances of four substrates [river gravel (control), recycled glass, expanded clay, expanded shale] and two plant species [blue flag iris (*Iris versicolor*) and lizard's tail (*Saururus cernuus*)] under two flow conditions: free water surface and subsurface flow. At the end of the experiment, except for the recycled glass group, all other substrate groups reduced water nitrate (NO_3) levels to less than 30 $mg{\cdot}L^{-1}$, the standard of the 2011 Forschungsgesellschaft Landschaftsentwicklung Landschaftsbau (FLL) guidelines. However, only the expanded clay group closely approached the P standard ($\leq$0.01 $mg{\cdot}L^{-1}$). Expanded clay and expanded shale demonstrated potential as substrates for NSP bogs. The final aboveground biomass dry weight was strongly negatively correlated with the final NO_3 and P water concentrations. However, direct plant uptake proved insufficient to remove all nutrient inputs, especially for P. Except for the recycled glass group (34%), a significant portion of N (79%–90%) from total N added was removed by aboveground biomass. However, P uptake by biomass was substantially lower (18%–37%). With acceptable vigor and biomass accumulation, blue flag iris may be a suitable species for vegetated NSPs, whereas lizard's tail is not because of uncertain establishment. Compared with controlling N, managing P for FLL standards in NSPs will be more challenging. Our work begins to fill the essential gap in the NSP literature regarding nutrient removal capabilities of substrates and vegetation. Future work should continue to study alternative substrates and plant species for P removal, particularly in field conditions and over longer periods.

Although widely applied in Europe, chemical-free, natural swimming pools (NSPs) are a new technology recently introduced to North America in the 2000s. Based on constructed wetland (CW) technology, NSPs use a designed hydraulic system and an impermeable liner to separate the system from the surrounding natural hydrologic cycle. Using plants and substrate to replicate natural wetland environments, NSPs encourage the establishment of zooplankton, plants, and beneficial bacteria to filter excess nutrients and harmful bacteria from aquatic systems (Littlewood 2016). Various sizes of NSPs have been installed in both public and private settings in the United States (Fig. 1). These systems offer a variety of ecosystem services, such as eliminating ecologically deleterious chemicals, supporting a diverse vegetative community, supplying wildlife habitat, providing four-season landscape aesthetics, and fostering community stewardship (Littlewood 2001).

To date, NSP-related literature remains very limited. Primarily from Europe, existing studies have explored aspects such as water quality and its public health risks (Bruns and Peppler 2018; Petterson et al. 2021) and testing guidelines (Schets et al. 2020), the effect of the hydraulic loading rate and vegetation on phytoremediation [e.g., elimination of nitrate (NO_3), ammonium (NH_4), biological oxygen demand, and chemical oxygen demand)] (Guardia-Puebla et al. 2019), effectiveness of phosphorus-reactive materials (Bus and Karczmarczyk 2015), microbial population characterization (Casanovas-Massana and Blanch 2013), overall condition assessments (Poloprutská et al. 2021), and a living wall as a potential filter (Farb 2020).

Of particular concern for NSPs are nitrogen (N) and phosphorus (P) levels because of their association with eutrophication and subsequent high algae levels. Specifically, P has a greater impact than N on algal growth, causing a rapid increase in eutrophication (Correll 1999). Total phosphorus (TP) levels in freshwater as low as 0.02 $mg{\cdot}L^{-1}$ can be implicated in nuisance algal blooms (Daniel et al. 1998). In the United States, the Environmental Protection Agency (USEPA) mandates that NO_3-N levels in natural freshwater bodies should be less than 10 $mg{\cdot}L^{-1}$, and that TP

Received for publication 4 May 2023. Accepted for publication 28 Aug 2023.

Published online 16 Oct 2023.

[1]Department of Plant Science, Tyson Building, The Pennsylvania State University, University Park, PA 16802, USA

[2]Department of Landscape Architecture, Stuckeman Family Building, The Pennsylvania State University, University Park, PA 16802, USA

M.C.H. is the corresponding author. E-mail: mch7@psu.edu.

https://doi.org/10.21273/HORTTECH05192-23

Units

To convert U.S. to SI, multiply by	U.S. unit	SI unit	To convert SI to U.S., multiply by
29,574	fl oz	μL	3.3814×10^{-5}
29.5735	fl oz	mL	0.0338
0.3048	ft	m	3.2808
3.7854	gal	L	0.2642
2.54	inch(es)	cm	0.3937
25.4	inch(es)	mm	0.0394
28.3495	oz	g	0.0353
7.0053	oz/acre	$mg{\cdot}m^{-2}$	0.1427
305.1517	oz/ft^2	$g{\cdot}m^{-2}$	0.0033
1	ppm	$mg{\cdot}L^{-1}$	1
0.9464	qt	L	1.0567
(°F − 32) ÷ 1.8	°F	°C	(°C × 1.8) + 32

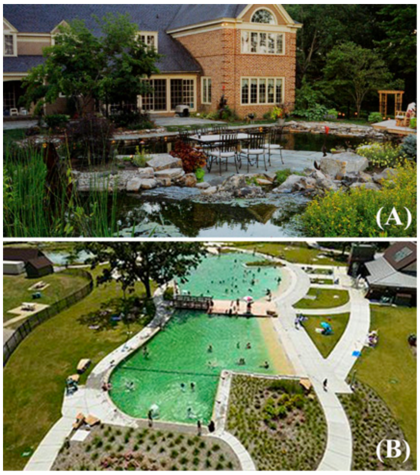

Fig. 1. (A) A private natural swimming pool (NSP) in Pennsylvania and (B) the public Webber NSP in Minneapolis, MN, USA (image courtesy of Webber Pool).

levels should be less than 0.1 mg·L^{-1} (USEPA 1988). In Europe, NSP operation standards developed by Forschungsgesellschaft Landschaftsentwicklung Landschaftsbau (FLL) specify that the levels of N and P should be 30 mg·L^{-1} NO_3 and 0.01 mg·L^{-1} P, respectively (FLL 2011).

Despite the importance of studying the nutrient removal capabilities of NSP biofilters, minimal literature exists regarding this topic. Similar systems, such as stormwater bioretention facilities (Muerdter et al. 2018; Read et al. 2008) and low-nutrient-load CWs, such as those for tertiary wastewater (DeBusk et al. 1995), nursery runoff (Berghage et al. 1999; Huett et al. 2005; Polomski et al. 2007; Taylor et al. 2006), or fish farm runoff treatment (Naylor et al. 2003), may offer insights into the types of vegetation or substrates that could potentially perform well in NSPs. However, NSP-specific studies are required to verify those assumptions. Additionally, it is important to note that NSPs differ from other systems because they are closed, low-nutrient-load systems (Table 1) that are sealed to prevent infiltration and exchange with the surrounding hydrology. Moreover, they are recreational and aesthetic landscape amenities; therefore, the biofilter vegetation should fulfill multiple purposes such as efficient nutrient reduction, ornamental enhancement, supporting biodiversity, and fostering stewardship.

Based on previous CW studies, choosing an effective substrate is critical to ensuring the nutrient removal capability of NSPs. Constructed wetlands are often efficient at removing N, but P can be problematic (Vymazal 2007). Achieving a P concentration target <0.01 mg·L^{-1} (FLL 2011) in NSPs can be especially challenging. Previous studies of alternative media showed that P removal improves when clay-based materials, such as light-expanded clay aggregates (Brix et al. 2001; Mlih et al. 2020) and calcined clay (White et al. 2011), materials high in aluminum, iron, or calcium (Brix et al. 2001; Hsieh and Davis 2005), and expanded shales (Vohla et al. 2011) are incorporated. For example, Glaister et al. (2014) demonstrated excellent P removal capabilities of stormwater biofilter columns with a saturated zone and Skye sand found in Australia with a coating comprising high iron and aluminum oxide levels. Similarly, wollastonite, a naturally occurring mineral that is also high in iron and aluminum oxide, showed good P removal capabilities in swimming and fishpond systems (82%–96% removal) (Fondu et al. 2015) and vertical upflow columns of secondary wastewater (80%–96% removal) (Brooks et al. 2000). However, even with very high P removal levels, the effluent P in most studies exceeded the FLL-suggested levels of 0.01 mg·L^{-1} (e.g., 0.03 mg·L^{-1} with fly ash-sand bioretention media blend) (Vogel et al. 2021) and 0.015 mg·L^{-1} with sandy loam amended with 15% air-dried water treatment residual (Qiu et al. 2019). In a few cases, sand and compost bioretention media mixed with activated alumina achieved 0.01 mg·L^{-1} (Ali and Pickering 2023).

In addition to the P removal capability, substrate selection for NSPs should also consider plant survivability, hydraulic conductivity, and saturation potential. Not all media can support plant growth. Some may increase the pH levels above those that are able to sustain a healthy aquatic system (Naylor et al. 2003). For example, electric arc furnace slag high in calcium, aluminum, and iron (Cameron et al. 2003) and a silica-calcite material called opoka (Bus and Karczmarczyk 2015) could increase the pH levels to those detrimental to vegetation (e.g., >11) despite excellent (78% to almost 100%) P removal performance. Most natural wetlands operate with a pH between 7 and 8, whereas bacteria needed for nitrification and denitrification prefer a pH between 6.5 and 7.5 (Kadlec and Wallace 2009). Low hydraulic conductivity caused by clogging, which is often related to the fine grain size of P-adsorbing materials, can lead to surface channeling, insufficient contact between water and substrate, and, therefore, low nutrient removal efficiency (Yan et al. 2018).

Table 1. Potential nitrogen (N) and phosphorus (P) deposition sources and loads for natural swimming pools.

Deposition source	N load[i]	Reference	P load[i]	Reference
Atmospheric deposition	100–1900 mg·m^{-2} per year	Simkin et al. 2016	2–30 μL per year	Koelliker et al. 2004
Pollen	0.12 mg·L^{-1}	Nicholls and Cox 1978	15 mg·m^{-2} per year	Nicholls and Cox 1978
Urban surface runoff	2.5 mg·L^{-1}	Brezonik and Stadelmann 2002	0.41 mg·L^{-1}	Brezonik and Stadelmann 2002
Fill water	2.2 mg·L^{-1}	Measurements by the authors	0.26–0.77 mg·L^{-1}	Measurements by the authors
User/sunscreen	Unknown		0.46 mg·L^{-1}	BioNova® Natural Pools Inc. (Chester, NJ, USA) unpublished
Plant litter	Unknown		0.97 ppm per year	BioNova® Natural Pools Inc. unpublished

[i] 1 mg·m^{-2} = 0.1427 oz/acre; 1 μL = 3.3814 × 10^{-5} fluid oz; 1 mg·L^{-1} = 1 ppm.

Finally, as substrates approach saturation, nutrient removal capabilities may decrease over time (Arias and Brix 2005; Davis et al. 2006; White et al. 2011). Therefore, ideal NSP substrates should have high hydraulic conductivity, high P adsorption abilities, and long-term P saturation potential, should maintain adequate plant growth and a reasonable pH, and should be readily available and cost-effective.

This study addressed the literature gaps by conducting a vegetated NSP greenhouse study to assess the nutrient removal performance of different plant species and substrates under contrasting hydrological conditions. We specifically explored the following three research questions: 1) How do different substrates support biomass establishment and growth and recover N and P? 2) How do different plant species perform to recover N and P? and 3) How well does the overall system maintain pool water quality below the FLL thresholds?

Materials and methods

A 15-week greenhouse experiment was conducted at The Pennsylvania State University, Department of Plant Science greenhouse complex, University Park, PA, USA (lat. 40°47′53.5704″N, long. 77°51′35.6724″W), from 2 Feb to 5 Jun 2012. We adopted the general approach of a mass balance analysis widely used for CW (Borin and Salvato 2012; Breen 1990; Johengen and LaRock 1993) and bioretention research (Rycewicz-Borecki et al. 2017), which estimates system nutrient inputs, removal, and storage through various compartments.

Hydrological conditions, substrates, and vegetation. We chose two contrasting hydrological conditions based on the two types of CWs most often used for nutrient remediation. The free water surface (FWS) CWs are characterized by a water level of varying depths over the wetland substrate (Brix 2003). Traditionally, FWS substrate consists of coarse river gravel, providing good hydraulic conductivity and some physical support for emergent plants (Kadlec and Wallace 2009). In contrast, in subsurface flow (SSF) CWs, the water level is below the substrate surface, and emergent plants are established in a shallow mulch or soil layer on top of the substrate. Nutrient removal capabilities are similar for both, but they differ in surface area requirements, aesthetics, control of algae and other nuisance organisms, and accompanying ancillary values (Kadlec 2009; Knight et al. 2001; Tanner 1996).

Four substrates (Table 2) were selected, including local river gravel, recycled glass (Growstones; Earthstone Corp., Santa Fe, NM, USA), expanded clay, and expanded shale (Norlite LLC, Cohoes, NY, USA). Local river gravel (size range 1/2 to 1 inch) was the control because of its broad application in NSPs, CWs, and bioretention systems. The other three have been commonly applied in green roofs, bioswales, aquaculture, and CWs, and they show the potential for use in NSPs because of their nutrient removal abilities, support for biomass establishment and growth, cost-effectiveness, and availability (Hill et al. 2000; Wang et al. 2017). More specifically, the high surface area to volume ratio, calcium carbonate, and extremely lightweight nature of recycled glass (size range 1/2 to 1 inch) makes it an attractive candidate for P removal. Expanded clay (size range 1/4 to 1 inch) has a chemical content that ensures P adsorption. Expanded shale is a lightweight coarse aggregate of fired shale, clay, and slate (size range 3/8 to 3/4 inch).

We chose to test obligate wetland plants native to the United States because their establishment and nutrient removal capabilities have not yet been assessed in NSPs despite being studied in bioretention systems (Chen et al. 2009; Lenhart et al. 2012; Maschinski et al. 1994; Moore and Kröger 2011) and CWs (DeBusk et al. 1995; Polomski et al. 2007; Read et al. 2008). Initially, three plant species were selected: blue flag iris (*Iris versicolor*), lizard's tail (*Saururus cernuus*), and pickerelweed (*Pontederia cordata*). However, only the former two survived. Blue flag iris is commonly found on lakeshores and in wetlands, and it is often found growing in shallow water. It can grow 2 to 3 ft in height and has ornamental blue flowers. Lizard's tail is often found along river shores and the margins of marshes and streams. It can grow 3 to 4 ft in height, has fragrant, white inflorescences, and frequently produces high-standing biomass. Both plants are recommended for CWs treating stormwater and agricultural wastewater, produce substantial biomass, and are attractive and commercially available, thus meeting the criteria for an NSP biofilter.

Experiment set-up. To mimic the ecological processes of an NSP and clearly separate the nutrient removal compartments, the experiment unit was designed to be a closed system with one bog unit (with substrate and vegetation) and one pool unit (Fig. 2). Water continually recirculated between the two units.

The experiment consisted of 32 units, with four replicates for each of the eight treatments (4 substrates × 2 hydraulic conditions) planted with the same vegetation. Through a completely randomized design, each substrate and hydraulic condition combination was assigned to four units. All units were randomly assigned their location on two benches. The greenhouse temperature was kept at 82 °F during the day and 76 °F at night. Artificial lights controlled by a timer were on from 5:00 PM to 10:00 PM daily to augment natural lighting.

Each experiment unit was built with two durable polyethylene containers [18 qt for the bog, 18 1/8 inches × 12 1/4 inches × 7 inches; 18 gal for the pool, 24 inches × 18 1/2 inches × 15 3/4 inches; Rubbermaid; Newell Brands, Atlanta, GA, USA). Bogs were filled with the same depth of respective substrates and elevated 1 ft above the pool units. Each set was initially filled with 12 gal of deionized water. Water was circulated at 6 gal per 12 h by a pool pump connected to a 5/8-inch-diameter flexible vinyl tube in the bog, resulting in a total turnover once every 24 h. To create the two hydraulic conditions, the hole in the bog container allowing water to flow to the pool was drilled at different heights. The FWS units had a consistent water level of 2 inches above the substrate surface, whereas the SSF drainage hole was below the substrate surface under 2 inches of 3/8 to 1/2-inch pea gravel. Water levels were checked weekly and kept consistent for each unit. Deionized water was added as necessary, with the volume and timing of each addition recorded. Each bog container was fertilized monthly with 6 g of 20N–4.4P–16.6K water-soluble fertilizer (JR Peters, Inc., Allentown, PA, USA) four times on 13 Feb, 2 Mar, 20 Apr, and 23 May 2012.

Table 2. Composition of the four substrates used for a greenhouse experiment to study substrate performance in supporting biomass establishment and growth and nutrient recovery.

Group	Substrate	Composition	Proportion (%)
1	River gravel (control)	Neutral	
2	Recycled glass	Recycled glass	95–99
		Calcium carbonate ($CaCO_3$)	0.5–5
3	Expanded clay	Fired shale	
		Iron oxide (Fe_2O_3)	9.6
		Alumina (Al_2O_3)	19.4
		Calcium oxide (CaO)	3.4
		Magnesium oxide (MgO)	5.6
		Silica (SiO_2)	57.6
4	Expanded shale	Fired shale, clay, slate silica (SiO_2)	64.2
		Alumina (Al_2O_3)	20.2
		Iron oxide (Fe_2O_3)	4.9
		Titanium oxide (TiO_2)	0.7
		Calcium oxide (CaO)	2.0
		Magnesium oxide (MgO)	3.6
		Alkalies	3.2
		Sulfur trioxide (SO_3)	0.7

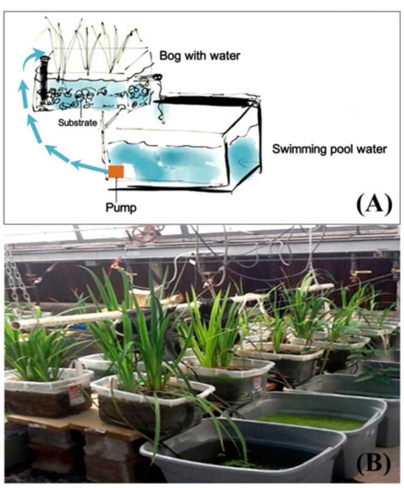

Fig. 2. (A) Conceptual diagram and (B) greenhouse picture of the experiment units. The unit was designed as a closed system with one bog and one pool unit, with water continually recirculated between the two units by a pump in the pool. A total of 32 units were included. Each unit was built with two durable polyethylene containers (18 qt for the bog and 18 gal for the pool) connected via a 5/8-inch flexible vinyl tube. Bogs were filled with the same depth of respective substrates and elevated 1 ft above the pools. Two hydraulic conditions were created by setting the hole in the bog container at different heights. Each set was initially filled with 12 gal of deionized water and planted with two blue flag irises and one lizard's tail; 1 qt = 0.9464 L; 1 gal = 3.7854 L; 1 inch = 2.54 cm; 1 ft = 0.3048 m.

All plants were obtained from Maryland Aquatics Nursery (Jarrettsville, MD, USA) as plugs on 29 Sep 2011. Before the experiment, they were potted and placed in water-filled tubs and fertilized monthly with 15N–6.5P–12.5K water-soluble fertilizer (JR Peters, Inc.). All plants were planted in the units on 9 Nov 2011, approximately 3 months before the experiment start date, to allow plants time to acclimatize. All units initially received one plant each of lizard's tail, pickerelweed, and blue flag iris. However, as previously mentioned, the pickerelweed failed in all units before the experiment started, likely because of difficulty breaking dormancy; therefore, each was replaced on 2 Feb 2012, by one backup blue flag iris that had grown in the same environment since the beginning of the study, resulting in a final mix of two blue flag iris and one lizard's tail per unit. Additionally, during the experiment, aphids became a problem and were controlled with horticultural soap, which may have caused a spike in pool water nutrients in late April. Iron chelate was applied to address chlorosis on 11 and 23 Apr 2012.

To determine plant baseline weights, percent N, and percent P, dried aboveground plant tissues of representative samples for both species were measured at the Penn State Agriculture Analytical Laboratory (University Park, PA, USA). Plant tissue P was determined using the dry ash method reported by Miller (2019), and N was determined using the combustion methods reported by Horneck and Miller (2019).

Water, substrate, and vegetation measurements. Water samples (50 mL/sample) were drawn from all 32 pool units weekly and at the end of the experiment to measure how pool water quality changed over time. Nitrate, pH, and temperature were tested approximately weekly, whereas TP was tested approximately every 2 weeks with additional samples. Specifically, NO_3, pH, and temperature were measured using handheld probes (Forstan LabNavigator; Synaptic Sensors LLC, Fort Collins, CO, USA). Probe calibration occurred before each sampling event, and measurement accuracy was verified using spiked samples. For biweekly TP measurements, additional water samples were filtered through a 45-mm filter and then measured using a spectrophotometer (Lambda 25; PerkinElmer, Waltham, MA, USA) with ultraviolet WINLAB version 2.85 software (PerkinElmer) based on the colorimetric molybdenum blue method (Murphy and Riley 1962). Additionally, relative algae levels were measured with a spectrophotometer (Spectronic 20 Colorimeter; Bausch & Lomb, Rochester, NY, USA), although they were unmeasurable in terms of chlorophyll most of the time.

To measure changes in the substrate total nitrogen (TN) and TP, 50-mL substrate samples were collected from each bog unit at the beginning and end of the experiment and analyzed at the same laboratory where the vegetation was measured. The TN was measured using the combustion method (Bremner 1996), and TP was analyzed using the Mehlich 3 extraction method (Wolf and Beegle 2011).

To measure plant health and growth, plant heights were measured four times on 12 Apr, 18 Apr, 5 May, and 24 May 2012. Growth indices (GIs) were calculated for both species in every unit. The GI of blue flag iris was calculated by height × number of fans; the GI of lizard's tail was calculated by height × number of shoots. The unit total GI was calculated by summing the GIs of the two species.

Finally, to measure the biomass and nutrient content for each treatment, we harvested all plants on 5 Jun 2012, recorded their heights and weights, and dried only the aboveground plant material because of the difficulty separating the roots from substrates. Then, dried tissue samples from each unit were analyzed to determine the percent N and percent P. Plant tissue N was determined using the combustion methods (Horneck and Miller 2019), and P was determined using the dry ash method reported by Miller (2019).

Data analysis. All the data were analyzed using statistical software (IBM SPSS Statistics version 27.0; IBM Corp., Armonk, NY, USA). We compared the final water quality (NO_3 and P concentrations), plant growth (final plant height, GI, and aboveground biomass dry weight), and final percent N and percent P in aboveground biomass by flow condition and substrate. We began with two-way analysis of variance (ANOVA) tests to detect whether a statistically significant interaction existed between flow and substrate for these variables. Because no significant interactions emerged, we used the Mann-Whitney *U* test (nonparametric test for comparing two groups) because of violation of data normality assumptions to detect differences by flow and the Kruskal-Wallis test (nonparametric ANOVA for comparing two or more groups) with post hoc pairwise comparisons for differences by substrate. The significance level to reject the null hypotheses for all tests was $P < 0.05$.

We used a series of formulas to calculate the final percentage removal of NO_3-N and P by the entire system (Eqs. [1] and [2]), N and P uptake efficiency by biomass only (Eqs. [3] and [4]), and N and P uptake efficiency by biomass and substrate (Eqs. [5] and [6]) for each unit (Henderson et al. 2007; Polomski et al. 2007). The same Mann-Whitney *U* and Kruskal-Wallis tests were applied to explore differences in these six variables by flow and substrate.

$$\text{Final percent removal of } NO_3 = \left(1 - \frac{Final\ N(g)\ in\ water}{Total\ added\ N(g)\ in\ fertilizer}\right) \times 100\% \quad [1]$$

$$\text{Final percent removal of P} = \left(1 - \frac{Final\ P(g)\ in\ water}{Total\ added\ P(g)\ in\ fertilizer}\right) \times 100\% \quad [2]$$

$$\text{N uptake efficiency by biomass} = \frac{\text{Delta biomass N(g)}}{Total\ added\ N(g)\ in\ fertilizer} \times 100\% \quad [3]$$

$$\text{P uptake efficiency by biomass} = \frac{\text{Delta biomass P(g)}}{Total\ added\ P(g)\ in\ fertilizer} \times 100\% \quad [4]$$

$$\text{N uptake efficiency by biomass and substrate} = \frac{\text{Delta biomass N(g)} + \text{Delta substrate N(g)}}{Total\ added\ N(g)\ in\ fertilizer} \times 100\% \quad [5]$$

$$\text{P uptake efficiency by biomass and substrate} = \frac{\text{Delta biomass P(g)} + \text{Delta substrate P(g)}}{Total\ added\ P(g)\ in\ fertilizer} \times 100\% \quad [6]$$

The total added N and P (grams) in fertilizer for all units were 4.8 and 2.4 g, respectively. Delta biomass N or P (grams) was determined by the difference in N or P (grams) between harvested (final aboveground biomass dry weight × final percent N or percent P) and initial aboveground biomass dry weight (estimated as 0.66 g N and 0.19 g P for all units). Delta substrate N or P (grams) was determined by the difference in N or P (grams) between the final (substrate initial dry weight × final percent N or percent P) and initial substrate (substrate initial dry weight × initial percent N or percent P).

Next, we calculated the nutrient content of the water, substrate, and vegetation compartments for each unit's mass balance quantification. Insects, algae cycling, sediments, and water loss through splash and evaporation were unmeasured compartments because of their small amounts. Bacterial transformation and root exudates were also unmeasured because of a lack of resources at the time of the experiment.

Finally, bivariate correlations were conducted to explore the relationships between aboveground biomass dry weight and GI, final NO_3 and P water concentrations, water pH and nutrient concentrations, water temperature and nutrient concentrations, and final aboveground biomass and nutrient water concentrations. When assumptions of normality were met, Pearson's correlation tests were performed; otherwise, the nonparametric correlation Spearman's rho test was performed.

Results

Because there were no significant differences in nutrient removal or biomass for flow conditions, results were pooled for each substrate group for eight units per group.

Water quality. Pool water quality fluctuated over time in response to the addition of fertilizers (Fig. 3). Spikes in N and P levels closely followed fertilizer additions, followed by reductions in nutrient levels. The larger P spike in late April may be related to the application of horticultural soap spray for aphid control.

Regarding the final pool water quality, three out of four substrate groups fulfilled the NO_3 standard ($\leq 30\ mg \cdot L^{-1}$) of the FLL, and only one group closely approached the P standard ($\leq 0.01\ mg \cdot L^{-1}$). First, the final mean NO_3 concentration in the pool water ranged from 5.83 (expanded clay) to 53.04 $mg \cdot L^{-1}$ (recycled glass) (Table 3). Percentage N removal by the system ranged from 49.61% (recycled glass) to 94.47% (expanded clay). The recycled glass group was the only one with an average NO_3 level that failed the FLL standard. The Kruskal-Wallis test revealed significant differences in average final NO_3 levels among the four groups [H statistic $H(3) = 17.778$; $P < 0.001$], but only one pairwise difference between the recycled glass and expanded clay group. The final P water concentrations ranged from 0.011 (expanded clay) to 0.074 $mg \cdot L^{-1}$ (recycled glass). Percentage P removal by the system ranged from 99.86% (recycled glass) to 99.98% (expanded clay) (Table 3). The expanded clay group was the only one with an average P level that almost fulfilled the FLL standard. The Kruskal-Wallis test revealed significant differences in average water P levels among the four groups [$H(3) = 19.526$, $P < 0.001$], with three pairwise differences (recycled glass vs. expanded shale, recycled glass vs. expanded clay, and river gravel vs. expanded clay) (Table 3). Additionally, the final NO_3 water concentration was strongly correlated with the final P water concentration [Spearman's correlation coefficient $r_s(30) = 0.772$; $P < 0.001$].

Water temperatures during the experiment varied between 13.7 and 29.1 °C [mean ± *SD* (21.0 ± 2.67 °C)], and the pH varied between 5.7 and 8.9 (7.05 ± 0.51). The bivariate correlation tests showed that water pH was significantly negatively correlated with the NO_3 water concentration [$r_s(477)$ =

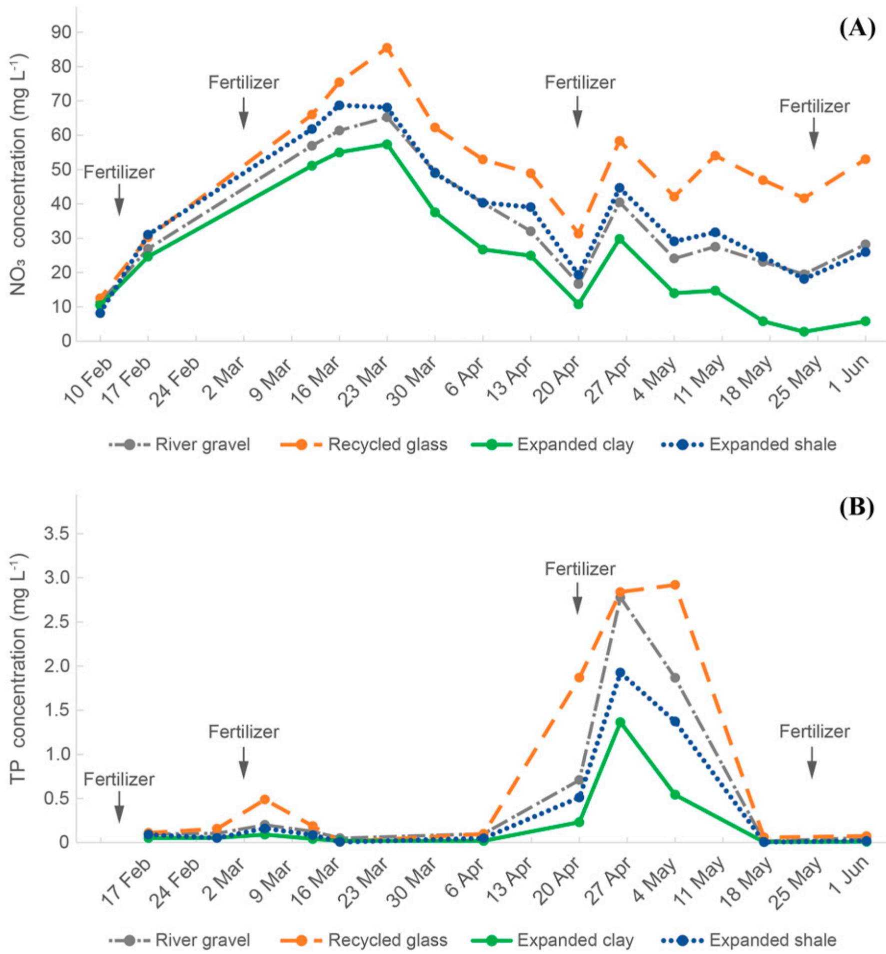

Fig. 3. Mean pool water (A) nitrate (NO_3) and (B) total phosphorus (TP) concentrations of each substrate group from 10 Feb to 1 Jun 2012. Nitrate and TP were tested 15 times and 11 times throughout the experiment, respectively. 1 $mg \cdot L^{-1}$ = 1 ppm.

−0.360, $P < 0.001$], but that the P water concentration was not [$r_s(349) = -0.037$; $P = 0.484$]. Water temperature was not correlated with either NO_3 or P water concentration (r_s = 0.652 and 0.329, respectively), likely because of the temperatures remaining within optimum ranges for biological activity for nutrient removal.

Plant growth and nutrient uptake by biomass

Plant growth and health. Blue flag iris showed vigorous growth, whereas lizard's tail showed slow growth. The lizard's tails did not start growing until late during the experiment, likely because of the lower-nutrient environment of the study. Therefore, the total unit biomass was mostly contributed by blue flag iris. The final average blue flag iris height ranged from 55.6 (recycled glass) to 97.2 cm (expanded shale), whereas lizard's tail height ranged from 16.2 (expanded shale) to 32.5 cm (recycled glass or expanded clay) (Table 4). The Kruskal-Wallis tests revealed significant differences among the four substrates in both height and GI for blue flag iris, but not lizard's tail. For blue flag iris, post hoc pairwise comparisons with Bonferroni correction revealed that the recycled glass group produced significantly lower height than the expanded shale and expanded clay groups. However, only one pairwise difference in the GI was detected between the recycled glass and expanded clay group. Similarly, the substrate type exhibited a significant effect on the final total aboveground biomass dry weight [$H(3) = 11.875$; $P = 0.008$]. Post hoc pairwise comparisons with Bonferroni correction revealed that the recycled glass group produced significantly lower biomass than the river gravel and expanded clay groups. Finally, the final aboveground biomass dry weight and unit total GI were correlated [$r_s(30) = 0.580$; $P < 0.001$], indicating that GI is a reasonable means of comparing plant health.

Nutrient uptake by aboveground biomass. The N concentrations in the aboveground biomass ranged from 1.63% to 1.93% (Table 5), with no significant difference across substrates. In contrast, P concentrations ranged from 0.30% to 0.41%

Table 3. Final pool water nitrate (NO_3) and total phosphorous (TP) concentrations on 1 Jun 2012 and percent nitrogen (N) and phosphorous (P) removal by each group. Final NO_3 water concentrations were measured using a handheld probe and final TP concentrations using a spectrophotometer based on the colorimetric molybdenum blue method. Kruskal-Wallis tests with post hoc pairwise comparisons were performed to determine group differences in final NO_3 and TP concentrations.

Group	Substrate	Final NO_3 concn ($mg \cdot L^{-1}$)[i]	Final N removal from water (%)	Final TP concn ($mg \cdot L^{-1}$)	Final P removal from water (%)
		Mean (*SD*)			
1	River gravel (control)	28.24 (14.99) ab[ii]	73.17 (14.24) ab	0.057 (0.025) ab	99.89 (0.05) ab
2	Recycled glass	53.04 (20.89) a	49.61 (19.84) a	0.074 (0.038) a	99.86 (0.07) a
3	Expanded clay	5.83 (8.50) b	94.47 (8.08) b	0.011 (0.013) c	99.98 (0.02) c
4	Expanded shale	26.01 (19.30) ab	75.29 (18.33) ab	0.020 (0.020) bc	99.96 (0.04) bc
	Kruskal-Wallis test statistic (H)	$H(3) = 17.778$, $P < 0.001$		$H(3) = 19.526$, $P < 0.001$	

[i] 1 $mg \cdot L^{-1}$ = 1 ppm.
[ii] Different letters indicate significant differences among substrates discovered by post hoc tests with Bonferroni correction.

Table 4. Final average heights and growth indices (GIs) of blue flag iris and lizard's tail measured on 24 May 2012, and total aboveground biomass dry weight of each group measured on 5 Jun 2012. Kruskal-Wallis tests with post hoc pairwise comparisons were performed to determine group differences in height and GI for both species, as well as the final total aboveground biomass dry weight.

Group	Substrate	Blue flag iris ht (cm)[i]	Blue flag iris GI (cm)[ii]	Lizard's tail ht (cm)	Lizard's tail GI (cm)[ii]	Final aboveground biomass dry wt (g)
		Mean (*SD*)				
1	River gravel (control)	80.3 (13.0) ab[iii]	807.9 (423.9) ab	17.6 (14.0)	17.5 (16.7)	301.4 (118.2) b
2	Recycled glass	55.6 (29.2) a	349.9 (225.4) a	32.5 (22.9)	65.9 (80.5)	142.0 (115.0) a
3	Expanded clay	96.0 (16.7) b	924.4 (389.6) b	32.5 (37.8)	48.0 (78.8)	312.8 (86.0) b
4	Expanded shale	97.2 (10.9) b	824.5 (357.1) ab	16.2 (18.0)	19.1 (29.0)	235.0 (46.2) ab
	Kruskal-Wallis test statistic (H)	H(3) = 15.204, P = 0.002	H(3) = 12.156, P = 0.007	P = 0.528	P = 0.364	H(3) = 11.875, P = 0.008

[i] 1 cm = 0.3937 inch.
[ii] Blue flag iris GI = height × number of fans; lizard's tail GI = height × number of shoots.
[iii] Different letters indicate significant differences among substrates discovered by post hoc tests with Bonferroni correction.

and showed significant differences across substrates. The expanded clay group generated a significantly lower biomass percent P than all the other three groups.

The N uptake efficiency by aboveground biomass ranged from 33.78% (recycled glass) to 90.46% (expanded clay). It showed significant differences by substrate, with the recycled glass group having significantly lower biomass N uptake efficiency than the three other groups (Table 5). The P uptake efficiency by aboveground biomass ranged from 18.05% (recycled glass) to 36.62% (river gravel), with no significant differences by substrate. The bivariate correlation tests showed a strong negative correlation between final aboveground biomass dry weight and final pool nutrient concentrations [$r_s(30) = -0.703$ and $P < 0.001$ for N; $r_s(30) = -0.605$ and $P < 0.001$ for P].

System nutrient uptake. The N uptake efficiency by the system (both aboveground biomass and substrate) ranged from 14.86% (recycled glass) to 91.07% (expanded shale) (Table 6). System P uptake efficiency ranged from −82.04% (recycled glass) to 36.62% (river gravel). Both measures showed significant differences by substrate, with the recycled glass group showing significantly lower N and P uptake efficiency than the other three groups.

Mass balance analysis of P and N. The mass balance calculations showed that the system accounted for 65% (recycled glass) to 116% (expanded shale) of total added N compared with −82% (recycled glass) to 37% (river gravel) of total added P (Fig. 4). The percentage of unaccounted N (−16% to 35%) was much lower than that of unaccounted P (63%–182%). The recycled glass substrate showed substantial decreases of N (19% of added N) and P storage (100% of added P), whereas expanded clay showed a slight decrease of N storage (3% of added N). The river gravel substrate showed no N or P content changes, whereas expanded shale showed a small increase in N storage (12% of added N).

Discussion

Substrate performance. The two substrates of expanded clay and expanded shale demonstrated higher potential than river gravel and recycled glass for NSP bogs. Both groups reduced the final water N and P concentrations to below or close to the FLL values (mean P concentration for the expanded shale group was 0.02 $mg \cdot L^{-1}$ which was close to the 0.01 $mg \cdot L^{-1}$ target) (Table 3). Additionally, expanded shale showed the highest N media storage. Recycled glass, however, is an unsuitable substrate for NSPs. It showed the highest initial N and P contents (Table 6) and negative storage of both nutrients (Fig. 4). Avoiding recycled glass is consistent with the work of Hatt et al. (2009) because selecting media with a low initial

Table 5. Final average nitrogen (N) and phosphorus (P) concentrations in aboveground biomass and uptake efficiency by aboveground biomass of each group. Kruskal-Wallis tests with post hoc pairwise comparisons were performed to determine group differences in N and P concentrations and uptake efficiency in aboveground biomass.

Group	Substrate	N in aboveground biomass (%)	N uptake efficiency by aboveground biomass only (%)	P in aboveground biomass (%)	P uptake efficiency by aboveground biomass only (%)
		Mean (*SD*)			
1	River gravel (control)	1.66 (0.23)	87.67 (36.41) b[i]	0.40 (0.07) b	36.62 (17.14)
2	Recycled glass	1.90 (0.54)	33.78 (22.00) a	0.41 (0.10) b	18.05 (18.47)
3	Expanded clay	1.63 (0.32)	90.46 (27.86) b	0.30 (0.04) a	30.03 (7.33)
4	Expanded shale	1.93 (0.33)	79.18 (16.24) b	0.41 (0.07) b	30.86 (3.13)
	Kruskal-Wallis test statistic (H)	P = 0.282	H(3) = 13.347, P = 0.004	H(3) = 11.560, P = 0.009	P = 0.131

[i] Different letters indicate significant differences among substrates discovered by post hoc tests with Bonferroni correction.

Table 6. Average substrate nitrogen (N) and phosphorus (P) content changes and N and P uptake efficiency by aboveground biomass and substrate of each group. Kruskal-Wallis tests with post hoc pairwise comparisons were performed to determine group differences in N and P content changes in substrate and uptake efficiency by aboveground biomass and substrate.

Group no.	Substrate	Initial N in substrate (g)[i]	N change in substrate (g)[ii]	N uptake efficiency (aboveground biomass and substrate) (%)	Initial P in substrate (g)	P change in substrate (g)[ii]	P uptake efficiency (aboveground biomass and substrate) (%)
			Mean (*SD*)			Mean (*SD*)	
1	River gravel (control)	0	0 (0)	87.67 (36.41) b[iii]	5.46	0 (0)	36.62 (17.14) b
2	Recycled glass	2.64	−0.91 (0.55)	14.86% (24.86) a	8.10	−2.40 (0.88)	−82.04 (39.79) a
3	Expanded clay	1.40	−0.15 (0.34)	87.25% (31.77) b	6.86	0.01 (0.03)	30.57 (6.59) b
4	Expanded shale	1.78	0.57 (2.87)	91.07% (59.99) b	7.24	0.05 (0.05)	32.84 (3.31) b
	Kruskal-Wallis test statistic (H)		$H(3) = 11.368$, $P = 0.01$	$H(3) = 15.344$, $P = 0.002$		$H(3) = 23.376$, $P < 0.001$	$H(3) = 17.628$, $P < 0.001$

[i] 1 g = 0.0353 oz.
[ii] N or P change in substrate was determined by the difference in N or P (grams) between the final and initial substrate.
[iii] Different letters indicate significant differences among substrates discovered by post hoc tests with Bonferroni correction.

nutrient content is essential for effective nutrient removal.

Despite failing the FLL standard for P, the river gravel group showed high system uptake efficiency (ranked second for N and first for P) (Fig. 3), perhaps because of its lowest surface area ratio and least reactive sites for adsorption. During future experiments, using smaller-diameter river gravel (e.g., 16 mm) (Akratos and Tsihrintzis 2007) may help increase P removal efficiency because adsorption is the main P removal mechanism, although the source and chemical characteristics of the river gravel would determine exchange capacity. Moreover, river gravel and expanded shale could possibly be amended with clay-based materials (e.g., calcined clay) to improve P removal (White et al. 2011). More research is needed to determine the effectiveness of these strategies. If proven effective, then adding calcined clay to river gravel and reducing gravel diameter can become a highly cost-effective alternative to expanded clay or expanded shale.

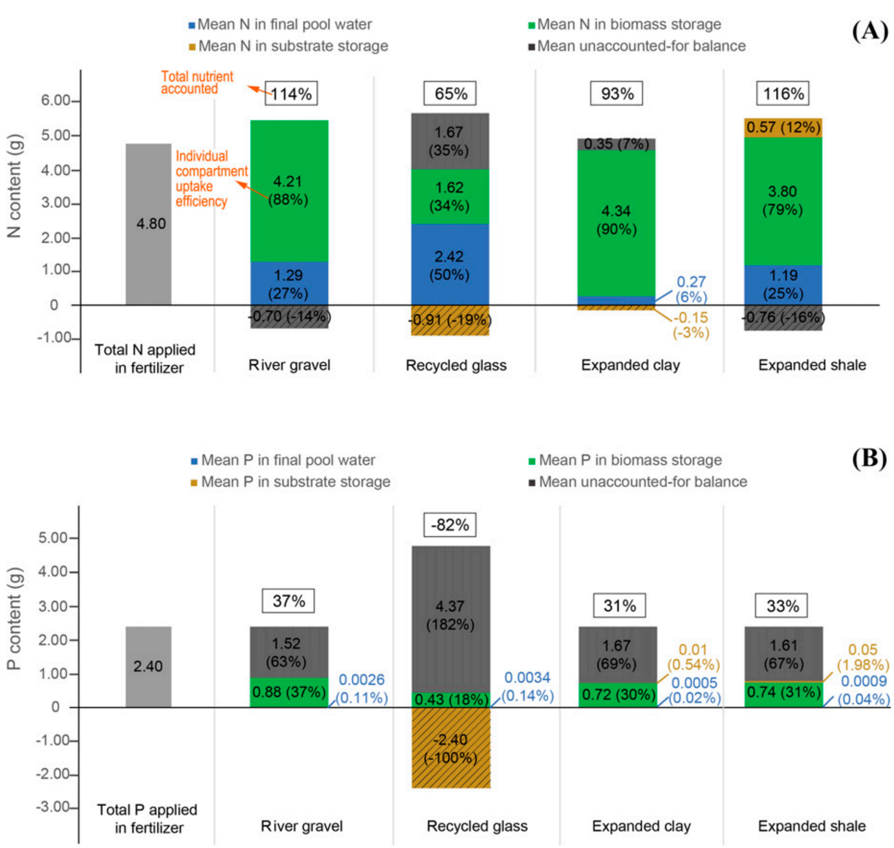

Fig. 4. Mass balance analyses of (A) nitrogen (N) and (B) phosphorus (P) by substrate group. The three system components included were pool water, substrate, and aboveground biomass. 1 g = 0.0353 oz.

Ability of aboveground biomass to recover nutrients. The significant portion of N removed by aboveground biomass from total N added (34%–90%) (Fig. 3), in most cases, was at a level similar to or greater than that of previous studies of similar systems. Except for the recycled glass group (34%), aboveground biomass in the other three groups assimilated 79% to 90% of N, in agreement with the level of >70% reported by Biswal et al. (2022), who reviewed the performance of stormwater bioretention. Regarding other low-load CW studies, DeBusk et al. (1995) reported 16% to 129% N uptake efficiency by biomass from dairy wastewaters, Naylor et al. (2003) reported 21% to 28% from fish farm runoff, and Huett et al. (2005) reported 76% from nursery runoff.

When investigating pathways for N removal, it is important to remember that NO_3-N removal is generally performed through either denitrification or plant uptake rather than sorption. The literature reporting N mass balance amounts stored by plants is divided. Some studies reported that N removal by plants is a minor fraction of TN removal, with denitrification representing the largest removal process (Maine et al. 2007). However, we could expect plant N uptake to be more significant during this study because water was consistently circulating, thereby oxygenating the substrate and reducing optimal denitrification conditions. Because of the low level of introduced N, plants were forced to mine the system for nutrients to support active growth, thus contributing the most to TN removal (Kadlec and Wallace 2009). The high aboveground biomass N uptake efficiencies in the expanded clay (90%), river gravel (88%), and expanded shale (79%) groups were probably reflected by the improved biomass production over any chemical substrate characteristics, as was the poor performance in the recycled glass group (34%), with the lowest biomass production. Finally, based on the estimate of potential N uptake of up to 51 $g \cdot m^{-2}$ per year by plants in treatment wetlands reported by Kadlec and Wallace (2009), biomass in NSPs could potentially remove a large portion of total added N because the N loads were substantially lower than 51 $g \cdot m^{-2}$ per year (Table 1).

Compared with N, the lower P uptake efficiency by aboveground biomass (18%–37%) reflects the significant variance in P removal reported in the bioretention and CW literature. Although biomass is frequently cited as a crucial compartment for N removal, studies regarding the importance of vegetation in P removal are divided. Many studies reported a smaller portion of removed P stored in the standing stock. For instance, regarding CWs, Chung et al. (2008) suggested that planted treatments significantly increased P removal over unplanted treatments; however, regarding the mass balance, substrate was the main removal pathway, with less than 1% of P removed by plant uptake. Similarly, Browning and Greenway (2003) reported 1% to 3% biomass P removal, and Tanner et al. (1995) reported 1.9% to 5.3% biomass P removal. In contrast, other studies indicated that both vegetation uptake and adsorption are important mechanisms of P removal, specifically TP. For example, as reported by Barrett et al. (2013), vegetated columns removed significantly more TP (77%–94%) from synthetic stormwater than unvegetated treatments (58%–80%). Using a greywater biofiltration system, Fowdar et al. (2017) indicated plant storage as the main P storage compartment, with >95% removal efficiency, although the fact that only tall sedge (*Carex appressa*), a plant with a high nutrient removal capability, was used could have skewed the results.

Regarding species, the blue flag iris appears to be a suitable candidate for a vegetated NSP biofilter with acceptable vigor and biomass accumulation. Muerdter et al. (2020) found similarly excellent survivability of blue flag iris in a bioretention container study. Additionally, unlike the yellow flag iris (*Iris pseudacorus*) commonly used for CW research, the blue flag iris is native to the United States and does not have invasive tendencies. Furthermore, another native iris very similar to the blue flag iris, the ‘Full Eclipse’ Louisiana hybrid iris (*Iris* hybrid), was also highly efficient for P uptake (Polomski et al. 2007). In contrast, although some literature showed success with lizard’s tail in higher-nutrient environments (Lenhart et al. 2012), it may be unsuitable because of uncertain establishment and low GI, despite a higher tissue nutrient concentration. Other studies that measured nutrient reduction in vegetated filters also discovered that lizard’s tail had a limited contribution to biofilter performance. For example, Moore and Kröger (2011) reported that lizard’s tail removed the lowest percentage of NO_3-N and ranked third for P removal among five species in stormwater ditches. Finally, the potential unsuitability of lizard’s tail should be verified by a field-scale NSP study during a longer period.

Implications for NSP water quality maintenance. Controlling N is a lesser issue than P in NSP management. The FLL limits are higher and easier to reach for NO_3 than P, and there is substantial evidence indicating that a well-vegetated bog would sufficiently control NO_3 and NH_4 (Biswal et al. 2022; Vymazal and Kröpfelová 2009). The N from fill water or stormwater runoff likely will not cause an issue either. As mandated by the USEPA, nitrate should be less than 10 $mg \cdot L^{-1}$ in drinking water. The National Stormwater Quality Database and other national studies have identified typical event mean concentrations of total Kjeldahl nitrogen (the sum of NH_4-N plus organically bound N) of 1.67 to 1.74 $mg \cdot L^{-1}$ (Yang and Lusk 2018). Circumstances remain in which N uptake can become an issue, for example, when vegetation and biomass are insufficiently established or if the NSP is improperly designed with a high N loading, such as from fertilizers or lawn trimmings.

In contrast, managing P for FLL standards can be a major challenge. Because P levels are not of particular concern for drinking water, significant amounts of P can be introduced by adding tap or well water. Early experiments at the Pennsylvania State University greenhouse measured 0.26 to 0.77 $mg \cdot L^{-1}$ of P in tap water. An earlier field study of three central Pennsylvania NPSs also found tap water and well water P input measurements of 0.15 $mg \cdot L^{-1}$ (Hoffman 2013). Therefore, a rainwater reservoir may be a preferable source for fill water. Additionally, testing for nutrients, especially P, before fill water is added to the system is essential. Other major sources of P must be properly managed. For example, typical stormwater runoff TP input ranges from 0.32 to 0.37 $mg \cdot L^{-1}$ (Yang and Lusk 2018). Therefore, avoiding draining adjacent surfaces, especially high P-loading ones, into the NSP is imperative. Other potential strategies include encouraging pool users to wear P-free sunscreen and shower before entering the pool and reducing pollen-producing plant species in adjacent areas.

Furthermore, proper maintenance of the substrate and vegetation is critical to long-term nutrient uptake. Regarding substrate maintenance, although this study did not measure sediment formation, minor sediment deposition did occur in some units and possibly was the source of some unaccounted P outputs. It seems likely that the bog filter would form sediments as it ages, thus increasing P removal caused by sediment accretion. Periodic removal of sediments would permanently remove adsorbed P, whereas the sediments could be easily treated as soil amendments and added to compost piles or gardens. However, the need to replace substrate because of saturation may be minimal because nutrient loading is much lower in NSPs than in most CWs and stormwater technologies. The media will possibly function beyond the expected lifetime of the other NSP components, thus qualifying biofiltration as a sustainable technology with much lower carbon costs than a CW. Therefore, occasional testing of the media after a long period of use (e.g., approximately 5 years based on the CW literature) (Vohla et al. 2011) is recommended. Adding external P filters (e.g., with expanded clay and alumina P) could also eliminate that issue (Hoffman 2013).

Finally, proper vegetation maintenance by harvesting and appropriate disposal of aboveground biomass is essential to limiting system N and P. Because of the significant amount of N and P sequestered in aboveground biomass, harvesting plant materials should be a critical component of the maintenance plan to minimize N and P accumulation through plant tissue decay (Davis et al. 2006; Muerdter et al. 2018; Rycewicz-Borecki et al. 2017; Vymazal 2020).

Study limitations. Despite the critical findings and management implications, there were several study limitations. First, because of resource constraints, this study was not replicated, which could have compromised the rigor of the findings. Second, nonvegetated units could have been included in the study design, which would have better-separated the substrates' contribution to nutrient removal. Third, more advanced testing techniques, such as the equilibrium isotherm experiments, would have been a more rigorous way to assess the substrates' role in adsorbing P. Fourth, even though biofilms have essential consequences for nutrient transformations (especially N), we were unable to quantify them because of resource constraints. Fifth, although we could confidently say that this research accounts for most of the decreases and losses of N in the units, P cycling was less well-represented by the mass balance model.

In the case of high unaccounted N (35%) with the recycled glass treatment, N may have been bound to some compartment in unrecoverable form or was exported from the system. Possible export mechanisms would be denitrification or biomass loss.

Regarding the overall high percentages of unaccounted P (63%–182%), it appears likely that P was tightly bound to the recycled glass, expanded clay, and expanded shale substrates. The Mehlich 3 extraction method was unable to remove the P for measurement, although the river gravel, with less reactive sites, also showed a large portion of unaccounted P (63%). The unmeasured contribution of sediments accumulated in the units could have contributed to P removal, as did roots and algae uptake. Because water samples for P analyses were filtered through a 45-mm filter, the measurements could not account for the P lost to algal sequestration. Because of the small amounts of sediments and algae, we suspect the unmeasured root biomass could be where much of the unaccounted P was stored. As previously mentioned, because of the extreme difficulty separating plant roots from the substrates (especially the more porous recycled glass), our study only measured plant contributions based on aboveground biomass. Further examinations of plant root structure and exudates may provide insight regarding characteristics of N and P removal and inform planting designs accordingly (Read et al. 2009). Moreover, if possible, future studies should explore other plant-mediated pathways beneficial to nutrient removal, including plant effects on microbial processes, species effects on denitrification by producing carbon in the rhizosphere, and decomposition of plant material for sediment formation (Tanner 2001).

Finally, future studies should be conducted to determine the NSP biofilter performance under field conditions and over a longer period instead of during only one season. Because of concerns that greenhouse research may show artificial inflation of nutrient removal efficiencies caused by increased vegetative densities, field studies can better represent the realities of an actual NSP environment. Moreover, NSPs need to operate efficiently throughout the year, with peak performance during spring and summer, when pool use is high and additional nutrient loading stress exists, such as pollen deposition and spring storm runoff. Fluctuations in seasonal performance and the effects of plant senescence and cold temperatures on nutrient removal capabilities still need to be quantified.

Conclusions

Through a vegetated NSP greenhouse study, we assessed the nutrient removal performance of two plant species, blue flag iris and lizard's tail, and four substrates, river gravel, recycled glass, expanded clay, and expanded shale, under FWS and SSF flow conditions. Except for the recycled glass group, the substrate groups reduced system water to target NO_3 levels of 30 $mg{\cdot}L^{-1}$. However, only the expanded clay group closely approached the P standard ($\leq$0.01 $mg{\cdot}L^{-1}$). The final aboveground biomass dry weight was strongly negatively correlated with final NO_3 and P water concentrations, thus demonstrating the significant role of plant uptake in nutrient removal. Expanded clay and expanded shale are substrates with high potential for NSP bogs because of their ability to support plant growth and remove nutrients, whereas recycled glass should be avoided. Blue flag iris may be a suitable candidate for vegetated NSPs with acceptable vigor and biomass accumulation, whereas lizard's tail is not because of uncertain establishment.

Although controlling N may be less difficult for NSPs, managing P according to FLL standards can be a major challenge. In addition to selecting and properly maintaining suitable vegetation and substrate, testing fill water for nutrients and properly managing other major P sources (e.g., stormwater runoff, sunscreens, and pollen) are essential practices. As one of the first NSP greenhouse studies, our work contributes to the very limited NSP literature by addressing the critical topic of nutrient removal capabilities of biofilter substrates and vegetation. Future work should continue exploring more effective vegetation and substrate choices for P removal, particularly in a field setting and over longer periods.

References cited

Akratos CS, Tsihrintzis VA. 2007. Effect of temperature, HRT, vegetation and porous media on removal efficiency of pilot-scale horizontal subsurface flow constructed wetlands. Ecol Eng. 29:173–191. https://doi.org/10.1016/j.ecoleng.2006.06.013.

Ali MA, Pickering NB. 2023. Systematic evaluation of materials to enhance soluble phosphorus removal using biofiltration or bioswale stormwater management controls. J Sustain Water Built Environ. 9:04022017. https://doi.org/10.1061/JSWBAY.0001004.

Arias CA, Brix H. 2005. Phosphorus removal in constructed wetlands: Can suitable alternative media be identified? Water Sci Technol. 51:267–273. https://doi.org/10.2166/wst.2005.0335.

Barrett ME, Limouzin M, Lawler DF. 2013. Effects of media and plant selection on biofiltration performance. J Environ

Eng. 139:462–470. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000551.

Berghage RD, MacNeal EP, Wheeler EF, Zachritz WH. 1999. "Green" water treatment for the green industries: Opportunities for biofiltration of greenhouse and nursery irrigation water and runoff with constructed wetlands. HortScience. 34:50–54. https://doi-org.ezaccess.libraries.psu.edu/10.21273/HORTSCI.34.1.50.

Biswal BK, Vijayaraghavan K, Adam MG, Lee Tsen-Tieng D, Davis AP, Balasubramanian R. 2022. Biological nitrogen removal from stormwater in bioretention cells: A critical review. Crit Rev Biotechnol. 42:713–735. https://doi.org/10.1080/07388551.2021.1969888.

Borin M, Salvato M. 2012. Effects of five macrophytes on nitrogen remediation and mass balance in wetland mesocosms. Ecol Eng. 46:34–42. https://doi.org/10.1016/j.ecoleng.2012.04.034.

Breen PF. 1990. A mass balance method for assessing the potential of artificial wetlands for wastewater treatment. Water Res. 24:689–697. https://doi.org/10.1016/0043-1354(90)90024-Z.

Bremner JM. 1996 Nitrogen-total, p 1085–1122. In: Sparks DL, Page AL, Helmke PA, Loeppert RH, Soltanpour PN, Tabatabai MA, Johnston CT, Sumner ME (eds). Methods of soil analysis part 3: Chemical methods. SSSA Book Series 5. Soil Science Society of America, Madison, WI, USA. https://doi.org/10.2136/sssabookser5.3.c37.

Brezonik PL, Stadelmann TH. 2002. Analysis and predictive models of stormwater runoff volumes, loads, and pollutant concentrations from watersheds in the Twin Cities metropolitan area, Minnesota, USA. Water Res. 36:1743–1757. https://doi.org/10.1016/S0043-1354(01)00375-X.

Brix H. 2003. Plants used in constructed wetlands and their functions. 1st International Seminar on the use of Aquatic Macrophytes for Wastewater Treatment in Constructed Wetlands. Lisboa, Portugal, p 81–109.

Brix H, Arias CA, del Bubba M. 2001. Media selection for sustainable phosphorus removal in subsurface flow constructed wetlands. Water Sci Technol. 44:47–54. https://doi.org/10.2166/wst.2001.0808.

Brooks AS, Rozenwald MN, Geohring LD, Lion LW, Steenhuis TS. 2000. Phosphorus removal by wollastonite: A constructed wetland substrate. Ecol Eng. 15:121–132. https://doi.org/10.1016/S0925-8574(99)00056-7.

Browning K, Greenway M. 2003. Nutrient removal and plant biomass in a subsurface flow constructed wetland in Brisbane, Australia. Water Sci Technol. 48:183–189. https://doi.org/10.2166/wst.2003.0315.

Bruns S, Peppler C. 2018. Hygienic quality of public natural swimming pools (NSP). Water Supply. 19:365–370. https://doi.org/10.2166/ws.2018.076.

Bus AZ, Karczmarczyk AA. 2015. Kinetic and sorption equilibrium studies on phosphorus removal from natural swimming ponds by selected reactive materials. Fresenius Environ Bull. 24:2736–2741.

Cameron K, Madramootoo C, Crolla A, Kinsley C. 2003. Pollutant removal from municipal sewage lagoon effluents with a free-surface wetland. Water Res. 37:2803–2812. https://doi.org/10.1016/S0043-1354(03)00135-0.

Casanovas-Massana A, Blanch AR. 2013. Characterization of microbial populations associated with natural swimming pools. Int J Hyg Environ Health. 216:132–137. https://doi.org/10.1016/j.ijheh.2012.04.002.

Chen Y, Bracy RP, Owings AD, Merhaut DJ. 2009. Nitrogen and phosphorous removal by ornamental and wetland plants in a greenhouse recirculation research system. HortScience. 44:1704–1711. https://doi.org/10.21273/HORTSCI.44.6.1704.

Chung AKC, Wu Y, Tam NFY, Wong MH. 2008. Nitrogen and phosphate mass balance in a sub-surface flow constructed wetland for treating municipal wastewater. Ecol Eng. 32:81–89. https://doi.org/10.1016/j.ecoleng.2007.09.007.

Correll DL. 1999. Phosphorus: A rate limiting nutrient in surface waters. Poult Sci. 78:674–682. https://doi.org/10.1093/ps/78.5.674.

Daniel TC, Sharpley AN, Lemunyon JL. 1998. Agricultural phosphorus and eutrophication: A symposium overview. J Environ Qual. 27:251–257. https://doi.org/10.2134/jeq1998.00472425002700020002x.

Davis AP, Shokouhian M, Sharma H, Minami C. 2006. Water quality improvement through bioretention media: Nitrogen and phosphorus removal. Water Environ Res. 78:284–293. https://doi.org/10.2175/106143005x94376.

DeBusk TA, Peterson JE, Ramesh Reddy K. 1995. Use of aquatic and terrestrial plants for removing phosphorus from dairy wastewaters. Ecol Eng. 5:371–390. https://doi.org/10.1016/0925-8574(95)00033-X.

Farb A. 2020. A deep dive into natural swimming pool filtration: Living walls as technical wetland filters (MLA Thesis). Utah State Univ, Logan, UT, USA. https://doi.org/10.26076/b46b-1450.

Forschungsgesellschaft Landschaftsentwicklung Landschaftsbau. 2011. Recommendations for planning, construction, servicing and operating of outdoor swimming pools with biological water purification (swimming and bathing ponds). FLL, Bonn, Germany.

Fondu L, De Bo I, Van Hulle SWH. 2015. Phosphate adsorption capacity testing of natural and industrial substrates in view of application in swimming and fish pond water treatment systems. Desalination Water Treat. 54:2461–2467. https://doi.org/10.1080/19443994.2014.899930.

Fowdar HS, Hatt BE, Cresswell T, Harrison JJ, Cook PLM, Deletic A. 2017. Phosphorus fate and dynamics in greywater biofiltration systems. Environ Sci Technol. 51:2280–2287. https://doi.org/10.1021/acs.est.6b04181.

Glaister BJ, Fletcher TD, Cook PLM, Hatt BE. 2014. Co-optimisation of phosphorus and nitrogen removal in stormwater biofilters: The role of filter media, vegetation and saturated zone. Water Sci Technol. 69:1961–1969. https://doi.org/10.2166/wst.2014.117.

Guardia-Puebla Y, Pérez-Quintero F, Rodríguez-Pérez S, Sánchez-Girón V, Llanes-Cedeño E, Rocha-Hoyos J, Peralta-Zurita D. 2019. Effect of hydraulic loading rate and vegetation on phytoremediation with artificial wetlands associated to natural swimming pools. J Water Land Dev. 40:39–51. https://doi.org/10.2478/jwld-2019-0004.

Hatt BE, Fletcher TD, Deletic A. 2009. Hydrologic and pollutant removal performance of stormwater biofiltration systems at the field scale. J Hydrol (Amst). 365:310–321. https://doi.org/10.1016/j.jhydrol.2008.12.001.

Henderson C, Greenway M, Phillips I. 2007. Removal of dissolved nitrogen, phosphorus and carbon from stormwater by biofiltration mesocosms. Water Sci Technol. 55:183–191. https://doi.org/10.2166/wst.2007.108.

Hill CM, Duxbury J, Geohring L, Peck T. 2000. Designing constructed wetlands to remove phosphorus from barnyard runoff: A comparison of four alternative substrates. J Environ Sci Health Part A Tox Hazard Subst Environ Eng. 35:1357–1375. https://doi.org/10.1080/10934520009377040.

Hoffman MC. 2013. Nutrient removal in natural swimming pools: A mass balance analysis (PhD Diss). The Pennsylvania State Univ, University Park, PA, USA.

Horneck DA, Miller RO. 2019. Determination of total nitrogen in plant tissue, p 75–83. In: Kalra YP (ed). Handbook and reference methods for plant analysis. CRC Press, New York, NY, USA.

Hsieh C, Davis AP. 2005. Evaluation and optimization of bioretention media for treatment of urban storm water runoff. J Environ Eng. 131:1521–1531. https://doi.org/10.1061/(ASCE)0733-9372(2005)131:11(1521).

Huett DO, Morris SG, Smith G, Hunt N. 2005. Nitrogen and phosphorus removal from plant nursery runoff in vegetated and unvegetated subsurface flow wetlands. Water Res. 39:3259–3272. https://doi.org/10.1016/j.watres.2005.05.038.

Johengen TH, LaRock PA. 1993. Quantifying nutrient removal processes within a constructed wetland designed to treat urban stormwater runoff. Ecol Eng. 2:347–366. https://doi.org/10.1016/0925-8574(93)90003-X.

Kadlec RH. 2009. Comparison of free water and horizontal subsurface treatment wetlands. Ecol Eng. 35:159–174. https://doi.org/10.1016/j.ecoleng.2008.04.008.

Kadlec RH, Wallace SD. 2009. Treatment wetlands (2nd ed). CRC Press, Boca Raton, FL, USA. https://doi.org/10.1201/9781420012514.

Knight RL, Clarke RA, Bastian RK. 2001. Surface flow (SF) treatment wetlands as a habitat for wildlife and humans. Water Sci Technol. 44:27–37. https://doi.org/10.2166/wst.2001.0806.

Koelliker Y, Totten LA, Gigliotti CL, Offenberg JH, Reinfelder JR, Zhuang Y, Eisenreich SJ. 2004. Atmospheric wet deposition of total phosphorus in New Jersey. Water Air Soil Pollut. 154:139–150. https://doi.org/10.1023/B:WATE.0000022952.12577.c5.

Lenhart HA, Hunt WF, Burchell MR. 2012. Harvestable nitrogen accumulation for five storm water wetland plant species: Trigger for storm water control measure maintenance? J Environ Eng. 138:972–978. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000550.

Littlewood M. 2001. Natural swimming pools: Inspiration for harmony with nature (1st ed). Schiffer Publishing, Ltd., Atglen, PA, USA.

Littlewood M. 2016. A guide to building natural swimming pools (1st ed). Schiffer Publishing, Ltd., Atglen, PA, USA.

Maine MA, Suñe N, Hadad H, Sánchez G. 2007. Temporal and spatial variation of phosphate distribution in the sediment of a free water surface constructed wetland. Sci Total Environ. 380:75–83. https://doi.org/10.1016/j.scitotenv.2006.11.036.

Maschinski J, Southam G, Hines J, Strohmeyer S. 1994. Efficiency of a subsurface constructed wetland system using native southwestern U.S. plants. J Environ Qual. 28:225–231. https://doi.org/10.2134/jeq1999.00472425002800010027x.

Miller RO. 2019. High-temperature oxidation: Dry ashing, p 53–56. In: Kalra YP (ed). Handbook and reference methods for plant analysis. CRC Press, New York, NY, USA.

Mlih R, Bydalek F, Klumpp E, Yaghi N, Bol R, Wenk J. 2020. Light-expanded clay aggregate (LECA) as a substrate in constructed wetlands–A review. Ecol Eng. 148:105783. https://doi.org/10.1016/j.ecoleng.2020.105783.

Moore MT, Kröger R. 2011. Evaluating plant species-specific contributions to nutrient mitigation in drainage ditch mesocosms. Water Air Soil Pollut. 217:445–454. https://doi.org/10.1007/s11270-010-0599-2.

Muerdter CP, Smith DJ, Davis AP. 2020. Impact of vegetation selection on nitrogen and phosphorus processing in bioretention containers. Water Environ Res. 92:236–244. https://doi.org/10.1002/wer.1195.

Muerdter CP, Wong CK, LeFevre GH. 2018. Emerging investigator series: The role of vegetation in bioretention for stormwater treatment in the built environment: Pollutant removal, hydrologic function, and ancillary benefits. Environ Sci Water Res Technol. 4:592–612. https://doi.org/10.1039/C7EW00511C.

Murphy J, Riley JP. 1962. A modified single solution method for determination of phosphate in natural waters. Anal Chim Acta. 27:31–36. https://doi.org/10.1016/S0003-2670(00)88444-5.

Naylor S, Brisson J, Labelle MA, Drizo A, Comeau Y. 2003. Treatment of freshwater fish farm effluent using constructed wetlands: The role of plants and substrate. Water Sci Technol. 48:215–222. https://doi.org/10.2166/wst.2003.0324.

Nicholls KH, Cox CM. 1978. Atmospheric nitrogen and phosphorus loading to Harp Lake, Ontario, Canada. Water Resour Res. 14:589–592. https://doi.org/10.1029/WR014i004p00589.

Petterson S, Li Q, Ashbolt N. 2021. Screening level risk assessment (SLRA) of human health risks from faecal pathogens associated with a natural swimming pond (NSP). Water Res. 188:116501. https://doi.org/10.1016/j.watres.2020.116501.

Polomski RF, Bielenberg DG, Whitwell T, Taylor MD, Bridges WC, Klaine SJ. 2007. Nutrient recovery by seven aquatic garden plants in a laboratory-scale subsurface-constructed wetland. HortScience. 42:1674–1680. https://doi.org/10.21273/HORTSCI.42.7.1674.

Poloprutská T, Nováček M, Oppeltová P. 2021. Case study of selected nature swimming pools in the south Moravian region. Ekologia (Bratisl). 40:312–324. https://doi.org/10.2478/eko-2021-0033.

Qiu F, Zhao S, Zhao D, Wang J, Fu K. 2019. Enhanced nutrient removal in bioretention systems modified with water treatment residuals and internal water storage zone. Environ Sci Water Res Technol. 5:993–1003. https://doi.org/10.1039/C9EW00093C.

Read J, Fletcher TD, Wevill T, Deletic A. 2009. Plant traits that enhance pollutant removal from stormwater in biofiltration systems. Int J Phytoremediation. 12:34–53. https://doi.org/10.1080/15226510902767114.

Read J, Wevill T, Fletcher T, Deletic A. 2008. Variation among plant species in pollutant removal from stormwater in biofiltration systems. Water Res. 42:893–902. https://doi.org/10.1016/j.watres.2007.08.036.

Rycewicz-Borecki M, McLean JE, Dupont RR. 2017. Nitrogen and phosphorus mass balance, retention and uptake in six plant species grown in stormwater bioretention microcosms. Ecol Eng. 99:409–416. https://doi.org/10.1016/j.ecoleng.2016.11.020.

Schets FM, van den Berg HHJL, Lynch G, de Rijk S, de Roda Husman AM, Schijven JF. 2020. Evaluation of water quality guidelines for public swimming ponds. Environ Int. 137:105516. https://doi.org/10.1016/j.envint.2020.105516.

Simkin SM, Allen EB, Bowman WD, Clark CM, Belnap J, Brooks ML, Cade BS, Collins SL, Geiser LH, Gilliam FS, Jovan SE, Pardo LH, Schulz BK, Stevens CJ, Suding KN, Throop HL, Waller DM. 2016. Conditional vulnerability of plant diversity to atmospheric nitrogen deposition across the United States. Proc Natl Acad Sci USA. 113:4086–4091. https://doi.org/10.1073/pnas.1515241113.

Tanner CC. 1996. Plants for constructed wetland treatment systems — A comparison of the growth and nutrient uptake of eight emergent species. Ecol Eng. 7:59–83. https://doi.org/10.1016/0925-8574(95)00066-6.

Tanner CC. 2001. Plants as ecosystem engineers in subsurface flow treatment wetlands. Water Sci Technol. 44:9–17. https://doi.org/10.2166/wst.2001.0804.

Tanner CC, Clayton JS, Upsdell MP. 1995. Effect of loading rate and planting on treatment of dairy farm wastewaters in constructed wetlands – II. Removal of nitrogen and phosphorus. Water Res. 29:27–34. https://doi.org/10.1016/0043-1354(94)00140-3.

Taylor MD, White SA, Chandler SL, Klaine SJ, Whitwell T. 2006. Nutrient management of nursery runoff water using constructed wetland systems. HortTechnology. 16:610–614. https://doi.org/10.21273/HORTTECH.16.4.0610.

US Environmental Protection Agency. 1988. Water quality standards criteria summaries: A compilation of state/federal criteria. USEPA, Office of Water Regulations and Standards, Washington, DC, USA.

Vogel JR, Chavez RA, Kandel S, Brown GO. 2021. Field scale demonstration of fly ash amended bioretention cells for stormwater phosphorus removal: A review of 12 years of work. Water. 13:1632. https://doi.org/10.3390/w13121632.

Vohla C, Kõiv M, Bavor HJ, Chazarenc F, Mander Ü. 2011. Filter materials for phosphorus removal from wastewater in treatment wetlands – A review. Ecol Eng. 37:70–89. https://doi.org/10.1016/j.ecoleng.2009.08.003.

Vymazal J. 2007. Removal of nutrients in various types of constructed wetlands. Sci Total Environ. 380:48–65. https://doi.org/10.1016/j.scitotenv.2006.09.014.

Vymazal J. 2020. Removal of nutrients in constructed wetlands for wastewater treatment through plant harvesting – Biomass and load matter the most. Ecol Eng. 155:105962. https://doi.org/10.1016/j.ecoleng.2020.105962.

Vymazal J, Kröpfelová L. 2009. Removal of nitrogen in constructed wetlands with horizontal sub-sureface flow: A review. Wetlands. 29:1114–1124. https://doi.org/10.1672/08-216.1.

Wang W, Han Y, Liu H, Zhang K, Yue Q, Bo L, Wang X. 2017. Pollutant removal performance of an integrated upflow-constructed wetland filled with haydites made of Al-based drinking water treatment residuals. Environ Technol. 38:1111–1119. https://doi.org/10.1080/09593330.2016.1220428.

White SA, Taylor MD, Albano JP, Whitwell T, Klaine SJ. 2011. Phosphorus retention in lab and field-scale subsurface-flow wetlands treating plant nursery runoff. Ecol Eng. 37:1968–1976. https://doi.org/10.1016/j.ecoleng.2011.08.009.

Wolf AM, Beegle DB. 2011. Recommended soil tests for macronutrients, p 39–47. In: Sims JT, Wolf A (eds). Recommended soil testing procedures for the northeastern United States. Northeast Reg Bull 493 (3rd ed). Agric Exp Sta, Univ Delaware, Newark, DE, USA.

Yan Q, James BR, Davis AP. 2018. Bioretention media for enhanced permeability and phosphorus sorption from synthetic urban stormwater. J Sustain Water Built Environ. 4:04017013. https://doi.org/10.1061/JSWBAY.0000836.

Yang Y-Y, Lusk MG. 2018. Nutrients in urban stormwater runoff: Current state of the science and potential mitigation options. Curr Pollut Rep. 4:112–127. https://doi.org/10.1007/s40726-018-0087-7.